

Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

**INSTITUTO OSWALDO CRUZ**  
**Pós-Graduação em Biologia Celular e Molecular**

***VANESSA DE VASCONCELOS SINATTI CASTILHO***

Clonagem, expressão heteróloga e caracterização molecular das enzimas gliconato cinase e uracil fosforribosiltransferase de *Trypanosoma cruzi*.

Dissertação apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Mestre em Biologia Celular e Molecular

**Orientadores:** Dr. Wim Maurits S. Degraeve  
Dr. Marcelo Alves-Ferreira

**RIO DE JANEIRO**

Abril de 2013

Ficha catalográfica elaborada pela  
Biblioteca de Ciências Biomédicas/ ICICT / FIOCRUZ - RJ

C352 Castilho, Vanessa de Vasconcelos Sinatti

Clonagem, expressão heteróloga e caracterização molecular das enzimas gliconato cinase e uracil fosforribosiltransferase de *Trypanosoma cruzi* / Vanessa de Vasconcelos Sinatti Castilho. – Rio de Janeiro, 2013.

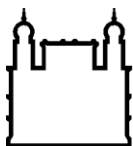
xxi, 108 f. : il. ; 30 cm.

Dissertação (Mestrado) – Instituto Oswaldo Cruz, Pós-Graduação em Biologia Celular e Molecular, 2013.

Bibliografia: f. 94-105

1. *Trypanosoma cruzi*. 2. Gliconatocinase. 3. Uracil fosforribosiltransferase. I. Título.

CDD 616.9363



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

**INSTITUTO OSWALDO CRUZ**  
**Pós-Graduação em Biologia Celular e Molecular**

***VANESSA DE VASCONCELOS SINATTI CASTILHO***

Clonagem, expressão heteróloga e caracterização molecular das enzimas gliconato cinase e uracil fosforribosiltransferase de *Trypanosoma cruzi*.

**ORIENTADORES:** Dr. Wim Maurits S. Degrave  
Dr. Marcelo Alves-Ferreira

**Aprovada em: 03/05/2013**

**EXAMINADORES:**

Dr<sup>a</sup>. Claudia Masini d'Ávila Levy - Presidente  
Dr. Antonio Ferreira Pereira  
Dr. Antonio Jorge Tempone  
Dr. Patricia Cuervo Escobar  
Dr. Fabio Faria da Mota

Rio de Janeiro, 03 de maio de 2013

Trabalho realizado no Laboratório de Genômica Funcional e Bioinformática (LAGFB/IOC), sob a orientação dos doutores Marcelo Alves-Ferreira e Wim Maurits S. Degrave.

Dedico a minha mãe por todo amor e incentivo e ao meu namorado Rodrigo por acreditar nos meus sonhos.

*“A ciência humana de maneira nenhuma nega a existência de Deus. Quando considero quantas e quão maravilhosas coisas o homem compreende, pesquisa e consegue realizar, então reconheço claramente que o espírito humano é obra de Deus, e a mais notável.”*

Galileu Galilei

*“A mais bela experiência que podemos ter é o mistério. É a emoção fundamental existente na origem da verdadeira arte e ciência. Aquele que não a conhece e não pode se maravilhar com ela está praticamente morto e seus olhos estão ofuscados.”*

Albert Einstein

## **AGRADECIMENTOS**

Para a realização de um grande sonho, o apoio, a confiança e a amizade são fundamentais e é por isso que tenho muito a agradecer:

Agradeço primeiramente a Deus por ter sempre me guiado e me dado força e determinação para alcançar meus objetivos e realizar meus sonhos.

À minha mãe, que sempre esteve ao meu lado me incentivando e tornando o meu caminho mais ameno. Você sempre foi e sempre será minha melhor amiga. Tenho muito orgulha de ser a sua filha, você me impulsiona para frente. Não sei o que seria de mim sem o seu apoio. Te amo!

À minha avó Magaly, minha segunda mãe, e ao meu pai por terem acreditado em mim e terem me apoiado em todas as minhas decisões, permitindo que eu prosseguisse com esse sonho. Amo vocês!

Ao meu namorado e melhor amigo Rodrigo, que sempre esteve presente me dando todo o seu amor, carinho e paciência sem fim. Você é o meu grande amor! Obrigada por fazer parte da minha vida, me dando conforto e fazendo eu me sentir especial. Você foi fundamental para a realização desse trabalho, quero sempre continuar crescendo ao seu lado.

À Angela, ao Odir e ao Ramon pelos bons momentos e por todo afeto.

À Vick, a minha pequenina linda e amada, por sempre estar ao meu lado.

Agradeço aos meus orientadores pelo grande aprendizado:

Ao meu orientador Dr. Marcelo A. Ferreira por ter me dado a oportunidade de ser sua aluna. Mesmo sendo “estressadinho”, te considero o meu pai científico. Obrigado por sempre ter acreditado no meu potencial e por toda a paciência e dedicação em me ensinar a fazer ciência.

Ao meu orientador Dr. Wim Degrave por toda disposição e empenho em me ajudar.

Essa jornada no Mestrado sem dúvida não trouxe só o aprendizado, mas amizades muito importantes que quero levar pela vida. Andressinha, agradeço por todo apoio, carinho e dedicação. Descobri em você uma amiga, irmã e que ainda faz papel de

anjo da guarda nas horas vagas. Obrigada por tudo! Claudinha, meu xodó! Acho que você é minha irmã gêmea, é incrível como nós somos parecidas e entendemos uma a outra. Adorei seguir esse projeto do seu lado. Estou no lucro, ganhei duas irmãs em dois anos! Já vale por tudo.

Agradeço à Renata, pela grande amizade e por todos os momentos que marcaram o início dessa história, lá na iniciação científica. Toda a sua ajuda e carinho foram muito importantes para mim.

Agradeço a todos do time 105:

À Pati, obrigada por todos os momentos divertidos e por toda ajuda que você me ofereceu. Se não fosse você não teria conhecido o mundo fascinante da imunofluorescência. Adorei!

Ao grande pesquisador Leo por todas as ideias geniais, os conselhos e dicas. Tudo foi muito valioso para mim, sem dúvida aprendi muito contigo.

Ao Beto por ter trazido um ambiente de descontração para a 105, bom humor é fundamental na ciência e, cá para nós, você é muito divertido.

Agradeço a todos da 101:

À Talitita por toda amizade e simpatia. Mesmo toda atarefada, você sempre esteve disposta a ajudar, conversar e dar conselhos. Obrigada pelo carinho!

Ao Leandrinho e a Palominha, grande a jornada a nossa na BCM, hein? Adorei acompanhar vocês nesse grande desafio que é o mestrado. Muitas disciplinas feitas juntos, muitos momentos compartilhados, amei tudo.

À Neiva por todo o carinho e disposição em contribuir para a realização de todos os projetos do laboratório, inclusive o meu.

Agradeço a todos da 201:

À Carol por ter um coração enorme e sempre ter me ajudado no desenvolvimento deste projeto. Considero você uma das minhas orientadoras! Sou grata a você por ter me orientado e me ensinado, com toda a paciência do mundo, tudo o que sei de bioinformática. Acho que sem a sua ajuda isso tudo não seria possível.



Ao Nicolas, por ser sempre simpático e atencioso comigo. As nossas conversas foram muito enriquecedoras para mim.

À Monete, pela sua companhia que é sempre muito agradável. Uma das únicas que entendem o que é morar em Jacarepaguá. Obrigada por todas as caronas.

Ao Marcos por estar sempre disposto a me dar dicas de bioinformática e me ajudar a aprimorar o meu conhecimento.

À Lalá, que mesmo pelo pouco tempo no laboratório já conquistou a amizade de todos, inclusive a minha. Adoro conversar contigo.

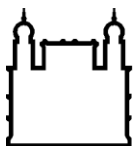
Agradeço a todos os pesquisadores do Laboratório de Genômica Funcional e Bioinformática:

À Dr<sup>a</sup>. Leila Medonça Lima e a Dr<sup>a</sup>. Mariana Caldas Waghbi por terem me recebido no laboratório e sempre serem atenciosas comigo. E a todos os outros integrantes dessa grande família por estarem sempre tão presentes e dispostos a ajudar: Ana, Cris, Aline, Renata, Dr<sup>a</sup>. Teca, Sônia. Vocês foram importantes para o meu crescimento como cientista.

Ao Dr. Antonio Basílio de Miranda por ter feito a revisão deste trabalho. As suas dicas e correções foram muito importantes a parte escrita desse projeto.

À todos aqueles que acreditaram em mim e me ajudaram de alguma maneira a alcançar esse importante passo em minha.

Ao Instituto Oswaldo Cruz pelo apoio financeiro pela oportunidade de realizar este trabalho, me através do qual recebi minha bolsa de mestrado e, que tanto acrescentou no meu amadurecimento profissional.



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

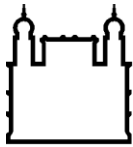
## INSTITUTO OSWALDO CRUZ

### Clonagem, expressão heteróloga e caracterização molecular das enzimas gliconato cinase e uracil fosforibosiltransferase de *Trypanosoma cruzi*

#### RESUMO

**Vanessa de Vasconcelos Sinatti Castilho**

A doença de Chagas é um problema de saúde pública na América Latina, com um forte impacto no Brasil. De acordo com estimativas da Organização Panamericana de Saúde (OPAS), em torno de 8 milhões de pessoas nas Américas estão infectados pelo *Trypanosoma cruzi*, com uma taxa de mortalidade de 12.000 indivíduos por ano. Apesar desta situação, os medicamentos utilizados no tratamento desta doença são ativos somente na fase aguda, apresentam uma baixa eficácia e uma série de efeitos colaterais. Portanto, a busca por novos alvos terapêuticos, através do estudo do metabolismo de *T. cruzi*, é uma alternativa no desenvolvimento de novos fármacos. Neste contexto, utilizou-se o AnEnPi, uma ferramenta computacional capaz de identificar e classificar proteínas com uma potencial atividade enzimática, utilizando dados genômicos previamente publicados, permitindo a reconstrução de diferentes vias metabólicas. Por meio dessa abordagem foram identificados em *T. cruzi* duas proteínas, as quais apresentam as seguintes atividades enzimáticas: uracil fosforibosiltransferase (UPRT) e gliconato cinase (GK). A UPRT é uma enzima da via de salvação de pirimidinas que converte uracil e D-5-fosforibosil-1-fosfato (PRPP) a monofosfato de uridina (UMP) e pirofosfato (PPi). E a GK é uma enzima da via das pentoses-fosfato que catalisa a fosforilação de gliconato a 6-fosfogliconato. Análises das sequências genômicas codificante para essas proteínas, utilizando diferentes ferramentas de bioinformática, permitiram a identificação de dois genes para cada uma dessas atividades enzimáticas em *T. cruzi* e confirmaram essas atividades pela identificação de seus motivos estruturais. Posteriormente, foi realizada a caracterização molecular das enzimas GK e UPRT de *T. cruzi*, através da clonagem dos respectivos genes, utilizando os vetores pBAD-TOPO TA e pET28a, e da expressão heteróloga em *Escherichia coli*, cepa BL21(DE3). As proteínas recombinantes GK e UPRT foram expressas na fração solúvel e insolúvel com pesos moleculares aproximados de 23 kDa e 28 kDa, respectivamente. Essas proteínas recombinantes foram purificadas sob condições desnaturantes através de cromatografia de afinidade. Os soros policlonais produzidos contra as proteínas purificadas foram eficazes no reconhecimento das proteínas recombinantes e de suas formas nativas presentes na forma epimastigota de *T. cruzi*. O ensaio de imunofluorescência revelou que as enzimas GK e UPRT estão localizadas possivelmente no citoplasma e no interior de vesículas citoplasmáticas ao longo do parasita. Além disto, a estrutura de ambas as enzimas foi deduzida através de modelagem por homologia. Este trabalho visa contribuir para o melhor entendimento do metabolismo deste parasita, de forma a auxiliar na busca de novos alvos terapêuticos contra a doença de Chagas.



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

## INSTITUTO OSWALDO CRUZ

### Cloning, heterologous expression and molecular characterization of the enzymes gluconate kinase and uracil phosphoribosyltransferase from *Trypanosoma cruzi*

#### ABSTRACT

**Vanessa de Vasconcelos Sinatti Castilho**

Chagas' disease is a public health problem in Latin America with a strong impact in Brazil. According to estimates by the Pan American Health Organization (PAHO), around 8 million people in the Americas are infected with *Trypanosoma cruzi* with rate of 12,000 deaths per year. Despite this situation, the drugs used for treating this disease are active only in the acute phase and have low efficiency and many side effects. Therefore, search for new therapeutic targets, by studying the metabolism of *T. cruzi*, is an alternative for the development of new drugs. In this context, we used AnEnPi, a software tool able to identify and classify proteins with a potential enzymatic activity using genomic data previously published, allowing the reconstruction of different metabolic pathways. Through this approach two proteins were identified in *T. cruzi*, which have the following enzymatic activities: uracil phosphoribosyltransferase (UPRT) and gluconate kinase (GK). UPRT is a pyrimidine salvage enzyme that converts uracil and 5-phosphoribosyl- $\alpha$ -1-pyrophosphate (PRPP) to uridine monophosphate (UMP) and pyrophosphate (PPi). GK is an enzyme of the pentose phosphate pathway that catalyses the phosphorylation of D-gluconate to 6-phospho-D-gluconate. Analysis of genomic sequences encoding for these proteins using various bioinformatic tools allowed the identification of two genes for each enzyme in *T. cruzi* and confirmed these activities by identifying their structural motifs. Subsequently, we performed the molecular characterization of the enzymes GK and UPRT of *T. cruzi*, by cloning the respective genes using the vector pBAD-TOPO and pET28a, and heterologous expression in *E. coli*, strain BL21 (DE3). The recombinant proteins GK and UPRT were expressed in the insoluble and soluble fraction with approximate molecular weights of 23 kDa and 28 kDa, respectively. These recombinant proteins were purified under denaturing conditions by affinity chromatography. The polyclonal antisera raised against the purified proteins were effective in the recognition of recombinant proteins and their native forms present in the epimastigote form of *T. cruzi*. The immunofluorescence assay revealed that GK and UPRT enzymes were possibly located in the cytoplasm and inside cytoplasmic vesicles along the parasite. Moreover, the structure of both enzymes was deduced by homology modeling. This work aims to contribute to a better understanding of the parasite's metabolism in order to help with the search for new therapeutic targets against Chagas' disease.

## LISTA DE FIGURAS

### INTRODUÇÃO

**Figura 1.1:** Mapa das rotas de migração de indivíduos infectados com a doença de Chagas e as estimativas do número de infectados em países não endêmicos

**Figura 1.2:** Sinais característicos da porta de entrada da doença de Chagas

**Figura 1.3:** Ilustração em corte longitudinal mostrando as principais estruturas celulares das três formas evolutivas de *T. cruzi*

**Figura 1.4:** Ciclo evolutivo de *T. cruzi*

**Figura 1.5:** Fluxograma do funcionamento do programa AnEnPi

**Figura 1.6:** Via das pentoses-fosfato

**Figura 1.7:** Reação catalisada pela enzima GK

**Figura 1.8:** Estrutura tridimensional de GK de *E. coli* resolvida por difração de raios-x

**Figura 1.9:** Estrutura de nucleotídeos

**Figura 1.10:** Síntese *de novo* das pirimidinas

**Figura 1.11:** Via de salvação de pirimidinas

**Figura 1.12:** Reação catalisada pela enzima UPRT

**Figura 1.13:** Estrutura tridimensional de UPRT de *Sulfolobus solfataricus*, *T. gondii*, e *Bacillus caldolyticus* resolvidas por difração de raios-x

### METODOLOGIA

**Figura 3.1:** Representação esquemática do vetor de clonagem pBAD TOPO TA e da reação catalisada pela topoisomerase I

**Figura 3.2:** Representação esquemática do vetor de expressão pET28a

**Figura 3.3:** Representação esquemática da disposição do papel 3mm, gel de poliacrilamida, membrana de nitrocelulose no sistema “Semi-dry” (BioRad) para a transferência

## RESULTADOS

**Figura 4.1:** Sequência do gene *TcGK*- Tc00.1047053508177.20 de *T. cruzi*

**Figura 4.2:** Sequência do gene *TcGK*- Tc00.1047053511693.60 de *T. cruzi*

**Figura 4.3:** Alinhamento das estruturas primárias de *TcGK* identificadas no genoma de *T. cruzi*

**Figura 4.4:** Sequências do gene *TcUPRT*- Tc00.1047053507739.40 de *T. cruzi*

**Figura 4.5:** Sequência do gene *TcUPRT*- Tc00.1047053506435.410 de *T. cruzi*

**Figura 4.6:** Alinhamento das estruturas primárias de *TcUPRT* identificadas no genoma de *T. cruzi*

**Figura 4.7:** Análise em gel de agarose 1% da amplificação por PCR dos genes *TcGK* e *TcUPRT*

**Figura 4.8:** Análise do PCR de colônia para confirmação da clonagem dos genes *TcGK* e *TcUPRT* no vetor pBAD-TOPO TA

**Figura 4.9:** Digestão das construções plasmidiais *TcGK*-pBAD –TOPO TA e *TcUPRT*-pBAD-TOPO TA e do plasmídeo pET28a

**Figura 4.10:** Análise do PCR de colônia para a confirmação da clonagem dos genes *TcGK* e *TcUPRT* no vetor pET28a

**Figura 4.11:** Análise da expressão total em pequena escala das proteínas recombinantes *TcGK* e *TcUPRT*

**Figura 4.12:** Análise dos testes de expressão em pequena escala da proteína *TcGK* recombinante de 23 kDa

**Figura 4.13:** Análise dos testes de expressão em pequena escala da proteína *TcUPRT* recombinante de aproximadamente 28,7 kDa

**Figura 4.14:** Análise da expressão em larga escala da proteína *TcGK* recombinante de 23,2 kDa

**Figura 4.15:** Análise da expressão em larga escala da proteína *TcUPRT* recombinante de 28,7 kDa

**Figura 4.16:** Cromatograma da purificação da proteína recombinante *TcGK*

**Figura 4.17:** Análise da purificação por cromatografia de afinidade da proteína recombinante TcGK. SDS-PAGE 15%

**Figura 4.18:** Cromatograma da purificação da proteína recombinante TcUPRT

**Figura 4.19:** Análise da purificação por cromatografia de afinidade da proteína recombinante TcUPRT

**Figura 4.20:** Titulação dos anticorpos policlonais anti-TcGK e anti-TcUPRT

**Figura 4.21:** Identificação por *western blot* das enzimas nativas TcGK e TcUPRT nas frações solúveis das formas epimastigotas de *T. cruzi*, cepa CL Brener

**Figura 4.22:** Identificação da enzima nativa GK na forma epimastigota de *T. cruzi* por imunofluorescência

**Figura 4.23:** Identificação da enzima nativa UPRT na forma epimastigota de *T. cruzi* por imunofluorescência

**Figura 4.24:** Alinhamento múltiplo entre as sequências de aminoácidos da GK de *E. coli* (Eco/GK) e da *T. cruzi* (Tcr/GK)

**Figura 4.25:** Alinhamento entre as estruturas secundárias de GK de *E. coli* (Eco/GK) e de *T. cruzi* (Tcr/GK)

**Figura 4.26:** Alinhamento múltiplo entre as sequências de aminoácidos da UPRT de *T. gondii* (Tgo/UPRT), de *T. cruzi* (Tcr/UPRT1) e de humano (Hsa/UPRT)

**Figura 4.27:** Alinhamento entre as estruturas secundárias de UPRT de *T. gondii* (Tgo/UPRT), de *T. cruzi* (Tcr/UPRT) e de humano (Hsa/UPRT)

**Figura 4.28:** Estrutura tridimensional de enzima TcGK determinada por modelagem por homologia

**Figura 4.29:** Gráfico Ramachandran da modelagem de TcGK

**Figura 4.30:** Estrutura tridimensional da enzima TcUPRT determinada por modelagem por homologia

**Figura 4.31:** Gráfico Ramachandran da modelagem de TcUPRT

## LISTA DE TABELAS

### MATERIAS E MÉTODOS

**Tabela 3.1:** Solução inorgânica (2x) do meio de cultura LIT

**Tabela 3.2:** Solução orgânica (2x) do meio de cultura LIT

**Tabela 3.3:** Meio de cultura LIT completo

**Tabela 3.4:** Componentes utilizados na reação de PCR

**Tabela 3.5:** Reagentes da reação de ligação ao vetor pBAD-TOPO TA

**Tabela 3.6:** Componentes utilizados na reação de PCR de colônia

**Tabela 3.7:** Reação de digestão com as enzimas *XhoI* e *NdeI*

**Tabela 3.8:** Reagentes da reação de ligação ao vetor pET28a

**Tabela 3.9:** Reagentes para preparação do SDS-PAGE

**Tabela 3.10:** Tampões para purificação por cromatografia de afinidade

**Tabela 3.11:** Informações sobre os anticorpos primários e os anticorpos secundários de cada experimento de *western blot* realizado

**Tabela 3.12:** Relação de anticorpos primários e as suas respectivas diluições nos ensaios de imunofluorescência

### RESULTADOS

**Tabela 4.1:** Relação dos potenciais domínios e assinaturas de famílias nas sequências de aminoácidos de TcGK e TcUPRT

**Tabela 4.2:** Predição de peptídeo sinal, região transmembranar e localização celular das enzimas TcGK e TcUPRT

**Tabela 4.3:** Predição pelo PROSITE de motivos característicos de sítios de ligação e de modificação pós-traducionais nas sequências de aminoácidos de TcGK e TcUPRT

## LISTA DE SIGLAS E ABREVIATURAS

C - Coulomb

°C - grau Celsius

µm - micrômetro

µg - micrograma

µL - microlitro

ADP- adenosina difosfato

ATP - adenosina trifosfato

Abs- absorvância

dNTP- desoxirribonucleotídeo 5' trifosfato

BSA - albumina de soro bovino

CMP, monofosfato de citidina;

CTP, trifosfato de citidina;

DAPI- 4',6-diamidino-2-phenylindol

DABCO - 1,4-diacibiciclo-octano

dCMP, monofosfato de desoxicitidina;

DNA - ácido desoxirribonucléico

dR-1-P, desoxirribose 1-fosfato;

dTMP, monofosfato de desoxitimidina.

dUMP- monofosfato de desoxiuridina

DTT - ditioneitol

E.C- *Enzyme comission*

EDTA- ácido etilenodiamino tetra-acético

FAD - flavina-adenina dinucleotídeo

g - aceleração da gravidade

GK - gliconato cinase

Gln, glutamina;

Glu, ácido glutâmico

GTP- trifosfato de guanosina

h - hora

IgG - imunoglobulina G

IPTG - isopropil-β-D-Galactopiranosídeo

kb - quilobase



kDa - quilodalton  
LB- Luria Bertani  
M - molar  
mA - miliamper  
mL - mililitro  
mM - milimolar  
NAD<sup>+</sup>, dinucleotídeo de nicotinamida-adenina oxidado;  
NADPH, dinucleotídeo de nicotinamida-adenina-fosfato.  
NADH+H<sup>+</sup>, dinucleotídeo de nicotinamida-adenina reduzido;  
OD- *optical density* (densidade ótica)  
OMP- orotidilato  
pb (s) - par(es) de base  
PBS - “phosphate buffered saline” (salina tamponada com fosfato)  
PCR- Reação em cadeia polimerase  
pH - potencial hidrogeniônico  
Pi, fosfato inorgânico;  
PPi, pirofosfato;  
PMSF- Fluoreto fenilmetilsulfônico  
PRPP, D-5-fosforibosil-1-fosfato;  
qsp - quantidade suficiente para  
R- Resistência  
R-1-P - ribose 1-fosfato  
rpm - rotações por minuto  
SDS - dodecil sulfato de sódio  
SDS-PAGE- Eletroforese em gel de poliacrilamida com SDS  
TAE- tris-Acetato-EDTA  
TBS- tampão salina de tris  
TCA - ácido tricloro acético  
TEMED- N, N, N', N'- tetrametiletano, 1, 2- diamina  
TRIS - tris-hidroxi-metil-amino-metano  
TTP- trifosfato de timidina  
UDP- difosfato de uridina  
UMP, monofosfato de uridina;  
UPRT- uracil fosforribosiltransferase

UTP, trifosfato de uridina;

V - voltagem

v - volts

v/v - volume/volume

WHO - "World Health Organization" (Organização Mundial da Saúde"

## ÍNDICE

Agradecimentos.....	vi
RESUMO .....	ix
ABSTRACT .....	x
LISTA DE FIGURAS .....	xi
LISTA DE TABELAS.....	xiv
LISTA DE SIGLAS E ABREVIATURAS .....	xv
<b>1. INTRODUÇÃO</b>	<b>1</b>
1.1. Doença de Chagas .....	1
1.2. <i>T. cruzi</i> .....	4
1.3. <i>T. cruzi</i> e seu ciclo de vida .....	6
1.4. Tratamento da doença de Chagas .....	7
1.5. AnEnPi (Analogous Enzyme Pipeline).....	9
1.6. Via das pentoses-fosfato .....	11
1.6.1. Gliconato cinase (GK).....	14
1.7. Metabolismo de nucleotídeos .....	15
1.7.1. Síntese <i>de novo</i> das pirimidinas .....	17
1.7.2. Via de salvação das pirimidinas.....	19
1.7.3. Uracil fosforribosiltransferase (UPRT) .....	22
<b>2. OBJETIVOS</b>	<b>25</b>
2.1. Objetivo geral.....	25
2.2. Objetivos específicos: .....	25
<b>3. MATERIAIS E MÉTODOS</b>	<b>26</b>
3.1. Identificação <i>in silico</i> das enzimas GK e UPRT de <i>T. cruzi</i> .....	26
3.2. Análise das estruturas primárias de TcGK e TcUPRT .....	26
3.3. Cultivo das formas epimastigotas de <i>T. cruzi</i> .....	27
3.4. Extração do DNA genômico de <i>T. cruzi</i> .....	28
3.5. Amplificação dos genes <i>TcGK</i> e <i>TcUPRT</i> por PCR (Reação da Polimerase em Cadeia) .....	28
3.6. Purificação dos produtos de PCR.....	30
3.7. Eletroforese em gel de agarose .....	30
3.8. Clonagem dos alvos selecionados em vetor pBAD-TOPO TA .....	30

3.9. PCR de colônia .....	32
3.10. Extrações de DNA plasmidial ( <i>Miniprep</i> ), sequenciamento e análise dos resultados .....	33
3.11. Dosagem do DNA plasmidial.....	34
3.12. Digestão do DNA plasmidial com as enzimas de restrição <i>XhoI</i> e <i>NdeI</i> .....	34
3.13. Subclonagem dos genes alvos no vetor pET28a .....	35
3.14. Expressão das proteínas recombinantes TcGK e TcUPRT em pequena escala.....	36
3.15. Eletroforese em gel de poliacrilamida em condições desnaturantes (SDS-PAGE).....	38
3.16. Testes de expressão das proteínas recombinantes em pequena escala .....	38
3.17. Lise bacteriana e separação das proteínas recombinantes TcGK e TcUPRT em fração solúvel e insolúvel.....	39
3.18. Expressão das proteínas recombinantes TcGK e TcUPRT em larga escala.....	40
3.19. Purificação por cromatografia de afinidade.....	41
3.20. Dosagem de proteína .....	42
3.21. Produção de anticorpos policlonais anti-TcGK e anti-TcUPRT .....	43
3.22. Titulação dos anticorpos policlonais .....	43
3.23. Purificação de imunoglobulina G a partir dos soros imunes .....	44
3.24. Western blot.....	44
3.25. Lise e extração de proteínas solúveis da forma epimastigota de <i>T. cruzi</i> .....	46
3.26. Identificação das enzimas nativas TcGK e TcUPRT nas frações solúveis da forma epimastigota de <i>T. cruzi</i> por <i>western blot</i> .....	47
3.27. Identificação das enzimas nativas TcGK e TcUPRT na forma epimastigota de <i>T. cruzi</i> por imunofluorescência .....	48
3.28. Modelagem estrutural das enzimas TcGK e TcUPRT e UPRT de <i>Homo sapiens</i> .....	49
<b>4. RESULTADOS</b>	<b>50</b>

<b>4.1. Análises das sequências nucleotídicas e das estruturas primárias das proteínas GK e UPRT .....</b>	<b>50</b>
4.1.1. Sequências nucleotídicas e protéicas referentes à GK .....	50
4.1.2. Alinhamento entre as sequências de aminoácidos preditas de TcGK.....	51
4.1.3. Sequências nucleotídicas e protéicas referentes à UPRT .....	52
4.1.4. Alinhamento entre as sequências protéicas preditas de TcUPRT .....	53
<b>4.2. Análises das sequências de aminoácidos das proteínas TcGK e TcUPRT para identificação de possíveis domínios, motivos e assinaturas de famílias de proteínas.....</b>	<b>54</b>
<b>4.3. Amplificação dos genes <i>TcGK</i> e <i>TcUPRT</i> por PCR.....</b>	<b>57</b>
<b>4.4. Clonagem dos genes <i>TcGK</i> e <i>TcUPRT</i> no vetor pBAD-TOPO TA .....</b>	<b>58</b>
<b>4.5. Sequenciamento das construções <i>TcGK</i> e <i>TcUPRT</i> em vetor pBAD-TOPO TA .....</b>	<b>59</b>
<b>4.6. Digestão do DNA plasmidial com as enzimas de restrição <i>XhoI</i> e <i>NdeI</i>.....</b>	<b>59</b>
<b>4.7. Subclonagem dos genes alvos no vetor pET28a .....</b>	<b>60</b>
<b>4.8. Expressão em pequena escala das proteínas recombinantes TcGK e TcUPRT .....</b>	<b>61</b>
<b>4.9. Sequenciamento das construções <i>TcGK</i>-pET28a e <i>TcUPRT</i>-pET28a.....</b>	<b>62</b>
<b>4.10. Testes de expressão e solubilidade das proteínas recombinantes TcGK e TcUPRT em pequena escala, seguidos de lise bacteriana e fracionamento das proteínas totais .....</b>	<b>63</b>
<b>4.11. Expressão em larga escala das proteínas recombinantes TcGK e TcUPRT .....</b>	<b>65</b>
<b>4.12. Purificação das proteínas recombinantes por cromatografia de afinidade.....</b>	<b>67</b>
<b>4.13. Titulação dos anticorpos policlonais anti-TcGK e anti-TcUPRT .....</b>	<b>70</b>
<b>4.14. Identificação das enzimas nativas TcGK e TcUPRT nas formas epimastigotas de <i>T. cruzi</i> por <i>western blot</i>.....</b>	<b>71</b>
<b>4.15. Identificação das enzimas nativas TcGK e TcUPRT nas formas epimastigotas de <i>T. cruzi</i> por imunofluorescência .....</b>	<b>72</b>

4.16. Análise das regiões de ligação ao substrato na estrutura primária de GK de <i>T. cruzi</i> .....	75
4.17. Predição da estrutura secundária de TcGK.....	75
4.18. Análise das possíveis regiões de ligação ao substrato e ao produto na estrutura primária da TcUPRT e da UPRT de humano ....	76
4.19. Predição da estrutura secundária da TcUPRT e da UPRT de humano .....	77
4.20. Modelagem estrutural por homologia da enzima TcGK.....	79
4.21. Modelagem estrutural por homologia da enzima TcUPRT e da UPRT de humano.....	80
<b>5. DISCUSSÃO</b>	<b>83</b>
<b>6. CONCLUSÕES</b>	<b>93</b>
<b>7. REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>94</b>
<b>8. ANEXOS</b>	<b>106</b>
8.1. Alinhamento das sequências de <i>TcGK</i> contra as sequências de referência de um dos clones de pBAD-TOPO TA, utilizando o programa ClustaW2.....	106
8.2. Alinhamento das sequências de <i>TcUPRT</i> contra as sequências de referência de um dos clones de pBAD-TOPO TA, utilizando o programa ClustaW2.....	107

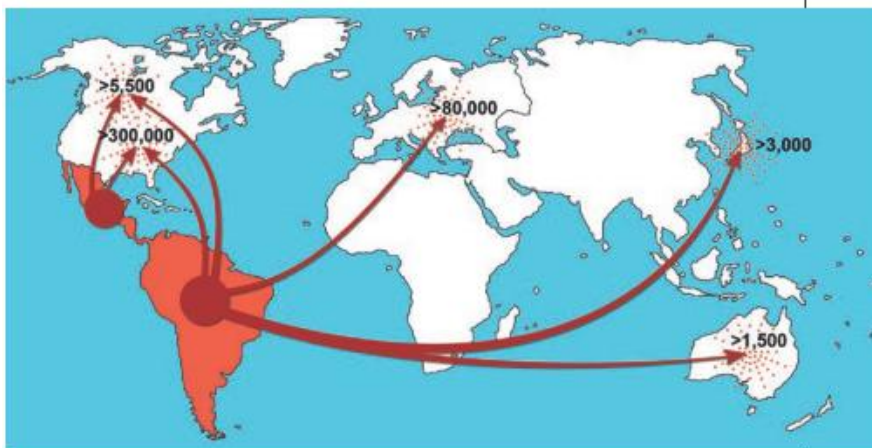
## 1. INTRODUÇÃO

### 1.1. Doença de Chagas

A doença de Chagas ou tripanossomíase americana é uma infecção causada pelo protozoário flagelado *Trypanosoma cruzi*, e está entre uma das doenças tropicais negligenciadas mais importantes. Atualmente, estima-se que mais de 100 milhões de pessoas estejam na área de risco dessa doença, que afeta em torno de 8 milhões de pessoas na América Latina e é a responsável por 12.000 mortes por ano (OPAS, 2012).

O *T. cruzi* pode ser transmitido não só para os seres humanos como também para animais domésticos e silvestres, que são os reservatórios naturais desse parasito. Uma das principais formas de transmissão é a vetorial causada por insetos hemípteros hematófagos infectados da subfamília *Triatominae*. Dentre as espécies classificadas nessa subfamília, apenas um pequeno número pode ser considerado vetor competente para o *T. cruzi*, nos quais se destacam as espécies *Triatoma infestans*, *Rhodnius prolixus*, *Triatoma dimidiata*, *Panstrongylus megistus* e *Triatoma brasiliensis* (Rassi *et al.*, 2012). Outras formas de transmissão da doença de Chagas envolvem mecanismos não vetoriais, tais como transfusão de sangue, transmissão congênita (via placenta) e transplante de órgãos contaminados, que são as principais causas de infestação urbana, além da transmissão oral e contaminação ocasionada por acidentes laboratoriais (Pérez-Molina *et al.*, 2012)

A distribuição geográfica da doença se localiza principalmente em países latino americanos onde a doença de Chagas é endêmica. No entanto, nas últimas décadas, os movimentos populacionais migratórios de indivíduos contaminados a partir dos países endêmicos têm levado a doença de Chagas para países inicialmente não afetados, tais como EUA, Canadá, Austrália, Japão e alguns países da Europa (figura 1.1). Em consequência disso, houve um considerável aumento no número de indivíduos infectados nesses países (Schmunis, 2007; Pérez-Molina *et al.*, 2012). Com a inexistência da transmissão vetorial, as maiores causas da infecção nessas regiões são a transfusão sanguínea e a transmissão vertical de mãe para filho, seguido do transplante de órgãos e de acidentes laboratoriais em menor escala (Coura & Borges-Pereira, 2010).



**Figura 1.1:** Mapa das rotas de migração de indivíduos infectados com a doença de Chagas e as estimativas do número de infectados em países não endêmicos. Fonte: imagem retirada de Coura & Viñas, 2010.

O Brasil e alguns outros países da América Latina têm realizado medidas de controle da doença de Chagas que incluem combate ao vetor, controle dos bancos de sangue, instrução sobre os mecanismos da doença a populações de risco e, em alguns casos, a observação da migração de indivíduos de áreas rurais endêmicas para as cidades com baixa taxa de transmissão vetorial (Martins-Melo *et al.*, 2012). Controles desse tipo têm sido bastante eficazes na redução da transmissão da doença e na eliminação, em áreas tipicamente endêmicas, de algumas populações de vetores domésticos, como as espécies *Triatoma infestans*, *Rhodnius prolixus* e outras espécies largamente distribuídas (Dias *et al.*, 2002). Como resultado da implementação dessas medidas no Brasil, em 2006 ele foi declarado livre da transmissão por *Triatoma infestans* (Gurgel-Gonçalves *et al.*, 2012). Entretanto, esta doença permanece um importante problema de saúde pública, apresentando um índice de 56.000 novos casos por ano (OPAS, 2012).

A doença de Chagas pode ser dividida em duas fases de manifestação: a fase aguda e a fase crônica. A fase aguda se inicia logo após a infecção com o *T. cruzi* e pode se estender de 4 a 8 semanas. Os parasitos se replicam ativamente nessa fase e podem ser detectados através de microscopia do sangue. Na maioria dos indivíduos, a fase aguda é assintomática; caso contrário, os sintomas se desenvolvem em torno de 1 a 2 semanas após o contato inicial com os parasitas ou após períodos maiores quando a contaminação se dá por meio de transfusão de sangue. Os sintomas são geralmente inespecíficos e incluem febre, dor de cabeça, náusea, vômito e diarreia. Em casos mais graves pode ocorrer inchaço do fígado e linfonodos. Na transmissão vetorial, é comum o chagoma de inoculação, em que um



edema é formado na região cutânea onde o parasito foi inoculado. Se a rota de entrada do parasito é a conjuntiva ocular, o edema causado é denominado sinal de Romaña (figura 1.2) (Rassi *et al.*, 2010; Parker & Sethi, 2011).



**Figura 1.2:** Sinais característicos da porta de entrada da doença de Chagas. **(A)** Sinal de Romaña (Fonte: <http://www.fac.org.ar/fec/chagas/c03lugon/c03lugon.htm>). **(B)** Chagoma de inoculação (Fonte: <http://www.bioscience.org/2003/v8/e/948/fulltext.asp?bframe=figureshtm&doi=yes>).

A fase crônica se desenvolve de dois a três meses após a infecção inicial e persiste ao longo da vida do paciente. Nessa fase, há uma diminuição no número de parasitos circulantes, que passam a se localizar nos tecidos alvos, principalmente nos músculos lisos do sistema digestivo e nos músculos cardíacos, já não sendo mais detectáveis por microscopia do sangue. A fase crônica possui três formas de manifestação clínica: a forma indeterminada, a forma cardíaca e a forma digestiva. (Coura & Borges-Pereira, 2010)

Os pacientes na forma indeterminada da fase crônica permanecem completamente assintomáticos e, em 60% a 70% dos casos, nunca evoluem para as outras formas crônicas da doença. A forma indeterminada é caracterizada por uma sorologia positiva para anticorpos contra *T. cruzi* e um eletrocardiograma e radiologias do esôfago, colón e tórax sem alterações notáveis. Apesar disso, pequenas lesões inflamatórias focais podem ser detectadas em amostras de tecido do coração, músculo esquelético e do trato gastrointestinal de pacientes assintomáticos (Prata, 2001; Rassi *et al.*, 2010; Muñoz-Saravia *et al.*, 2012).

Após décadas de latência, um terço desses pacientes desenvolvem as formas mais graves apresentando manifestações clínicas relacionadas a patologias cardíacas, digestivas ou uma combinação delas. A cardiopatia chagásica é a forma

clínica mais severa e frequente da fase crônica e é responsável por um acometimento cardíaco, podendo levar a insuficiência e arritmias cardíacas, dilatação das quatro cavidades do coração, tromboembolismo e morte súbita. Nessa forma, o eletrocardiograma é marcado por anormalidades e as radiografias do tórax podem mostrar diferentes graus de cardiomegalia (Rassi *et al.*, 2000; Acquatella, 2007).

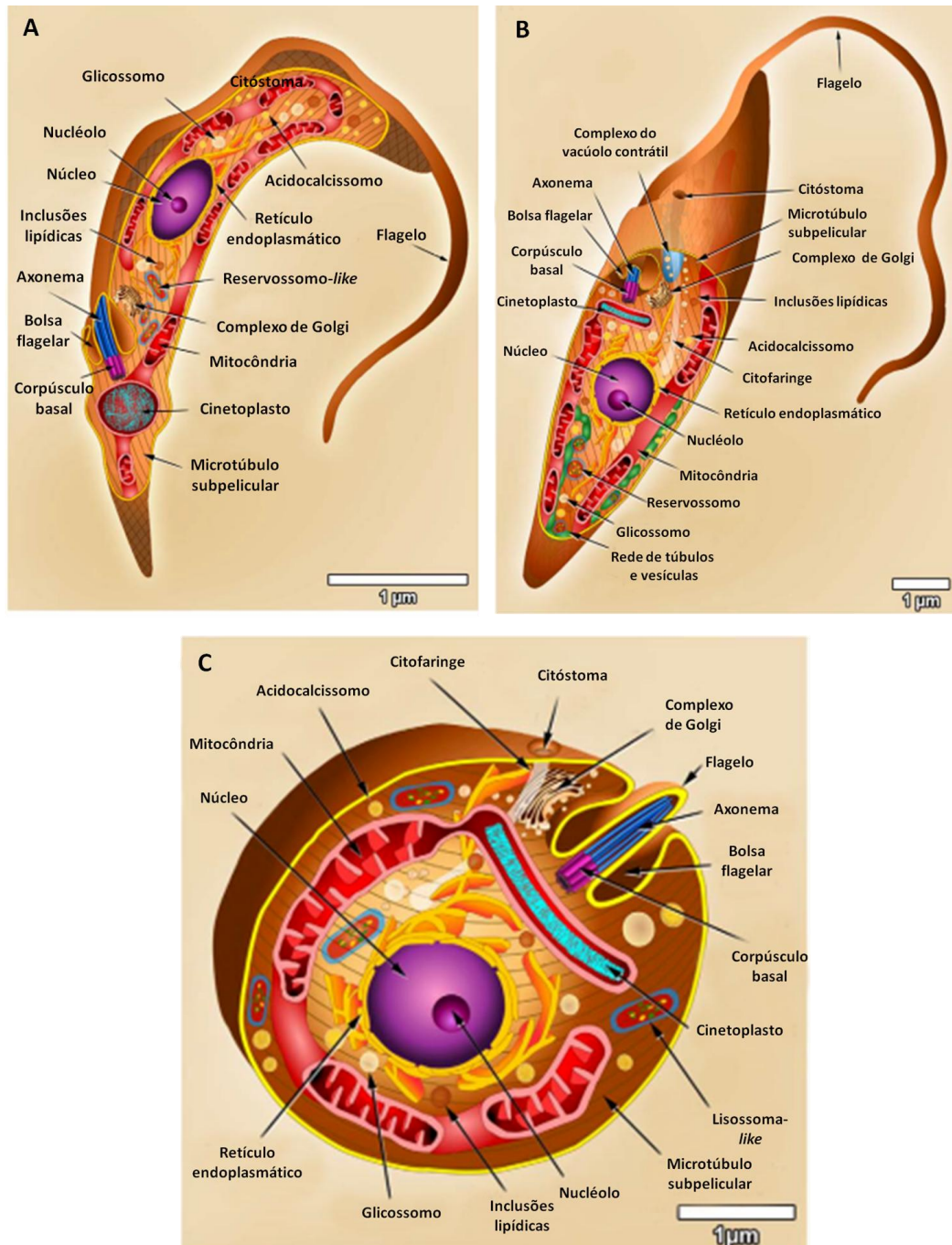
Já a forma digestiva se desenvolve em indivíduos que sofreram destruição das células do sistema nervoso entérico pela presença do parasito, causando alterações fisiológicas dos órgãos do sistema gastrointestinal. Essas alterações incluem problemas motores relacionados aos movimentos peristálticos, deficiência na secreção do suco gástrico e na absorção do alimento ingerido. Além disso, alterações morfológicas também são observadas principalmente nas regiões do cólon e do esôfago, podendo originar a formação do megacólon e do megaesôfago, respectivamente. Em consequência disso, sintomas de disfagia, excesso de salivação, regurgitação, dor abdominal, constipação e perda de peso são relatados pelos pacientes. Nessa forma da doença, a eletrogastrografia mostra alta incidência de disritmias gástricas e os exames radiológicos podem indicar a dilatação dos órgãos afetados (Matsuda *et al.*, 2009).

## 1.2. *T. cruzi*

O *T. cruzi*, o agente etiológico da doença de Chagas, é um protozoário pertencente à ordem Kinetoplastida, família Trypanosomatidae. As principais características desta ordem são a presença de flagelo e do cinetoplasto, uma estrutura formada por DNA mitocondrial condensado, localizada próxima ao corpúsculo basal do flagelo. Esse DNA mitocondrial (mtDNA) se organiza em maxi e mini círculos interligados e está contido em uma mitocôndria única que se ramifica por toda a célula do parasito (Brusca & Brusca, 2007).

Por ser um organismo eucarioto, o *T. cruzi* possui organelas comuns a esse grupo, como núcleo, retículo endoplasmático, complexo de Golgi e mitocôndria. Além disso, ele apresenta algumas organelas e estruturas características da família Trypanosomatidae, como, por exemplo, os microtúbulos subpeliculares, a estrutura paraflagelar, os reservossomos, o cinetoplasto, o citóstoma, os glicossomos e os acidocalcisomos (de Souza, 2002).

O *T. cruzi* é um parasita heteroxênico, possuindo um hospedeiro vertebrado e um invertebrado, e ao longo do seu desenvolvimento possui três formas evolutivas: tripomastigota, amastigota e epimastigota (figura 1.3). Essas formas se diferenciam morfológicamente pela posição do cinetoplasto em relação ao núcleo da célula e à emergência do flagelo (Brener, 1997).



**Figura 1.3:** Ilustração em corte longitudinal mostrando as principais estruturas celulares das três formas evolutivas de *T. cruzi*. (A) Forma tripomastigota. (B) Forma epimastigota. (C) Forma amastigota. Adaptada de Teixeira *et al.*, 2012.

A forma tripomastigota (figura 1.3A) é alongada e possui em torno de 25  $\mu\text{m}$  de comprimento. Ela se caracteriza por um cinetoplasto muito proeminente na parte posterior da célula. Além disso, ela possui um flagelo e uma membrana ondulante em toda a extensão lateral do parasito que emerge na região posterior do corpo celular. A forma epimastigota (figura 1.3B) pode apresentar de 20 a 40  $\mu\text{m}$  de comprimento, é fusiforme e o cinetoplasto se encontra junto ao núcleo. A membrana ondulante é curta e o flagelo, bem desenvolvido, se torna livre na região anterior do parasito. Os amastigotas (figura 1.3C), com 4  $\mu\text{m}$  de diâmetro, são caracterizados pelo formato arredondado, flagelo interno e ausência de membrana ondulante. O cinetoplasto é bem visível, sendo um pouco menor que o núcleo, e se localiza entre o flagelo e o núcleo. (Rey, 2002; Teixeira *et al.*, 2011).

### 1.3. *T. cruzi* e seu ciclo de vida

O ciclo de vida do *T. cruzi* (figura 1.4) é bastante complexo e possui, como visto anteriormente, diferentes estágios de desenvolvimento em dois organismos hospedeiros: um hospedeiro mamífero, podendo este ser o homem, e um hospedeiro invertebrado que é o inseto triatomíneo hematófago (barbeiro). Esses insetos se tornam infectados através da ingestão do sangue de mamíferos contaminados com as formas tripomastigotas circulantes do *T. cruzi*. Uma vez ingeridos, esses parasitos são direcionados para o intestino médio do inseto e se diferenciam nas formas epimastigotas não infectantes que têm a capacidade de se dividir por fissão binária. Na porção final do intestino, os parasitos se diferenciam novamente nas formas tripomastigotas metacíclicas não replicativas. A infecção no mamífero pode ocorrer durante o repasto sanguíneo do inseto infectado, que logo após sugar o sangue de um indivíduo, elimina fezes sobre a região da picada contendo as formas tripomastigotas metacíclicas infectantes do *T. cruzi*. Essas formas possuem alta motilidade e são capazes de invadir as células do hospedeiro vertebrado por meio de lesões na pele ou por meio da mucosa, incluindo o trato digestivo. A invasão da célula hospedeira leva a formação de um vacúolo endocítico, denominado vacúolo parasitóforo, que conterá o tripomastigota invasor. Esse vacúolo, após sucessivas fusões com lisossomas recrutados pela célula hospedeira, se tornará mais ácido, o que induzirá a produção de enzimas pelo *T. cruzi* que levarão à ruptura da membrana vacuolar e a liberação do parasito no citoplasma. No

citoplasma, os tripomastigotas diminuem de tamanho e mudam sua forma se transformando nas formas amastigotas. Os amastigotas se multiplicam por divisão binária até a célula ficar repleta destas formas que, então, se diferenciam novamente nas formas tripomastigotas sanguíneas pelo crescimento do flagelo. Os tripomastigotas rompem a célula, caem na corrente sanguínea através do qual alcançam diferentes tipos de células, principalmente células cardíacas e dos músculos lisos e esqueléticos. Além de invadir novas células, essas formas tripomastigotas podem ser ingeridas pelos insetos triatomíneos durante o repasto sanguíneo, fechando o ciclo (Lima *et al.*, 2010; de Souza *et al.*, 2010).

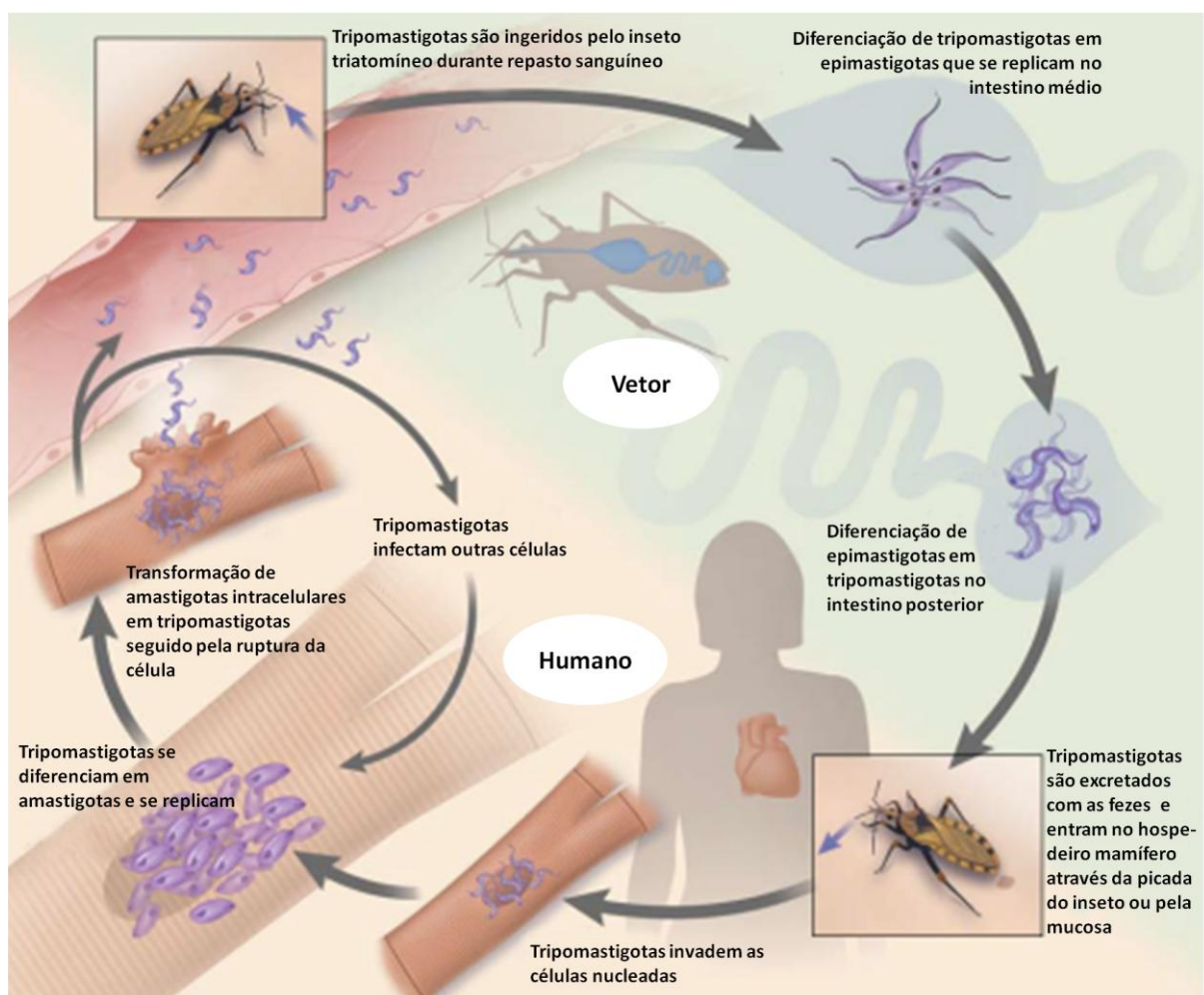


Figura 1.4: Ciclo evolutivo de *T. cruzi*. Adaptada de Bern, 2011.

#### 1.4. Tratamento da doença de Chagas

Desde o final da década de 60, dois medicamentos têm sido usados no tratamento da doença de Chagas: benznidazol que é capaz de modificar

covalentemente algumas macromoléculas do *T. cruzi*, e o nifurtimox, que atua produzindo radicais livres danosos a esse parasito (Polak & Richle, 1978; Guedes *et al.*, 2011). Ambas as drogas são mais ativas na fase aguda da doença, apresentando uma baixa eficiência na fase crônica. Além disso, outras desvantagens podem ser observadas com a administração desses medicamentos, como uma alta toxicidade, levando a graves efeitos colaterais; variações na resposta ao tratamento entre indivíduos com idades diferentes e de regiões endêmicas distintas; e alterações na susceptibilidade do *T. cruzi* aos medicamentos principalmente pela diversidade genética dos parasitos ao longo das áreas endêmicas. Associado a esses problemas devemos ressaltar a necessidade da administração dessas drogas por longos períodos de tempo e o consequente alto custo do tratamento (Steindel *et al.*, 2005; Neves *et al.*, 2005).

Entre reações adversas desses dois medicamentos estão manifestações digestivas, náuseas, vômitos, diarreia, perda de peso, anorexia, alterações psíquicas, hipersensibilidade, sonolência ou excitação, dermatite, linfadenopatia. Devido à gravidade desses efeitos colaterais, desde 1980 o nifurtimox teve sua comercialização descontinuada em alguns países, entre eles o Brasil, enquanto que o benznidazol continua a ser prescrito na fase aguda da doença (Coura & Castro, 2002).

A eficácia limitada desses medicamentos e o fato de a doença de Chagas ter se tornado uma doença de alcance mundial tem aumentado a urgência no desenvolvimento de fármacos que sejam mais potentes no combate à doença de Chagas. Entre os pré-requisitos de uma droga ideal está a capacidade de atuar em todas as fases da doença, principalmente no que se refere ao progresso das formas crônicas, com pouco ou nenhum efeito colateral (Bustamante & Tarleton, 2011).

Nos últimos anos, várias abordagens têm sido empregadas com esse objetivo, entre elas a que se baseia nos dados obtidos a partir do genoma de *T. cruzi*, visando um planejamento racional de fármacos. Nesse contexto, a bioinformática assume papel importante na análise das sequências genômicas, fornecendo informações sobre os possíveis genes ali codificados, sobre suas proteínas correspondentes e suas possíveis funções no metabolismo celular. Esse conjunto de informações permite uma maior compreensão da bioquímica do parasito e uma inferência final de potenciais alvos terapêuticos para o desenvolvimento de drogas (Kramer & Cohen, 2004).

Dentro da bioinformática, a genômica comparativa tem crescido e possibilita a identificação de genes conservados e específicos de relevância funcional em genomas de espécies próximas e/ou distantes (Glaucius, 2008). A partir desse contexto, o nosso laboratório desenvolveu uma nova abordagem computacional, o AnEnPi, que possui uma estratégia de classificação funcional de enzimas baseada em análises de similaridades entre sequências, permitindo a identificação de proteínas análogas, homólogas e específicas entre diferentes organismos, entre eles o *T. cruzi* e os seres humanos (Otto *et al.*, 2008).

### 1.5. AnEnPi (Analogous Enzyme Pipeline)

O AnEnPi é uma ferramenta computacional que permite a identificação de enzimas homólogas, análogas e específicas através do agrupamento das estruturas primárias de proteínas de acordo com suas atividades enzimáticas. Entre as funcionalidades do AnEnPi está também a inferência de funções utilizando buscas por similaridade via *BLAST* (Basic Local Alignment Search Tool) (<http://www.ncbi.nlm.nih.gov/BLAST/>) e/ ou *HMMER* (<http://hmmer.wustl.edu/>) e a reconstrução de vias metabólicas. Essa ferramenta foi utilizada neste estudo para inferir atividades enzimáticas nos genomas de organismos como o *T. cruzi* e selecionar as atividades que possuem papel relevante no metabolismo desse parasito (Otto *et al.*, 2008; Guimarães *et al.*, 2008; Alves-Ferreira *et al.*, 2009; Capriles *et al.*, 2010).

A metodologia utilizada pelo AnEnPi (figura 1.5) baseia-se na comparação de sequências protéicas de todos os organismos anotadas pelo KEGG (Kyoto Encyclopedia of Genes and Genomes), excluindo sequências pequenas (menos que 100 aminoácidos). Cada uma das sequências é comparada individualmente com todas as outras proteínas anotadas como possuindo a mesma atividade enzimática, - definida através de seu número EC (*Enzyme Commission Numbers*) - através do programa *blastp*, sendo feitas comparações par-a-par, todas contra todas. Os índices de similaridade obtidos após essas comparações são utilizados para a construção de *clusters* (grupos) através da aplicação de pontos de corte nos quais pares com uma pontuação (*score*) maior que 120 (*e-value* <0,0001) são reunidas em um único grupo. Sequências protéicas alocadas dentro de um mesmo grupo são consideradas homólogas enquanto que sequências protéicas anotadas como

possuindo a mesma atividade enzimática, mas localizadas em grupos diferentes são consideradas análogas (Otto *et al.*, 2008).

Enzimas homólogas são enzimas que podem ou não desempenhar a mesma atividade enzimática e possuem estruturas primárias similares entre si, devido a mesma origem evolutiva. Em muitos casos, as enzimas homólogas possuem estrutura tridimensional semelhante. Enzimas análogas são enzimas que possuem a mesma função enzimática, mas são oriundas de ancestrais diferentes, o que reflete pouca ou nenhuma similaridade entre as suas estruturas primárias e provavelmente estruturas tridimensionais distintas (Fitch, 1970).

A partir da formação de grupos para cada número EC, o AnEnPi pode ser usado também para comparar todas as sequências protéicas do genoma de um organismo contra os grupos previamente obtidos (através do *BLAST* e/ou *HMMER*). Essa comparação das sequências de um genoma contra as sequências dentro dos grupos permite a identificação de atividades enzimáticas no genoma do organismo analisado, podendo incluir genes até então não anotados no genoma em questão (Otto *et al.*, 2008).

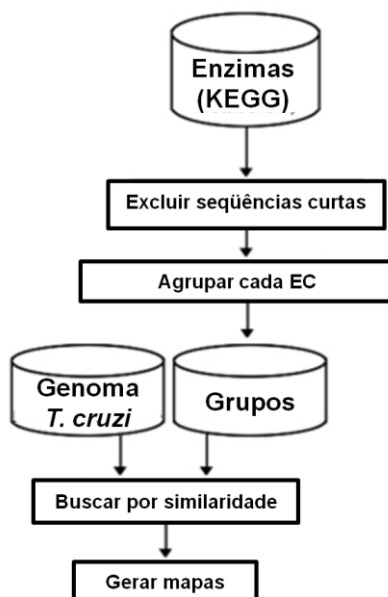
Para a reconstrução das vias metabólicas do *T. cruzi*, o AnEnPi, como descrito acima, primeiramente agrupou as proteínas de mesma atividade enzimática de acordo com a similaridade a partir das sequências protéicas de vários organismos depositadas no KEGG, formando grupos para cada número EC. Depois disso, foi realizada uma análise de similaridade por meio do programa *blastp*, utilizando como sequências a serem analisadas (*query*) o genoma de *T. cruzi* contra todas as sequências de cada grupo previamente obtido.

Utilizando essa abordagem pós genômica, o AnEnPi identificou as enzimas gliconato cinase (EC 2.7.1.12) pertencente à via das pentoses-fosfato e a uracil fosforribosiltransferase (EC 2.4.2.9) pertencente à via de salvação de pirimidinas. Pelo fato de a gliconato cinase não ter sido identificada em humano pelo AnEnPi, essa enzima foi classificada como específica de *T. cruzi* em relação a este hospedeiro. Já a uracil fosforribosiltransferase foi identificada em humano e alocada, pelo AnEnPi, no mesmo grupo que a uracil fosforribosiltransferase de *T. cruzi*, sendo essas enzimas, portanto, classificadas como homólogas entre si.

Uma vez que as enzimas gliconato cinase e a uracil fosforribosiltransferase participam de vias metabólicas relevantes de *T. cruzi*, a caracterização funcional e estrutural dessas enzimas é importante para uma melhor compreensão do



metabolismo desse parasito e para o conseqüente fornecimento de subsídios ao desenvolvimento de novos fármacos.



**Figura 1.5:** Fluxograma do funcionamento do programa AnEnPi. As estruturas primárias das proteínas depositadas no KEGG são agrupadas de acordo com EC. Os grupos formados são submetidos à análise por busca de similaridade com o genoma alvo gerando dados para a reconstrução dos mapas metabólicos. Adaptada do artigo Otto *et al*, 2008.

## 1.6. Via das pentoses-fosfato

A via das pentoses-fosfato tem como função a produção de dois compostos fundamentais ao metabolismo celular: ribose-5-fosfato e NADPH. A ribose-5-fosfato é um precursor essencial para a síntese de nucleotídeos e ácidos nucleicos, entre eles coenzimas, como ATP, NADH, FADH<sub>2</sub>, importantes principalmente em células em divisão. Já o NADPH atua como doador de elétrons, fornecendo a força redutora necessária às reações biossintéticas. Por manter um ambiente redutor na célula, o NADPH pode prevenir ou recuperar moléculas sensíveis, como proteínas e lipídeos, de danos oxidativos gerados pelas espécies reativas de oxigênio (Nelson & Cox, 2011).

A via das pentoses-fosfato (figura 1.6) é encontrada em organismos procariontes e eucariontes e pode ser dividida em duas etapas, também chamadas de ramos: uma oxidativa e outra não oxidativa. Essa via metabólica geralmente se inicia com a glicose-6-fosfato, que também pode ser metabolizada pela via glicolítica a piruvato. A entrada de glicose-6-fosfato na via glicolítica ou na via das pentoses-

fosfato é basicamente determinada pelas concentrações relativas de  $\text{NADP}^+$  e  $\text{NADPH}$  (Nelson & Cox, 2011).

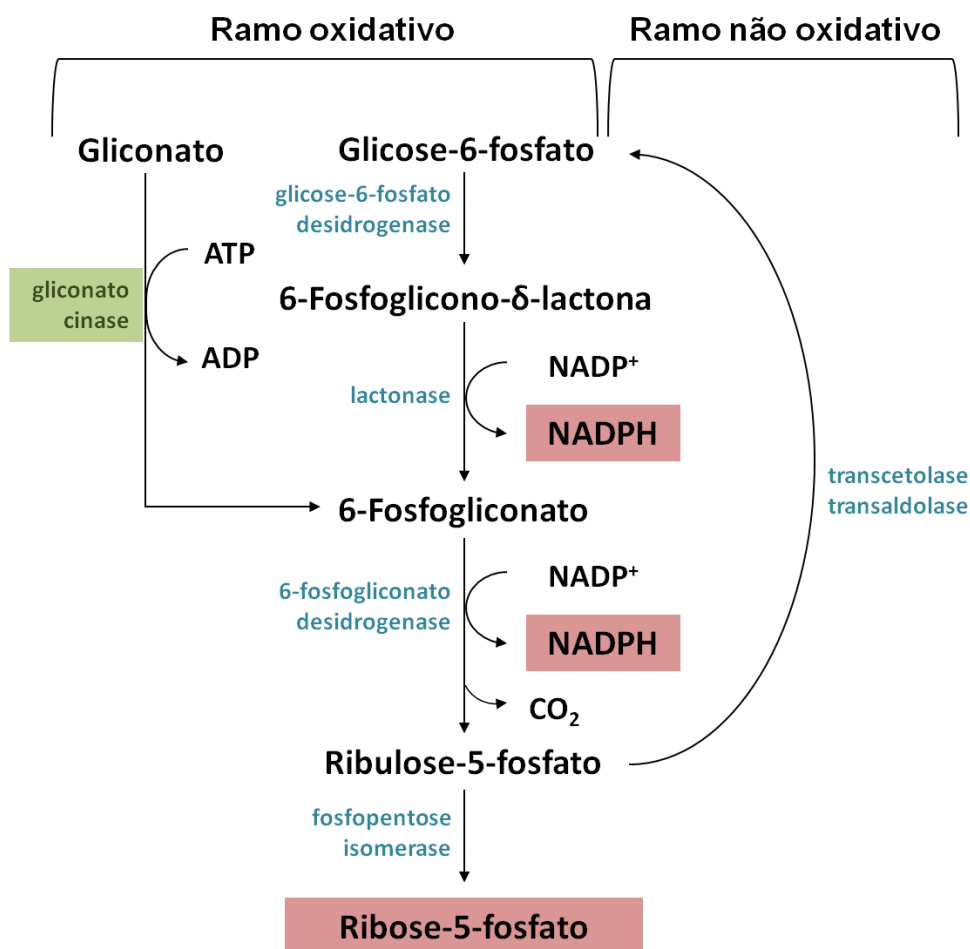
A primeira etapa da via das pentoses-fosfato envolve a oxidação de glicose-6 fosfato pela glicose-6-fosfato desidrogenase, reduzindo  $\text{NADP}^+$ . O produto dessa reação é hidrolisado a 6-fosfogliconato que, então, sofre oxidação e descarboxilação pela 6-fosfogliconato desidrogenase, levando a formação final de ribulose-5-fosfato e de mais uma molécula de  $\text{NADPH}$ . A ribulose-5-fosfato é convertida por uma isomerase a ribose-5-fosfato, que será utilizada nas reações de síntese de nucleotídeos (Nelson & Cox, 2011).

A segunda etapa da via das pentoses-fosfato envolve reações não oxidativas que convertem pentoses-fosfato a frutose-6-fosfato e a gliceraldeído-3-fosfato ou a glicose-6-fosfato. Para a recuperação dos produtos finais do ramo oxidativo, as enzimas transcetolase e a transaldolase catalisam a conversão reversível de seis pentoses-fosfato a cinco hexoses-fosfato. Essa etapa não oxidativa está relacionada a metabolismos celulares que requerem  $\text{NADPH}$ , mas não as pentoses-fosfato geradas, reciclando-as a frutose-6-fosfato e gliceraldeído-3-fosfato ou a glicose-6-fosfato. A frutose-6-fosfato e gliceraldeído-3-fosfato podem entrar na via glicolítica para produção de energia. Já a glicose-6-fosfato é capaz de reiniciar o ciclo das pentoses-fostato (Nelson & Cox, 2011).

Em tripanosomatídeos, a via das pentoses-fosfato tem sido largamente estudada e, assim como na maioria dos organismos, as enzimas a ela relacionadas estão localizadas essencialmente no citoplasma e em alguns casos nos glicossomos, organelas características desses parasitos (Barrett, 1997; Eicks *et al.*, 2002). Nas formas tripomastigotas de *Trypanosoma brucei* e nas formas promastigotas de *Leishmania mexicana*, demonstrou-se que todas as enzimas dessa via estão presentes, tendo sido algumas delas caracterizadas (Cronín *et al.*, 1989; Maugeri *et al.*, 2003). Do mesmo modo, essa via metabólica já foi bem descrita em *Leishmania donovani* (Ghosh, 1971). Em *T. cruzi*, foi confirmada a presença das enzimas da via das pentoses-fosfato em todas as formas evolutivas do parasito, tendo sido a maioria delas amplamente estudadas (Maugeri & Cazzulo, 2004).

Além disso, em *T. cruzi*, o programa AnEnPi identificou a presença de dois genes potencialmente codificantes para a enzima específica gliconato cinase que já foi descrita e caracterizada em outros organismos, como nas bactérias *Escherichia coli* (Izu *et al.*, 1996) e em *Gluconobacter oxydans* (Rauch *et al.*, 2010). Essa enzima

é capaz de catalisar a reação de fosforilação, em presença de ATP, de gliconato a 6-fosfogliconato, um dos principais intermediários do ramo oxidativo da via das pentoses-fosfato. Dessa forma, o gliconato poderia ser utilizado diretamente pela via das pentoses-fosfato, após ser fosforilado pela gliconato cinase, caracterizando assim uma rota alternativa de entrada de carbono nessa via em adição a rota tradicional mediada pela glicose-6-fosfato (figura 1.6).



**Figura 1.6:** Via das pentoses-fosfato. Os principais produtos dessa via metabólica estão coloridos em rosa e as enzimas envolvidas na via em azul, com o destaque para a enzima gliconato cinase em verde. Sigla: ATP, trifosfato de adenosina; ADP, adenosina difosfato; NADPH, dinucleotídeo de nicotinamida-adenina-fosfato. Adaptada de Nelson & Cox, 2011.

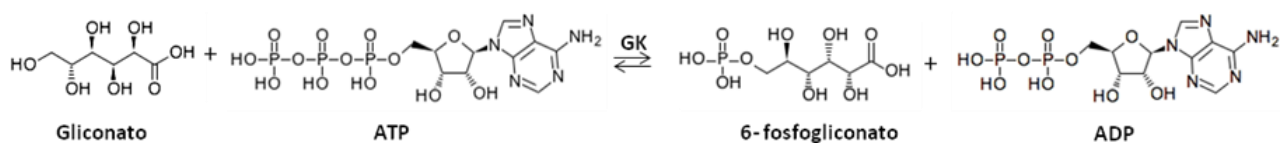
A importância dessa via metabólica nesses parasitos, e conseqüentemente, o interesse no estudo mais aprofundado dela, está relacionada à essencialidade da produção, já descrita anteriormente, de ribose-5-fosfato e, principalmente, ao poder redutor gerado na forma de NADPH. Em tripanosomatídeos, a presença de NADPH é fundamental para que a tripanotona redutase mantenha reduzida a tripanotona,

um reagente anti-oxidante característico dessa família. A tripanotiona juntamente com NADPH é o principal mecanismo de defesa do parasito em resposta ao estresse gerado pelas espécies reativas de oxigênio (Augustyns *et al.*, 2001).

O papel chave da via das pentoses-fosfato na proteção do *T. cruzi* contra agentes oxidantes é indicado principalmente pelos altos níveis de atividade enzimática dos componentes dessa via nas formas evolutivas presentes no mamífero hospedeiro, já que é nesse momento que o parasito é exposto a um alto estresse oxidativo. Isso se torna particularmente importante nas formas tripomastigotas metacíclicas e amastigotas localizadas dentro de macrófagos, um ambiente em que se tem alta concentração de radicais livres liberados pelo próprio macrófago para matar o parasito intracelular (Maugeri *et al.*, 2003). Dessa forma, o protozoário necessitaria de um grande fluxo de glicose ou gliconato através da via das pentoses-fosfato de forma a produzir NADPH e assim sobreviver ao estresse oxidativo. Por outro lado, a produção de ribose-5-fosfato é muito importante para as formas replicativas epimastigotas e amastigota.

### 1.6.1. Gliconato cinase (GK)

A gliconato cinase (GK, EC 2.7.1.12) é uma enzima que participa de uma parte inicial alternativa da via das pentoses-fosfato e catalisa a reação de fosforilação, em presença de ATP, gerando 6-fosfogliconato a partir de gliconato (figura 1.7).



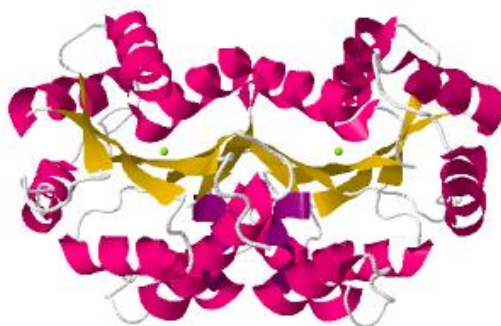
**Figura 1.7:** Reação catalisada pela enzima GK. Siglas: ATP, trifosfato de adenosina; ADP, difosfato de adenosina. Adaptada de <http://www.genome.jp/Fig/reaction/R01737.gif>.

Embora essa enzima tenha sido pouco estudada, ela já foi caracterizada em micro-organismos como *Pseudomonas* (Coffee & Hu, 1972). *E. coli* (Izu *et al.*, 1996), *Corynebacterium glutamicum* (Porco *et al.*, 2008) e em *Gluconobacter oxydans* (Rauch *et al.*, 2010). Em eucariotos, especialmente em mamíferos, a GK tem sido pouco estudada, embora sua atividade tenha sido descrita em fígado de porco

(Leder, 1957). Em *T. cruzi*, o programa AnEnPi identificou a presença de dois genes potencialmente codificantes para a enzima GK. As anotações preditas pelo KEGG e pelo TriTrypDB para ambas as sequências também identificam-nas como putativas GK.

Até o momento somente a cristalografia de raios-x da GK de *E. coli* foi relatada (Kraft *et al.*, 2002) (figura 1.8), embora estruturas de GKs de outros organismos, como *Lactobacillus acidophilus*, *Sinorhizobium meliloti* e *Janibacter sp.* também estejam disponíveis no PDB (<http://www.rcsb.org/pdb/results/results.do?grid=990A689E&tabtoshow=Current>). Entre as características descritas na estrutura de GK de *E. coli* está uma topologia similar às nucleosídeo monofosfato (NMP) cinases, embora o sítio de ligação em GK seja localizado em um posição diferente ao sítio de ligação de nucleosídeos dessa NMP cinases. Ainda nessa estrutura, uma cadeia lateral conservada de lisina 21 é responsável por interagir com o grupo fosfato, promovendo a fosforilação. O 6-fosfogliconato se liga com seu grupo fosfato em uma posição similar ao fosfato do ATP (Kraft *et al.*, 2002).

Além disso, a estrutura quaternária de GK em *E. coli* se comporta em solução como um dímero composto por duas subunidades idênticas, parecendo ser a forma ativa cataliticamente dessa enzima (Kraft *et al.*, 2002). Em *Zymomonas mobilis* (Zachariou & Scopes, 1985) e em *Gluconobacter oxydans* (Rauch *et al.*, 2010), GK é descrita como sendo um monômero.



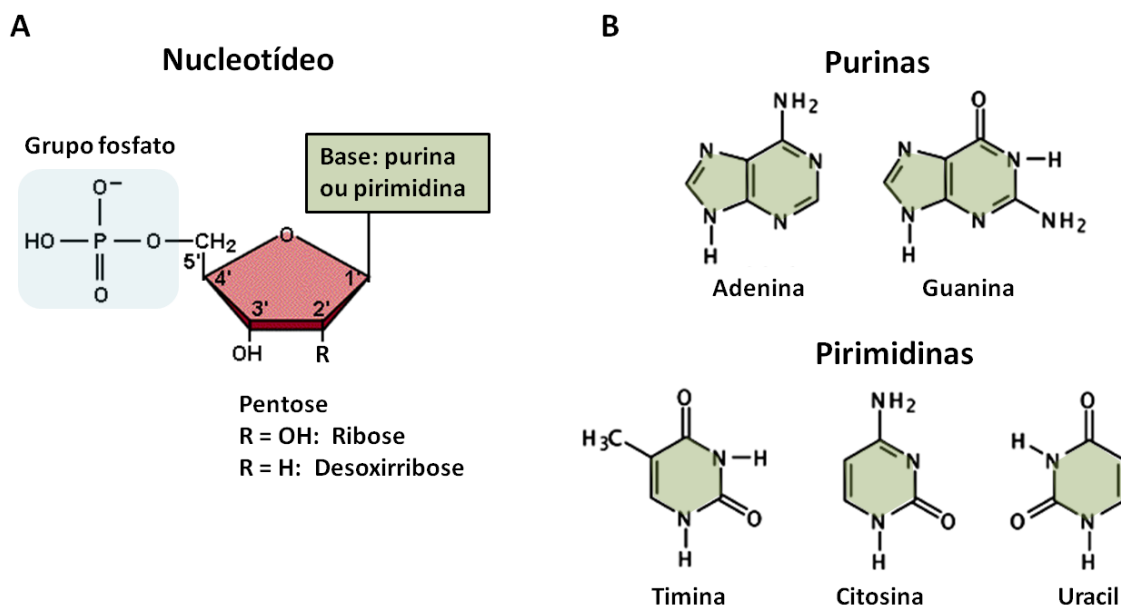
**Figura 1.8:** Estrutura tridimensional de GK de *E. coli* resolvida por difração de raios-x. Resolução de 2,80 Å (PDBid: 1KOF).

## 1.7. Metabolismo de nucleotídeos

Os nucleotídeos participam de diversos processos bioquímicos fundamentais à manutenção do metabolismo celular, entre eles os anabólicos. A importância

dessas moléculas no contexto celular está relacionada à sua capacidade de armazenar grande quantidade de energia em suas ligações químicas, como no caso do ATP, GTP e outros nucleotídeos fosforilados. Além disso, os nucleotídeos são os componentes estruturais dos ácidos nucleicos e de cofatores enzimáticos, como NAD<sup>+</sup> e FAD, e podem atuar como intermediários biossintéticos ativados (Nelson & Cox, 2011).

Estruturalmente, os nucleotídeos (figura 1.9A) são formados por uma base nitrogenada, uma pentose e um grupamento fosfato. A molécula sem o grupamento fosfato é conhecida como nucleosídeo. A base nitrogenada é um composto cíclico que contém nitrogênio e pode ser classificada em pirimidinas e purinas (figura 1.9B). As pirimidinas, como a timina (T), citosina (C) e uracil (U), possuem um anel heterocíclico simples de seis átomos, no qual dois átomos de nitrogênio substituem o carbono nas posições 1 e 3. As bases púricas, representadas pela adenina (A) e a guanina (G), apresentam um anel duplo composto por um anel de pirimidina fundido a um anel imidazólico (Voet & Voet, 2011).



**Figura 1.9:** Estrutura de nucleotídeos. **(A)** Estrutura geral mostrando a convenção numérica do anel de pentose, que pode ser uma ribose, com radical R igual ao grupo –OH, ou uma desoxirribose, com radical R igual a –H. O grupo fosfato está colorido em azul, a pentose está colorida em rosa e a base nitrogenada em verde. Adaptada de <http://www.biology.iupui.edu/biocourses/N100/images/3nucleotide.gif>. **(B)** Principais bases púricas e pirimídicas coloridas em verde. Adaptada de <http://toxamb.pharmacy.arizona.edu/images/f1-1-1-d.gif>

São conhecidos dois tipos de vias responsáveis pela síntese de nucleotídeos: a vias *de novo* e as vias de salvação. Essas vias não estão correlacionadas e são

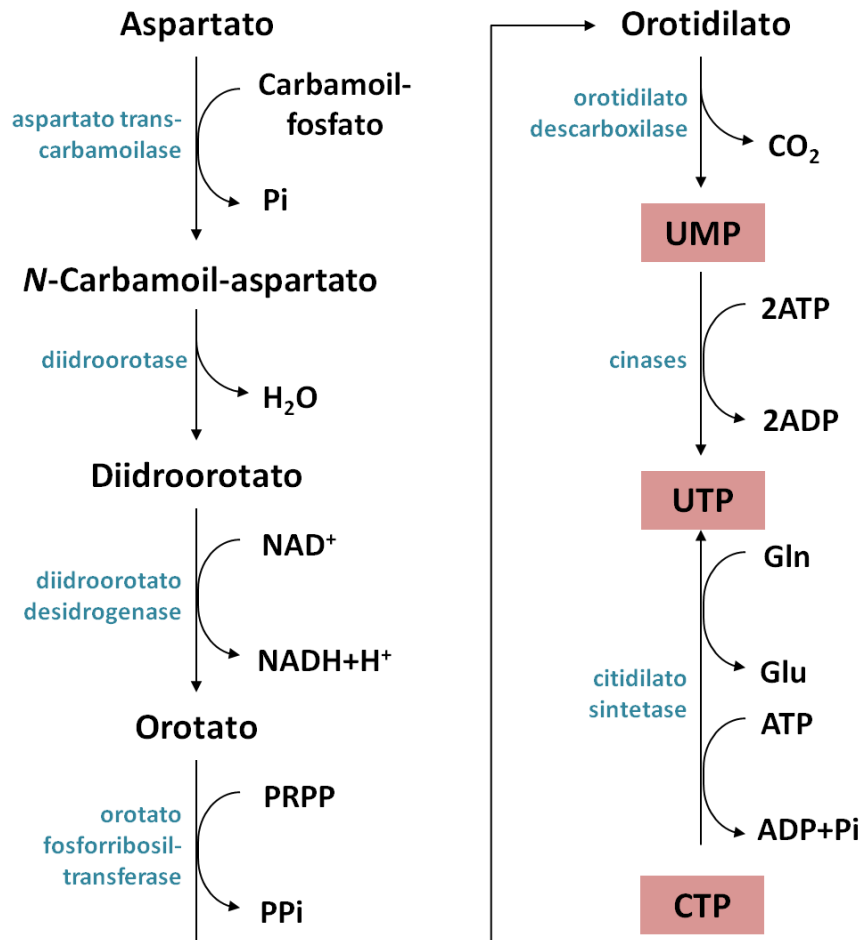
independentes uma da outra principalmente no que se refere a seus mecanismos e em sua regulação. A via de síntese *de novo* dos nucleotídeos inicia com precursores metabólicos simples como aminoácidos, ribose-5-fosfato,  $\text{CO}_2$  e  $\text{NH}_3$ . O que é importante de se notar é que as bases livres guanina, adenina, timina, citidina e uracil não são intermediários nas vias *de novo*, ao invés disso, a estrutura dos anéis das bases, provenientes dos aminoácidos, é construída e ligada à ribose durante o processo. Já as vias de salvação reciclam os produtos liberados a partir da degradação de ácidos nucleicos, as bases livres e os nucleosídeos, usando-os como precursores para a síntese de nucleotídeos (Nelson & Cox, 2011). A coexistência dessas duas vias em diversos organismos salienta a importância da síntese de nucleotídeos nos processos bioquímicos.

### 1.7.1. Síntese *de novo* das pirimidinas

Os nucleotídeos de pirimidina podem ser representados principalmente pelos compostos monofosfato de citidina (CMP) e monofosfato de uridina (UMP), os quais contêm as bases nitrogenadas citosina e uracil. A biossíntese *de novo* dos nucleotídeos de pirimidina (figura 1.9) é obtida através de uma série de reações enzimáticas e ocorre a partir de aspartato, carbamoil-fosfato e PRPP (D-5-fosforibosil-1-fosfato). O precursor do PRPP é a ribose-5-fosfato, obtida da via das pentoses-fosfato, e de ATP, em reação catalisada pela enzima PRPP sintase (Nelson & Cox, 2011).

A via começa com a formação de carbamoil-fosfato a partir do grupo amino da glutamina, carbonato e duas moléculas de ATP, em uma reação catalisada pela carbamoil-fosfato sintetase II. Em seguida, o carbamoil-fosfato se condensa ao aspartato pela ação da aspartato transcarbamoilase, formando *N*-carbamoil-aspartato. A ciclização do *N*-carbamoil-aspartato é catalisada pela enzima diidroorotase e produz o anel de pirimidina. O diidroorotato formado é irreversivelmente oxidado a orotato, que então é condensado a PRPP, pela ação da orotato fosforibosiltransferase, produzindo orotidilato (OMP). Esse composto sofre descarboxilação para formação final de UMP, pela ação da orotidilato descarboxilase. O nucleotídeo UMP formado pode então ser fosforilado a UDP e UTP, que nessa via, juntamente com a glutamina que atua como doador de nitrogênio é o precursor de CTP pela ação da citidilato sintetase, assim como de

outros nucleotídeos de pirimidina (figura 1.10) (Voet & Voet, 2011; Nelson & Cox, 2011).



**Figura 1.10:** Síntese *de novo* das pirimidinas. Biossíntese de UTP e CTP via orotidilato. Os principais produtos dessa via metabólica estão coloridos em rosa e as enzimas em azul. Siglas: Pi, fosfato inorgânico; PPi, pirofosfato; NAD<sup>+</sup>, dinucleotídeo de nicotinamida-adenina oxidado; NADH+H<sup>+</sup>, dinucleotídeo de nicotinamida-adenina reduzido; PRPP, D-5-fosforibosil-1-fosfato; ATP, trifosfato de adenosina; ADP, difosfato de adenosina; UMP, monofosfato de uridina; UTP, trifosfato de uridina; CTP, trifosfato de citidina; Gln, glutamina; Glu, ácido glutâmico. Adaptada de Nelson & Cox, 2011.

A partir de estudos da via de síntese *de novo* de pirimidina em bactérias, em fungos e em mamíferos, tem-se demonstrado que essa via pode ser universal, apresentando os mesmos passos sequenciais de reações enzimáticas em todos os organismos (O'Donovan & Neuhard, 1970). No entanto, algumas diferenças podem ser notadas entre eles como, por exemplo, nos mamíferos, em que os três passos iniciais são catalisados por uma única enzima multifuncional chamada carbamoil-fosfato sintetase-aspartato transcarbamoilase-diidroorotase (Iwahana *et al.*, 1996). Em *E. coli* e em outras bactérias, ao contrário, cada reação enzimática é catalisada por uma proteína diferente (Neuhard *et al.*, 1987).



Em tripanosomatídeos, essa via também parece estar ativa, uma vez que todas as seis enzimas para a síntese *de novo* de pirimidina foram identificadas a partir do extrato total de *Leishmania amazonensis* (Hammond & Gutteridge, 1982), de *T. brucei* (Hammond & Gutteridge, 1982) e de *T. cruzi* (Gao *et al.*, 1999). Atualmente, já se tem conhecimento de que o gênero *Leishmania* como um todo possui a maquinaria enzimática necessária a essa via. Em *L. donovani*, observou-se que um nocaute<sup>1</sup> da primeira enzima da via *de novo* pode ser letal para a forma promastigota em meio de cultura livre de pirimidinas exógenas, que é o substrato da via de salvação pela uracil fosforribosiltransferase. Na presença dessas pirimidinas, o nocaute não tem efeito devido ao suprimento de nucleotídeos conseguido alternativamente pela via de salvação (Wilson *et al.*, 2012). Já em *Toxoplasma gondii*, além de terem sido identificadas as enzimas da via de síntese *de novo* das pirimidinas, foram encontradas evidências de que essa via pode estar relacionada com a virulência desse parasito (Fox & Bzik, 2002).

### 1.7.2. Via de salvação das pirimidinas

As bases nitrogenadas e nucleosídeos livres obtidos do meio extracelular, ou a partir da degradação de nucleotídeos, podem ser reutilizados para sintetizar novos nucleotídeos de pirimidina ou de purina por meio da via de salvação de nucleotídeos. Essa via, por ser muito mais simples e energeticamente econômica do que a síntese *de novo* de nucleotídeos, é considerada uma estratégia celular de reciclagem de nucleosídeos e bases livres, visando a manutenção do aporte necessário de nucleotídeos em contrabalanço a degradação deles (Zrenner *et al.*, 2006).

A via de salvação de pirimidinas (figura 1.11) pode ocorrer através de diversas rotas metabólicas, variando de acordo com o substrato inicial e a maquinaria enzimática dessa via presente no organismo em questão. Em uma dessas rotas, a salvação de nucleotídeo pode ocorrer pela conversão direta de bases uracil livres no seu nucleotídeo correspondente (UMP), pela ação da enzima uracil fosforribosiltransferase em presença de PRPP. No entanto, a salvação de pirimidinas a partir da base citosina não é direta e ocorre através da deaminação a

---

<sup>1</sup> Nocaute (do inglês, *knockout*) é a inativação de um gene específico através de uma mutação natural ou por meio de técnicas de inserção aleatória ou direcionada de um fragmento de DNA, levando a formação de uma versão não funcional desse gene (Griffithis *et al.*, 2008).

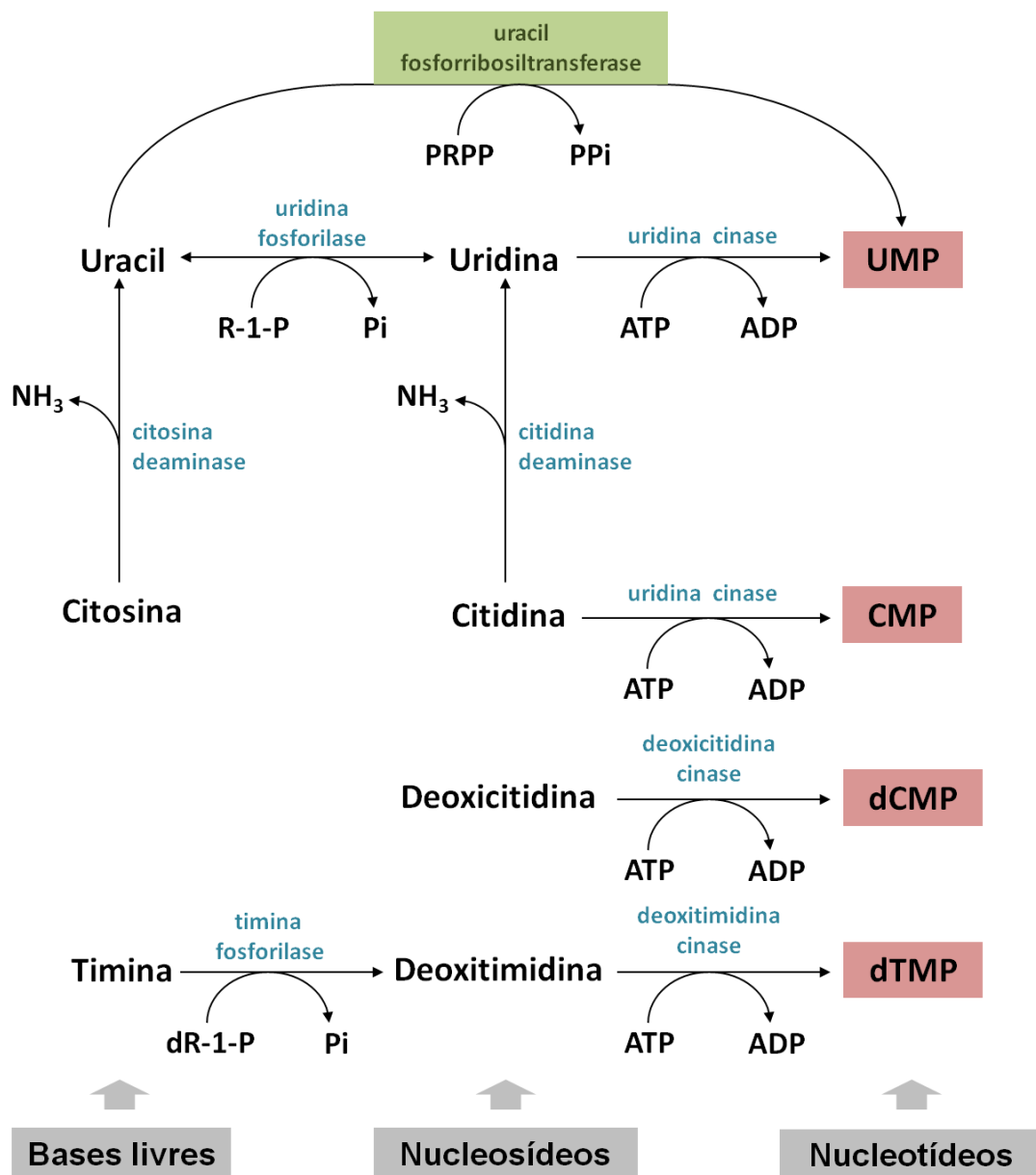
amônia e uracil, pela citosina deaminase, e a posterior conversão do uracil em nucleotídeo. A via de salvação das bases pirimidínicas pode ocorrer também pela conversão de bases nitrogenadas livres em seus respectivos nucleosídeos por meio de nucleosídeo fosforilases, seguida pela conversão dos nucleosídeos em nucleotídeos pela ação de nucleosídeo cinases específicas (Hyde, 2007).

Já na salvação dos nucleosídeos de pirimidinas livres, como a uridina, citidina e deoxitimidina, eles são convertidos aos seus respectivos nucleotídeos pela ação de nucleosídeo cinases, como a uridina cinase e a deoxitimidina cinase. No caso da citidina, ela pode ser deaminada a uridina pela citidina deaminase e, então, essa uridina é diretamente convertida a UMP pela uridina cinase. A citidina pode também ser fosforilada a CMP pela uridina cinase. Outra rota possível dos nucleosídeos é, por exemplo, a clivagem da uridina a uracil e ribose-1-fosfato através da uridina fosforilase e posterior conversão do uracil ao UMP pela uracil fosforribosiltransferase caracterizadas (Wilson *et al.*, 2012).

A via de salvação de pirimidinas pode estar presente em organismos procariotos e eucariotos, abrangendo de bactérias a mamíferos. Diferentes estudos têm demonstrado que a via de salvação de pirimidinas é vital para as células bacterianas, podendo essa via possuir diferenças entre as espécies, uma vez que várias rotas de salvação de pirimidinas são descritas. Por exemplo, a bactéria *Pseudomonas aeruginosa* possui a citosina deaminase, mas não possui a citidina deaminase e a uridina fosforilase, enquanto que a *Pseudomonas mendocina* possui a citidina deaminase, a citosina deaminase e a uridina fosforilase (Beck & O'Donovan, 2008). Já em *E. coli* foram descritas as enzimas uridina fosforilase, uridina cinase e timidina cinase, além da uracil fosforribosiltransferase (O'Donovan & Neuhard, 1970; Andersen *et al.*, 1992). Em humano, a uracil fosforribosiltransferase foi caracterizada, mas não foi identificada a sua atividade catalítica (Li *et al.*, 2007); apesar disso, o UMP pode ser obtido pela via de salvação a partir da atividade da uracil fosforilase (Renck *et al.*, 2010), seguida pela reação catalisada pela uridina cinase (Suzuki *et al.*, 2004).

Em estudos realizados com algumas espécies da família Trypanosomatidae, como *T. cruzi*, *Leishmania mexicana* e *T. brucei*, observou-se que algumas enzimas de via de salvação de UMP não estão presentes ou não foram identificadas e, conseqüentemente, não anotadas como sendo tais enzimas. A atividade da enzima uridina cinase não foi encontrada nesses organismos, sugerindo uma salvação de pirimidina por meio das enzimas uridina fosforilase e uracil fosforribosiltransferase

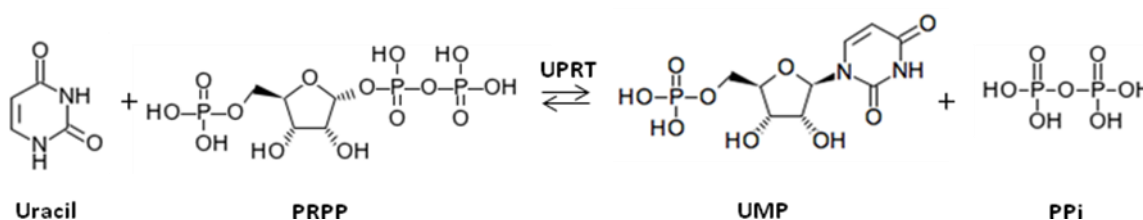
ambas identificadas (Gao *et al.*, 1999). Da mesma forma, em *T. gondii*, a principal rota de salvação de pirimidina é mediada pela uracil fosforribosiltransferase (Carter *et al.*, 1997). No gênero *Leishmania*, de um modo geral, foram identificadas as atividades das enzimas uracil fosforribosiltransferase, uracil fosforilase e citidina deaminase, tendo algumas delas sido caracterizadas (Wilson *et al.*, 2012).



**Figura 1.11:** Via de salvação de pirimidinas. Os principais produtos dessa via metabólica estão coloridos em rosa e as enzimas em azul, com destaque para a enzima uracil fosforribosiltransferase em verde. Siglas: R-1-P, ribose-1-fosfato; dR-1-P, desoxirribose-1-fosfato; PRPP, D-5-fosforibosil-1-fosfato; PPi, pirofosfato; Pi, fosfato inorgânico; ATP, trifosfato de adenosina; ADP, difosfato de adenosina; UMP, monofosfato de uridina; CMP, monofosfato de citidina; dCMP, monofosfato de desoxicidina; dTMP, monofosfato de desoxitimidina. Adaptada de Borges, 2011.

### 1.7.3. Uracil fosforribosiltransferase (UPRT)

A uracil fosforribosiltransferase (UPRT, EC 2.4.2.9) é uma enzima que participa da via de salvação de pirimidinas e catalisa a conversão de uracil e PRPP a UMP e pirofosfato (PPi) (figura 1.12). O UMP produzido é de grande importância nessa via, uma vez que ele é o precursor de todos os nucleotídeos de pirimidina. Trabalhos com *T. gondii* demonstram que a UPRT é específica para uracil e não consegue utilizar como substrato citosina, citidina, deoxicitidina, ou timina já que o tamanho dos grupos exocíclicos na posição 5 da base pirimidina impedem a ligação (Iltzsch & Tankersley, 1994; Carter *et al.*, 1997).



**Figura 1.12:** Reação catalisada pela enzima UPRT. Siglas: PRPP, D-5-fosforribosil-1-fosfato; UMP, monofosfato de uridina; PPi, pirofosfato; Pi, fosfato inorgânico. Adaptada de <http://www.genome.jp/Fig/reaction/R00966.gif>.

Essa enzima tem sido caracterizada em vários micro-organismos tais como *E. coli*, *Mycobacterium tuberculosis* (Villela *et al.*, 2013), *Lactococcus lactis* (Martinussen & Hammer, 1994), e em alguns eucariotos inferiores como *Saccharomyces cerevisiae* (Kern *et al.*, 1990), *Candida albicans* (Alloush & Kerridge, 1994), *Giardia intestinalis* (Dai *et al.*, 1995) e *T. gondii* (Carter *et al.*, 1997). Em eucariotos superiores, especialmente em mamíferos, a UPRT tem sido pouco descrita, embora ela tenha sido identificada e caracterizada em humano (Li *et al.*, 2007).

Em *T. cruzi*, o programa AnEnPi identificou por similaridade dois genes que teriam o potencial de codificar para a enzima UPRT, classificando-as como homólogas a enzima UPRT de humano. Da mesma forma, as sequências de ambos os genes também foram anotadas como sendo putativas UPRT pelo KEGG e pelo TriTrypDB. Além disso, estudos realizados na década de 80 demonstraram a atividade dessa enzima em extratos celulares das formas evolutivas epimastigota e tripomastigota de *T. cruzi*, exceto na forma amastigota em que a UPRT parece estar ausente (Hammond & Gutteridge, 1982). No entanto, trabalhos mostrando a

expressão por microarranjo de diversos genes anotados no genoma de *T. cruzi* indicam que os genes potencialmente codificantes para a enzima UPRT estariam sendo expressos também na forma amastigota desse parasito (Minning *et al.*, 2009).

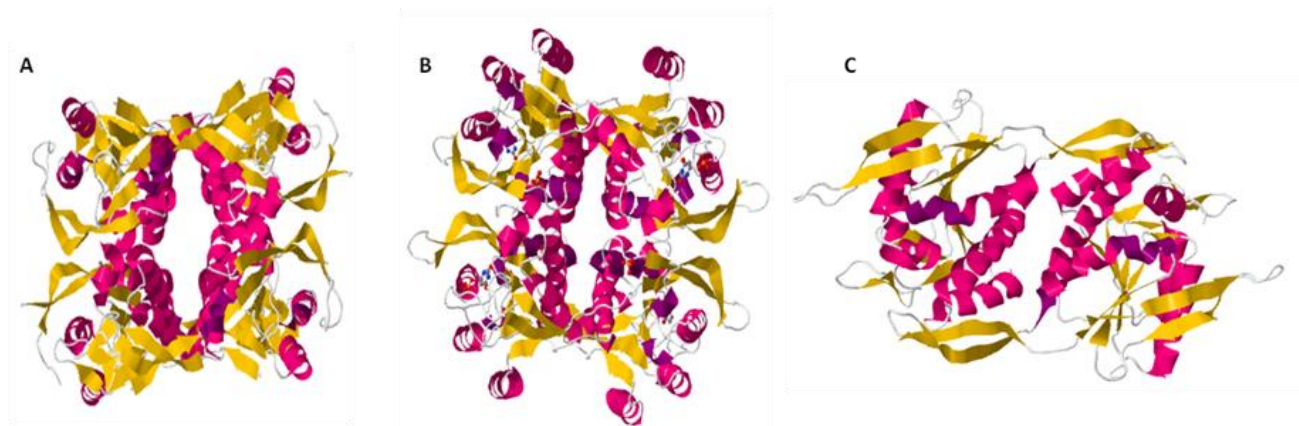
As UPRTs de diferentes organismos demonstram mecanismos de regulação distintos. Um exemplo disso é a ativação e estabilização da enzima UPRT pelo GTP, como ocorre em *E. coli* (Jensen & Mygind, 1996) e *T. gondii* (Schumacher *et al.*, 2002), enquanto que em outros organismos, como em *M. tuberculosis* (Villela *et al.*, 2013), a UPRT não é modulada pelo GTP e em muitos casos não possui ativação alostérica conhecida (Villela *et al.*, 2013). Do mesmo modo, mecanismos distintos de inibição dessa enzima também têm sido descritos, como por exemplo, a inativação por CTP, em *Sulfolobus solfataricus*, ou pelo excesso do produto UMP, em *M. tuberculosis*, além da inibição pelo UMP, dUMP, dCMP, UDP, TTP descritos em leveduras (Natalini *et al.*, 1979).

Além disso, diferentes estados oligoméricos são encontrados em UPRTs. A UPRT de *Sulfolobus solfataricus* está presente na forma de tetrâmero (Jensen *et al.*, 2005). A UPRT de *T. gondii* se comporta em solução como um dímero, enquanto que na presença de GTP a enzima assume a forma tetramérica (Schumacher *et al.*, 2002). Em *E. coli*, a UPRT é um dímero ou trímero na ausência de ligantes, enquanto que na presença de PRPP e GTP passa a ser um pentâmero ou um hexâmero com ambas as formas presentes em equilíbrio dinâmico (Rasmussen *et al.*, 1986; Jensen & Mygind, 1996). O PRPP e/ou GTP atuam na estabilização desses oligômeros. Já a UPRT de *Giardia intestinalis* (Dai *et al.*, 1995) e *Bacillus caldolyticus* (Jensen *et al.*, 1997) são enzimas diméricas.

As estruturas cristalográficas das UPRTs de *T. gondii* (Schumacher *et al.* 1998), de *B. caldolyticus* (Kadziola *et al.* 2002), de *Thermus thermophilus* (Kukimoto-Niino *et al.* 2005) e de *Sulfolobus solfataricus* (Christoffersen *et al.*, 2009) já foram determinadas (figura 1.13). Dessa forma, sabe-se que a enzima UPRT pertence ao grupo das fosforribosiltransferases tipo I caracterizadas por conterem uma região de 13 resíduos conservados que formam quatro ou cinco folhas  $\beta$  pregueadas paralelas que são o motivo de ligação de PRPP. Próximo à região C terminal da UPRT existe outra região também conservada relacionada à ligação do uracil (Schumacher *et al.* 1998).

A UPRT, quando presente, realiza um papel central na via da salvação das pirimidinas e pode ser considerada um alvo interessante para o desenho racional de fármacos contra doenças parasitárias. O potencial dessa enzima como alvo

terapêutico é reforçado por estudos recentes que demonstram que o duplo nocaute da carbamoil fosfato sintetase e de UPRT na forma amastigota de *L. donovani* consegue atenuar a infecção e eliminar o parasito em semanas. Isso demonstra que a presença de vias *de novo* e de salvação de pirimidina intactas são necessárias à sobrevivência desse parasito e a infecção causada no hospedeiro (Wilson *et al.*, 2012). Em *T. gondii*, a UPRT é capaz de se ligar a vários análogos de pirimidina que tem atividade anti-*T. gondii*. Entre esses análogos está o 5-fluoruracil que, ao ser convertido pela UPRT ao nível de nucleotídeo, se torna tóxico ao *T. gondii* (Carter *et al.*, 1997). Dessa forma, estudos mais aprofundados da UPRT em *T. cruzi* são importantes para aumentar o conhecimento sobre esta enzima e sobre o metabolismo a qual ela pertence nesse parasito e também como uma tentativa de identificar alvos potenciais para o desenvolvimento de novas drogas.



**Figura 1.13:** Estrutura tridimensional de UPRT de *Sulfolobus solfataricus*, *T. gondii*, e *Bacillus caldolyticus* resolvidas por difração de raios-x: **(A)** estrutura cristalográfica da UPRT de *Sulfolobus solfataricus* na resolução de 1.80 Å (PDBid: 1XTT); **(B)** estrutura cristalográfica da UPRT de *T. gondii* na resolução de 2.20 Å (PDBid:1BD4); **(C)** estrutura cristalográfica da UPRT de *Bacillus caldolyticus* na resolução de 3.00 Å (PDBid:1I5E).

## 2. OBJETIVOS

### 2.1. Objetivo geral

O projeto visa à caracterização molecular das enzimas uracil fosforribosiltransferase (UPRT) e gliconato cinase (GK) de *T. cruzi*.

### 2.2. Objetivos específicos:

- Análise por programas de bioinformática visando: a) análise das estruturas primárias e secundárias das proteínas gliconato cinase e uracil fosforribosiltransferase de *T. cruzi*; b) identificação e comparação com as respectivas enzimas humanas, se existirem;
- Clonagem e expressão heteróloga em pET28a das enzimas alvos gliconato cinase e uracil fosforribosiltransferase;
- Purificação por cromatografia de afinidade das proteínas recombinantes gliconato cinase e uracil fosforribosiltransferase expressas;
- Produção de anticorpos policlonais em camundongos para a identificação das enzimas alvos nativas na forma epimastigota de *T. cruzi*, por meio de *western-blot* e imunofluorescência;
- Modelagem por homologia das enzimas gliconato cinase e uracil fosforribosiltransferase de *T. cruzi*.

### 3. MATERIAIS E MÉTODOS

#### 3.1. Identificação *in silico* das enzimas GK e UPRT de *T. cruzi*

A busca por enzimas metabólicas de *T. cruzi* que pudessem representar potenciais alvos terapêuticos foi realizada através da inferência de funções enzimáticas e da reconstrução metabólica desse parasito utilizando a ferramenta computacional AnEnPi (Otto *et al.*, 2008) (<<http://bioinfo.pdtis.fiocruz.br/AnEnPi/>>). Foram analisadas vias metabólicas relevantes ao parasito, o que permitiu a identificação e a seleção de duas enzimas: gliconato cinase (GK, EC 2.7.1.12), pertencente à via das pentoses fosfato e a uracil fosforribosiltransferase (UPRT, EC 2.4.2.9), pertencente à via de salvação de pirimidinas. A enzima GK e a enzima UPRT foram classificadas pelo AnEnPi como sendo específica e homóloga, respectivamente, em relação aos humanos. Cada uma das enzimas é codificada por duas cópias gênicas em *T. cruzi* foram denominadas *TcGK* (Tc00.1047053508177.20 e Tc00.1047053511693.60) referentes a GK e *TcUPRT* (Tc00.1047053506435.410 e Tc00.1047053507739.40) para UPRT.

#### 3.2. Análise das estruturas primárias de *TcGK* e *TcUPRT*

As sequências de proteínas das enzimas *TcGK* e *TcUPRT*, cujas atividades foram identificadas pelo AnEnPi, foram submetidas a análises utilizando ferramentas de bioinformática disponíveis online. Para o alinhamento das sequências foi utilizado o programa ClustalW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). Já os programas Pfam (<http://pfam.sanger.ac.uk/>) e ProDom (<http://prodom.prabi.fr>) permitiram a inferência de possíveis domínios funcionais e motivos protéicos conservados nessas sequências. Para a predição de peptídeo sinal, identificação de regiões transmembranares e predição da localização subcelular foram utilizados os programas SignalP (<http://www.cbs.dtu.dk/services/SignalP/>), TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>) e PSORT (<http://www.psort.org/>), respectivamente. Além disso, utilizou-se o PROSITE (<http://www.expasy.ch/prosite>) para a identificação nas sequências de motivos potenciais para o sítio ativo, sítio de ligação e modificações pós-traducionais.



### 3.3. Cultivo das formas epimastigotas de *T. cruzi*

As formas epimastigotas de *T. cruzi* cepa CL Brener foram cultivadas em meio LIT (*Liver Infusion Tryptose*) (tabelas 3.1, 3.2 e 3.3) a 28°C. A manutenção das culturas foi realizada a cada 7º dia de crescimento, obtendo-se ao final do processo uma concentração de aproximadamente  $6,0 \times 10^6$  células/mL. A partir da cultura crescida era feita uma diluição padrão 1:10 com volume final de 10ml.

**Tabela 3.1:** Solução inorgânica (2x) do meio de cultura LIT.

Reagente	Concentração
NaCl	8 g/L
KCl	0,8 g/L
Na <sub>2</sub> HPO <sub>4</sub>	16 g/L

**Tabela 3.2:** Solução orgânica (2x) do meio de cultura LIT.

Reagente	Concentração
Triptose	10 g/L
<i>Liver Broth Infusion</i>	10 g/L

**Tabela 3.3:** Meio de cultura LIT completo.

Reagentes	Volume/Litro
Sol. inorgânica 2X	450 mL
Sol. orgânica 2X	450 mL
Glicose (0,2 g/ml)	10 mL
Hemina (450 mg/ml)	0,5 mL
Soro Fetal Bovino	100 mL
H <sub>2</sub> O	qsp 1L

### 3.4. Extração do DNA genômico de *T. cruzi*

O DNA genômico de *T. cruzi* cepa CL Brener foi extraído segundo o método de extração com fenol/clorofórmio, conforme descrito em Sambrook & Russell (2001). Os resultados foram analisados por eletroforese em gel de agarose, conforme descrito no item 3.7.

### 3.5. Amplificação dos genes *TcGK* e *TcUPRT* por PCR (Reação da Polimerase em Cadeia)

A partir das sequências obtidas para os genes que codificam as enzimas *TcGK* e *TcUPRT*, os oligonucleotídeos (*primers*) foram desenhados manualmente e analisados utilizando o programa NET-PRIMER (<http://www.premierbiosoft.com/netprimer/netprlaunch/netprlaunch.html>). Os oligonucleotídeos senso (*forward*) e anti-senso (*reverse*) foram desenhados baseados nas extremidades 5' e 3' dos genes alvos. Para posterior clonagem no vetor de expressão pET28a, sítios de restrição para as enzimas *NdeI* (CATATG) e *XhoI* (CTCGAG) foram inseridos nas extremidades 5' das sequências dos oligonucleotídeos, como mostrados abaixo:

Gliconato cinase:

Oligonucleotídeo *TcGK forward*: 5'- CCATATGGCTGTCCCCCCTCT 3'

Oligonucleotídeo *TcGK reverse*: 5'- CCTCGAGTCAAAGATGCAGCCAC 3'

Uracil fosforribosiltransferase:

Oligonucleotídeo *TcUPRT forward*: 5'- CCATATG(G/A)AGAATTGCGGTGAC 3'

Oligonucleotídeo *TcUPRT reverse*: 5'- CCTCGAGTTAATCATGTGTCCC(G/A)AA 3'

Para a amplificação dos genes alvos *TcGK* e *TcUPRT*, a reação de PCR foi feita no termociclador *Mastercycler gradient* (Eppendorf), utilizando os respectivos oligonucleotídeos para cada sequência. As reações foram preparadas para um volume final de 50 µL (tabela 3.4) e continham a *Platinum Taq DNA polymerase High Fidelity* (Invitrogen) que contém uma mistura das enzimas *Taq DNA polymerase* e *Pyrococcus species GB-D polymerase* e do anticorpo *Platinum Taq Antibody*. Essa

mistura, portanto, tem a capacidade de adicionar uma desoxiadenosina (A) ao final 3' dos produtos de PCR, além de possuir atividade *proofreading* e *Hot start* automático. Os produtos gerados ao final da reação de PCR foram analisados por eletroforese em gel de agarose 1% (item 3.7).

**Tabela 3.4:** Componentes utilizados na reação de PCR

Reagentes	Volume
Tampão 10X <i>High Fidelity</i> PCR	5 µL
10 mM dNTP	1 µL
50 mM MgSO <sub>4</sub>	2,0 µL
Iniciador <i>Reverse</i> (10 pmol)	2,0 µL
Iniciador <i>Forward</i> (10 pmol)	2,0 µL
DNA genômico (1:10) <i>T. cruzi</i> cepa CL Brener	2,0 µL
<i>Platinum Taq DNA polymerase High Fidelity</i>	0.2 µL
H <sub>2</sub> O	10,5 µL
Volume final	50 µL

A programação padrão executada no termociclador, modificando-se apenas as temperaturas de anelamento de acordo com o gene a ser amplificado, foi à seguinte:

- Desnaturação inicial - 94 °C ..... 2 min
  - Desnaturação - 94 °C ..... 30 seg
  - Anelamento - *TcGK* 65,1 °C ..... 30 seg
  - *TcUPRT* 56,4 °C .. 30 seg
  - Extensão - 68 °C ..... 2 min
  - Extensão final - 68 °C ..... 7min e 30seg
- } 35 vezes

### 3.6. Purificação dos produtos de PCR

Os produtos de PCR obtidos foram purificados por meio do kit *QIAquick PCR Purification* (Qiagen). A purificação foi feita conforme as instruções do protocolo do kit. Por meio dessa purificação foi possível eliminar do produto de PCR os oligonucleotídeos, enzima, dNTPs e sais para, então, proceder a clonagem dos fragmentos.

### 3.7. Eletroforese em gel de agarose

Para a análise das amostras de DNA amplificadas por PCR, bem como das posteriores mini-preparações e digestões de DNA plasmidial, foram feitas eletroforeses em géis de agarose 1% em tampão TAE (Tris-Acetato-EDTA, pH 8,0). As amostras foram submetidas à voltagem constante de 70 V por 1 hora. Após a corrida, os géis foram corados com brometo de etídio, permitindo assim a visualização das bandas de DNA em transiluminador com radiação U.V. e, em seguida, fotografados.

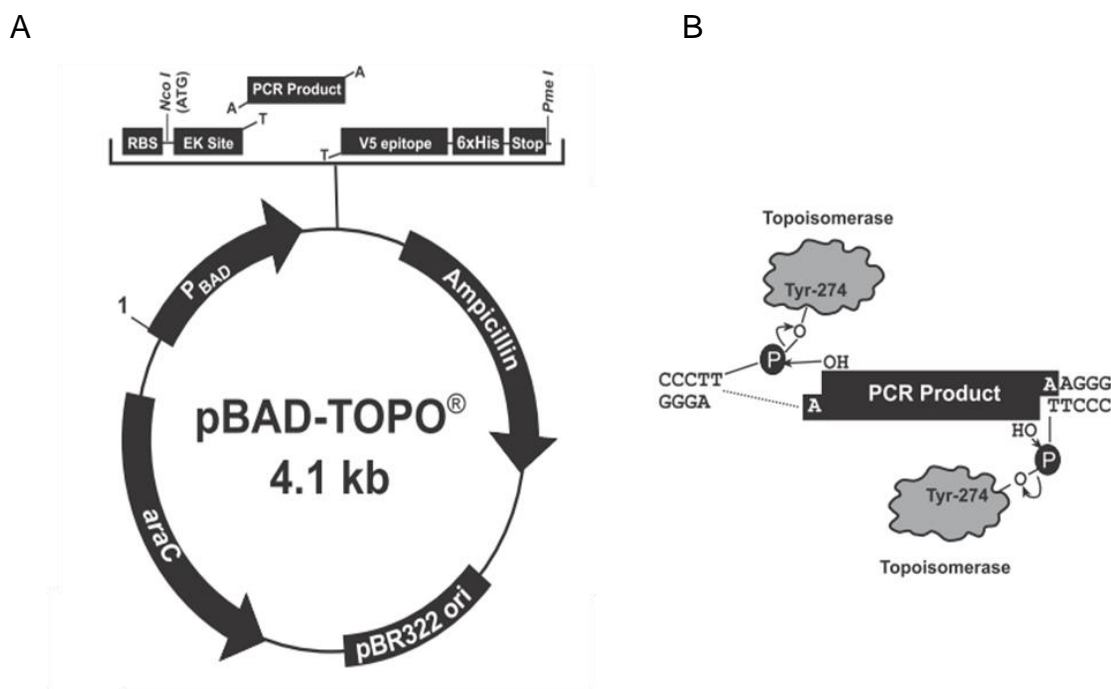
### 3.8. Clonagem dos alvos selecionados em vetor pBAD-TOPO TA

O vetor pBAD-TOPO TA (Invitrogen) (figura 3.1A) é um plasmídeo linear que contém uma desoxitimina (T) livre em cada uma das suas extremidades 3'. Essa desoxitimina é capaz de se ligar aos produtos de PCR adenilados através da reação catalisada pela enzima topoisomerase I (figura 3.1B) que vem acoplada ao vetor. O plasmídeo circular resultante pode ser introduzido em *E. coli*, visando a propagação plasmidial.

Os alvos *TcGK* e *TcUPRT* amplificados por PCR foram utilizados para a reação de ligação (tabela 3.5) ao vetor pBAD-TOPO TA (Invitrogen). Essa reação foi incubada a temperatura ambiente por 30 minutos.

Para a clonagem de cada alvo, foram utilizados 2µL de reação de ligação para transformar 50µL de *E. coli*, cepa TOP10, eletrocompetente através do eletroporador *Gene Pulser* da BioRad, seguindo os seguintes parâmetros: Coulomb=500µFD; Resistência=200Ω; Voltagem=2,5V. Todo o procedimento de eletroporação foi realizado em cubetas estéreis, previamente resfriadas. Após a

eletroporação, foi adicionado 1 mL de meio de cultura LB às bactérias que foram, então, mantidas sob agitação de 100 rpm, 37 °C, por 1 hora. Para o plaqueamento, foram utilizadas placas de petri com LB Ágar, contendo ampicilina à 100 µg/mL. As placas foram colocadas na estufa, à 37 °C por 18 horas.



**Figura 3.1:** Representação esquemática do vetor de clonagem pBAD-TOPO TA (A) e da reação catalisada pela topoisomerase I (B) A topoisomerase I se liga ao DNA dupla fita em sítios específicos e cliva em uma das fitas a ligação fosfodiéster depois de uma sequência 5'-CCCTT. A energia liberada com essa clivagem é conservada através da formação de uma ligação covalente entre o fosfato 3' da fita clivada e o resíduo de tirosina (Tyr-274) da topoisomerase I. A ligação fosfo-tirosil entre o DNA e a enzima pode ser subsequentemente atacada pela hidroxila 5' da fita originalmente clivada, invertendo a reação e libertando topoisomerase (Shuman, 1994).

**Tabela 3.5:** Reagentes da reação de ligação ao vetor pBAD-TOPO TA.

Componentes	Volume
Produto de PCR	2,0 µL
Vetor pBAD-TOPO TA	1,0 µL
H <sub>2</sub> O	2,0 µL
Solução de sal diluído (1:4)	1,0 µL
Volume final	6,0 µL

### 3.9. PCR de colônia

A partir das colônias de *E. coli* cepa TOP10 transformadas, fez-se um PCR de colônia para a confirmação da presença dos insertos *TcGK* e *TcUPRT* nas mesmas. Para isso, as colônias isoladas foram repassadas com auxílio de um palito estéril para a solução de PCR (tabela 3.6). Em seguida, cada colônia foi inoculada em 5 mL de meio LB com ampicilina 100 µg/mL, sendo incubada sob agitação a 200 rpm, 37 °C, por 18 horas.

O PCR de colônia segue o mesmo princípio do PCR comum, descrito no item 3.5. No entanto, no PCR de colônia não há a necessidade de se utilizar uma polimerase *high fidelity*, como a *Platinum Taq DNA polymerase High Fidelity*, visto que o produto de PCR é usado apenas para fins de análise do resultado da clonagem. Dessa forma, para o PCR de colônia utilizou-se a *Master Mix* (Promega) que é uma solução concentrada 2x de *Taq polymerase*, dNTPs e todos os componentes necessários para a reação de PCR, exceto o DNA genômico e os oligonucleotídeos. O resultado do PCR de colônia foi analisado em gel de agarose 1% em corrida de 1 h a 70V, como descrito no item 3.7.

**Tabela 3.6:** Componentes utilizados na reação de PCR de colônia.

Reagentes	Volume
<i>Master Mix</i> (2x)	12,5 µL
H <sub>2</sub> O	11,5 µL
Iniciador <i>Reverse</i> (10 pmol)	0,5 µL (5 pmol)
Iniciador <i>Forward</i> (10 pmol)	0,5 µL (5 pmol)
Clone de <i>E. coli</i> cepa TOP10	-
Volume final	25 µL

A programação padrão executada no termociclador, modificando-se apenas as temperaturas de anelamento de acordo com o gene a ser amplificado, foi à seguinte:

- Desnaturação inicial - 95 °C ..... 4 min
  - Desnaturação - 95 °C ..... 30 seg
  - Anelamento - *TcGK* 65,1 °C ..... 30 seg
  - *TcUPRT* 56,4 °C ..... 30 seg
  - Extensão - 72 °C ..... 1min e 30 seg
  - Extensão final - 72° ..... 7min e 30seg
- } 35 vezes

### 3.10. Extrações de DNA plasmidial (*Miniprep*), sequenciamento e análise dos resultados

Das culturas de TOP10 positivas para a presença do inserto pela análise do PCR de colônia foram centrifugados 3 mL a 9.300xg por 5 minutos. Após a centrifugação, os sobrenadantes foram descartados e os precipitados foram submetidos à extração e purificação de seus plasmídeos por meio do kit *QIAprep Spin Miniprep* (QIAGEN), conforme as instruções do fabricante.

Com o DNA plasmidial extraído, preparou-se o meio para a reação de sequenciamento contendo de 200-500 ng de DNA plasmidial e 3,2 pmol de Iniciador “*forward*” ou “*reverse*”, conforme indicado pela Plataforma Genômica - Sequenciamento de DNA PDTIS/FIOCRUZ. Para o sequenciamento dos insertos clonados em pBAD-TOPO TA, utilizou-se os oligonucleotídeos pBAD *reverse* e pBAD *forward* nas concentrações de que acompanham o kit do referido vetor. Todas as etapas que se sucedem até o sequenciamento, exceto a preparação das amostras, foram realizadas na plataforma de sequenciamento de DNA PDTIS/FIOCRUZ.

A partir do resultado do sequenciamento, realizou-se um alinhamento das sequências obtidas contra as respectivas sequências teóricas identificadas pelo programa AnEnPi. Este alinhamento foi feito com o auxílio do programa ClustalW2 (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>). Em seguida, foi escolhido apenas um clone de cada alvo para dar prosseguimento às etapas posteriores do projeto.

### 3.11. Dosagem do DNA plasmidial

A partir do DNA plasmidial referente ao clone escolhido para cada alvo e do vetor pET28a circular, foi feita a quantificação de DNA com o kit *Qubit dsDNA BR Assay* (Invitrogen) seguindo as instruções do protocolo do kit. O DNA foi dosado no *Qubit 2.0 Fluorometer* (Invitrogen).

### 3.12. Digestão do DNA plasmidial com as enzimas de restrição *XhoI* e *NdeI*

Após o vetor pET28a circular e o DNA plasmidial de cada clone escolhido terem sido devidamente quantificados, foram feitas reações de dupla digestão (tabela 3.7) utilizando as enzimas de restrição *NdeI* e *XhoI*. As digestões foram incubadas por 3 horas a 37 °C. Posteriormente, as enzimas de restrição foram inativadas a 65 °C em banho maria por 20 minutos. Essas digestões visaram a linearização do vetor pET28a e a separação dos insertos alvos do vetor pBAD-TOPO TA, gerando pontas coesivas em ambos os casos. As análises das digestões foram feitas em gel de agarose 1%, conforme descrito no item 3.7. Os fragmentos *TcGK* e *TcUPRT* liberados com a digestão foram purificados a partir do gel de agarose utilizando o kit *QIAquick Gel Extraction* (Qiagen).

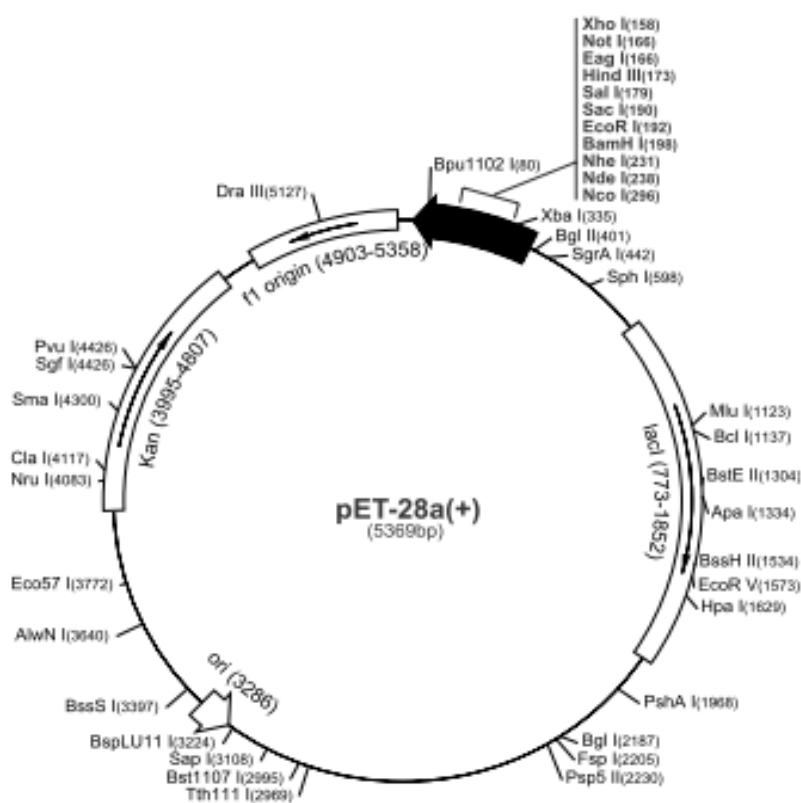
**Tabela 3.7:** Reação de digestão com as enzimas *XhoI* e *NdeI*

Reagentes	Condição final
DNA plasmidial	1,25 µg
Tampão 10X NEB4 (Biolabs)	1X
BSA 100X	1X
<i>NdeI</i>	20 U
<i>XhoI</i>	10 U
H <sub>2</sub> O	qsp 50 µL
Volume final	50 µL



### 3.13. Subclonagem dos genes alvos no vetor pET28a

O vetor pET28a (Invitrogen) (figura 3.2) é um plasmídeo de expressão que confere resistência à kanamicina e possui o promotor T7, permitindo altos níveis de expressão a partir da indução com IPTG. Além disso, esse vetor admite uma construção plasmidial capaz de conter uma sequência nucleotídica na região N- ou C-terminal do sítio de clonagem que codifica uma cauda de seis histidinas fusionada à proteína recombinante. Essa cauda de histidina permite a purificação da proteína recombinante através de cromatografia de afinidade em metal imobilizado e a detecção dessa proteína com anticorpos anti-histidina (anti-His).



**Figura 3.2:** Representação esquemática do vetor de expressão pET28a

A partir dos insertos alvos *TcGK* e *TcUPRT* e do vetor pET28a linearizado resultantes da digestão, foi feita a reação de ligação (tabela 3.8) dos insertos ao vetor. Essa reação foi mantida à temperatura ambiente por 2 horas. Os plasmídeos de expressão construídos foram denominados *TcGK*-pET e *TcUPRT*-pET, os quais tem o potencial de codificar para as proteínas *TcGK* e *TcUPRT*, respectivamente.

Para a clonagem dos plasmídeos *TcGK-pET* e *TcUPRT-pET* gerados, foram utilizados 2µL de reação de ligação para transformar em 50µL de *E. coli*, cepa TOP10, eletrocompetente através do eletroporador *Gene Pulser* da BioRad, seguindo os seguintes parâmetros: Coulomb=500µFD; Resistência=200Ω; Voltagem=2,5V. Após a eletroporação, foi adicionado 1mL de meio de cultura LB puro às bactérias que foram mantidas sob agitação de 100 rpm, à 37 °C, por 1 hora. Foram utilizadas placas de petri com LB Ágar com kanamicina 50 µg/mL para o plaqueamento das bactérias transformadas.

Para a confirmação da presença dos insertos *TcGK* e *TcUPRT* nas colônias crescidas, fez-se um PCR de colônia, seguindo os mesmos princípios descritos no item 3.9. As colônias em que se verificou resultado positivo para esses insertos foram submetidos a extrações de DNA plasmidial (Miniprep).

**Tabela 3.8:** Reagentes da reação de ligação ao vetor pET28a.

Componentes	Volume
Inserto alvo digerido	4,0 µL
Vetor pET28a digerido	1,0 µL
T4 DNA ligase	0,75 µL
Tampão 10X	1,0 µL
H <sub>2</sub> O	qsp 10 µL
Volume final	10,0 µL

### 3.14. Expressão das proteínas recombinantes *TcGK* e *TcUPRT* em pequena escala

As amostras de DNA plasmidial *TcGK-pET* e *TcUPRT-pET* extraídas de *E. coli* cepa TOP10, as quais por PCR de colônia apresentaram resultado positivo para a presença dos insertos *TcGK* e *TcUPRT*, respectivamente, foram utilizadas para a transformação em *E. coli* linhagem BL21(DE3). As bactérias *E. coli* linhagem BL21(DE3) são células competentes para a expressão de proteínas recombinantes. A transformação em BL21 foi feita por eletroporação, conforme descrito no item 3.13. Após a transformação, foi feito o plaqueamento utilizando placas de petri contendo LB/ Ágar com kanamicina 50µg/mL. As placas foram colocadas na estufa,

à 37 °C por 18 horas. Ao final desse período, foram escolhidas quatro colônias isoladas para o alvo TcGK e cinco colônias isoladas para o alvo TcUPRT, visando à análise de expressão heteróloga.

Cada uma das colônias escolhidas após a transformação em BL21 foi inoculada em 5 mL de meio LB com kanamicina 50 µg/mL e incubada sob agitação de 200 rpm, 37 °C, durante 18 horas. Desse pré-inóculo foram retirados 500 µL e repassados a novos 5 mL de meio LB com kanamicina 50 µg/mL e mantidas nas mesmas condições anteriores até apresentarem uma Abs<sub>600nm</sub> entre 0,4 e 0,6. Depois de atingida a densidade ótica (OD) ideal, foi feita a indução com adição de IPTG para uma concentração final de 1 mM, incubando-se as culturas por 3 horas nas condições adequadas. A partir do pré-inóculo, foi feito também um controle negativo que não foi submetido à indução com IPTG, mas foi mantido nas mesmas condições que as colônias induzidas. Ao término do período de incubação, duas alíquotas de 1mL foram retiradas de cada solução induzida e da não induzida (controle negativo) e foram repassadas cada uma a um microtubo para centrifugação a 9.300xg por 5 minutos. Após a centrifugação, os sobrenadantes foram totalmente descartados e os dois precipitados resultantes de cada amostra foram guardados. Simultaneamente, todas as soluções induzidas foram diluídas 4x com meio LB para a medição de suas absorvâncias a 600 nm com auxílio do espectrofotômetro.

Uma alíquota da duplicata de precipitados, referente a cada cultura induzida, foi ressuspensa com tampão de amostra 1X (125mM Tris-HCl, pH 6,8; SDS 3 %, glicerol 20 % e azul de bromofenol), contendo 2mM de β-mercaptoetanol. A quantidade de tampão de amostra adicionada ao precipitado levou em consideração a absorvância obtida de cada cultura após a indução, de forma que todas as amostras tivessem as mesmas concentrações finais de bactéria, e que variações na quantidade de proteína observadas no gel de SDS-PAGE fossem devido unicamente a diferença de expressão. Após as amostras terem sido preparadas, foram aplicados 10µL de cada uma (corresponde a uma Abs = 0,1 da suspensão de bactérias) em gel de poliacrilamida em condições desnaturantes, que foi feito segundo os procedimentos descritos no item 3.15.

### 3.15. Eletroforese em gel de poliacrilamida em condições desnaturantes (SDS-PAGE)

As proteínas foram analisadas por eletroforese, utilizando-se SDS-PAGE, sendo as concentrações do gel de concentração 4% e a do gel de separação 15%, conforme indicado na tabela 3.9. A todas as amostras foram adicionados tampão de amostra contendo  $\beta$ -mercaptoetanol 2mM. Em seguida, as amostras foram aquecidas por 5 minutos à 100 °C e posteriormente aplicadas no gel de poliacrilamida. Além das amostras, foram aplicados no gel 5  $\mu$ L de padrão de peso molecular (*Broad range* – BioRad; *Broad Range Prestained* – BioRad ou *PageRuler Prestained* - Thermo Scientific). A eletroforese em gel de poliacrilamida foi realizada a 200 V por 1 hora e, após seu término, o gel foi corado com solução de Coomassie-blue 0,5% (0,5% de Coomassie blue R-250; 1:1 Etanol 100%/H<sub>2</sub>O), sob agitação branda por 30 minutos. Foi, em seguida, aplicada a solução descorante (etanol 30% e ácido acético 7%) para a retirada do corante e depois a solução de secagem (etanol 30% e glicerol 2%).

**TABELA 3.9:** Reagentes para preparação do SDS-PAGE

Gel de concentração 4%		Gel de separação 15%	
Reagentes	Volume	Reagentes	Volume
H <sub>2</sub> O	1,2 mL	H <sub>2</sub> O	1,38 mL
1,5M Tris-HCl, pH 6,8	500 $\mu$ L	1,5M Tris-HCl, pH 8,8	1,5 mL
30% acrilamida/ 0,8% bisacrilamida	267 $\mu$ L	30% acrilamida/ 0,8% bisacrilamida	3,0 mL
10% SDS (duodecil sulfato de sódio)	20 $\mu$ L	10% SDS (duodecil sulfato de sódio)	60 $\mu$ L
10% persulfato de amônio	20 $\mu$ L	10% persulfato de amônio	60 $\mu$ L
TEMED	2 $\mu$ L	TEMED	6 $\mu$ L

### 3.16. Testes de expressão das proteínas recombinantes em pequena escala

Todos os clones submetidos à expressão em pequena escala tiveram o seu DNA plasmidial sequenciado (item 3.10) por meio dos oligonucleotídeos do vetor pET28a T7 *forward* e T7 *reverse*, visando à confirmação da integridade das

construções plasmidiais. Com o resultado do sequenciamento foi escolhido apenas um clone de cada alvo para dar continuidade aos testes de expressão. Os testes de expressão em pequena escala das proteínas recombinantes TcGK e TcUPRT foram feitos variando o tempo de indução e a temperatura de indução, utilizando as seguintes condições: à 20 °C, por 3h e 24h, e à 27 °C e 37 °C por 3h, a partir da indução com 1mM de IPTG, seguindo os mesmos procedimentos citados no item 3.14. Ao término do período de incubação, alíquotas de 1 mL foram retiradas de cada solução induzida e da não induzida (controle negativo) e foram repassadas a microtubos para centrifugação a 9.300xg por 5 minutos. Após a centrifugação, os sobrenadantes foram totalmente descartados e os precipitados resultantes de cada amostra foram guardados para ensaios posteriores de lise e separação em fração solúvel e insolúvel. Simultaneamente, todas as soluções induzidas foram diluídas 4x com meio LB para aferição das absorbâncias a 600 nm.

### **3.17. Lise bacteriana e separação das proteínas recombinantes TcGK e TcUPRT em fração solúvel e insolúvel**

As alíquotas dos precipitados de cada cultura resultantes dos testes de expressão em pequena escala foram ressuspendidas com 1 mL de tampão de lise (20 mM Tris-HCl, pH 7,5, 100 mM KCl, 5 mM EDTA) e passados para tubos contendo pérolas de vidro de 0,1 mm . Esses tubos com o precipitado, pérolas e tampão de lise foram colocados em equipamento de alta rotação (*Bead Beater*), sendo submetidos a três pulsos de um minuto intercalados com imersão também de 1 minuto no gelo. Após isso, separou-se por decantação as pérolas de vidro da solução que foi, então, centrifugada por 10 minutos a 16.100xg. Com essa centrifugação é possível separar cada amostra em uma fração contendo as proteínas na forma solúvel e outra na forma insolúvel. O precipitado (referente à fração de proteína insolúvel) obtido foi guardado a -20 °C e o sobrenadante (referente à fração de proteína solúvel) foi precipitado com ácido tricloroacético (TCA) a 17% e centrifugado à 16.100xg por 10 minutos. O TCA é capaz de precipitar as proteínas solúveis através da desnaturação ácida das mesmas. Ao haver precipitação, o novo sobrenadante obtido foi descartado e ao precipitado, contendo as proteínas, foi adicionado acetona/trietanolamina para o qual se fez uma nova centrifugação e descarte de sobrenadante. O precipitado de proteína solúvel

referente a cada alvo foi guardado a -20 °C. Após ter sido feita a lise celular e a separação das proteínas solúveis das insolúveis, fez-se um gel de poliacrilamida para a análise dos resultados, conforme descrito no item 3.15, no qual foram aplicados 20 µL de cada amostra preparada com tampão de amostra contendo 2mM de β-mercaptoetanol. Além disso, para uma análise mais sensível dos resultados, foi feito um *western blot* (item 3.24) tendo como anticorpo primário anti-histidina (Invitrogen) na diluição de 1:2.500 e anticorpo secundário conjugado à peroxidase na diluição de 1:10.000 (tabela 3.11).

### **3.18. Expressão das proteínas recombinantes TcGK e TcUPRT em larga escala**

A partir do clone escolhido para cada alvo, foi feita a expressão em larga escala de TcGK e TcUPRT recombinantes seguindo-se os parâmetros de indução que tiveram melhor resultado na expressão em pequena escala das proteínas recombinantes na forma solúvel. Para tal, 25 mL da colônia referente a cada alvo (derivada de pré-inoculo de 1% crescido por 12 horas) foram inoculados em 500 mL de meio LB com kanamicina 50 µg/mL. Quando as culturas atingiram uma  $Abs_{600nm}$  entre 0,4 e 0,6, foi feita a indução da expressão com IPTG 1mM à 20 °C por 24h sob agitação de 100 rpm.

Após a indução, foram separadas amostras de 1 mL do extrato bruto de bactérias para, posteriormente, serem submetidas à lise em pequena escala e separação das proteínas solúveis das insolúveis, descrito no item 3.17. O volume restante das culturas foi centrifugado à 22.100xg à 4 °C, por 20 minutos, e os precipitados celulares foram ressuspensos em tampão de lise (20 mM de Tris-HCl, pH 7,5, 100 mM de KCl, 5 mM de EDTA, 1 mM de PMSF, 1 mM de DTT). A lise celular foi realizada com auxílio de um equipamento denominado *Bead Beater*, seguindo os mesmos princípios da lise em pequena escala.

Os lisados obtidos foram centrifugados à 22.100xg, à 4 °C por 20 minutos. A partir da centrifugação, para cada alvo, obteve-se um sobrenadante correspondente a fração de proteína solúvel que foi armazenada à -20 °C e um precipitado que é referente à porção de proteínas insolúveis, o qual foi submetido a sucessivas lavagens com tampão de lavagem (50 mM de Tris-HCl, pH 8,5, 5% de Triton X-100, 5 mM de EDTA e 150 mM de NaCl). Para a lavagem do precipitado, adicionou-se ao

mesmo 5 mL de tampão de lavagem e fez-se a homogeneização da suspensão. Transferiu-se, então, a suspensão formada para um homogeneizador de cristal/vidro (*Douncer*), sendo assim o precipitado foi macerado cerca de 10 vezes de forma a retirar restos de membrana celular, células não lisadas e outros contaminantes. Centrifugou-se o precipitado lavado a 32.900xg, por 15 minutos, à 4 °C. Após esse procedimento, leu-se a  $Abs_{405nm}$  do sobrenadante obtido, tendo como branco o tampão de lavagem com Triton. O procedimento de lavagens com tampão de lavagem do homogeneizado com o auxílio do *Douncer* foi feita tantas vezes quanto necessárias de forma a obter valores de  $Abs_{405nm}$  dos sobrenadantes constantes, indicando que não havia mais extração de resíduos contaminantes.

Os precipitados obtidos a partir das lavagens com tampão de lavagem e contendo as proteínas recombinantes TcGK e TcUPRT na forma insolúvel foram solubilizados em 10 mL de solução de ureia 8 M em 50 mM Tris-HCl, pH 8,5 com auxílio de um *Douncer* e do vórtex. Depois, centrifugou-se as amostras à 32.900xg por 15 minutos à 4 °C. O resultado foi analisado em gel de proteína SDS-PAGE 15% a 200V por 1 hora (item 3.15).

Para analisar se haveria precipitação das proteínas recombinantes em concentrações mais baixas de ureia, foram feitas diluições seriadas dessas proteínas utilizando tampão de diluição (50 mM de Tris-HCl, pH 8,0, 600 mM de NaCl, 10 mM de imidazol) de forma a obter concentrações finais de 1 M, 2 M e 4 M de ureia. Os resultados foram analisados em SDS-PAGE (item 3.15).

### **3.19. Purificação por cromatografia de afinidade**

O precipitado de cada enzima recombinante, previamente solubilizado com tampão de ureia 8M, foi aplicado em uma coluna HisTrapHP (GE) contendo volume de 1 mL e carregada com íons  $Ni^{2+}$  para purificação por cromatografia de afinidade das proteínas recombinantes em função da cauda de histidina inserida nas mesmas. As proteínas ligadas à coluna por afinidade foram eluídas em gradiente crescente de imidazol (20 mM – 250 mM).

O procedimento utilizado para a purificação das proteínas recombinantes baseou-se em: injeção da amostra na coluna; lavagem da coluna com tampão de ligação (tabela 3.10) para eluir proteínas que não se ligaram à coluna; eluição das proteínas adsorvidas a coluna por meio de um gradiente descontínuo de imidazol

gerado pela mistura do tampão de ligação com o tampão de eluição; e finalmente lavagem da coluna com tampão de eluição (tabela 3.10).

Diversas purificações das proteínas recombinantes foram realizadas utilizando-se volumes de amostra que variaram de 500µL a 1,5 mL.

**TABELA 3.10:** Tampões para purificação por cromatografia de afinidade

<b>Tampão de Ligação</b>	<b>Tampão de Eluição</b>
Tampão fosfato de sódio 20mM, pH 7,5	Tampão fosfato de sódio 20mM, pH 7,5
NaCl 0,5 M	NaCl 0,5 M
Ureia 8 M	Ureia 8 M
Imidazol 20 mM	Imidazol 500 mM

Para a avaliação do processo de purificação das proteínas recombinantes, o volume de exclusão (*flow-through*), referente às proteínas que não se ligaram na coluna e as frações eluídas, que apresentaram alguma leitura à 280nm, foram submetidas à análise por SDS-PAGE (item 3.15). Após a análise do resultado, as frações que apresentaram a proteína recombinante foram selecionadas e juntadas em um único estoque (*pool*). A escolha do método de purificação foi baseada no fato de que, em vetores pET, a proteína recombinante é expressa fusionada a uma cauda de poli-histidina justamente para auxiliar o processo de purificação, como citado anteriormente. Essa sequência de poli-histidina é pequena e geralmente não altera as propriedades das proteínas as quais esta ligada, além de possuir uma alta afinidade pelo níquel ( $Ni^{2+}$ ) e outros cátion divalentes. Consequentemente, as proteínas que possuem uma cauda de poli-histidina são seletivamente ligadas à resina carregada devido aos íons de  $Ni^{2+}$  imobilizados, enquanto outras proteínas presentes no extrato formam ligações fracas ou são facilmente eluídas com os tampões específicos.

### 3.20. Dosagem de proteína

A dosagem de proteínas foi realizada através do Kit de dosagem RCDC - BioRad (modificado de Lowry et al., 1951), segundo instruções do fabricante. Para a curva padrão, utilizou-se albumina de soro bovino (BSA) em diferentes concentrações. As curvas padrões foram feitas em duplicata. Já as amostras foram



dosadas em diferentes diluições, sempre utilizando triplicata. A leitura foi feita em espectrômetro com  $\lambda=595$  nm.

### 3.21. Produção de anticorpos policlonais anti-TcGK e anti-TcUPRT

Com o intuito de produzir soro policlonal anti-TcGK e anti-TcUPRT, essas proteínas recombinantes purificadas e dosadas foram inoculadas em dois grupos de 5 camundongos Balb-C fêmeas, com 4 semanas de vida. Ao todo, duas doses de 10  $\mu$ g de proteína por camundongo foram administradas por via intraperitoneal, com intervalo de 15 dias entre as aplicações. A fim de potencializar a resposta imune, utilizou-se adjuvante incompleto de Freund em volume igual à solução de proteína. Antes da imunização, um camundongo foi sacrificado para obtenção do soro pré-imune que foi utilizado como controle negativo em ensaios de *western blot* (item 3.24).

Ao final de um mês, a partir da administração da primeira dose, fez-se a extração do sangue total. Para a obtenção dos soros imunes, primeiramente o sangue foi incubado à temperatura ambiente por 30 minutos e depois foi incubado à 4°C por 10 minutos, para que ocorresse a retração do coágulo. Após este procedimento, o sangue coagulado foi centrifugado a 1.000xg por 3 minutos. Com cuidado retirou-se, com auxílio de uma pipeta, o sobrenadante referente ao soro imune.

Para testar se houve produção de anticorpos policlonais a partir das proteínas recombinantes utilizadas, fez-se um *western blot*. Para isso, tanto para TcGK quanto para TcUPRT, foi feito um gel preparativo SDS-PAGE, no qual foram aplicados 35  $\mu$ g de proteína recombinante purificada. Após a separação das proteínas por eletroforese, elas foram transferidas para uma membrana de nitrocelulose, seguindo-se o procedimento de *western blot* descrito no item 3.24. Como anticorpos primários foram utilizados os soros imunes produzidos nas diluições de 1:1.000 e de 1:10.000.

### 3.22. Titulação dos anticorpos policlonais

Após a confirmação da produção de anticorpos policlonais anti-TcGK e anti-TcUPRT pelos camundongos imunizados, fez-se um *pool* para cada soro imune.

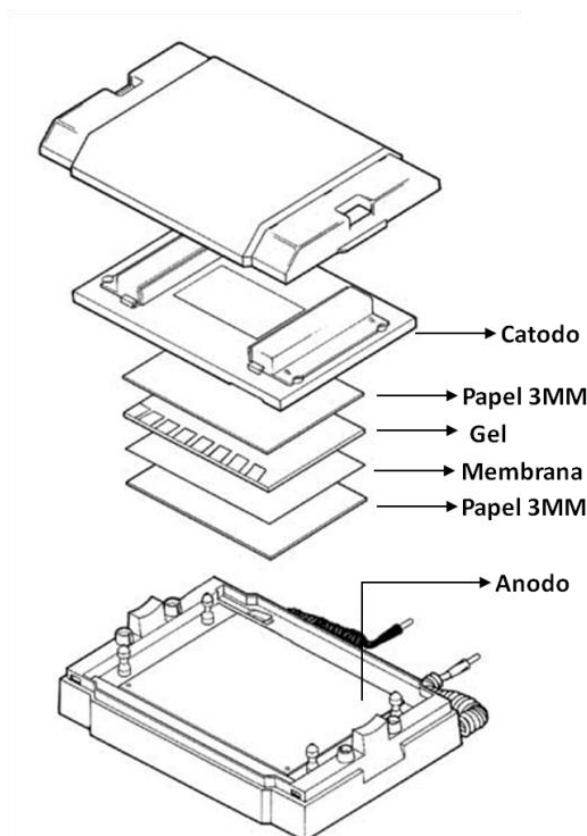
Esses *pools* de soro imune foram titulados e testados por *western blot*. Para tal, o gel preparativo SDS-PAGE e o *western blot* foram feitos seguindo os procedimentos mostrados no item 3.24, tabela 3.11, nos quais se utilizou como anticorpos primários os *pools* de soros imunes anti-TcGK e anti-TcUPRT nas diluições de 1:1.000, 1:2.000, 1:10.000, 1:20.000, 1:50.000 e 1:100.000. O anticorpo secundário conjugado à peroxidase foi diluído 1:10.000. Como controle negativo, utilizou-se o soro pré-imune na diluição 1:1.000 e como controle positivo o anticorpo anti-His na diluição de 1:20.000.

### 3.23. Purificação de imunoglobulina G a partir dos soros imunes

Amostras dos *pools* de soro imune, contendo os anticorpos policlonais TcGK e TcUPRT, foram submetidos à purificação de imunoglobulinas G (IgG) através do kit *Melon Gel IgG Spin Purification* (Thermo Scientific). Os IgGs purificados, assim como os *pools* de soros imunes não submetidos à purificação, foram utilizados em ensaios de *western blot* e imunofluorescência.

### 3.24. Western blot

O *western blot* é uma técnica que permite a detecção de uma proteína a partir de um anticorpo que a reconheça. Para isso, foi feita a separação da amostra de proteínas por eletroforese em gel SDS-PAGE (item 3.15). As proteínas foram então transferidas do gel para uma membrana de nitrocelulose através do sistema *Semi-dry* (BioRad). Para a transferência, papel 3MM, gel de poliacrilamida, membrana de nitrocelulose foram incubados em tampão de transferência gelado (39 mM de glicina, 48 mM de Tris Base, 20% etanol) e, posteriormente, dispostos, conforme a figura 3.3, sobre o anodo de platina do sistema *Semi-dry*. A transferência foi realizada durante 50 minutos sob voltagem constante de 15V e amperagem limitada a 500 mA. Após a transferência, a membrana foi corada com vermelho *Ponceau* para a análise da qualidade da transferência e, posteriormente, descorada com TBS 1X (100 mM de Tris-HCl pH 7,5, 150 mM de NaCl). Em seguida, a membrana de nitrocelulose foi deixada em solução de bloqueio (5% de leite em pó desnatado em TBS) *overnight*, sob leve agitação. Em seguida, a membrana foi lavada três vezes com TBS 1X e três vezes com TBS-Tween (10% TBS 10X, 0,05% Tween 20).



**Figura 3.3:** Representação esquemática da disposição do papel 3MM, gel de poliacrilamida, membrana de nitrocelulose no sistema "Semi-dry" (BioRad) para a transferência. Fonte: figura adaptada de [www.bio-rad.com/prd/pt/BR/LSR/PDP/b92801aa-76b7-45c9-825e-672589d01b00/Trans-Blot-SD-Semi-Dry-Transfer-Cell](http://www.bio-rad.com/prd/pt/BR/LSR/PDP/b92801aa-76b7-45c9-825e-672589d01b00/Trans-Blot-SD-Semi-Dry-Transfer-Cell).

Após a lavagem, a membrana de nitrocelulose foi incubada com anticorpo primário diluído em solução de bloqueio. De acordo com o experimento realizado, as incubações da membrana com anticorpo primário variaram nos seguintes parâmetros: anticorpo primário utilizado, diluição do mesmo e o tempo de incubação, conforme mostrado na tabela 3.11. Depois da primeira incubação, a membrana foi lavada três vezes com TBS 1X e três vezes com TBS-Tween e incubadas por 1 hora com anticorpo secundário conjugado à peroxidase. A diluição do anticorpo secundário variou de acordo com o experimento e com o anticorpo primário utilizado, conforme mostrado na tabela 3.11. Depois dessa incubação o procedimento das lavagens foi repetido e o resultado foi revelado com o kit *SuperSignal West Pico Chemiluminescent Substrate* (Thermo Scientific) segundo protocolo fornecido pelo fabricante, em câmara escura. O filme de raio-x foi exposto à membrana por diferentes períodos de tempos, sendo o máximo de 1 hora.

**Tabela 3.11:** Informações sobre os anticorpos primários e os anticorpos secundários de cada experimento de *western blot* realizado

Experimento	Anticorpo primário			Anticorpo secundário
	Nome do anticorpo	Diluição	Tempo de incubação	Diluição
Testes de expressão da proteína recombinante em pequena escala (item 3.16)	anti-His	1:2.500	2 horas	1:10.000
Lise celular e separação das proteínas recombinantes TcGK e TcUPRT em fração solúvel e insolúvel (item 3.17)	anti-His	1:2.500	2 horas	1:10.000
Titulação dos anticorpos policlonais (item 3.22)	Soro pré-imune (Controle negativo)	1:1.000	2 horas	1:10.000
	anti-His (Controle positivo)	1:2.500	2 horas	1:10.000
	Soros imunes	1:1.000	2 horas	1:10.000
		1:2.000		
		1:10.000		
1:20.000				
	1:50.000			
	1:100.000			
Identificação da enzima nativa TcGK em extrato total de <i>T. cruzi</i> (item 3.25)	anti-His (Controle positivo)	1:20.000	<i>overnight</i>	1:10.000
	IgG purificado/ soro imune anti-TcGK	1:10/ 1:100	<i>overnight</i>	1:5.000
Identificação da enzima nativa TcUPRT em extrato total de <i>T. cruzi</i> (item 3.25)	anti-His (Controle positivo)	120.000	<i>overnight</i>	1:10.000
	IgG purificado/ soro imune anti-TcUPRT	1:10/	<i>overnight</i>	1:5.000
		1:100		

### 3.25. Lise e extração de proteínas solúveis da forma epimastigota de *T. cruzi*

Amostras da forma epimastigota de *T. cruzi*, na concentração de  $6 \times 10^8$  células/mL, foram coletadas no quinto dia de cultivo do parasito, visando à lise e extração de suas proteínas solúveis. Essas amostras passaram por dois ciclos de lavagens com PBS 1X pH 7,2 (137mM de NaCl, 2,7mM de KCl, 10mM de  $\text{Na}_2\text{HPO}_4$ ,

2mM de  $\text{KH}_2\text{PO}_4$ ) alternados com centrifugações a 1.000xg por 20 minutos a 4 °C. Após a última centrifugação, as células foram ressuspensas em PBS 0,1X contendo inibidor de fosfatase e coquetel de inibidores de proteases nas concentrações indicadas pelo fabricante (Sigma). A relação de células por mL foi mantida em  $6 \times 10^8$ . A lise celular foi feita por meio de choque térmico, no qual os parasitos foram submetidos a cinco ciclos consecutivos de congelamento em  $\text{N}_2$  líquido e descongelamento em banho maria à 37 °C. Ao final do procedimento de lise foi feita a centrifugação do lisado total à 16.000xg por 30 minutos. O pellet foi descartado e o sobrenadante, fração onde se encontravam as proteínas solúveis, foi armazenado para experimentos posteriores.

### **3.26. Identificação das enzimas nativas TcGK e TcUPRT nas frações solúveis da forma epimastigota de *T. cruzi* por western blot**

Para identificar as enzimas nativas TcGK e TcUPRT nas frações solúveis da forma epimastigota de *T. cruzi*, realizou-se um ensaio por *western blot* com o soro policlonal obtido após a imunização dos camundongos. As proteínas recombinantes foram utilizadas como controle positivo.

Para esse fim, foi realizado um SDS-PAGE (item 3.15) contendo como amostras 10 µg/poço de cada proteína recombinante e 100 µg/poço de fração solúvel da forma epimastigota de *T. cruzi*. A cada amostra foi acrescentado tampão de amostra de proteína 4X com β-mercaptoetanol. No *western blot* (item 3.24), o anticorpo primário anti-His, correspondente ao controle positivo, foi diluído 1:20.000. Os soros imunes anti-TcGK e anti-TcUPRT foram utilizados na diluição de 1:100, enquanto que os IgGs purificados anti-TcGK e anti-TcUPRT foram utilizados na diluição de 1:10. Todos os anticorpos primários foram incubados com as membranas *overnight*. O anticorpo secundário acoplado à peroxidase foi usado na diluição 1:10.000 no controle positivo e na diluição de 1:5.000 nas outras amostras (tabela 3.11).

### 3.27. Identificação das enzimas nativas TcGK e TcUPRT na forma epimastigota de *T. cruzi* por imunofluorescência

No 5º dia de proliferação,  $6 \times 10^7$  células totais, das formas epimastigotas de *T. cruzi*, foram centrifugados à 2.500xg por 5 minutos. Em seguida foram feitas duas lavagens com PBS 1X seguidas de centrifugação à 2.500xg por 5 minutos e descarte do sobrenadante. Após as lavagens, os parasitos foram fixados em solução de paraformaldeído 4% diluído em PBS 1X, durante 5 minutos, sob agitação. Depois de fixados, os parasitos foram submetidos a um novo ciclo de lavagens com PBS 1X e aderidos à lamínulas previamente revestidas com 0,1% de poli-L-lisina (1 mg/mL). Em seguida, os parasitos foram permeabilizados com solução de 0,1% Triton X 100 em PBS 1X durante 10 minutos. Para o bloqueio foram feitas três incubações com BSA 2% em PBS 1X por 10 minutos à temperatura ambiente.

Após devidamente aderidos às lamínulas, os parasitos foram incubados *overnight* com os anticorpos primários com diferentes diluições, como mostrado na tabela 3.12. No controle negativo, os parasitos foram incubados com solução de BSA 2% em PBS 1X, sem a presença do anticorpo primário. As lamínulas foram então lavadas três vezes por 10 minutos em solução de bloqueio antes de serem incubadas por 1h com anticorpo secundário *Alexa Fluor 488 goat anti-mouse IgG* (Invitrogen) na diluição de 1:1.000 em solução de bloqueio. Em seguida, as lamínulas passaram por um novo ciclo de lavagens somente com PBS e foram incubadas por 10 minutos com DAPI (Invitrogen) diluído 1:1000 em PBS para marcação de DNA (núcleo). Em seguida, as lamínulas foram lavadas com PBS por 10 minutos e montadas sobre as lâminas com DABCO. Nas lamínulas controle, a incubação com o anticorpo primário foi omitida. O material foi observado ao microscópio de fluorescência APOTOME (Zeiss).

**Tabela 3.12:** Relação de anticorpos primários e as suas respectivas diluições nos ensaios de imunofluorescência:

<b>Anticorpo Primário</b>	<b>Diluição</b>
Anticorpo Anti-TcGK	1:25
Anticorpo Anti-TcGK	1:50
Anticorpo Anti-TcGK	1:100
IgG purificado anti-TcGK	1:25
Anticorpo Anti-TcUPRT	1:25
Anticorpo Anti-TcUPRT	1:50
Anticorpo Anti-TcUPRT	1:100
IgG purificado anti-TcUPRT	1:25

### 3.28. Modelagem estrutural das enzimas TcGK e TcUPRT e UPRT de *Homo sapiens*

As análises de estrutura secundária das sequências protéicas foram feitas através do programa PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred/>). A predição da estrutura terciária das enzimas alvos de *T. cruzi* e das enzimas com as mesmas funções enzimáticas preditas em humano foram resolvidas, utilizando o método de modelagem por homologia. O processo de construção dos modelos envolveu quatro etapas: identificação das proteínas molde, alinhamento das sequências alvos contra as sequências molde, construção do modelo e validação do mesmo. Para a identificação das proteínas molde tridimensionalmente resolvidas, foi feito um BLAST contra o banco de dados PDB (Protein Data Bank) a partir das sequências de aminoácidos das enzimas a serem modeladas. Posteriormente, foram feitos alinhamentos entre as sequências molde e as sequências a serem modeladas. Para o processo de modelagem estrutural das enzimas foi utilizado o programa Modeller (Fiser & Sali, 2003) que permite a modelagem por restrições espaciais. Uma avaliação precisa do modelo foi feita através do programa PROCHEK (Laskowski *et al.*, 1993) e por meio do gráfico de Ramachandran, utilizando o MolProbity (Davis *et al.*, 2004).

## 4. RESULTADOS

### 4.1. Análises das sequências nucleotídicas e das estruturas primárias das proteínas GK e UPRT

As sequências nucleotídicas e protéicas referentes às atividades enzimáticas GK (EC 2.7.1.12) e UPRT (EC 2.4.2.9) de *T. cruzi*, cepa CL Brener, foram obtidas a partir do TriTrypDB ([www.tritrypdb.org/](http://www.tritrypdb.org/)). Cada uma dessas atividades enzimáticas foi identificada pelo AnEnPi, assim como pelo TriTrypDB e KEGG, como possuindo duas cópias de genes putativos no genoma de *T. cruzi*. Para a análise dessas sequências, utilizaram-se programas de bioinformática disponíveis na internet, como ClustalW2, Pfam, ProDom, SignalP e PSORT, como descrito no item 3.2.

#### 4.1.1. Sequências nucleotídicas e protéicas referentes à GK

As sequências nucleotídicas de *T. cruzi* identificadas pelo AnEnPi, (Tc00.1047053508177.20 e Tc00.1047053511693.60), como codificantes para proteínas com possível atividade enzimática gliconato cinase foram denominadas *TcGK*. Cada uma dessas sequências possui 585 pb. A sequência nucleotídica *TcGK*-Tc00.1047053508177.20 codifica para proteínas de 194 aminoácidos com massas moleculares estimadas de 21,07 kDa (figura 4.1). A sequência nucleotídica *TcGK*-Tc00.1047053511693.60 codifica para proteínas de 194 aminoácidos com massas moleculares de 21,10 kDa (figura 4.2).



(A)> Tc00.1047053508177.20 gliconato cinase [EC 2.7.1.12];  
 ATGGCTGTCCCCCTCTCCGTCGCATCCTACGCTGTCTGTGCCGGTGGCTGTGGTTGTCATGGGC  
 CCCTCGGGCACGGGCAAGAGCACAATTGGTCTCCGCCTTGCACAGGAGCTCGGCTGGGAATTTGCA  
 GAAGGTGACAATTATCACCCCGCGGGCAACGTAGAAAAGATGGCGTCGGGTGTGCCTCTTAATGAT  
 GAGGACCGCTTGCCGTGGCTGAAAATTCTGCGCACCGAGGCCATAGAGGGTGCCTCGAGGGCGGCG  
 GCAGGAAGCGGTGTGCTGCTTGCATGTTCCGCGTTGCGAGAGTCGTACAGGGACGTGCTTCGCTCG  
 GACCACTCGGGCAACGGTGAAGAGTGAAGTGTCTTTGTGGAGCTTTTTGGCGACATCGGGTTC  
 ATTTCCAAGCGCTGCAGGGGCGTCATCATGCGTATATGCCGGCTTGTTTACTTGCATCGCAGTAC  
 GCCACGTTGGAGCCGCTGCGGCCGAAAAGGGCGAGCTTGGGGTGCAGATATCCGCGGCGTTGGAA  
 CCCGAGAAGATTGTGGAACGTGTGCTGGAGGAACTGAGGCGGTGGCTGCATCTTTGA

(B)> Tc00.1047053508177.20 gliconato cinase [EC 2.7.1.12];  
 MAVPPLPSHPTLSVPVAVVVMGSPGTTTCGKSTIGLRLAQELGWEFAEGDNYHPAGNVEKMASGVPL  
 NDEDRLPWLKILRTEAIEGASRAAAGSGVVLACALRESYRDVLRSDHSGNGGRVKVFFVELFGDI  
 GFISKRLQGRHHAYMPACLLASQYATLEPLRPEKGELGVRVSAALEPEKIVERVVEELRRWLHL

**Figura 4.1:** Sequência do gene *TcGK*- Tc00.1047053508177.20 de *T. cruzi*. (A) Sequência nucleotídica. (B) Sequência deduzida de aminoácidos.

(A)> Tc00.1047053511693.60 gliconato cinase [EC 2.7.1.12];  
 ATGGCGGTCTCCCCCTTCCGTCGCATCCTCCGCTCTCTGTGCCGGTGGCTGTGGTTGTCATGGGC  
 CCCTCGGGCACGGGCAAGAGCACAATTGGTCTCCGCCTTGCACAGGAGCTCGGCTGGGACTTTGCA  
 GAAGGTGACAATTATCACCCCGCGGGCAACGTGAAAAGATGGCGTCGGGTGTGCCTCTTAATGAT  
 GAGGACCGCTTGCCGTGGCTGAAAATTCTGCGCACTGAGGCCATAGAGGGTGCCTCGAGGGCGGCG  
 GCAGGAGGCGGTGTGCTGCTTGCATGTTCCGCGTTGCGAGAGTCGTACAGGGACGTGCTTCGCTCG  
 GACCACTTGGGCGACAGTGAAGAGTGAAGTGTCTTTGTGGAGCTTTTTGGCGACATCGGGTTC  
 ATTTCCAACGCCTGCAGGGGCGTCATCATGCGTATATGCCGGCTTGTTTACTTGCATCGCAGTAC  
 GCCACGTTGGAGCCGCTGCGGCCGAGAAGGGCGAGTTTGGGGTGCAGATATCCGCGGCGTTGGAA  
 CCCGAGAAGATTGTGGAACGTGTGCTGGAGGAACTGAGGCGGTGGCTGCATCTTTGA

(B)> Tc00.1047053511693.60 gliconato cinase [EC 2.7.1.12];  
 MAVSPLPSHPPLSVPVAVVVMGSPGTTTCGKSTIGLRLAQELGWDFAEAGDNYHPAGNVEKMASGVPL  
 NDEDRLPWLKILRTEAIEGASRAAAGGGVVLACALRESYRDVLRSDHLGDSGRVKVFFVELFGDI  
 GFISKRLQGRHHAYMPACLLASQYATLEPLRPEKGEFVVRVSAALEPEKIVERVVEELRRWLHL

**Figura 4.2:** Sequência do gene *TcGK*- Tc00.1047053511693.60 de *T. cruzi*. (A) Sequência nucleotídica. (B) Sequência deduzida de aminoácidos.

#### 4.1.2. Alinhamento entre as sequências de aminoácidos preditas de *TcGK*

A partir das sequências de aminoácidos das proteínas codificadas por ambas as cópias dos genes *TcGK* foi feito um alinhamento múltiplo utilizando o programa ClustalW2 (figura 4.3). O alinhamento indicou uma identidade de 95,8% e similaridade de 96,9% entre as sequências analisadas. Esse alto grau de identidade e similaridade permitiu o prosseguimento do estudo de caracterização da atividade

enzimática GK em *T. cruzi* utilizando apenas a cópia gênica TcGK-Tc00.1047053508177.20.

```

Tc00.1047053508177.20      MAVPPLPSHPTLSVPVAVVVMGPSGTGKSTIGLRLAQELGWFEAGDNYH 50
Tc00.1047053511693.60      MAVSPLPSHPPLSVPVAVVVMGPSGTGKSTIGLRLAQELGWDFAEGDNYH 50
                              *** .*****.*****:*****

Tc00.1047053508177.20      PAGNVEKMASGVPLNDEDRLPWLKILRTEAIEGASRAAAGSGVVLACSAL 100
Tc00.1047053511693.60      PAGNVEKMASGVPLNDEDRLPWLKILRTEAIEGASRAAAGGGVVLACSAL 100
                              *****.*****

Tc00.1047053508177.20      RESYRDVLRSDHSGNGGRVKVFFVELFGDIGFISKRLQGRHHAYMPACLL 150
Tc00.1047053511693.60      RESYRDVLRSDHLDGSGRVKVFVFFVELFGDIGFISKRLQGRHHAYMPACLL 150
                              ***** *:*****

Tc00.1047053508177.20      ASQYATLEPLRPEKGEFGVRSAALEPEKIVERVVEELRRWLHL 194
Tc00.1047053511693.60      ASQYATLEPLRPEKGEFGVRSAALEPEKIVERVVEELRRWLHL 194
                              *****:*****

```

**Figura 4.3:** Alinhamento das estruturas primárias de TcGK identificadas no genoma de *T. cruzi*. Alinhamento das sequências TcGKs (Tc00.1047053508177.20 e Tc00.1047053511693.60) referentes a atividade enzimática GK, utilizando o programa ClustalW2. Aminoácidos idênticos estão indicados por asterisco (\*), substituições conservadas, por dois pontos (:), e substituições semi-conservadas por um ponto (.).

#### 4.1.3. Sequências nucleotídicas e protéicas referentes à UPRT

As sequências nucleotídicas de *T. cruzi* identificadas pelo AnEnPi, (Tc00.1047053507739.40 e Tc00.1047053506435.410), como codificantes para proteínas com a possível atividade enzimática uracil fosforribosiltransferase foram denominadas *TcUPRT*. As sequências de ambos os genes possuem 711 pb e codificam proteínas de 236 aminoácidos com massas moleculares teóricas de 26,5 kDa (figuras 4.4 e 4.5).

(A)> Tc00.1047053507739.40 uracil fosforribosiltransferase [EC 2.4.2.9];  
 ATGGAGAATTGCGGTGACGGCAGTTCTTTGGAGCAGTCCCTTCTGAAGGAATTTCCGCAACTGCAC  
 CTGGTCCCGCAAACAAATCAGCTGAGGTTTCTATTTACCGTCATTCGCGACAAGAAGACAACCCGA  
 CAGATTTTGTGTTTTACTGCGAGCGGATCATTTCGTCTTGTGGTGGAGGCGGGGCTTGATTTGGTTC  
 CTGTTCTGCCAAAGGATGTGACGACACCAACCGGCGCCACCTACAAGGGCTGCATGCCTGATCCAC  
 AAGGCATTATTGGCATCAGCATCCTTCGCGCCGGCGAGTCTATGGAACGCGTTCTTCGAGAACTT  
 GCCGCGGTATTCGCATCGGAAAGATTCTGGTTCAGCGGGACGAGAAAACAGTTGAGAAGTGTCCGG  
 ATGAGCGGTATAACTACAGCAAAGTTCCAAGGGATGTGGCTGGACGTCGCGTGCTGCTTCTTGATC  
 CTATGATTGCGTTCGGGCGGCAGCGCCATCCACGCCACAGATATTCTTATAACGGGAATACAAGGTGC  
 CGGAGGAGAACATTATTTTTCTTAACCTTATTACGTGCCAGAGGGAATCCGCGGTATATGAGTA  
 GATTTCCAAAGGTCCACGTAGTGACCGCGGCGATCGATGGGGCTCTAGACGATTCCAATACATC  
 GTCCCCGGACTGGGTGATTTTGGCGATCGCTACTTCGGGACACATGATTAA

(B)> Tc00.1047053507739.40 uracil fosforribosiltransferase [EC 2.4.2.9];  
 MENC D G S S L E Q S L L K E F P Q L H L V P Q T N Q L R F L F T V I R D K K T T R T D F V F Y C E R I I R L V V E A G L D L V  
 P V L P K D V T T P T G A Y K G C M P D P Q G I I G I S I L R A G E S M E R V L R E T C R G I R I G K I L V Q R D E K T V E K C P  
 D E R Y N Y S K V P R D V A G R R V L L L D P M I A S G G S A I H A T D I L I R E Y K V P E E N I I F L N L I T C P E G I R R Y M S  
 R F P K V H V V T A A I D G A L D D S K Y I V P G L G D F G D R Y F G T H

**Figura 4.4:** Sequências do gene *TcUPRT*- Tc00.1047053507739.40 de *T. cruzi*. (A) Sequência nucleotídica. (B) Sequência protéica deduzida.

(A)> Tc00.1047053506435.410 uracil fosforribosiltransferase [EC 2.4.2.9];  
 ATGAAGAATTGCGGTGACGGCAGTTCTTTGGAGCAGTCCCTTCTGAAGGAATTTCCGAAACTGCAC  
 CTGGTCCCGCAAACAAATCAGCTGATTTTTCTATTTACCATCATTCGCGACAAGAAGACAACCCGC  
 ACAGATTTTGTGTTTTACTGCGAGCGGATCATTTCGTCTTGTGGTGGAGGCGGGGCTTGATTTGGTT  
 CCTGTTCTGCCAAAGGATGTGACGACACCAACCGGCGCCATCTACAAGGGCTGCATGCCTGATCCA  
 CAAGGCATTATTGGCATCAGCATCCTTCGCGCCGGCGAGTCTATGGAACGCGTTCTCCGGGAACT  
 TGCCGCGGTATTCGCATCGGAAAGATTCTGGTCCAGCGGGACGAGAAAACAGTTGAGAAGTGTCCG  
 GATGAGCGGTATAACTACAGCAAAGTTCCAAGGGATGTGGCTGGACGTCGCGTGCTGCTTCTTGAT  
 CCTATGATTGCGTTCGGGCGGCAGCGCCATCCACGCCACAGATATCCTTATAACGGGAATACAAGGTG  
 CCGGAGGAAAACATTATTTTTCTTAACCTTATTACGTGCCAGAGGGAATCCGCGGTATATGAGT  
 AGATTTCCAAAGGTCCACGTAGTGACCGCGGCGATCGATGGGGCTCTAGACGATTCCAATACATC  
 GTCCCCGGACTGGGTGATTTTGGCGATCGCTACTTTGGGACACATGATTAA

(B)> Tc00.1047053506435.410 uracil fosforribosiltransferase [EC 2.4.2.9];  
 M K N C D G S S L E Q S L L K E F P K L H L V P Q T N Q L I F L F T I I R D K K T T R T D F V F Y C E R I I  
 R L V V E A G L D L V P V L P K D V T T P T G A I Y K G C M P D P Q G I I G I S I L R A G E S M E R V L R E T C R G I R  
 I G K I L V Q R D E K T V E K C P D E R Y N Y S K V P R D V A G R R V L L L D P M I A S G G S A I H A T D I L I R E Y K  
 V P E E N I I F L N L I T C P E G I R R Y M S R F P K V H V V T A A I D G A L D D S K Y I V P G L G D F G D R Y F G T H

**Figura 4.5:** Sequência do gene *TcUPRT*- Tc00.1047053506435.410 de *T. cruzi*. (A) Sequência nucleotídica. (B) Sequência protéica deduzida.

#### 4.1.4. Alinhamento entre as sequências protéicas preditas de *TcUPRT*

Para a análise da similaridade e identidade entre as sequências de aminoácidos das *TcUPRTs*, foi feito um alinhamento múltiplo dessas sequências

utilizando o programa ClustalW2 (figura 4.6). O alinhamento indicou uma similaridade de 99,1% e uma identidade de 97,8% entre as sequências analisadas. Essa alta porcentagem de identidade e similaridade possibilitou o prosseguimento da caracterização da UPRT de *T. cruzi* utilizando apenas a cópia gênica *TcUPRT-Tc00.1047053507739.40*.



**Figura 4.6:** Alinhamento das estruturas primárias de TcUPRT identificadas no genoma de *T. cruzi*. Alinhamento das sequências TcUPRT (Tc00.1047053507739.40 e Tc00.1047053506435.410) referentes a atividade enzimática UPRT, utilizando o programa ClustalW2. Aminoácidos idênticos estão indicados por asterisco (\*), substituições conservadas, por dois pontos (:), e substituições semi-conservadas por um ponto (.).

#### 4.2. Análises das sequências de aminoácidos das proteínas TcGK e TcUPRT para identificação de possíveis domínios, motivos e assinaturas de famílias de proteínas

Análises com as ferramentas de bioinformática ProDom e Pfam foram realizadas visando a predição dos principais domínios funcionais e assinaturas de famílias de proteínas nas sequências de aminoácidos referentes às putativas enzimas TcGK e TcUPRT. Como se pode observar na tabela 4.1, a maioria dos domínios e famílias identificados para cada uma das sequências de aminoácidos analisadas é coerente com as atividades enzimáticas preditas pelo AnEnPi, apresentando *e-value* significativos. Para a sequência de aminoácidos referente à TcGK, foi identificado, com *e-value* de  $9,0 \times 10^{-30}$ , um domínio conservado

característico de gliconato cinase termorresistente. No entanto, uma região de 118 aminoácidos característica da família de chiquimato cinases foi identificada com *e-value*  $1,8 \times 10^{-9}$ . Na sequência de aminoácidos da putativa TcUPRT foram identificados potenciais domínios e assinaturas de famílias das uracil fosforribosiltransferases, embora domínios característicos das uridina cinases também tenham sido identificados.

**Tabela 4.1:** Relação dos potenciais domínios e assinaturas de famílias nas sequências de aminoácidos de TcGK e TcUPRT. As análises foram feitas através dos programas PFAM e ProDom.

Enzima	Programa	Domínio / Família	Posição dos aminoácidos	<i>e-value</i>
TcGK	Pfam	Chiquimato cinase	24-142	$1,8 \times 10^{-9}$
	ProDom	Gliconato cinase termorresistente	18-189	$9,0 \times 10^{-30}$
TcUPRT	Pfam	UPRTase- Uracil fosforribosiltransferase	24-234	$4,3 \times 10^{-70}$
	ProDom	Uridina cinase transferase	97-148	$1,0 \times 10^{-18}$
		Uracil glicosiltransferase fosforribosiltransferase	150-200	$7,0 \times 10^{-17}$
		Uridina cinase transferase	40-96	$2,0 \times 10^{-16}$
		Uracil transferase fosforribosiltransferase	20-96	$6,0 \times 10^{-07}$
		Uracil transferase fosforribosiltransferase	11-38	$2,0 \times 10^{-05}$
		Uracil glicosiltransferase fosforribosiltransferase	203-232	$2,0 \times 10^{-05}$
		Uridina cinase transferase	19-73	$2,0 \times 10^{-05}$
		Uridina cinase transferase	199-235	$8,0 \times 10^{-05}$

As predições de peptídeo sinal e de sítios de clivagem de peptídeo sinal nas sequências de aminoácidos de TcGK e TcUPRT foram realizadas através do programa signalP. A partir dessas análises, nenhum peptídeo sinal foi identificado nas proteínas TcGK e TcUPRT (tabela 4.2).

O programa TMHMM foi utilizado para a predição de regiões transmembrana nas sequências alvos e indicou a ausência dessas regiões nas enzimas TcGK e TcUPRT (tabela 4.2).

As predições de localização subcelular foram feitas com o programa Psort e indicaram que as proteínas TcGK e TcUPRT tem a probabilidade de 52,2% e 60,9%, respectivamente, de serem citoplasmáticas (tabela 4.2). Esses resultados se referem a não identificação, por esse programa, de peptídeos sinais responsáveis pelo endereçamento dessas proteínas a organelas, semelhante aos resultados obtidos com as análises dessas proteínas utilizando o SignalP.

**Tabela 4.2:** Predição de peptídeo sinal, região transmembranar e localização celular das enzimas TcGK e TcUPRT.

<b>Enzima</b>	<b>Peptídeo sinal (SignalP)</b>	<b>Região transmembranar (TMHMM)</b>	<b>Localização celular (Psort)</b>
<b>TcGK</b>	Não identificado	Não identificada	Citoplasmática
<b>TcUPRT</b>	Não identificado	Não identificada	Citoplasmática

Para a predição de motivos característicos de sítio ativo, sítios de ligação a substrato e/ou cofatores, modificações pós-traducionais e formação de pontes dissulfeto nas sequências de aminoácidos das proteínas alvos, foram realizadas análises com o PROSITE (tabela 4.3). No que se refere à modificações pós-traducionais, esse algoritmo permitiu a identificação na sequência de TcGK de seis prováveis sítios de N-miristoilação, três prováveis sítios de fosforilação por proteína cinase C e um provável sítio de fosforilação por caseína cinase II. Além de sítios de modificações pós-traducionais, foi identificado também um provável sítio de ligação a ATP/GTP nas posições 22 – 29 da sequência de aminoácidos.

Para a sequência de TcUPRT (tabela 4.3), foi identificado um provável sítio de N-miristoilação, um de fosforilação por tirosina cinase, três de fosforilação por proteína cinase C, três de fosforilação por caseína cinase II e um de fosforilação por proteína cinase dependente de cAMP e cGMP, além de um possível sítio para N-glicosilação e um sítio de amidação.

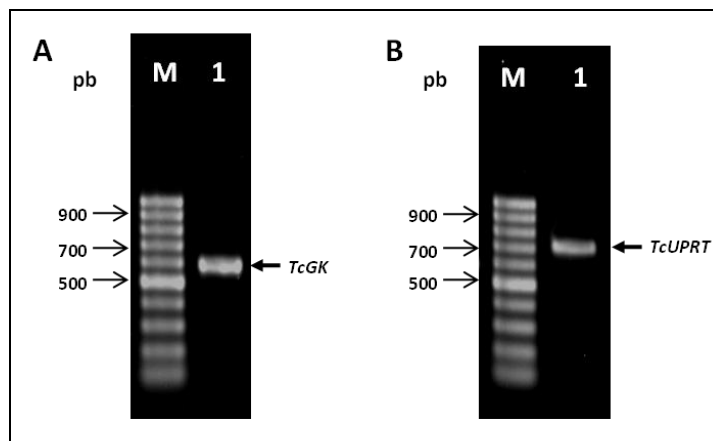
**Tabela 4.3:** Predição pelo PROSITE de motivos característicos de sítios de ligação e de modificação pós-traducionais nas sequências de aminoácidos de TcGK e TcUPRT. Os aminoácidos em letras maiúsculas representam posições conservadas enquanto os aminoácidos em letras minúsculos representam as posições mais variáveis.

Enzima	PROSITE		
	Sítios identificados	Aminoácidos envolvidos	Posição dos aminoácidos
TcGK	N-miristoilação	GTgkST GLrIAQ GVpIND GAsrAA GVvIAC GVrvSA	25 – 30 32 – 37 61 – 66 83 – 88 92 – 97 168 – 173
	Fosforilação por proteína cinase C	TgK SyR SkR	26 – 28 103 – 105 134 – 136
	Fosforilação por caseína cinase II	SyrD	103 – 106
	Sítio de ligação ATP/GTP (P-loop)	GpsgtGKS	22 - 29
TcUPRT	N-miristoilação	GlriGK	113 – 118
	Fosforilação por tirosina cinase	KcpdErynY	130 – 138
	Fosforilação por proteína cinase C	TtR TyK TcR	42 – 44 80 – 82 110 – 112
	Fosforilação por caseína cinase II	SsIE TrtD TcpE	8 – 11 43 - 46 188 - 191
	fosforilação por proteína cinase dependente de cAMP e cGMP	KKtT	40 – 43
	N-glicosilação	NYSK	137 – 140
	Sítio de amidação	aGRR	146 – 149

#### 4.3. Amplificação dos genes *TcGK* e *TcUPRT* por PCR

Os genes alvos *TcGK* e *TcUPRT* foram amplificados a partir do DNA genômico de *T. cruzi* cepa CL Brener através da reação de PCR. Essas amplificações foram realizadas utilizando a enzima *Platinum Taq DNA polymerase High Fidelity* (Invitrogen) e *primers* específicos contendo os sítios de restrição para as enzimas *NdeI* e *XhoI*. A análise das amplificações foi feita por eletroforese em gel

de agarose 1% e evidenciou que o tamanho dos fragmentos amplificados correspondia ao peso molecular esperado de 585pb para o gene *TcGK* e de 711pb para o gene *TcUPRT* (figura 4.7). Os fragmentos *TcGK* e *TcUPRT* amplificados foram purificados com o kit *QIAquick PCR Purification* (Qiagen).

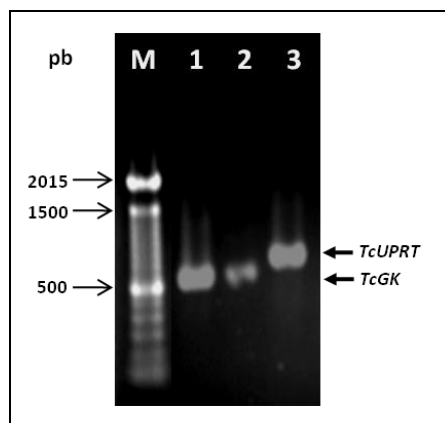


**Figura 4.7:** Análise em gel de agarose 1% da amplificação por PCR dos genes *TcGK* e *TcUPRT*. **(A)** Poço M - marcador de peso molecular 100pb DNA (Biolabs); poço 1 - fragmento *TcGK* de peso molecular 585pb. **(B)** poço M - marcador de peso molecular 100pb DNA (Biolabs); poço 1 - alvo *TcUPRT* de peso molecular 711pb.

#### 4.4. Clonagem dos genes *TcGK* e *TcUPRT* no vetor pBAD-TOPO TA

A partir dos fragmentos *TcGK* e *TcUPRT* purificados, realizou-se a reação de ligação ao vetor pBAD-TOPO TA (Invitrogen) originando as construções *TcGK*-pBAD-TOPO TA e *TcUPRT*-pBAD-TOPO TA, respectivamente. Em seguida, essas construções foram eletroporadas em *E. coli* eletrocompetentes, cepa TOP10, que foram plaqueadas em meio LB Ágar com ampicilina. As colônias crescidas foram submetidas a um PCR de colônia, conforme descrito no item 3.9, visando à confirmação da presença dos insertos alvos em questão. Os resultados do PCR de colônia foram analisados em gel de agarose 1% (figura 4.8). Ao todo foram obtidas duas colônias positivas para o fragmento *TcGK* e uma colônia positiva para o fragmento *TcUPRT*.





**Figura 4.8:** Análise do PCR de colônia para confirmação da clonagem dos genes *TcGK* e *TcUPRT* no vetor pBAD-TOPO TA. Gel de agarose 1%. Poço M: marcador de peso molécula 100 bp DNA Ladder (Invitrogen); poços 1 e 2 - colônias positivas para *TcGK*; poço 3 - colônia positiva para *TcUPRT*.

#### 4.5. Sequenciamento das construções *TcGK* e *TcUPRT* em vetor pBAD-TOPO TA

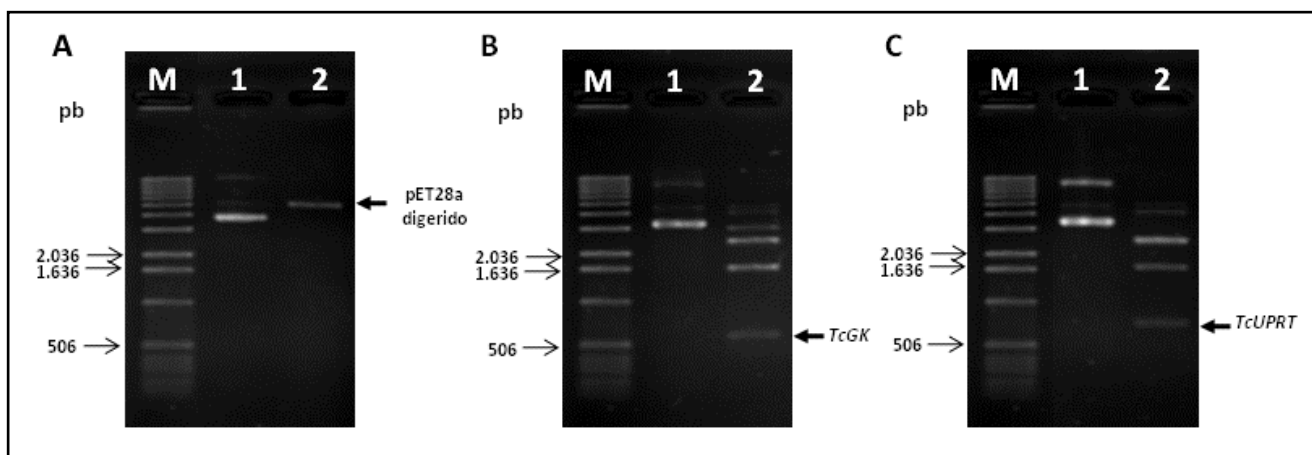
As colônias que apresentaram resultado positivo para a presença do inserto alvo, *TcGK* ou *TcUPRT*, foram submetidas à extração de DNA plasmidial utilizando o kit *QIAprep Spin Miniprep* (Qiagen). Para a confirmação da integridade das sequências dos insertos *TcGK* e *TcUPRT* clonados, os DNAs plasmidiais extraídos de cada colônia foram sequenciados pela Plataforma Genômica – Sequenciamento de DNA (RPT01-A), do PDTIS/FIOCRUZ.

A partir dos resultados do sequenciamento, foram feitos alinhamentos das sequências obtidas contra as sequências de referência utilizando o programa ClustalW2, o que permitiu verificar a integridade dos genes *TcGK* e *TcUPRT* clonados (Anexos 8.1 e 8.2, respectivamente). Com isso foi possível a seleção de uma colônia, com a sua respectiva construção plasmidial, para cada um dos alvos de forma a dar prosseguimento às etapas de digestão, subclonagem em pET28a e a posterior expressão das proteínas recombinantes.

#### 4.6. Digestão do DNA plasmidial com as enzimas de restrição *XhoI* e *NdeI*

Os DNAs plasmidiais *TcGK*-PBAD-TOPO TA e *TcUPRT*-PBAD-TOPO TA, referentes a cada colônia escolhida, e o plasmídeo pET28a circular foram

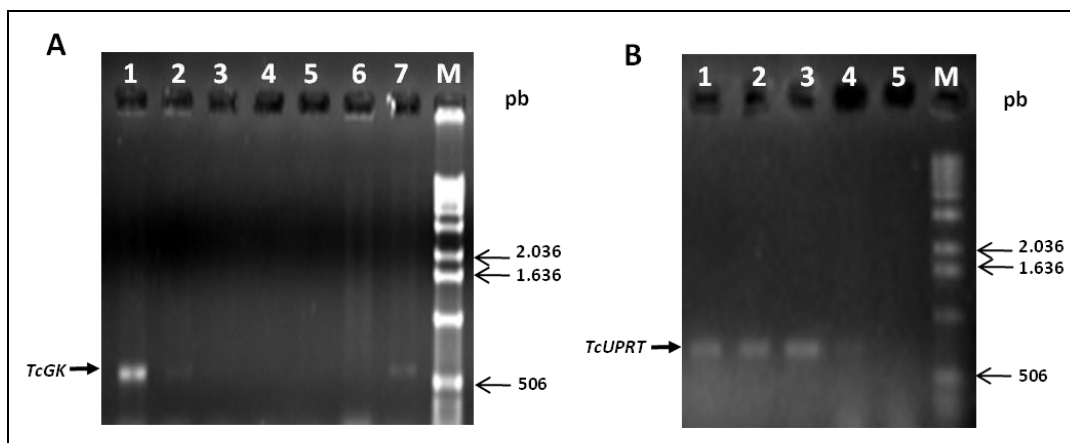
quantificados e submetidos a reação de dupla digestão com as enzimas de restrição *NdeI* e *XhoI*. As análises das digestões foram feitas em gel de agarose 1% e evidenciaram a liberação dos fragmentos alvos *TcGK* e *TcUPRT* nos tamanhos esperados, bem como a linearização do vetor pET28a (figura 4.9). Os fragmentos *TcGK* e *TcUPRT* resultantes da digestão foram purificados a partir do gel de agarose utilizando o kit *QIAquick Gel Extraction* (Qiagen).



**Figura 4.9:** Digestão das construções plasmidiais *TcGK*-pBAD –TOPO TA e *TcUPRT*-pBAD-TOPO TA e do plasmídeo pET28a. Gel de agarose 1%. **(A)** Poço M - marcador de peso molecular 1kb DNA Ladder (Invitrogen); poço 1 - vetor pET28a não digerido; poço 2: vetor pET28a digerido. **(B)** Poço M: marcador de peso molecular 1kb DNA Ladder (Invitrogen); poço 1 - DNA plasmidial *TcGK*-pBAD-TOPO TA não digerido; poço 2 - DNA plasmidial *TcGK*-pBAD-TOPO TA digerido. **(C)** Poço M - marcador de peso molecular 1kb DNA Ladder (Invitrogen); poço 1 - DNA plasmidial *TcUPRT*-pBAD-TOPO-TA não digerido; poço 2 - DNA plasmidial *TcUPRT*-pBAD-TOPO TA digerido.

#### 4.7. Subclonagem dos genes alvos no vetor pET28a

Os fragmentos *TcGK* e *TcUPRT* resultantes das digestões foram ligados ao vetor de expressão pET28a linearizado, conforme descrito no item 3.12. Essas construções, denominadas *TcUPRT*-pET28a e *TcGK*-pET28a, foram transformadas em *E. coli* eletrocompetente, cepa TOP10,. A partir das colônias crescidas, realizou-se um PCR de colônia, conforme descrito no item 3.9, seguido da eletroforese em gel de agarose 1%, para a confirmação da presença dos insertos alvos nas respectivas colônias. Como pode ser observado na figura 4.10, foram isoladas três colônias positivas para *TcGK* e quatro colônias positivas para *TcUPRT*.



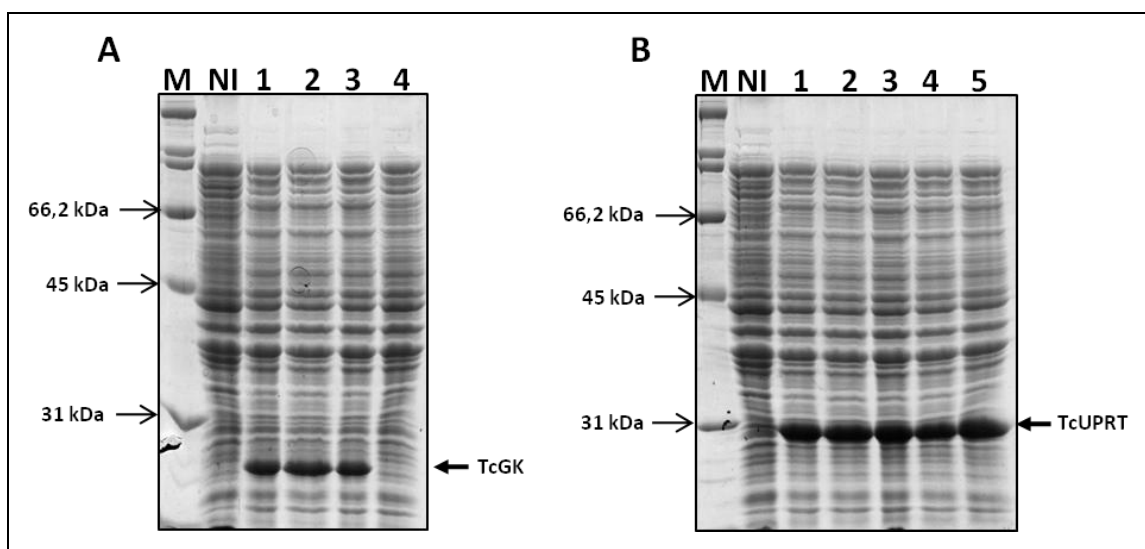
**Figura 4.10:** Análise do PCR de colônia para a confirmação da clonagem dos genes *TcGK* e *TcUPRT* no vetor pET28a. Gel de agarose 1%. **(A)** Poço M - marcador de peso molecular 1kb DNA Ladder (Invitrogen); poços 1, 2 e 7 - colônias positivas para a presença de *TcGK*; poços 3 a 6 - colônias negativas para *TcGK*. **(B)** Poço M - marcador de peso molecular 1kb DNA Ladder (Invitrogen); poços 1 a 4 - colônias positivas para *TcUPRT*; poço 5 - colônia negativa para *TcUPRT*;

#### 4.8. Expressão em pequena escala das proteínas recombinantes TcGK e TcUPRT

As colônias que apresentaram resultado positivo para a presença do inserto alvo foram submetidas à extração de DNA plasmidial. Procedeu-se, em seguida, a transformação com essas construções em *E. coli*, cepa BL21 (DE3), e posterior plaqueamento em meio LB Ágar com kanamicina. A partir das colônias obtidas, foram escolhidas quatro colônias para o alvo TcGK e cinco colônias para o alvo TcUPRT para a expressão em pequena escala das proteínas recombinantes TcGK e TcUPRT, respectivamente. As colônias foram induzidas com 1mM de IPTG em meio de LB com kanamicina à 37 °C por 3 horas. Para o controle não induzido, uma colônia representante de cada alvo foi cultivada nas mesmas condições que as colônias induzidas, mas sem a indução com IPTG. Os resultados das expressões foram analisados por meio de SDS-PAGE 15%, como pode ser visualizado na Figura 4.11.

Na figura 4.11, pode-se observar que as colônias de ambos os alvos induzidas com IPTG apresentaram expressão das proteínas recombinantes, exceto a colônia referente ao alvo TcGK localizada no poço 4 (fig. 4.11A). A falta de expressão da proteína recombinante nessa colônia pode indicar a ausência do inserto *TcGK* no vetor pET28a transformado na colônia em questão. Comparações com os controles não induzidos indicam que as proteínas recombinantes TcGK e

TcUPRT foram expressas em grandes quantidades. O peso molecular da proteína recombinante TcGK foi similar ao peso molecular esperado de 23 kDa. Da mesma forma, a proteína recombinante TcUPRT foi expressa com peso molecular compatível ao predito (28 kDa). Esses pesos moleculares se referem aos tamanhos aproximados das proteínas recombinantes fusionadas à cauda de poli-histidina.



**Figura 4.11:** Análise da expressão total em pequena escala das proteínas recombinantes TcGK e TcUPRT. SDS 15%. **(A)** Poço M - padrão de peso molecular *Broad Range* (BIO-RAD); poço NI - colônia não induzida com IPTG; poço 1 a 3 - extrato total de colônias induzidas com 1 mM IPTG, exibindo a expressão da proteína recombinante TcGK em torno de 23,2 kDa; poço 4 - colônia induzida com 1 mM IPTG, porém não há indício da expressão da proteína recombinante TcGK. **(B)** Poço M - padrão de peso molecular *Broad Range* (BIO-RAD); poço NI - colônia não induzida com IPTG; poços 1 a 5 - extrato total de colônias induzidas com 1 mM IPTG, exibindo a expressão da proteína recombinante TcUPRT em torno de 28,7 kDa.

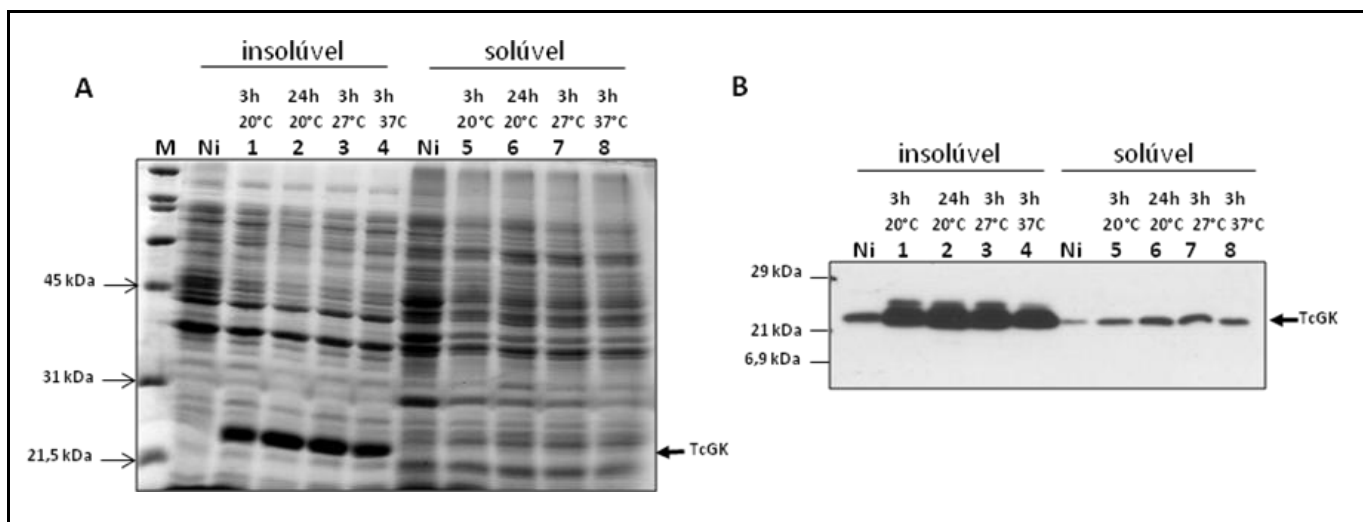
#### 4.9. Sequenciamento das construções *TcGK*-pET28a e *TcUPRT*-pET28a

Todas as colônias submetidas à expressão em pequena escala tiveram seus DNAs plasmidiais sequenciados pela Plataforma Genômica – Sequenciamento de DNA (RPT01A), do PDTIS/FIOCRUZ, utilizando os *primers* T7 *forward* e T7 *reverse* para o vetor pET28a. O sequenciamento visou avaliar a integridade das construções plasmidiais, assegurando a clonagem dos genes *TcGK* e *TcUPRT* no vetor pET28a. A partir disso, foi possível escolher apenas uma colônia de cada alvo para dar continuidade aos testes de expressão. Os insertos alvos das colônias escolhidas apresentaram 100% de identidade com as suas sequências de referência e se ligaram corretamente ao vetor pET28a.

#### **4.10. Testes de expressão e solubilidade das proteínas recombinantes TcGK e TcUPRT em pequena escala, seguidos de lise bacteriana e fracionamento das proteínas totais**

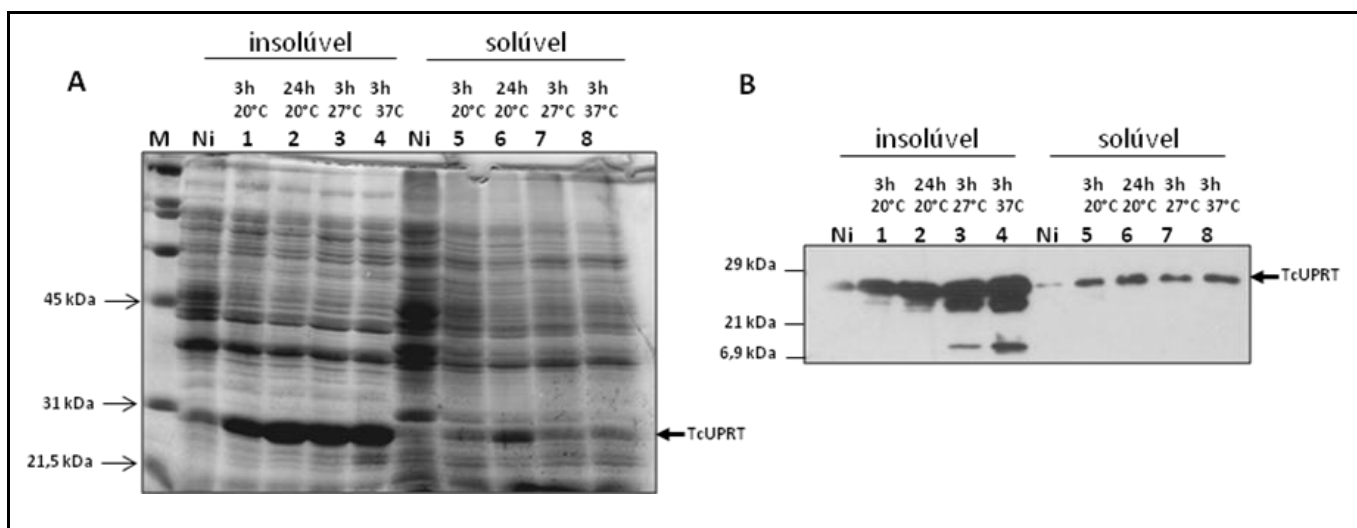
As colônias selecionadas foram submetidas a testes de solubilidade a fim de determinar se as proteínas recombinantes TcGK e TcUPRT estavam sendo expressas na forma solúvel e quais as melhores condições de temperaturas para esse tipo de expressão. Para a realização desses testes, as bactérias referentes às colônias selecionadas foram induzidas com 1 mM de IPTG por três horas à 20 °C, 27 °C e 37 °C e por 24 horas à 20 °C. Posteriormente, as bactérias foram lisadas mecanicamente e separadas em frações de proteínas solúveis e insolúveis por meio de centrifugação, conforme descrito no item 3.17, sendo ambas as frações analisadas por SDS-PAGE 15% (figura 4.12). Para uma análise mais apurada do perfil de expressão das proteínas recombinantes nas diferentes frações, realizou-se também um *western blot*, no qual se utilizou um anticorpo monoclonal anti-histidina como anticorpo primário. A utilização do anticorpo primário anti-histidina foi possível uma vez que as proteínas recombinantes expressas em pET28a contém uma cauda de poli-histidina fusionada a suas regiões N-terminal.

A partir da análise em SDS-PAGE 15% do teste de solubilidade do alvo TcGK (figura 4.12A), observa-se que, em todas as condições de temperatura, as amostras das frações de proteínas insolúveis apresentaram uma grande expressão da proteína recombinante TcGK, com peso molecular estimado de 23 kDa, quando comparadas ao controle não induzido com IPTG (poço Ni, figura 4.12A). Ao contrário do que é observado nas frações insolúveis, nas frações solúveis a proteína recombinante TcGK foi de difícil detecção. Com o auxílio do *western blot* (figura 4.12B), que é um método mais específico e sensível na detecção de proteínas, verificou-se uma expressão basal da proteína recombinante no controle não induzido. Além disso, foi possível observar que a proteína recombinante TcGK também estava sendo expressa na forma solúvel, não havendo diferenças significativas nos níveis de expressão dessa proteína na forma solúvel com a alteração das condições de indução. Apesar disso, escolheu-se como condição para a expressão da proteína recombinante TcGK em larga escala a indução com 1 mM de IPTG à 20 °C por 24 horas.



**Figura 4.12:** Análise dos testes de expressão em pequena escala da proteína TcGK recombinante de 23 kDa. **(A)** Análise por SDS-PAGE 15%. Poço M - padrão de peso molecular *Broad Range* (BIO-RAD); poço Ni - controle não induzido com IPTG correspondente às frações solúvel e insolúvel; poço 1 - fração insolúvel do induzido com IPTG à 20 °C por 3 horas; poço 2 - fração insolúvel do induzido com IPTG à 20 °C por 24 horas; poço 3 - fração insolúvel do induzido com IPTG à 27 °C por 3 horas; poço 4 - fração insolúvel do induzido com IPTG à 37 °C por 3 horas; poço 5 - fração solúvel do induzido com IPTG à 20 °C por 3 horas; poço 6 - fração solúvel do induzido com IPTG à 20 °C por 24 horas; poço 7 - fração solúvel do induzido com IPTG à 27 °C por 3 horas; poço 8 - fração solúvel do induzido com IPTG à 37 °C por 3 horas. **(B)** Análises por *western blot*: as amostras estão dispostas conforme a figura A. O padrão de peso molecular utilizado no *western blot* foi o *Broad Range Prestained* (BIO-RAD).

As análises das diferentes condições de expressão da proteína recombinante UPRT por SDS-PAGE 15% (figura 4.13A) e por *western blot* (figura 4.13B), indicam que a proteína recombinante UPRT com o peso molecular estimado de 28 kDa foi expressa majoritariamente em corpos de inclusão, se localizando, por isso, na fração insolúvel. Embora a proteína recombinante UPRT nas frações solúveis estivesse em pequena quantidade, ela é visível em SDS-PAGE 15% principalmente na condição de indução com 1 mM de IPTG à 20 °C por 24 horas (figura 4.13A, poço 6). Os ensaios de *western blot* (figura 4.13B), da mesma forma que para a proteína recombinante TcGK, indicaram uma expressão basal da proteína recombinante TcUPRT no controle não induzido, que pode acontecer mesmo na ausência de IPTG, segundo o manual do vetor pET. Com os resultados do *western blot* e do SDS-PAGE foi possível escolher como a melhor condição de expressão da proteína recombinante TcUPRT na forma solúvel a indução com 1 mM de IPTG à 20 °C por 24 horas.

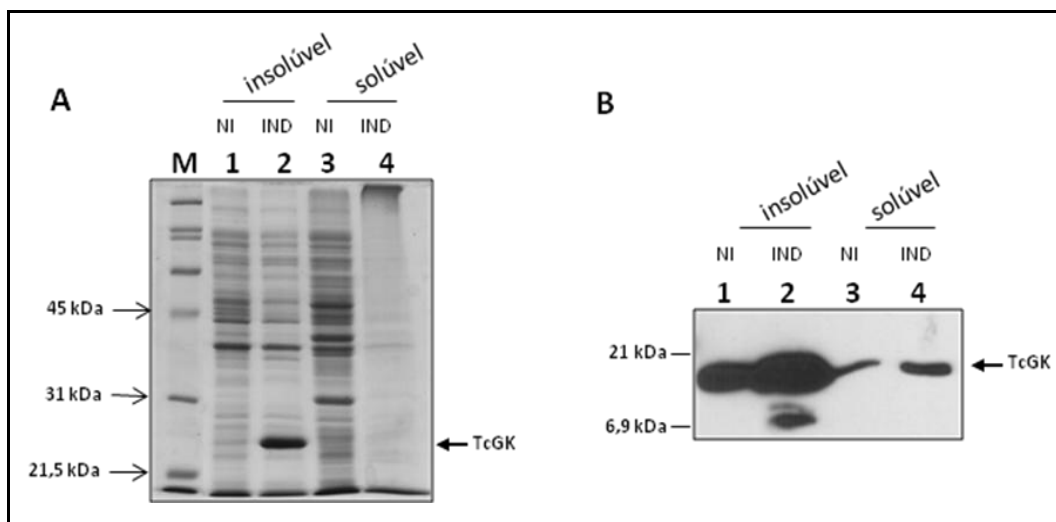


**Figura 4.13:** Análise dos testes de expressão em pequena escala da proteína TcUPRT recombinante de aproximadamente 28,7 kDa. **(A)** Análises por SDS-PAGE 15%. Poço M - padrão de peso molecular *Broad Range* (BIO-RAD); poço Ni - controle não induzido com IPTG correspondente as frações solúvel e insolúvel; poço 1 - fração insolúvel do induzido com IPTG à 20 °C por 3 horas; poço 2 - fração insolúvel do induzido com IPTG à 20 °C por 24 horas; poço 3 - fração insolúvel do induzido com IPTG à 27 °C por 3 horas; poço 4 - fração insolúvel do induzido com IPTG à 37 °C por 3 horas; poço 5 - fração solúvel do induzido com IPTG à 20 °C por 3 horas; poço 6 - fração solúvel do induzido com IPTG à 20 °C por 24 horas; poço 7 - fração solúvel do induzido com IPTG à 27 °C por 3 horas; poço 8 - fração solúvel do induzido com IPTG à 37 °C por 3 horas. **(B)** Análises por *western blot*: as amostras estão dispostas conforme a figura A. O padrão de peso molecular utilizado no *western blot* foi o *Broad Range Prestained* (BIO-RAD).

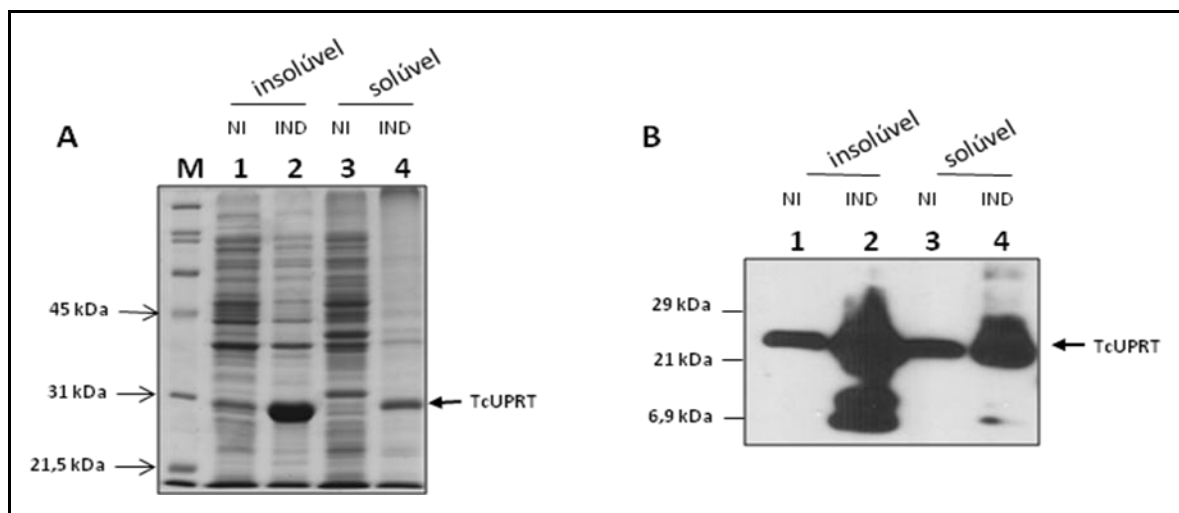
#### 4.11. Expressão em larga escala das proteínas recombinantes TcGK e TcUPRT

Uma vez escolhidas as melhores condições de expressão das proteínas recombinante TcGK e TcUPRT nas formas solúveis, realizou-se a expressão em larga escala dessas proteínas. Posteriormente, foi feita a separação das proteínas em frações solúvel e insolúvel. Os resultados foram analisados, assim como na expressão em pequena escala, através de SDS-PAGE 15% e *western blot*.

Mantendo o mesmo perfil observado na expressão em pequena escala, a proteína recombinante TcGK expressa em larga escala apresentou o peso molecular esperado de 23 kDa e se localizou majoritariamente na fração insolúvel como corpos de inclusão (figura 4.14A e figura 4.14B). Embora a proteína recombinante TcGK não tenha sido detectada na fração solúvel analisada por SDS-PAGE 15% (figura 4.14A), as análises por *western blot* revelaram que a proteína recombinante TcGK também foi expressa na forma solúvel (figura 4.14B, poço 4).



**Figura 4.14:** Análise da expressão em larga escala da proteína TcGK recombinante de 23,2 kDa. **(A)** Análise da expressão por SDS-PAGE 15%. Poço M - padrão de peso molecular *Broad Range* (BIO-RAD); poço 1 - controle não induzido com IPTG correspondente a fração insolúvel; poço 2 - induzido com IPTG referente a fração insolúvel; poço 3 - controle não induzido com IPTG referente a fração solúvel; poço 4 - induzido com IPTG referente a fração solúvel. **(B)** Análises por *western blot*: as amostras estão dispostas conforme a figura A. O padrão de peso molecular utilizado no *western blot* foi o *Broad Range Prestained* (BIO-RAD).



**Figura 4.15:** Análise da expressão em larga escala da proteína TcUPRT recombinante de 28,7 kDa. **(A)** Análise da expressão por SDS-PAGE 15%. Poço M - padrão de peso molecular *Broad Range* (BIO-RAD); poço 1 - controle negativo correspondente a fração insolúvel; poço 2 - induzido com IPTG referente a fração insolúvel; poço 3 - controle não induzido com IPTG referente a fração solúvel; poço 4 - induzido com IPTG referente a fração solúvel. **(B)** Análises por *western blot*: as amostras estão dispostas conforme a figura A. O padrão de peso molecular utilizado no *western blot* foi o *Broad Range Prestained* (BIO-RAD).

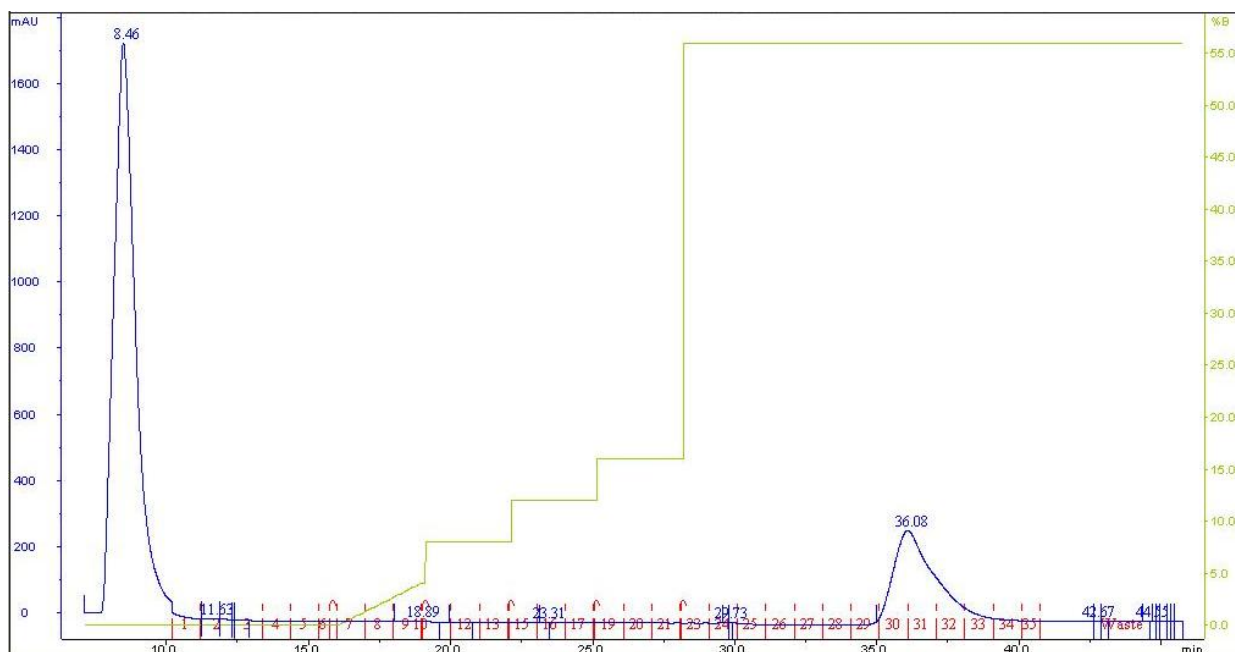


#### 4.12. Purificação das proteínas recombinantes por cromatografia de afinidade

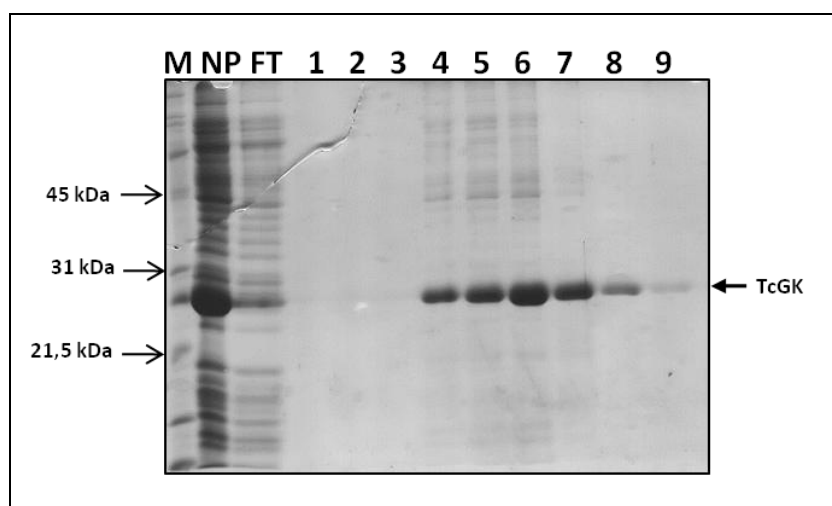
Como as proteínas recombinantes foram obtidas majoritariamente na forma insolúvel, elas foram submetidas à solubilização com ureia 8M. Após a solubilização, foi feita a purificação das proteínas recombinantes TcGK e TcUPRT por cromatografia de afinidade, utilizando a coluna HisTrap HP (GE), carregada com íons  $\text{Ni}^{2+}$ . Os cromatogramas obtidos nas purificações das proteínas recombinantes TcGK e TcUPRT estão representados, respectivamente, nas figuras 4.16 e 4.18. A partir dos perfis cromatográficos obtidos, foram analisados em SDS-PAGE 15% as frações em que se obteve detecção de proteínas eluídas, representadas pelos picos no cromatograma, além do volume de exclusão (*flow through*).

Com o cromatograma da purificação de TcGK (figura 4.16), escolheu-se as frações 5, 25 e de 28 à 34 para serem analisadas por SDS-PAGE 15% (figura 4.17). Por meio dessa análise, confirmou-se que houve um bom grau de pureza da proteína recombinante TcGK purificada, principalmente nas frações 32, 33 e 34, em que se obteve a proteína recombinante com menores níveis de proteínas contaminantes, sendo dessa forma reunidas em um único *pool*.

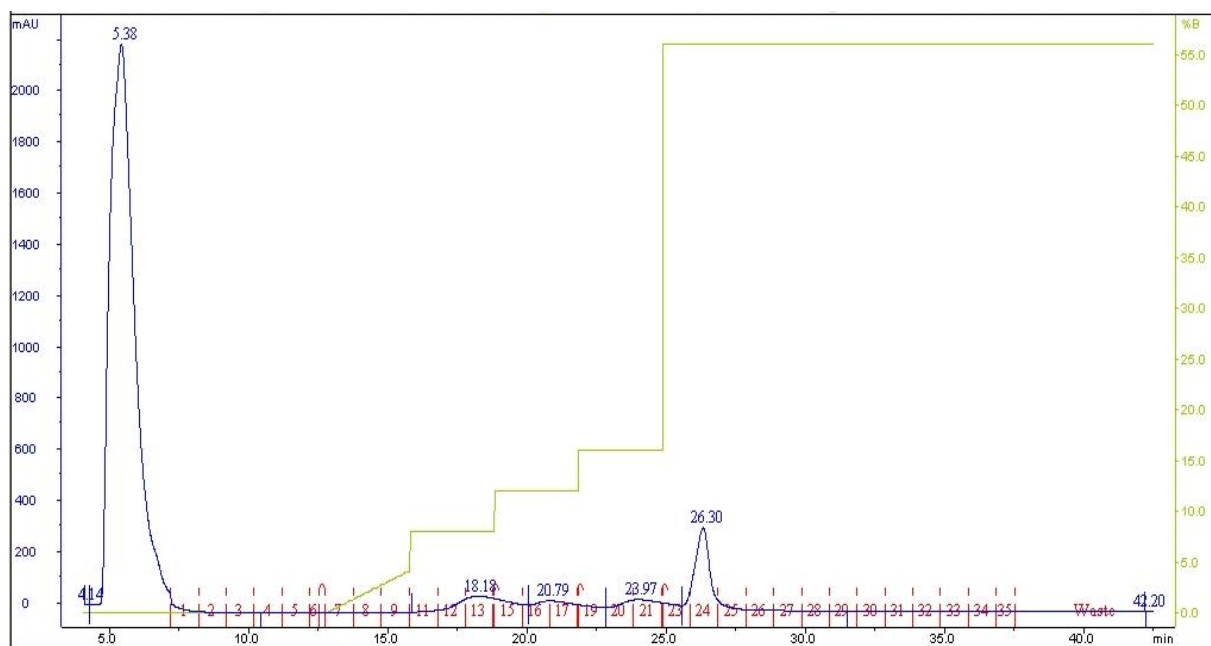
O perfil de cromatografia obtida com a purificação de TcUPRT (figura 4.18) permitiu a escolha das frações 12, 13, 15, 20, 21 e de 23 à 25, para serem analisadas por SDS-PAGE 15%. Como se pode ver na figura 4.19, a maioria das frações analisadas apresentou a proteína recombinante TcUPRT em um bom grau de pureza, exceto a fração 23. A partir dessas frações que apresentaram a proteína recombinante com pureza satisfatória, fez-se um único *pool* de proteína recombinante TcUPRT purificada.



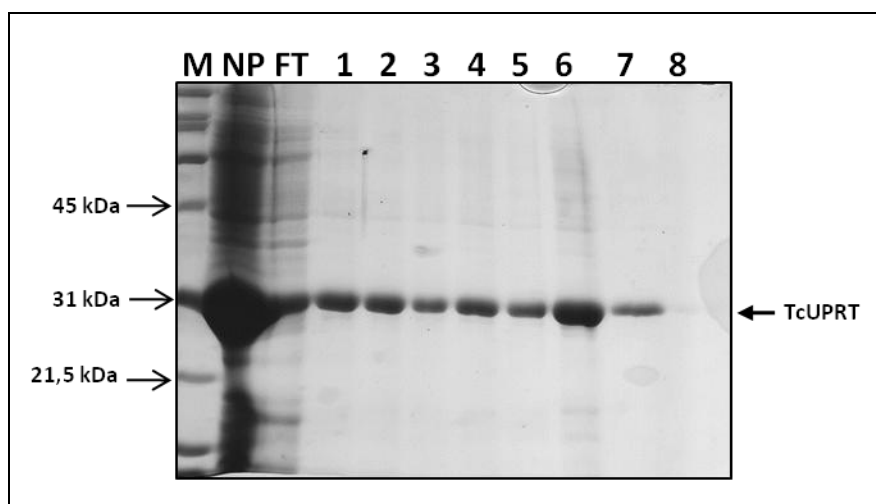
**Figura 4.16:** Cromatograma da purificação da proteína recombinante TcGK. A proteína recombinante TcGK purificada por cromatografia de afinidade, utilizando a coluna HisTRAP-HP carregada com  $\text{Ni}^{2+}$ , utilizando o AKTA Purifier (GE). O eixo x corresponde ao tempo de retenção das proteínas na coluna, e os números acima dos picos correspondem ao tempo de retenção referente aos mesmos; o eixo y da esquerda, representa a absorvância em 280 nm, assim como a curva em azul demonstra o perfil de eluição ao longo da purificação; já o eixo y da direita, corresponde a porcentagem do tampão de eluição, o qual contém 500 mM de imidazol inicialmente. A curva em verde representa o gradiente descontínuo de imidazol utilizado na eluição das proteínas. Em vermelho estão numeradas as frações (1 mL) coletadas.



**Figura 4.17:** Análise da purificação por cromatografia de afinidade da proteína recombinante TcGK. SDS-PAGE 15%. Poço M - padrão *Broad range* (BIO-RAD); poço NP - proteína recombinante solubilizada com ureia 8M ainda não submetida à purificação; poço FT - *flow through*; poços 1, 2, 3, 4, 5, 6, 7, 8 e 9 - amostras referentes às frações 05, 25, 28, 29, 30, 31, 32, 33 e 34, respectivamente.



**Figura 4.18:** Cromatograma da purificação da proteína recombinante TcUPRT. A proteína recombinante TcUPRT purificada por cromatografia de afinidade, utilizando a coluna HisTrap HP, carregada com Ni<sup>2+</sup>, utilizando o AKTA Purifier (GE). O eixo x corresponde ao tempo de retenção das proteínas na coluna, e os números acima dos picos correspondem ao tempo de retenção referente aos mesmos; o eixo y da esquerda, representa a absorbância em 280 nm, assim como a curva em azul demonstra o perfil de eluição ao longo da purificação; já o eixo y da direita, corresponde a percentagem do tampão de eluição, o qual contém 500 mM de imidazol inicialmente. A curva em verde representa o gradiente descontínuo de imidazol utilizado na eluição das proteínas. Em vermelho estão numeradas as frações (1 mL) coletadas.

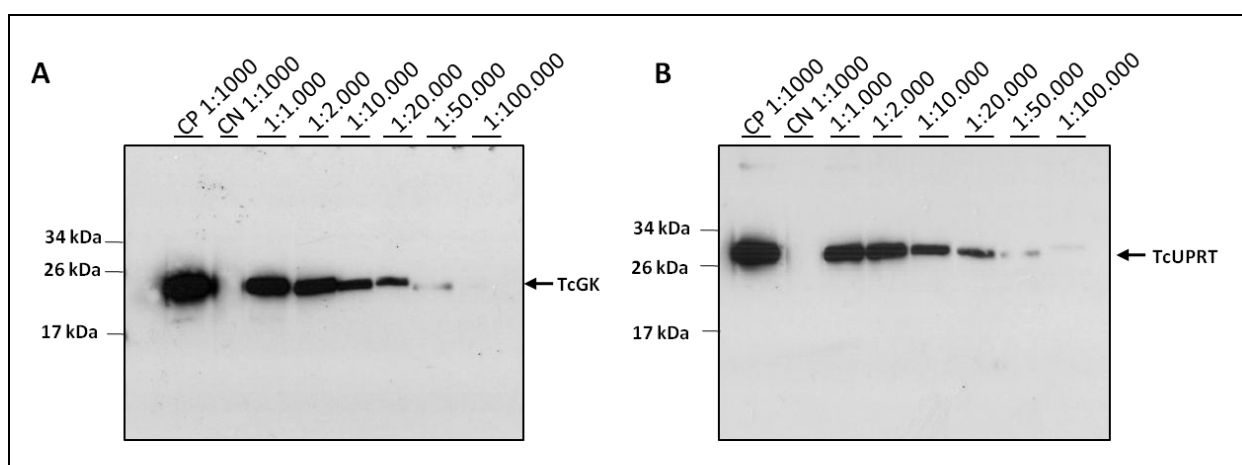


**Figura 4.19:** Análise da purificação por cromatografia de afinidade da proteína recombinante TcUPRT. SDS-PAGE 15%. Poço M - padrão *Broad range* (BIO-RAD); poço NP - proteína recombinante solubilizada com ureia 8M ainda não submetida à purificação; poço FT - *flow through*; poços 12, 13, 15, 20, 21, 23, 24 e 25 - amostras referentes as frações 12, 13, 15, 20, 21, 23, 24 e 25, respectivamente.

#### 4.13. Titulação dos anticorpos policlonais anti-TcGK e anti-TcUPRT

Após a dosagem pelo método de RCDC das concentrações das proteínas recombinantes TcGK e TcUPRT purificadas, foram realizadas imunizações em camundongos Balb-C, como descrito no item 3.21. Após a imunização, a análise da produção de anticorpos policlonais anti-TcGK e anti-TcUPRT em cada camundongo foi realizada por meio de *western blot* contra a respectiva proteína recombinante purificada. O resultado do *western blot* (dado não mostrado) permitiu a confirmação da produção dos anticorpos policlonais anti-TcGK e anti-TcUPRT de forma semelhante em todos os camundongos. Dessa forma, para cada alvo, os soros imunes obtidos a partir da imunização dos camundongos foram reunidos em um único *pool*.

A titulação do *pool* dos soros imunes, referente a cada alvo, foi realizada nas diluições de 1:1.000, 1:2.000, 1:10.000, 1:20.000, 1:50.000 e 1:100.000 e testada por *western blot* contra a proteína alvo recombinante purificada. As figuras 4.20A e 4.20B demonstram os resultados das titulações utilizando as diferentes diluições dos soros imunes, conforme descrito anteriormente. Pode-se notar que os anticorpos policlonais anti-TcGK conseguem identificar a proteína recombinante TcGK até uma diluição de 1:50.000. Por outro lado, os anticorpos policlonais anti-TcUPRT identificaram a proteína recombinante TcUPRT até uma diluição de 1:100.000.



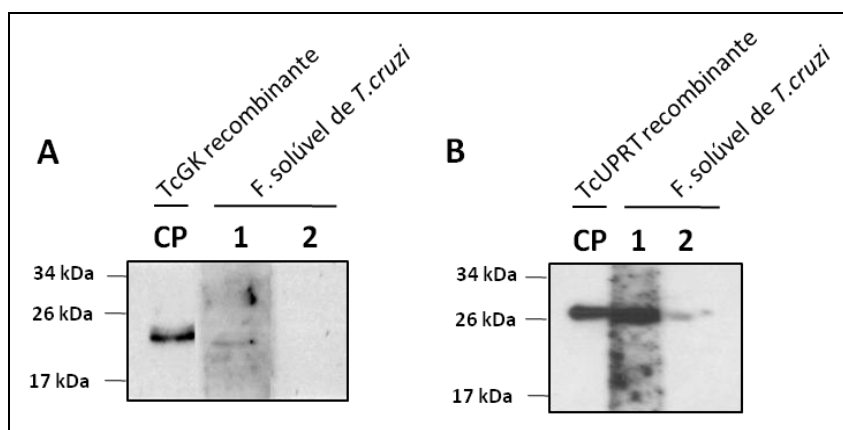
**Figura 4.20:** Titulação dos anticorpos policlonais anti-TcGK e anti-TcUPRT. Análise da titulação dos anticorpos policlonais anti-TcGK e anti-TcUPRT por *western blot* utilizando como controle positivo (CP) anticorpo anti-His, como controle negativo (CN) o soro pré-imune. As diluições estão descritas acima de cada poço em ambos os painéis. O padrão de peso molecular utilizado foi o *Broad Range Prestained* (BIO-RAD). **(A)** Titulação do anticorpo policlonal anti-TcGK em diferentes diluições contra a proteína recombinante TcGK. **(B)** Titulação do anticorpo policlonal anti-TcUPRT em diferentes diluições contra a proteína recombinante TcUPRT.

#### 4.14. Identificação das enzimas nativas TcGK e TcUPRT nas formas epimastigotas de *T. cruzi* por *western blot*

Para identificar as proteínas TcGK e TcUPRT nativas nas formas epimastigotas de *T. cruzi*, realizou-se à lise e extração das proteínas solúveis dessas formas evolutivas, seguida de um ensaio por *western blot* com os soros imunes contendo os anticorpos policlonais anti-TcGK e anti-TcUPRT produzidos. Os soros imunes anti-TcGK e anti-TcUPRT foram utilizados na diluição de 1:100 contra as frações solúveis do parasito e na diluição de 1:20.000 contra as respectivas proteínas recombinantes purificadas. As IgGs policlonais purificadas, contendo as IgGs anti-TcGK e anti-TcUPRT, foram utilizadas na diluição de 1:10.

Uma proteína nativa com peso molecular de 21 kDa, como predito para TcGK nativa, foi identificada na fração solúvel de *T. cruzi* no ensaio em que se utilizou o soro imune contendo anticorpos policlonais anti-TcGK (figura 4.21A, poço 1). Já no ensaio com os IgGs purificados anti-TcGK, não houve detecção da proteína nativa TcGK e nem de nenhuma outra proteína nativa de *T. cruzi* (figura 4.21A, poço 2). Esse resultado provavelmente se deve à baixa concentração de anticorpos policlonais anti-TcGK após as etapas de purificação das imunoglobulinas.

Com relação à identificação de TcUPRT, uma proteína com peso molecular de 26 kDa, correspondente ao estimado para a proteína TcUPRT nativa, foi identificada na fração solúvel de *T. cruzi* (figura 4.21B) nos ensaios utilizando tanto o soro imune quanto os IgGs purificados, ambos contendo os anticorpos policlonais anti-TcUPRT. Os ensaios com os IgGs purificados anti-TcUPRT permitiram a detecção de uma proteína que provavelmente corresponde a TcUPRT nativa, mas com uma marcação bem menos intensa do que quando comparada a detecção pelo soro imune anti-TcUPRT não purificado.

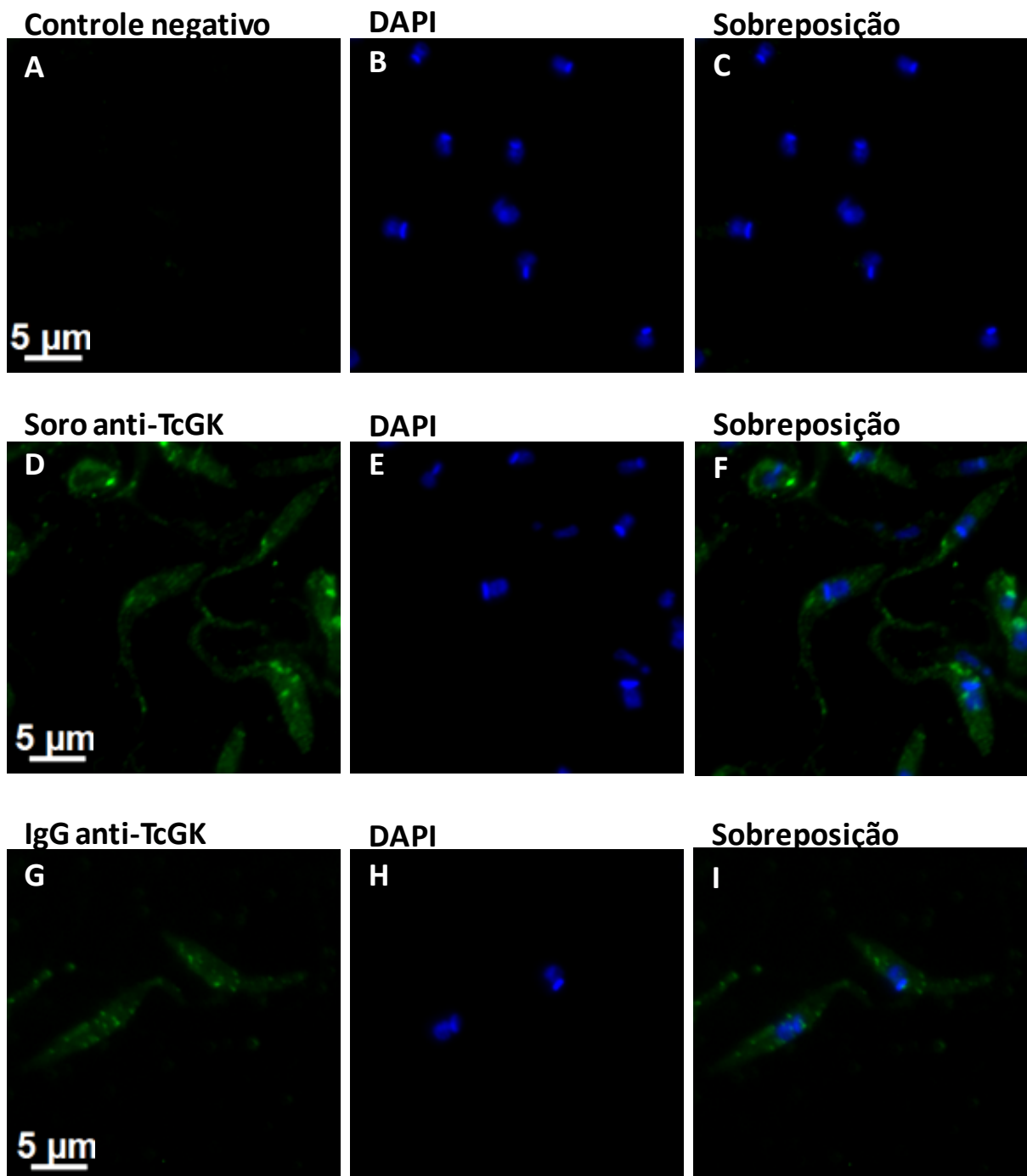


**Figura 4.21:** Identificação por *western blot* das enzimas nativas TcGK e TcUPRT nas frações solúveis das formas epimastigotas de *T. cruzi*, cepa CL Brener. O padrão de peso molecular utilizado foi o *Broad Range Prestained* (BIO-RAD). **(A)** Identificação por *western blot* da enzima nativa TcGK de 21 kDa: poço CP: controle positivo utilizando anticorpo primário anti-TcGK na diluição de 1:20.000 contra a proteína recombinante TcGK purificada de 23,2 kDa; poço 1- anticorpo primário anti-TcGK na diluição 1:100 contra a fração solúvel de *T. cruzi*; poço 2- anticorpo primário anti-TcGK purificado na diluição 1:10 contra a fração solúvel de *T. cruzi*. **(B)** Identificação por *western blot* da enzima nativa TcUPRT de 26 kDa: poço CP: controle positivo utilizando anticorpo primário anti-TcUPRT na diluição de 1:20.000 contra a proteína recombinante TcUPRT purificada de 28,7 kDa; poço 1 - anticorpo primário anti-TcUPRT na diluição 1:100 contra a fração solúvel de *T. cruzi*; poço 2 - anticorpo primário anti-TcUPRT purificado na diluição 1:10 contra a fração solúvel de *T. cruzi*.

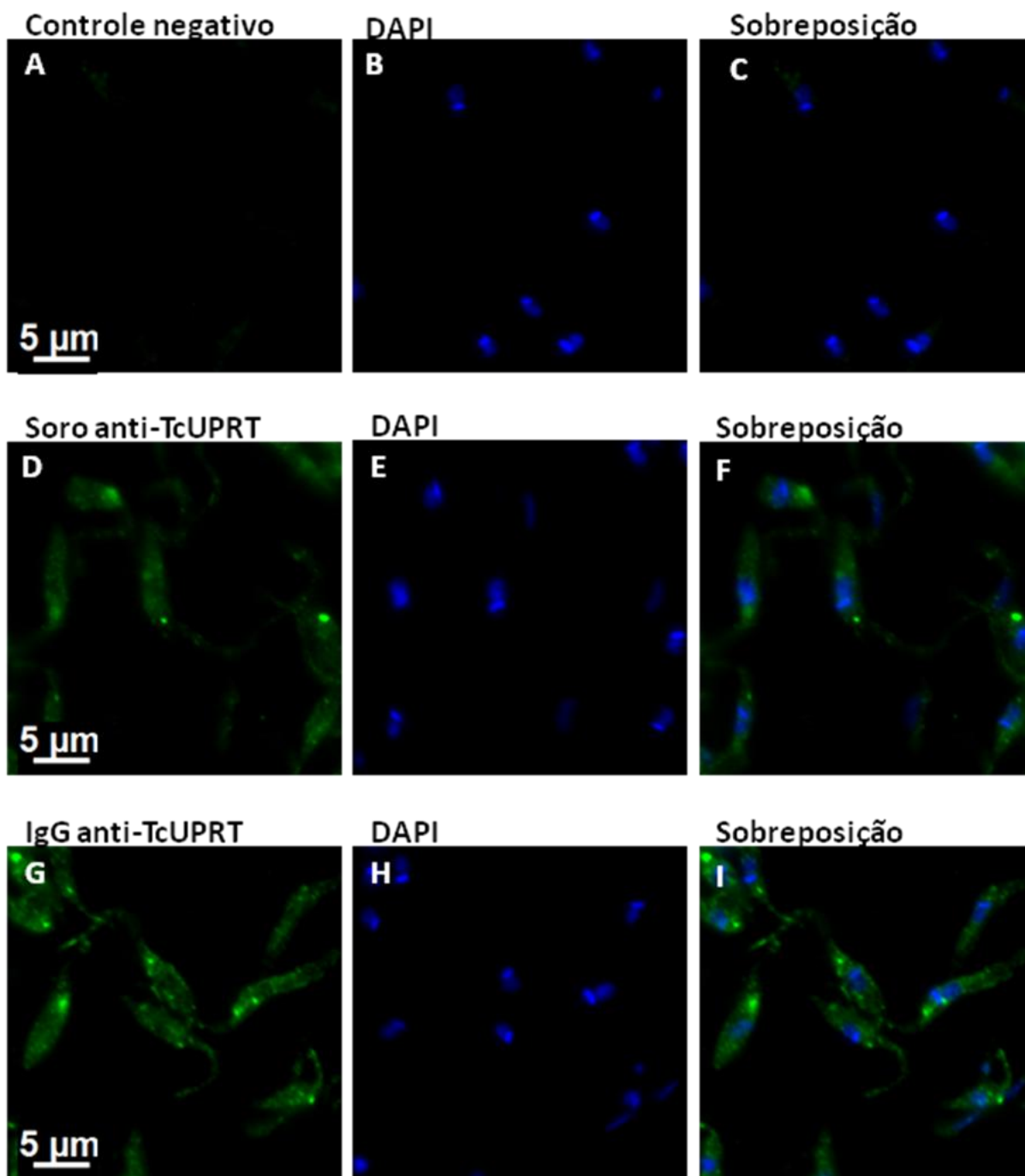
#### 4.15. Identificação das enzimas nativas TcGK e TcUPRT nas formas epimastigotas de *T. cruzi* por imunofluorescência

Com o objetivo de identificar a possível localização celular das proteínas TcGK e TcUPRT nativas na forma epimastigota de *T. cruzi*, deu-se início a ensaios de imunofluorescência indireta. Para esses ensaios, os parasitos foram marcados com os soros imunes contendo os anticorpos policlonais anti-TcGK e anti-TcUPRT nas diluições de 1:25, 1:50 e 1:100 e com os IgGs purificados anti-TcGK e anti-TcUPRT na diluição de 1:25, conforme descrito no item 3.27. As melhores imagens foram obtidas com as diluições de 1:100 dos soros imunes anti-TcGK e anti-TcUPRT e com as diluições 1:25 dos anticorpos obtidos após a purificação desses soros.

Para a proteína TcGK nativa (figura 4.22), observou-se uma marcação difusa no citoplasma e algumas marcações concentradas em vesículas. No caso da proteína TcUPRT nativa (figura 4.23) também se observou uma marcação ao longo do citoplasma e em vesículas, além de uma marcação proeminente próxima à região da bolsa flagelar. Em todos os ensaios foi possível verificar a marcação com o DAPI do núcleo e do cinetoplasto.



**Figura 4.22:** Identificação da enzima nativa GK na forma epimastigota de *T. cruzi* por imunofluorescência. **(A, B, C)** Controle negativo, sem incubação do parasito com soro/ IgG anti-TcGK; **(D, E, F)** Marcação do *T. cruzi* com o soro anti-TcGK; **(G, H, I)** Marcação do *T. cruzi* com a IgG purificadas anti-TcGK; **(B, E, H)** Marcação dos epimastigotas com DAPI; **(C, F, I)** Sobreposição das imagens A e B, D e E, G e H, respectivamente.

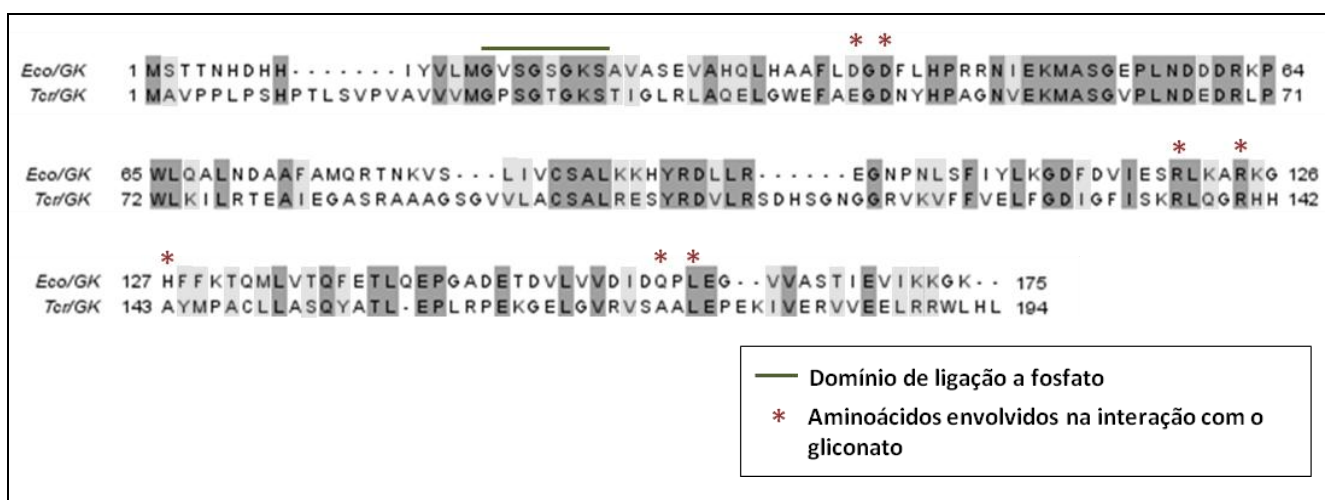


**Figura 4.23:** Identificação da enzima nativa UPRT na forma epimastigota de *T. cruzi* por imunofluorescência. **(A, B, C)** Controle negativo, sem incubação do parasito com soro/ IgG anti-TcUPRT; **(D, E, F)**. Marcação do *T. cruzi* com o soro anti-TcUPRT; **(G, H, I)** Marcação do *T. cruzi* com a IgG purificadas anti-TcUPRT; **(B, E, H)** Marcação dos epimastigotas com DAPI; **(C, F, I)** Sobreposição das imagens A e B, D e E, G e H, respectivamente.



#### 4.16. Análise das regiões de ligação ao substrato na estrutura primária de GK de *T. cruzi*

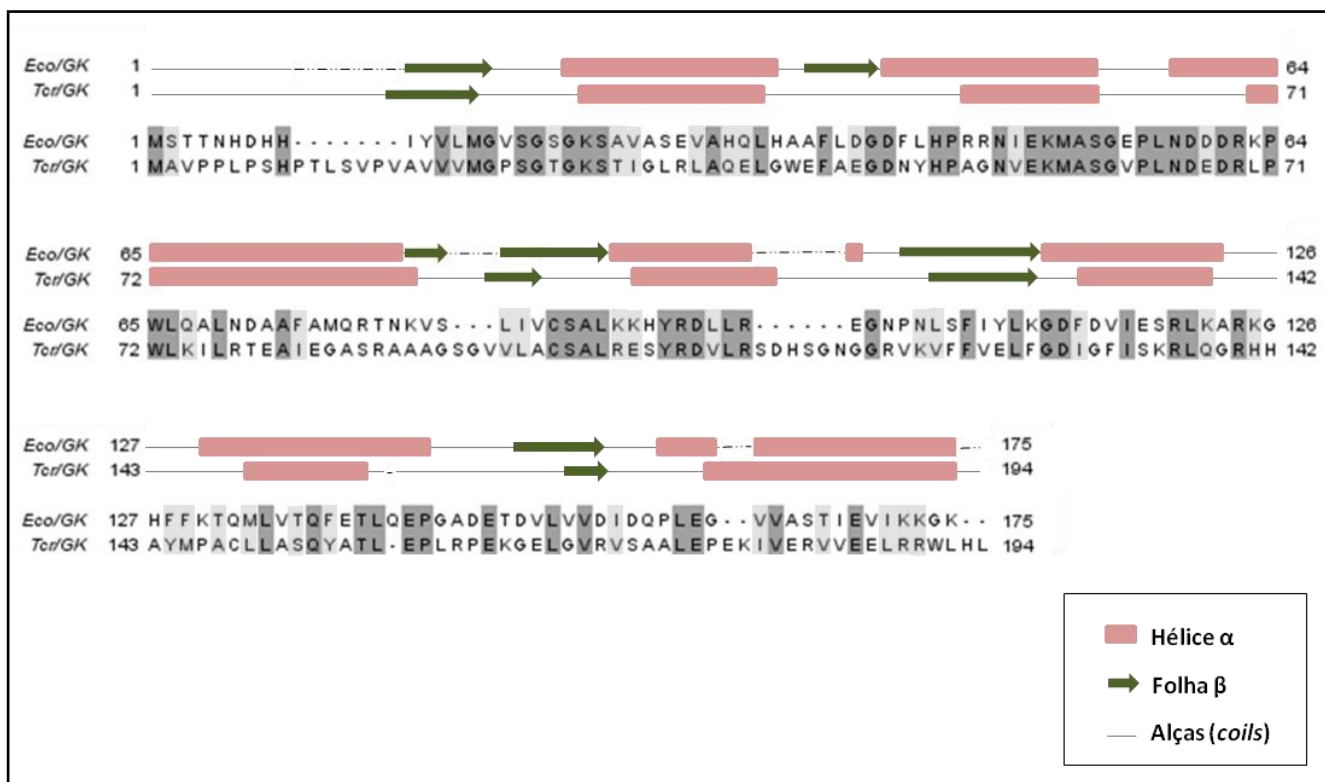
Um alinhamento múltiplo foi realizado entre a enzima TcGK e a GK de *E. coli* (figura 4.24), uma vez que essa enzima já teve sua estrutura resolvida por cristalografia. Além disso, nessa estrutura os aminoácidos envolvidos na interação com o substrato gliconato e o domínio de ligação a fosfato já foram descritos (Kraft *et al.*, 2002). A TcGK apresentou uma identidade de 38,2% com a GK de *E. coli*.



**Figura 4.24:** Alinhamento múltiplo entre as seqüências de aminoácidos da GK de *E. coli* (Eco/GK) e da *T. cruzi* (Tcr/GK). Foi utilizado para o alinhamento o ClustalW2 através do programa Jalview (Waterhouse *et al.*, 2009). Aminoácidos idênticos estão realçados em cinza escuro, substituições conservadas estão realçadas por cinza claro. O domínio de ligação a fosfato (linha verde) e os aminoácidos envolvidos na interação com gliconato (asteriscos rosa) estão marcados acima das seqüências.

#### 4.17. Predição da estrutura secundária de TcGK

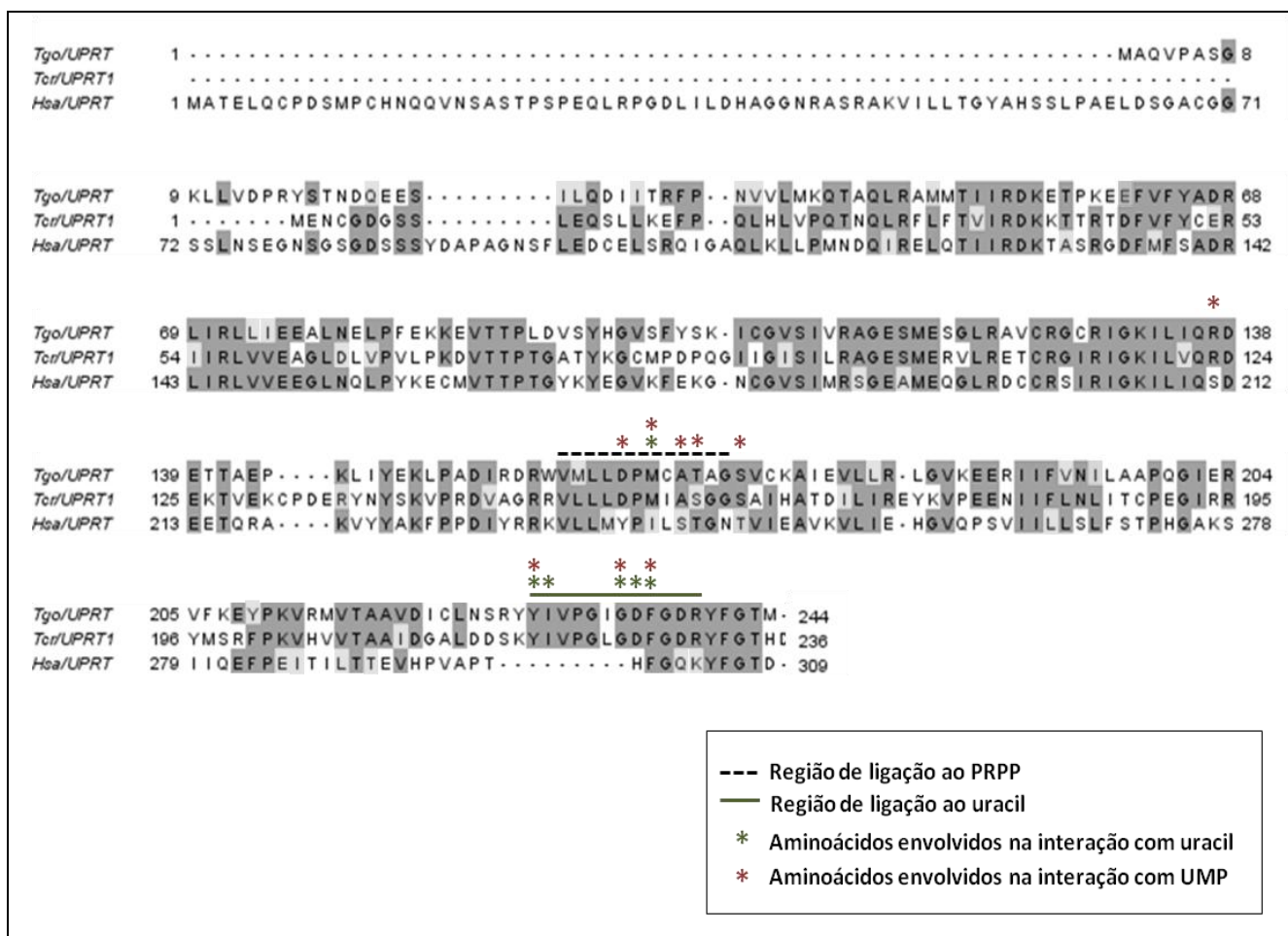
A predição da estrutura secundária da proteína TcGK foi gerada a partir da seqüência de aminoácidos dessa proteína, utilizado o programa PSIPRED. As predições para proteína TcGK apontam a existência de sete prováveis domínios hélice  $\alpha$  e quatro domínios de folha beta  $\beta$ , ambos alternados por alças (*coils*) (figura 4.25). Além disso, foi realizada a comparação entre a estrutura secundária de TcGK com a GK de *E. coli*, cuja estrutura secundária já foi descrita (Kraft *et al.*, 2002).



**Figura 4.25:** Alinhamento entre as estruturas secundárias de GK de *E. coli* (Eco/GK) e de *T. cruzi* (Tcr/GK). A estrutura secundária de GK de *E. coli* está apontada como descrito na literatura. A predição da estrutura secundária de TcGK (Tcr/GK) foi feita utilizando o programa PSIPRED. O alinhamento múltiplo das estruturas primárias foi feito com o ClustalW2, através do programa Jalview. Aminoácidos idênticos estão realçados em cinza escuro, substituições conservadas estão realçadas por cinza claro.

#### 4.18. Análise das possíveis regiões de ligação ao substrato e ao produto na estrutura primária da TcUPRT e da UPRT de humano

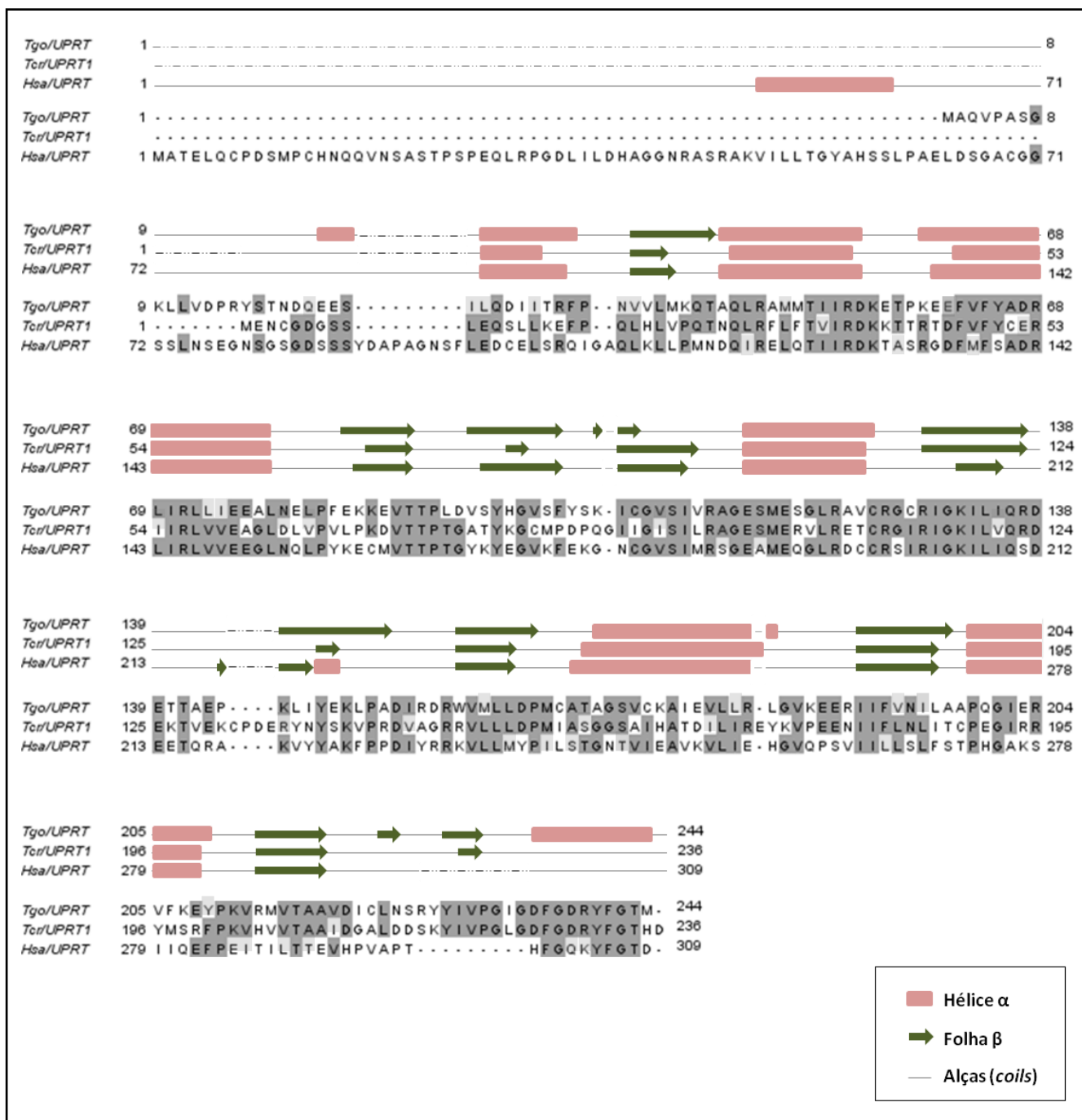
Uma vez que a proteína TcUPRT foi classificada pelo AnEnPi como sendo homóloga à UPRT de humano, foi realizado um alinhamento múltiplo, através do ClustalW2, entre as estruturas primárias da TcUPRT e da UPRT de humano e de *T. gondii*, que teve sua estrutura resolvida por cristalografia (Schumacher *et al.*, 1998). Esse alinhamento visou analisar o grau de identidade entre as sequências dos três organismos e o mapeamento das principais regiões de ligação aos substratos PRPP e uracil e ao produto UMP na TcUPRT, a partir do que foi descrito para *T. gondii* (Schumacher *et al.*, 1998) e comparar essas regiões com a UPRT de humano. A sequência de *T. gondii* apresentou uma identidade de 52,2% com a TcUPRT e de 43,7% de identidade com a UPRT de humano. As sequências de aminoácidos da UPRT de humano e da TcUPRT apresentaram uma identidade de 41,9%.



**Figura 4.26:** Alinhamento múltiplo entre as sequências de aminoácidos da UPRT de *T. gondii* (*Tgo/UPRT*), de *T. cruzi* (*Tcr/UPRT1*) e de humano (*Hsa/UPRT*). Foi utilizado para o alinhamento o ClustalW2 através do programa Jalview. Aminoácidos idênticos estão realçados em cinza escuro, substituições conservadas estão realçadas por cinza claro. A região de ligação a PRPP (linha preta tracejada), a região de ligação a uracil (linha verde) e os aminoácidos envolvidos na interação com uracil (asteriscos verdes) e UMP (asteriscos rosa) estão marcados acima das sequências.

#### 4.19. Predição da estrutura secundária da TcUPRT e da UPRT de humano

Para a predição das estruturas secundárias das proteínas TcUPRT e UPRT de humano foi utilizado o programa PSIPRED. As estruturas secundárias foram geradas a partir das sequências de aminoácidos dessas proteínas. Para a TcUPRT, foram identificados seis possíveis domínios hélice  $\alpha$  e 11 domínios folha  $\beta$ , enquanto a UPRT de humano apresentou oito potenciais domínios hélice  $\alpha$  e nove folhas  $\beta$ . Em *T. gondii* são descritas na literatura sete hélice  $\alpha$  e 11 folhas  $\beta$  (Schumacher *et al.*, 1998).

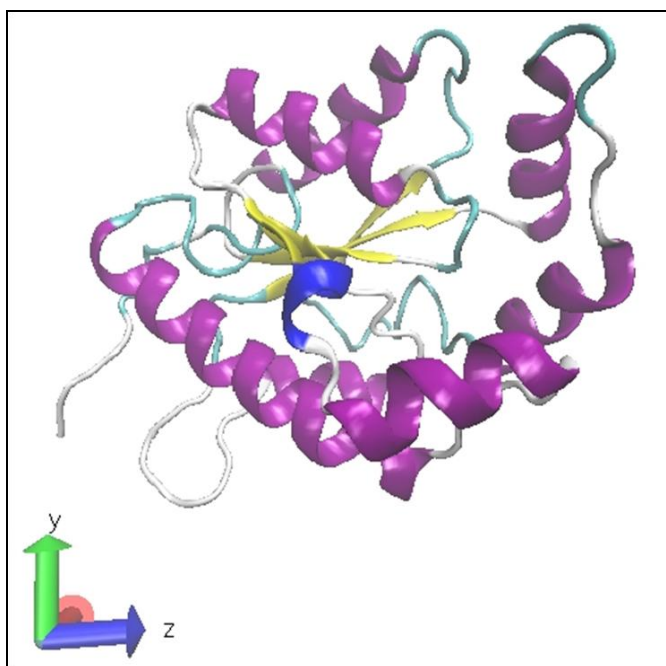


**Figura 4.27:** Alinhamento entre as estruturas secundárias de UPRT de *T. gondii* (Tgo/UPRT), de *T. cruzi* (Tcr/UPRT) e de humano (Hsa/UPRT). A estrutura secundária de *T. gondii* (Tgo/UPRT), está apontada como descrito na literatura. A predição das estruturas secundárias de *T. cruzi* (Tcr/UPRT) e *Homo sapiens* (Hsa/UPRT), foram feitas utilizando o programa PSIPRED. O alinhamento múltiplo das estruturas primárias foi feito com o ClustalW2, através do programa Jalview. Aminoácidos idênticos estão realçados em cinza escuro, substituições conservadas estão realçadas por cinza claro.

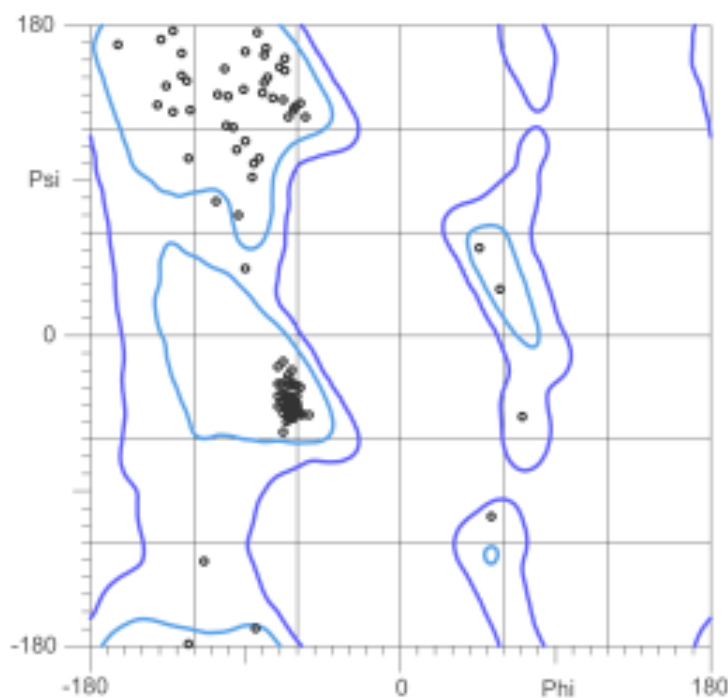
#### 4.20. Modelagem estrutural por homologia da enzima TcGK

A modelagem por homologia da enzima TcGK e TcUPRT foi feita pelo método de restrições espaciais utilizando o programa MODELLER. O molde utilizado para a modelagem da enzima TcGK foi a enzima GK de *E. coli* com estrutura resolvida por difração de raios-x, depositada no PDB sob o número de acesso 1KNQ. Esse molde apresentou uma identidade maior que 35% e cobertura entre 70% e 90% com a TcGK, o que indica que o modelo a ser gerado provavelmente possuirá boa qualidade.

A modelagem realizada para TcGK resultou em uma estrutura terciária satisfatória (figura 4.28), uma vez que a validação da estrutura com o gráfico de Ramachandran (figura 4.29), que avalia os ângulos de torções *Phi* e *Psi* nos resíduos da cadeia peptídica, apresentou 99,5% dos aminoácidos localizados nas áreas permitidas e 95,3% do total de aminoácidos localizados na região favorável. Apenas o aminoácido valina, localizado na posição 94, ficou fora dessa região. Da mesma forma, a validação baseada na verificação dos parâmetros estereoquímicos da estrutura, utilizando o programa PROCHECK, que analisou a geometria total e resíduo por resíduo, confirmou a qualidade do modelo gerado.



**Figura 4.28:** Estrutura tridimensional de enzima TcGK determinada por modelagem por homologia.



**Figura 4.29:** Gráfico Ramachandran da modelagem de TcGK. Esse gráfico representa os aminoácidos em geral da estrutura TcGK. O eixo x representa os valores dos ângulos de torção Phi. O eixo y representa os valores dos ângulos torção de Psi. As linhas coloridas em azul claro delimitam as regiões favoráveis, indicando as conformações que não envolvem sobreposição estérica. As áreas delimitadas pelas linhas coloridas em azul escuro são as regiões em que as conformações são permitidas. As regiões não delimitadas correspondem a conformações não permitidas.

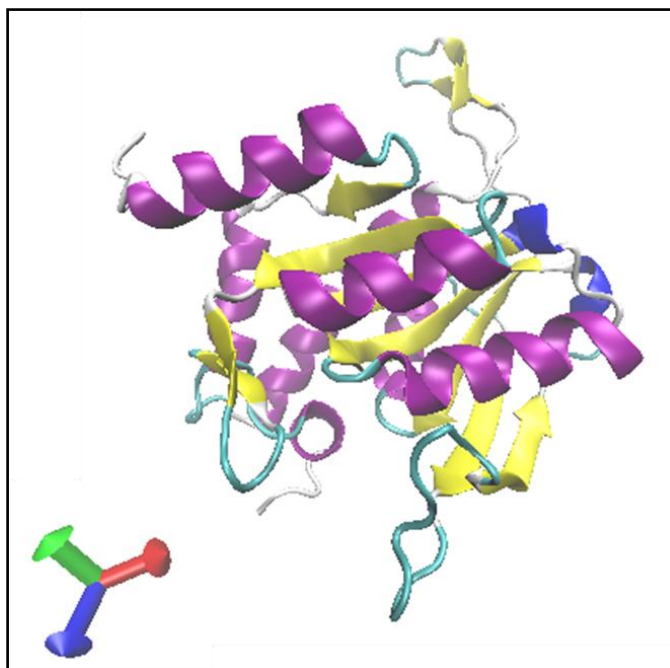
#### 4.21. Modelagem estrutural por homologia da enzima TcUPRT e da UPRT de humano

Para a modelagem de TcUPRT, utilizou-se como molde a UPRT de *T. gondii* com estrutura resolvida por difração de raios-x e depositada no PDB sob o número de acesso 1UPF. O molde escolhido apresentou uma identidade entre 50% e 75% e cobertura maior que 90% com a TcUPRT o que indica que o modelo a ser gerado provavelmente será de alta qualidade.

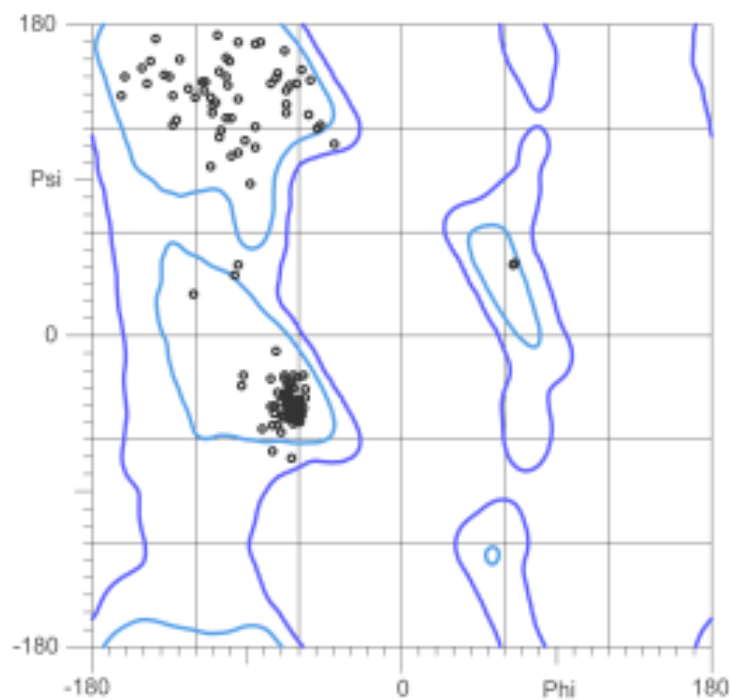
A estrutura terciária resultante da modelagem por homologia da TcUPRT pelo método de restrições espaciais foi satisfatória (figura 4.30), o que era esperado devido a alta identidade e cobertura do molde com a TcUPRT. O gráfico de Ramachandran (figura 4.31). gerado para esse modelo permitiu verificar a presença de um resíduo de prolina, na posição 132, fora da região permitida. No entanto, a prolina, por possuir uma cadeia lateral que lhe confere maior rigidez, é o único aminoácido capaz de ter enlaces CIS e TRANS, podendo, portanto, ocupar qualquer

área no gráfico de Ramachandran, inclusive as áreas não permitidas. A avaliação da qualidade estereoquímica, a partir da análise pelo PROCHECK da geometria total e resíduo por resíduo da estrutura da TcUPRT gerada também foi satisfatória.

Para a modelagem de UPRT de humano, utilizou-se como molde a UPRT de *T. gondii* (PDB 1UPF), por ser a proteína com estrutura resolvida que apresentou a maior homologia entre as estruturas depositadas no PDB. No entanto, a pouca identidade e cobertura da UPRT de humano com o molde não permitiu gerar um modelo de qualidade para essa proteína.



**Figura 4.30:** Estrutura tridimensional da enzima TcUPRT determinada por modelagem por homologia.



**Figura 4.31:** Gráfico Ramachandran da modelagem de TcUPRT. Esse gráfico representa os aminoácidos em geral da estrutura TcUPRT. O eixo x representa os valores dos ângulos de torção Phi. O eixo y representa os valores dos ângulos de torção Psi. As linhas coloridas em azul claro delimitam as regiões favoráveis, indicando as conformações que não envolvem sobreposição estérica. As áreas delimitadas pelas linhas coloridas em azul escuro são as regiões em que as conformações são permitidas. As regiões não delimitadas correspondem a conformações não permitidas. .



## 5. DISCUSSÃO

A doença de Chagas, antes restrita aos países da América Latina, tornou-se um problema socioeconômico e de saúde pública de alcance mundial. Os movimentos migratórios de pacientes infectados de países endêmicos para países não endêmicos foram os principais responsáveis pela expansão epidemiológica dessa doença, representando 56.000 novos casos por ano (OPAS, 2012). Este quadro é reforçado pela alta taxa de mortalidade da doença, a capacidade do parasita desenvolver resistência aos medicamentos utilizados e a alta toxicidade destes (Wilkinson *et al.*, 2008). A doença de Chagas permanece na categoria de doenças tropicais negligenciadas e ainda não existem terapias adequadas que levem a sua cura na fase crônica. Dessa forma, existe urgência no desenvolvimento de novos fármacos que sejam capazes de combater a doença de Chagas em todas as suas fases de manifestação.

Estudos focados no metabolismo do *T. cruzi*, utilizando novas técnicas e metodologias, podem ser de grande importância na obtenção de informações que contribuam para a procura de potenciais alvos terapêuticos e no planejamento racional de fármacos (El-Sayed *et al.*, 2005; Minning *et al.*, 2009; Capriles *et al.*, 2010; Guimarães *et al.*, 2012). Uma abordagem propõe que, a partir de um alvo molecular essencial ao parasito, nesse caso uma enzima metabólica, seguida da elucidação de sua estrutura tridimensional, possam ser planejados ou identificados inibidores de alta afinidade, capazes de agir nesse alvo, mas não em um eventual homólogo humano.

A busca de alvos moleculares para o desenvolvimento de fármacos tem sido impulsionada pelo amplo desenvolvimento da bioinformática (Glaucius, 2008). A partir da análise do genoma de organismos patogênicos, como o *T. cruzi*, tem sido possível obter informações sobre os genes ali codificados e as funções das suas proteínas correspondentes no metabolismo celular (Glaucius, 2008). Esse conjunto de informações permite, através da reconstrução *in silico* de vias metabólicas, uma maior compreensão sobre o metabolismo e outros processos celulares deste parasito, possibilitando uma inferência de potenciais alvos terapêuticos.

A abordagem utilizada neste trabalho baseou-se na identificação e caracterização de atividades enzimáticas relevantes no metabolismo energético de *T. cruzi*. Um dos programas utilizado neste trabalho foi o AnEnPi, em que uma das funções é a identificação de atividades enzimáticas e a classificação delas como

homólogas, análogas e específicas, em relação a outros organismos, através do agrupamento das estruturas primárias de proteínas.

Nas proteínas preditas do genoma de *T. cruzi* foram identificadas, com o AnEnPi, várias atividades enzimáticas participando em sua maioria do metabolismo de lipídeos e carboidratos. Algumas destas já vêm sendo estudadas como potenciais alvos terapêuticos, assim como outras atividades ainda não caracterizadas (Alves-Ferreira *et al.*, 2009). Dentre os alvos apontados pelo AnEnPi, foram selecionadas para a realização deste estudo as enzimas gliconato cinase (GK, EC 2.7.1.12) e a uracil fosforribosiltransferase (UPRT, EC 2.4.2.9) de *T. cruzi*.

A via das pentoses-fosfato apresenta como principal função a produção de ribose-5-fosfato e NADPH. O NADPH é essencial a inúmeras reações biossintéticas e na manutenção de um ambiente redutor na célula. Em tripanosomatídeos, o NADPH é o cofator necessário para a tripanotiona redutase catalisar a redução da tripanotiona, um reagente antioxidante característico dessa família (Augustyns *et al.*, 2001). Já a ribose-5-fosfato é um componente fundamental dos nucleotídeos e ácidos nucleicos (Nelson & Cox, 2011). Na via das pentoses-fosfato, a GK catalisa a fosforilação de gliconato a 6-fosfogliconato, utilizando ATP, possibilitando uma rota alternativa na produção de 6-fosfogliconato em adição à rota tradicional mediada pela glicose-6-fosfato desidrogenase. A GK, apesar de pouco estudada, foi identificada em fígado de porco (Leder, 1957), *Pseudomonas* spp. (Coffee & Hu, 1972), *Corynebacterium glutamicum* (Porco *et al.*, 2008), *Gluconobacter oxydans* (Rauch *et al.*, 2010) e em *E. coli* (Izu *et al.*, 1996), a qual teve a sua estrutura terciária resolvida através de cristalografia e difração de raios-X (Kraft *et al.*, 2002; Rajaram *et al.*, 2002).

Rolfsson e cols. (2013) demonstraram a presença no genoma humano de um gene codificante para uma proteína com possível atividade enzimática de GK, até então não anotado no genoma. A abordagem utilizada neste trabalho é baseada na identificação de metabólitos órfãos no metabolismo humano, assim como na busca por atividades enzimáticas que possibilitassem a incorporação dos mesmos em vias conhecidas, utilizando a ferramenta RECON 1 (Rolfsson *et al.*, 2011). A partir de metabólitos encontrados em humanos cujos metabolismos permanecem incompletos, foi possível identificar computacionalmente soluções metabólicas capazes de explicar a existência desses compostos utilizando bancos de dados de reações bioquímicas. O gliconato é um caso desses metabólitos órfãos, e uma das soluções apontadas para sua presença no organismo humano é a reação catalisada

pela enzima GK. A GK foi posteriormente apontada como sendo codificada pelo gene candidato de supressor de tumor, o C9orf103, no qual se observou uma similaridade com o gene codificante para a GK de *E. coli*. A superexpressão do gene C9orf103 em células *HeLa* permitiu a confirmação da atividade enzimática GK da proteína recombinante humana. Esses resultados colocam em avaliação a classificação feita pelo AnEnPi que aponta a GK como uma enzima específica de *T. cruzi* em relação ao ser humano. A ausência desta atividade enzimática no AnEnPi se deve ao fato da não identificação da enzima GK de humano na etapa de anotação funcional nos bancos de dados do genoma humano disponíveis. Como o AnEnPi utiliza a informação de atividades enzimáticas contidas no KEGG que por sua vez não possui tal informação, conseqüentemente a atividade enzimática correspondente a GK humana não foi identificada, sendo considerada específica de *T. cruzi*. Uma vez que a enzima GK de humano foi identificada pelo RECON 1, outras análises *in silico* comparando as sequências de aminoácidos de GK de *T. cruzi* e de humano devem ser realizadas a fim de determinar uma possível homologia ou analogia entre essas enzimas, visando um mapeamento de regiões diferenciadas em suas estruturas tridimensionais. A partir dessas análises será possível avaliar se esta enzima de *T. cruzi* pode ser considerada um potencial alvo terapêutico.

Já a enzima UPRT é uma das principais enzimas da via de salvação de UMP a partir de uracil, em presença de PRPP. A via de salvação de pirimidinas permite que nucleosídeos livres e bases nitrogenadas, provenientes da degradação de nucleotídeos de pirimidina ou do meio extracelular, sejam reutilizadas na síntese de novos nucleotídeos de pirimidina. Essa via de salvação, juntamente com a via *de novo* de síntese de pirimidina, compõe o metabolismo das pirimidinas. Esse metabolismo, de forma geral, é extremamente dependente da via das pentoses-fosfato, uma vez que um dos principais intermediários, o PRPP, é produzido a partir de ribose-5-fosfato e de ATP, em reação catalisada pela enzima PRPP sintase (Nelson & Cox, 2011).

A atividade da UPRT já foi demonstrada em extratos celulares das formas evolutivas epimastigota e tripomastigota de *T. cruzi* (Hammond & Gutteridge, 1982). Em *T. gondii* essa enzima foi amplamente caracterizada e teve sua estrutura tridimensional resolvida (Schumacher *et al.*, 1998). O grande interesse na UPRT de *T. gondii* está no fato dessa enzima ser capaz de se ligar a vários análogos de pirimidina, como o 5-fluoruracil, que tem atividade tóxica ao *T. gondii* (Carter *et al.*,

1997). Estudos mostraram que o nocaute da UPRT e da uridina fosforilase em *T. gondii*, ambas as enzimas da via de salvação de pirimidinas, não afetavam o desenvolvimento da forma cística desse parasito, não sendo, portanto, requeridas no estabelecimento do estágio latente da infecção, o que sugeria que a via *de novo* de síntese de pirimidina era funcional e essencial nesse estágio da doença (Fox *et al.*, 2011). Na forma amastigota de *L. donovani*, no entanto, demonstrou-se que o duplo nocaute de UPRT e da carbamoil fosfato sintetase, pertencente à via *de novo* das pirimidinas, consegue atenuar a infecção e eliminar o parasito em semanas. Esses resultados indicam que a presença da via de salvação de pirimidina e da via *de novo* de síntese de pirimidina intactas são indispensáveis a infecção causada no hospedeiro e à sobrevivência desse parasito (Wilson *et al.*, 2012). Em *M. tuberculosis*, a função crucial da UPRT na via de salvação de pirimidinas também pode sugerir alguma relação ao estado de latência e/ou virulência desse micro-organismo (Villela *et al.*, 2013).

A UPRT de *T. cruzi* foi apontada pelo AnEnPi como sendo homóloga à UPRT do ser humano. Apesar disso, o isolamento da UPRT a partir da biblioteca de cDNA de cérebro fetal humano, seguido da sua clonagem, expressão e purificação rendeu uma proteína recombinante que não teve atividade catalítica de UPRT detectável (Li *et al.*, 2007). Por outro lado, a UMP em humano, ao contrário do que acontece em *T. cruzi*, pode ser obtida por meio da reação sequencial da uridina fosforilase (Renck *et al.*, 2010) e da uridina cinase (Suzuki *et al.*, 2004). Esses resultados permitem apontar a UPRT de *T. cruzi* como um alvo terapêutico potencial, uma vez que essa enzima apresenta um papel importante na via de salvação de pirimidina e não existe ainda evidência experimental da atividade do seu homólogo funcional em humano.

Os contextos metabólicos em que estão inseridas as enzimas GK e UPRT tornaram interessante a caracterização bioquímica e molecular, assim como a identificação e localização dessas proteínas em *T. cruzi*. Para este estudo, como as duas cópias gênicas identificadas como codificantes das atividades enzimáticas GK em *T. cruzi* apresentaram alta similaridade e identidade entre si, apenas a cópia Tc00.1047053508177.20 foi utilizada para caracterização da TcGK. De maneira semelhante a putativa UPRT de *T. cruzi* também foi apontada como sendo codificada por duas cópias gênicas, com alta identidade e similaridade, sendo somente a cópia Tc00.1047053507739.40 caracterizada.

Análises de bioinformática com os programas Pfam e ProDom foram realizadas a fim de identificar possíveis domínios funcionais e assinaturas de

famílias de proteína. O resultado obtido com o ProDom para TcGK ratificou a atividade enzimática predita pelo AnEnPi, através da identificação de um domínio característico de GK. Já o programa Pfam (Tabela 4.1) identificou uma assinatura da família das chiquimato cinases. A alta homologia presente entre a família das chiquimato cinases com a família das gliconato cinases pode ter sido a responsável por esta predição.

Na sequência de aminoácidos de TcUPRT foram identificados tanto pelo Pfam quanto pelo ProDom assinaturas de famílias e domínios conservados característicos da função uracil fosforribosiltransferase, novamente confirmando a identificação predita pelo AnEnPi (Tabela 4.1). Nessa sequência de TcUPRT também foram identificados domínios característicos das uridina cinases pelo ProDom, o que provavelmente é explicado pela íntima relação entre as atividades uracil fosforribosiltransferase e uridina cinase, uma vez que ambas as enzimas possuem o nucleotídeo UMP como ligante. A uracil fosforribosiltransferase realiza em um único passo a formação de UMP a partir de uracil que, de outra forma, é catalisada sequencialmente pela uridina fosforilase e uridina cinase (como observado na figura 1.11). Esses fatos sugerem a presença de domínios de ligação semelhantes entre essas enzimas.

A ausência de peptídeo sinal e/ou sítio de clivagem de peptídeo sinal nas sequências de aminoácidos da TcGK e da TcUPRT (analisadas com o SignalP) pode indicar que as enzimas em questão não são endereçadas para nenhuma organela, uma vez que o peptídeo sinal está diretamente relacionado com o endereçamento da proteína para diferentes compartimentos celulares. As predições de localização subcelular de TcGK e TcUPRT, com o programa Psort, indicaram que ambas as enzimas se localizam no citoplasma (tabela 4.2). Esses resultados sugerem a não identificação, por esse programa, de peptídeos sinais responsáveis pelo endereçamento dessas proteínas a organelas, semelhante aos resultados obtidos com o SignalP.

Inúmeros sítios de modificação pós-traducional foram preditos pelo PROSITE (tabela 4.3) para ambas as proteínas estudadas. Para TcGK foram identificados prováveis sítios de N-miristoilação e fosforilação, enquanto que para TcUPRT, além de sítios de N-miristoilação e fosforilação, foram identificados também sítios de amidação e N-glicosilação. Apesar desse possível sítio de N-glicosilação na sequência de aminoácidos de TcUPRT, proteínas sem peptídeo sinal são improváveis de serem expostas a esse tipo de modificação, em teoria, mesmo

possuindo motivos potenciais de glicosilação. Já as fosforilações, ao contrário das glicosilações, não dependem da presença de peptídeo sinal, uma vez que podem ocorrer no citoplasma. A importância da identificação de sítios de modificação pós-traducional está relacionada a capacidade de essas modificações influenciarem a localização, estabilidade e conformação da proteína e até mesmo na sua função e regulação.

Após as análises *in silico* das estruturas primárias de TcGK e TcUPRT, os genes referentes a essas proteínas foram amplificados por PCR. Os resultados das amplificações apresentaram fragmentos correspondentes aos pesos moleculares preditos de 711pb para *TcUPRT* (cópia Tc00.1047053507739.40) e de 585pb para *TcGK* (cópia Tc00.1047053508177.20). Os fragmentos obtidos foram posteriormente clonados no vetor pBAD TOPO TA e subclonados no sistema de expressão pET28a, visando a expressão heteróloga das proteínas recombinantes TcGK e TcUPRT.

A escolha do sistema de expressão pET28a foi principalmente devida à expressão da proteína recombinante estar sob o controle do promotor T7 que, na presença do indutor IPTG e da T7 RNA polimerase proveniente da célula hospedeira, no caso BL21 (DE3), é capaz de promover a expressão da proteína recombinante em grandes quantidades. Além disso, esse sistema possibilita a expressão da proteína recombinante fusionada a uma cauda de poli-histidina na sua região C- ou N-terminal. A cauda de poli-histidina permite a purificação da proteína recombinante por cromatografia de afinidade, visto que a histidina possui alta afinidade a íons divalentes.

Como pode ser observado nas figuras 4.12 e 4.13, as proteínas recombinantes foram expressas em altas concentrações, resultando na formação de corpos de inclusão e a consequente precipitação dessas proteínas. Variando as condições de indução, a expressão de TcUPRT recombinante apresentou um aumento significativo da solubilidade, o que não foi observado para a TcGK recombinante. Os melhores resultados de expressão de TcUPRT solúvel se devem provavelmente ao fato de que, em temperaturas menores, como a 20°C e a 27°C, as bactérias crescem mais lentamente e expressam a proteína recombinante em menor quantidade, diminuindo a formação de corpos de inclusão. Isso oferece ainda mais tempo para essas proteínas completarem seu enovelamento e, conseqüentemente, se tornarem solúveis. Outros métodos de aumento da expressão das proteínas recombinantes na forma solúvel envolvem mudanças na concentração do indutor ou

a utilização para a expressão em outras cepas bacterianas como Rosetta (DE3) e Origami (DE3).

Como as proteínas recombinantes TcGK e TcUPRT foram expressas em grandes quantidades na forma insolúvel, elas foram solubilizadas utilizando o agente caotrópico ureia, porém esse método promove a desnaturação das proteínas. Esse procedimento de solubilização, apesar de não viabilizar a realização de ensaios enzimáticos, permite a purificação da proteína recombinante visando a produção de anticorpos policlonais. Deste modo, as proteínas recombinantes TcGK e TcUPRT foram purificadas por cromatografia de afinidade em coluna contendo  $\text{Ni}^{+2}$ , e apresentaram um grau de pureza adequado para as etapas de imunização.

Através da imunização de camundongos Balb-C com as proteínas recombinantes purificadas, obtiveram-se soros imunes contendo anticorpos policlonais anti-TcGK e anti-TcUPRT. Os soros imunes anti-TcGK foram capazes de identificar a proteína recombinante TcGK até uma diluição de 1:50.000 (figura 4.20A). Por outro lado, os soros anti-TcUPRT identificaram a proteína recombinante TcUPRT até uma diluição de 1:100.000 (figura 4.20B). Ambos os soros, mesmo em baixas diluições, não reconheceram proteínas nativas de *E. coli* no *western blot*. Esta produção de anticorpos policlonais foi de grande importância para a realização de metodologias visando à identificação das enzimas UPRT e GK nativas em *T. cruzi*, tendo sido utilizadas as técnicas de *western blot* e de imunofluorescência.

Os ensaios de *western blot* utilizando os soros imunes e as IgGs purificadas anti-TcGK e anti-TcUPRT, contra a fração de proteínas solúveis extraídas da forma epimastigota de *T. cruzi*, cepa CL Brener, permitiram a identificação de proteínas nativas apresentando pesos moleculares em torno de 21 kDa e de 26 kDa (figura 4.21) conforme esperado, respectivamente, para as proteínas GK e UPRT. A diferença observada no peso molecular das proteínas recombinantes quando comparado aos pesos das proteínas nativas GK e UPRT de *T. cruzi* tem como principal motivo a presença da cauda de histidina fusionada as proteínas recombinantes.

Esse resultado positivo para a expressão da proteína nativa UPRT na forma epimastigota de *T. cruzi* corrobora dados obtidos em estudos realizados na década de 80, que demonstraram a atividade dessa enzima em extratos celulares das formas evolutivas epimastigota e tripomastigota de *T. cruzi* (Hammond & Gutteridge, 1982). Além disso, estudos obtidos com a técnica de microarranjo de diversos genes anotados no genoma de *T. cruzi* indicaram que os genes potencialmente

codificantes para a enzima UPRT e para a enzima GK estavam sendo expressos em todas as formas evolutivas desse parasito (Minning *et al.*, 2009).

A imunofluorescência utilizando os soros imunes e as IgGs purificadas, anti-TcGK e anti-TcUPRT, contra a forma epimastigota de *T. cruzi* permitiu inferir a possível localização das enzimas nativas GK e UPRT. Para ambas as enzimas, observou-se uma marcação difusa ao longo do citoplasma e marcações dentro de vesículas (figuras 4.22 e 4.23). Além desta marcação difusa, foi possível observar também uma marcação proeminente com a utilização do soro e/ou IgG purificadas anti-TcUPRT próximo à região da bolsa flagelar (figura 4.23).

A localização no citoplasma para ambas as enzimas nativas já era esperada, uma vez que as análises das sequências proteicas utilizando o programa PSORT indicaram uma alta probabilidade de essas enzimas serem citoplasmáticas, o que confere com a marcação difusa observada no citoplasma em ambas os casos (figuras 4.22 e 4.23). Além disso, o programa SignalP não identificou sequências sinais nessas enzimas. O fato das enzimas alvos também estarem localizadas em vesículas, mesmo não tendo sido identificadas sequências sinais pelo programa em questão, indica a presença de uma possível sequência sinal de endereçamento em *T. cruzi* ainda não caracterizada (Baldi *et al.*, 2000; Hiss & Schneider, 2009).

A localização citoplasmática de TcUPRT evidenciada pelos ensaios de imunofluorescência e pelas predições *in silico* ratificaram ensaios realizados por Hammond e cols. (1982), que demonstraram em vários kinetoplastidas, entre eles *T. cruzi*, a atividade enzimática da UPRT na fração citoplasmática. Já a UPRT de humano apresentou não só uma localização citoplasmática como também nuclear (Li *et al.*, 2007).

Embora pouco se saiba sobre a localização subcelular de GK nos organismos em que essa enzima foi estudada, ensaios de permeabilização com digitonina mostraram que as enzimas relacionadas à via convencional das pentoses-fosfato na forma epimastigota de *T. cruzi* se localizam essencialmente no citoplasma. Apesar disso, a ribulose-5-fosfato epimerase, também pertencente a via convencional das pentoses-fosfato, apresentou uma localização distinta, provavelmente em vesículas ligadas a membrana plasmática (Maugeri & Cazzulo, 2004).

A GK de *E. coli* teve sua estrutura tridimensional resolvida por cristalografia e o seu mecanismo de catálise descrito por Kraft e cols. (2002). Com o alinhamento obtido entre a estrutura primária de GK de *E. coli* e a estrutura primária de TcGK, foi observado que o domínio de ligação a fosfato em TcGK, referente aos aminoácidos



22 à 29, se manteve conservado (figura 4.24). Entretanto, a TcGK apresenta dois resíduos distintos em relação a mesma região na GK de *E. coli*. Outro indício de que essa região de TcGK ainda corresponde ao sítio de ligação a fosfato foi a identificação pelo PROSITE dessa mesma região como sendo um provável sítio de ligação de ATP/GTP (Tabela 4.3). Por outro lado, dos sete aminoácidos envolvidos na ligação do gliconato na GK de *E. coli*, a GK de *T. cruzi* apresentou apenas quatro aminoácidos conservados. Apesar de haverem três substituições dos aminoácidos responsáveis por essa ligação, a substituição observada da histidina pela alanina, na posição 143, em TcGK, provavelmente é atenuada pela existência da histidina 142. Já a comparação entre a estrutura secundária predita da GK de *T. cruzi* com a GK de *E. coli* (figura 4.25), cuja estrutura secundária já foi descrita (Kraft *et al.*, 2002), aponta que essas estruturas possuem um perfil semelhante de distribuição das hélices  $\alpha$  e folhas  $\beta$ .

A modelagem por homologia pelo método de restrições espaciais possibilitou a definição de um modelo tridimensional da TcGK (figura 4.28), utilizando como molde a GK de *E. coli*. Uma vez que o modelo gerado apresentou qualidade satisfatória, as próximas etapas envolvem o mapeamento dos sítios de ligação aos substratos gliconato e ATP na estrutura terciária de TcGK, como realizado na sua estrutura primária, assim como ensaios de atracamento e dinâmica molecular desses sítios. Além disso, a modelagem por homologia e o mapeamento da estrutura terciária da GK de humano também seriam importantes para a comparação com a estrutura terciária da TcGK.

O alinhamento das estruturas primárias de UPRT de *T. cruzi* com a UPRT de *T. gondii* e do ser humano (figura 4.26) permitiu analisar o grau de identidade entre essas três estruturas e mapear as principais regiões de ligação a substrato e ao produto UMP na TcUPRT de *T. cruzi* com base no que foi descrito para a UPRT de *T. gondii* (Schumacher *et al.*, 1998). Pode-se observar que, em relação à UPRT de *T. gondii*, a TcUPRT manteve um alto grau de conservação indicado pela identidade entre essas estruturas, principalmente pelo fato de as principais regiões e aminoácidos diretamente envolvidos na ligação a PRPP, ao uracil e ao UMP terem se mantido praticamente idênticos com apenas algumas diferenças pontuais. Essa grande semelhança indica que provavelmente o mecanismo de catálise da UPRT de *T. cruzi* é similar ao da UPRT de *T. gondii*.

Apesar da identidade entre a estrutura primária de TcUPRT de *T. cruzi* com a UPRT do ser humano ser expressiva, a proteína UPRT de humano possui em sua

região N-terminal uma longa sequência de aminoácidos que não está presente na proteína de *T. cruzi* (figura 4.26). Além disso, a região possivelmente envolvida na ligação ao uracil em *T. cruzi* (posição 219 à 230) juntamente com os aminoácidos diretamente envolvidos nessa ligação, se apresentam pouco conservados no ser humano, com a ausência da maioria dos aminoácidos do domínio de ligação ao uracil. A metionina na posição 166 em *T. gondii*, apontada como fundamental na interação com uracil e na ligação a PRPP (Schumacher *et al.*, 1998), se mantém conservada em *T. cruzi* (metionina 155), enquanto que no ser humano o sítio correspondente apresenta uma isoleucina (posição 240). Quanto a provável região de ligação a PRPP de 12 aminoácidos, referente às posições 150 à 161 em *T. cruzi*, nota-se que a mesma região em humanos apresenta sete aminoácidos distintos aos de *T. cruzi*, embora alguns deles tenham propriedades químicas semelhantes. Já os aminoácidos em *T. cruzi* que estão provavelmente envolvidos na interação com o produto UMP, localizados principalmente nas regiões de ligação de uracil e PRPP, são bastante distintos no ser humano, sendo o resíduo fenilalanina (posição 301 na proteína humana) o único conservado no alinhamento com as duas outras espécies.

A comparação entre o que é descrito para a estrutura secundária de UPRT de *T. gondii* (Schumacher *et al.*, 1998) e o que foi predito para as UPRTs de *T. cruzi* e do ser humano indica que essas estruturas secundárias são muito semelhantes entre si, mantendo um padrão de hélices  $\alpha$  e folhas  $\beta$  similar em várias regiões (figura 4.27). No entanto, em humanos algumas pequenas diferenças são notadas, principalmente na região N-terminal dessa proteína, enquanto a região C-terminal da estrutura secundária da UPRT de *T. gondii* possui um padrão distinto das demais.

A modelagem por homologia utilizando o método de restrições espaciais da estrutura tridimensional da TcUPRT (figura 4.30) foi realizada utilizando como molde a UPRT de *T. gondii* resolvida por cristalografia. As avaliações da estrutura tridimensional de TcUPRT pelo gráfico de Ramachandran (figura 4.31) e por meio de avaliações estereoquímicas indicaram a alta qualidade desse modelo gerado. Dessa forma, o próximo passo para a caracterização da estrutura tridimensional de TcUPRT é mapear nessa estrutura os sítios de ligação aos substratos uracil e PRPP e ao produto UMP, assim como foi realizado para a estrutura primária de TcUPRT. Como não se conseguiu gerar um modelo satisfatório para a estrutura terciária da UPRT de humano por modelagem por homologia, é de grande importância a resolução dela por cristalografia de forma a permitir comparações entre a estrutura tridimensional gerada com a predita para TcUPRT.

## 6. CONCLUSÕES

Após a realização desse estudo, podemos concluir que:

- ✓ A clonagem e a expressão heteróloga das proteínas TcGK e TcUPRT foram realizadas satisfatoriamente;
- ✓ As proteínas TcGK e TcUPRT recombinantes foram expressas nas frações solúveis e insolúveis de *E. coli*, embora a quantidade dessas proteínas nas frações insolúveis tenha sido maior;
- ✓ As proteínas recombinantes foram purificadas por cromatografia de afinidade e apresentaram bons graus de pureza possibilitando a produção satisfatória de anticorpos policlonais anti-TcGK e anti-TcUPRT por imunização;
- ✓ Os anticorpos policlonais produzidos em camundongo foram capazes reconhecer as proteínas TcGK e TcUPRT recombinantes e as proteínas TcGK e TcUPRT nativas presentes nas frações solúveis da forma epimastigota de *T. cruzi*, cepa CL Brener;
- ✓ Os resultados de imunofluorescência apontaram para uma localização citoplasmática e no interior de vesículas das enzimas TcGK e TcUPRT de *T. cruzi*, o que corroborou em parte com as predições de localização subcelular feitas *in silico*;
- ✓ As estruturas tridimensionais das enzimas TcGK e TcUPRT de *T. cruzi* resolvidas por meio de modelagem por homologia apresentaram bons índices de confiabilidade de acordo com os métodos empregados de validação de modelos.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

Acquatella H. Echocardiography in Chagas heart disease. *Circulation*. 2007 Mar 6;115(9):1124–31.

Alloush HM, Kerridge D. Characterisation of a partially purified uracil phosphoribosyltransferase from the opportunistic pathogen *Candida albicans*. *Mycopathologia*. 1994. 125:129–141

Alves-Ferreira M, Guimarães ACR, Capriles PVDSZ, Dardenne LE, Degrave WM. A new approach for potential drug target discovery through in silico metabolic pathway analysis using *Trypanosoma cruzi* genome information. *Memórias do Instituto Oswaldo Cruz*. 2009 Dec;104(8):1100–10.

Andersen PS, Smith JM, Mygind B. Characterization of the upp gene encoding uracil phosphoribosyltransferase of *Escherichia coli* K12. *The Federation of European Biochemical Societies Journal*. Wiley Online Library; 1992;204(1):51–6.

Augustyns K, Amssoms K, Yamani A, Rajan PK, Haemers A. Trypanothione as a target in the design of antitrypanosomal and antileishmanial agents. *Current Pharmaceutical Design*. 2001;7(Netherlands PT - Journal Article PT - Review LG - English):1117–41.

Baldi P, Brunak S, Chauvin Y, Andersen CA, Nielsen H. Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics*. Oxford Univ Press; 2000;16(5):412–24.

Barrett MP. The pentose phosphate pathway and parasitic protozoa. *Parasitology today Personal ed*. Elsevier Sci Ltd; 1997;13(1):11–6.

Beck DA, O'Donovan GA. Pathways of pyrimidine salvage in *Pseudomonas* and former *Pseudomonas*: detection of recycling enzymes using high-performance liquid chromatography. *Current Microbiology*. 2008;56(2):162–7.

Bern C. Antitrypanosomal therapy for chronic Chagas' disease. *The New England journal of medicine*. 2011 Jun 30;364(26):2527–34.

Borges CB. Fosforribosilpirofosfato sintase de mycobacterium tuberculosis tipo selvagem: uma PRS classe II bacteriana? Porto alegre. Dissertação [Mestrado em Biologia Celular e Molecular]- Pontifícia Universidade Católica do Rio Grande do Sul; 2011

Brener Z. Trypanosoma cruzi: morfologia e ciclo evolutivo. In *Clínica e terapêutica da doença de Chagas. Uma abordagem prática para o clínico geral*. Rio de Janeiro: Fiocruz, 1997.

Brusca RC, Brusca GJ. *Invertebrados*. 2ªed, Rio de Janeiro: Guanabara Koogan, 2007.

Bustamante JM, Tarleton RL. Methodological advances in drug discovery for Chagas disease. *Expert Opinion On Drug Discovery*. 2011;6(6):653–61.

Capriles PVSZ, Guimarães ACR, Otto TD, Miranda AB, Dardenne LE, Degraeve WM. Structural modelling and comparative analysis of homologous, analogous and specific proteins from Trypanosoma cruzi versus Homo sapiens: putative drug targets for chagas' disease treatment. *BMC genomics*. 2010 Jan;11:610.

Carter D, Donald RG, Roos D, Ullman B. Expression, purification, and characterization of uracil phosphoribosyltransferase from Toxoplasma gondii. *Molecular and biochemical parasitology*. 1997 Aug;87(2):137–44.

Christoffersen S, Kadziola A, Johansson E, Rasmussen M, Willemoës M, Jensen KF. Structural and kinetic studies of the allosteric transition in Sulfolobus solfataricus uracil phosphoribosyltransferase: Permanent activation by engineering of the C-terminus. *Journal of molecular biology*. Elsevier Ltd; 2009 Oct 23;393(2):464–77.

Coffee CJ, Hu AS. The kinetic characterization of gluconokinase from a pseudomonad. *Arch Biochem Biophys*. 1972 Apr;149(2):549-59.

Coura JR, de Castro SL. A critical review on Chagas disease chemotherapy. *Memórias do Instituto Oswaldo Cruz*. 2002 Jan;97(1):3–24.

Coura JR, Viñas PA. Chagas disease: a new worldwide challenge. *Nature*. 2010;465(7301):S18–S20.

Coura JR, Borges-Pereira J. Chagas disease: 100 years after its discovery. A systemic review. *Acta tropica*. 2010;115(1-2):5–13.

Cronin CN, Nolan DP, Voorheis HP. The enzymes of the classical pentose phosphate pathway display differential activities in procyclic and bloodstream forms of *Trypanosoma brucei*. *FEBS Letters*. 1989;244(1):26–30.

Dai YP, Lee CS, O'Sullivan WJ. Properties of uracil phosphoribosyltransferase from *Giardia intestinalis*. *International Journal for Parasitology*. 1995;25(2):207–14

Davis IW, Murray LW, Richardson JS, Richardson DC. MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res*. 2004 Jul 1;32:615-9.

de Souza W. Basic cell biology of *Trypanosoma cruzi*. *Current Pharmaceutical Design BENTHAM SCIENCE PUBL LTD*; 2002;8(4):269–85.

de Souza W, de Carvalho TMU, Barrias ES. Review on *Trypanosoma cruzi*: Host Cell Interaction. *International journal of cell biology*. 2010 Jan;2010.

Dias JCP, Silveira a C, Schofield CJ. The impact of Chagas disease control in Latin America: a review. *Memórias do Instituto Oswaldo Cruz*. 2002 Jul;97(5):603–12.

Eicks M, Maurino V, Knappe S, Flügge U-I, Fischer K. The Plastidic Pentose Phosphate Translocator Represents a Link between the Cytosolic and the Plastidic Pentose Phosphate Pathways in Plants<sup>1</sup>. *Plant Physiology*. *American Society of Plant Physiologists*; 2002;128(2):512–22.

El-Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G, Tran AN, et al. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 2005 Jul 15;309(5733):409-15.

Fiser A, Sali A. "Modeller: generation and refinement of homology-based protein structure models". *Meth. Enzymol.* 2003; 374: 461–91

Fitch WM. Distinguishing homologous from analogous proteins. *Syst Zool.* 1970 Jun;19(2):99-113

Fox B a, Bzik DJ. De novo pyrimidine biosynthesis is required for virulence of *Toxoplasma gondii*. *Nature.* 2002 Feb 21;415(6874):926–9.

Fox BA, Falla A, Rommereim LM, Tomita T, Gigley JP, Mercier C, et al. Type II *Toxoplasma gondii* KU80 Knockout Strains Enable Functional Analysis of Genes Required for Cyst Development and Latent Infection. *Eukaryotic Cell.* 2011;10(9):1193–206.

Gao G, Nara T, Nakajima-Shimada J, Aoki T. Novel organization and sequences of five genes encoding all six enzymes for de novo pyrimidine biosynthesis in *Trypanosoma cruzi*. *Journal of Molecular Biology.* 1999;285(1):149–61.

Glaucius O. Bioinformática: Perspectivas na Medicina. *Gazeta Médica da Bahia* 2008; 78 (Suplemento 1):52-58

Ghosh K. *Leishmania donovani*: Assay Phosphate for a Functional Pathway Pentose. *Exp Parasitol.* 1971; 29(1):103-9.

Guedes PMM, Silva GK, Gutierrez FRS, Silva JS. Current status of Chagas disease chemotherapy. *Expert review of anti-infective therapy.* 2011 May;9(5):609–20.

Guimarães a CR, Otto TD, Alves-Ferreira M, Miranda a B, Degraive WM. In silico reconstruction of the amino acid metabolic pathways of *Trypanosoma cruzi*. *Genetics and molecular research : GMR*. 2008 Jan;7(3):872–82.

Griffiths AJF, Wessler SR, Lewontin RC, Carroll SB. *Introdução a Genética*. 9ªed, Rio de Janeiro: Guanabara Koogan, 2009.

Gurgel-Gonçalves R, Galvão C, Costa J, Peterson a T. Geographic distribution of chagas disease vectors in Brazil based on ecological niche modeling. *Journal of tropical medicine*. 2012 Jan;2012:705326.

Hammond DJ, Gutteridge WE. UMP synthesis in the kinetoplastida. *Biochimica et Biophysica Acta*. 1982;718(1):1–10.

Hiss JA, Schneider G. Architecture, function and prediction of long signal peptides. *Briefings in Bioinformatics*. Oxford Univ Press; 2009;10(5):569–78.

Hyde JE. Targeting purine and pyrimidine metabolism in human apicomplexan parasites. *Current Drug Targets*. BENTHAM SCIENCE PUBL LTD; 2007;8(1):31–47.

Iltzsch MH, Tankersley KO. Structure-activity relationship of ligands of uracil phosphoribosyltransferase from *Toxoplasma gondii*. *Biochemical Pharmacology*. 1994;48(4):781–92.

Izu H, Adachi O, Yamada M. Purification and characterization of the *Escherichia coli* thermoresistant gluconokinase encoded by the *gntK* gene. *FEBS Letters*. 1996 Sep;394(1):14–6.

Iwahana H, Fujimura M, Ii S, Kondo M, Moritani M, Takahashi Y, *et al*. Molecular cloning of a human cDNA encoding a trifunctional enzyme of carbamoyl-phosphate synthetase-aspartate transcarbamoylase-dihydroorotase in de Novo pyrimidine synthesis. *Biochemical and Biophysical Research Communications*. 1996;219(1):249–55.



Jensen KF, Mygind B. Different oligomeric states are involved in the allosteric behavior of uracil phosphoribosyltransferase from *Escherichia coli*. *The Federation of European Biochemical Societies Journal*. 1996;240(3):637–45.

Jensen HK, Mikkelsen N, Neuhard J. Recombinant uracil phosphoribosyltransferase from the thermophile *Bacillus caldolyticus*: expression, purification, and partial characterization. *Protein Expression and Purification*. 1997;10(3):356–64.

Jensen KF, Arent S, Larsen S, Schack L. Allosteric properties of the GTP activated and CTP inhibited uracil phosphoribosyltransferase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. *The FEBS journal*. 2005;272(6):1440–53.

Kadziola A, Neuhard J, Larsen S. research papers Structure of product-bound *Bacillus caldolyticus* uracil phosphoribosyltransferase confirms ordered sequential substrate binding research papers. 2002;936–45.

Kern L, de Montigny J, Jund R, Lacroute F. The *FUR1* gene of *Saccharomyces cerevisiae*: cloning, structure and expression of wild-type and mutant alleles. *Gene*. 1990;88:149–157.

Kraft L, Sprenger Ga, Lindqvist Y. Conformational changes during the catalytic cycle of gluconate kinase as revealed by X-ray crystallography. *Journal of molecular biology*. 2002 May 10;318(4):1057–69.

Kramer R, Cohen D. Functional genomics to new drug targets. *Nature reviews. Drug discovery*. 2004 Nov;3(11):965–72.

Kukimoto-Niino M, Shibata R, Murayama K, Hamana H, Nishimoto M, Bessho Y, *et al.* Crystal structure of a predicted phosphoribosyltransferase (TT1426) from *Thermus thermophilus* HB8 at 2.01 Å resolution. *Protein Science*. Cold Spring Harbor Laboratory Press; 2005;14(3):823–7.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK - a program to check the stereochemical quality of protein structures. *J. App. Cryst.* 1993; 26:283-91.

Leder IG. Hog kidney gluconokinase. *The Journal of biological chemistry.* 1957 Mar;225(1):125–36.

Li J, Huang S, Chen J, Yang Z, Fei X, Zheng M, *et al.* Identification and characterization of human uracil phosphoribosyltransferase (UPRTase). *Journal of Human Genetics.* 2007;52(5):415–22.

Lima FM, Oliveira P, Mortara R a, Silveira JF, Bahia D. The challenge of Chagas' disease: has the human pathogen, *Trypanosoma cruzi*, learned how to modulate signaling events to subvert host cells? *New biotechnology.* Elsevier B.V.; 2010 Dec 31;27(6):837–43.

Martins-Melo FR, Ramos ANJ, Alencar CH, Heukelbach J. Mortality due to Chagas disease in Brazil from 1979 to 2009: trends and regional differences Original Article Mortality due to Chagas disease in Brazil from 1979 to 2009. *J Infect Dev Ctries* 2012;6(11):817-824.

Martinussen J, Hammer K. Cloning and characterization of *ufpp*, a gene encoding uracil phosphoribosyltransferase from *Lactococcus lactis*. *J Bacteriol* 1994;176:6457–6463.

Matsuda NM, Miller SM, Evora PRB. The chronic gastrointestinal manifestations of Chagas disease. *Clinics (São Paulo, Brazil).* 2009 Jan;64(12):1219–24.

Maugeri DA, Cazzulo JJ, Burchmore RJS, Barrett MP, Ogbunode POJ. Pentose phosphate metabolism in *Leishmania mexicana*. *Molecular and Biochemical Parasitology.* 2003 Aug;130(2):117–25.

Maugeri D a, Cazzulo JJ. The pentose phosphate pathway in *Trypanosoma cruzi*. *FEMS microbiology letters.* 2004 May 1;234(1):117–23.

Minning T a, Weatherly DB, Atwood J, Orlando R, Tarleton RL. The steady-state transcriptome of the four major life-cycle stages of *Trypanosoma cruzi*. *BMC genomics*. 2009 Jan;10:370.

Muñoz-Saravia SG, Haberland A, Wallukat G, Schimke I. Chronic Chagas' heart disease: a disease on its way to becoming a worldwide health problem: epidemiology, etiopathology, treatment, pathogenesis and laboratory medicine. *Heart failure reviews*. 2012 Jan;17(1):45–64.

Natalini P, Ruggieri S, Santarelli I, Vita A, Magni G. Baker's yeast UMP: pyrophosphate phosphoribosyltransferase. Purification, enzymatic and kinetic properties. *J Biol Chem*. 1979; 254:1558–1563.

Nelson DL, Cox MM. *Princípios da Bioquímica de bioquímica de Lehninger*. 5ªed. Porto Alegre: Artmed, 2011.

Neves DP, Melo AL, Linardi PM *et al*. *Parasitologia Humana*. 11ª ed. São Paulo: Atheneu, 2005.

Neuhard J, Nygaard P. Purines and pyrimidines. In *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, ed. FC Neidhardt, JL Ingraham, BK Low, B Magasanik, M Schaechter, HE Umbarger, pp. 445–73. Washington, DC: ASM Press, 1987.

Neves D.P.; Melo A.L.; Linardi, PM *et al*. *Parasitologia Humana*. 11ª ed. São Paulo: Editora Atheneu, 2005

O'Donovan G a, Neuhard J. Pyrimidine metabolism in microorganisms. *Bacteriological reviews*.1970 Sep;34(3):278–343.

Otto TD, Guimarães ACR, Degraive WM, de Miranda AB. AnEnPi: identification and annotation of analogous enzymes. *BMC bioinformatics* . 2008 Jan; 9:544.

OPAS. Organização Mundial de Saúde; 2012. Disponível em: <http://www.opas.org.br/>.

Parker ER, Sethi A. Chagas disease: coming to a place near you. *Dermatologic clinics*. 2011 Jan;29(1):53–62.

Pérez-Molina JA, Norman F, López-Vélez R. Chagas disease in non-endemic countries: epidemiology, clinical presentation and treatment. *Current infectious disease reports*. 2012 Jun;14(3):263–74

Polak A, Richle R. Mode of action of 2-nitroimidazole derivative benznidazole. *Ann Trop Med Parasitol*. 1978;72: 228-232.

Porco A, Gamero EE, Mylonás E, Istúriz T. Gluconate as suitable potential reduction supplier in *Corynebacterium glutamicum*: cloning and expression of gntP and gntK in *Escherichia coli*. *Biological Research*. 2008;41(3):349–58.

Prata A. Clinical and epidemiological aspects of Chagas disease. *The Lancet Infectious Diseases*. 2001;1: 92-100.

Rajaram V, Rajaganapathi J, Khan F, Savithri HS, Murthy MRN. Crystallization and preliminary X-ray crystallographic studies of recombinant thermoresistant gluconate kinase GntK from *Escherichia coli*. *Acta Crystallographica Section D Biological Crystallography*. 2002;58:359–61.

Rasmussen UB, Mygind B, Nygaard P. Purification and some properties of uracil phosphoribosyltransferase from *Escherichia coli* K12. *Biochimica et Biophysica Acta*. 1986;881(2):268–75.

Rassi A, Rassi A Jr, Rassi GG. Fase aguda da doença de Chagas. In: Brener Z, Andrade ZA, Barral-Netto M, editors. *Trypanosoma cruzi e doença de Chagas*. 2ªed, Rio de Janeiro: Guanabara Koogan, 2000.

Rassi A, Rassi A, Marin-Neto JA. Chagas disease. Lancet. Elsevier Ltd; 2010 Apr 17;375(9723):1388–402.

Rassi A, Rassi A, Marcondes de Rezende J. American trypanosomiasis (Chagas disease). Infectious disease clinics of North America. 2012 Jun;26(2):275–91.

Rauch B, Pahlke J, Schweiger P, Deppenmeier U. Characterization of enzymes involved in the central metabolism of *Gluconobacter oxydans*. Applied microbiology and biotechnology. 2010 Oct;88(3):711–8.

Renck D, Ducati RG, Palma MS, Santos DS, Basso LA. The kinetic mechanism of human uridine phosphorylase 1: Towards the development of enzyme inhibitors for cancer chemotherapy. Archives of Biochemistry and Biophysics. 2010;497(1-2):35–42.

Rey L. Parasitologia. 3ª ed. Rio de Janeiro: Guanabara Koogan, 2002.

SAMBROOK J, RUSSEL DW. Molecular cloning: a laboratory manual. 2ªed. Vol. 1,2 e 3. Cold spring harbor: Cold Spring Harbor laboratory press, 2001.

Schmunis GA. Epidemiology of Chagas disease in non-endemic countries: the role of international migration. Memórias do Instituto Oswaldo Cruz . 2007 Oct 30;102 Suppl 1(PAHO 2006):75–85.

Schumacher M a, Carter D, Scott DM, Roos DS, Ullman B, Brennan RG. Crystal structures of *Toxoplasma gondii* uracil phosphoribosyltransferase reveal the atomic basis of pyrimidine discrimination and prodrug binding. The EMBO journal . 1998 Jun 15;17(12):3219–32.

Schumacher M a, Bashor CJ, Song MH, Otsu K, Zhu S, Parry RJ, *et al*. The structural mechanism of GTP stabilized oligomerization and catalytic activation of the *Toxoplasma gondii* uracil phosphoribosyltransferase. Proceedings of the National Academy of Sciences of the United States of America. 2002 Jan 8;99(1):78–83.

Shuman, S. Novel Approach to Molecular Cloning and Polynucleotide Synthesis Using Vaccinia DNA Topoisomerase. *J. Biol. Chem.* 1994; 269:32678-32684.

Steindel M, Dias JCP, Romanha AJ. Doença de Chagas: mal que ainda preocupa. *Ciência Hoje.* 2005;37: 32–38.

Suzuki NN, Koizumi K, Fukushima M, Matsuda A, Inagaki F. Structural basis for the specificity, catalysis, and regulation of human uridine-cytidine kinase. *Structure London England* 1993. Elsevier Ltd; 2004;12(5):751–64.

Teixeira DE, Benchimol M, Crepaldi PH, De Souza W. Atlas didático: Ciclo de vida do *Trypanosoma cruzi*. 2ªed, Rio de Janeiro: Fundação CECIERJ, Consórcio CEDERJ, 2011.

Teixeira DE, Benchimol M, Crepaldi PH, de Souza W. Interactive multimedia to teach the life cycle of *Trypanosoma cruzi*, the causative agent of Chagas disease. *PLoS neglected tropical diseases.* 2012 Aug;6(8):e1749.

Villela AD, Ducati RG, Rosado LA, Bloch CJ, Prates MV, Gonçalves DC, *et al.* Biochemical characterization of uracil phosphoribosyltransferase from *Mycobacterium tuberculosis*. *PloS one.* 2013 Jan; 8(2):e56445.

Voet D, Voet JG. *Biochemistry.* 4ªed. EUA: Wiley, 2011

Zachariou M, Scopes RK. Gluconate kinase from *Zymomonas mobilis*: isolation and characteristics. *Biochemistry International.* 1985;10(3):367–71.

Zrenner R, Stitt M, Sonnewald U, Boldt R. Pyrimidine and purine biosynthesis and degradation in plants. *Annual review of plant biology.* 2006 Jan;57:805–36.

Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics.* Oxford University Press; 2009; 25(9):1189–91.

Wilson ZN, Gilroy C a, Boitz JM, Ullman B, Yates P a. Genetic dissection of pyrimidine biosynthesis and salvage in *Leishmania donovani*. The Journal of biological chemistry. 2012 Apr; 287(16):12759–70.

## 8. ANEXOS

### 8.1. Alinhamento das sequências de *TcGK* contra as sequências de referência de um dos clones de pBAD-TOPO TA, utilizando o programa ClustaW2.

CLUSTAL 2.1 multiple sequence alignment

```

GK/Tc00.1047053508177.20    ---ATGGCTGTCCCCCTCTTCCGTCGCATCCTACGCTGTCTGTGCCGGT 47
SequenciaGK-pBAD           CATATGGCTGTCCCCCTCTTCCGTCGCATCCTACGCTGTCTGTGCCGGT 50
                             *****

GK/Tc00.1047053508177.20    GGCTGTGGTTGTTCATGGGCCCTCGGGCACGGGCAAGAGACAATTGGTC 97
SequenciaGK-pBAD           GGCTGTGGTTGTTCATGGGCCCTCGGGCACGGGCAAGAGACAATTGGTC 100
                             *****

GK/Tc00.1047053508177.20    TCCGCCTTGCAGCAGGAGCTCGGCTGGGAATTTGCAGAAGGTGACAATTAT 147
SequenciaGK-pBAD           TCCGCCTTGCAGCAGGAGCTCGGCTGGGAATTTGCAGAAGGTGACAATTAT 150
                             *****

GK/Tc00.1047053508177.20    CACCCCGCGGGCAACGTAGAAAAGATGGCGTCGGGTGTGCCTCTTAATGA 197
SequenciaGK-pBAD           CACCCCGCGGGCAACGTAGAAAAGATGGCGTCGGGTGTGCCTCTTAATGA 200
                             *****

GK/Tc00.1047053508177.20    TGAGGACCGCTTGCCGTGGCTGAAAATCTGCGCACCGAGGCCATAGAGG 247
SequenciaGK-pBAD           TGAGGACCGCTTGCCGTGGCTGAAAATCTGCGCACCGAGGCCATAGAGG 250
                             *****

GK/Tc00.1047053508177.20    GTGCGTCGAGGGCGCGGCAGGAAGCGGTGTCGTGCTTGCATGTTCCGCG 297
SequenciaGK-pBAD           GTGCGTCGAGGGCGCGGCAGGAAGCGGTGTCGTGCTTGCATGTTCCGCG 300
                             *****

GK/Tc00.1047053508177.20    TTGCGAGAGTCGTACAGGGACGTGCTTCGCTCGGACCACTCGGGCAACGG 347
SequenciaGK-pBAD           TTGCGAGAGTCGTACAGGGACGTGCTTCGCTCGGACCACTCGGGCAACGG 350
                             *****

GK/Tc00.1047053508177.20    TGAAGAGTGAAAGTGTCTTTTGTGGAGCTTTTGGCGACATCGGGTTCA 397
SequenciaGK-pBAD           TGAAGAGTGAAAGTGTCTTTTGTGGAGCTTTTGGCGACATCGGGTTCA 400
                             *****

GK/Tc00.1047053508177.20    TTCCAAGCGCCTGCAGGGCGTCATCATGCGTATATGCCGGCTTGTTTA 447
SequenciaGK-pBAD           TTCCAAGCGCCTGCAGGGCGTCATCATGCGTATATGCCGGCTTGTTTA 450
                             *****

GK/Tc00.1047053508177.20    CTTGCATCGCAGTACGCCACGTTGGAGCCGCTGCGGCCGAAAAGGGCGA 497
SequenciaGK-pBAD           CTTGCATCGCAGTACGCCACGTTGGAGCCGCTGCGGCCGAAAAGGGCGA 500
                             *****

GK/Tc00.1047053508177.20    GCTTGGGGTGCAGTATCCGCGGCGTTGGAACCCGAGAAGATTGTGGAAC 547
SequenciaGK-pBAD           GCTTGGGGTGCAGTATCCGCGGCGTTGGAACCCGAGAAGATTGTGGAAC 550
                             *****

GK/Tc00.1047053508177.20    GTGTCGTGGAGAACTGAGGCGGTGGCTGCATCTTTGA----- 585
SequenciaGK-pBAD           GTGTCGTGGAGAACTGAGGCGGTGGCTGCATCTTTGACTCGAG 594
                             *****

```



## 8.2. Alinhamento das seqüências de *TcUPRT* contra as seqüências de referência de um dos clones de pBAD-TOPO TA, utilizando o programa ClustaW2.

CLUSTAL 2.1 multiple sequence alignment

```

UPRT/Tc00.1047053507739.40    ---AATACGGACTACTGACGCTTTTTATCGCACTCTCTACTGTTTCTCCA 47
sequenciaUPRT-pBAD           CATAATACGGACTACTGACGCTTTTTATCGCACTCTCTACTGTTTCTCCA 50
                               *****

UPRT/Tc00.1047053507739.40    TACCCGTTTTTTTGGGCTAGAAATAATTTGGTTTAACTTTAAGAAGGAGA 97
sequenciaUPRT-pBAD           TACCCGTTTTTTTGGGCTAGAAATAATTTGGTTTAACTTTAAGAAGGAGA 100
                               *****

UPRT/Tc00.1047053507739.40    TATACATACCCATGGGCTCTGGATCCGGTGATGACGATGACAAGCTTGAA 147
sequenciaUPRT-pBAD           TATACATACCCATGGGCTCTGGATCCGGTGATGACGATGACAAGCTTGAA 150
                               *****

UPRT/Tc00.1047053507739.40    GGTAAGCCTATCCCTAACCCCTCTCCTCGGTCTCGATTCTACGCGTACCGG 197
sequenciaUPRT-pBAD           GGTAAGCCTATCCCTAACCCCTCTCCTCGGTCTCGATTCTACGCGTACCGG 200
                               *****

UPRT/Tc00.1047053507739.40    TCATCATCACCATCACCATTGAGTTTAAACGGTCTCCAGCTTGCTGTTT 247
sequenciaUPRT-pBAD           TCATCATCACCATCACCATTGAGTTTAAACGGTCTCCAGCTTGCTGTTT 250
                               *****

UPRT/Tc00.1047053507739.40    TGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAATCAGAACGCA 297
sequenciaUPRT-pBAD           TGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAATCAGAACGCA 300
                               *****

UPRT/Tc00.1047053507739.40    GAAGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTC 347
sequenciaUPRT-pBAD           GAAGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTC 350
                               *****

UPRT/Tc00.1047053507739.40    CCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCCGTAGCGCCGATGG 397
sequenciaUPRT-pBAD           CCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCCGTAGCGCCGATGG 400
                               *****

UPRT/Tc00.1047053507739.40    TAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAAGTCCAGGCATCAAATA 447
sequenciaUPRT-pBAD           TAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAAGTCCAGGCATCAAATA 450
                               *****

UPRT/Tc00.1047053507739.40    AAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTT 497
sequenciaUPRT-pBAD           AAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTT 500
                               *****

```

## CLUSTAL 2.1 multiple sequence alignment

```

UPRT/Tc00.1047053507739.40      ACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAA 597
sequenciaUPRT-pBAD              ACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATAA 600
*****

UPRT/Tc00.1047053507739.40      ACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTT 647
sequenciaUPRT-pBAD              ACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTT 650
*****

UPRT/Tc00.1047053507739.40      TCGGTTTCTACAAACTCTTTGTTTATTTTCTAAATACATTCAAATATG 697
sequenciaUPRT-pBAD              TCGGTTTCTACAAACTCTTTGTTTATTTTCTAAATACATTCAAATATG 700
*****

UPRT/Tc00.1047053507739.40      TATCCGCTCATGAGACAATAACCCGTGATAAATGCTTCAATAATATTGAAA 747
sequenciaUPRT-pBAD              TATCCGCTCATGAGACAATAACCCGTGATAAATGCTTCAATAATATTGAAA 750
*****

UPRT/Tc00.1047053507739.40      AAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTT 797
sequenciaUPRT-pBAD              AAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTT 800
*****

UPRT/Tc00.1047053507739.40      TTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAA 847
sequenciaUPRT-pBAD              TTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAA 850
*****

UPRT/Tc00.1047053507739.40      GTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAAAC 897
sequenciaUPRT-pBAD              GTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAAAC 900
*****

UPRT/Tc00.1047053507739.40      TGATCTCAACAGGCGGTAAGATCCTTGAGAGTTTTTCCGCCCGGAGACG 947
sequenciaUPRT-pBAD              TGATCTCAACAGGCGGTAAGATCCTTGAGAGTTTTTCCGCCCGGAGACG 950
*****

UPRT/Tc00.1047053507739.40      TTTTCAATGATGACACTTAGTTCTGCTATGTTGGGCGCGTAATTATATCC 997
sequenciaUPRT-pBAD              TTTTCAATGATGACACTTAGTTCTGCTATGTTGGGCGCGTAATTATATCC 1000
*****

UPRT/Tc00.1047053507739.40      CCGGTGG----- 1004
sequenciaUPRT-pBAD              CCGGTGGCTCGAG 1013
*****

```