AMÉRICO TAVARES RANZANI

# **"ESTUDOS ESTRUTURAIS E FUNCIONAIS DA ENZIMA GLICOSE-6-**

# FOSFATO DESIDROGENASE"

CAMPINAS 2013



# UNIVERSIDADE ESTADUAL DE CAMPINAS INSTITUTO DE BIOLOGIA

AMÉRICO TAVARES RANZANI

## **"ESTUDOS ESTRUTURAIS E FUNCIONAIS DA ENZIMA GLICOSE-6-**

## FOSFATO DESIDROGENASE"

Est	e exe	mpia	IL CONTR	:50011	NG C	1 secondary	
da	tese	defe	endida	peic	)(a)	candidate	o (a)
A	MERI	10	TAVA	RES	RA	NZADU	
1	T	_	CI	1		)	
	Ant	m	Lz G	ody			
0 1	- Marina Marina	de r	oela Co	miss	ao j	uigadora	

**Orientador: Dr. Artur Torres Cordeiro** 

Dissertação apresentada ao Instituto de Biologia para obtenção do Título de Mestre em Biociências e Tecnologia de Produtos Bioativos, na área de Fármacos, Medicamentos e Insumos para Saúde.

CAMPINAS 2013

## FICHA CATALOGRÁFICA ELABORADA POR MARA JANAINA DE OLIVEIRA – CRB8/6972 BIBLIOTECA DO INSTITUTO DE BIOLOGIA - UNICAMP

R178e	Ranzani, Américo Tavares, 1988- Estudos estruturais e funcionais da enzima glicose-6- fosfato desidrogenase / Américo Tavares Ranzani. – Campinas, SP: [s.n.], 2013.
	Orientador: Artur Torres Cordeiro. Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Biologia.
	<ol> <li>Chagas, Doença de. 2. Glicose-6-fosfato desidrogenase. 3. Desidroepiandrosterona. 4. Cristalografia. I. Cordeiro, Artur Torres. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.</li> </ol>

## Informações para Biblioteca Digital

Título em Inglês: Structural and functional studies of the enzyme glucose-6phosphate dehydrogenase Palavras-chave em Inglês: Chagas' disease Glucose-6-phosphate dehydrogenase Dehydroepiandrosterone Crystallography Área de concentração: Fármacos, Medicamentos e Insumos para Saúde Titulação: Mestre em Biociências e Tecnologia de Produtos Bioativos Banca examinadora: Artur Torres Cordeiro [Orientador] Rafael Victorio Carvalho Guido André Luis Berteli Ambrosio Data da defesa: 19-02-2013 Programa de Pós Graduação: Biociências e Tecnologia de Produtos Bioativos Campinas, 19 de fevereiro de 2013.

**BANCA EXAMINADORA** 

Dr. Artur Torres Cordeiro (Orientador)

Dr. Andre Luis Berteli Ambrosio

Prof. Dr. Rafael Victório Carvalho Guido

Dr. Marcio Vinicius Bertacine Dias

Dra. Sandra Martha Gomes Dias

Arem F. Colif Assinatura
ANDRE AM
Assinatura
20/00 Cquido Assinatura
'

Assinatura

Assinatura

# Dedicatória

Aos meus pais, Américo e Jacira,

e irmãos, Olívia e Otávio.

## Agradecimentos

Aos meus pais, Jacira e Américo, e irmãos Olívia e Otávio, pelo carinho, incentivo e apoio.

À toda minha família.

Ao Dr. Artur Torres Cordeiro pela orientação, ensinamentos, paciência, apoio e confiança.

À Dra. Juliana Oliveira, Dra. Elisandra Rodrigues e Dra. Patrícia Castellen, pelos ensinamentos, apoio e incentivo diário.

Aos membros da pré-banca, Dr. André Ambrosio, Dra. Juliana Oliveira e Dr. Marcio Dias, pelas sugestões e críticas.

Aos colegas de laboratório, Alessandra, Alexandre, Aline Sampaio, Amanda, Ana Letícia, Ana Lú, Andrey, Ângela Fala, Ângela Saito, Ariane, Bia, Camila Fornezari, Carla, Carol Cassago, Cássia, Cristiane, Daniel, Emerson, Fabiana, Fernanda, Germanna, Gustavo, Henrique, Igor, Jéssica Faria, Jéssica Lois, Joyce, Juliana Bueno, Juliana Fattori, Juliana Smetana, Kaliandra, Kelven, Leandro, Malú, Marcel, Mariana, Mariela, Maruska, Melissa, Michelle, Patrick, Paula, Plínio, Renata Baroni, Roberta, Rosa, Silvana, Vanessa Pegos, pela experiência compartilhada e os momentos de descontração.

Aos técnicos, Andreia, Celisa, Givanil, Jackeline, Sami, Tereza e Wanderley, sempre dispostos a ajudar no que fosse possível.

Aos funcionários do laboratório.

Ao programa de pós-graduação, Biociências e Tecnologia de Produtos Bioativos.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo financiamento do projeto.

Ao Laboratório Nacional de Biociências e Laboratório Nacional de Luz Síncrotron, pela infraestrutura oferecida.

# Sumário

Lista de Figuras	viii
Lista de Tabelas	x
Lista de abreviações	xi
Resumo	xiii
Abstract	xv
Capitulo 1: Introdução	1
1.1 Tripanossomíase Humana Americana	1
1.2 Desenvolvimento de fármacos baseado em estrutura	3
1.3 Via de Pentose Fosfato e G6PD	4
1.4 Interação G6PD com DHEA	8
1.5 Objetivos:	9
Capítulo 2: Materiais e Métodos	11
2.1 Subclonagem da Hs $\Delta$ G6PD e Tc $\Delta$ G6PD	11
2.2 Teste de expressão	14
2.3 Expressão em meio de auto-indução	15
2.4 Purificação Hs∆G6PD	15
2.4.a Lise celular	15
2.4.b Afinidade em resina de níquel e clivagem na resina por sumo protease	15
2.4.c Troca iônica	16
2.4.d Cromatografia de exclusão molecular	16
2.5 Estudo de estado oligomérico por cromatografia analítica de exclusão molec	cular 16
2.6 Cinética enzimática	17
2.7 Ensaio de termofluorescência	
2.8 Mutação sítio dirigida	
2.9 Ensaios de cristalização	19
2.10 Coleta de dados de difração de Raio-X e processamento dos dados	20

Capitulo 3: Subclonagem, produção heteróloga e caracterização cinética da G6PD22
3.1 Subclonagem da HsΔG6PD e TcΔG6PD22
3.2 Teste de expressão em meio auto-indutor 22
3.3 Purificação24
3.3.a Afinidade com resina de Níquel24
3.3b Troca iônica da Hs $\Delta$ G6PD27
3.3c Gel filtração com Superdex 200 16/6027
3.4. Avaliação da atividade da enzima purificada27
Capítulo 4: Formulação de complexo estável
4.1 Avaliação de inibidores competitivos
4.2 Mutante D200N
Capítulo 5: Estudos cristalográficos da Hs∆G6PD40
Capítulo 6: Mutantes A277C e E347A 48
6.1 A HsΔG6PD DNEA
6.2 HsΔG6PD DNAC
Capítulo 7: Conclusão
Referências
Anexo

# Lista de Figuras

Figura 1: Ciclo de vida do <i>Trypanosoma cruzi</i> 1
Figura 2: Fármacos utilizados atualmente para o tratamento da fase aguda da doença de Chagas.
Figura 3: Via de síntese de pentose fosfato5
Figura 4: Estrutura da G6PD humana em seus diferentes estados oligoméricos
Figura 5: Estrutura química dos esteroides inibidores da G6PD e tabela contendo as constantes de
inibição9
Figura 6: Mapa dos vetores de expressão utilizados13
Figura 7: Gel de agarose 1% do produto de PCR22
Figura 8: Gel de agarose 1% da digestão com enzimas de restrição
Figura 9: Curva de crescimento de BL21 com plasmídeo da Hs∆G6PD em meio ZYM5052 24
Figura 10: A: Gel SDS-PAGE 10% da clivagem da SUMO com ULP-1 à 25°C por 4 horas da Tc∆G6PD
e Hs∆G6PD
Figura 11: SDS-PAGE 15% da clivagem na coluna de afinidade da Hs $\Delta$ G6PD26
Figura 12: A: cromatograma da troca iônica, B: cromatograma da gel filtração em Superdex200
16/60 C: SDS-PAGE 15% das diferentes etapas de purificação da Hs∆G6PD
Figura 13: Curvas de Michaelis-Menten para determinação de K <sub>m</sub> para os substratos
Figura 14: Curvas dose resposta para determinação de IC <sub>50</sub> dos esteroides para a construção
HsΔG6PD
Figura 15: Compostos utilizados para formulação de complexo ternário
Figura 16: IC <sub>50</sub> para glucosamina 6-fosfato para Hs $\Delta$ G6PD
Figura 17: Exemplo de dados obtidos com o ensaio de termofluorêscencia estabelecido 32

Figura 18: Deslocamento no $\rm T_m$ causado pelos diferentes ligantes para a Hs $\Delta \rm G6PD$	em ensaio de
termofluorescência	33
Figura 19: Esquema geral do mecanismo de reação da G6PD	
Figura 20: Inibição do mutante D200N pelo esteroides	
Figura 21: Cavidade próxima ao D200.	
Figura 22: Alinhamento de diferentes sequências de G6PD	
Figura 23: Cristais da Hs $\Delta$ G6PD	43
Figura 24: Conteúdo da unidade assimétrica	
Figura 25: Mapa diferença ( $F_0$ - $F_c$ ) em verde com cotorno de $3\sigma$	46
Figura 26: Interface do tetrâmero. 5	
Figura 27: Perfil do cromatograma de gel filtração dos mutantes Hs∆G6PD D2	200N E347A e
HsΔG6PD D200N A277C em <i>Superdex 200 10/30</i>	
Figura 28: Fotos dos cristais obtidos com o mutante Hs∆G6PD DNEA em scree	<i>ning</i> inicial de
cristalização	50
Figura 29: Termofluorescência do mutante Hs∆G6PD DNEA	51
Figura 30: Mapas de densidade eletrônica para o modelo da Hs $\Delta$ G6PD D2001	N A277C após
primeiro ciclo de refinamento	53
Figura 31: Sobreposição das estruturas da Hs $\Delta$ G6PD e Hs $\Delta$ G6PD DNAC	54

# Lista de Tabelas

Tabela 1: Ciclo de PCR para amplificação Hs∆G6PD e Tc∆G6PD12
Tabela 2 : Oligonucleotídeos utilizados para mutação pontual19
Tabela 3: Constante de Michaelis-Menten ( $K_M$ ) para os substratos da G6PD
Tabela 4: Parâmetros cinéticos para o mutante Hs∆G6PD D200N
Tabela 5: Resumo dos conjuntos coletados para Hs $\Delta$ G6PD44
Tabela 6: Dados estatísticos referentes ao processamento dos dados
Tabela 7: Parâmetros estatísticos do refinamento da estrutura da Hs∆G6PD – G6P
Tabela 8: Dados estatísticos referentes ao processamento dos dados do cristal da Hs $\Delta$ G6PD D200N
A277C crescidos na presença de G6P+NADPH+16BrEA52

# Lista de abreviações

- °C Grau Celsius
- 16BrEA- 16-bromo-epiandrosterona
- DHEA- Desidroepiandrosterona
- DNA- Ácido desoxirribonucleico
- EA- Epiandrosterona
- EDTA Ácido etilenodiamino tetra-acétic
- F6P- Frutose 6-fosfato
- G6P- Glicose 6-fosfato
- G6PD Glicose 6-fosfato desidrogenase
- GIn6P- Glucosamina 6-fosfato
- HsG6PD Glicose 6-fosfato desidrogenase de Homo sapiens
- IC<sub>50</sub> Concentração de inibição de 50% de atividade enzimática
- IPTG- Isopropil-β-D-tiogalactosídeo
- $k_{\text{cat}}$  Constante catalítica
- k<sub>d</sub> Constante de dissociação
- *k*<sub>i</sub> Constante de inibição
- K<sub>m</sub> Constante de Michaelis
- L. mexicana Leishmania mexicana
- LB- Luria-Bertani
- LD<sub>50</sub> Dose letal mediana
- LmG6PD- Glicose 6-fosfato desidrogenase de Leishmania mexicana

LNLS- Laboratório Nacional de Luz Síncrotron

- NADP<sup>+</sup>- Nicotinamida adenina dinucleótido fosfato
- NADPH- Nicotinamida adenina dinucleótido fosfato reduzida
- PAGE PolyAcrylamide Gel Electrophoresis Eletroforese em gel de Poliacrilamida
- PCR- Polymerase Chain Reaction reação em cadeia da polimerase
- PDB- Protein Data Bank Banco de dados de proteína
- RNA- Ácido ribonucleico
- RNAi- RNA de interferência
- SBDD- Structure-based Drug Design- Desenho de fármaco baseado em estrutura
- S-DHEA Desidroepiandrosterona sulfato
- SDS- Sodium Dodecyl Sulfate- Dodecil sulfato de sódio
- T. brucei Trypanosoma brucei
- T. cruzi Trypanosoma cruzi
- TAE Tampão Tris-Acetato-EDTA
- TbG6PD- Glicose 6-fosfato desidrogenase de Trypanosoma brucei
- TcG6PD- Glicose 6-fosfato desidrogenase de Trypanosoma cruzi
- V<sub>max</sub> Velocidade máxima

# Resumo

Tripanossomíases são enfermidades associadas à infecção por protozoários do gênero Trypanosoma. Estas doenças são normalmente tratadas com medicamentos de alta toxicidade e baixa eficiência. A pesquisa de novos compostos que atuem de maneira específica contra alvos metabólicos pré-estabelecidos pode levar ao desenvolvimento de fármacos mais eficientes no tratamento destas enfermidades. A enzima glicose-6-fosfato desidrogenase (G6PD) é um alvo metabólico validado experimentalmente em Trypanosoma brucei. A G6PD catalisa o primeiro passo da via de síntese de pentoses que supre a célula com ribose-5-fosfato para síntese de bases nitrogenadas e NADPH para biosíntese de lípideos, colesterol e neutralização de espécies reativas de oxigênio. O esteroide desidroepiandrosterona (DHEA) e seus análogos são conhecidos inibidores incompetitivos da enzima humana e recentemente foram caracterizados também como potentes inibidores da G6PD de T. brucei e T. cruzi. Estes esteroides também apresentam efeito citotóxico à forma sanguínea de T. brucei e à forma epimastigota de T. cruzi. Em conjunto, estes resultados sugerem que esteroides análogos do DHEA devam ser explorados como uma nova classe de fármacos anti-parasitários. Apesar de já haver estruturas da enzima, o sítio de ligação destes esteroides não é conhecido, informação que ajudaria no desenho racional de novos inibidores. Assim, este projeto teve como objetivo principal a identificação do sítio de inibição por cristalografia. Conseguiu-se estabelecer um protocolo de expressão, purificação e cristalização para a G6PD humana. Testou-se a co-cristalização, seeding e soaking com os esteroides DHEA, epiandrosterona (EA) e 16-bromo-epiandrosterona (16BrEA) com os substratos da enzima, além de outros complexos com frutose-6-fosfato, glucosamina-6-fosfato e NADPH. Na tentativa de diminuir a atividade da enzima para facilitar a cristalização com os substratos e os esteroides, fezse o mutante D200N, que além de ser menos ativo, também apresentou uma menor inibição pelos esteroides. A enzima humana se apresenta como um equilíbrio de dímero e tetrâmero, não se sabendo a influência destes estados na atividade da enzima. Assim, para favorecer somente um estado oligomérico, foram avaliadas mutações na interface do tetrâmero. A mutação A277C permite a formação de uma ponte dissulfeto com a C294, estabilizando o tetrâmero. A mutação E347A é capaz de desestabilizar o tetrâmero, favorecendo a forma dimérica. Embora não tenha sido possível obter a estrutura da enzima com os esteroides, conseguiu-se pela primeira vez apontar para um resíduo importante para a inibição, o que abre oportunidade para investigações futuras que podem levar à descrição do sítio de inibição.

# Abstract

Trypanosomiases are infectious diseases caused by protozoan parasites from genus Trypanosoma. Such diseases are currently treated with low effective and highly toxic drugs. The research of novel compounds which inhibits validated metabolic targets might lead to the development of more efficient drugs to such neglected diseases. The enzyme glucose-6-phosphate dehydrogenase (G6PD) was experimentally validated as a metabolic target in T. brucei bloodstream form. G6PD catalyzes the first step of the pentose phosphate pathway (PPP), which supplies the cell with ribose-5-phosphate for synthesis of nucleotides and NADPH for biosynthesis of lipids, cholesterol and detoxification of reactive oxygen species. The steroid Dehydroepiandrosterone (DHEA) and analogues are known uncompetitive inhibitors of the human G6PD and recently they were also characterized as potent inhibitors of *T. brucei* and *T. cruzi* G6PD. It was also demonstrated that these steroids are toxic to cultured *T. brucei* bloodstream form and T. cruzi epimastigote form. Together these results suggest that DHEA derivatives must be explored as a novel anti-parasite drug class. Although the structure of the enzyme is solved, the binding site of these steroids is unknown, information that would help in the rational design of new inhibitors. So, the main goal of this project was the identification of the inhibition site by crystallography. It was established an expression, purification and crystallization protocol for the human G6PD. It was tried the co-crystallization, seeding and soaking of the steroids DHEA, epiandrosterone (EA) and 16-bromo-epiandrosterone (16BrEA) with the substrates of the enzyme, beyond other complexes with fructose-6-phosphate, glucosamine-6-phosphate and NADPH. Attempting to diminish the enzyme activity to facilitate the crystallization with the substrates and the steroids, it was made the mutant D200N, which not only is less active, but it's also less inhibited by the steroids. The human enzyme is in equilibrium between dimer and tetramer, being unknown the effect of these states to the enzyme activity. This way, to favor one oligomer state, it was assessed mutations at the tetramer interface. The mutation A277C allows the formation of a disulfide bond with the C294, stabilizing the tetramer. The mutation E347A is able to destabilize the tetramer, favoring the

dimer. Although it was not possible to determine the structure of the enzyme with the steroids, for the first time it was shown the involvement of a residue in the steroid inhibition, which makes possible future investigations that can lead to the inhibition site description.

# Capítulo 1: Introdução

## 1.1 Tripanossomíase Humana Americana

A tripanossomíase americana, ou Mal de Chagas, é uma doença negligenciada parasitária causada pelo protozoário *Trypanosoma cruzi*, família *Trypanosomatidae*, ordem *Kinetoplastida*, e transmitida por insetos dos gêneros *Triatoma*, *Rhodnius* e *Panstrongylus*, conhecidos como barbeiros. A doença foi descoberta no começo do século XX pelo médico brasileiro Carlos Chagas.<sup>1</sup>

O ciclo de vida do parasita é complexo e possui diferentes estágios, divididos entre o vetor e o hospedeiro mamífero. No triatomíneo, encontra-se a forma replicativa epimastigota, que se multiplica no intestino do inseto, e a forma tripomastigota metacíclica, a forma infectante, que sai nas fezes do vetor durante o repasse sanguíneo. No hospedeiro mamífero, o tripanossomatídeo é encontrado na corrente sanguínea na forma de tripomastigota, que é capaz de invadir todos os tipos de células nucleadas, diferenciando-se na forma replicativa amastigota. Dentro da célula, estas se multiplicam e se transformam novamente em tripomastigotas, destruindo as células e voltando à corrente sanguínea. O ciclo se fecha com a picada do barbeiro, que é infectado com a forma tripomastigota<sup>1</sup> (Figura 1).



Figura 1: Ciclo de vida do *Trypanosoma cruzi*. Figura adaptada do *Centers for Diseases Control and Prevention* (www.cdc.gov/).

A transmissão da doença também pode ocorrer por transfusão de sangue, transplante de órgãos infectados, ingestão de alimentos contendo fezes contaminadas do inseto e por transmissão congenital.

Estima-se que 8 milhões de pessoas estejam infectadas e 120 milhões estejam em área de risco<sup>2</sup>, sendo a América Latina a área mais afetada. A fase aguda, que pode durar entre 4 a 8 semanas após a infecção, é geralmente assintomática, o que dificulta o diagnóstico. Quando apresenta sintomas, estes vão de febre, mal-estar, aumento do baço, fígado e linfonodos. No caso de transmissão por vetor, o local de entrada do parasita pode apresentar um edema inflamatório, chamado de chagoma, que quando ocorre na pálpebra, é chamado de sinal de Romaña.

A fase crônica pode ser assintomática durante anos, sendo diagnosticada quando já ocorreram sérios danos a tecidos parasitados. O protozoário se multiplica lentamente no interior de órgãos como fígado, baço, sistema nervoso e coração. Assim, após décadas, pode haver o aparecimento de demência, mega-cólon, megaesôfago e cardiomiopatia, causados por danos nas células nervosas e musculares destes órgãos. O mais frequente e severo é a doença crônica cardíaca, que acomete em torno de 20% a 30% dos casos e leva a insuficiência cardíaca, podendo causar a morte.<sup>1</sup>

No momento só há tratamento eficaz para a fase aguda da doença, que é feita com benzonidazol e/ou nifurtimox. Porém, estas drogas apresentam longo período de tratamento (60 e 60-90 dias, respectivamente), e efeitos colaterais, como anorexia, náusea, alucinação e hipersensibilidade, levando à necessidade de acompanhamento médico durante o tratamento, o que é complicado em regiões pobres, onde justamente há o maior número de casos. Além disso, há casos de cepas resistentes ao tratamento. <sup>3</sup> Assim, drogas que sejam eficazes também à forma crônica e mais seguras são desejadas para estas doença, sendo necessária a pesquisa de novos fármacos.



Benzonidazol

Nifurtimox

Figura 2: Fármacos utilizados atualmente para o tratamento da fase aguda da doença de Chagas.

#### 1.2 Desenvolvimento de fármacos baseado em estrutura

Uma abordagem para o desenvolvimento de fármacos que combinem efeitos colaterais reduzidos e potência é o desenho racional de uma pequena molécula que module seletivamente a atividade de um alvo terapêutico. Este alvo geralmente é caracterizado por uma macromolécula biológica, comumente proteínas, que se apresenta essencial à manutenção da doença, sendo que sua modulação seja capaz de alterar o estado senil. Para o caso de doenças parasitárias, um alvo terapêutico seria uma proteína indispensável para a sobrevivência, proliferação e/ou infectividade do parasita. É interessante que o alvo seja exclusivo ao parasita ou que existam diferenças estruturais entre a proteína do parasita e do hospedeiro para criação de compostos seletivos. Assim, o desenho de novas moléculas é guiado para atuar especificamente neste alvo, diminuindo as chances de efeitos colaterais e tendo potência. Outros parâmetros também devem ser avaliados durante o processo do desenvolvimento destes novos compostos, considerando que não só estes devem atuar no alvo, como também devem possuir propriedades farmacocinéticas para que consigam chegar ao seu destino dentro do organismo.<sup>4</sup>

Uma ferramenta para validação de alvos terapêuticos é o RNA de interferência, na qual se utiliza uma maquinaria de defesa própria do organismo com o auxílio de técnicas de biologia molecular para a redução transitória do nível de expressão de uma determinada proteína, podendo assim avaliar a importância desta para o parasita. No caso dos tripanossomatídeos, *Trypanosoma brucei* e *Trypanosoma congolense* possuem tal maquinaria, que está ausente em *T. cruzi.*<sup>5</sup> Assim, geralmente se usa o *T. brucei* como modelo para validação de alvos contra tripanossomíase africana e americana.

No cenário de desenho racional de fármacos, um conceito utilizado é o Desenho de Fármaco Baseado na Estrutura (*Structure-Based Drug Design*, SBDD), no qual se utiliza a informação estrutural do alvo para guiar o desenvolvimento de novas entidades químicas, avaliando a interação entre ligantes e a estrutura da proteína. Exemplos de fármacos que chegaram ao mercado e que foram desenvolvidos com este conceito são a dorzolamida, inibidor da anidrase carbônica, utilizada como anti-glaucoma, e o saquinavir, utilizado no tratamento contra o vírus da imunodeficiência humana-1 (HIV-1).<sup>6</sup>

Para aplicação do SBDD é interessante ter um ligante conhecido que module especificamente o alvo para que este seja um molde para se fazer alterações químicas. No caso de enzimas, é interessante ter um inibidor com inibição de 50% da atividade ( $IC_{50}$ ) no ordem de

3

micromolar como composto inicial, chamado *hit*. A partir deste hit são feitos análogos, podendo chegar numa série de compostos líderes, com IC<sub>50</sub> inferior de 100 nM. Deve-se também avaliar as propriedades farmacocinéticas destes compostos: absorção, distribuição, metabolismo, excreção e toxicidade (ADMET). A avaliação das propriedades de ADMET requer ensaios em cultura de células e em modelos animais bem estabelecidos.<sup>7</sup> Os compostos que possuem potências e boas propriedades farmacocinéticas são chamados de candidatos a fármaco, que então seguem para ensaios pré-clínicos. Apesar de não ser essencial, informações estruturais podem ajudar em todas estas etapas do desenvolvimento, desde na sugestão de mudanças químicas que aumentem a afinidade pelo alvo, até modificações que melhorem as propriedades farmacocinéticas, mas que não alterem a ligação no sítio de inibição.

A enzima glicose 6-fosfato desidrogenase (G6PD), primeira enzima do ciclo de pentose fosfato, é um alvo validado em *T. brucei* por RNAi.<sup>8</sup> Além disso, esta enzima de *T. cruzi* e *T. brucei* é inibida por esteroides humanos análogos de desidroepiandrosterona (DHEA) com IC<sub>50</sub> abaixo de 30 μM, tendo estes compostos atividade contra a forma da corrente sanguínea de *T. brucei*, e contra a forma epimastigota de *T. cruzi*.<sup>8,9</sup> Assim, a G6PD se apresenta como um interessante alvo para o desenvolvimento de fármacos contra tripanossomíase, tendo inibidores esteroides como compostos moldes para o desenho racional de novas moléculas. Porém, o sítio de ligação destes inibidores não é conhecido, dificultando estudos de SBDD.

#### 1.3 Via de Pentose Fosfato e G6PD

A via de pentose fosfato utiliza a glicose 6-fosfato (G6P) para produção de ribose-5fosfato, açúcar constituinte de nucleotídeos, e NADPH, principal agente redutor das células. Este via pode ser separada em duas fases, uma oxidativa, onda há a redução de NADP<sup>+</sup> e produção de ribulose-5-fosfato, e outra não-oxidativa, na qual há a interconversão de pentoses fosfatadas, formando frutose-6-fosfato (F6P), que pode voltar à via glicolítica ou gliconeogênica<sup>10</sup> (Figura 3-A).

A G6PD (E.C.1.1.1.49) é a primeira enzima da via, catalisando a oxidação da glicose 6fosfato (G6P) em 6-fosfoglucono- $\delta$ -lactona com a concomitante redução de NADP<sup>+</sup> a NADPH. O NADPH produzido é utilizado como doador de elétrons por enzimas da biossíntese de diferentes moléculas biológicas, como ácidos graxos<sup>11</sup> e esteróis<sup>12</sup>. Além disso, é usado pela enzima tripanotiona redutase para manter o composto tripanotiona (*N1,N8- bisglutathionyl spermidine*) no seu estado reduzido (T(SH)<sub>2</sub>).<sup>13</sup> A principal função da T(SH)<sub>2</sub> é neutralizar peróxidos produzidos pelo sistema de defesa do hospedeiro mamífero contra a invasão celular. Este processo de desintoxicação é realizado por enzimas do tipo peroxidases que utilizam elétrons doados diretamente da tripanotiona ou através do intermediário triparedoxina (Tpx). Outra familia de enzimas que utiliza o potencial redox da tripanotiona (e da Tpx) são as ribonucleotídeo redutases. Estas enzimas catalizam o primeiro passo na via de síntese *de novo* de precursores de DNA<sup>13,14</sup> (Figura 3-B).



Figura 3: A: Via de síntese de pentose. As enzimas participantes são: glucose 6-phosphate dehydrogenase (G6PD), 6-phosphogluconolactonase (6-PGL), 6-phosphogluconate dehydrogenase (6PGDH), ribose 5-phosphate isomerase (RPI), ribulose 5-phosphate epimerase (RPE), transaldolase (TAL) and transketolase (TKT). Figura adaptada de Igollio-Esteve *et al*,2007. B: Sistema Tripanotiona Peroxidase. As enzimas participantes são: Tripanotiona redutase, Triparedozina (Tpx), Ribonucleotideo redutase e Peroxidase. O composto tripanotiona pode alternar-se entre o estado reduzido (T(SH)2) e oxidado (TS2). Figura adaptada de Krauth-Siegel *et al*, 2003.

Assim, dada a importância de manter a tripanotiona reduzida para defesa dos tripanossomatídeos, as duas desidrogenases da via de pentose fosfato são interessantes alvos terapêuticos para tripanossomíases humanas. A inibição da G6PD, por ser a primeira enzima do ciclo, acaba por inibir toda a produção de NADPH por esta via. Sem esquecer também que a produção de nucleotídeos, ácidos graxos e esteróis também são afetadas.

Outra característica da G6PD que potencializa seu uso como alvo para desenvolvimento de fármacos contra doenças infecciosas é o fato do ser humano ser tolerante a deficiência de sua própria G6PD. Isso reduz a chance de efeitos colaterais decorrentes de um fármaco inibidor da G6PD com baixa seletividade para a enzima do parasita. A deficiência de G6PD é uma doença genética hereditária ligada ao cromossomo X causada por mutações no gene da G6PD, sendo identificadas por volta de 140 mutações, resultando em variações da proteína com diferentes níveis de atividade enzimática, que são associadas com uma ampla variedade de fenótipos bioquímicos e clínicos, que podem ser classificados em 5 classes:

- Classe 1: deficiência severa, associada com anemia crônica
- Classe 2: deficiência severa (1-10% atividade) anemia aguda
- Classe 3: moderada (10-60% atividade)
- Classe 4: atividade normal (60-150%)
- Classe 5: atividade aumentada (>150%)

O sintoma mais comum é icterícia neonatal e anemia hemolítica aguda, a qual na maioria das vezes é iniciada por agentes exógenos. A alta similaridade entre as áreas onde a deficiência de G6PD é comum e a malária é endêmica evidencia que a deficiência confere resistência à malaria. A deficiência é predominante na África, Ásia, mediterrâneo e oriente médio.<sup>16</sup> As principais evidências sugerem que deve haver um retardo no processo de maturação do *Plasmodium spp.* em hemácias deficientes de G6PD. Isso aumentaria a eficiência do processo de fagocitose das hemácias infectadas antes da eclosão de parasitas maduros na corrente sanguínea.<sup>17</sup>

A enzima geralmente é constituída de 500 a 600 resíduos, com duas regiões altamente conservadas: a sequência GxxGDLx, referente à ligação ao NADP; e RIDHYLGK, referente à ligação da G6P. A enzima pode ser encontrada na forma dimérica e/ou tetramérica, dependendo da espécie. Estudos de ultracentrifugação analítica da G6PD humana (HsG6PD) mostram que esta se encontra em equilíbrio entre dímero e tetrâmero, sendo este equilíbrio influenciado por força iônica e pH.<sup>18,19</sup>

Atualmente, há 14 estruturas depositadas no PDB (*Protein Data Bank*), sendo 9 estruturas de *Leuconostoc mesenteroides*, 3 de *Homo sapiens* e 2 de *Trypanosoma cruzi*. A primeira estrutura obtida (PDB:1DPG) foi de *L. mesenteroides* em 1994<sup>20</sup>, mostrando que a enzima apresenta dois domínios: um menor no N-terminal com uma conformação clássica  $\beta$ - $\alpha$ - $\beta$  para ligação de dinucleotídeos (do tipo *Rossman fold*); e um domínio maior  $\beta$ - $\alpha$  referente à ligação a G6P, que é

principalmente formado por uma folha- $\beta$  antiparalela de 9 fitas. A enzima de *L. mesenteroides* é somente encontrada na forma dimérica. A interface de dimerização é formada pela interação entre as folhas- $\beta$ , sendo predominantemente constituída por resíduos hidrofóbicos. Porém, apresenta duas ligações de hidrogênio e algumas pontes salinas, sendo conservado o E183 e a K386 entre diversos organismos.<sup>20</sup>

A primeira estrutura cristalográfica humana (PDB: 1QKI), obtida em 2000, é de um mutante da deficiência humana (variante Canton, R459L) e se apresenta como um tetrâmero, ajudando na compreensão da interface de tetramerização. Este é formado por dois dímeros e é predominantemente formado por interações eletrostáticas. Além disto, essa estrutura comprovou que a enzima humana possui um sítio estrutural para NADP<sup>+</sup>, formado por folhas- $\beta$  próximo à interface de dimerização<sup>21,22</sup> (Figura 4). Estudos posteriores mostraram que a constante de dissociação do NADP<sup>+</sup> estrutural (k<sub>d</sub> = 37 nM) é por volta de 200 vezes menor que a do NADP<sup>+</sup> catalítico e que a ligação deste está associado ao aumento da estabilidade da enzima, e não à sua atividade.<sup>23</sup>



Figura 4: Estrutura da G6PD humana (PDB:2BH9) em seus diferentes estados oligoméricos. Os estados ativos são o dímero e o tetrâmero. Em esferas no dímero está o NADP<sup>+</sup> estrutural, próximo à interface de dimerização.

Das mutações referentes à classe 1 da deficiência humana de G6PD, a forma mais severa da doença, 26 estão dentro ou próximas da interface de dimerização e 14 estão próximas do sítio estrutural de NADP<sup>+</sup>. Isto corrobora que a enzima precisa estar na forma dimérica para ser ativa, apesar de possuir o sítio catalítico completo no monômero.<sup>22</sup> A função do NADP<sup>+</sup> estrutural em dar estabilidade à enzima também é reforçada, pois como as hemácias não possuem núcleo, não podem reabastecer seu repertório proteico, requerendo que suas proteínas permaneçam estáveis e ativas durante todo o tempo de vida da célula (em torno de 120 dias). Por isso, alterar a estabilidade da G6PD com mutações na região do NADP<sup>+</sup> estrutural causa anemia crônica, pois estas estão em constante estresse oxidativo.<sup>22,23</sup> Talvez estas mutações não sejam tão relevantes às células nucleadas, uma vez que estas podem reciclar suas enzimas.

Em 2005 foi publicada a estrutura da humana nativa (PDB: 2BH9, 2BHL), sendo que o cristal só foi obtido com a retirada de parte do N-terminal. As estruturas da humana mostram que não há grandes diferenças estruturais na cadeia principal com relação à de procarioto.<sup>24</sup> Em 2012 foram depositadas as estruturas de *Trypanosoma cruzi* (PDB: 4EM5, 4E9I), que também apresentam uma deleção no N-terminal.<sup>25</sup>

#### 1.4 Interação G6PD com DHEA

Em 1960, Marks e Banks demonstraram que a HsG6PD é inibida por esteroides como DHEA e pregnenolona, sendo que a enzima de levedura ou de planta não são, sugerindo que estes inibidores são específicos para enzimas de mamíferos.<sup>26</sup> Porém, em 2009, Cordeiro e colaboradores mostraram que a enzima de *Trypanosoma brucei* (TbG6PD) também é inibida<sup>8</sup>, quebrando com o conceito que somente enzimas de mamíferos seriam inibidas. Hoje se sabe que além da enzima de *T. brucei*, a enzima de *T. cruzi* (TcG6PD) também é inibida<sup>9</sup>, e curiosamente a enzima de *Leishmania mexicana* não apresenta inibição. Os estudos cinéticos mostram que a inibição é do tipo incompetitiva em relação aos dois substratos, ou seja, os esteroides não se ligam à enzima livre, mas sim ao complexo ternário enzima:G6P:NADP<sup>+</sup>. <sup>8,27</sup>

A desidroepiandrosterona (DHEA) e sua forma sulfatada S-DHEA são pró-hormônios esteroidais sintetizados pelas glândulas adrenais de mamíferos. Eventualmente, eles são requisitados em tecidos específicos para serem convertidos em hormônios sexuais como estrógeno e testosterona. Em homens e mulheres (até os 25 anos) os valores de concentração sanguínea de DHEA e S-DHEA variam entre 1,5 a 25 nM e 2,2 a 9 μM, respectivamente.<sup>28</sup> Após os 25 anos ocorre uma queda

contínua na produção destes pró-hormônios. O DHEA tem sido frequentemente usado em terapias de reposição hormonal. Sua administração está correlacionada com a diminuição da incidência de doenças como câncer, diabetes e arteriosclerose. Ainda não está claro como o DHEA interfere no surgimento de doenças associadas ao envelhecimento, mas evidências sugerem que o controle exercido sobre a via de síntese de pentoses poderia estar relacionado com estes efeitos.<sup>29</sup>

A DHEA, a epiandrosterona (EA) e o ánalogo 16-Bromo-epiandrosterona (16BrEA) são tóxicos a forma sanguínea de *T. brucei* e a forma epimastigota de *T. cruzi*, mas inertes a culturas de *L. mexicana*.<sup>8,9</sup> Além disso, parasitas mutantes de *T. brucei* que expressam a G6PD de *L. mexicana* também são inertes a DHEA e EA<sup>30</sup>, evidenciando que a toxicidade dos esteroides é causada pela inibição da G6PD. Abaixo na figura 5 é possível ver os valores das constantes de inibição dos esteroides para as diferentes enzimas e a atividade tóxica contra culturas de tripanossomas.



	HsGe	5PD <sup>19</sup>	TcG	6PD <sup>9</sup>	TbG	6PD <sup>8</sup>	T. cruzi <sup>9</sup>	Т.
								brucei <sup>8</sup>
	k <sub>i</sub> <sup>NADP</sup> (μM)	k <sub>i</sub> <sup>G6P</sup> (μM)	k <sub>i</sub> <sup>NADP</sup> (μM)	k <sub>i</sub> <sup>G6P</sup> (μM)	k <sub>i</sub> <sup>NADP</sup> (μM)	k <sub>i</sub> <sup>G6P</sup> (μM)	LD <sub>50</sub> (μM)	LD <sub>50</sub> (μM)
DHEA	6,2±0,5	8,9±0,3	22,3±0,5	21,5±0,5	1,1±0,1	1,7±0,1	nd <sup>*</sup>	43,8±2
EA	3,4±0,3	3,0±0,1	4,8±0,3	4,8±0,3	0,46±0,03	0,49±0,03	36,5±3,6	24,5±0,7
16BrEA	0,93±0,03	0,53±0,01	-	-	-	-	12,4±1,2	-
16BrEA	0,60:	±0,08	0,086	±0,008	0,013	±0,03		
IC50								

Figura 5: Estrutura química dos esteroides inibidores da G6PD e tabela contendo as constantes de inibição contra a enzima humana (HsG6PD), de *Trypanosoma cruzi* (TcG6PD), de *Trypanosoma brucei* (TbG6PD) e dose letal mediana (LD<sub>50</sub>) para a forma epimastigota de *T. cruzi* e para forma *bloodstream* de *T. brucei*. DHEA: 5-androsten-3β-ol-17-ona, EA: 5α-androstan-3β-ol-17-ona, 16BrEA: 5α-androstan-16α-bromo-3β-ol-17-ona.\*-, não publicado, nd, não detectado no ensaio utilizado.

## 1.5 Objetivos:

O projeto de mestrado teve como objetivo geral mapear o sítio de ligação do DHEA na enzima G6PD por cristalografia de raios X. Um dos problemas do estudo deste complexo por cristalografia é o fato do DHEA ser um inibidor incompetitivo, ou seja, este só se liga à enzima complexada com seus substratos, sendo difícil obter uma solução homogênea na qual não ocorra catálise alguma. Além disso, o equilíbrio entre dímero e tetrâmero também contribui para a formação de uma solução heterogênea, dificultando a cristalização e/ou a obtenção de cristais homogêneos com ocupância dos esteroides.

Por estes motivos, o objetivo principal deste projeto foi a otimização do processo de cristalização, através do estudo da formação de um complexo da enzima formado com análogos do substrato que não sejam metabolizados, mas que ainda permitem a ligação do inibidor. Ou ainda, o uso de mutantes que diminuem a atividade catalítica da enzima sem alterar a ligação de seus substratos para a estabilização do complexo.

# Capítulo 2: Materiais e Métodos

## 2.1 Subclonagem da Hs∆G6PD e Tc∆G6PD

Foi a feita a amplificação das construções Hs $\Delta$ G6PD e Tc $\Delta$ G6PD a partir dos moldes HsG6PD-pET30b e TcG6PD-pET28, cedidos gentilmente por Dr. Paul Engel (*University College of Dublin*, Irlanda) e Dr. Mariana Igoillo-Esteve (*Christian De Duve Institute of Cellular Pathology*, Bruxelas, Bélgica) respectivamente, com os seguintes *primers*:

Hs∆G6PDfw-EcoRI

5'-AG<mark>GAATTC</mark>TCGGATACACACATATTCATCATCATG

Hs∆G6PDrv-Xhol

5'- AG<mark>CTCGAG</mark>TTAGTTCACCCACTTGTAGGTGCC

Tc∆G6PDfw-BamHI

5'-AG<mark>GGATCC</mark>GCCCTCACCATCGTGGTACT

Tc∆G6PDrv-Xhol

#### 5'- AG<mark>CTCGAG</mark>TTACGCGTTTGAGCACTTATACGC

Em amarelo estão destacados os sítios das enzimas de restrição (EcoRI, XhoI, BamHI e XhoI, respectivamente) que foram usadas para ligação em vetores de expressão.

A reação de PCR utilizada era constituída de:

5 μL Tampão de PCR (*Invitrogen*) 10X

1 µL dNTP (desoxirribonucleotídeos Fosfatados )(10 mM)

2 μL MgSO<sub>4</sub> (50 mM)

 $1 \,\mu\text{L}$  Primer fw ( $1 \,0\mu\text{M}$ )

 $1 \ \mu L Primer rv (10 \ \mu M)$ 

 $0,5 \ \mu L DNA Molde (150 \ ng/\mu L)$ 

0,5  $\mu$ L Taq HiFi (*Invitrogen*) (250 U/  $\mu$ L)

39 µL água

A amplificação foi feita usando um programa de PCR com 31 repetições no termociclador *Thermocycler GeneAmp PCR System 9700 (Applied Biosystems)*, sendo a temperatura de abertura das fitas de 94°C, hibridização à 62°C e elongamento à 68°C. O ciclo está representado abaixo:

ΕΤΑΡΑ	Desnaturação	Desnaturação	Anelamento	Extensão	Extensão	Finalização
Nº de ciclos	1x	31x			1x	1x
Temperatura (°C)	94	94	62	68	68	4
Tempo (min)	5	0.5	2	0.5	7	-

Tabela 1: ciclo de PCR para amplificação Hs $\Delta$ G6PD e Tc $\Delta$ G6PD.

O produto do PCR foi aplicado em eletroforese em gel de agarose 1% com 0,5 µg/ml Brometo de Etídeo em TAE e a banda amplificada foi purificada com kit comercial (*Wizard SV Gel and PCR Clean-Up System – Promega*). Os insertos e os vetores de expressão (pET-SUMO e pET-TEV, ambos vetores engenheirados a partir do pET28a *-Novagen-* para codificar a SUMO no Nterminal; e possuir o sítio de clivagem da TEV protease, Figura 6) foram digeridos com as enzimas de restrição apropriadas. As reações de digestão usadas foram:

Reação de digestão EcoRI/XhoI

Reação de digestãoBamHI/Xhol

DNA (100 ng/μL)	30 µL	DNA (100 ng/μL)	30 µL
Tampão 4 (10X)	5 μL	Tampão 3 (10X)	5 μL
EcoRI(20 U/μL)	1 μL	BSA (10X)	5 μL
Xhol (20 U/μL)	1 μL	EcoRI(20 U/μL)	1 μL
Água	13 µL	Xhol (20 U/μL)	1 μL
		Água	8 μL

#### Mapa do vetor pET-TEV

His tag ccaccaccaccacctgagatccggctgctaacaaagcccgaa ggtggtggtggtggtgactctaggccgacgattgtttcgggctt ProProProProProLeuArgSerGlyCysEnd

#### Mapa do pET-SUMO



Figura 6: Mapa dos vetores de expressão utilizados.

As enzimas e os tampões utilizados foram da *New England BioLabs (NEB)*. As reações foram mantidas à 37°C por 5 h. As amostras foram analisadas em gel de agarose (agarose 1% com 0,5  $\mu$ g/ml Brometo de Etídeo em TAE) e as bandas digeridas foram cortadas e purificadas como descrito anteriormente.

Os plasmídeos e os insertos digeridos e purificados foram então ligados, considerando uma proporção de 1 de plasmídeo para 3 de inserto (em concentração molar). A reação de ligação usada foi:

Vetor digerido (60 ng/µL)	1,6 µL
Inserto (100 ng/μL)	0,8 μL
Tampão (5X)	2,0 μL
T4 DNA ligase (400 U/μL)	1,0 μL
Água	4,6 μL

A enzima e o tampão foram da *NEB*. A reação foi conduzida à 25°C por 3 horas. O produto da ligação foi transformado em TOP10 competente por choque térmico, seguindo as seguintes etapas: incubação com o DNA por 30 minutos a 4°C, 2 minutos à 42°C, 2 minutos à 4°C, adição de 0,5 mL de meio LB com incubação por uma hora à 37°C sob agitação, centrifugação por 3 minutos à 5000 rpm com posterior retirada de 300 μL de sobrenadante, seguido de plaqueamento do *pellet* ressuspendido em placas LB ágar com 50 μg/mL de canamicina. As placas foram mantidas por 18 horas em estufa a 37°C.

Escolheu-se 6 colônias diferentes para extração do plasmídeo e confirmação da correta ligação. Para isso, cada colônia foi inoculada em 5 mL de meio LB com 50 µg/mL de canamicina e mantida sob agitação à 37°C por 18 horas. Em seguida, centrifugou-se o meio por 10 minutos à 5000 rpm e extraiu-se o plasmídeo com kit de miniprep (*Promega*). Para confirmação da inserção do inserto, foi feito um teste de PCR com os plasmídeos extraídos, usando a mesma reação e ciclo descrito na amplificação do gene. Dois plasmídeos positivos foram seqüenciados e transformados em BL21(DE3), seguindo o mesmo protocolo descrito anteriormente.

#### 2.2 Teste de expressão

Inoculou-se 100  $\mu$ L de pré-inóculo (crescido por 18 horas a 37°C) em 10 mL de LB com 50  $\mu$ g/mL de canamicina. O ensaio foi feito a 37°C. Quando a densidade ótica a 600 nm atingiu valores entre 0,8 e 1,0 foi adicionado à cultura 1 mM final de IPTG (Isopropil  $\beta$ -D-1-thiogalactopiranoside), coletando uma amostra de 1 mL antes da indução e depois de 4 horas de indução, centrifugando-as à 5000rpm.

Os *pellets* foram ressuspensos em 100 µL de tampão (100 mM Tris-HCl pH 8) e analisados por SDS-PAGE em gel de poliacrilamida 15%.

## 2.3 Expressão em meio de auto-indução

Meio litro de meio de auto-indução ZYM5052<sup>50</sup> com 50 µg/mL de canamicina foi inoculado com 5mL de pré-inóculo de cultura e mantido sob agitação (200 rpm) à 20°C por 48 horas em erlenmeyer de 2 litros. Em diferentes tempos após o inoculo se coletou uma amostra de 1 mL da cultura, além da medição da densidade ótica(D.O.) à 600 nm, a fim de analisar o crescimento e o conteúdo proteico da cultura ao longo do tempo e estabelecer um bom período de incubação. As amostras assim que coletadas foram centrifugadas a 5000 rpm e os *pellets* foram congelados. O volume final foi centrifugado a 6000 rpm por 15 minutos e o *pellet* foi congelado à -20°C.

As amostras foram ressuspensas em tampão 100 mM Tris-HCl pH 8,0 na proporção 10  $\mu$ L tampão /0,1 D.O. e submetidas a SDS PAGE em gel 15%.

## 2.4 Purificação Hs∆G6PD

#### 2.4.a Lise celular

Ressuspendeu-se o *pellet* em tampão de lise (100 mM Tris-HCl pH 8, 300 mM NaCl, 10 mM imidazol) na proporção de 5mL de tampão/grama de *pellet*, adicionando 400µg lisozima/mL de tampão, deixando em repouso por meia hora em gelo. Em seguida, sonicou-se a 40% de amplitude com ciclos de 10 segundos de sonicação com 10 segundos de repouso por 8 minutos, em gelo. A amostra foi então centrifugada a 16000 rpm a 4°C por 30 minutos. O sobrenadante foi separado e mantido a 4°C.

#### 2.4.b Afinidade em resina de níquel e clivagem na resina por sumo protease

Equilibrou-se a resina (*Ni-NTA Superflow-Qiagen*) (1 mL de resina para cada litro de meio de cultura) com 10 vezes em volume de tampão de lise. Em seguida, passou-se o extrato celular, seguido de uma lavagem de 50 volumes de coluna com tampão de lise. A resina foi ressuspensa em 3 volumes de coluna de tampão e adicionou-se ULP-1 (protease específica para SUMO, produzida no laboratório) na proporção de 100 µg de ULP-1 /mL de resina. A resina foi mantida sob agitação por 5 horas e amostras de 20µL foram coletadas em tempos diferentes e já preparadas para SDS-PAGE, para avaliar o melhor tempo de clivagem.

Em seguida, empacotou-se a resina novamente, coletando o *flowthrough* e fazendo 3 lavagens com tampão de lise, cada lavagem de 5 volumes de coluna. Enfim, foi feita a eluição com 10 volumes de coluna com tampão de eluição (100 mM Tris-HCl pH 8, 300 mM imidazol).

#### 2.4.c Troca iônica

Juntaram-se as amostras contendo a Hs∆G6PD (*flowthrough* e lavagens após clivagem na afinidade) e para retirada do sal para posterior aplicação em troca iônica, concentrou-se a amostra (*Amicon 10kDa - Milipore*), seguido de diluição de 10 vezes em tampão A (100 mM Tris-HCl pH 8), seguida de nova concentração e diluição, tendo ao fim uma amostra em 100 mM Tris-HCl pH 8, 3 mM NaCl e 0,1 mM imidazol.

A amostra com baixa força iônica foi então aplicada em uma coluna de troca aniônica HiTrap Q FF (*GE Healthcare*) em sistema cromatográfico (*ÄKTA FPLC – GE Healthcare*) previamente equilibrada com tampão A. Após a aplicação da amostra, fez-se a lavagem com 5 volumes de coluna com tampão A , iniciando a eluição com gradiente linear, alcançando 100% de tampão B (100 mM Tris-HCl pH 8, 500 mM NaCl) em 10 volumes de coluna. A corrida foi toda feita a 0,5 mL/min, com coleta de frações de 2 mL. As amostras correspondentes aos picos de absorbância em 280 nm foram recolhidas e 20 μL de cada amostra foram usadas para análise em SDS-PAGE.

#### 2.4.d Cromatografia de exclusão molecular

Este passo da purificação foi feito se utilizando a coluna *Superdex* 200 16/60 (*GE Healthcare*), com a aplicação de 0,5 ml a 1 ml de amostra com a coluna previamente equilibrada em tampão de GF (50 mM Tris-HCl pH 8 , 150 mM NaCl), sendo a corrida conduzida com fluxo de 1 mL/min, com frações de 2 mL.

# 2.5 Estudo de estado oligomérico por cromatografia analítica de exclusão molecular

Este estudo foi feito utilizando a coluna analítica *Superdex* 200 10/30 (*GE Healthcare*). Primeiramente, foi feito uma calibração da coluna utilizando proteínas com peso molecular conhecido para se ter um curva padrão relacionando o coeficiente de partição com o peso molecular. Os padrões utilizados são do kit MGW200 (*Sigma-Aldrich*), sendo:  $\beta$ -amilase (200kDa), álcool desidrogenase (150kDa), albumina bovina (66kDa), anidrase carbônica (29kDa) e citocromo C (12,4 kDa). O volume morto (V<sub>0</sub>) da coluna foi obtido utilizando *blue dextran*. Os padrões foram

16

dissolvidos em 50 mM Tris-HCl pH 8,0 e 150 mM NaCl, com injeção de 100  $\mu$ L, fluxo de 0,5 mL/min. Para cálculo do coeficiente de partição, divide-se o volume de eluição (V<sub>e</sub>) de cada marcador pelo volume morto.

Tendo a curva de calibração, aplicou-se a G6PD nas mesmas condições de tampão, injeção e fluxo, sendo possível calcular o peso molecular de cada pico da eluição, tendo seu estado oligomérico.

## 2.6 Cinética enzimática

Para caracterização enzimática, acompanhou-se a formação de NADPH (nicotinamida adenina dinucleótido fosfato reduzida) por sua fluorescência em 450 nm, não presente nos substratos (NADP<sup>+</sup> e G6P), esteróides e F6P. Os experimentos foram realizados no leitor de placa *EnVision (PerkinElmer)* em placa de 96 poços, com filtro de excitação de 340 nm e filtro de emissão de 450nm. Os experimentos foram realizados a 25°C, em tampão de 50 mM Tris-HCl pH 8, 150 mM NaCl , com volume de reação de 100  $\mu$ L. A reação foi disparada com a adição da enzima, tendo esta uma concentração final entre 5 nM e 20 nM. Para determinação do K<sub>m</sub> aparente da G6P, fezse uma diluição seriada de 1:1 da G6P começando com 2 mM (concentração final) e com uma concentração constante de NADP<sup>+</sup> de 400  $\mu$ M. Para determinação do K<sub>m</sub> aparente do NADP<sup>+</sup>, fezse uma diluição seriada de 1:1 de NADP<sup>+</sup> começando com 400  $\mu$ M e com uma concentração constante de 1:1 de NADP<sup>+</sup> começando com 400  $\mu$ M e com uma concentração constante de 1:1 de NADP<sup>+</sup> começando com 400  $\mu$ M e com uma concentração constante de 1:1 de NADP<sup>+</sup> começando com 400  $\mu$ M e com uma concentração constante de 1:1 de NADP<sup>+</sup> começando com 400  $\mu$ M e com uma concentração constante de G6P de 2 mM. Os experimentos foram feitos em triplicata.

A medida feita foi da fluorescência de NADPH com o tempo. Assim, a velocidade inicial V<sub>i</sub> pode ser obtida derivando a fluorescência em relação ao tempo, em um intervalo em que a velocidade seja constante. Esta etapa é realizada com o próprio programa do leitor de placa. Em seguida, para obtenção dos valores de K<sub>m</sub> e V<sub>max</sub>, foi feito o ajuste das curvas de velocidade inicial em relação à concentração de substrato com a equação de Michaelis-Menten:

$$V_{i} = \frac{Vmax.[S]}{Km+[S]}$$
eq. 1

Sendo V<sub>i</sub> velocidade inicial, Vmax velocidade máxima, Km, constante de Michaelis e [S], concentração do substrato.

Para determinação do IC<sub>50</sub>, foi feita uma diluição do inibidor de 1:1,5, começando com 100  $\mu$ M, mantendo a concentração dos substratos fixa (2 mM G6P e 200  $\mu$ M NADP<sup>+</sup>). A concentração

final de DMSO foi de 2%. O valor de  $IC_{50}$  foi obtido com o ajuste de uma sigmoide ao gráfico de porcentagem de atividade pelo logaritmo da concentração de inibidor.

## 2.7 Ensaio de termofluorescência

Neste experimento é utilizada uma sonda fluorescente (*Sypro Orange*) que possui baixa fluorescência em solução aquosa, porém, quando interage com grupos hidrofóbicos, apresenta alta fluorescência. A amostra da proteína com a sonda é aquecida, desnaturando a proteína e expondo os resíduos hidrofóbicos, que passam a interagir com a sonda, aumentando significativamente a fluorescência. Assim, é possível observar a temperatura na qual ocorre a transição entre enovelada/desenovelada. A temperatura na qual 50% da proteína se encontra desenovelada é chamada de  $T_M$ . Quando uma pequena molécula se liga à proteína, as novas interações realizadas entre o ligante e o sítio estabilizam a estrutura, aumentando o valor de  $T_M$ . Assim, esta técnica pode ser utilizada para avaliação da formação de complexos proteína/ligante, que deverão apresentar um  $T_M$  maior que a proteína sozinha, além de estudar mutantes e/ou tampões que estabilizem a enzima, aumentando as chances de cristalização.<sup>31</sup>

Os ensaios foram realizados utilizando entre 5  $\mu$ M e 20  $\mu$ M de proteína, incubadas com diferentes concentrações de *Sypro Orange (Sigma-Aldrich)* (entre 5X e 20X, considerando que o estoque comprado é de 5000X). Os ensaios foram realizados em placa de 96 poços em um equipamento de PCR quantitativo (*7500 PCR Real-Time System – Applied Biosystems*), começando com 10°C chegando à 96°C com passo de 1°C/min, em volume de 25  $\mu$ L.

Os dados obtidos foram tratados utilizando o programa *GraphPad Prism 5.* O valor de  $T_M$  foi obtido com o ajuste de uma sigmoide aos dados normalizados de fluorescência.

## 2.8 Mutação sítio dirigida

Para avaliação da importância de diferentes resíduos no estado oligomérico da proteína ou na sua atividade, foram feitas mutações pontuais. Para isso, foi utilizado o kit de mutagênese *QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies)*, seguindo as instruções do manual. O molde utilizado foi a construção Hs∆G6PD em pET-SUMO e os oligonucleotídeos utilizados foram: Tabela 2 : Oligonucleotídeos utilizados para mutação pontual. Em amarelo, as bases mutadas.

Mutação	Oligo forward	Oligo reverse
D200N	CAGATCTACCGCATC <mark>A</mark> ACCACTACCTGGGC	GCCCAGGTAGTGGT <mark>T</mark> GATGCGGTAGATCTG
A277C	GGCCATGGAGAAGCCC <mark>TG</mark> CTCCACCAACTCAG	CTGAGTTGGTGGAG <mark>CA</mark> GGGCTTCTCCATGGCC
E347A	CTATGTGGAGAATG <mark>C</mark> GAGGTGGGATGGGG	CCCCATCCCACCTC <mark>G</mark> CATTCTCCACATAG

#### 2.9 Ensaios de cristalização

Foram feitos *screenings* iniciais com kits comerciais de cristalização (*Crystal Screen and Crystal Screen 2* e *SaltRx* da *Hampton Research; Wizard I e II, Precipitant Synergy* da *Emerald BioSystems* ; e *PACT* e *JCSG+* da *Nextal/Qiagen*). Os ensaios foram feitos em placas de 96 poços, com método de difusão de vapor com gota sentada, montadas pelo robô de pipetagem *HoneyBee 963* (*Genomic Solutions* ), com gotas de 1  $\mu$ L de proteína e 1  $\mu$ L solução de poço, com 80  $\mu$ L de solução no poço. As placas foram mantidas a 18°C em robô de fotodocumentação *Rock Imager UV* (*Formulatrix*), sendo fotodocumentadas em tempos específicos com luz visível e UV.

As condições que apresentaram cristais foram submetidas a refinamento, variando diferentes parâmetros, como concentração da proteína, precipitante, proporção proteína/poço e pH. Além disso, fez-se a reprodução da condição já publicada para Hs $\Delta$ G6PD<sup>24</sup>, que é de 100 mM Tris-HCl pH 8,5, 200 mM MgCl<sub>2</sub>, 12% PEG 4000, 5% glicerol, com 5 mg/ml de proteína em 100 mM Tris-HCl pH 7,5, 75 mM NADP<sup>+</sup>, 60 mM G6P. Esta condição foi refinada também.

Nas etapas de refinamento, utilizou-se o método de gota pendurada em placa de 24 poços, montadas à mão. As gotas eram formadas entre 1  $\mu$ L e 3  $\mu$ L de proteína para 1  $\mu$ L e 3  $\mu$ L de solução de poço, sendo o poço preenchido com 500  $\mu$ L.

Para tentar cristalizar a proteína com os diferentes ligantes ou introduzi-los no cristal já formado, foram utilizadas a co-cristalização, *seeding* e *soaking*, descritos a seguir:

#### 1. Co-cristalização

Neste ensaio, os ligantes em estudo foram incubados com a proteína e submetidos aos ensaios de cristalização. Quando feita com esteroides, a concentração final de DMSO foi de 2% a 5%, sendo a concentração do inibidor por volta 100 μM.

19
#### 2. Soaking

Cristais formados foram transferidos para outra gota contendo diferentes combinações dos substratos e/ou análogos com esteroide em solução idêntica à solução do poço. Os cristais foram coletados após diferentes tempos de incubação, variando entre 30 segundos a 2 horas.

#### 3. Seeding

Neste método, pegou-se uma gota com cristais e acrescentou 50µL da solução de poço e uma conta de plástico (*Seed Bead- Hampton*) para triturar os cristais com a ajuda de um vórtex. Em seguida, molhou-se um fio de rabo de cavalo na amostra e passou-se em gotas previamente equilibradas pelo método de gota dependurada. As gotas que recebem as sementes foram feitas variando a concentração do precipitante entre uma concentração que se tem cristais e outra em que a gota fica limpa. Para saber este intervalo, sempre que se produzia um novo lote da proteína, montaram-se algumas condições variando entre 10-17% PEG4000. Assim, verificava-se em qual concentração de precipitante começa a aparecer cristais. Tipicamente, fizeram-se as gotas para *seeding* com 9,10, 11 e 12% de PEG4000.

## 2.10 Coleta de dados de difração de Raio-X e processamento dos dados

A coleta de dados foi realizada na linha W01B-MX2 do LNLS.<sup>32,33</sup> O comprimento de onda utilizado foi de 1,46 Å e o padrão de difração foi medido com um detector de CCD (*Rayonix MarMosaic 225 CCD*). A distância entre o cristal e o detector foi ajustada conforme a resolução de cada cristal, ficando em torno de 200 mm.

As reflexões das imagens foram indexadas e integradas utilizando o programa *iMOSFLM*<sup>34</sup> e escalonadas com *SCALA*<sup>35</sup>. Foi feita a substituição molecular com *MOLREP*<sup>36</sup> utilizando o monômero da estrutura humana (PDB: 2BH9), retirando os ligantes e as águas. O sucesso da substituição foi avaliado pelo contraste obtido (razão entre o *score* da melhor solução pelo *score* médio das soluções, sendo esperado um contraste acima de 2,5 para a solução), pelo

empacotamento cristalino, observando se não havia choques estereoquímicos e se havia a formação do dímero ou tetrâmero já descritos. O refinamento no espaço recíproco foi feito utilizando o programa *REFMAC*5<sup>37</sup> e o refinamento manual foi feito com o *WinCoot*.<sup>38</sup>

# Capitulo 3: Subclonagem, produção heteróloga e caracterização cinética da G6PD

## 3.1 Subclonagem da Hs∆G6PD e Tc∆G6PD

As estruturas da GGPD humana selvagem só foram obtidas quando se fez uma deleção no N-terminal de 25 aminoácidos<sup>24,39</sup>, sendo essa excisão guiada com o alinhamento com a sequência da GGPD de *L. mesenteroides*, que foi a primeira estrutura resolvida. Considerando que no começo do projeto a única enzima com condição de cristalização publicada e que fosse inibida por esteroides era a humana, resolveu-se trabalhar com ela também, já que uma vez encontrado o sítio, este poderia ser estudado nos parasitas por homologia. Quanto à enzima de *T. cruzi*, foi adotada a mesma estratégia de fazer uma deleção baseada no alinhamento da sequência com as estruturas resolvidas (LmG6PD - PDB: 1QKI, HsG6PD – PDB: 2BH9) para tentar melhorar o processo de cristalização. Assim, decidiu-se clonar a G6PD tanto de *T. cruzi* e *H. sapiens* com a deleção no N-terminal, obtendo as construções Hs $\Delta$ G6PD (resíduo 29 até 511) e Tc $\Delta$ G6PD (resíduo 68 até 555). Lembrando que até então, não se tinha publicada a estrutura do parasita, mas em setembro de 2011 foi publicada<sup>25</sup> uma condição de cristalização para a enzima de *T. cruzi*, sendo de uma construção sem os primeiros 37 aminoácidos.

Fez-se o PCR para amplificação das construções de Hs $\Delta$ G6PD e Tc $\Delta$ G6PD. Abaixo se pode ver o gel de agarose da amplificação dos genes da Hs $\Delta$ G6PD (1465 pares de base) e Tc $\Delta$ G6PD (1460 pares de base). Observa-se na Figura 7 que houve amplificação em grande quantidade, que apesar de haver uma banda inespecífica, esta se encontra bem separada da banda de interesse (destacada na caixa branca), permitindo a separação pelo corte da banda.



Figura 7: Gel de agarose 1% do produto de PCR. 1- marcador (Gene Ruler 1Kb DNA Ladder – SM0313 – *Fermentas*); 2- Hs $\Delta$ G6PD ; 3 - Tc $\Delta$ G6PD.

Após a purificação das bandas e digestão com as enzimas de restrição, pode-se ver o gel da digestão dos insertos e do pET-SUMO (Figura 8). A digestão do pET-TEV não é mostrada.



Figura 8: Gel de agarose 1% da digestão com enzimas de restrição. A1 e B1: marcador (Gene Ruler 1Kb DNA Ladder – SM0313 – *Fermentas*); A2: pET-SUMO digerida com BamHI e XhoI; A3: pET-SUMO digerida com EcoRI e XhoI; B2: Tc∆G6PD digerida com BamHI e XhoI; B3: Hs∆G6PD digerida com EcoRI e XhoI.

Primeiramente, a Tc $\Delta$ G6PD foi ligada somente em pET-SUMO, mas devido a dificuldade em clivar a SUMO durante a purificação, como será apresentado posteriormente, esta foi clonada em pET-TEV. Já a Hs $\Delta$ G6PD foi clonada somente em pET-SUMO.

Depois de purificar os insertos e vetores digeridos, fez-se a reação de ligação seguida de transformação em células competentes TOP10 em meio com canamicina, para selecionar aquelas bactérias que possuem o plasmídeo circularizado. Escolheu-se 6 colônias diferentes para cada construção, fazendo extração de plasmídeo de 5 mL de cultura de cada colônia. Fez-se teste de PCR para comprovar a ligação do inserto. Um plasmídeo de cada clone positivo foi seqüenciado e analisado pelo programa BioEdit е 0 seqüenciamento feito no BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) contra as sequências da HsG6PD (gi:31542) e TcG6PD (gi:71666195), sendo observada a correta clonagem, sem mutações. Os plasmídeos então foram transformados em BL21(DE3) para prosseguir com a expressão e purificação.

### 3.2 Teste de expressão em meio auto-indutor

Fez-se um teste inicial de expressão a 37°C com indução com IPTG como descrito em materiais e métodos. Neste teste foi possível observar a expressão em 4 horas, podendo seguir os experimentos com a construção e cepa escolhidas. Em seguida, foi feita a caracterização da expressão em meio auto-indutor ZYM-5052. Na Figura 9 se pode ver o SDS-PAGE das amostras coletadas em diferentes tempos em meio de auto-indução junto com a curva de crescimento da

Hs $\Delta$ G6PD. Pode-se observar que com 30 horas a cultura já está na fase estacionária de crescimento e a expressão também já não aumenta com o tempo. Assim, ficou estabelecido o tempo de 30 horas de cultivo a 20°C sob agitação de 200 rpm. A Tc $\Delta$ G6PD tem o mesmo comportamento, tanto na construção em pET-SUMO como em pET-TEV.



Figura 9: A: Curva de crescimento de BL21 com plasmídeo da Hs $\Delta$ G6PD em meio ZYM5052; B: SDS-PAGE em gel 15% de amostras dos diferentes tempos de crescimento da cultura; M – marcador de peso molecular (*Page ruler protein ladder SM0661-Fermentas*).

### 3.3 Purificação

#### 3.3.a Afinidade com resina de Níquel

Inicialmente, fez-se a afinidade sem clivagem na coluna, tendo a proteína eluída com 300 mM de imidazol, seguindo da clivagem com a ULP-1 (SUMO-protease). O observado é que a proteína humana clivava muito bem com 2 horas de incubação a 25°C, mas que a de *T. cruzi* não clivava totalmente, como mostrado na figura 10-A. Assim, testaram-se diferentes condições de clivagem, como diferentes proporções de proteína/protease, tampões, temperatura e tempo de incubação. Não se conseguiu melhorar a clivagem da SUMO para a proteína do parasita. Assim, concluiu-se que a conformação da SUMO fusionada com a Tc $\Delta$ G6PD dificultava o acesso ao sítio de clivagem pela protease, não permitindo a total clivagem da SUMO. Com isto, decidiu-se clonar a Tc $\Delta$ G6PD em pET-TEV, que adiciona somente a cauda de histidina, para poder prosseguir nos passos de purificação. Assim, fez-se a afinidade da Tc $\Delta$ G6PD, como pode ser visto na Figura 10-B. A proteína obtida não possui muitos contaminantes, mas se apresenta instável, precipitando em algumas horas após a eluição. Tentou-se diferentes condições para estabilizar a proteína, como utilizar glicerol nos tampões de purificação e realizar todo o processo a 4°C, mas não se obteve nenhuma melhora significativa. Comparando os procedimentos para obtenção da enzima do

parasita com os utilizados por Ortíz *et al*<sup>25</sup>, publicada posteriormente com a condição de cristalização, observa-se que a principal diferença é a construção, sendo a utilizada neste projeto começando no aminoácido 68, e o publicado, no 38. Analisando a estrutura (PDB: 4EM5), os aminoácidos do N-terminal presentes começam no resíduo 52, sendo que entre V52 e V60 não apresenta nenhuma estrutura secundária definida, e entre os resíduos 61-67 há uma pequena  $\alpha$  hélice. Assim, provavelmente os resíduos de 52-67 dão estabilidade à enzima, explicando o insucesso da construção utilizada.



Figura 10: A: Gel SDS-PAGE 10% da clivagem da SUMO com ULP-1 à 25°C por 4 horas da Tc $\Delta$ G6PD e Hs $\Delta$ G6PD eluídas em 300mM imidazol; 1 e 3 amostras antes da adição da protease, 2 e 4 amostras após clivagem. A banda da SUMO não aparece porque o gel é 10%. B: Gel SDS-PAGE 15% de afinidade da Tc $\Delta$ G6PD com cauda de histidina 1- lavagem com 100mM Tris-HCl pH 8, 300mM NaCl , 10mM imidazol , 2 e 3 eluições.

Quanto à Hs∆G6PD-SUMO, observou-se uma boa clivagem, mas também há uma pequena quantidade que permanece sem clivar. Tentativas posteriores para separar a proteína clivada da não clivada por troca iônica ou exclusão molecular falharam, tendo a proteína não clivada eluída sempre juntamente com a clivada. Isso provavelmente acontece devido à formação de dímeros e tetrâmeros, acontecendo da proteína com SUMO estar oligomerizando com a proteína livre. Para contornar esta situação, decidiu-se testar a clivagem da SUMO enquanto a proteína permanecia ligada à resina da afinidade. Obteve-se sucesso neste caso, pois a proteína clivada ficou em solução enquanto a proteína não clivada permaneceu ligada à resina. Uma pequena quantidade da proteína clivada permaneceu na eluição com imidazol, provavelmente oligomerizada com a não clivada, mas essa perda não é significativa, considerando a quantidade de proteína que se consegue purificar. Além disso, essa foi a única forma que se conseguiu separar a porção clivada da não clivada. Na Figura 11 é possível observar que apesar da clivagem da Hs $\Delta$ G6PD na coluna não ser total, consegue-se separar a porção não clivada, que sai na eluição. A estabilidade da Hs $\Delta$ G6PD obtida é bem melhor que a Tc $\Delta$ G6PD, sendo estável a 4°C por uma semana, sem precipitação a altas concentrações (a maior concentração já alcançado foi de 40 mg/mL, apresentando-se sem sinal de precipitação), permitindo fazer outros passos de purificação e estudos estruturais.



Figura 11: SDS-PAGE 15% da clivagem na coluna de afinidade da Hs∆G6PD. 1: amostra da resina ressuspensa após clivagem com ULP-1, 2- *flow-through*, 3, 4,5 e 6 lavagens, 7- eluição.

#### 3.3b Troca iônica da Hs∆G6PD

A amostra concentrada da afinidade foi desalinizada como descrito anteriormente e aplicada na troca aniônica. A proteína eluiu na faixa de 15mS/cm e 29mS/cm, tendo um pico com 22mS/cm (Figura 12-A).

#### 3.3c Gel filtração com Superdex 200 16/60

A proteína apresentou um pico de eluição em 62 mL e um pequeno pico em 69 mL. Os dois picos não são totalmente separados com a resolução da coluna. Correndo a mesma amostra na coluna analítica e utilizando a curva de calibração, pode-se observar que o primeiro pico é

referente a forma tetramérica e o segundo, a forma dimérica da enzima. Nos estudos de cristalografia, separou-se as frações dos centros dos picos.

Ao fim de todos os passos de purificação, conseguiu-se em média 3 mg de Hs∆G6PD/litro de cultura referente ao centro do pico de gel filtração e mais 2 mg das outras regiões do pico principal. Na Figura 12 é mostrado o cromatograma da troca iônica e da gel filtração, além de gel SDS das amostras após as diferentes etapas de purificação.



Figura 12: A: Cromatograma da troca iônica, B: Cromatograma da gel filtração em Superdex200 16/60 C: SDS-PAGE 15% das diferentes etapas de purificação da Hs $\Delta$ G6PD. 1: após clivagem na coluna de afinidade, 2: após troca iônica, 3: após gel filtração.

## 3.4. Avaliação da atividade da enzima purificada

Ensaios cinéticos foram usados para avaliar se a enzima purificada possuía atividade e se a deleção causou alguma mudança na inibição pelos esteroides.

Fez-se a curva de Michaelis-Menten para determinação do K<sub>m</sub> aparente da G6P e NADP<sup>+</sup>, obtendo valores condizentes com o reportado na literatura para esta construção, como mostrado na Tabela 3 e Figura 13. O valor de k<sub>cat</sub> obtido foi inferior ao da literatura, isto talvez seja causado pelo tampão utilizado no ensaio, que não é o mesmo (tampão do ensaio da literatura<sup>24</sup>: 100 mM Tris-HCl pH 8,0 ; 10 mM MgCl<sub>2</sub>, tampão utilizado: 50 mM Tris-HCl pH 8, 150 mM NaCl), o que pode influenciar na atividade da enzima.



Figura 13: Curvas de Michaelis-Menten para determinação de K<sub>m</sub> para os substratos.

Tabela 3: Constante de Michaelis-Menten (K<sub>M</sub>) para os substratos da G6PD.

Hs∆G6PD	Κ <sub>Μ</sub> <sup>G6P</sup> (μΜ)	K <sub>M</sub> <sup>NADP</sup> (μM)	k <sub>cat</sub> (1/s)
Purificada	58,3±0,7	7,5±0,8	70±10
Literatura <sup>24</sup>	42	6,45	208

Para avaliar a inibição da Hs $\Delta$ G6PD pelos esteroides, fez-se uma curva de dose resposta para o DHEA, EA e 16BrEA, determinando o IC<sub>50</sub>, que são mostrados abaixo na Figura 14.



Figura 14: Curvas dose resposta para determinação de  $IC_{50}$  dos esteroides para a construção Hs $\Delta$ G6PD. A direita, tabela comparando os valores obtidos com dados da literatura. \*( Gordon et al, 1995)

Os valores obtidos são condizentes com os da literatura, considerando que para um inibidor incompetitivo em condição de saturação dos substratos, o  $IC_{50 e}$  aproximadamente igual

ao valor de K<sub>i</sub>.<sup>40</sup> Isso mostra que a deleção no N-terminal não alterou a inibição pelo esteroides, podendo esta construção ser usada nos estudos.

## Capítulo 4: Formulação de complexo estável

O DHEA e seus análogos são inibidores incompetitivos em relação aos dois substratos da G6PD, ou seja, estes só se ligam ao complexo G6PD-G6P-NADP<sup>+</sup>, não tendo afinidade pela enzima livre. Assim, para conseguir cristalizar a enzima com estes inibidores, é necessário colocar os dois substratos, além do esteroide. Porém, como a inibição não é completa e o processo de cristalização pode durar dias, a reação provavelmente acontece durante o experimento, diminuindo as chances de co-cristalização, pois os substratos podem se esgotar, além da enzima estar sofrendo mudanças conformacionais, o que atrapalha no processo de cristalização.

Assim, adotou-se 2 estratégias para tentar contornar este problema:

- 1- Encontrar inibidores competitivos dos substratos que pudessem gerar um complexo ternário não reativo que ainda permitisse a ligação do esteroide. Os inibidores testados eram análogos de G6P ou o próprio produto da reação (NADPH)
- 2- Fazer uma mutação que diminuísse significativamente a atividade da enzima, sem alterar a sua afinidade pelos substratos.

### 4.1 Avaliação de inibidores competitivos

Buscando na literatura, encontraram-se duas hexoses fosfatadas que atuam como inibidores competitivos em relação a G6P, a glucosamina-6-fosfato (Gln6P) e frutose-6-fosfato (F6P).<sup>41,42</sup>(Figura 15) Para a enzima humana, tem-se dados sobre a Gln6P, que atua como inibidor competitivo, tendo um K<sub>i</sub> de 1,08 mM<sup>42</sup>. Não foram encontrados dados sobre a F6P para a enzima humana. Outro inibidor competitivo conhecido é o NADPH, tendo K<sub>i</sub> de 9  $\mu$ M<sup>42</sup>. Assim, decidiu-se testar estes 3 inibidores para a formulação do complexo.



Figura 15: compostos utilizados para formulação de complexo ternário. Análogos do substrato glicose 6-fosfato: glucosamina 6-fosfato e frutose 6-fosfato; NADP<sup>+</sup> e NADPH.

Primeiramente, avaliou-se se a F6P atuava como um substrato alternativo, realizando ensaio de atividade com 100  $\mu$ M NADP<sup>+</sup> e diferentes concentrações de F6P, variando de 50 mM até 100  $\mu$ M. Observou-se a formação de NADPH quando se utilizou F6P na faixa de 3 mM a 50 mM. Nos controles, não se observou fluorescência para a F6P nos comprimentos de onda utilizados no ensaio, nem a formação de NADPH no controle negativo (todos os reagentes com exceção da enzima), indicando que a enzima utiliza a F6P como um fraco substrato alternativo, o que não é interessante para a formação do complexo.

Quanto à Gln6P, não se observou a formação de NADPH quando esta é incubada com NADP<sup>+</sup> e G6PD, não atuando como substrato alternativo, como esperado. Foi feita a curva de dose resposta para avaliar se a inibição para a nossa construção é parecida com a da literatura. O valor obtido de  $IC_{50}$  foi de 1,02 mM (com o intervalo de 95% de confiança entre 0,56 mM e 1,88 mM), feito em ensaio com a concentração de G6P igual ao seu valor de K<sub>m</sub> e saturação de NADP<sup>+</sup> (Figura 16). Nesta condição, e considerando que se trata de um inibidor competitivo, o K<sub>i</sub> é metade do valor de  $IC_{50}$ ,<sup>40</sup> sendo então em torno de 0,5mM (entre 0,28mM e 0,94mM). Este valor esta condizente com o da literatura.

#### IC50 Glucosamina-6-fosfato



Figura 16: IC<sub>50</sub> para glucosamina 6-fosfato para Hs∆G6PD.

Para estudar o quanto estes inibidores estabilizam a proteína e para avaliar se há a formação do complexo com 16BrEA, utilizou-se a técnica de termofluorescência. Primeiramente foi feita uma padronização do ensaio, variando-se a concentração da proteína e da sonda. Avaliou-se a variação entre a fluorescência enquanto a proteína está enovelada com a fluorescência ao fim do desenovelamento. Além disso, observou-se um bom perfil de desnaturação, com um comportamento sigmoidal, e um bom discernimento entre o platô inicial e final. O ensaio padrão estabelecido foi com 10  $\mu$ M de G6PD e 10X de *Sypro Orange*. Na figura 17 se pode ver um exemplo dos dados medidos no qPCR e depois de tratados para o cálculo de T<sub>m</sub> (descrito no ítem 2.7), validando o ensaio para estudo da ligação de pequenas moléculas.



Figura 17: Exemplo de dados obtidos com o ensaio de termofluorêscencia estabelecido.

Este ensaio foi utilizado com as diferentes combinações entre os ligantes. Todas as condições continham 2% DMSO para poder comparar as amostras na ausência e presença do esteroide. Na tabela 4 está representado o  $\Delta T_m$  em relação à enzima apo para os diferentes ligantes.



Figura 18: Deslocamento no  $T_m$  causado pelos diferentes ligantes para a Hs $\Delta$ G6PD em ensaio de termofluorescência. As concentrações utilizadas para cada ligante foram: 80 $\mu$ M 16BrEA, 100 $\mu$ M NADP<sup>+</sup>, 100 $\mu$ M NADPH, 40mM G6P, 40mM Gln6P e 40mM F6P. Todas as condições possuem 2% DMSO.

Os inibidores competitivos utilizados apresentam um deslocamento menor que os substratos e não se pode afirmar que há a formação de complexo ternário (enzima:substrato:inibidor competitivo), nem de complexo quaternário, uma vez que nenhuma amostra dos complexos apresentam maior T<sub>m</sub> que os complexos de ordem menor (por exemplo, a amostra enzima+NADPH+G6P é mais estável que enzima+NADPH, porém não é mais estável que enzima+G6P, não podendo concluir se há a formação do complexo ternário). A F6P se apresenta como o pior candidato para a formulação do complexo, pois é a que menos estabiliza a proteína, comparando com Gln6P e NADPH.

Durante os ensaios de padronização foi observado que quando se utiliza 600  $\mu$ M G6P, concentração 10 vezes maior que o K<sub>m</sub>, o  $\Delta$ T<sub>m</sub> é de somente 0,7°C. Como visto na Figura 18, quando se utiliza 40 mM G6P o deslocamento sobe para 8,7 °C. Assim, talvez esteja ocorrendo a formação do complexo com os esteroides, porém a concentração utilizada de 16BrEA não seja o suficiente para dar um deslocamento significativo. Mas acima de 100  $\mu$ M em 2% DMSO há a precipitação do esteroide e não se pode aumentar muito a concentração de DMSO porque este aumenta o sinal de fluorescência do platô inicial, provavelmente por estar interagindo com a sonda fluorescente. Ensaios de cristalização com estes complexos foram feitos, tanto *screening* inicial, como com a condição publicada e otimizada, que será mostrada em capítulo posterior. Não se obteve nenhum indício de cristalização. Assim, partiu-se para a estratégia de gerar um mutante para diminuir a atividade da enzima.

### 4.2 Mutante D200N

O mecanismo de reação da G6PD aceito consiste em uma díade Asp-His, onde a histidina funciona como uma base para abstrair o próton do grupo hidroxil C1 da G6P, com a concomitante saída de um grupo hidreto do carbono C1 da G6P para o anel nicotina do NADP<sup>+</sup>. Assim, após a saída dos produtos, a histidina é reciclada doando o próton para uma molécula de água, restaurando a enzima. O ácido aspártico tem a função de aumentar a basicidade da histidina catalítica <sup>43,44</sup>.



Figura 19: esquema geral do mecanismo de reação da G6PD. Os índices dos resíduos são para a sequência da G6PD de *Leuconostoc mesenteroides.(* Figura de *Cosgrove et al., Biochemistry,* 2000)

Para elucidar o mecanismo de reação e o papel de cada resíduo, *Cosgrove et. al* utilizaramse de mutação sítio-dirigida e estruturas cristalográficas. Nestes estudos, realizou-se a mutação do ácido aspártico para asparagina (D177N, em *Leuconostoc mesenteroides*), obtendo uma enzima mutante com k<sub>cat</sub> duas ordens de grandeza menor que a selvagem, mas com pouco alteração nos valores de K<sub>m</sub> para os substratos G6P e NADP, indicando que o ácido é importante para a catálise (aumentando a basicidade da histidina 240), mas que este não altera a afinidade do sítio por G6P<sup>43</sup>. Com este mutante, conseguiram resolver uma estrutura cristalográfica onde os cristais cresceram na presença de G6P e NADPH, observando uma boa densidade para a G6P e densidade parcial do NADPH, tendo densidade para a adenina e os fosfatos do co-fator. Assim, conseguiram com este mutante um complexo ternário G6PD-G6P-NADPH.<sup>44</sup>

Deste modo, decidiu-se fazer este mutante na enzima humana com o intuito de formar o complexo ternário estável capaz de ligar os esteroides, aumentando a chance de cristalização.

A mutação D200N na enzima humana foi realizada por mutação sítio-dirigida, como descrita em materiais e métodos. A enzima foi expressa e purificada do mesmo modo que a selvagem, apresentando o mesmo perfil de expressão e purificação.

Foi feita a caracterização cinética do mutante, obtendo uma diminuição do  $k_{cat}$  de 30 vezes, um aumento de 5 vezes no  $K_m^{G6P}$  e um aumento de 3 vezes no  $K_m^{NADP}$  em relação a selvagem, como pode ser visto na tabela 4. O aumento nas constantes de Michaelis é pequeno comparado à diminuição da atividade da enzima, o que era esperado para este mutante.

Enzima	K <sub>m</sub> <sup>G6P</sup> (μM)	K <sub>m</sub> <sup>NADP</sup> (μM)	k <sub>cat</sub> (s <sup>-1</sup> )
Hs∆G6PD	58,3±0,7	7,5±0,8	70±10
Hs∆G6PD D200N	283±9	22±2	2,3±0,5

Tabela 4: Parâmetros cinéticos para o mutante Hs∆G6PD D200N.

Para avaliar a inibição deste mutante pelos esteroides, mediu-se a atividade da enzima na presença das maiores concentrações de solubilidade destes em 2%DMSO, sendo de 100  $\mu$ M para o DHEA e 80  $\mu$ M 16BrEA. Como pode ser visto na Figura 20, não se observou inibição para o DHEA e EA nestas concentrações. Já 80  $\mu$ M de 16BrEA apresenta uma inibição em torno de 60%, uma notável diferença em relação à nativa, que apresenta uma inibição >96% nesta concentração (em mesmas condições de ensaio). No caso da nativa, 1 $\mu$ M de 16BrEA apresenta 63±1% de inibição ,

podendo estimar assim que o mutante é 80 vezes menos inibido por 16BrEA. Não foi feita a curva de dose resposta, pois teria que ter concentrações mais altas do inibidor para ajustar a curva.



Inibição pelos esteroides



O fato do mutante continuar sendo inibido por 16BrEA, mesmo que com menor potência, e ter menor atividade em relação a enzima ativa, torna-o interessante para a formação do complexo quaternário, sendo um promissor complexo para os ensaios de cristalização.

Mais interessante ainda é o fato que houve uma grande diminuição do poder inibitório destes esteroides, indicando que o aspartato 200 não é tão somente importante para a catálise, como também para a ligação destes inibidores. Até então não se foi mostrado nenhum indício de algum resíduo que pudesse influenciar na inibição de esteroides análogos de DHEA, sendo este um primeiro dado que pode levar a uma hipótese sobre o provável sítio de ligação destes inibidores à HsG6PD.

Olhando a região próxima do D200, é possível observar uma cavidade no interior da proteína com um pequeno acesso à região do solvente. Esta cavidade possui resíduos hidrofóbicos e uma dimensão compatível com a de um esteroide (Figura 21). O posicionamento do DHEA foi feito sem a ajuda de métodos de docagem, sendo que este poderia estar em outra conformação e orientação, somente para mostrar que as dimensões da cavidade são compatíveis com a do esteroide. Além disto, a estrutura mostrada é referente à enzima humana ligada a G6P (PDB:2BHL), sendo que este bolsão pode ainda sofrer mudanças conformacionais com a ligação do esteroide. Analisando o alinhamento de sequências de G6PD provenientes de diferentes organismos que são ou não inibidas por DHEA, observa-se que há uma conservação da maioria dos

resíduos pertencentes à cavidade, indicados com os círculos amarelos na Figura 22. Porém, há resíduos não conservados entre as espécies que poderia explicar a não inibição, como por exemplo, a T265 e C269. Esta é só uma hipótese que deverá ser testada com outras mutações nesta cavidade.

Avaliando outras regiões do alinhamento, não se observou nenhum resíduo nitidamente conservado entre as G6PDs inibidas e não conservado entre inibidas/não inibidas, o que poderia explicar a seletividade na inibição pelos esteroides. Para esta análise, avaliaram-se os resíduos que estavam a 12 Å do NADP<sup>+</sup> catalítico e da G6P. Talvez a melhor forma de fazer este estudo seria, além de comparar as sequências, comparar o modelo da estrutura da G6PD de *Leishmania mexicana* com a estrutura de *T. cruzi*, pois se deve levar em consideração também a conformação dos resíduos na estrutura, não somente a identidade em si.



Figura 21: A: cavidade próxima ao D200. B: DHEA em uma posição dentro da cavidade, mostrando que a dimensão é compatível com o tamanho do esteroide. C: Resíduos hidrofóbicos que compõem a cavidade.

		30	40	5 0	60	70
Homo sapiens		HQSDT	HIFIIMCASCI	LAKKKIYPTIW	WLFRDGLLPEN	TFIVGYARSR
Schistosoma mansoni		DKSHVI	HVLVIIGASGI	LAKKKTYPTLW	WEFRDGLLPPR	TYFVGFARSD
Trypanosoma_bruce1	ADTLRKDSCDIIPDA	AAPELKERAI	LTIVVLCASCI	LARNETFPALF	QLFCNGLIPRT	INIVGYARTK
Trypanosoma cruzi	AAELRGEVCERIPDA	VSPELRSRAI	LTIVVLCASCI	LAKKKTFPALF	OLYCNGMLPRD	VNILGYARST
Arabidonsis thaliana	RASLODSCDO LTR	FHUTKCEST	LITTVVCASC	TAKKETFPALF	ARVECCLEDED	PSUPEVALTK
Spinacia oleracea	PIVAKSIDPSADLOLLPL	LESVKERPTI	LSILVVCASCO	KKRIFPALF	AFYENCLPEN	FTVFGFSBTE
Escherichia coli		MAVTOTAOV	CDLVIFCAKGD	LARRELLPSLY	OLEKAGOLNPD	TRIIGVGRAD
Leuconostoc mesenteroides		MVSEIK	TLVTFFGGTGI	LAKRKLYPSVF	NLYKKCYLQEH	FAIVGTARQQ
		-				
	8.0		100	10 12	0	120
Wana and and		NAMES DE LA CONTRACTA				130
Homo sapiens	LIVADIR KOS EPFF	KATPEERLKI	VERENNONEVA	AGQYDDAASIQ	INDEPTER	HLGSQAN
Trypaposoma brucei	MODVROW KKRSLAKHE	DRTKDRCDH	TRAPLETTY	SGSVDGARDER	RENDUTTERER	SEDGKOKGGN
Trypanosoma cruzi	MEDVEKW KKDTLAGFF	TRLDERGCH	VGNFLRRISYN	TGSYDRDEDFA	RMNERILRMEE	AFOGPEKGGN
Leishmania mexicana	VDDAEKW KHETLMKHF	SNLSERECH	AEDFLKHISY	CGSYDKVEDFK	RDIMIREKEN	AFKGPEKGGN
Arabidopsis_thaliana	LTHEELRDMISSTLTCRI	DQREKCGDK	MEQFLKRCFYH	ISGQYNSEEDFA	ELNKKLKEK	EAGKISN
Spinacia oleracea	MNDEELRTMISKTLTCRI	DQRENCGEK	MDHFLQRCFYH	ISGQYNSEDDFS	GLDCKLKEK	EAGRLQN
Escherichia coli	WDKAAYTKVVREALETFM	KETIDEG. LI	WDTLSARLDFC NBNBINUECWI	NLDVNDTAAFS	KIGAMLDQK	NRI
Leuconoscoc_mesenceroides	LSDDEFKQLVKDSI	KUFIEDQAQ	ABAFIARTSI	CANDVIDAASIG	I KSALEBAAT	KFDIDGM
	140 150	160	170	180	190	200
Homo sapiens	RLFYLALPPTVYEAVTKN	IHESCMSQIC	G WNRIIVER	PFGRDLQSSDR	LSNHISSLFRE	QIYRIDHYL
Schistosoma_mansoni	RIFYYAIPPSVYKPVSSS	IKEYCTNKNI	PDTWTRLIIER	PFGRDLESFNE	LNCELTKLFTE	EQIYRIDHYL
Trypanosoma_bruce1	RLFYLALPOSVFMHACTG	IRTHVMQKPC	GLGWVRIIIER	FGHDTESSNE	LSRQLEPLFEE	SQIFRIDHYL
Trypanosoma cruzi	RLFYLALPPSVFVGVCRG	LSKGAMQKPI	ELGWVRLIVER VCCWWRWIIFR	PFCRDTETSEQ	LSNQLKPLFNE	ROVERIDHYL
Arabidoneie thaliana	PL VVL STDENT PUDUUPC	ACLDACCPM		PADDOBCOCO	TPCLEOVITE	POTE PTDUVI.
Spinacia oleracea	RLFVLSIPONIFVDVVRC	VSHRASSASC	G. WTOVIVE	FGRDSDSSRE	LTRSFKOYLSE	DOIFRIDHYL
Escherichia coli	TINYFAMPOSTFGAICKG	LGEAKLNAKI	P ARVVMB	PLGTSLATSOE	INDOVGEYFEE	OVYRIDHYL
Leuconostoc mesenteroides	RIFYMSVAPRFFGTIAKY	LKSEGLLAET	TG.YNRLMIER	PFGTSYATAEE	LQSDLENAFDD	OLFRIDHYL
						_
	210 220	230	240	250	260	270
Homo_sapiens	GKEMVONLMVLRFANRIF	GPIWNRDNI	ACVILTEREPF	GTEGRGGYFDE	FGIIRDVMQNH	LLOMLCLVAM
Schistosoma mansoni	GKEMVQNLLILRFCNQVF	SPLUNRENI	DNITISFKEPF	GTEGRGGYFDQ	FGIIRDVVQNH	LIGILSLVAN
Trypanosoma brucei	GKEMVQNIVVTRFANRVF	SALWNNNI	ACVRITFKESI	GTEGRGGYFDK.	AGIIRDVVQNH	LTOILSLLAM
Trypanosoma cruzi	GKEMVQNIIVTRFANRVF	SALWNSNSI	ACVQITEKSKI	GTAGRGGYFDS	IGIIRDVIQNH	LTOILSLLTM
Arabidoneis thaliana	GKENVONI ITTREANEIF GVENVENLGVI PEGNIJVE	F DLUG PNV	NUOLTESODE	GTEGEGGTFDS	VETTPDTMONH	LIGILALPAN
Spinacia oleracea	GKELVENLSVLRFSNLVF	EPLOSRNY	RNVOLIFSODF	GTEGEGOVEDN	YGIIRDIMONH	LUILALFAM
Escherichia coli	GKETVLNLLALRFANSLF	VNNNDNRTII	HVEITVASEV	GIEGRWGYFDK	AGOMROMIONH	LLOILCMIAM
Leuconostoc mesenteroides	GKEMVONIAALREGNPIE	DAAMNKDYI	KNVQVTLABVL	GVEERAGYYDT	TGALLDMIONH	TMOIVGWLAM
-			-			
	280 290		300	310	320 3	зņ
Homo sapiens	280 290 EK <b>PAST</b> NSDD <b>VEDEKVKV</b>	LKCISEVQ.	300 ANNV <b>VLG</b>	310 YVGNPDGEGE.	320 3 ATKG <b>YL</b> DDPTV	3 0 PRGSTTATFA
Homo_sapiens Schistosoma_mansoni	280 290 EKPASTNSDDVRDEKVKV EKPVSVNADDIRDEKVKV	LKCISEVQ. LRSIEPLT.	300 ANNV <mark>VLG</mark> IDDI <mark>VIG</mark>	310 YVGNPDGEGE. YVADPNATNPP	320 3 ATKG <mark>YL</mark> DDPTV ASLS <b>YT</b> DDPSV	30 PRGSTTATFA PKDSITPTYV
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei	280 290 EKPASTNSDDVRDEXVKV EKPVSVNADDIRDEKVRV EKPRSLSPEDIRDEKVRV	LKCISEVQ. LRSIEPLT. LRHVNPVT.	300 ANNVVLC IDDIVIC PADCVLC	310 YVGNPDGEGE. YVADPNATNPP TRSEDGSIP.	320,3 ATKG¥LDDPTV ASLS¥TDDPSV G¥LEDPTV	3 0 PRGSTTATFA PKDSITPTYV PRGSKCATFV
Homo sapiens schistosoma mansoni Trypanosoma brucei Trypanosoma cruzi	280 EKPASTNSDDVRDEKVKV EKPVSVNADDIEDEKVKV EKPRSLSPEDIEDEKVKV EKPRSLSAEDIEDEKVOV	LKCISEVQ. LRSIEPLT. LRHVNPVT. LRQVVPAN.	300 ANNV <mark>VLCC</mark> IDDI <b>VIC</b> PADC <b>VLC</b> PAEC <b>VLC</b>	310 YVGNPDGEGE. YVADPNATNPP TRSEDGSIP. YTASADGSTP.	320 ATKG¥LDDPTV ASLS¥TDDPSV G¥LEDPTV G¥LDDPSV	3 Q PRGSTTATFA PKDSITPTYV PRGSKCATFV PKGSHCPTFA
Homo sapiens Schistosoma mansoni Trypanosoma brucei Trypanosoma cruzi Leishmania mexicana Vrabidancia bhaliana	280 290 EKDASTNSDDVEDEKVKV EKDVSVNADDIRDEKVRV EKDRSLSAEDIRDEKVVV EKDRSLSAEDIRDEKVVV EKDRTLDARCIRDEKVVV	LKCISEVQ. LRSIEPLT. LRHVNPVT. LROVVPAN. LKHIEPVT.	300 ANNVVLCC IDDIVIC PADCVLCC PADCVLCC KENCVLCC	310 VGNPDGEGE. VADPNATNPP VTRSEDGSIP. VTASADGSTP. VTASADGSIP.	320 ATKGYLDDPTV ASLSYTDDPSV GYLEDPTV GYLEDETV GYLEDETV	3 0 PRGSTTATFA PKDSITPTYV PRGSKCATFV PKGSHCPTFA PKGSTCPTFA
Homo sapiens Schistosoma mansoni Trypanosoma brucei Trypanosoma cruzi Leishmania mexicana Arabidopsis thaliana Spinacia oleracea	280 290 EKASTNSDDVEDEKVKV EKDVSVNADDIRDEKVKV EKDRSLSPEDIRDEKVVV EKDRSLSPEDIRDEVVV EKDRSLSABDIRDEKVVV ETVSLDAEDIRSEKVKV ETDVSLDAEDIRSEKVKV	LKCISEVQ. LRSIEPLT. LRHVNPVT. LROVVPAN. LKHIEPVT. LRSMKPLK. LRSMKPLK.	300 ANNVVLG IDDIVIG PADCVLG KENCVLG KENCVLG LEDVVVG	310 VGNPDGEGE. VADPNATNPP TRSEDGSIP. TASADGSIP. VTASADGSIP. VKGHNKGGKT.	320 ATKGYLDDPTV ASLSYTDDPSV GYLEDPTV GYLEDETV GYLEDETV .YSGYTDDPTV YSGYTDDPTV	30 PRGSTTATFA PKDSITPTYV PRGSKCATFV PKGSHCPTFA PKGSTCPTFA PNHSLTPTFA
Homo sapiens Schistosoma mansoni Trypanosoma brucei Trypanosoma cruzi Leishmania mexicana Arabidopsis thaliana Spinacia oleracea Bscherichia coli	280 290 EKDASTNSDDVEDEKVKV EKDASTNSDDVEDEKVKV EKDASLSPEDIEDEKVIV EKDRSLSAEDIEDEKVIV EKDRTLGAECIEDEKVSV ETDVSLDAEDIENEKVKV STDVSLDAEDIENEKVKV	LKCISEVQ. LRSIEPLT. LRHVNPVT. LROVVPAN. LKHIEPVT. LRSMKPLR. LRSMKPLK.	300 ANNVVLG IDDIVIG PADCVLG KENCVLG LEDVVVG LQDVVVG	310 VGNPDGEGE. VADPNATNPP TRSEDGSIP. TASADGSTP. KGHNKGGKT. KGHSKGNKS. VTAGFAOGKK	320 ATKGYLDDPTV ASLSYTDDPSV GYLEDPTV GYLEDETV GYLEDETV .YGYTDDPTV .YGYTDDPTV .YGYTDDPTV .YGYTDDPTV	3 0 PRGSTTATFA PKDSITPTYV PRGSKCATFY PKGSTCPTFA PKGSTCPTFA PNHSLTPTFA PNHSVTFTFA NKSSNTFTFA
Homo sapiens Schistosoma mansoni Trypanosoma brucei Trypanosoma cruzi Leishmania mexicana Arabidopsis thaliana Spinacia oleracea Escherichia coli Leuconostoc mesenteroides	280 290 EKJASTNSDDVEDEKVKV EKJNSDDIRDEKVKV EKJRSLSPEDIRDEKVIV EKJRSLSPEDIRDEKVIV EKJRSLSARDIRDEKVVV EKJRSLDARDIRSEKVKV ETJVSLDARDIRNEKVKV SPSDLSADSIRDEKVKV	LKCISEVQ. LRSIEPLT. LRHVNPVT. LROVVPAN. LKHIEPVT. LRSMKPLR. LRSMKPLR. LKSLRRIDRS	300 ANNVVLG PACVLG PACVLG PACVLG LCVVVG LQVVVG SNVREKTVRG EEVNKYFVRG	310 VGNPDGEGE. VADPNATNPP TRSEDGSIP. TASADGSIP. VTASADGSIP. VGNKGKT. KGHSKGNKS. TAGFAQGKK. VGAGDADYK.	ATKGYLDDPTV ASLSYTDDPSV GYLDDPSV GYLDDPSV GYLEDPTV .YPGYTDDPTV .YSGYTDDPTV .VPGYLEEEGA PYLEEAD	3 0 PRGSTTATFA PKDSITPTYV PRGSKCATFV PKGSTCPTFA PKGSTCPTFA PNHSLTPTFA PNNSVTPTFA NKSSNTETFY
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides	280 290 EKASTNSDDVEDEKVKV EKDVSVNADDIRDEKVKV EKDRSLSPEDIRDEKVVV EKDRSLSAEDIRDEKVVV EKDRSLSAEDIRDEKVVV ETVSLDAEDIRSEKVKV ETVSLDAEDIRNEVKV ETVSLDAEDIRDEKVKV ETVSLDAEDIRDEKVKV ETVSLDAEDIRDEKVKV	LKCISEVO LRSIEPLT. LRHVNPVT. LKHVNPVT. LKHIEPVT. LKSKPLR. LKSLRRIDRS PNALKIYNEI	300 ANNVULG PADCULG PACULG LEDVUVG SNVREXTURG EEVNKYFURA	310 VVGNPDGEGE. VADPNATNPP TRSEDGSIP. VTASADGSIP. VTASADGSIP. VKGHNKGGKT. VGHSKGNKS. VGHSKGNKS. VGAGDTADYK.	ATKGYLDDPTV ASLSYTDDPSV GYLDDPSV GYLDDPSV GYLEDPTV GYLEDETV .YPGYTDDPTV .YSGYTDDPTV .VPGYLEEGA PYLEBADV	3 Q PRCSTTATFA PRDSITPTYV PRCSKCATFV PRCSKCPTFA PRCSTCPTFA PNHSLTPTA PNNSVTPTFA NKSSNTETFV PADSKNTFTFI
Homo_sapiens Schistosoma_brucei Trypanosoma_brucei Trypanosoma cruzi Leishmania mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides	280 290 EK-ASTNSDDVRDEKVKV EK-VSVNADDIZDEKVKV EK-VSVNADDIZDEKVVV EK-RSLSPEDIRDEKVVV EK-RSLSADIRDEKVVV ET-VSLDARDIZSEVKV ET-VSLDARDIZSEVKV ET-VSLDARDIZDEKVKV EF-SFNDKDIZAAKNAA	LKCISEVQ. LRSIEPLT. LRHVNPVT. LKHIEPVT. LKHIEPVT. LKSKPLK. LKSLRRIDRS FNALKIYNEI	300 ANNYULGC IDDIUGC PADCULGC PACULGC LEDVUVGC LEDVUVGC LQDVUVGC SNVREKTVRG EEVNKYFVRA	310 VADPDATNP TRSEDGSIP. TASADGSIP. TASADGSIP. KGHNKGKT. KGHSKGNKS. TAGFAQGKK. GAGDTADYK.	ATKCILDDTV ASLSTDDTV CLEDPTV CLEDPTV CLEDPTV CLEDPTV CLEDETV .YPGITDPTV .YGGITDDPTV .YPGITEESA PILEESA	9 PRGSTTATPA PRGSKCATPV PRGSKCATPV PRGSKCATPV PRGSCOTPA PRGSCOTPA PRGSCOTPA PRGSCOTPA PRNSVTPTPA PNNSVTPTPA PADSKNNTPI
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_truzi Leishmania mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides	280 290 EKDASTNSDDVEDEVKV EKDASTNSDDVEDEVKV EKDASTNSDDVEDEVKV EKDASLSPEDIEDEKVVV EKRELSAEDIEDEKVVV EKDATLDAECIEDEKVVV ETDVSLDAEDIENEKVVV SPOSDLSADSIEDEKVVV EKDEFNDKDIEAAKNAA	LKCISEVQ. LRSIEPLT. LRVVPVT. LRVVPVT. LKUVVPAN. LKSIERTDRS FNALKIYNEI	300 ANNYULG IDDIULG PARCYLG RENCYLG KENCYLG LEDVYVG LEDVYVG SNVREKTYRC EEVNKYFYRA	310 VGNPDGEGE. VADPNATNPP TRSEDGSIP. TASADGSIP. KGHNKGGKT. KGHSKGNKS. GAGDTADYK.	320 ATKCILDDPTV ASLSTDDPSV .CILEDPTV .CILEDPTV .CILEDPTV .YSCITDDPTV .YSCITDDPTV .YSCITDDPTV .VPCLEEGA	DEGSTTATPA PROSITPTYV PROSKCATPV PROSKCATPV PROSKCATPV PROSKCATPV PROSKCATPV PROSKCATPV PROSKCATPV PROSK PRO
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides	280 290 EKDVSVNADDIZDEXVKV EKDVSVNADDIZDEXVKV EKDRSLSPEDIZDEXVVV EKDRSLSPEDIZDEXVVV EKDRSLSAEDIZDEXVVV ENDRTLDAECIZDEXVVV ETDVSLDAEDIZSEXVKV SPISDLSADSIZDEXVKV EKDESPNDKDIZAAXNA	LKCISEVQ. LRSIEPUT. LRHVNPVT. LKHUVPPT. LKHIEPVT. LKSKFPVT. LKSKFPVT. LKSKFPVT. LKSLRRIDRS FNALKIYNEI	300 ANNYULGC IDDIYGC PADCULGC PADCULGC KENKULGC LEDVYVGC SNVREKTURGC EEVNKYFVRAC 370	310 VAGNPDGEGE. VADPNATNPP TRSEDGSIP. TASADGSIP. VTASADGSIP. XGHNKGGKT. KGHSKGNKS. TAGFAQGKK. GAGDTADYK.	320 ATKCIDDPSV ASISTDDPSV CLEDPTV CLDDPSV CLEDPTV CLDPSV CLDPSV VPGTDDPTV .VPGTDPTV .VPGLEEEGA PLEEADV	30 PRGSTTATPA PRGSTTPTYV PRCSTCPTYV PRGSTCPTYA PRGSTCPTPA PNRSTPTPA PNNSVTPTPA NKSSNTETPV PADSKNNTPT
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_cruzi Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni	280 290 EKASTNSDDVEDEKVKV EKDVSVNADDIEDEKVKV EKDVSUSPEDIEDEKVVV EKDRSLSPEDIEDEKVVV EKDRSLSPEDIEDEKVVV ETDVSLDAEDIESEKVKV ETDVSLDAEDIENEVKV SPDSDLSADSIEDEKVKV EKDESFNDKDIEAAKNAA 340 250 AVVEYVEREKDCVPPIE	LKCISEVQ. LRSIEPLT. LRVVPVAN. LRVVPVAN. LRSWKPLK. LKSIKPIR. LKSLKIDRS FNALKIYNEI 360 RCEALNER	300 ANNYULGC IDDIUIG PADCULGC PACULGC LEDVUYGC LEDVUYGC SNVREKTVRG EEVNKYFYRAC	310 VADPDATNP TRSEDGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VKGHNKGGKT. VGASKGNKS. VGAGDTADYK. 380 VAGDIFHQ	320 ATKCIDDTV ASLSTDDSV CLEDPTV CLEDPTV CLEDPTV CLEDPTV CLEDPTV CLEDPTV CLEDTV VPCIDDTV CLEEEGA PILEEADV 390 QCKENEV	PROSITATPA PROSITATPA PROSICATPY PROSKCATPY PROSKCATPY PROSTOTATA PROSTOTATA PROSTOTATA PROSTOTATA PROSTOTATA
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_truzi Leishmania mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei	280 290 EKJASTNSDDVRDEKVKV EKJVSVNADDIRDEKVKV EKJVSVNADDIRDEKVKV EKJRSLSPEDIRDEKVVV EKJRTLDARCIRDEKVQV EKJRTLDARCIRDEKVVV EKJRSLSADSIRDEKVKV EFJVSLDARDIRDEKVKV EKJESFNDKDIRAAKNAA 340 350 AVVLYVENERNDVJVFIL CAVLYVKSDRKKCVPFIL	LKCISEVQ. LRSIEPLT. LRVNPVT. LRVVPPN. LRSMKPLK. LRSMKPLK. LRSMKPLK. LKSLRTDRS PNALKIYNEP 260 RCCALNER RACAVER	300 ANNYULGC IDDIUIG PARCVLGC LEDVUVG LEDVUVG LEDVUVG EEVNKYFYRA 370 KAEVRLOHDV KAEVRVQ KEE	310 VADPDATNP VADPNATNPP TRSEDGSIP. TASADGSIP. KGHNKGGKT. KGHSKGNKS. XTAGFAQGKK. GAGDTADYK. 380 AGDIFHQ HIHLFGS HIHLFGS	320 ATKCLDDPTW ASLSTDDPSW CLEDPTW CLEDPTW CLEDPTW CLEDTW CLEDTW CLEDTW CLEETW DPTLEEGA PLEEGA 	3 0 PRGSTTATPA PRDSITPYV PRGSKCATPV PRGSKCATPV PRGSKCATPV PRGSTCPTPA PRGSTCPTPA PNNSVTPTPA NKSSNTSPV PADSKNNTFI 400 IRVQPDBAVY IRVQPDBAVY IRVQPDBAVY
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei	290 290 EKASTNSDDVEDEKVKV EKOVSVNADDIRDEKVKV EKORSLSPEDIRDEKVVV EKORSLSPEDIRDEKVVV EKORTLDABCIRDEKVVV EROPTLDABCIRDEKVVV SPOSDLSADSIRDEKVKV SPOSDLSADSIRDEKVKV EKORSPNDKDIRAANAA 340 350 AVVLYVERERDOVPPIL VIRLENNDEKDGVPPIL VIRLENNDEKDGVPPIL	LKCISEVQ. LRSIEPUT. LRHVNPVT. LKHIVPVT. LKHIVPVT. LKSIKPUR. LKSIKPUR. LKSIRRIDR PNALKIVNPI RCOLAINER RAGRALNER RAGRALNER RAGRALEER	300 IDDIYGC PADCULGC PADCULGC EDVUVGC LEDVUVGC SNVREKTVRC SNVREKTVRC BEVNKYFVRA 370 KAEVRLOTHD KAEVRCOTKET YLGIRICZKD.	310 VAGNPDGEGE. VADPNATNPP TRSEDGSIP. TASADGSIP. TASADGSIP. KGHSKGNKS. KGHSKGNKS. GAGDTADYK. 380 AGDIFHQ HIHLFGS EIRPFGV EIRPFGV	320 ATKCIDDPSV ATKCIDDPSV CILEDPTV CILEDPTV CILEDPTV CIDDPSV CIDDPSV CIDDPTV YSCIDPTV .VSCIDPTV .VSCIEEEGA PILEEADV CKRNEV KEG LPRNEV AAQRNEIV	9 PRGSTTATFA PRCSTTPTYV PRCSTCPTPA PRGSTCPTPA PRGSTCPTPA PNRSVTPTPA NKSSNTETPV PADSKNNTFTP 1 400 1 RVQPDEAVY 1 RAQPSEAMY
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Irypanosoma_cruzi Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_cruzi Leishmania_mexicana	280 290 EKD X5 TNSDDVEDEKVKV EKD X5 NADDIRDEKVKV EKD X5 NADDIRDEKVVV EKD X5 LSPEDIRDEKVVV EKD X5 LSPEDIRDEKVVV EKD X5 LDAEDIRDEKVVV ETD V5 LDAEDIRDEKVVV SP SDLSADSIEDEKVVV EKD 250 AVVLYVENE KO SVF JL CAVLYVKEREKO SVPFIL VLRLFINNDEKO SVPFIL VLRLFINNDEKO SVPFIL	LKCISEVQ. LRSIEPLT. LRSVPVPAN. LROVVPAN. LKHIEPVT. LKSKPUK. LKSLRLDRS FNALKIVNEI S60 RCGMALNER RAGKALNER RAGKALEER RAGKALEER	300 ANNY VLGC IDDIVLGC PADCVLGC PAECVLGC LEDVVVGC LEDVVVGC SNVREKTVRGC BEVNKYFVRAC 370 KAEVRVCCKEE VLGIRIOCKD. LDDIRCCKD.	310 VAGNPDGEGE. VADPNATNP TRSEDGSIP. VTASADGSIP. VTASADGSIP. XGHNKGGKT. XGASKGNKS. VGAGDTADYK. 380 VAGDIFHQ HILFGS EIRPFGV EIRPFGE	320 ATKCIDDTV ASLSIDDTV ASLSIDDTV CCLEDTV CCLEDTV VYOCTDDTV VYOCTDDTV VYOCTDDTV VYOCTEEEGA 	PROSITATPA PROSITPTY PROSICATPY PROSICATPY PROSICATPY PROSICATPY PROSICATPY PROSICATPT PROSICATPY PROSICATPY PROSICATPY PROSICATPY PADEXINITY IRVQPDEAVY IRVQPSEAMY IRAQDSEAMY
Homo_sapiens Schistosoma_brucei Trypanosoma_brucei Irypanosoma_brucei Leishmania mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Irypanosoma_cruzi Leishmania_mexicana Arabidopsis_thaliana	2 8 0 2 9 0 E K A 5 TN S D D V R D E K V K V E K V NA D I D E K V K V E K V S NA D I D E K V K V E K P S L S P D I E D E K V V V E K P R L S A D I E D E K V V V E T V S L D A E D I E S E V K V E T V S L D A E D I E D E K V K V S P S D L S A D I E D E K V K V E T V S L D A E D I E D E K V K V E T V S L D A E D I E D E K V K V E T V S L D A E D I E D E K V K V E T V S L D A E D I E D E K V K V E T V S L D A E D I E D E K V K V E T V S L D A E D I E D E K V K V E T V S L D A E D I E D E K V K V E T V S L D A E D I E D E K V K V E T V S L D A E D I E D E K V K V E T V S L D A E D I E D E K V K V E T V S L D A E D I E D E V K V E T D I D E V K V F I L V I R L F I N N D K M E V P F I L V M E L N N D K M E V P F I M V M E L N N D K M E V P F I M V M E L N N D K M E V P F I M V M E L N N D K M E V P F I M V M E L N N D K M E V P F I M V M E L N N D K M E V P F I M V M E L N N D K M E V P F I M V M E L N N D K M E V P F I M V M E L N N D K M E V P F I M V M E L N N D K M E V P F I M V M E L N N D K M E V P F I M V M E L N N D K M E V P F I M V M E L N N D K M E W D E V P F I M K M E M M E M M E M E M E M E M E M E M	LKCISEVQ. LRSIEPLT. LRSVPVPVT. LRVVPVAN. LRSWFPLR. LKSIEPVT. LKSLKIENS FNALKIYNEF 360 RCIALNER RAGIALNER RAGIALEER RAGIALEER RAGIALEER	300 ANNYULGC IDDIUIG PADCULGC PARCULGC LEDVUVG LEDVUVG SNVREXTVGG EEVNKYFYRA 370 KAEVRLOTHDU KAEVRCOKEF YVAIRICKD SAEIRVCKPR	310 VADPDATNP TRSEDGSIP. TASADGSIP. TASADGSIP. KGHNKGGKT. KGHSKGNKS. TAGFAQGKK. GAGDTADYK. 380 AGDIFHQ HIHLFGS EIRPFGV EIRPFGE YGNLYKKSFAT	320 ATKCIDD PTW ASLSTDD PSW CLED PTW CLED PTW CLED PTW CLED PTW CLED PTW CLEDETW VPCIEESQA PULEESQA 	9 PROSITATPA PROSITATPA PROSICATPY PROSICATPY PROSICATPY PROSICATPY PROSICATPA PROSICATPA PROSICATPA PROSICATPA PROSICATPA A000 IRVQENEAVY IRVQEDEAVY IRVQEDEAVY IRAQESEAMY IRAQESEAMY IRAQESEAMY IRAQESEAMY
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_truzi Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea	280 290 EKASTNSDDVEDEKVKV EKOVSVNADDIRDEKVKV EKOVSVNADDIRDEKVKV EKORSLSPEDIRDEKVVV EKORSLSPEDIRDEKVVV EKORSLSAEDIRDEKVVV ENDVSLDAEDIRDEKVVV SPOSDLSADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVVV AVVLYVENEKNDGVPFIL VIRLENNNDKMGVPFIL VIRLENNNDKAGVPFIL VIRLENNNDKAGVPFIL AAAMFINNARSNOVPFIL AAAMFINNARSNOVPFIL	LKCISEVQ. LRSIEPUT. LRHVNPVT. LRHVNPVT. LKHIEPVT. LKSLRRIDRS PNALKIVNPT S60 RCGIAINER RAGKALNER RAGKALNER KAGKALNER KAGKALNER KAGKALNER KAGKALNER KAGKALNER	300 NNNYULGC IDDIYGC PADCULGC PADCULGC EENYUYGC SNVREXTYRC SNVREXTYRC SNVREXTYRC SNVREXTYRC VYGC KAEVROCKEF YLGIRICKD YMAIRICKD MC REVCKEF SNVREXTYRC AEEVROCKEF SNVREXTYRC AEEVROCKEF SNVREXTYRC AEEVROCKEF SNVREXTYRC	310 VAGNPDGEGE. VADPNATNPP TRSEDGSIP. TASADGSIP. KGHSKGNKT. KGHSKGNKS. GAGDTADYK. 380 AGDIFHQ AGDIFHQ B. EIRPFGV EIRPFGV EIRPFGV EVHPYGE VGNLYKKSFAT	320 ATKCILDDPSW ATKCILDPSW ATKCILDPSW CILEDPTW CILEDPTW CILEDETW .YSCIDDPTW .YSCIDDPTW .YSCIDPTW .VSCILEEEGA PILEEADW CKENEV KEG LPRNEV ACKENEV TORNEV ATORNEV ATORNEV ATORNEV ATORNEV ATORNEV ATORNEV	9 PRGSTTATFA PRGSTTATFA PRCSTTPTYV PRCSKCATFV PRGSTCPTFA PRGSTCPTFA PNNSVTPTA NKSSNTETFV PADSKNNTFT PADSKNNTFT IRVQPDEALVY IRVQPDEALVY IRAQPSEAMY IRAQPSEAMY IRAQPSEAMY IRAQPSEAMY IRAQPSEAMY IRAQPSEAMY IRAQPSEAMY IRAQPSEAMY IRAQPSEAMY IRAQPSEAMY IRAQPSEAMY IRAQPSEAMY IRAQPSEAMY IRAQPSEAMY
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Irypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Irypanosoma_brucei Irypanosoma_brucei Irypanosoma_brucei Spinacia_oleracea Escherichia_coli	290 290 EKASTNSDDVEDEKVKV EKASTNSDDVEDEKVKV EKASLSPEDI PDEVVV EKASLSPEDI PDEVVV EKASLSPEDI PDEVVV EKASLSPEDI PDEVVV EKASLSPEDI PDEVVV ETAVSLDAEDI SEKVVV ETAVSLDAEDI SEKVVV SPSDLSADSI PDEVVV EKASEPNDKDI PAAKNAA 340 350 AVVLYVENERD KGVPFIL VLRLPINNDENG GVPFIL VLRLPINNDENG GVPFIL AAAMFINNAEND GVPFLM AAALFINNAEND GVPFLM AAALFINNAEND GVPFLM AAALFINNAEND GVPFLM AAALFINNAEND GVPFLM	LKCISEVQ. LRSIEPLT. LRSVPPT. LRVVPPN. LRVVPPN. LRSMKPLR. LSSMKPLK. LKSLRVF NALKIVNE 260 RCGALNER RAGAANER RAGAANER RAGAANER RAGAANER RAGAANER RAGAANER RAGAANER RAGAANER RAGAANER RAGAANER RAGAANER	300 ANNYULGC IDDIYIC PADCULGC PACULGC EDVYYC SNVREKTVRGC EEVNKYFVRG EEVNKYFVRG EEVNKYFVRG LUDIRIG KD JUDIRIG KD JUDIRIG KD JUDIRIG KD JUDIRIG KD	310 VAGNPDGEGE. VADPNATNP TRSEDGSIP. TASADGSIP. XTASADGSIP. XGASKGNKS. GAGDTADYK. GAGDTADYK. GAGDTADYK. BIRPFGV AGDIFHQ EIRPFGV EIRPFGV EVHPYGE PGNLYKKTFGT ELNLFKE	320 ATKCIDDTY ASLSIDDTY ASLSIDDTY ASLSIDDTY CCLEDTY CCLEDTY CCLEDTY YPOTDDTY YSOTDDTY VSOTDDTY ASLSID ASS ASS ASS ASS ASS ASS ASS AS	9 PRGSTTATFA PRGSTTATFA PRGSTCATFY PRGSKCATFY PRGSCCTFA PRGSTCFFA PRGS
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Irypanosoma_cruzi Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_cruzi Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides	290 290 EKASTNSDDVRDEKVKV EKDVSVNADDIRDEKVKV EKDVSVNADDIRDEKVVV EKDRSLSPEDIRDEKVVV EKDRSLSPEDIRDEKVVV ETDVSLDARDIRSEKVKV ETDVSLDARDIRDEKVKV SPDSDLSADSIRDEKVKV EKDESFNDKDIRAAKNAA 340 350 AVVLYVENERDOKVPJIL VLRLENNDRMACVPJIL VMRLNNNDRMGVPJIL AAMFINNARN GVPJLM AAALFIDNARN GVPJLM AIRVDDNWRMGVPJYV	LKCISEVQ. LRSIEPLT. LRSVPVT. LRVVPAN. LROVVPAN. LRSMKPLR. LKSIEVT. LKSLEIDRS FNALKIYNE S60 RCIALNER RACKAVER RACKAVER RACKAVER KACKAVER KACKAVER KACKAVER KACKAVER	300 ANNYULGC IDDIUIG PADCULGC PACULGC LEDVUYGC LEDVUYGC LEDVUYGC LQUVYGC BEVNKYFYRAC 370 KÄEVRLOTHDU KÄEVRLOTHDU GAEIROCKD GAEIROCKD GAEIRVORHV RAEIRVORHV KAEVRVORHV SEVVYKTF QTRVDIVGKAG	310 VADPDATNP TRSEDGSIP. TASADGSIP. TASADGSIP. KGHNKGGKT. KGHSKGNKS. GAGDTADYK. BAGDTADYK. BAGDTADYK. BIRPFGE EIRPFGE EIRPFGE VGNLYKKTFGT DTFNFGS	320 ATKCIDDTW ASLSTDDSV CILDDTV CILDDTV CILDDTV CILDDTV CILDTV CILDTV CILDTV CIERTU 	PROSTATATA PROSTATATA PROSTATATA PROSTOFA
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides	280 290 EK2 VS WADD IK DEK VK V EK2 VS WADD IK DEK VK V EK2 VS WADD IK DEK VK V EK2 RS LS PED I 2 DEK VK V EK2 RS LS PED I 2 DEK VK V EK2 RS LD AED I 2 DEK VK V EK2 RS LD AED I 2 DEK VK V SP35 DLS AD SI 2 DEK VK V EK2 SS PND KD I 2 AAKNAA 340 350 AVVLYVENE 20 CVP7 I L VLR LFINND 20 CVP7 I L VLR LFINND 20 CVP7 I L VLR LNINND 20 CVP7 I L VLR LNINND 20 CVP7 I L VLR LNINND 20 CVP7 I L AAAM FINNA 20 VP7 LM AAALFIDNA 20 CVP7 LM AARVDIDNA 20 VP7 LM	LKCISEVQ. LRSIEPUT. LRHVNPVT. LRHVNPVT. LRHVNPVT. LRSMKPLR. LKHIEPVT. LRSMKPLR. LKSLRRIDRS PNALKIYNER RGANALNER RAGKAVERK KAGALNER KAGALNER KAGALHTRK KAGALHTRK RTGKRLPAK	300 ANNY ULGC IDDIYG PADCULGC PADCULGC EEVVYG SNVREXTYRC SNVREXTYRC SNVREXTYRC BEVNKYFVRA 370 KAEVRVQ KEF YLGIRIQKD YLGIRIQKD YVAIRICKD SNVREXTYRC BEVNKYFVRA SNVREXTYRC BEVNKYFVRA SNVREXTYRC SNV	310 VACONPDGEGE. VADPNATNPP TASADGSIP. TASADGSIP. XTASADGSIP. XTASADGSIP. XTASADGSIP. XTAGFAQKT. XGAGDTADYK. GAGDTADYK. BIRPFGV 	320 ATKCIDDPSV ATKCIDDPSV ATKCIDDPSV CILEDPTV CILEDPTV CILEDPTV CILEDSV CILEDSV CILEDSV CILEDSV   	Image: Solution of the second state of the second
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Leishmania mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides	200 290 EKASTNSDDVEDEKVKV EKOVSVNADDIRDEKVKV EKOVSVNADDIRDEKVVV EKORSLSAEDIRDEKVVV EKORSLSAEDIRDEKVVV ENDRIDAECIRDEKVVV ETOVSLDAEDIRDEKVVV SPSDLSADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNADSIRDEKVKV EKOSSNA AVVLVVENEENDOV VIRLINNDENGVPPIL AAALFINNARNDOVPFLM AIRVDIDNKVAGVPFYL AGELOPDLPKKEVPYVE	LECISEVQ. LRSIEPUT. LRHVNPVT. LRHVNPVT. LRHVENVT. LKHIEPVT. LRSMKPLR. LKSLRRIDRS FNALKIVNE 260 RCGLAINER RAGTALNER RAGTALNER KAGTALHTR KAGTALHTR RTGERLPTK KAGTALHTR	300 ANNYULGC IDDIYGC PADCULGC PACULGC EDVYYGC SNVREKTURGC EEVWKYFVRA 370 KAEVRVC KEF YLGIRICSD YLGIRICSD YVAIRICSD AERVCRHU CSEVVYYKTF GEVVYXKTF GEVVYXKTF GEVVYXKTF GEVVYYKTF GEVVYXKTF GEVVYXKTF GEVVYXKTF GEVVYXKTF GEVVYXKTF GEVVYXKTF GEVVYXKTF GEVVYXKTF GEVVXX GEVVYXKTF GEVVXX GEVVXX GEVVXX GEVVXX GEXX GEX GE	310 VACONPDGEGE. VADPNATNPP TRSEDGSIP. TASADGSIP. VAGHNKGGKT. KGHSKGNKS. GAGDTADYK. GAGDTADYK. GAGDTADYK. BIRDFGV EIRPFGV EIRPFGV EIRPFGV COLYKKSFAT COLYKKFGS. EINLFKE	320 ATKC IDD PTW ASLS TOD PSW ASLS COLOPSW CLEDPTW CLEDTW .YPC TDDPTW .YPC TDDPTW .YPC IDETW .YPC IDETW .YPC LEEEGA PILEEADW QCKENEJW AAORNEJU AAORNEJW ATORNEJW ATORNEJW ATORNEJW ATORNEJW ATORNEJW ATORNEJW ATORNEJW	9 PRGSTTATFA PRGSTTATFA PRGSTTATFA PRGSTCATFA PRGSTCPTA PRGSTCPTA PRGSTCPTA PNNSVTPTA NKSSNTETFV PADSKNNTFT IRVQPDEAVY IRAQPSEAMY IRAQPSE
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Schistosoma_brucei Trypanosoma_brucei Schistosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Schistosoma_brucei teishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides	290 290 EXASTNSDDVEDEXVKV EKDVSVNADDIRDEXVKV EKDVSVNADDIRDEXVVV EKDRSLSPEDIRDEXVVV EKDRSLSPEDIRDEXVVV EXDRSLSABJRDEXVVV ETDVSLDAEDIRSEXVKV ETDVSLDAEDIRDEXVKV ETDVSLDAEDIRDEXVKV EXDRSLSADSI EXDRSLSADSIRDEXVKV EXDRSLSADSIRDEXVKV EXDRSL	LKCISEVQ. LRSIEPLT. LRSVVPAN. LRVVPVT. LRSVVPLAN. LRSMKPLK. LKSIERIDRS FNALKIVNER RCGALNER RAGKALEBRI RAGKALEBRI RAGKALEBRI RAGKALEBRI KAGKAVEQKI KAGKAVEQKI KAGKAVEQKI KAGKAVEQKI KAGKAVEQKI KAGKAVEQKI KAGKALEBRI RAGKALEBRI CALANA CA	300 ANNY VLGC IDDIVIG PADOVLGC PARCVLGC LEDVVYGC LEDVVYGC SNVREKTVRGC BEVNKYFVRA 370 KAEVRVOTKEF VLGIRIQEKD LDIRICKD LOVYGC BEVNKYFVRA BEVNKYFVRA CEVVYKF QAEIRVOTKE QAEIRVOTKE QAEIRVOTKE QAEIRVOTKE QAEIRVOTKE QAEIRVOTKE	310 VADPDATNP TRSEDGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VGHNKGGKT. VGAGDTADYK. VGAGDT	320 ATKCIDDTSV ASLSIDDTSV ASLSIDDTSV CILEDTV CILEDTV CILEDTV .YSCIDDTV SIS 	PROSTATATA PROSTATATATA PROSTATATA PROSTATATATA PROSTATATATA PROSTATATATA PROSTATATATA PROSTATATATA PROSTATATATATA PROSTATATATA PROSTATATATATATATATATATATATATATATATATATATA
Homo sapiens Schistosoma mansoni Trypanosoma brucei Trypanosoma cruzi Leishmania mexicana Arabidopsis thaliana Spinacia oleracea Escherichia coli Leuconostoc_mesenteroides Homo sapiens Schistosoma mansoni Trypanosoma cruzi Leishmania mexicana Arabidopsis thaliana Spinacia oleracea Escherichia_coli Leuconostoc_mesenteroides	290 290 EKASTNSDDVEDEKVKV EKOVSVNADDIEDEKVKV EKOVSVNADDIEDEKVKV EKORSLSPEDIEDEKVSV EKORSLSPEDIEDEKVSV EKORSLSAEDIEDEKVSV EKOVSLDAEDIENEKVKV SPOSDLSADSIEDEKVKV EKOSPNDKDIEAAENAA 340 350 AVVEYVENEROVYPIL AVVEYVENEROVYPIL ALPEINNEN AVVPPIL VIRLEINNEN AVVPPIL AAAMFINNARNOVPPIL AAAMFINNARNOVPPIL AAALPIDNARNOVPPIL AAALPIDNARNOVPPIL AABLOPDLPENEGVPPYV 410 TKMMTEKPOMFFNPE	LKCISEVQ. LRSIEPUT. LRSVVPAN. LRHVNPVT. LRHVNPVT. LRSMKPLR. LKSIERTDRS PNALKIYNER S60 RCGIALNER RAGKAVERK KAGALNER RAGKAVERK KAGALHTRK KAGALHTRK RTGIRLPAK SGRLAAL	300 ANNY UGC IDDIYG PADCULG PACULG EDVUYG SNVREXTYRG SNVREXTYRG BEVNKYFVRA 370 KAEVRLOTHD KAEVRVOKEF YLGIRIGKD YVAIRIOKE YVAIRIOKE SNVRYFKAVKLPDI 430 NRYKNVKLPDI	310 VACONPDGEGE. VADPNATNPP TASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTAGFAQKK. SGAGDTADYK. SGAGDTADYK. SGAGDTADYK. BIRPFGV EIRPFGV EIRPFGV EIRPFGV BILYKSFAT VPGNLYKKSFAT VPGNLYKKSFAT SGAGDTADYKSFAT	320 3100 390 390 0 CKENEJV XEC LOPAN 390 0 CKENEJV XEC LOPAN 390 0 CKENEJV XEC LOPAN 390 0 CKENEJV XEC LOPAN 390 0 390 0 CKENEJV XEC LOPAN 390 0 CKENEJV XEC LOPAN 390 0 CKENEJV XEC LOPAN 390 0 390 0 CKENEJV XEC LOPAN 390 0 CKENEJV XEC LOPAN 390 0 CKENEJV XEC LOPAN 390 0 CKENEJV XEC LOPAN 450 450 450 450 450 450 450 450	9 PRGSTTATFA PRGSTTATFA PRCSTTPTYV PRCSKCAFPV PRGSTCPTPA PRGSTCPTPA PRGSTCPTPA PNRSVTPTA NKSSNTEIFV PADSKNNTFI PADSKNNTFI IRVQPDEALVY IRVQPDEALVY IRAQPSEAMY IRAQPSEAMY IRQQPDEGIY IRQQPDEGIY IRQQPDEGIY IRQQPDEGIY IRQQPDEGIY IRQQPDEGIY REAVEIFTPL 6 0
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Irypanosoma_cruzi Leishmania mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Irypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei	200 290 EKASTNSDDVEDEKVKV EKOVSVNADDIRDEKVKV EKOVSVNADDIRDEKVVV EKORSLSPEDIRDEKVVV EKORSLSPEDIRDEKVVV EKORSLSAEDIRDEKVVV ETOVSLDAEDIRSEKVKV ETOVSLDAEDIRNEKVKV SPOSDLSADSIRDEKVKV EKOSPNDKDIRAAKNAA 340 350 AVVLYVENEENDOVPFIL VLRLFUNNDRNGVPFIL VLRLFUNNDRNGVPFIL AAAMFINNARNDVPFIL AAAMFINNARNDVPFIL AAAMFINNARNDVPFIL AAAMFINNARNDVPFIL AAAMFINNARNDVPFIL AAAMFINNARNDVPFIL AAAFONDRNGVPFIL AAAFONDRNGVPFIL AAAFONDRNGVPFIL AAAFONDRNGVPFIL AAAFONDRNGVPFIL AAAFONDRNGVPFIL AAAFONDRNGVPFIL AAAFONDRNGVPFIL AAAFONDRNGVPFIL AAAFONDRNGVPFIL AAAFONDRNGVPFIL AAAFONDRNGVPFIL AAAFONDRNGVPFIL AAFONDRNGVPFIL AAAFONDRNGVPFIL AAFONDLONNRNGVPFIL AAAFONDLONNRNGVPFIL AAFONDLONNRNGVPFIL AAFONDLONNRNGVPFIL AGELOFDLPRNEGVPFYL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL AGELOFDLFREGMAAFONDRNGVPFIL	LKCISEVQ. LRSIEPUT. LRIVNPVT. LRIVNPVT. LRINKPVT. LKHIEPVT. LKSIERIDRS FNALKIVNE 260 RCGALNERI RAGALNERI LSURI LANGLINI LSURI	300 ANNY VIGC IDDIY GC PADCVLGC PACVLGC EDVVYGC SNVREKTVRGC EEVWKYFVRA KAEVRVCFKEF VIGIRICSKD YVAIRICSKD YVAIRICSKD YVAIRICSKD GAEIRVCFKH CSEVVYYFKF GAEIRVCFKH CSEVVYYFKF 430 NRYKNVKLPDJ HRYKAIKLPDJ	310 VACONPDGEGE. VADPNATNPP TRSEDGSIP. TASADGSIP. VAGHNKGGKT. KGHSKGNKS. TAGFAQGKK. GAGDTADYK. GAGDTADYK. GAGDIFHQ EIRPFGV EIRPFGV EIRPFGV ELNLFKE ELNLFKE TFNFGS 440 VGRLILDVFCC VGRLILDVFCC VGRLILDVFCC	320 31000 3100 3100 3100 3100 3100 3100 3100	9 PRGSTTATFA PRGSTTATFA PRGSTTATFA PRGSTCATFA PRG
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania mexicana Arabidopsis_thaliana Spinacia oleracea Escherichia coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Descharichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei	290 290 EKASTNSDDVEDEKVKV EKDVSVNADDIRDEKVKV EKDVSVNADDIRDEKVVV EKDRSLSPEDIRDEKVVV EKDRSLSPEDIRDEKVVV EKDRSLSABDIRDEKVVV ETDVSLDAEDIRSEKVKV ETDVSLDAEDIRSEKVKV ETDVSLDAEDIRSEKVKV ETDVSLDAEDIRSEKVKV ETDVSLDAEDIRSEKVKV EKDESFNCKDIRAKVAA 340 350 AVVLYVENERVCVPFIL VLRLFVNNDENGVPFIL VLRLFVNNDENGVPFIL AAAMFINNAENGVPFIL AAAMFINNAENGVPFIL AAAMFINNAENGVPFIL AAAMFINNAENGVPFIL AAAMFINNAENGVPFIL AAAMFINNAENGVPFVV 410 TKMMTKKPGMFPNPE IKLNVSSCMKPQTE LRLTAETPGVLNDTH LKLTAETPGVLNDTH	LK CISEVQ. LRSIEPLT. LRSVNPVT. LRSVNPVT. LRSNKPLR. LKHIEPVT. LRSMKPLK. LKSLRLDRS FNALKIVNE S60 RCGMALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER CALENT SGR LAR (CLENC) ESE DLTYG CTEDLTYE	300 ANNYULGC IDDIUIG PADULGC PACULGC LOUVUGC LOUVUGC SNVEKTVRGC SNVEKTVRGC SNVEKTVRGC EEVNKYFVRA (AEVRVG/KEI YLGIRIQCKD LOURUGCKEI YLGIRIQCKD LOURUGCKEI YLGIRIQCKEI YLGIRIQCKEI GAEIRVCHUCKEI GAEIRVCHUCKEI GAEIRVCHUCKEI GAEIRVCHUCKEI (NEYKNVKLPDI HRYKAIKLEDDI HRYKAIKLEDDI	310 VADPNATNP TRSEDGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VGHNKGGKT. VGAGDTADYK. VGAGDTADYK. VGAGDTADYK. VGAGDTADYK. VGAGDTADYK. VGAGDTADYK. VGAGDTADYK. VGAGDTADYK. VGAGDTADYK. VGALUKTPGT 	320 310 310 310 310 310 310 310 31	9 PRGSTTATFA PRGSTTATFA PRGSTTATFA PRGSTCATFY PRG
Homo sapiens Schistosoma mansoni Trypanosoma Drucei Trypanosoma Cruzi Leishmania mexicana Arabidopsis thaliana Spinacia oleracea Escherichia coli Leuconostoc mesenteroides Homo sapiens Schistosoma mansoni Trypanosoma Cruzi Leishmania mexicana Spinacia oleracea Escherichia coli Leuconostoc mesenteroides Homo sapiens Schistosoma mansoni Trypanosoma Cruzi Leishmania mexicana	290 290 EKASTNSDDVEDEKVKV EKASTNSDDVEDEKVKV EKASNADDIRDEKVKV EKARLSPEDIRDEKVIV EKARLSPEDIRDEKVVV EKARLSPEDIRDEKVVV EKARLSAEDIRDEKVVV SPSDLSASIRDEKVKV EKASSNADIRSEKVKV EKASSNADIRSEKVKV EKASSNADIRSEKVKV EKASSNADIRSEKVKV EKASSNADIRSEKVKV EKASSNADIRSEKVKV EKASSNADIRSEKVKV EKASSNADIRSEKVKV EKASSNADIRSEKVKV EKASSNADIRSEKVKV EKASSNADIRSEKVKV EKASSNA AVVEVVENENDENGVPFIL CAVLYVKSDEKKOVFIL VIRLENNNDENGVPFIL VIRLENNNDENGVPFIL AAAMFINNARDOVPFIL AAAMFINNARDOVPFIL AAAMFINNARDOVPFIL AAAMFINNARDOVPFIL AARVONDENGVPFVV 410 TKMMTKKPGM.FFNPE IKLINVSPGM.KF0TE IKLINVSPGM.KF0TE	LKCISEVQ. LKSIEPUT. LRSVVPAN. LRHVNPVT. LKHIEPVT. LKSKKPLK. LKSKKPLK. LKSKKPLK. LKSKKPLK. LKSKAVENK RCGIALNER RAGIALNER RAGIALNER RAGIALNER RAGIALNER RAGIALEEN KAGIALEEN	300 ANNY ULGC IDDIYG PARCULGC PARCULGC ERVYYG SNVREXTYRC SNVREXTYRC SNVREXTYRC SNVREXTYRC SNVREXTYRC AF VYALIRICKD YVAIRICKD YVAIRICKD AF YVAIRICKD AF AF YVAIRICKD AF AF AF AF AF AF AF AF AF AF	310 VVGNPDGEGE. VVADPNATNPP VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTAGFAQGKT. VGAGDIFHQ AGDIFHQ	320 3100 3	9 PRGSTTATFA PRGSTTATFA PRCSTTATFA PRCSTCATFY PRGSTCPTPA PRGSTCPTPA PRGSTCPTPA PNNSVTPTA NKSSNTETFY PADSKNNTFT PADSKNNTFT IRVQPDEALVY IRVQPDEALVY IRVQPDEALVY IRVQPDEALVY IRVQPDEALVY IRVQPDEALVY IRVQPDEAL IRVQPDEA
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Schistosoma_mansoni Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei	290 290 EKASTNSDDVEDEKVKV EKOVSVNADDIRDEKVKV EKOVSVNADDIRDEKVKV EKORSLSPEDIRDEKVVV EKORSLSAEDIRDEKVVV EKORSLSAEDIRDEKVVV ETOVSLDAEDIRSEKVKV ETOVSLDAEDIRSEKVKV SPOSDLSADSIRDEKVKV EKOSSNADSIRDEVKV EKOSSNADSIRDEVKV EKOSSNADSIRDEVKV EKOSSNADSIRDEVKV EKOSSNADSIRDEVKV EKOSSNADSIRDEVKV EKOSSNADSIRDEVKV EKOSSNADSIRDEVKV EKOSSNADSIRDEVVV EKOSSNADSIRDEVVV EKOSSNADSIRDEVVV EKOSSNADSIRDEVVV EKOSSNA AVVIVVENEEN OVPPIL VIRLEVNNDEN AVVIVENEEN AAANFINNARN AIRVDIDNANDEVPIL AAAMFINNARN EVPPIL AAAMFINNARN EVPPIL AAAMFINNARN EVPPIL AAAMFINNARN EVPPIL AAAMFINARN EVPPIL AAAMFINARN EVPPIL AAAMFINARN EVPPIL AAAMFINARN EVPPIL AAAMFINARN EVPPIL AAAMFINARN EVPPIL AAAMFINARN EVPPIL AAAMFINARN EVPPIL AAAMFINARN EVPPIL AAAMFINARN EVPPIL AAAMFINARN EVPPIL AAAMFINARN EVPIL EKINVESPGMFFNPE IKLINVESPGMFPNPE IKLINVESPGMNETVPE VKITTVPCLSGDLRQTH UKITTVPCLSGDLRQTH	LKCISEVQ. LRSIEPUT. LRSIEPUT. LRHVNPVT. LRHVNPVT. LRSMKPUR. LKSIERIDRS FNALKIYNEP 260 RCGIALNER RAGIALNER RAGIALNER RAGIALNER RAGIALNER RAGIALNER RAGIALNER RAGIALNER RAGIALNER RAGIALNER RAGIALNER RAGIALNER CIEDLTYG ESEDLTYG ESEDLTYG QTEDLTYH RGDINLLYR	300 ANNY VIGC IDDIY PADCVLG PADCVLG PACVLG EDVVYG SNVREKTVRG EEVNKYFVRG EEVNKYFVRG EEVNKYFVRG KAEVRVGKEF YLGIRIOKD LODVYG SNVREKTVRG EEVNKYFVRG AEVNGK VIGC INCON SNVREKTVRG AEVNGK SNVREKTVRG AEVNGK SNVREKTVRG AEVNGK SNVREKTVRG AEVNGK SNVREKTVRG AEVNGK A	310 VACONPDGEGE. VADPNATNPP TRSEDGSIP. TASADGSIP. VAGHNKGGKT. KGHSKGNKS. GAGDTADYK. GAGDTADYK. GAGDTADYK. GAGDTADYK. GAGDIFHQ EIRPFGV EIRPFGV EIRPFGS EIRPFGS EIRPFGS 	320 31000 3100 3100 3100 3100 3100	9 PROSITATPA PROSITATPA PROSITATPA PROSITATPA PROSITATPA PROSICATPA PROSICATPA PROSICATPA PROSICATPA PROSICATPA PROSICATPA PROSICATPA PROSICATPA PROSICATPA PROSICATPA PROSICATION PROS
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Dishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Arabidopsis_thaliana Spinacia_oleracea	280 290 EKASTNSDDVEDEKVKV EKASTNSDDVEDEKVKV EKASTNSDDVEDEKVKV EKASTSPEDI PDEVVV EKASTSPEDI PDEVVV EKASTDAECIEDEKVSV ETAVSLDAEDI SEKVKV ETAVSLDAEDI SEKVKV SPSDLSADSI EDEKVKV SPSDLSADSI EDEKVKV EKASFNCKOVPII VLRLPINNDEN GVPPII VLRLPINNDEN GVPPII AAAVFINNAEN GVPPII AAAVFINNAEN GVPPIM AIRVDIDNKKA GVPPIM AAALFINNAEN GVPPIM AAALFINNAEN GVPPIM AAALFINNAEN GVPPIM AAALFINNAEN GVPPIM EKINKYPGLSCDLROTH LKLTATPGL.NNEW	LKCISEVQ. LRSIEPLT. LRSVVPAN. LRVVPVT. LRVVPVT. LRSVKPVT. LRSKFPLR. LKSIEPLT. LSSKFPLR. LSSERIDRS FNALKIVNE S60 RCGMALNER RAGEALEBR RAGEALEBR RAGEALEBR RAGEALHTRI RAGEALHTRI RAGEALHTRI RAGEALHTRI RAGEALHTRI RAGEALHTRI RAGEALHTRI RAGEALHTRI RAGEALHTRI RAGEALHTRI RAGEALHTRI RAGEALHTRI RAGEALHTRI RAGEALHTRI REFEDLTYER QTEDDLTYER QTEDDLTYER RAGEALTYER REFEDL	300 ANNYULGC IDDIYLGC PADCVLGC PACVLGC PACVLGC LDVVYGC SNVREKTVRGC BEVNKYFVRG BEVNKYFVRG BEVNKYFVRG LDDIRCGKD. LDIRCGKD. LDIRCGKD. GAEIRVCHUK	310 VACONPDGEGE. VADPNATNP TRSEDGSIP. VTASADGSIP. VTASADGSIP. VGASKGNKS. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGASCAGSIP. VGSCAGS	320 310 310 310 310 310 310 310 31	9 PRGSTTATFA PRGSTTATFA PRGSTTATFA PRGSTCATFY PRG
Homo sapiens Schistosoma mansoni Trypanosoma brucei Trypanosoma cruzi Leishmania mexicana Arabidopsis thaliana Spinacia oleracea Escherichia coli Leuconostoc mesenteroides Homo sapiens Schistosoma mansoni Trypanosoma cruzi Leishmania mexicana Arabidopsis thaliana Spinacia oleracea Escherichia coli Leuconostoc mesenteroides Homo sapiens Schistosoma mansoni Trypanosoma cruzi Leishmania mexicana Arabidopsis thaliana Spinacia oleracea Escherica oleracea Escherica oleracea Escherica oleracea Escherichia coli	290 290 EKDVSVNADDIZDEKVKV EKDVSVADDIZDEKVKV EKDRSLSPEDIZDEKVKV EKDRSLSPEDIZDEKVVV EKDRSLSPEDIZDEKVVV EKDRSLSAEDIZDEKVVV ETDVSLDAEDIZDEKVVV SPSDLSASIZDEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNA	LKCISEVQ. LKSIEPUT. LRSVVPAN. LRHVNPVT. LKHIEPVT. LKSKFPLR. LKSKFPLR. LKSKFPLR. LKSKFPLR. LKSKFPLR. LKSKFPLR. LKSLRRIDRS PNALKIYNER RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR RACKAVERR LARK LARK LEST LINK LEST LINK LEST LINK LEST LINK LEST LINK LEST LINK LEST LINK LEST LINK LEST LINK LINK LEST LINK	300 ANNY ULGC IDDIYG PARCULGC PARCULGC ERVYYG SNVREXTYRC SNVREXTYRC SNVREXTYRC SNVREXTYRC SNVREXTYRC AF VYALIRCORN VYALIRCORN VYALRICKD YVALRICKD SEVVYY VYALRICKD AF SEVVYY VYALRICKD MRYKNVKLPDD AF AF AF AF AF AF AF AF AF AF	310 VVQNPDGEGE. VVADPNATNPP VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTAGFAQGKT. VGAGDIFHQ  EIRPFGV  EIRPFGV  EIRPFGV  EIRPFGV  EIRPFGS 	320 ATKCILDD PTW ATKCILDD PTW ATKCILDD PSW CLED PTW CLED PTW CLED ETW CLED ETW YSCIDD PTW YSCIDD PTW YSCIDD PTW YSCIDD PTW YSCID PT	9 PRGSTTATFA PRGSTTATFA PRCSTTATFA PRCSTCATFY PRCSKCATFY PRGSTCPTFA PRGSTCPTFA PNNSVTPTA NKSSNTETFY PADSKNNTFT IRVQPDEALVY IRVQPDEALVY IRAQDSEAMY IRAQDSEAMY IRAQDSEAMY IRAQDSEAMY IRAQDSEAMY IRQQEDEALVY IRQQEDEALVY IRQQEDEALVY IRQQEDEAL IIIDPKGAIE 6 0 REAWRIFTPL DAAWRIYTPL DAAWRIYTPL DAAWLFTPL DAAWLFTPL DAAWLFTPL DAAWLFTPL
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania mexicana Arabidopsis thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Scherichia_coli Leuconostoc_mesenteroides	290 290 EXASTNSDDVEDEKVKV EKOVSNADDIZDEKVKV EKOVSNADDIZDEKVKV EKORSLSAEDIZDEKVVV EKORSLSAEDIZDEKVVV ERORIDABCIRDEKVSV ETOVSLDAEDIRDEKVSV SPOSDLSADSIZDEKVKV EKOSSNASSIZDEKVKV SPSSSIZDEKVKV SPSSSIZDEKVKV EKOSSNASSIZDEKVKV SPSSSIZDEKVKV EKOSSNASSIZDEKVKV EKOSSNASSIZDEKVKV SPSSSSIZDEKVKV SPSSSIZDEKVKV SPSSSSIZDEKVKV SPSSSSIZDEKVKV SPSSSIZDEKVKV SPSSSSIZDEKVKV SPSSSSIZDEKVKV SPSSSSIZDEKVKV SPSSSSIZDEKVKV SPSSSSIZDEKVKV SPSSSSSSZZZZZZZZZZZZZZZZZZZZZZZZZZZZZZ	LKCISEVQ. LKSIEPUT. LRIVNPVT. LRIVNPVT. LKHIEPVT. LKSIERUT. LKSIERUT. LKSIERUT. LSSKEPUT. LKSIERUT. LSSKEVT. LKSLERUT. SCORTALNER ACTALNER KACKALNER K	300 ANNY VIGC IDDIYG PADCVLG PADCVLG PACVLG EDVVYG SNYREXTYRG EVNKYFVRG EVNKYFVRG EVNKYFVRG SNVREKTYRG SNVREKTYRG SNVREKTYRG AEVNYG KEF YLGIRICKD YLGIRICKD SAEIRVORHV CSEVVYY KTF GAEIRVORHV CSEVVYY KTF GTRVDIV KAG 430 NRYKNVKLPDJ HRYKAIKLPDJ TRYD.VRLPDJ	310 VVGNPDGEGE. VADPNATNPP TRSEDGSIP. TASADGSIP. VTASADGSIP. XGGNKGKT. XGAGDTADYK. GAGDTADYK. GAGDTADYK. TAGFAQGKK. GAGDTADYK. EIRPFGV EIRPFGV EIRPFGV EIRPFGV EIRPFGV EIRPFGS EIRPFGS EINLFKE TFNFGS EINLFKE TFNFGS EINLFKE TFNFGS 	320 31000 3100 3100 3100 3100 3100 3100 3100 3100 3100 3100	9 PRGSTTATPA PRCSTTATPA PRCSTTATPA PRCSTTATPA PRCSTCATPA PRCSTCATPA PR
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia oleracea Escherichia coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Schistosoma_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides	290     290       EKD VS WADDINEDEKY VKY       EKD VS WADDINEDEKY VKY       EKD VS WADDINESKY VY       EKD VSLOBELS PEDINEVY VY       EKD VSLOBELS PEDINEVY VY       EKD VSLOBELS PEDINEVY VY       ETD VSLOBEN VSV       ETD VSLOBEN VSV       SP SOLSADSIEDEKY VY       ETD VSLOBEN VSV       SP SOLSADSIEDEKY VY       BKD RSLADSIEDEKY VY       SP SOLSADSIEDEKY VY       VENEVY VY       VENEVY VY       VENEVY VY       VENEVY VY       AA ALPINNNEND VA GVPPIN AAALPINNNEND VY VY       AGELOPDE PREEVY VY       410       TKMMTKKPGM	LKCISEVQ. LRSIEPLT. LRSIEPLT. LRSVNPVT. LRSNKPUR. LRSMKPUR. LSSKKPUR. LSSKKPUR. LSSKKPUR. S60 RCGMALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER CALLANS C	300 ANNY VLGC IDDIY PADCVLGC PADCVLGC PACVLGC EDVVYGC SNVREKTVRGC EEVNKYFVRG EEVNKYFVRG ULDIRIG KD LODVYGC SNVREKTVRGC EEVNKYFVRG 100 KD LODVYGC SNVREKTVRGC HDVYGC SNVREKTVRGC HDVYGC SNVREKTVRGC HDV SNVREKTVRGC SNVREKTV	310 VADPNATNPY TRSEDGSIP. VADPNATNPY TRSEDGSIP. VTASADGSIP. VGHNKGGKT. KGHSKGNKS. GAGDTADYK. GAGDTADYK. GAGDTADYK. CAGDIFHQ C HILPGS C BIRPFGV EIRPFGV EIRPFGV EIRPFGV C EVHPYGS C EVHPYGS C EVHPYGS C ELNLFKE C ELLLDAIEC C SLINDALC C SLINDALC C SLINDALC	320 ATKG LDD PTW ASLS TOD PSW ASLS TOD PSW GLED PTW GLED TW GLED TW GLE	9 PRGSTTATFA PRCSTTATFA PRCSTTATFA PRCSTCATFY PRCSKCATFY PRCSKCATFY PRCSTCATFA PRC
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_truzi Leishmania mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Spinacia oleracea Escherichia_coli Leuconostoc_mesenteroides	290 290 EKDVSVNADDIZDEKVKV EKDVSVADDIZDEKVKV EKDRSLSPEDIZDEKVKV EKDRSLSPEDIZDEKVVV EKDRSLSPEDIZDEKVVV EKDRSLSAEDIZDEKVVV ETDVSLDAEDIZDEKVVV SPSDLSADSIZDEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNADJENEKVKV EKDSSNAJENEKVKV EKDSSNADJENEKVKV EKDSSNAJENEKVKV EKDSS	LKCISEVQ. LKSIEPUT. LRSVVPAN. LRSVVPAN. LKHIEPVT. LKHIEPVT. LKSIKPLR. LKSIKPLR. LKSIKPIDR S60 RCGIALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER LKAKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER RAGKALNER LKAKALNER LKAKALNER LKAKALNER LKAKALNER LKAKALNER LKAKALNER LKAKALNER LKAKALNER RAGKAL	300 ANNY ULGC IDDIY PARCULGC PARCULGC PARCULGC ERVYUGC SNVREXTYRC SNVREXTYRC SNVREXTYRC BEVNKYFYRA VYALIRICKD YYALRICKD YYALRICKD YYALRICKD YYALRICKD ABERVOZHU RAEIRVOZHU RAEIRVOZHU RAEIRVOZHU RAEIRVOZHU RAEIRVOZHU RAEIRVOZHU RAEIRVOZHU SEYVYYKA 430 NRYKNVKLPDJ SRYP.REIPDJ SRYP.REIPDJ SRYP.REIPDJ EIPNQTHLADJ DEDK.KNTPEI	310 VACONPDGEGE. VADPNATNP VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. CACONF. SCONTACT SCON	320 ATKCILDDPT ATKCILDDPSV ATKCILDDPSV CLEDPTV CLEDTV CLEDTV CLEDTV CLEDTV YSCIDDPTV YSCIDDTV YSCIDP	9 PRGSTTATFA PRGSTTATFA PRCSTTATFA PRCSTCATFY PRGSTCPTFA PRGSTCPTFA PNRSVTFFA NKSSNTETFY PADSKNNTFTFA NKSSNTETFY PADSKNNTFTFA NKSSNTETFY PADSKNNTFTFA NKSSNTETFY PADSKNNTFTFA 1 RAQPSEAMY 1 RAX 1 R
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania mexicana Arabidopsis thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis thaliana Spinacia oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Escherichia_coli Leuconostoc_mesenteroides	290     290       EXASTNSDDVEDEKVKV       EKCVSVNADDIZDEKVKV       EKCVSVNADDIZDEKVVV       EKDVSDZDEKVVV       EKDVSDZDEKVVV       EKDVSDZDEKVVV       EKDVSDZDEKVVV       EKDVSDZDADIZDEKVVV       ETVSLDAEDIRDEVVV       ETVSLDAEDIRDEVVV       SPSDJSADSIZDEKVVV       ETVSLDAEDIRDEVVV       ETVSLDAEDIRDEVVV       SPSDJSADSIZDEVVV       EKDSFNDKDIZAANA       340     350       AVVLYVENEERDOVVV       EKDSFNDKDIZAANA       CAVLYVENEERDOVVV       VIRLENNNDRMGVPFIL       VARLNNNRAKGVPFIL       AAAMFINNARN GVPFIL       AAAMFINNARN GVPFYL       AAAMFINNARN GVPFYL       AAAMFINNARN GVPFYL       AAAMFINNARN GVPFYL       AGBLOPDLPWEGVV       LKLNVSPGMFFNPE       IKKMTKKPGMFFNPE       IKKMTVSPGLSCHNRLDUP       LKLNNVSPGLSCHNRLDUP       LKLNNVSPGLSCHNRLDUP       LKLNNVSPGLSCHNRLDUP       LKLNNVSPGLSCHNRLDUP       LKLNNVSPGLSCHNRLDUP       LKLNNVSPGLSCHNRLDUP       LKLNNKSVEDAFNTR </td <td>LKCISEVQ. LKSIEPUT. LRSVVPAN. LRSVVPAN. LKSVVPAN. LKSVFPVT. LKSVFPVT. LKSLFRIDRS PNALKIYNEP 260 RCGALNER RAGYALNER LANG LA</td> <td>300  ANNY VIGC  IDDIYG  PADCVLG  PADCVLG  PACVLG  EDVVYG SNVREXTVRG EVNKYFVRG EVNKYFVRG BEVNKYFVRG SNVREXTVRG SNVREXTVRG SNVREXTVRG AEVRYGYF VIGIRIG SNVREXTVRG AEVRYGYF VIGIRIG SNVREXTVRG AEVVYY SEV SEVVYY SEV SEVVYY SEV SEV SEV SEV SEV SEV SEV SEV</td> <td>310 VACONPDGEGE. VADPNATNPP TRSEDGSIP. TASADGSIP. VAGHNKGGKT. KGHSKGNKS. GAGDTADYK. GAGDTADYK. BIAGAOGKK. GAGDTADYK. BIRPFGV BIRPFGV BIRPFGV BIRPFGV CHILLFFG GAGNTADYKSPAT PGNLYKKSPAT PGNLYKKFFGT CHILLDVFCG MIRLILDVFCG MIRLILDVFCG MIRLILDVFCG MISLINDALLG MISLINDALLG MISLINDALLG MISLINDALLG MISLLETMRG MIRLLEMMRG</td> <td>320 ATKCILDD PTW ATKCILDD PSW ATKCILDD PSW CLED PTW CLED ETW CLED ETW YPGIDD PTW .YSGIDD PTW .YSGIDD PTW .YSGID PTW .YSGID PTW .YSGID PTW YSGID PTW YSGI</td> <td>9 PRGSTTATPA PRCSTTATPA PRCSTTATPA PRCSTCATPA PRCSTCATPA PRGSTCPTPA PRGSTCPTPA PNNSVTPA NKSSNTETPY PADSKNNTPT AUX SSNTETPY AUX IRVQPDEAVY IRVQV IRVQV IRVQV IRVQV IRVQV IRVQV IRVQV IRVQV IRVQV IRVV IRVQV IRVV</td>	LKCISEVQ. LKSIEPUT. LRSVVPAN. LRSVVPAN. LKSVVPAN. LKSVFPVT. LKSVFPVT. LKSLFRIDRS PNALKIYNEP 260 RCGALNER RAGYALNER LANG LA	300 ANNY VIGC IDDIYG PADCVLG PADCVLG PACVLG EDVVYG SNVREXTVRG EVNKYFVRG EVNKYFVRG BEVNKYFVRG SNVREXTVRG SNVREXTVRG SNVREXTVRG AEVRYGYF VIGIRIG SNVREXTVRG AEVRYGYF VIGIRIG SNVREXTVRG AEVVYY SEV SEVVYY SEV SEVVYY SEV SEV SEV SEV SEV SEV SEV SEV	310 VACONPDGEGE. VADPNATNPP TRSEDGSIP. TASADGSIP. VAGHNKGGKT. KGHSKGNKS. GAGDTADYK. GAGDTADYK. BIAGAOGKK. GAGDTADYK. BIRPFGV BIRPFGV BIRPFGV BIRPFGV CHILLFFG GAGNTADYKSPAT PGNLYKKSPAT PGNLYKKFFGT CHILLDVFCG MIRLILDVFCG MIRLILDVFCG MIRLILDVFCG MISLINDALLG MISLINDALLG MISLINDALLG MISLINDALLG MISLLETMRG MIRLLEMMRG	320 ATKCILDD PTW ATKCILDD PSW ATKCILDD PSW CLED PTW CLED ETW CLED ETW YPGIDD PTW .YSGIDD PTW .YSGIDD PTW .YSGID PTW .YSGID PTW .YSGID PTW YSGID PTW YSGI	9 PRGSTTATPA PRCSTTATPA PRCSTTATPA PRCSTCATPA PRCSTCATPA PRGSTCPTPA PRGSTCPTPA PNNSVTPA NKSSNTETPY PADSKNNTPT AUX SSNTETPY AUX IRVQPDEAVY IRVQV IRVQV IRVQV IRVQV IRVQV IRVQV IRVQV IRVQV IRVQV IRVV IRVQV IRVV
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Irypanosoma_cruzi Leishmania mexicana Arabidopsis_thaliana Spinacia oleracea Escherichia coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Irypanosoma_brucei Deishmania_mexicana Arabidopsis_thaliana Spinacia oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Irypanosoma_brucei Irypanosoma_brucei Irypanosoma_brucei Irypanosoma_brucei Irypanosoma_brucei Escherichia_coli Leishmania_mexicana Arabidopsis_thaliana Spinacia oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens	280     290       EKDVSVNADDIZDEKVKV       EKDVSVNADDIZDEKVKV       EKDVSVNADDIZDEKVKV       EKDVSVADDIZDEKVVV       EKDVSLDAEDIZDEKVVV       EKDVSLDAEDIZDEKVVV       ETDVSLDAEDIZDEKVVV       ETDVSLDAEDIZDEKVVV       SPSSDLSADSIZDEKVKV       ETDVSLDAEDIZSEKVKV       SPSSDLSADSIZDEKVKV       SPSSDLSADSIZDEKVKV       EKDSFNCKDIZAALNAA       340     350       AVVLYVENERDOKGVPPIL       VLRLFINNDENDENGVPPIL       VLRLHVINDENGVPPIL       VLRLHVINDENGVPPIL       VLRLHVINDENGVPPIL       AVVLYVENERD       AVDISTONANDENGVPPIL       VLRLHVINDENGVPPIL       VLRLHVINDENGVPPIL       VLRLHVINDENGVPPIL       AGELOPDLPENECVPPYL       AGELOPDLPENECVPPYL       AGELOPDLPENECVPYL       AGELOPDLPENECVPYL       LKLIATPGLLNDTH       VKITTVPGLGMRLD       IQVLNVPGLD.HKHNLQ       LKLINASVEDAFNTR       470     480       LKLUNASVEDAFNTR	LK CISEVQ. LRSIEPLT. LRSIEPLT. LRSWPVT. LRSWFPVT. LRSMKPLK. LSSLRIDRS PNALKIYNEP 260 RCGMALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER RAGTALNER CALLANER C	300 ANNY VLGC IDDIY PADCVLGC PACVLGC PACVLGC EDVVYGC SNVREKTVRGC EEVNKYFVRG EEVNKYFVRG 1000000000000000000000000000000000000	310 VACONDATNP TRSEDGSIP. VAADAATNP TRSEDGSIP. VAASADGSIP. VAASADGSIP. VAASADGSIP. VAASADGSIP. VAASADGSIP. VAASAGSIP. VAA	320 ATKCILDDPTV ASLETTDPSV ASLETTDPSV ASLETDPSV CLEDPTV CLEDPTV CLEDTV CLEDTV CLEDTV CLEDTV YSCIDDPTV .VSCIDDPTV .VSCIDPTV .VSCIDPTV SV	9 PRGSTTATFA PRGSTTATFA PRGSTTATFA PRGSTCATFY PRGSKCATFY PRGSTCFFA PRGSTCFFA PRGSTCFFA PRGSTCFFA PRGSTCFFA PRGSTCFFA PRGSTCFFA PRGSTCFFA PRGSTCFFA PRGSTCFFA PRGSTCFFA PRGSTCFFA PRGSTCFFA PRGSTCFFA PRGSTCFFA CO CO CO CO CO CO CO CO CO CO
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_truzi Leishmania mexicana Arabidopsis_thaliana Spinacia oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Spinacia oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni	290 290 EKD VENNADDIEDEKVEV EKD VENADDIEDEKVEV EKD RELSPEDIEDEKVEV EKD RELSPEDIEDEKVEV EKD RELSPEDIEDEKVEV EKD RELDAEDIEDEKVEV EKD RELDAEDIEDEKVEV EKD RELDAEDIEDEKVEV SPOSDISADSIEDEKVEV SPOSDISADSIEDEKVEV EKD RESPNDEDIE AAKNAA 340 350 AVVEYVENER MOGVPFIL CAVLYVENER MOGVPFIL CAVLYVENER MOGVPFIL VIRLENNNDEN GVPFIL VIRLENNNDEN GVPFIL AAAMFINNAR OGVPFIL AAAMFINNAR OGVPFIL AAAMFINNAR OGVPFIL AAAMFINNAR OGVPFIL AAAMFINNAR OGVPFIL AAAMFINNAR OGVPFIL AAAMFINNAR OGVPFIL AAAMFINNAR OGVPFIL LELTATFOL. LNDTH UKLING VPGLG. MELD IQULN VPGLD.HENNE KENNE VFRG. MELD IQULN VPGLD.HENNE	LKCISEVQ. LKSIEPUT. LRSVVPAN. LRSVVPAN. LKHIEPVT. LKHIEVT. LKSIKPLR. LKSIKPLR. LKSIKPLR. LKSIKPUR. LKSIKPUR. SGO RCGIALNER RACKALNER RACKALNER RACKALNER RACKALNER RACKALNER RACKALNER RACKALNER RACKALNER RACKALNER RACKALNER RACKALNER RACKALNER RACKALNER RACKALNER RACKALNER LKACKALNER RACKALNER RACKALNER RACKALNER LISS LI	300 ANNY ULGC IDDIY PARCULGC PARCULGC PARCULGC ERVYUGG SNVREXTYRC SNVREXTYRC SNVREXTYRC BEVNKYFYRA LCDVYGC SNVREXTYRC BEVNKYFYRA LCDVYGC SNVREXTYRC AF 1000 SNVREXTYRC AF	310 VVQNPDGEGE. VVADPNATNPP VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. STAGFAQGKT. VTAGFAQGKT. VGNLYKGPAQGK. SGAGDTADYK.	320 ATKCILDPTW ATKCILDPTW ATKCILDPTW CLEPTW CLEPTW CLEDTV CLEDTV CLEDTV CLEDTV YSCILDPTW YSCILDPTW YSCILEECA PLEEADW QENEY   	9 PRGSTTATFA PRGSTTATFA PRCSTTATFA PRCSTCATFY PRCSKCATFY PRGSTCPTFA PRGSTCPTFA PNNSVTFA NKSSNTETFY PADSKNNTFTFA NKSSNTETFY PADSKNNTFTFA NKSSNTETFY IRVQPDEALVY IRVQPDEALVY IRVQPDEALVY IRAQPSEAMY IRQQPSEAMY IRVQPDEALVY IRQQPSEAM
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei	290 290 EXASTNSDDVEDEKVKV EKOVSNADDIZDEKVKV EKOVSNADDIZDEKVKV EKORSLSAEDIZDEKVVV EKORSLSAEDIZDEKVVV EKORTUABCIEDEKVSV ETOVSLDAEDIRDEKVKV SPOSDLSADSIZDEKVKV EKOSSNASIZDEKVKV EKOSSNASIZDEVK EKOSSNASIZDEVK EKOSSNAS	LKCISEVQ. LKCISEVQ. LRSIEPUT. LRSIEPUT. LRSNKPVT. LKHIEVVPAN. LKSIERIDRS FNALKIYNEI S60 RCGALNERI RAGYALNERI REGINALIYEN REGINAL	300 ANNY VIGC IDDIY PADCVLGC PADCVLGC PACVLGC EDVVYGC SNVREXTVRGC SNVREXTVRGC SNVREXTVRGC SNVREXTVRGC SNVREXTVRGC SNVREXTVRGC SNVREXTVRGC SNVREXTVRGC SNVREXTVRGC SNVREXTVRGC SNVREXTVRGC SNVREXTVRGC SNVREXTVRGC ABEIRVGC REVVYY SEVVVY SEVV SEVVVY SEVVV SEVVVY SEVVV SEVVVY SEVVV	310 VVGNPDGEGE. VVADPNATNPP TRSEDGSIP. VTASADGSIP. VTASADGSIP. VGHSKGKT. VGAGDTADYK. GAGDTADYK. GAGDTADYK. CAGDIFHQ CAGDIFHQ CBIRPFGV CBIRPFGV CBIRPFGV CBIRPFGV CBIRPFGV CBIRPFGV CBIRPFGV CBIRPFGS CBIRPFG	320 ATKCILDD PTW ATKCILDD PSW ATKCILDD PSW CLEDPTW CLEDETW CLEDETW CLEDETW .YSCILDETW .YSCIDDPTW .YSCIDDPTW .YSCIDPTW .YSCIDPTW .YSCIDPTW .YSCIDPTW YSCIDPTW	9 PRGSTTATFA PRGSTTATFA PRCSTTPTYV PRCSTCPTFA PRGSTCPTFA PNNSVTPTPA NKSSNTETPV PADSKNNTFTP NKSSNTETPV PADSKNNTFTP IRVQPDEAVY IRVQDEAVY IRVQDDEAVY IRVVDDEAVY IR
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Frypanosoma_brucei Trypanosoma_brucei	290     290       EKDVSVNADDIRDEKVKV       EKDVSVNADDIRDEKVKV       EKDVSVNADDIRDEKVVV       EKDVSVADDIRDEKVVV       EKDVSLAEDIRDEKVSV       EKDVSLAEDIRDEKVSV       EKDVSLAEDIRDEKVSV       ETDVSLDAEDIRDEKVSV       ETDVSLDAEDIRDEKVSV       ETDVSLDAEDIRDEKVSV       ETDVSLDAEDIRDEKVSV       ETDVSLDAEDIRDEKVSV       SPSSDLSADSIRDEKVKV       EKDSFNCKDIRAANA       340     350       AVVLYVENERDOKGVPPIL       VLRLFINNDENDENDEVPYIL       VLRLFINNDENDENDEVPYIL       VLRLFINNDENDEVPYIL       AAMFINNDENDEVPYIL       AAAMFINNDENDEVPYIL       AAAMFINNDENDEVPYIL       AAAMFINNDENDEVPYIL       AAAMFINNDENDEVPYIL       AAAMFINNENDENDEVPYIL       AAAMFINNENDENDEVPYIL       AAAMFINNENDENDEVPYIL       AAAMFINNENDENDEVPYIL       AAAMFINNENDENDEVPYIL       AAALFINNENDENDEVPYIL       AAALFINNENDENDEVPYIL       AAALFINNENDENDEVPYIL       AAALFINNENDENDEVPYIL       AAALFINNENDENDEVPYIL       AAALFINNENDENDEVPYIL       AAALFINNEN	LK CISEVQ. LRSIEPLT. LRSIEPUT. LRSWPVT. LRSWKPVT. LRSMKPLK. LSSKKPLK. LSSKKPLK. LSSKKPLK. LSSKKPLK. SG Q CG Q ESC Q	300 ANNY VLGC IDDIY PADCVLGC PACVLGC PACVLGC EDVVYGC SNVREKTVRGC EEVNKYFVRA 370 KAEVRVCKEF VLGIRIOKD LODVYGC SNVREKTVRGC EEVNKYFVRA (AEVRVCKEF VLGIRIOKD LODVYGC SNVREKTVRGC IDDIY SNVREKTVRGC 430 NRYKNVKLPDD ALIRVCKH (SEVVYYKTF (SEVVYYKTF (SEVVYYKTF (SEVVYYKTF (SEVVYYKTF (SEVVYYKTF SO LKKRVGFQYEG LCKKAGFQYEG LCKKAGFQYEG LCKKAGFQYEG LCKKAGFQYEG LCKKAGFQYEG LCKKAGFQYEG LCKKAGFQYEG LCKKAGFQYEG LCKKAGFQYEG LCKKAGFQYEG LCKKAGFQYEG LCKKAGFQYEG SO CCC	310 VACONPDGEGE. VADPATNP TRSEDGSIP. TRSEDGSIP. VASADGSIP. KGHSKGNKS. TAGAQGKT. GAGDTADYK. GAGDTADYK. GAGDTADYK. CAGDIFHQ C AGDIFHQ C BIRPFGV EI	320 ATKCILDD PTW ASLS TOD PSW ASLS TOD PSW GLEDPTW GLEDTW GLEDTW .YPGTDD PTW .YSG TOD PTW .YSG TOD PTW .YSG TOD PTW .YSG TOD PTW .YSG TOD PTW .STST PREFY AAQRNEJU STQRNEJV AAQRNEJU STQRNEJV AAQRNEJU STQRNEJV SWQDLPQNKJT EQE.AQESVIS 450 450 450 450 450 450 450 450	9 PRGSTTATFA PRGSTTATFA PRGSTTATFA PRGSTCATFY PRGSKCATFY PRGSTCFFA PRGSTCFFA PNNSVTFA NKSSNTETFY PADSKNNTFT PADSKNNTFT IRVQPDEAVY IRAQPSEAMY IRAQPSE
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_truzi Leishmania mexicana Arabidopsis_thaliana Spinacia oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Schistosoma_mansoni Leishmania_mexicana Arabidopsis_thaliana Spinacia oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei	290 290 EKDVSVNADDIZDEKVKV EKDVADDIZDEKVKV EKDRSLSPEDIZDEKVKV EKDRSLSPEDIZDEKVVV EKDRSLSPEDIZDEKVVV EKDRSLSAEDIZDEKVVV ERDVSLDAEDIZDEKVVV SPSDLSADSIZDEKVKV EKDSSNADKDIZAAKNAA 340 350 AVVEYVENEMOCVPTIL CAVLYVKSDEKKOVPTIL CAVLYVKSDEKKOVPTIL CAVLYVKSDEKKOVPTIL VLRLENNNDENGVPTIL VLRLENNNDENGVPTIL VLRLENNNENGVPTIL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVPTYL AAAMFINNARDOVTYTYL AACVTYLYSSPGMFFNPE IKLNVKSPGMFFNPE IKLNVKSPGLLNDTH VKITKVPGLGMRLD LKINNKVPGLD.HKHNLQ LKLNALSVEDAFNTR	LKCISEVQ. LKCISEVQ. LRSIEPUT. LRSIEPUT. LRSIEVT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. Social LRSMEPUT. LRSMEPUT. LRSMEPUT. Social LRSMEPUT. LRSMEPU	300 ANNY UGC IDDIY PADCULG PADCULG PADCULG PADCULG PADCULG PADCULG SNVEXTYRC SNVEXTYRC SNVEXTYRC SNVEXTYRC SNVEXTYRC ACCOUNTS LODY VYAGE SNVEXTYRC SNVEXTYRC ACCOUNTS LODY VYAGE SNVEXTYRC ACCOUNTS LODY SNVEXTYRC LODY VYAGE SNVEXTYRC LODY SNVEXTYRC LODY LO	310 VAUNDATNPP VADPNATNPP VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTAGFAQGKT. VGAGDIFHQ SCORPTSIC VIENT VGALVKSPAT VGALVKSPAT VGALVKSPAT VGALVKSPAT VGALVKSPAT VGALVKSPAT VSLIHEALLOVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLILDVFCC VIENLIDVFCC	320 ATKCILDPTV ATKCILDPTV ATKCILDPTV CLEPTV CLEPTV CLEDTV CLEDTV CLEDTV CLEDTV YSCILDPTV YSCILDPTV YSCILEECA PLEECA PLEECA PLEECA   	9 PRGSTTATFA PRGSTTATFA PRCSTTATFA PRCSTCATFY PRCSKCATFY PRGSTCPTFA PRGSTCPTFA PNNSVTFFF PNNSVTFFF PNNSVTFFF PADSKNNTFFF IRVQPDEALVY IRVQPDEALVY IRVQPDEALVY IRAQPSEAMY IRAQPSEAMY IRQQPDEALVY IRQQPDEALVY IRQQPDEALVY IRQQPDEALVY IRQQPDEALVY IRQQPDEALY IRQQPSEAM
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_cruzi Leishmania mexicana Arabidopsis thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Arabidopsis thaliana Spinacia oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Escherichia_coli Leuconostoc_mesenteroides Bochistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei	290 290 EKASTNSDDVEDEKVKV EKOVSNADDIZDEKVKV EKOVSNADDIZDEKVKV EKORSLSAEDIZDEKVKV EKORSLSAEDIZDEKVVV EKORTUDABCIRDEKVVV ETOVSLDABDIRSEKVKV SPOSDLSADSIZDEKVKV ETOVSLDABDIRSEKVKV ETOVSLDABDIRSEKVKV ETOVSLDABDIRSEKVKV ETOVSLDABDIRSEKVKV EKOSONSJENEKV EKOSONSJENEKVSE ELENEKVKV EKOSONSJENEKVKV EKOSONSJENEKVSE ENKOSONSJENEKV	LKCISEVQ. LKCISEVQ. LRSIEPUT. LRSIEPUT. LRSIEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. LRSMEPUT. SCIENCE LLSSEE LL	300 ANNY VIGC IDDIYG PADCVLGC PACVLGC PACVLGC EDVVYGC SNVREXTVRG SNVREXTVRG EEVNKYFVRG EEVNKYFVRG SNVREXTVRG KAEVRYGTKEF YLGIRICKD YLGIRICKD YLGIRICKD YLGIRICKD SNVREXTVRG 430 NRYKNVKLPD ALEIRYGRHU CSEVVYY KTF GTRVDIVKAC 430 NRYKNVKLPD TRYD.VTLPDD CCKKAGQYSG FVDDTGFRHOR LCKKAGQYSG FVDDTGFRHOR LASKYNVRWG	310 VACONPDGEGE. VADPDATNPP TRSEDGSIP. TASADGSIP. VAGHNKGGKT. KGHSKGNKS. GAGDTADYK. GAGDTADYK. GAGDTADYK. GAGDTADYK. BIRDFGV 	320 320 ATKCILDD PTV ATKCILDD SV ATKCILDD SV CLED PTV CLED ETV CLED ETV CLED ETV .YSCILDD PTV .YSCILD PTV .YSCILD PTV .YSCILD PTV .YSCILD PTV .YSCILD PTV .YSCILD PTV .YSCILD PTV .YSCILD PTV .YSCILD PTV TORNEJV ATORNEJV ATORNEJV ATORNEJV ATORNEJV ATORNEJV ATORNEJV ATORNEJV ATORNEJV TO	9 PRGSTTATFA PRGSTTATFA PRCSTTATFA PRCSTTATFA PRST
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_cruzi Leishmania mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Jeishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_cruzi Leishmania_mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli	290 290 EKASTNSDDVEDEKVKV EKASTNSDDVEDEKVKV EKASKNADDIRDEKVKV EKASLSPEDIRDEVVV EKARLSPEDIRDEVVV EKARLSPEDIRDEVVV ERARTDAECIEDEKVSV ETVSLDAEDIRSEKVKV ETVSLDAEDIRSEKVKV SPSSDSADSIRDEVKV EKASSNO AVVLYVENEENDOVPIL VLRLFUNNDRNGVPFIL VLRLFUNNDRNGVPFIL VLRLFUNNDRNGVPFIL AAAMFINNARNGVPFIL AAAMFINNARNGVPFIL AAAMFINNARNGVPFIL AAAMFINNARNGVPFIL AAAMFINNARNGVPFIL AAAMFINNARNGVPFIL AAAMFINNARNGVPFIL AAAMFINNARNGVPFIL AAAMFINNARNGVPFIL AAAMFINNARNGVPFIL AAAMFINNARNGVPFIL AAAMFINNARNGVPFIL AAAMFINNARNGVPFIL AAAMFINNARNGVPFIL AAAMFINNARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINNARNGVPFIL AAAFINNARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNGVPFIL AAFINARNGVPFIL AAAFINARNGVPFIL AAFINARNGVPFIL AAAFINARNGVPFIL AAAFINARNOVPFIL AAAFINARNGVPFIL AAAFINARNOVPFIL AAAFINARNOVPFIL AAAFINARNOVPFIL AAAFINARNOVPFIL AAAFINARNOVPFIL AAAFINARNOVPFIL AAAFINARNOVPFIL AAAFINARNOVPFIL AAAFINARNOVPFIL AAAFINARNOVPFIL AAAFINARNOVPFIL AAAFINARNOVPFIL AAAFINARNOVPFIL AAAFINARNOVPFIL AAFINARNOVPFIL AAAFINARNOVPFIL AAFINAR	LK CISEVQ. LR SIEPLT. LR SIEPLT. LR SIEPLT. LR SIEPLT. LR SIEPLT. LR SIEPLT. LR SIEPLT. LR SIEPLT. LR SIEPTT. LR SIEPTT. LR SIEPTT. LR SIEPTT. LR SIEPTT. LR SIEPTT. LR SIEPTT. LR SIEPTT. LR SIEPTT. LR SIEPTT. SIERT. 260 RCGMALNER RAGIALER RAGIALER RAGIALER RAGIALER RAGIALER RAGIALER LA VERE RAGIALER RAGIALER LA VERE LA V	300 ANNY VLGC IDDIY PADCVLGC PADCVLGC PACVLGC PACVLGC EDVY VGC SNVREKTVRGC EEVNKY FVRA 370 KAEVRVCKEF VLGIRIOKD VVGCKEF VLGIRIOKD SNVREKTVRGC EEVNKY FVRA 430 NRYKNVKLPDJ MRYKNKKLPDJ HRYKAIKLPDJ RYP. GEIPDJ TRYDIVKAG 430 NRYKNVKLPDI KAEVRVCKEF SO0 LMKRVGF FINGTHLADJ DEDK.KNTPEI FIISGYKTT FIISGYKTT FIISGYKTT FIISGYKTT FIISGYKTT FIISGYKTT FIISGYKTT FIISGYKTT FIISGYKTT FIISGYKTT FIISGYKTT FIISGYKTT FIISGYKTT	310 VACONDATNP VADPATNP TRSEDGSIP. VASADGSIP. VASAGSIP.	320 ATKCILDD PTW ASLSTIDD PSW ASLSTIDD PSW GLED PTW GLED PTW GLED TW GLED TW GLED TW YSGTDD PTW YSGTDD PTW YSGTDD PTW YSGTDD PTW YSGTDD PTW YSGTDD PTW YSGTD PTW	9 PRGSTTATFA PRGSTTATFA PRGSTTATFA PRGSTCATFY PRGSKCATFY PRGSTCFFA PRGSTCFFA PNNSVTFFA NKSSNTETFY PADSKNNTFT PADSKNNTFT IRVQPDEAVY IRAQPSEAMY IRAQPS
Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_truzi Leishmania mexicana Arabidopsis_thaliana Spinacia_oleracea Escherichia_coli Leuconostoc_mesenteroides Homo_sapiens Schistosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Schistosoma_mansoni Trypanosoma_brucei Frypanosoma_brucei Frypanosoma_brucei Schistosoma_mansoni Trypanosoma_brucei Homo_sapiens Schistosoma_mansoni Trypanosoma_brucei Trypanosoma_brucei Homo_sapiens Schistosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Trypanosoma_brucei Escherichia_coli Leuconostoc_mesenteroides	290 290 EKDVSVNADDIEVEDEKVKV EKDRSLSPEDIEDEKVKV EKDRSLSPEDIEDEKVKV EKDRSLSPEDIEDEKVSV EKDRSLSPEDIEDEKVSV EKDRSLSAEDIEDEKVSV ETDVSLDAEDIENEKVKV SPSDLSADSIEDEKVKV EKDRSLAAEDIESEKVKV EKDRSPNDKDIEAAENAA 340 350 AVVEYVENERADGVPFIL CAVLYVKSDENKGVPFIL CAVLYVKSDENKGVPFIL VLRLENNNDENGVPFIL VLRLENNNDENGVPFIL VLRLENNNENGGVPFIL AAAMFINNARDGVPFIL AAAMFINNARDGVPFIL AAAMFINNARDGVPFIL AAAMFINNARDGVPFIL AAAMFINNARDGVPFIL AAAMFINNARDGVPFIL AAAMFINNARDGVPFIL EKLENKSPGMFFNPE IKLNVSPGMFFNPE IKLNVSPGLLSDTH VKITKVPGLGMRLD IQVLNVYPGLD.HKHNLQ LKEINAKVEDAFNTR 470 480 EHQIELEKPKPIPYIYGS LEAIERGETTYPYSAGS LHQIELEKVYPYYSE	LKCISEVQ. LKCISEVQ. LRSIEPUT. LRSIEPUT. LRSIEPUT. LRSIEVT. L	300 ANNY VLGC IDDIY PARCVLGC PARCVLGC PARCVLGC ERVVVGC SNVREXTVRC SNVREXTVRC SNVREXTVRC SNVREXTVRC SNVREXTVRC AEVNVVG NRYKNVKLPDJ YVAIRICKD LCVVC KAEVRVC HRVKNVKLPDJ YVAIRICKD TLDY YVAIRICKD LCVVC KAEVRVC KAEVRVC HRVKNVKLPDJ SNVF TLDDI SRYP.REIPDJ SRYP.REIPDJ ETPNQTHLADJ DEDK.KNVEPT SOO LMKRVGFQYEC LASKYNVRMGI LASKYNVRMGI LASKYNVRMGI LASKYNVRMGI	310 VAUNDATNPP VADPNATNPP TASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTASADGSIP. VTAGFAQGKT. VGAGDIFHQ SCORPAGE SCORPACE VITAGFAQGKK. SCORPACE VITAGFAQGKK. SCORPACE VITAGFAQGKK. SCORPACE VITAGFAQGKK. SCORPACE VITAGFAQGKK. SCORPACE VITAGFAQGKK. SCORPACE VITAGFAQGKK. SCORPACE VITAGFAQGKK. VITAGFAQGKK. SCORPACE VITAGFAQGKK. VITAGFAQGKK. SCORPACE VITAGFAQGKK. VITAGFAQGKKK. VITAGFAQGKK. VITAGFAQGKKK. VITAGFAQGKKK. VITAGFAQGKKK. VITAGFAQGKKK. VITAGFAQGKKK. VITAGFAQGKKK. VITAGFAQGKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK	320 ATKCILDPTV ATKCILDPTV ATKCILDPTV CLEPTV CLEPTV CLEDTV CLEDTV CLEDTV YSCILDPTV YSCILDPTV .VSCILEEGA PLEEADV QCKENEVV KEG.LPRNEVV AAQRNEV AAQRNEV ATORNEVV KEG.LPRNEV AAQRNEV ATORNEV KEG.LPRNEV EQE.AQESVIS SWODLPQNKT EQE.AQESVIS STNVVKKDEL ERRLISISDEL ERRLISISDEL ERRLISISDE ERRLISISSE QAL VARDEV      	9 PRGSTTATFA PRGSTTATFA PRCSTTATFA PRCSTCATFY PRCSKCATFY PRGSTCPTFA PRGSTCPTFA PNNSVTFFF PNNSVTFFF PNNSVTFFF PNNSVTFFF PNNSVTFFF PADSKNNTFFF IRVQPDEALVY IRVQPDEALVY IRVQPDEALVY IRAQDSEAMY IRAQDSEAMY IRVQPDEALVY IRVQPDEALVY IRVQPDEALVY IRVQPDEALVY IRVQPDEALY

Figura 22: Alinhamento de diferentes sequências de G6PD. Acima da linha estão sequências de organismos que são inibidos pelos esteroides, e abaixo, sequências de organismos que não são inibidos. Em destaque azul, sequência referente à ligação de NADP<sup>+</sup>, em verde, região de ligação a G6P, em vermelho, resíduos que

estão a 12Å de distância do NADP<sup>+</sup> ou G6P, e em amarelo, resíduos pertencentes à cavidade próxima ao D200, descrita na figura 21. A numeração dos resíduos é referente à sequência humana.

Foi feita uma varredura de condições de cristalização com kits inicias deste mutante, tanto para o pico referente ao dímero quanto ao tetrâmero e não se obteve formação de cristais. Também se testou na condição conhecida para HsΔG6PD, além de *seeding* com sementes de cristais da enzima selvagem, não obtendo cristais.

## Capítulo 5: Estudos cristalográficos da Hs∆G6PD

Foram feitos *screenings* iniciais da enzima na busca de novas condições de co-cristalização, mesmo porque a ligação do esteroide pode gerar mudanças conformacionais que alterem a condição de cristalização publicada para a forma com NADP ou G6P. Não se obteve nenhuma condição inicial nova, sendo explorada a condição conhecida.

Conseguiu-se reproduzir a condição de cristalização da estrutura com G6P,<sup>24</sup> constituída de 100 mM Tris-HCl pH 8,5, 200 mM MgCl<sub>2</sub>, 12% PEG 4000 e 5% glicerol (estando a proteína à 5 mg/ml com 75 µM NADP<sup>+</sup> e 60 mM G6P). Estes cristais aparecem em 1 dia, tendo o formato de bastões, sendo que na sua maioria, constituídos de bastões ocos e frágeis, apresentando muitos *clusters* difíceis de serem manipulados, havendo ainda a formação de uma película na gota com proteína precipitada (Fig.23-A). Os cristais difratavam até 6 Å e era necessário fazer *soaking* em condição crioprotetora com 15% de glicerol, o que dificultava ainda mais a manipulação dos cristais durante a coleta.

Assim, foi feito o refinamento da condição variando a concentração da proteína, tamanho da gota, pH, concentração e outros sais com magnésio, concentração de glicerol, entre outros fatores. As lições aprendidas foram:

- Aumentar o glicerol para 15% - 20% produz cristais mais espalhados na gota e mais bem formados, sem serem ocos. Há também a diminuição de precipitante na gota, porém a gota não fica totalmente límpida. Uma vantagem também é que a condição com estas concentrações de glicerol já é crioprotetora, facilitando o manuseio durante a montagem dos cristais.

- É necessária a presença do magnésio, mas a variação da sua concentração não gerou cristais melhores do que a concentração inicial de 200 mM.

- Aumentar o volume da gota e a concentração da proteína para 20 mg/mL produz cristais maiores e em menor quantidade.

- O aumento do pH de 8,5 para 9 também melhora o formato dos cristais, favorecendo o formato mais cúbico, efeito também causado pelo aumento do glicerol.

- Outros PEGs foram utilizados, mas só houve cristalização com o PEG 4000.

- Os cristais não crescem a 4°C, somente a 18°C (foram as únicas temperaturas testadas).

- A presença de até 5% de DMSO não atrapalha a cristalização, o que é importante para a co-cristalização e *soaking* dos inibidores.

Os cristais também crescem quando se faz a co-cristalização com NADP<sup>+</sup>, G6P e esteroide
(DHEA, EA, 16BrEA)

Assim, conseguiu-se uma melhor condição de cristalização (100 mM Tris-HCl pH 9, 200 mM MgCl<sub>2</sub>, 13-15% PEG 4000, 15-20% glicerol/ proteína: 20 mg/mL, 2 mM NADP<sup>+</sup>, 40 mM G6P em 50 mM Tris-HCl pH 8,0 e 150 mM NaCl) que produz cristais maciços, espalhados na gota, em condição crioprotetora, resistentes a 5% DMSO e a *soaking*. Estes cristais difratam por volta de 3 Å, resolução suficiente para resolver a estrutura e localização dos sítios dos ligantes.

Um problema nos ensaios de cristalização com os esteroides é a baixa solubilidade destes em meio aquoso, comparada à concentração da proteína utilizada no ensaio. A solubilidade dos esteroides utilizados é por volta de 100  $\mu$ M em 2% DMSO. A concentração da proteína utilizada nos ensaios é entre 180  $\mu$ M (10 mg/mL) a 360  $\mu$ M (20 mg/mL). Ou seja, não se consegue preparar uma amostra com 1:1 de proteína ligante, sendo que o ideal seria ter um excesso de ligante, conforme sua constante de afinidade. Assim, tentaram-se diferentes técnicas de preparo de amostra para contornar este problema, inspiradas pela forma como foram resolvidas outras estruturas com esteroides.<sup>46</sup> Abaixo estão os métodos utilizados.

1. Co-cristalização

Neste caso, tentou-se estabelecer condições com maior concentração de precipitante para poder utilizar a proteína em concentrações mais baixas, entre 2 e 5 mg/mL. Porém, não se obteve sucesso neste caso.

Foram feitas amostras com 100  $\mu$ M em 2-5% DMSO em condição de subsaturação, sendo os cristais posteriormente submetidos a *soaking*.

Tentou-se também incubar a enzima com 0,5 mM a 1 mM DHEA ou EA em 5% DMSO por 1 hora, à 4°C, sempre colocando os ligantes na ordem NADP<sup>+</sup>, esteroide, G6P, para tentar diminuir a atividade da enzima, uma vez que deste modo ela não entre em contato com ambos os substratos na ausência de inibidor. Em seguida, a amostra foi centrifugada para retirada do precipitante. Neste caso, ainda observou-se o crescimento de cristais.

#### 2. Concentração da proteína na presença de esteroide.

Outra forma de tentar saturar a enzima com os esteroide foi de incubar a enzima diluída com um excesso de ligante, seguido de concentração. Depois, a proteína é novamente diluída com tampão contendo os ligantes, seguida de nova concentração. A ideia é que a enzima concentrada continua ligada ao inibidor depois da diluição, conseguido assim uma maior concentração de ligante na solução, parte ligada à proteína, parte em solução, aumentando a saturação da enzima. Este método foi utilizado com sucesso na resolução do complexo da 17β-hidroxiesteroide desidrogenase tipo 5 (17β-HSD5) com um inibidor esteroidal EM1404 (1ZQ5).<sup>46</sup> Brevemente, incubou-se 3 mL enzima diluída (10 μM – 0,5 mg/mL) com 2 mM NADP<sup>+</sup> e 100 μM EA ou 16BrEA à 5% de DMSO. Depois se acrescentou 10 mM G6P, tudo à 4°C. Incubou-se por uma hora e depois a amostra foi concentrada até a proteína estar a 15 mg/mL. Em seguida, a enzima foi diluída em tampão contendo as mesmas concentrações de NADP<sup>+</sup>, G6P e esteroide, sendo concentrada novamente, seguindo para cristalização. Também se obtiveram cristais com este método para o EA (Fig. 23-B).

#### 3. Esteroide adicionado na condição de cristalização

Observou-se que a solubilidade dos esteroides é maior na condição de cristalização, provavelmente devido ao PEG 4000, sendo que se consegue 0,5 mM de DHEA e EA em 0,1 mM Tris HCl pH 9,0; 13%PEG 4000, 15% glicerol, 2%DMSO e 200 mM MgCl<sub>2</sub>; Assim, foram feitas as condições de cristalização com a adição de 0,5 mM DHEA ou EA com 40 mM G6P (os ligantes estão só na solução usada na gota, sendo o poço preenchido com a condição idêntica sem os ligantes). Neste caso, montou-se a gota com a solução preparada com o esteroide+G6P e uma solução de proteína (20 mg/mL) com 2 mM NADP<sup>+</sup>. Neste caso cresceram cristais principalmente para DHEA, com cristais quase cúbicos e grandes (aproximadamente 0,5 mm) (Fig. 23-C).

#### 4. Adição de esteroide na gota com cristais

Este método foi utilizado com cristais que tinham crescido com NADP<sup>+</sup> e G6P. Colocouse 0,3  $\mu$ L de esteroide (10mM em 100% DMSO) diretamente na gota com os cristais

formados. Os cristais foram coletados com diferentes tempos (entre 10 segundos e 5 minutos) após a incubação.

5. Soaking

Neste caso, testou-se fazer o *soaking* dos cristais na condição de cristalização (com 20% glicerol) com 0,5 mM a 1 mM de DHEA, EA ou 16BrEA, com G6P+NADP<sup>+</sup>, G6P+NADPH e Gln6P+NADP<sup>+</sup>.

Abaixo se pode ver foto dos cristais inicias e de cristais obtidos com a condição refinada conjuntamente com diferentes técnicas de preparação de amostra.



Figura 23: A: Cristais da Hs $\Delta$ G6PD reproduzidos da literatura (condição: 100mM Tris-HCl pH 8,5, 200mM MgCl<sub>2</sub>, 12% PEG 4000 e 5% glicerol, proteína: 5mg/ml, 75  $\mu$ M NADP<sup>+</sup>, 60mM G6P). B: cristais na condição refinada, com a proteína sendo concentrada na presença de EA (condição: 100mM Tris-HCl pH 9, 200mM MgCl<sub>2</sub>, 13% PEG 4000 e 20% glicerol, proteína: 15mg/ml, 2mM NADP<sup>+</sup>, 10mM G6P, 100  $\mu$ M EA), C: cristal obtido com DHEA na condição de cristalização (condição: 100mM Tris-HCl pH 9, 200mM MgCl<sub>2</sub>, 13% PEG 4000 e 20% glicerol, 0,5mM DHEA, 2% DMSO, 40mM G6P; proteína: 20mg/ml + 2mM NADP<sup>+</sup>)

Em geral, não se obteve cristais por co-cristalização com F6P, GIn6P e NADPH, sendo estes ligantes usados para *soaking*. Quando se faz a co-cristalização somente com NADP<sup>+</sup>, o resultado são agulhas finas. Foi feito *seeding* para o complexo NADPH+G6P+16BrEA, mas não se obteve resultado. Estima-se que foram testados entre 100 a 200 cristais, sendo que destes, foi possível coletar 21 conjuntos de dados para diferentes complexos da Hs $\Delta$ G6PD, como mostrado na tabela 6. A exposição média por imagem foi de 30 s a 1 min na W01B-MX2 (LNLS), sendo os maiores cristais com 10 s de exposição.

Ligantes	Método de preparo de amostra	N° cristais coletados
NADP++G6P	1	2
NADP++G6P + DHEA	1	1
	3	3
	1	1
NADP++G6P + EA	2	3
	1 e 4	1
	2 e 4	2
	1	1
NADP++G6P + 16BrEA	1 e 4	3
	5	1
NADPH+G6P + 16BrEA	5	2
NADP++Gln6P + 16BrEA	5	1
NADP++F6P + 16BrEA	5	1

Tabela 5: Resumo dos conjuntos coletados para Hs∆G6PD, mostrando quantos cristais foram coletados para cada tipo de complexo e métodos de preparo de amostra, discutidos anteriormente.

1: co-cristalização; 2: Concentração da proteína na presença de esteroide; 3: Esteroide adicionado na condição de cristalização; 4: Adição de esteroide na gota com cristais; 5: *soaking*.

Os conjuntos foram processados e após um ciclo inicial de refinamento, observou-se o mapa diferença para avaliação de possíveis densidades para os esteroides. Não foi observada nenhuma densidade que pudesse indicar a presença do inibidor, tendo quase sempre a densidade bem definida para G6P e uma densidade parcial para o NADP<sup>+</sup> estrutural. Deste modo, apresentase o estudo de um conjunto coletado para apresentar os parâmetros do cristal, que são diferentes dos publicados, a substituição molecular e refinamento da estrutura. O conjunto usado como modelo foi um da gota da figura 23-B, estando na legenda a condição de cristalização. Apesar dos cristais não apresentarem o esteroides, estes permitiram o aprendizado da técnica de cristalografia, que deverá ser aprofundado durante o doutorado.

Os dados de difração de raios X foram coletados na linha W01B-MX2 do LNLS. Obteve-se um conjunto com 2,87 Å. As reflexões foram indexadas e integradas pelo programa iMOSFLM. O escalonamento dos dados foi feito utilizando SCALA, os fatores de estrutura foram calculados pelo CTRUNCATE, separando 5% de reflexões aleatoriamente para o calculo de R<sub>free</sub>. Os dados estatísticos deste conjunto estão apresentados na tabela 6.

	Hs∆G6PD
Comprimento de onda (Å)	1,46
Grupo espacial	P22 <sub>1</sub> 2 <sub>1</sub>
Cela unitária:	
a (Å)	53,76
b (Å)	124,78
c (Å)	166,85
Monômeros por UA	2
Resolução (Å)	55,62 – 2,87
N° reflexões	94499
N° reflexões únicas	22899
Ι/σ	9,9 (2,4)*
R <sub>pim</sub> (%)	4,2 (27,4)*
Completeza (%)	87,3 (79,6)*
Redundância	4,1 (3,3)*

Tabela 6: Dados estatísticos referentes ao processamento dos dados.

\* fora dos parênteses está o valor médio para o conjunto e dentro, o valor referente a faixa de maior resolução. UA: unidade assimétrica.

O grupo espacial pôde ser confirmado analisando as ausências sistemáticas nos eixos. Para isto, utilizaram-se também outros conjuntos que apresentavam a mesma cela unitária e no qual um eixo ou outro foi medido melhor, devido ao posicionamento do cristal em relação ao feixe. Assim, foram observadas ausências sistemáticas na direção dos eixo b e c (reflexões (0,k,0) ou (0,0,I) tinham intensidade nula quando k e I são impáres, respectivamente). Utilizando o programa *Matthews* do pacote CCP4 para o cálculo do volume de Matthews, estimou-se o conteúdo de dois monômeros na cela unitária, com 50,7% de solvente.

A substituição molecular foi feita pelo programa MOLREP com o monômero da estrutura humana 2BH9, sendo a estrutura com maior resolução para esta espécie. Foram retirados os ligantes e as moléculas de água. Dois monômeros foram posicionados na unidade assimétrica, formando um dímero não funcional (Figura 24). Considerando as operações de simetria, não haviam choques estéricos e o tetrâmero funcional é formado por uma rotação de ordem 2 (rotação na direção do eixo a).



Figura 24: Em laranja, conteúdo da unidade assimétrica formada por dois monômeros, sendo o tetrâmero funcional formado por rotação.

O refinamento no espaço recíproco foi feito utilizando o programa REFMAC5. Primeiramente se fez um refinamento de corpo rígido, seguido de refinamento de máxima verossimilhança com restrições. Após o primeiro ciclo de refinamento com restrição, obteve-se um modelo com R<sub>work</sub> de 29,59% e R<sub>free</sub> de 32,75%. Em seguida, foram analisadas os mapas de densidade eletrônica pelo programa WinCoot, analisando regiões com pouco densidade, além de procurar picos no mapa F<sub>o</sub>-F<sub>c</sub>, que devem ser correspondentes a ligantes. Foi possível observar com clareza a densidade da G6P no mapa diferença, visto na figura 25.



Figura 25: Mapa diferença ( $F_0$ - $F_c$ ) em verde com cotorno de  $3\sigma$ , evidenciando a presença da G6P.

O refinamento manual foi feito ajustando resíduos que estão fora da densidade eletrônica, apagando *loops* ou resíduos sem densidade e incluindo ligantes. Assim, alternou-se entre o refinamento manual com o refinamento do REFMAC5, sempre acompanhando os valores de  $R_{work}/R_{free}$ , que devem possuir o mesmo comportamento, ou seja, os dois devem diminuir juntamente, e não só o  $R_{work}$ . O refinamento foi concluído quando mudanças no modelo não geraram a diminuição dos valores de  $R_{work}$  e  $R_{free}$ , além da estrutura respeitar as restrições estereoquímicas, analisadas pelo *Rampage*. Os resíduos Y484 e E297 estão em regiões proibidas, porém apresentam uma boa densidade e conservação da cadeia principal entre as estruturas publicadas, sendo encontradas em *loops* curtos entre fita- $\beta$  e hélice- $\alpha$ . A estrutura final apresenta fator de temperatura médio de 75 Å<sup>2</sup> e uma correlação de 0,94 entre os fatores de estrutura calculados pelo modelo com os fatores de estrutura obtidos experimentalmente. Abaixo são apresentados os parâmetros estatísticos referentes ao modelo.

Parâmetros estatísticos	Hs∆G6PD – G6P
R <sub>work</sub> (%)	20,8%
R <sub>free</sub> (%)	26,8%
RMS comprimento de ligação (Å)	0,01
RMS ângulo de ligação (Å)	1,4
Resíduos em regiões favoráveis	no 98
diagrama de Ramachandran* (%)	
Resíduos em regiões permitidas	no 2
diagrama de Ramachandran* (%)	

Tabela 7: Parâmetros estatísticos do refinamento da estrutura da Hs∆G6PD – G6P.

\* calculado pelo programa *Rampage*.

Comparando a estrutura obtida com a estrutura do PDB complexada com G6P (2BHL), não se observam grandes diferenças, tendo um rmsd entre a cadeia principal das duas estruturas de 0,63 Å, e para a estrutura como um todo, um rmsd de 1,02 Å (valores calculados por *Superpose*).<sup>47</sup>

Neste capítulo foram descritos os esforços para tentar resolver a estrutura da Hs∆G6PD complexada com esteroide. Apesar de nenhum conjunto de dados coletado apresentar uma densidade que poderia ser referente ao inibidor, conseguiu-se estabelecer um protocolo para cristalização da G6PD que é reprodutivo e tendo cristais que são passíveis de *soaking*, criando uma ferramenta que poderá ser utilizado pelo grupo para caracterização estrutural de novos inibidores da enzima.

## Capítulo 6: Mutantes A277C e E347A

Na perspectiva de melhorar a resolução e qualidade dos cristais atuais, tornando-os ainda mais adequados aos estudos com inibidores, ou ainda encontrando outras condições de cristalização, avaliou-se em mutações que pudessem estabilizar a forma tetramérica da proteína ou o contrário, que pudesse rompê-la. Além disto, pouco se sabe sobre a importância fisiológica do equilíbrio entre dímero e tetrâmero, não se sabendo se uma forma é mais ou menos ativa que a outra. Assim, mutantes que tenham somente um estado oligomérico podem ser úteis nestas investigações.

A interface do tetrâmero é formada basicamente por interações eletrostáticas, sendo a região de contato entre as superfícies de 706 Å<sup>2</sup> entre os monômeros. As principais interações ocorrem entre resíduos da hélice 16 (considerando anotação de DSSP<sup>48</sup>), sendo que a hélice de cada monômero está uma de frente para outra (Figura 26). É possível observar também a ponte salina entra o E347 e a K275.



Figura 26: Interface do tetrâmero. Em destaque, interação entre E347 e K275 e hélice 16 com C294 e A277 próximas.

Como pode ser visto na Figura 26, no início da hélice 16 encontra-se a alanina 277 que está de frente com a cisteína 294, com o enxofre da cisteína distando 4Å do Cβ da alanina. Assim, pensou-se que ali poderia ser uma boa oportunidade para tentar realizar uma ponte dissulfeto que poderia manter o tetrâmero ligado covalentemente, mutando a alanina 277 por uma cisteína.

Foram checados também se a geometria entre os resíduos permitiriam tal ligação, que acontece com um ângulo de 105° entre C $\beta$ -S-S<sup>\*</sup>.<sup>49</sup>

Outra mutação estudada foi a de trocar o ácido glutâmico 347 por uma alanina, a fim de quebrar a interação deste com a cavidade positiva encontrada na interface, gerando um mutante que não formaria mais o tetrâmero.

Tentou-se fazer a mutação tanto no vetor da Hs $\Delta$ G6PD como para o mutante Hs $\Delta$ G6PD D200N. Não se conseguiu obter ainda os mutantes para a forma selvagem, tendo só conseguido os mutantes Hs $\Delta$ G6PD D200N A277C (Hs $\Delta$ G6PD DNAC) e Hs $\Delta$ G6PD D200N E347A (Hs $\Delta$ G6PD DNEA). Os mutantes foram expressos e purificados do mesmo modo que a forma selvagem.

Abaixo se pode ver o cromatograma de gel filtração destes mutantes. Para Hs∆G6PD DNAC, pode-se ver um pico simétrico correspondente ao tetrâmero. Já para o Hs∆G6PD DNEA, observa-se um pico para o dímero, com uma pequena formação do tetrâmero. Isto comprova que as mutações feitas conseguiram estabilizar a proteína em um estado oligomérico preferencial.



Figura 27: Perfil do cromatograma de gel filtração do mutante Hs∆G6PD D200N e dos mutantes Hs∆G6PD D200N E347A e Hs∆G6PD D200N A277C em *Superdex 200 10/30*.

## 6.1 A Hs∆G6PD DNEA

A Hs∆G6PD DNEA foi submetida a um *screening* de cristalização com kits inicias. Obtiveram-se 3 condições com aparecimento de agulhas em 2 dias após a montagem das placas, sendo as 3 condições muito relacionadas. As fotos com exposição de UV revelam que provavelmente as agulhas são de proteína, pois é visível a fluorescência destas, mostrado na figura 28.



Figura 28: Fotos dos cristais obtidos com o mutante Hs∆G6PD DNEA em *screening* inicial de cristalização. Em A, B e C, fotos retiradas com luz branca; D, F e G, fotos retiradas com exposição de luz UV. A-D: condição: 20% PEG 3350, 200mM NaBr, 100mM Bis-Tris propano pH 7,5; B-F: condição: 20% PEG 3350, 200mM Nal, 100mM Bis-Tris propano pH 7,5; C-G: condição: 20% PEG 3350, 200mM acetato de sódio, 100mM Bis-Tris propano pH 8,5. Em todas estas condições, a proteína estava à 10mg/mL com 2mM NADP<sup>+</sup>.

Estas condições serão refinadas futuramente, testando também outros complexos. Devido ao tempo, não foi feito nada a mais com este mutante.

## 6.2 Hs∆G6PD DNAC

Foi feita a termofluorescência deste mutante para avaliar se a mutação garante maior estabilidade à proteína. O resultado obtido foi que a mutação na alanina por cisteína aumentou em 10°C a T<sub>m</sub> da enzima em relação à proteína selvagem, corroborando com a devida formação da ponte dissulfeto entre os dímeros.



Figura 29: Termofluorescência do mutante Hs∆G6PD DNEA, mostrando um aumento de 10°C no T<sub>m</sub>.

Este mutante foi utilizado em ensaio de cristalização com 40mM G6P + 2mM NADP<sup>+</sup>, e com 40 mM G6P + 3mM NADPH + 0,2µM NADP<sup>+</sup> + 0,3mM 16BrEA, na condição refinada para a selvagem, variando a concentração de PEG e glicerol. Diferente do que acontecia com o mutante Hs $\Delta$ G6PD D200N, que nunca cristalizou, este novo mutante produziu cristais de ambos complexos, sendo também os primeiros cristais que cresceram na presença de NADPH. Os cristais são morfologicamente parecidos com os da selvagem, descritos em capítulo anterior.

Foram coletados dois conjuntos de dados para o complexo com G6P+NADPH+NADP<sup>+</sup>+ 16BrEA. Esperava-se um aumento na resolução dos cristais, uma vez que esta proteína é mais estável e apresenta somente um estado oligomérico. Porém, estes cristais continuam difratando por volta de 3 Å.

Estes cristais possuem cela e grupo espacial diferente da enzima selvagem. Abaixo podem se vistos os dados estatísticos referentes ao processamento de um dos conjuntos.

	Hs∆G6PD D200N A277C
Comprimento de onda (Å)	1,46
Grupo espacial	C222 <sub>1</sub>
Cela unitária:	
a (Å)	57,76
b (Å)	175,91
c (Å)	217,22
Monômeros por UA	2
Resolução (Å)	55,9 - 3,2
N° reflexões	66652
N° reflexões únicas	16449
Ι/σ	8,6 (3,3)*
R <sub>pim</sub> (%)	5,6 (21,5)*
Completeza (%)	89,1 (84,2)*
Redundância	4,1 (4,1)*

Tabela 8: Dados estatísticos referentes ao processamento dos dados do cristal da Hs∆G6PD D200N A277C crescidos na presença de G6P+NADPH+16BrEA.

\* fora dos parênteses está o valor médio para o conjunto e dentro, o valor referente a faixa de maior resolução. UA: unidade assimétrica

A substituição molecular e o refinamento foram feitos do mesmo modo que a selvagem. A unidade assimétrica também possui dois monômeros, correspondendo a um dímero não funcional, que quando aplicado a operação de simetria de ordem 2, gera o tetrâmero. Após o primeiro ciclo de refinamento, obteve-se um R<sub>work</sub> de 27,49% e um R<sub>free</sub> de 32,58%. Com este modelo, foi possível observar no mapa diferença: densidade no sítio catalítico para o NADPH (não sabendo em realidade se está na forma reduzida); densidade parcial para a G6P; e densidade parcial para a adenina do NADP<sup>+</sup> estrutural, como mostrado na figura 30. A ponte dissulfeto entre as cisteínas também pode ser vista. Não foi observada nenhuma densidade que pudesse ser do 16BrEA.



Figura 30: Mapas de densidade eletrônica para o modelo da Hs $\Delta$ G6PD D200N A277C após primeiro ciclo de refinamento. Em verde: mapa diferença (F<sub>0</sub>-F<sub>c</sub>) com cotorno de 3 $\sigma$ . Em azul: mapa (2F<sub>0</sub>-F<sub>c</sub>) com contorno de 2 $\sigma$ . A: densidade para G6P e NADPH no sítio catalítico. B: parte de NADP<sup>+</sup> estrutural. C: Interface do tetrâmero para estrutura refinada da Hs $\Delta$ G6PD. D: Interface do tetrâmero para estrutura Hs $\Delta$ G6PD D200N A277C, após primeiro ciclo de refinamento, sendo evidenciada a ponte dissulfeto.

O modelo possui um  $R_{work}$  de 21,8% e  $R_{free}$  de 27,1% (rmsd para ligação de 0,011Å e rmsd para ângulo de ligação de 1,51°), índice de correlação de 0,94 e fator de temperatura médio de 81,6 Å<sup>2</sup>. Analisando a geometria, 98% dos resíduos estão em regiões favoráveis no diagrama de Ramachandran e 2% estão em regiões permitidas.

A estrutura é muito parecida com a estrutura da selvagem. O rmsd para o domínio do NADP<sup>+</sup> é de 0,42 Å, para o domínio da G6P, 0,33 Å e de 0,49 Å entre os monômeros. Isto indica que o domínio do NADP<sup>+</sup> não se alterou significativamente com a ligação deste, considerando que o valor de rmsd é parecido com o valor obtido para o dominío de G6P e para o monômero como um todo. A formação do dímero também é conservada, tendo um rmsd de 0,83 Å.

Já em relação à organização entre os dímeros para formação do tetrâmero, observou-se que esta é um pouco diferente entres as estruturas, causada pela ponte dissulfeto, mas que as principais interações na interface continuam as mesmas. Utilizando o programa *DynDom*<sup>51</sup> para medição das diferenças, obteve-se que no tetrâmero da Hs∆G6PD DNAC os dímeros estão 1,4 Å mais próximos e há uma rotação de 3,3° se comparado com o tetrâmero da Hs∆G6PD (Figura 31).

Sobreposição





Pode-se observar que com este mutante foi possível cristalizar um complexo que apresenta tanto a G6P quanto NADPH, o que é necessário para a ligação dos esteroides. Esta estrutura também poderá ser utilizada para estudos de docagem do esteroide, sendo a mais adequada, por possuir os dois substratos.

## Capítulo 7: Conclusão

Neste trabalho se conseguiu estabelecer um protocolo de expressão e purificação para a HsΔG6PD que possui rendimento adequado para realização de ensaios biofísicos e estruturais. Apesar de não ter conseguido um cristal que apresentasse densidade eletrônica para os esteroides, estabeleceu-se um sistema de cristalização reprodutivo que permite o estudo estrutural de outros inibidores, o que será feito nos próximos anos. Além disto, estes cristais permitiram ter uma experiência inicial em cristalografia de proteínas.

Na tentativa de contornar os problemas inerentes à cristalização de um inibidor incompetitivo, testaram-se diferentes complexos não reativos, não tendo nenhum resultado favorável. Outra estratégia foi fazer um mutante (D200N) que diminuiu a atividade da enzima em 30 vezes para facilitar a cristalização do complexo. Porém, observou-se uma diminuição na inibição pelos esteroides, apontando que o aspartato 200 além de ser importante para a atividade da enzima, também é importante para a inibição destes compostos. Esta é a primeira evidência de algum resíduo que influencia na inibição da G6PD por esteroides e que deverá ser explorada futuramente, com o uso de técnicas computacionais e avaliação de outros mutantes no entorno do resíduo.

Concomitantemente, foram propostas mutações na interface do tetrâmero para obtenção de enzimas que só apresentassem um estado oligomérico, com o objetivo de melhorar a resolução dos cristais já obtidos ou obter outras condições de cristalização, além de obter mutantes que permitirão o estudo mais aprofundado da influência do estado oligomérico para a atividade da enzima. Não se obtiveram os mutantes para a enzima selvagem, somente para a construção já contendo a mutação D200N, tendo duplo mutantes. A mutação E347A se mostrou capaz de favorecer fortemente o equilíbrio para o dímero, mostrando que o glutamato 347 é uma peça chave para formação do tetrâmero. Conseguiu-se também novas condições iniciais de cristalização para esta construção (HsΔG6PD D200N E347A), que poderão facilitar a cristalização da enzima com os esteroides ou com outros inibidores. Obteve-se também com a mutação A277C a criação de pontes dissulfetos que estabilizam a forma tetramérica da enzima. Para esta construção (HsΔG6PD D200N ternário necessário à ligação dos esteroides. Nenhuma densidade para estes foi vista, mas outros cristais deverão ser coletados, além de outras formas de
preparo de amostra deverão ser feitas, podendo talvez conseguir o complexo quaternário. Estas mutações mostraram os resíduos que quando alterados, conseguem alterar o equilíbrio entre dímero/tetrâmero da enzima. Assim, uma vez conseguidas estas mutações na enzima selvagem, estudos da influência do estado oligomérico na atividade da enzima poderão ser feitos.

## Referências

[1] A. Rassi, A. Rassi, and J. A. Marin-Neto, "Chagas disease.," *Lancet*, vol. 375, no. 9723, pp. 1388–402, Apr. 2010.

[2] J. Utzinger, S. L. Becker, S. Knopp, J. Blum, a L. Neumayr, J. Keiser, and C. F. Hatz, "Neglected tropical diseases: diagnosis, clinical management, treatment and control.," *Swiss medical weekly*, vol. 142, no. November, pp. 19–22, Jan. 2012.

[3] J. R. Coura, "Present situation and new strategies for Chagas disease chemotherapy: a proposal.," *Memórias do Instituto Oswaldo Cruz*, vol. 104, no. 4, pp. 549–54, Jul. 2009.

[4] X. P. Chen and G. H. Du, "Target validation: A door to drug discovery.," *Drug discoveries & therapeutics*, vol. 1, no. 1, pp. 23–9, Aug. 2007.

[5] A. Vincent, K. Daniel, and K. Moses, "RNA interference : a pathway to drug target identification and validation in trypanosome," vol. 2, no. 3, pp. 66–73, 2008.

[6] T. T. Talele, S. a Khedkar, and A. C. Rigby, "Successful applications of computer aided drug discovery: moving drugs from concept to the clinic.," *Current topics in medicinal chemistry*, vol. 10, no. 1, pp. 127–41, Jan. 2010.

[7] W. N. Hunter, "Structure-based ligand design and the promise held for antiprotozoan drug discovery.," *The Journal of biological chemistry*, vol. 284, no. 18, pp. 11749–53, May 2009.

[8] A. T. Cordeiro, O. H. Thiemann, and P. a M. Michels, "Inhibition of Trypanosoma brucei glucose-6-phosphate dehydrogenase by human steroids and their effects on the viability of cultured parasites.," *Bioorganic & medicinal chemistry*, vol. 17, no. 6, pp. 2483–9, Mar. 2009.

[9] A. T. Cordeiro and O. H. Thiemann, "16-bromoepiandrosterone, an activator of the mammalian immune system, inhibits glucose 6-phosphate dehydrogenase from Trypanosoma cruzi and is toxic to these parasites grown in culture.," *Bioorganic & medicinal chemistry*, vol. 18, no. 13, pp. 4762–8, Jul. 2010.

[10] M. Igoillo-Esteve, D. Maugeri, A. L. Stern, P. Beluardi, and J. J. Cazzulo, "The pentose phosphate pathway in Trypanosoma cruzi: a potential target for the chemotherapy of Chagas disease.," *Anais da Academia Brasileira de Ciências*, vol. 79, no. 4, pp. 649–63, Dec. 2007.

[11] S. H. Lee, J. L. Stephens, and P. T. Englund, "A fatty-acid synthesis mechanism specialized for parasitism.," *Nature reviews. Microbiology*, vol. 5, no. 4, pp. 287–97, Apr. 2007.

[12] R. Hurtado-Guerrrero, J. Peña-Díaz, A. Montalvetti, L. M. Ruiz-Pérez, and D. González-Pacanowska, "Kinetic properties and inhibition of Trypanosoma cruzi 3-hydroxy-3-methylglutaryl CoA reductase.," *FEBS letters*, vol. 510, no. 3, pp. 141–4, Jan. 2002. [13] R. L. Krauth-Siegel and M. a Comini, "Redox control in trypanosomatids, parasitic protozoa with trypanothione-based thiol metabolism.," *Biochimica et biophysica acta*, vol. 1780, no. 11, pp. 1236–48, Nov. 2008.

[14] M. Dormeyer, N. Reckenfelderbäumer, H. Ludemann, and R. L. Krauth-Siegel, "Trypanothione-dependent synthesis of deoxyribonucleotides by Trypanosoma brucei ribonucleotide reductase.," *The Journal of biological chemistry*, vol. 276, no. 14, pp. 10602–6, Apr. 2001.

[15] R. L. Krauth-Siegel, S. K. Meiering , H. Schmidt "The parasite-specific trypanothione metabolism of trypanosoma and leishmania", *Biological Chemistry*, vol 384, no 4, pp. 539–549, Apr. 2003

[16] E. F. Irccs and F. Via, "Glucose-6-phosphate dehydrogenase deficiency," vol. 371, pp. 64–74, 2008.

[17] K. Ayi, F. Turrini, A. Piga, and P. Arese, "Enhanced phagocytosis of ring-parasitized mutant erythrocytes: a common mechanism that may explain protection against falciparum malaria in sickle trait and beta-thalassemia trait.," *Blood*, vol. 104, no. 10, pp. 3364–71, Nov. 2004.

[18] P. Cohen and M. a Rosemeyer, "Subunit interactions of glucose-6-phosphate dehydrogenase from human erythrocytes.," *European journal of biochemistry / FEBS*, vol. 8, no. 1, pp. 8–15, Mar. 1969.

[19] P. Cohen and M. A. Rosemeyer, "Human Glucose-6-Phosphate dehydrogenase: Purification of the Erythrocyte Enzyme and the Influence of Ions on its Activity," *European Journal of Biochemistry*, vol. 8, pp. 1–7, 1969.

[20] P. Rowland, a K. Basak, S. Gover, H. R. Levy, and M. J. Adams, "The three-dimensional structure of glucose 6-phosphate dehydrogenase from *Leuconostoc mesenteroides* refined at 2.0 A resolution.," *Structure (London, England : 1993)*, vol. 2, no. 11, pp. 1073–87, Nov. 1994.

[21] S. W. N. Au, C. E. Naylor, S. Gover, L. Vandeputte-Rutten, D. a. Scopes, P. J. Mason, L. Luzzatto, V. M. S. Lam, and M. J. Adams, "Solution of the structure of tetrameric human glucose 6-phosphate dehydrogenase by molecular replacement," *Acta Crystallographica Section D Biological Crystallography*, vol. 55, no. 4, pp. 826–834, Apr. 1999.

[22] S. W. Au, S. Gover, V. M. Lam, and M. J. Adams, "Human glucose-6-phosphate dehydrogenase: the crystal structure reveals a structural NADP(+) molecule and provides insights into enzyme deficiency.," *Structure*, vol. 8, no. 3, pp. 293–303, Mar. 2000.

[23] X. Wang, T. F. A. I. Chan, and V. M. S. Lam, "What is the role of the second " structural " NADP + -binding site in human glucose 6-phosphate dehydrogenase ?," *Protein Science*, no. 17, pp. 1403–1411, 2008.

[24] M. Kotaka, S. Gover, L. Vandeputte-Rutten, S. W. N. Au, V. M. S. Lam, and M. J. Adams, "Structural studies of glucose-6-phosphate and NADP+ binding to human glucose-6-phosphate dehydrogenase.," *Acta crystallographica. Section D, Biological crystallography*, vol. 61, no. 5, pp. 495–504, May 2005.

[25] C. Ortíz, N. Larrieux, A. Medeiros, H. Botti, M. Comini and A. Buschiazzo "Expression, crystallization and preliminary X-ray crystallographic analysis of glucose-6-phosphate dehydrogenase from the human pathogen *Trypanosoma cruzi* in complex with substrate", *Acta Crystallographica. Section F, Structural Biology and Crystallization Communications*, vol. 67, nov. 2011

[26] P. A. Marks, J. Banks "Inhibition of Mammalian glucose-6-phosphate dehydrogenase by steroids" *Proceedings of the National Academy of Sciences of the United States of America*, vol. 46, pp. 447-452, 1960

[27] G. Gordon, M. C. Mackow, and H. R. Levy, "On the Mechanism of Interaction of Steroids with Human Glucose 6-phosphate Dehydrogenase," *Archives of biochemistry and biophysics*, vol. 318, no. 1, pp. 25–29, 1995.

[28] P. D. Kroboth, F. S. Salek, A. L. Pittenger, T. J. Fabian, and R. F. Frye, "DHEA and DHEA-S: a review," *The Journal of Clinical Pharmacology*, vol. 39, pp. 327–348, 2012.

[29] A. G. Schwartz and L. L. Pashko, "Dehydroepiandrosterone, glucose-6-phosphate dehydrogenase, and longevity.," *Ageing research reviews*, vol. 3, no. 2, pp. 171–87, Apr. 2004.

[30] S. Gupta, A. T. Cordeiro, and P. a M. Michels, "Glucose-6-phosphate dehydrogenase is the target for the trypanocidal action of human steroids.," *Molecular and biochemical parasitology*, vol. 176, no. 2, pp. 112–5, Apr. 2011.

[31] U. B. Ericsson, B. M. Hallberg, G. T. Detitta, N. Dekker, and P. Nordlund, "Thermofluorbased high-throughput stability optimization of proteins for structural studies.," *Analytical Biochemistry*, vol. 357, no. 2, pp. 289–98, Oct. 2006.

[32] Guimaraes, B. G., Sanfelici, L., Neuenschwander, R. T., Rodrigues, F., Grizolli, W. C., Raulik, M. A., Piton, J. R., Meyer, B. C., Nascimento, A. S. & Polikarpov, I. "The MX2 macromolecular crystallography beamline: a wiggler X-ray source at the LNLS "*Journal of Synchrotron Radiation*", vol. 16, pp. 69-75. 2009

[33] Z. Dauter, "Data-collection strategies," *Acta Crystallographica Section D Biological Crystallography*, vol. 55, no. 10, pp. 1703–1717, Oct. 1999.

[34] T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, and A. G. W. Leslie, "iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM.," *Acta crystallographica*. *Section D, Biological crystallography*, vol. 67, no. Pt 4, pp. 271–81, Apr. 2011.

[35] P. R. Evans, "An introduction to data reduction: space-group determination, scaling and intensity statistics.," *Acta crystallographica. Section D, Biological crystallography*, vol. 67, no. Pt 4, pp. 282–92, Apr. 2011.

[36] A. Vagin and A. Teplyakov, "Molecular replacement with MOLREP.," *Acta crystallographica. Section D, Biological crystallography*, vol. 66, no. Pt 1, pp. 22–5, Jan. 2010.

[37] G. N. Murshudov, P. Skubák, A. a Lebedev, N. S. Pannu, R. a Steiner, R. a Nicholls, M. D. Winn, F. Long, and A. a Vagin, "REFMAC5 for the refinement of macromolecular crystal structures.," *Acta crystallographica. Section D, Biological crystallography*, vol. 67, no. Pt 4, pp. 355–67, Apr. 2011.

[38] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. a Potterton, H. R. Powell, R. J. Read, A. Vagin, and K. S. Wilson, "Overview of the CCP4 suite and current developments.," *Acta crystallographica. Section D, Biological crystallography*, vol. 67, no. Pt 4, pp. 235–42, Apr. 2011.

[39] S. W. N. Au, C. E. Naylor, S. Gover, L. Vandeputte-Rutten, D. A. Scopes, P. J. Mason, L. Luzzatto, V. M. S. Lam, and M. J. Adams, "Solution of the structure of tetrameric human glucose 6-phosphate dehydrogenase by molecular replacement," *Acta Crystallographica Section D Biological Crystallography*, vol. 55, no. 4, pp. 826–34, Apr. 1999.

[40] R. Z. Cer, U. Mudunuri, R. Stephens, and F. J. Lebeda, "IC50-to-Ki: a web-based tool for converting IC50 to Ki values for inhibitors of enzyme activity and ligand binding.," *Nucleic acids research*, vol. 37, no. Web Server issue, pp. W441–5, Jul. 2009.

[41] O. Ibraheem, O. A. Adewale, and A. Afolayan, "Purification and Properties of Glucose-6-phosphate Dehydrogenase from Aspergillus aculeatus," *Jornal of Biochemistry and Molecular Biology*, vol. 38, no. 5, pp. 584–590, 2005.

[42] X.-T. Wang, S. W. N. Au, V. M. S. Lam, and P. C. Engel, "Recombinant human glucose-6-phosphate dehydrogenase," *European Journal of Biochemistry*, vol. 269, no. 14, pp. 3417–3424, Jul. 2002.

[43] M. S. Cosgrove, C. Naylor, S. Paludan, M. J. Adams, and H. R. Levy, "On the Mechanism of the Reaction Catalyzed by Glucose 6-Phosphate," vol. 2960, no. 97, pp. 2759–2767, 1998.

[44] M. S. Cosgrove, S. Gover, C. E. Naylor, L. Vandeputte-rutten, M. J. Adams, and H. R. Levy, "Articles An Examination of the Role of Asp-177 in the His-Asp Catalytic Dyad of Leuconostoc mesenteroides Glucose 6-Phosphate Dehydrogenase : X-ray Structure and pH Dependence of Kinetic Parameters of the D177N Mutant Enzyme.," *Biochemistry*, vol. 39, pp. 15002–15011, 2000.

[45] A. Kozlenkov, M. Helene, L. Du, P. Cuniasse, T. Ny, M. F. Hoylaerts, and L. Milla, "Residues Determining the Binding Specificity of Uncompetitive Inhibitors to Tissue-Nonspecific Alkaline Phosphatase," vol. 19, no. 11, pp. 1862–1872, 2004.

[46] M. Mazumdar, M. Zhou, D.-W. Zhu, A. Azzi, and S.-X. Lin, "Crystallogenesis of Steroid-Converting Enzymes and Their Complexes: Enzyme–Ligand Interaction Studies and Inhibitor Design Facilitated by Complex Structures" *Crystal Growth & Design*, vol. 7, no. 11, pp. 2206–2212, Nov. 2007.

[47] R. Maiti, G. H. Van Domselaar, H. Zhang, and D. S. Wishart, "SuperPose: a simple server for sophisticated structural superposition.," *Nucleic acids research*, vol. 32, pp. W590–W594, Jul. 2004.

[48] W. Kabsch, C. Sander, "Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features.", *Biopolymers*, vol. 22, pp. 2577-637., Dec., 1983

[49] M. T. Petersen, P. H. Jonson, and S. B. Petersen, "Amino acid neighbours and detailed conformational analysis of cysteines in proteins.," *Protein engineering*, vol. 12, no. 7, pp. 535–48, Jul. 1999.

[50] F. W. Studier, "Protein production by auto-induction in high-density shaking cultures," *Protein Expression and Purification*, vol. 41, no. 1, pp. 207–234, May 2005.

[51] S. Hayward and H. J. Berendsen, "Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme.," *Proteins*, vol. 30, no. 2, pp. 144–54, Feb. 1998.

Uso exclusivo da CIBio:

Número de projeto / processo:

Formulário de encaminhamento de projetos de pesquisa para análise pela CIBio - Comissão Interna de Biossegurança da ABTLuS – Associação Brasileira de Tecnologia de Luz Síncrotron

<u>Título do projeto</u>: Formulação de um complexo estável entre a enzima *glicose-6-fosfato desidrogenase de Tripanosoma cruzi* e o inibidor esteróide dehidroepiandrosterona. <u>Pesquisador responsável</u>: Artur Torres Cordeiro <u>Experimentador</u>: Américo Tavares Ranzani Nível do treinamento do experimentador: []-Iniciação científica, [X]-mestrado, []-doutorado, []-doutorado direto, []-pós-doutorado, []-nível técnico, []-outro, especifique:

Resumo do projeto: Tripanossomíases e leishmaniose são enfermidades associadas à infecção por protozoários dos gêneros Trypanosoma e Leishmania, respectivamente. Estas doenças são normalmente tratadas com drogas de alta toxicidade e baixa eficiência. A pesquisa de novos compostos que atuem de maneira específica contra alvos metabólicos préestabelecidos, pode levar ao desenvolvimento de fármacos mais eficientes no tratamento destas enfermidades. A enzima glicose-6-fosfato desidrogenase (G6PDH) é um alvo metabólico validado experimentalmente em T. brucei. A G6PDH catalisa o primeiro passo da via de síntese de pentoses que supre a célula com ribose-5-fosfato para síntese de bases nitrogenadas e NADPH para biosíntese de lípideos, colesterol e neutralização de espécies reativas de oxigênio. O esteroíde dehidroepiandrosterona (DHEA) foi recentemente caracterizado como um potente inibidor incompetitivo (Ki = 1.7 µM) da G6PDH de T. brucei. O DHEA também apresentou efeito citotóxico à forma sanguínea de T. brucei em baixas concentrações ( $LD_{50} = 43.8 \mu M$ ). Em conjunto estes resultados sugerem que esteróides análogos do DHEA devam ser explorados como uma nova classe de fármacos anti parasitários. O atual projeto pretende estudar uma forma de otimizar o processo de cristalização da enzima com o inibidor, permitindo localizar o sítio de ligação deste. Informações derivadas deste estudo auxiliarão o desenvolvimento de esteróides inibidores da G6PDH, inertes ao metabolismo de mamíferos e eficientes no tratamento de doenças infecciosas.

A CIBio analisou este projeto em reunião realizada no dia:  $\underline{11, 12}$ . <u>Parecer final</u>:  $[\chi]$ -projeto aprovado, []-projeto recusado, []-projeto com deficiências, favor comentários anexo.

Presidente de CIBio - ABTLuS-LNBio

Jörg Kobarg

Membro da CIBio - ABTLuS-LNBio Prof. Dr. Celso Eduardo Benedetti

Membro da CIBio – ABTLuS-L'NBio Andrea Balan

Membro da CIBio – ABTLuS-SGT Carolina Barbosa Marini

Membro da CIBio da ABTLuS-CTBE Fabio Squina

Sindlelion Fuites Membro da CIBio da ABTLuS-CTBE

Sindelia Freitas Azzoni

Membro da CIBio da ABTLuS-CTBE Roberto Ruller