



Gene expression and molecular modeling of the HSP104 chaperone of *Trypanosoma cruzi*

R.A. Campos¹, M.L. da Silva¹, G.V. da Costa³, P.M. Bisch¹, J.M. Peralta³, R. Silva¹, E. Rondinelli^{1,2} and T.P. Ürményi¹

¹Programa de Biologia Molecular e Estrutural, Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil

²Departamento de Clínica Médica, Faculdade de Medicina, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil

³Instituto de Microbiologia Professor Paulo de Góes, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil

Corresponding author: T.P. Ürményi
E-mail: turmenyi@biof.ufrj.br

Genet. Mol. Res. 11 (3): 2122-2129 (2012)

Received January 26, 2012

Accepted July 2, 2012

Published August 6, 2012

DOI <http://dx.doi.org/10.4238/2012.August.6.15>

ABSTRACT. Heat shock protein (HSP) 104 is a highly conserved molecular chaperone that catalyzes protein unfolding, disaggregation and degradation under stress conditions. We characterized *HSP104* gene structure and expression in *Trypanosoma cruzi*, a protozoan parasite that causes Chagas' disease. The *T. cruzi* HSP104 is an 869 amino-acid protein encoded by a single-copy gene that has the highest sequence similarity (76%) with that of *T. brucei* and the lowest (23%) with that of the human protein. *HSP104* transcripts were detected at room temperature, and levels increased after incubation at 37° or 40°C. The HSP104 protein was found at low levels in non-heat-shocked cells, and accumulated continuously up to 24 h at elevated temperatures. We developed a predicted structural model of hexameric *T. cruzi* HSP104, which showed some conserved features.

Key words: HSP104; Gene expression; Protein structure; *Trypanosoma cruzi*

INTRODUCTION

Trypanosoma cruzi is a kinetoplastid protozoan parasite that causes Chagas' disease, affecting millions of people in Latin America (Coura and Borges-Pereira, 2010). No vaccines are available at present, and the drugs currently used for treatment exhibit undesirable side effects. The genome sequence of this parasite was recently completed, and proteomic analysis of the different stages of *T. cruzi* has been carried out, providing important evidence of stage-specific expression of numerous genes (Atwood et al., 2005). During its life cycle, *T. cruzi*, a digenetic parasite, undergoes several differentiation steps and a temperature shift when moving from the insect vector to the mammalian host, inducing heat shock proteins (HSPs) in the process (Rondinelli, 1994; Vonlaufen et al., 2008). HSPs are molecular chaperones and proteases that are induced under stress conditions to maintain protein homeostasis, and function in the intracellular environment (Hartl and Hayer-Hartl, 2002).

The HSP100 group of chaperone proteins, a member of the AAA+ superfamily of ATPases, is comprised of hexameric protein complexes that catalyze protein unfolding, disassembly, disaggregation, and degradation (reviewed in Doyle and Wickner, 2009). The group is subdivided into two classes based on domain organization. Class 1 proteins contain two nucleotide-binding domains (NBDs), and each member promotes either disaggregation or proteolysis. On the other hand, class 2 proteins have only one NBD and promote proteolysis. Bacterial ClpB and its eukaryotic ortholog, HSP104, are class 1 HSP100s that are characterized by the presence of a Walker-type NBD and that help organisms survive severe stress by being capable of disentangling even insoluble protein aggregates. Both ClpB and HSP104 are induced by heat shock and act in concert with DnaK and HSP70, respectively. In addition, these proteins are required for thermotolerance.

Several HSP genes have been characterized in trypanosomatids (Folgueira and Requena, 2007). In *T. cruzi*, members of the HSP90 (Dragon et al., 1987), HSP70 (Engman et al., 1989; de Carvalho et al., 1990), chaperonin (Giambiagi-de Marval et al., 1993; Fernandes et al., 2005) and small HSP (Pérez-Morales et al., 2009) families have been described. Of the HSP100 family, the HSP104 gene structure has been described in *T. brucei* (Gottesman et al., 1990), *Leishmania major* (Hübel et al., 1995), and *L. donovani* (Krobitsch et al., 1998), while HSP104 gene expression has been characterized only in *L. major* (Hübel et al., 1995) and *L. donovani* (Krobitsch et al., 1998). In this study, we investigated the gene structure and expression of *T. cruzi* HSP104.

MATERIAL AND METHODS

Parasites, culture conditions, and DNA sequence analysis

Epimastigote forms of the *T. cruzi* clone CL Brener (Cano et al., 1995) were maintained at 29°C in liver infusion tryptose medium supplemented with 10% bovine fetal serum, 0.025 mg/mL hemin, and 100 µg/mL ampicillin. All experiments described were performed on early exponential phase cultures. The *T. cruzi* HSP104 gene coding sequence (GenBank ID: HM014310) was obtained from the preliminary clone CL Brener genomic sequence data (<http://TcruziDB.org>) and completed by sequencing PCR-amplified genomic DNA fragments. Nucleotide and amino acid sequence analyses were performed using the Geneious Pro™ software (Drummond et al., 2011).

Quantitative RT-PCR analysis

Total RNA was extracted using the RNeasy® Mini-kit (Qiagen, Germany) according to manufacturer instructions. Real-time RT-PCR (qRT-PCR) assays were performed using the 7500 Real-Time PCR System (Applied Biosystems). Reverse transcription was performed from 1 µg total RNA treated with RQ1 RNase-free DNase (Promega) using random hexamer primers and M-MLV Reverse Transcriptase (Invitrogen). Amplifications of *HSP104* sequences were performed in triplicate in 25-µL reactions using the Power SYBR Master Mix (Applied Biosystems) under the following conditions: 55°C for 2 min, 95°C for 10 min, and 40 cycles of 95°C for 30 s and 58°C for 60 s. The oligonucleotides 5'-ACGAGCGTCTAACGCTGAAGG-3' and 5'-CCGCTGCGCCGCTGGGTAGAA-3' were used as primers. PCR efficiency and specificity were evaluated using serial dilutions of the template cDNA and analyzing melting curve data. Enolase mRNA (primers: 5'-AACGGCCATTGAGAAGAAGGC-3' and 5'-TGCAGAACTGGA TGCCTCGAT-3') was used as an endogenous control, since its levels remain unchanged during heat shock (data not shown). Relative quantifications were done to determine the fold change in mRNA levels during heat shock. Normalized *HSP104* mRNA levels were obtained for each temperature using the comparative CT method expressed by the formula $2^{-\Delta CT}$, where ΔCT is the difference between the respective CT of *HSP104* mRNA and enolase mRNA in each experiment.

Western blot analysis

Western blots were performed as previously described (Kim et al., 1994). Briefly, *T. cruzi* total protein extracts were separated on 10% SDS-PAGE and transferred to nitrocellulose membranes (Bio-Rad). Membranes were probed with anti-*L. major* HSP104 antibody (Hübel et al., 1995) (diluted 1:500) kindly provided by Dr. Joachim Clos (Bernhard Nocht Institute, Germany). Bound antibodies were detected with a rabbit anti-chicken secondary antibody coupled to peroxidase (diluted 1:10,000) and revealed using an ECL Plus kit (GE Healthcare) according to manufacturer instructions. Protein bands were stained with Coomassie brilliant blue.

Molecular modeling

The predicted 3-D model of *T. cruzi* HSP104 was obtained by comparative modeling based on the crystal structure of *Thermus thermophilus* ClpB (Lee et al., 2003); PDB ID: 1QVR. Sequence alignment and model building were carried out using MODELLER v6 (Sali and Blundell, 1993; Eswar et al., 2007). Models were ranked by MODELLER, checked and validated using the PROCHECK program (Laskowski et al., 1993), and analyzed with the Swiss-PDB Viewer (Guex and Peitsch, 1997), PyMOL (Schrödinger, 2011), and VMD (Humphrey et al., 1996) programs.

RESULTS AND DISCUSSION

Preliminary similarity searches of the publicly available *T. cruzi* genome revealed several HSP100-like sequences, two of which corresponded to partial HSP104-coding regions. Further analysis revealed a gap in the genome sequence, which was resolved by PCR amplification and sequencing of the corresponding genomic region (data not shown). The complete *T. cruzi* HSP104-

coding region is 2604-nucleotides long and encodes an 869-amino acid protein (Figure 1). Alignment with HSP104 sequences from other organisms showed that the *T. cruzi* sequence is conserved (Figure 1A) and contains domains characteristic of the HSP104 family, such as an N-terminal domain and Walker A- and B-type NBDs (Figure 1B). As expected, the *T. cruzi* sequence shared the highest degree of similarity with that of *T. brucei* (Gottesman et al., 1990) and *L. major* (Hübel et al., 1995) (76 and 73% identity, respectively), and the lowest degree with the human protein (23%) (Figure 1A). The observed degree of similarity among trypanosomatids is an intermediate between highly conserved HSPs such as HSP70 (Engman et al., 1989) and HSP10 (Fernandes et al., 2005), and less conserved HSPs such as HSP60 (Giambiagi-de Marval et al., 1993) and sHSP16 (Pérez-Morales et al., 2009), and results from the conservation are mostly limited to the NBD regions. Interestingly, the *T. cruzi* sequence was more similar to that of the bacterium *T. thermophilus* than to that of yeast (55 and 47%, respectively), previously described for the *Leishmania* sequence (Hübel et al., 1995) as well as for kinetoplastid HSP78 proteins (Folgueira and Requena, 2007), and possibly suggesting a mitochondrial function (Folgueira and Requena, 2007).

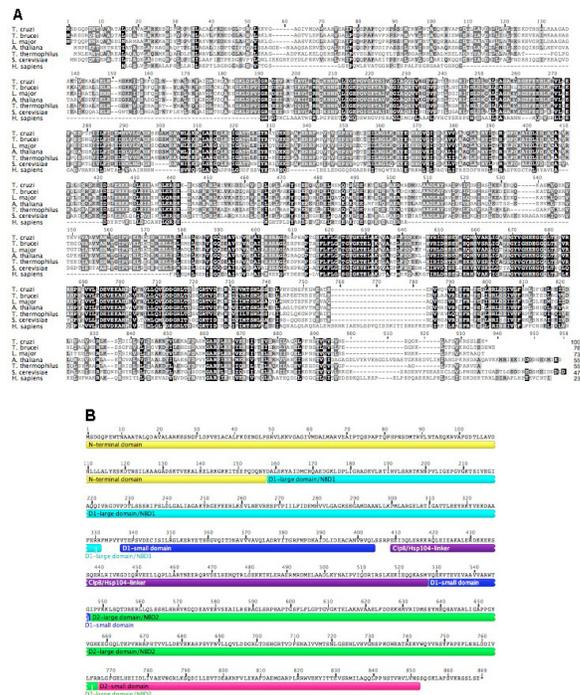


Figure 1. A. Alignment of the *Trypanosoma cruzi* HSP104 sequence with some orthologs from eukaryotic and prokaryotic species. Identical or highly similar amino acid residues are shown in black; similar residues are shaded. The degree of identity relative to the *T. cruzi* sequence is shown on the bottom right. The following HSP104 sequences were used: *T. cruzi* (GenBank ID: HM014310); *T. brucei* (GenBank ID: AAQ16055); *Leishmania major* (GenBank ID: CAA86116); *Arabidopsis thaliana* (GenBank ID: AAA67927); *Thermus thermophilus* (GenBank ID: AAS81465); *Saccharomyces cerevisiae* (GenBank ID: AAA50477); *Homo sapiens* (GenBank ID: AAH06404). B. Conserved domains of *T. cruzi*'s HSP104 are indicated as rectangles below the sequence: N-terminal domain in yellow, D1-large domain/NBD1 in light blue, D1-small domain in dark blue, ClpB/Hsp104-linker in purple, D2-large domain/NBD2 in green, and D2-small domain in pink. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.

To investigate *T. cruzi* *HSP104* gene expression, we detected *HSP104* transcript levels by qRT-PCR in epimastigote cells incubated at different temperatures. As is frequently the case for *HSP* mRNAs in trypanosomatids (de Carvalho et al., 1990; Hübel et al., 1995; Krobitch et al., 1998; Fernandes et al., 2005; Pérez-Morales et al., 2009), *HSP104* transcripts were detected at normal temperatures (Figure 2A, open bar). Transcript levels increased approximately 3-fold when the cells were incubated for 3 h at 37°C (Figure 2A, filled bar), but increased less than 2-fold at 40°C (Figure 2A, hatched bar). A similar increase (2- to 3-fold) in *HSP104* mRNA levels was detected at 34° and 37°C in *L. major* (Hübel et al., 1995). To investigate *HSP104* protein levels in heat-shocked epimastigote cells, we performed Western blot analysis of total protein extracts from cells incubated at 29°, 37°, and 40°C. As can be seen in Figure 2B, *HSP104* was detected at low levels at normal temperatures, in agreement with what has been described in *L. major* (Hübel et al., 1995) and *L. donovani* (Krobitch et al., 1998). Incubation at 37°C for 3 h increased protein levels by approximately 2-fold, somewhat lower than the observed increase in mRNA levels (Figure 2A and B). Further incubation for up to 24 h at the same temperature gradually increased the abundance of the protein, suggesting continuous synthesis throughout the incubation period. These results are also in agreement with what has been described in *L. major* (Hübel et al., 1995). Although *T. cruzi* cells cannot withstand prolonged incubation at 40°C, heat treatment at 40°C for 3 h led to increased levels of the *HSP104* protein, similar to the corresponding increase in mRNA level (Figure 2A and B). Induction of *HSP70* in *T. cruzi* also results from an increase in both mRNA and protein levels during heat shock (de Carvalho et al., 1990), and modulation of both mRNA stability and translation have been described as gene regulatory mechanisms (de Carvalho et al., 1990; Rodrigues et al., 2010). Further study will be required to identify the regulatory mechanisms of *HSP104* gene expression in this organism.

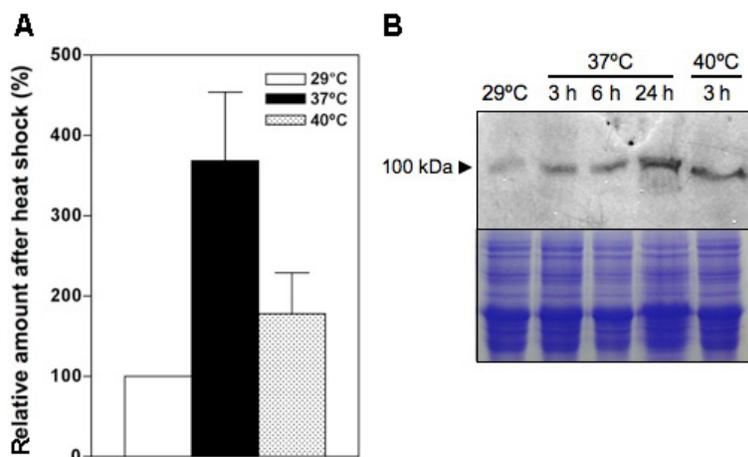


Figure 2. Analysis of *Trypanosoma cruzi* *HSP104* gene expression during heat shock. **A.** qRT-PCR was performed in total RNA from epimastigotes incubated for 3 h at 29°C (open bar), 37°C (filled bar) and 40°C (hatched bar). Transcript levels are shown as a percentage of the value at 29°C. Error bars indicate standard deviations of three independent experiments. **B.** Western blot of total protein extracts from 10^8 epimastigote cells incubated at 29°C, 37°C (3, 6 and 24 h) and 40°C (3 h) probed with an anti-*Leishmania major* *HSP104* antibody. Arrowhead indicates the migration of polypeptides with 100-kDa molecular mass (above). A Coomassie brilliant blue staining of the gel is shown as a protein load control (below).

The *T. cruzi* HSP104 is a large protein complex where each monomer comprises 869 amino acids, probably organized as a hexamer comprising more than 5000 amino acids. Since the protein of this parasite is markedly different from the human ortholog (see alignment in Figure 1A), investigating the structural aspects of *T. cruzi* HSP104 may be relevant to potential drug or vaccine targets. However, determination of the full-3-D structure of such a large complex by X-ray diffraction is not yet available in structure databases. We therefore decided to construct the *T. cruzi* HSP104 3-D model using the comparative modeling technique (Marti-Renom et al., 2000; Eswar et al., 2008). We used the crystallographic structure at 3.0-Å resolution of a subunit of the *T. thermophilus* ClpB hexamer as the template. The sequence alignment between *T. cruzi* HSP104 and *T. thermophilus* ClpB was calculated using the align2d module of MODELLER, a global dynamic programming algorithm that takes into account structural information from the template when constructing the alignment (Eswar et al., 2007). The sequences shared 51% amino acid identity, 70% amino acid similarity, and more than 99% sequence coverage. We used MODELLER to construct candidate structures, and the selected structures had the maximum GA341 score of 1.0. The candidate structures were ranked using the Discrete Optimized Protein Energy (DOPE) score and PROCHECK. The Ramachandran plot for the best 3-D model for *T. cruzi* HSP104 showed 91.3% of residues in the favored region, 7.3% in additional allowed regions, 1.2% in generously allowed regions, and 0.3% (two residues) in disallowed regions, which should be considered a very good model. Finally, using the crystal symmetry of the template, we were able to construct a model for the *T. cruzi* HSP104 hexamer, as shown in Figure 3A.

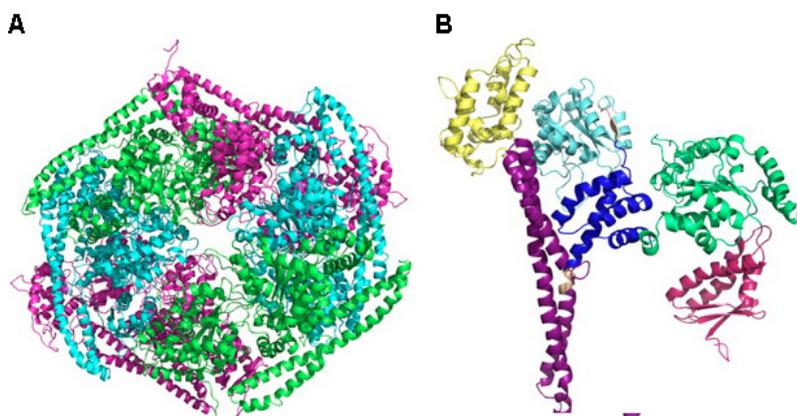


Figure 3. **A.** Ribbon model of the 3-D model of *Trypanosoma cruzi* HSP104 hexamer. For clarity, HSP104 subunits are differentially colored. **B.** Ribbon model of one *T. cruzi* HSP104 monomer. All secondary structure elements are labeled as in Figure 1B: N-terminal domain in yellow, D1-large domain in light blue, D1-small domain in dark blue, ClpB/Hsp104-linker in purple, D2-large domain in green, and D2-small domain in pink.

The structure of this ClpB chaperone may be divided into five distinct domains as proposed for the ClpB protein from *T. thermophilus* (Lee et al., 2003). These domains are represented in the protein sequence in Figure 1B and in the *T. cruzi* HSP104 monomer model in Figure 3B: N-terminal domain (yellow); D1-large domain or NBD1 (light blue); D1-small domain (dark blue); ClpB/Hsp104-linker (purple); D2-large domain or NBD2 (green), and

D2-small domain (pink), containing motif 1 and motif 2 leucine-rich regions, which have an important function in the disaggregation and solubilization of other proteins. The structural similarity implied by the amino acid sequence similarity suggests the expected functional conservation between *T. cruzi* HSP104 and *T. thermophilus* ClpB. Additional studies should improve our understanding of HSP104 function in this organism.

ACKNOWLEDGMENTS

We thank César Félix Schmidt and Cláudio Nunes Pereira for technical assistance, and Dr. J. Clos (Bernhard Nocht Institute for Tropical Medicine, Germany) for the HSP104-specific antibody. Research supported by Grants from the National Council for Scientific and Technological Development (CNPq), the Foundation for the Coordination of Improvement of Higher Education Personnel (CAPES), and the Foundation for Research Funding of Rio de Janeiro State (FAPERJ), Brazil.

REFERENCES

- Atwood JA III, Weatherly DB, Minning TA, Bundy B, et al. (2005). The *Trypanosoma cruzi* proteome. *Science* 309: 473-476.
- Cano MI, Gruber A, Vazquez M, Cortes A, et al. (1995). Molecular karyotype of clone CL Brener chosen for the *Trypanosoma cruzi* genome project. *Mol. Biochem. Parasitol.* 71: 273-278.
- Coura JR and Borges-Pereira J (2010). Chagas disease: 100 years after its discovery. A systemic review. *Acta Trop.* 115: 5-13.
- de Carvalho EF, de Castro FT, Rondinelli E, Soares CM, et al. (1990). HSP 70 gene expression in *Trypanosoma cruzi* is regulated at different levels. *J. Cell Physiol.* 143: 439-444.
- Doyle SM and Wickner S (2009). Hsp104 and ClpB: protein disaggregating machines. *Trends Biochem. Sci.* 34: 40-48.
- Dragon EA, Sias SR, Kato EA and Gabe JD (1987). The genome of *Trypanosoma cruzi* contains a constitutively expressed, tandemly arranged multicopy gene homologous to a major heat shock protein. *Mol. Cell Biol.* 7: 1271-1275.
- Drummond AJ, Ashton B, Buxton S, Cheung M, et al (2011). Geneious v5.4. Geneious. Available at [http://www.geneious.com]. Accessed July 24, 2011.
- Engman DM, Kirchhoff LV and Donelson JE (1989). Molecular cloning of mtp70, a mitochondrial member of the hsp70 family. *Mol. Cell Biol.* 9: 5163-5168.
- Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, et al. (2007). Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* Chapter 2: Unit.
- Eswar N, Eramian D, Webb B, Shen MY, et al. (2008). Protein structure modeling with MODELLER. *Methods Mol. Biol.* 426: 145-159.
- Fernandes M, Silva R, Rossle SC, Bisch PM, et al. (2005). Gene characterization and predicted protein structure of the mitochondrial chaperonin HSP10 of *Trypanosoma cruzi*. *Gene* 349: 135-142.
- Folgueira C and Requena JM (2007). A postgenomic view of the heat shock proteins in kinetoplastids. *FEMS Microbiol. Rev.* 31: 359-377.
- Giambiagi-de Marval M, Gottesdiener K, Rondinelli E and Van der Ploeg LH (1993). Predicted amino acid sequence and genomic organization of *Trypanosoma cruzi* hsp 60 genes. *Mol. Biochem. Parasitol.* 58: 25-31.
- Gottesman S, Squires C, Pichersky E, Carrington M, et al. (1990). Conservation of the regulatory subunit for the Clp ATP-dependent protease in prokaryotes and eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 87: 3513-3517.
- Guex N and Peitsch MC (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18: 2714-2723.
- Hartl FU and Hayer-Hartl M (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* 295: 1852-1858.
- Hübel A, Brandau S, Dresel A and Clos J (1995). A member of the ClpB family of stress proteins is expressed during heat shock in *Leishmania* spp. *Mol. Biochem. Parasitol.* 70: 107-118.
- Humphrey W, Dalke A and Schulten K (1996). VMD: visual molecular dynamics. *J. Mol. Graph.* 14: 33-38.
- Kim KS, Teixeira SM, Kirchhoff LV and Donelson JE (1994). Transcription and editing of cytochrome oxidase II RNAs

- in *Trypanosoma cruzi*. *J. Biol. Chem.* 269: 1206-1211.
- Krobitsch S, Brandau S, Hoyer C, Schmetz C, et al. (1998). *Leishmania donovani* heat shock protein 100. Characterization and function in amastigote stage differentiation. *J. Biol. Chem.* 273: 6488-6494.
- Laskowski RA, Moss DS and Thornton JM (1993). Main-chain bond lengths and bond angles in protein structures. *J. Mol. Biol.* 231: 1049-1067.
- Lee S, Sowa ME, Watanabe YH, Sigler PB, et al. (2003). The structure of ClpB: a molecular chaperone that rescues proteins from an aggregated state. *Cell* 115: 229-240.
- Martí-Renom MA, Stuart AC, Fiser A, Sanchez R, et al. (2000). Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 29: 291-325.
- Pérez-Morales D, Ostoa-Saloma P and Espinoza B (2009). *Trypanosoma cruzi* SHSP16: Characterization of an alpha-crystallin small heat shock protein. *Exp. Parasitol.* 123: 182-189.
- Rodrigues DC, Silva R, Rondinelli E and Urmenyi TP (2010). *Trypanosoma cruzi*: modulation of HSP70 mRNA stability by untranslated regions during heat shock. *Exp. Parasitol.* 126: 245-253.
- Rondinelli E (1994). Conservation of heat-shock proteins in *Trypanosoma cruzi*. *Parasitol. Today* 10: 172-176.
- Sali A and Blundell TL (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234: 779-815.
- Schrödinger LCC (2011). The PyMol Molecular Graphics System. Version 1.3. pymol.org. Available at [<http://www.pymol.org/>]. Accessed August 22, 2011.
- Vonlaufen N, Kanzok SM, Wek RC and Sullivan WJ Jr (2008). Stress response pathways in protozoan parasites. *Cell Microbiol.* 10: 2387-2399.