HIGH-CONFIDENCE STRUCTURAL MODEL OF THE *Trypanosoma cruzi* GP63-LIKE METALLOPROTEASE GENERATED BY ALPHAFOLD AND VALIDATED THROUGH MOLECULAR DYNAMICS SIMULATIONS

Emily Cristiny Silva Santos[1], Jessica Silva dos Santos[1], Alex Santos Macedo[1], Paulo Henrique Matayoshi Calixto[2]

[1] Goiano Federal Institute
[2] Computational Biology Unit - Goiano Federal Institute

E-mail para correspondência: paulo.calixto@ifgoiano.edu.br

Abstract

The 63 kDa glycoprotein (TcGP63) is the major surface metalloprotease of *Trypanosoma cruzi* and plays a central role in cell invasion and immune evasion processes. Despite its biological relevance, the three-dimensional structure of this enzyme has not yet been experimentally determined. In this study, TcGP63 was modeled using AlphaFold, and the obtained model was evaluated for stereochemical, energetic, and conformational quality. The predicted structure presented high confidence values (pTM = 0.93; pLDDT > 90), indicating the robustness and accuracy of the model. The conserved catalytic motif HE XX H and the coordinated Zn^{2+} ion were clearly identified, with average coordination distances compatible with other metalloproteases of the metzincin family. Dynamic validation was conducted through 25 ns simulations using the GROMOS54a7 force field. Analyses of RMSD, radius of gyration (Rg), and RMSF demonstrated conformational stability and structural compactness throughout the simulation, with fluctuations localized only at the C-terminal extremity. Collectively, the results confirm the structural integrity of TcGP63 and provide a solid basis for molecular docking studies and the rational design of inhibitors aimed at controlling Chagas Disease.

Keywords: *Trypanosoma cruzi*; GP63; AlphaFold; molecular dynamics; structural modeling.

MODELAGEM ESTRUTURAL DE ALTA CONFIANÇA DA METALOPROTEASE GP63-LIKE DE *Trypanosoma cruzi* OBTIDA POR ALPHAFOLD E VALIDADA POR DINÂMICA MOLECULAR

Resumo

A glicoproteína de 63 kDa (TcGP63) é a principal metaloprotease de superfície de *Trypanosoma cruzi* e desempenha papel central nos processos de invasão celular e evasão imunológica. Apesar de sua relevância biológica, a estrutura tridimensional dessa enzima ainda não foi determinada experimentalmente. Neste estudo, a TcGP63 foi modelada utilizando o AlphaFold, e o modelo obtido foi avaliado quanto à qualidade estereoquímica, energética e conformacional. A estrutura predita apresentou altos valores de confiança (pTM = 0,93; pLDDT > 90), indicando robustez e acurácia do modelo. O motivo catalítico conservado HE*XX*H e o íon Zn^{2+} coordenado foram claramente identificados, com distâncias médias de

coordenação compatíveis com outras metaloproteases da família das metzincinas. A validação dinâmica foi conduzida por meio de simulações de 25 ns com o campo de força GROMOS54a7. As análises de RMSD, raio de giro (Rg) e RMSF demonstraram estabilidade conformacional e compacidade estrutural ao longo da simulação, com flutuações localizadas apenas na extremidade C-terminal. Em conjunto, os resultados confirmam a integridade estrutural da TcGP63 e fornecem uma base sólida para estudos de docagem molecular e desenho racional de inibidores voltados ao controle da Doença de Chagas.

Palavras-chave: *Trypanosoma cruzi*; GP63; AlphaFold; dinâmica molecular; modelagem estrutural.

1. Introduction

Chagas Disease, a parasitic zoonosis caused by the protozoan *Trypanosoma cruzi*, represents a significant public health challenge, affecting millions of people, primarily in Latin America (WHO, 2025). The persistence of infection and progression to the chronic phases of the disease, which frequently involve severe cardiomyopathy, are intrinsically linked to the parasite's ability to invade a wide range of mammalian cells and evade the host's innate and adaptive immune responses. To mediate these complex processes, *T. cruzi* utilizes a diversified repertoire of surface molecules, among which metalloproteases stand out (BUNKOFSKE et al., 2025).

The 63 kDa glycoprotein (GP63), or TcGP63, is the major surface metalloprotease of *T. cruzi* and is considered a key virulence factor (BERNÁ et al., 2025). Homologous to the leishmanolysin of *Leishmania*, TcGP63 has been shown to be multifunctional, actively participating in cell adhesion, degradation of the extracellular matrix to facilitate tissue invasion, and cleavage of complement system components, serving as a crucial immune evasion mechanism (KULKARNI et al., 2009; BERNÁ et al., 2025).

Despite its central importance in pathogenesis, the determination of the three-dimensional structure of TcGP63 by experimental methods, such as X-ray crystallography or cryogenic electron microscopy, has not yet been achieved. Structural knowledge is indispensable for the rational design of inhibitors that can selectively block its active site (BARRETO; LIMA, 2020). Although the structure of its homolog in *Leishmania major* (LmGP63) has been resolved (SCHLAGENHAUF; ETGES; METCALF, 1998), sequence variations between species, especially in loop regions that influence substrate specificity, limit

traditional homology modeling for TcGP63 (RIGO et al., 2023).

Exponential advances in computational methods, driven by artificial intelligence, have drastically changed the landscape of structural biology (ABRAMSON et al., 2024). The advent of AlphaFold, in particular, provided a means for protein structure prediction with near-experimental accuracy, even for proteins without closely resolved structural templates (JUMPER et al., 2021). This methodology offers an unprecedented opportunity to bridge the structural gap of TcGP63.

In this work, we present the first high-fidelity structural model of the *T. cruzi* metalloprotease TcGP63, generated using AlphaFold. The 3D architecture of the model was subjected to rigorous computational validation to assess its global stereochemical and energetic quality (LASKOWSKI et al., 1993). Furthermore, to investigate the conformational behavior of the protein in solution, we performed molecular dynamics (MD) simulations using the GROMOS54a7 force field (SCHMID et al., 2011). The structural and dynamic characterization of TcGP63 provides crucial insights into its catalytic mechanism and establishes a robust foundation for future virtual docking studies in the development of novel trypanocidal agents.

2. Material and Methods

2.1 Acquisition of the TcGP63 sequence and primary analyses

The sequence of the TcGP63 surface glycoprotein was obtained from the UniProt database (accession number: Q4CP46). The region corresponding to the mature form of the protein was analyzed using the ProtParam program (ExPASy) to determine physicochemical parameters such as theoretical molecular mass, isoelectric point, and amino acid composition. The presence of a signal peptide was evaluated using SignalP 6.0 (TEUFEL et al., 2022), while the potential GPI anchoring site was predicted with PredGPI (PIERLEONI et al., 2008).

2.2 Comparative sequence analysis

The alignment of TcGP63 and its homolog LmGP63 sequences was performed using the Clustal Omega software, available on the EMBL-EBI platform. Global identity and similarity levels were evaluated, as well as the conservation of catalytic and structural residues, including the HE XX H motif and disulfide bridge-forming cysteine residues. The structural conservation analysis was interpreted based on previous studies on

trypanosomatid metalloproteases (CHAUDHURI; CHANG, 1989; YAO et al., 2003; STEINKRAUS et al., 1993; CALIXTO et al., 2013).

2.3 Three-dimensional structural modeling

Due to the low sequence identity between TcGP63 and the crystallographic structures available in the Protein Data Bank (PDB), homology modeling was considered inadequate. Thus, the three-dimensional structure of TcGP63 was predicted using AlphaFold (JUMPER et al., 2021), a deep learning tool based on convolutional neural networks capable of predicting protein structures with near-experimental accuracy. The generated model was evaluated based on the pTM (Predicted Template Modeling score) and pLDDT (Predicted Local Distance Difference Test) metrics. The stereochemistry and geometric quality of the model were verified with the PROCHECK program (LASKOWSKI et al., 1993), through Ramachandran plot analysis.

2.4 Structural analysis and identification of the catalytic site

The three-dimensional architecture of TcGP63 was explored in UCSF ChimeraX, highlighting the N-terminal, central, and C-terminal subdomains. The catalytic site was identified by comparison with the structure of LmGP63 (PDB ID: 1LML), evidencing the conserved HE*XX*H motif and the third histidine coordinating the Zn^{2+} ion. Zinc coordination distances were determined by direct geometric measurements, and disulfide bridges were identified and created using the covalent bond analysis module of UCSF ChimeraX.

2.5 Molecular Dynamics Simulation

The stability and conformational flexibility of TcGP63 in solution were evaluated by molecular dynamics (MD) simulations conducted in the GROMACS 2025.3 software, employing the GROMOS54a7 force field (SCHMID et al., 2011; PÁLL; ABRAHAM; HESS, 2022). The system was solvated in a dodecahedral box containing the SPC water model, maintaining a minimum distance of 10 Å between the protein and the box edges. Na^+ and Cl^- ions were added to neutralize the total charge of the system. After energy minimization for 5,000 steps of Steepest Descent, the system was equilibrated in two stages: (i) NVT ensemble for 100 ps, followed by (ii) NPT ensemble for 100 ps, both at 310 K. The production simulation was conducted for 25 ns, with a time integration step of 2 fs. Structural stability analyses were based on RMSD, RMSF, and

Radius of Gyration parameters, calculated with internal GROMACS tools.

2.6 Electrostatic analysis and structural comparison

The molecular surface electrostatic potential was calculated with APBS (Adaptive Poisson–Boltzmann Solver) (JURRUS et al., 2018). The structural alignment between TcGP63 and LmGP63 was performed using UCSF ChimeraX, and the superposition of the structures was evaluated by the root-mean-square deviation (RMSD).

3. Results and Discussion

3.1 Acquisition of the TcGP63 sequence and preliminary analyses

The analysis of the primary mature sequence of TcGP63, conducted by the ProtParam program, indicated an approximate molecular mass of 50 kDa. The prediction of signal regions performed by SignalP (TEUFEL et al., 2022) identified the segment comprised between Met1 and Ala22 as the signal peptide, responsible for directing the protein to the secretory pathway (Fig. 1). Complementarily, PredGPI (PIERLEONI et al., 2008) predicted the segment downstream of residue Ser522 as the cleavage point for the addition of the GPI group. Together, these evidences suggest that TcGP63 assembles all the structural elements necessary for its localization on the cell surface of *T. cruzi*.

The alignment of the amino acid sequence of TcGP63 with the *L. major* homolog (LmGP63) revealed the presence, in the N-terminal portion, of a conserved segment corresponding to the propeptide (Glu23–Ser60). This region contains a cysteine residue (Cys26) (Fig. 1), essential for the cysteine switch mechanism, responsible for regulating enzymatic activity. Such a mechanism prevents the newly synthesized enzyme from exerting proteolytic activity on intracellular components of the parasite, ensuring its activation only after proper processing (KIM et al., 2023).

MRHTMLLLVPLLCCVSGSVAVAEHH**C**ISDEIEKKVGPRTTAVVLELPTRGGGMIRALTAS
DPDWAPIRFQFFTEDLNDPSRYCTAEGQIRPDFTGGTVECKREDILKEEKKSIILKSLVP
RALKMHTDRLLVRPLMGRVIVPEFLSGVCAQFTIPSSHQIEGVTGADMYLYVSAAPVKGS
ALAWATSCSALPDGRPVVGVVNYGPSSVTDSEYSVRVVVHEIGHALGFAVEIMEERNMLK
EVKGVRGKAKVLQVSSPKTVEKTREHFNCVNATGMELEDEGGERTASSHWKRRNAKDELM
AGNEGIGYYTALTMAAFEDTGFYRANWGKEEPMSWGNNSGCALLTEKCVINGVTKYPEMF
CTAESRLFSCTSDRLGLGHCTIELYDAPLPPQYQYFSNPKLGGSPGFFMDFCPYIEAYFN
TWCTDGEADVWMRGSRVGPTSKCLKGDGLADFMGRIGDVCAEVSCDKGEVSVRYLGDDAW
HKCPEGSSITPTGLFTGGRILCPKYDDVCIVFNTINGTGDVSSLLSAFPPIPLVMLVLIF
ISMC

Figure 1. Elements of the TcGP63 amino acid sequence. The residues corresponding to the signal peptide are indicated in green. The amino acids highlighted in blue represent the propeptide, while the cysteine residue in bold indicates the site involved in the cysteine switch mechanism. The residues shown in red correspond to the region cleaved for GPI anchor addition. The catalytic site is formed by the amino acids highlighted in violet. Finally, the residues represented in black correspond to the mature form of the protein, following post-translational modifications. Source: the authors.

3.2 Conservation of GP63 sequences

Comparative analyses available in the literature demonstrate that GP63s from different *Leishmania* species exhibit a high degree of structural and functional conservation (GÜVENDI et al., 2025). The main exception occurs in *Leishmania guyanensis*, whose predicted sequence presents an absence of the five terminal cysteine residues; however, the activity or stability of this variant has not yet been experimentally evaluated (STEINKRAUS et al., 1993). Considering these observations, a local alignment involving GP63 sequences from *L. major* (LmGP63) and *T. cruzi* (Fig. 2) was performed. The alignment revealed a high degree of conservation along the sequence, including the residues of the catalytic motif HE XX H and the third histidine involved in active center coordination. Generally, about 40% of residues are identical among all analyzed proteins, with notable conservation of 18 cysteines, 21 glycines, and 11 prolines, amino acid residues that contribute to the maintenance of the enzyme's architecture and stability.



Figure 2. Alignment between the amino acid sequences of TcGP63 and LmGP63. The sequences presented correspond to the mature form of the proteins. Residues were colored according to identity (blue corresponds to identical residues) to facilitate visualization of the degree of conservation between the sequences. Source: the authors.

3.3 Optimization and validation of the model

Verification of structures homologous to TcGP63 was performed using BLAST (Basic Local Alignment Search Tool) and the Protein Data Bank. The only identified structure was that of LmGP63 (PDB code: 1LML), whose identity and similarity were 39% and 55%, respectively. The low degree of identity between the sequences rendered model construction using comparative modeling techniques unfeasible. The construction of the 3D model of TcGP63 was performed employing AlphaFold, which utilizes deep neural networks for model generation. The structure presented a pTM of 0.93. The pTM (Predicted Template-Modeling score) is a global confidence metric used by AlphaFold to evaluate protein structure quality. Its value ranges from 0 (low quality) to 1 (high quality) (MEIER; SÖDING, 2015). The pLDDT (Predicted Local Distance Difference Test) presented indices superior to 90, indicating the high quality and reliability of the modeled structure (MARIANI et al., 2013). Stereochemical quality was evaluated with the aid of the Procheck software (LASKOWSKI et al., 1993), which presented satisfactory results for this parameter, as no amino acid residue is found in unfavorable regions (Fig. 3).

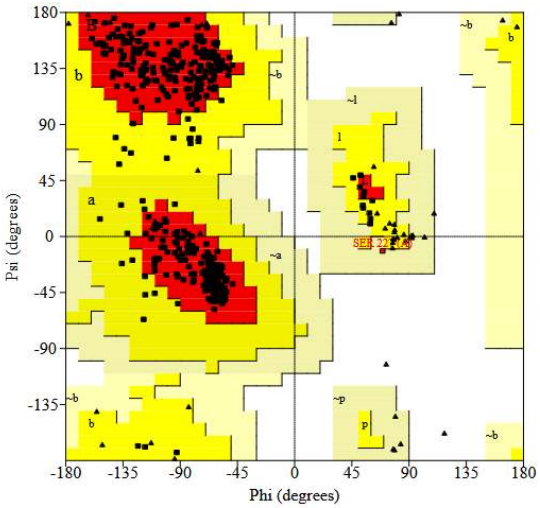

Figure 3. Ramachandran Plot. Residues in most favorable regions (A,B,L): 350 (90.2%); residues in additionally allowed regions (a,b,l,p): 37 (9.5%); residues in generously allowed regions (~a,~b,~l,~p): 1 (0.3%); residues in disallowed regions: 0 (0%). Source: the authors.

3.4 General structure of TcGP63

The TcGP63 protein presents a compact three-dimensional organization, composed of three main subdomains: N-terminal, central, and C-terminal (Fig. 4). The N-terminal subdomain, extending from residues Asp1 to Gly167, exhibits an architecture similar to the typical catalytic module of zinc-dependent proteases. In this region (Fig. 4C), the constituent residues of the catalytic site HE*XX*H are located His160, Glu161, and His164 which correspond to the conserved motif characteristic of metalloproteases.

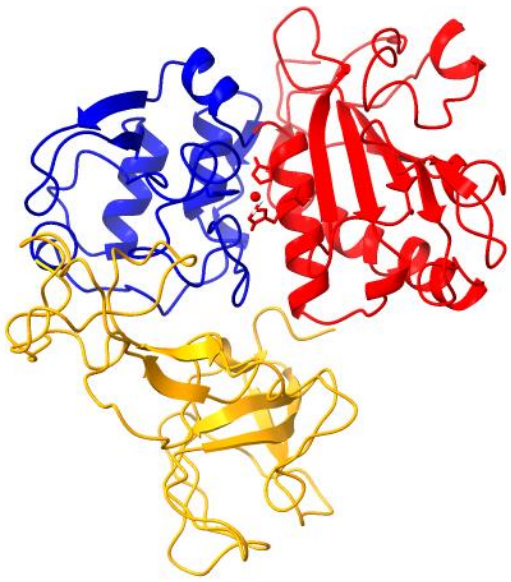

Figure 4. Three-dimensional structure of TcGP63 obtained using AlphaFold. Representation of the tertiary conformation of the protein, highlighting the N-terminal, Central, and C-terminal subdomains, indicated by red, blue, and yellow colors, respectively. Source: the authors.

The central subdomain, comprised between Phe168 and Thr286, is formed by a compact core constituted by α-helices and β-sheets arranged in an antiparallel manner. This region connects to the C-terminal subdomain through a single disulfide bond. In metzincins, the extended catalytic motif HE*XX*H*XX*G*XX*H contains a glycine residue that participates in a structural loop responsible for positioning the third histidine residue in the catalytic site, allowing its coordination to the zinc atom. In TcGP63, however, an insertion of 61 residues is observed between Gly167 and His229, the latter also being involved in metal coordination, a characteristic present in all described leishmanolysins (SCHLAGENHAUF et al., 1998; RAZZAZAN et al., 2008).

The configuration of residues interacting with zinc is comparable to that observed in other proteases of this type. The metal is coordinated to the side chains of histidines 160, 164, and 229, with distances of 2.1 Å, 2.2 Å, and 2.1 Å, respectively (Fig. 5). Furthermore, a water molecule participates in this coordination, acting as a nucleophile in the attack on the substrate during the catalytic process.

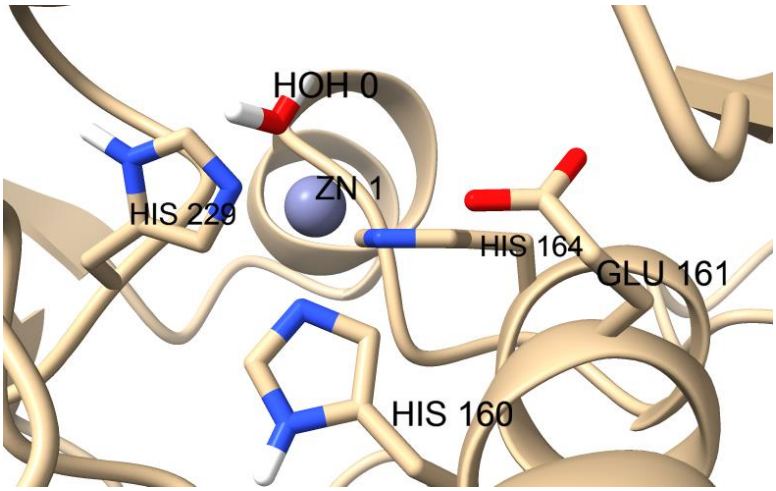

Figure 5. Catalytic center. Representation of the active site of TcGP63, evidencing the three histidine residues and the catalytic water molecule coordinated to the zinc ion. The glutamate residue acts in polarizing the water molecule, promoting nucleophilic attack during catalysis. This structural arrangement is characteristic of zinc-dependent metalloproteases of the metzincin family, being essential for peptide bond cleavage and the enzyme's hydrolytic activity. Source: the authors.

The C-terminal subdomain, delimited between residues Lys287 and Asn452, concentrates six of the nine disulfide bridges of the molecule (Fig. 6), conferring high structural stability. The final residues of this region compose the portion destined for GPI anchoring, organized in a spiral loop.

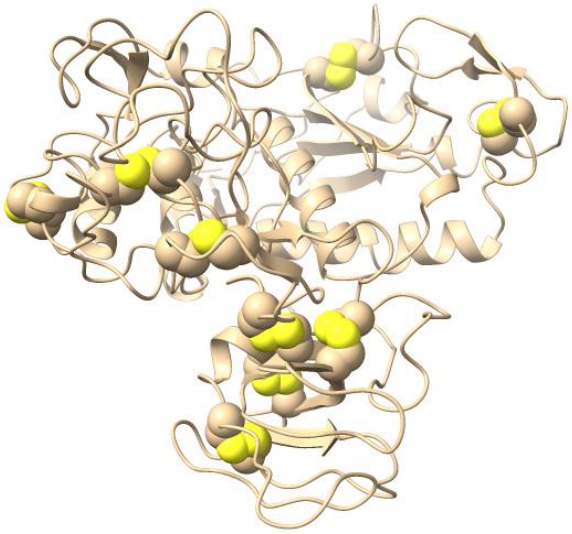

Figure 6. Three-dimensional structure of TcGP63. The yellow sphere clusters represent cysteine residues involved in disulfide bond formation. Of the nine disulfide bonds identified, six are located in the C-terminal domain, indicating greater structural rigidity in this region of the protein. Source: the authors.

To evaluate the degree of structural conservation between TcGP63 and LmGP63, a comparative analysis between the structures was conducted (Fig. 7). The obtained structural alignment revealed high correspondence between both proteins, especially in regions corresponding to the central and N-terminal subdomains. In contrast, the C-terminal subdomain presented lower conservation, although this variation is expected, as differences in this region have been consistently observed among GP63s, including in isoforms of the same species (YAO et al., 2003).



Figure 7. Structural alignment. The structures represented in green and red correspond to the LmGP63 and TcGP63 proteins, respectively. The alignment revealed an RMSD value of 0.180 Å, indicating a high degree of structural conservation. Source: the authors.

The RMSD (Root-mean-square deviation) value calculated between the two structures was 0.180 Å, confirming the high conformational similarity between them. Considering that enzyme-substrate affinity depends not only on active site topology but also on electrostatic complementarity at the interaction interface (GRASSMANN et al., 2023), an analysis of partial charge distribution on the molecular surface was proceeded (Fig. 8).

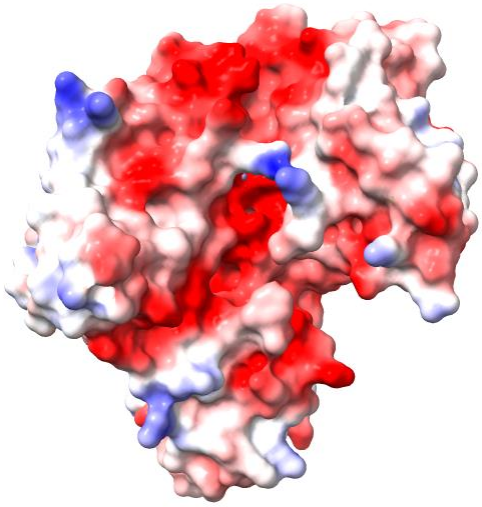

Figure 8. Surface charge distribution of TcGP63. Positively charged, negatively charged, and neutral regions are represented by blue, red, and white colors, respectively. Source: the authors.

Electrostatic mapping revealed that TcGP63 presents a predominance of negative charges in the region corresponding to the catalytic cleft, attributed to the presence of negatively charged residues. This partial charge distribution pattern is similar to that described for LmGP63. This correspondence suggests that both enzymes may share affinity for similar substrates present in vertebrate and invertebrate hosts.

3.5 Analysis of TcGP63 structural stability

The structural stability of the GP63 protein was evaluated via root-mean-square deviation (RMSD), radius of gyration (Rg), and root-mean-square fluctuation per residue (RMSF) analyses over 25 ns of molecular dynamics.

The RMSD plot (Fig. 9) indicates that the GP63 structure presented an initial increase in backbone deviation during the first ~5 ns of simulation, reaching values around 0.6–0.7 nm, corresponding to the

relaxation and conformational adjustment process after solvation and equilibration. From approximately 8–10 ns, the RMSD stabilized around 0.5–0.6 nm, with no significant growth trend until the end of the simulation, suggesting that the system reached conformational equilibrium. This behavior indicates that GP63 maintained its structural integrity without the occurrence of major rearrangements or denaturation during the trajectory.

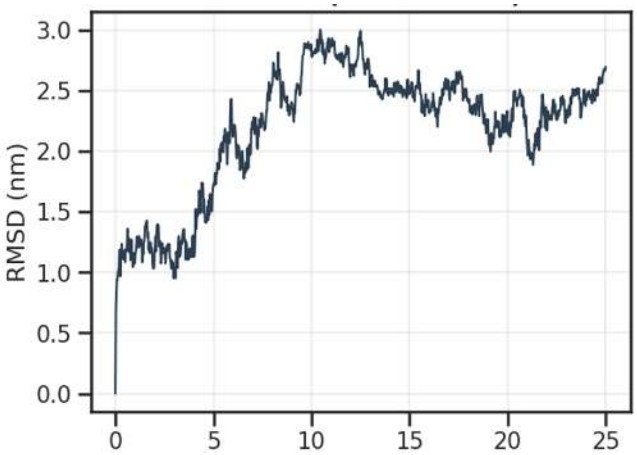

Figure 9. RMSD of TcGP63. The plot shows the variation of backbone RMSD relative to the initial structure over time. After an initial increase in the first 5 ns, the RMSD stabilizes around 0.5–0.6 nm, indicating that the protein reached conformational equilibrium and maintained structural stability during the simulation. Source: the authors.

The analysis of the radius of gyration (Rg) (Fig. 10) reinforces this stability, showing average values close to 2.4 nm with small oscillations over time. After the start of the simulation, a slight reduction in Rg is observed, followed by stabilization from about 10 ns, indicating a compact and consistent structure. The absence of significant fluctuations in Rg demonstrates that the protein did not undergo relevant structural expansions or collapses, maintaining its global compactness.

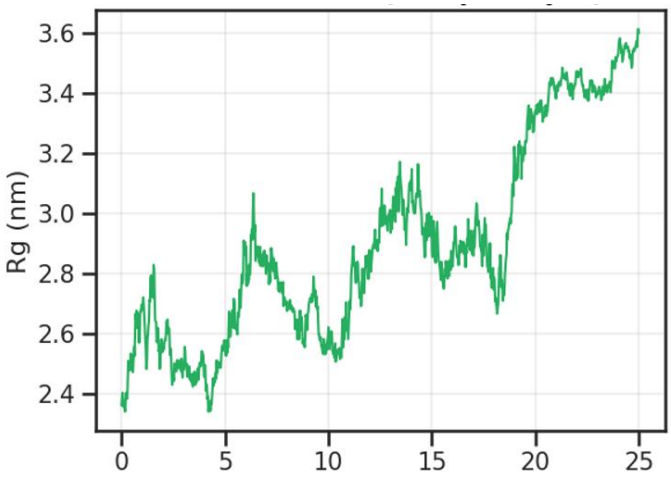

Figure 10. Radius of gyration of TcGP63. The Rg presented average values close to 2.4 nm, with small oscillations over time. The stabilization observed after approximately 10 ns demonstrates maintenance of compactness and absence of significant structural expansion, confirming the global stability of the protein in solution. Source: the authors.

Finally, the RMSF analysis (Fig. 11) revealed low average fluctuations (<0.3 nm) for most residues, indicating low local mobility and adequate structural rigidity. A sharp peak was observed at the C-terminal extreme, which is common in unstructured or solvent-exposed terminal regions and does not compromise the protein's global stability. Additional small fluctuations observed in some regions may be associated with loops or more flexible surface domains (ARGUETA; PARKINS; PANTOURIS, 2024).

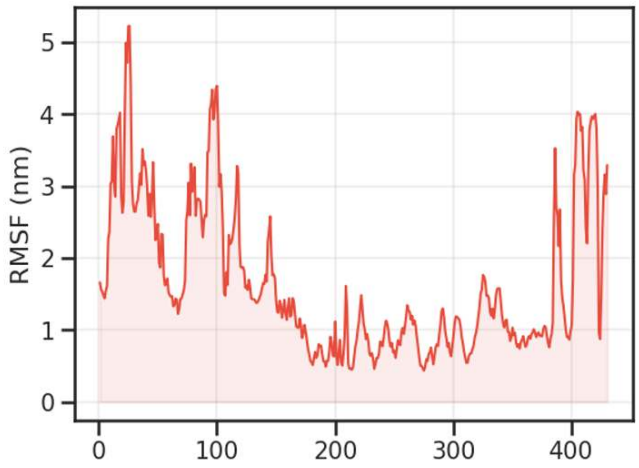

Figure 11. RMSF of TcGP63. The RMSF revealed low fluctuations in the majority of residues, reflecting structural rigidity and conformational stability. The sharp increase observed at the C-terminal extreme corresponds to a flexible region, possibly disordered or solvent-exposed, with no impact on the protein's global stability. Source: the authors.

Collectively, the RMSD, Rg, and RMSF analyses demonstrate that *T. cruzi* GP63 maintained high structural stability throughout the 25 ns simulation, preserving its compaction and global conformation, with only expected local variations in flexible regions.

4. Conclusion

The present study presents the first high-confidence structural model of the TcGP63 metalloprotease, integrating AlphaFold prediction and molecular dynamics validation. The consistent stereochemical quality, allied with robust pTM and pLDDT metrics, demonstrates that the obtained model faithfully represents the enzyme's global architecture, including subdomain organization, the HE*XX*H catalytic motif arrangement, and Zn^{2+} ion coordination. Comparative structural analysis evidences a high degree of conformational conservation relative to LmGP63, reinforcing the hypothesis that essential elements for proteolytic function were preserved throughout trypanosomatid evolution.

Molecular dynamics simulations confirm that TcGP63 maintains conformational stability and structural compactness in solution, with fluctuations restricted to naturally flexible regions, especially at the C-terminus. These results indicate that the model is suitable for subsequent applications, including enzyme-substrate interaction studies, electrostatic interface analyses, and, above all, molecular docking strategies for virtual inhibitor screening.

Besides filling a relevant structural gap for *T. cruzi* biology, the data presented herein establish a solid platform for investigations into the role of TcGP63 in cell invasion and immune evasion, contributing to the understanding of Chagas Disease pathogenesis. The findings also open perspectives for the development of selective inhibitors based on unique structural characteristics of the protein, enabling advances in the rational design of therapeutic agents against this important parasite.

5. References

ABRAMSON, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, v. 630, n. 8016, p. 493-500, 2024.

BARRETO, A. G.; LIMA, P. C. Planejamento racional de fármacos como estratégia para doenças negligenciadas. *Química Nova*, v. 43, n. 5, p. 620-635, 2020.

ARGUETA, C.; PARKINS, A.; PANTOURIS, G. Conformational flexibility of the C-terminal region influences distal active site residues across the tautomerase superfamily. *International Journal of Molecular Sciences*, v. 25, n. 23, p. 12617, 2024.

BERNÁ, L.; CHIRIBAO, M. L.; PITA, S.; ALVAREZ-VALIN, F.; PARODI-TALICE, A. Exploring the genomic landscape of the GP63 family in *Trypanosoma cruzi*: evolutionary dynamics and functional peculiarities. *PLoS Neglected Tropical Diseases*, v. 19, n. 3, p. e0012950, 2025.

BUNKOFSKE, M. et al. The importance of persistence and dormancy in Trypanosoma cruzi infection and Chagas disease. *Current Opinion in Microbiology*, v. 86, p. 102615, 2025.

CALIXTO, P. H. M. et al. Gene identification and comparative molecular modeling of a *Trypanosoma rangeli* major surface protease. *Journal of Molecular Modeling*, v. 19, n. 8, p. 3053-3064, 2013.

CHAUDHURI, G.; BRYANT, P.; GHOSH, S. Isolation and characterization of GP63, a major surface glycoprotein of *Leishmania donovani*. *Molecular and Biochemical Parasitology*, v. 33, n. 1, p. 85–96, 1989.

GRASSMANN, G. et al. Electrostatic complementarity at the interface drives transient protein-protein interactions. *Scientific Reports*, v. 13, n. 1, p. 10207, 2023.

GÜVENDI, M. et al. In silico identification of *Leishmania* GP63 protein epitopes to generate a new vaccine antigen against leishmaniasis. *PLoS Neglected Tropical Diseases*, v. 19, n. 6, p. e0013137, 2025.

JUMPER, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, v. 596, n. 7873, p. 583-589, 2021.

JURRUS, E. et al. Improvements to the APBS biomolecular solvation software suite. *Protein Science*, v. 27, n. 1, p. 112-128, 2018.

KIM, I.-S.; YANG, W.-S.; KIM, C.-H. Physiological properties, functions, and trends in the matrix metalloproteinase inhibitors in inflammation-mediated human diseases. *Current Medicinal Chemistry*, v. 30, n. 18, p. 2075-2112, 2023.

KULKARNI, M. M. et al. TcGP63-I, a *Trypanosoma cruzi* metalloprotease, is involved in host cell invasion. *Infection and Immunity*, v. 77, n. 11, p. 4859-4868, 2009.

LASKOWSKI, R. A. et al. PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, v. 26, n. 2, p. 283-291, 1993.

MARIANI, V. et al. lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics*, v. 29, n. 21, p. 2722-2728, 2013.

MEIER, A.; SÖDING, J. Automatic prediction of protein 3D structures by probabilistic multi-template homology modeling. *PLoS Computational Biology*, v. 11, n. 10, p. e1004343, 2015.

PÁLL, S.; ABRAHAM, M. J.; HESS, B. GROMACS: High-performance computing for molecular simulations. *Methods in Molecular Biology*, v. 2457, p. 51-73, 2022.

PIERLEONI, A.; MARTELLI, P. L.; CASADIO, R. PredGPI: a GPI-anchor predictor. *BMC Bioinformatics*, v. 9, n. 1, p. 392, 2008.

RAZZAZAN, A.; SABERI, M. R.; JAAFARI, M. R. Insights from the analysis of a predicted model of gp63 in *Leishmania donovani*. *Bioinformation*, v. 3, n. 3, p. 114, 2008.

RIGO, G. et al. Metallopeptidases as key virulence attributes of clinically relevant protozoa: new discoveries, perspectives, and frontiers of knowledge. *Current*

Protein and Peptide Science, v. 24, n. 4, p. 307-328, 2023.

SCHLAGENHAUF, E.; ETGES, R.; METCALF, P. The crystal structure of the *Leishmania major* surface proteinase leishmanolysin (GP63). *Structure*, v. 6, n. 8, p. 1035-1046, 1998.

SCHMID, N. et al. Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *European Biophysics Journal*, v. 40, n. 7, p. 843-856, 2011.

STEINKRAUS, H. B. et al. Sequence heterogeneity and polymorphic gene arrangements of the *Leishmania guyanensis* gp63 genes. *Molecular and Biochemical Parasitology*, v. 62, n. 2, p. 173-185, 1993.

TEUFEL, F. et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology*, v. 40, 2022.

WORLD HEALTH ORGANIZATION. *Chagas disease (also known as American trypanosomiasis)*. Geneva: WHO, 2 abr. 2025. Disponível em: https://www.who.int/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis
. Acesso em: 13 nov. 2025.

YAO, C.; DONELSON, J. E.; WILSON, M. E. The major surface protease (MSP or GP63) of *Leishmania* sp.: biosynthesis, regulation of expression, and function. *Molecular and Biochemical Parasitology*, v. 132, n. 1, p. 1-16, 2003.