# Exploring Binding Sites in Chagas Disease Protein TcP21 Using Integrated Mixed Solvent Molecular Dynamics Approaches

**William O. Soté (PG)**[1*], **Moacyr Comar Jr. (PQ)**[1]

[1] Institute of Chemistry, Universidade Federal de Uberlândia, Uberlândia, 38400-902, Brazil
* Corresponding email: woliveir@ufu.br

## ABSTRACT

Chagas disease remains a global health burden, particularly in Latin America. It primarily affects socioeconomically vulnerable populations, aggravating both economic inequality and marginalization. This study examines TcP21, a novel protein secreted by extracellular amastigotes, implicated in an alternative infective pathway. Three independent computational approaches—mixed solvent molecular dynamics simulations, fragment-based molecular docking, and pharmacophore model docking coupled with molecular dynamics simulations—were employed to identify potential binding sites and provide comprehensive insights into TcP21. The approaches converged on a common site located on the external surface of the protein, characterized by the key residues GLU55, ASP52, VAL70, ILE62, and TRP77. The protonated amino, acetamido, and phenyl groups of a hypothetical pharmacophore probe were consistently observed to interact with the site, suggesting an important role in ligand binding within the context.

*Keywords*: *Chagas disease*, *binding site*, *computational chemistry*, *molecular dynamics*, *pharmacophore probing*.

## Introduction

Chagas disease, or American trypanosomiasis, was discovered 115 years ago and still has no cure. It disproportionately affects socioeconomically vulnerable populations, aggravating economic inequality, contributing to marginalization, and further reducing political visibility (1). Current efforts to combat Chagas disease and other tropical diseases are aligned with the global 2030 Agenda for Sustainable Development, especially in light of the steady increase in population migration and climate changes—factors expected to significantly expand both the number of infected individuals and the range of disease vectors (2). The persistence of the parasite in the human body can be reasonably explained by the complexity of its biological cycles, which involve numerous signaling pathways. Infection mechanisms vary depending on factors such as the parasite's diverse strains, and the host cell types targeted (3). Moreover, not only the typical trypomastigote forms are infective, but the amastigote form—formerly regarded as solely replicative—also demonstrated infective capacity (4). Given this context, the TcP21 protein, secreted by extracellular amastigotes, has been identified as a ubiquitous protein with high tendency for membrane adhesion (5). It induces phagocytosis by directly interacting with transmembrane proteins and is also implicated with parasite replication and survival (6). While the potential of TcP21 for understanding Chagas disease is noteworthy, many aspects remain unexplored, particularly its three-dimensional structure and potential binding sites on its surface. Computational methods for assessing potential binding sites have become a common and successful practice in this context (7). Two commonly used approaches that parallel experimental techniques are mixed solvent molecular dynamics (MSMD) (8) and fragment-based molecular docking (9). Even though these methods provide statistical insights, they are faster, more cost-efficient, and less labor-intensive than experimental methods, enabling a more targeted use of experimental resources (10). Given the novelty of the TcP21 protein and its potential role in an unexplored alternative pathway for *T. cruzi* infection, alongside the versatility of computational approaches for mapping protein binding sites, this study used three independent approaches to better understand the TcP21 structure and identify potential surface binding sites. These include: i. MSMD, ii. fragment-based molecular docking, and iii. a combination of both, using as probe a hypothetical pharmacophore molecule. To our knowledge, this is the first comprehensive computational study of the TcP21 protein and also using these integrated methods.

## Experimental

This study comprises two main stages: 1) three-dimensional structure generation and evaluation, and then 2) binding site mapping. The general workflow used is illustrated in Figure 1. The protein structure was generated using the RoseTTAFold method, and its quality was assessed using Verify3D, ERRAT, PROCHECK, and Ramanchandran plots, and RMSD analyses. All molecular dynamics simulations were conducted in triplicates of 100 ns each. Six organic solvents were selected for the MSMD simulations: acetamide (ACM), acetonitrile (ACN), acetate ion (ACT), benzene (BNZ), ethanol (ETH), and methylammonium ion (MAM). The OPLS/CM1A forcefield was used for the molecular dynamics simulations. Molecular docking was conducted using max exhaustiveness level, a score function affinity energy threshold of 7 kcal $mol^{-1}$, and 20 unique runs with 20 binding modes generated for each run.

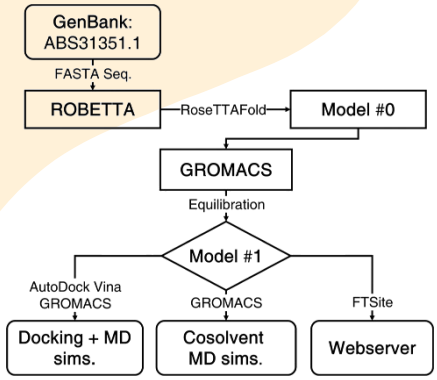


**Figure 1**. Flowchart of the experimental workflow.

## Results and Discussion

Qualitative and quantitative results are summarized in Figure 2 and Table 1. Volumetric maps in Figure 2A–B depict probe occupancy (organic solvents and pharmacophore model) during molecular dynamics sampling, while Figure 2C shows the top clusters identified via webserver-based molecular docking.

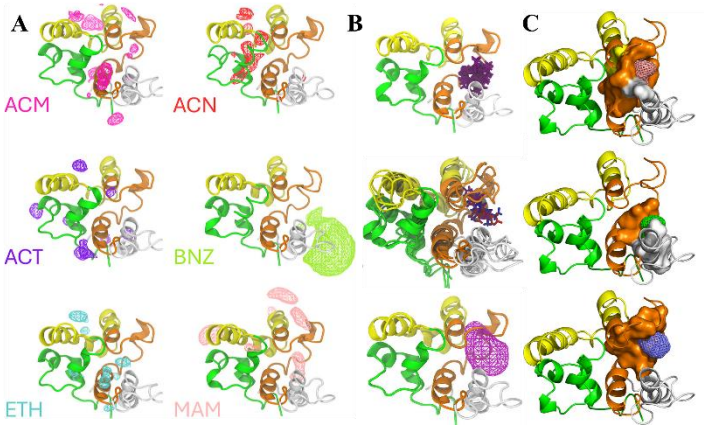


**Figure 2**. Mapped binding sites based on A. MSMD, B. docking + molecular dynamics, C. Webserver-based molecular docking.

Based on the three qualitative results, a single most probable binding site—located between the orange and white helices—was identified. In the MSMD results, this site appears for the ACM, ETH, and MAM probes. Table 1 lists the common interaction residues from the most representative conformational states of the pharmacophore probe systems and the webserver-based docking.

**Table 1**. Summary of Common Interaction Residues from TcP21 Across the Used Methodologies [a]

| MD simulations (Ph. probe) | | | Molecular Docking (webserver) | | |
|---|---|---|---|---|---|
| **R1** | **R2** | **R3** | **C1** | **C2** | **C3** |
| | | | | SER17 | |
| | | | GLU48 | | |
| | | | ARG51 | | |
| ASP52 | ASP52 | | ASP52 | ASP52 | |
| GLU55 | GLU55 | GLU55 | GLU55 | GLU55 | GLU55 |
| | | | ARG59 | ARG59 | ARG59 |
| ILE62 | | ILE62 | | | ILE62 |
| CYS66 | CYS66 | CYS66 | | | |
| VAL70 | VAL70 | VAL70 | VAL70 | | VAL70 |
| | | PHE73 | PHE73 | | |
| | | TRP77 | TRP77 | | TRP77 |

[a] Note. The letter codes Ph, R, and C represent "Pharmacophore", "Replica", and "Cluster", respectively.

Residue GLU55 (hydrophilic acidic) was consistently identified in both approaches, indicating it as the most promising hotspot. Other recurrent residues included ASP52 (hydrophilic acidic), VAL70 and ILE62 (hydrophobic aliphatic), and TRP77 (hydrophobic aromatic). Figure 3 presents a qualitative binding site model that incorporates the most frequent residues and those within 5 Å of an arbitrary pharmacophore binding mode for perspective. The model represents an external surface site with minimal steric hindrance, consistent with all approaches. Interactions between the protein and the protonated amino, acetamido, and phenyl groups of the pharmacophore model—via salt bridges, hydrogen bonds, charge–charge, and alkyl–π contacts—suggest a significant role in ligand binding.

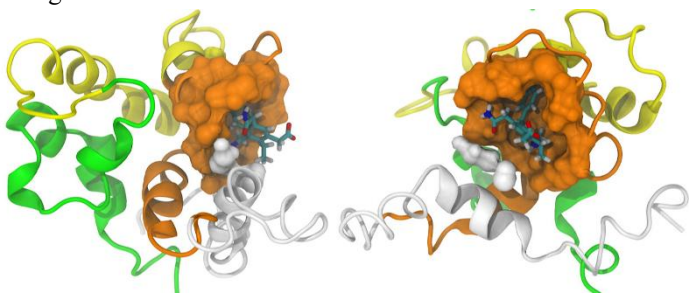

**Figure 3**. Qualitative binding site model for the pharmocophore probe. Form left to right: front and side views. Binding site is shown as a surface representations.

Without an experimentally determined 3D structure or known binding sites for TcP21, full validation of the proposed site and its potential allosteric modulation (activation or inhibition) remains uncertain. Still, the strong correlation across the mapping approaches supports this study's relevance for advancing research on TcP21 in Chagas disease.

## Conclusions

Chagas disease remains a longstanding burden that profoundly impacts socioeconomic conditions of thousands of individuals. Despite ongoing research efforts, an effective treatment has yet to be identified. Computational approaches for preliminary binding site assessment offer a versatile and robust strategy for protein investigation, thereby improving the efficiency of experimental resource use. This study employed three distinct computational approaches to explore potential binding sites in the novel TcP21 protein, within the context of Chagas disease. All approaches identified a common binding site featuring a hotspot formed by five residues (GLU55, ASP52, VAL70, ILE62, and TRP77) located on an accesible external pocket. These findings though preliminary, provide a new perspective on TcP21. The study does not aim to propose therapeutic candidates but to highlight a previously uncharacterized protein implicated in *T. cruzi* cell invasion.

## Acknowledgments

The authors acknowledge FAPEMIG (Proc. 5.10/2022) for the financial support and Prof. Dr. Luiz Guilherme Machado de Macedo (UFSJ) for the computational resources.

## References

1. de Oliveira Junior, W. A. et al., *Mem. Inst. Oswaldo Cruz* **2022**, *117*, e220066;
2. WHO *Ending the Neglect to Attain the Sustainable Development Goals: A Road Map for Neglected Tropical Diseases 2021–2030*; World Health Organization: Geneva, Swizterland, 2020;
3. Romano, P. S. et al., *IUBMB Life* **2012**, *64* (5), 387– 396;
4. Martín-Escolano, J. et al., *ACS Infect. Dis.* **2022**, *8* (6), 1107–1115;
5. da Silva, C. V. et al., *Microbes Infect.* **2009**, *11* (5), 563– 570;
6. Teixeira, S. C. et al., *Sci. Rep.* **2017**, *7* (1), 44978;
7. Feher, V. A. et al., *Curr. Opin. Struct. Biol.* **2014**, *25*, 98– 103;
8. Ilie, I. M. et al., *J. Chem. Inf. Model.* **2023**, *63* (12), 3878– 3891;
9. Grant, B. J. et al., *PLoS One* **2011**, *6* (10), e25711;
10. Ghanakota, P.; Carlson, H. A. *J. Med. Chem.* **2016**, *59* (23), 10383– 10399.