Mendeleev Communications

Mendeleev Commun., 2025, **35**, 133–135

Structural modeling of L-galactonolactone dehydrogenase from *Trypanosoma cruzi*

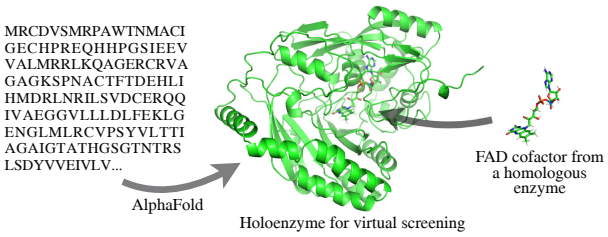
Vsevolod O. Shegolev,[a] Alexander V. Veselovsky[b] and Elena V. Kudryashova*[a]

[a] *Department of Chemistry, M. V. Lomonosov Moscow State University, 119991 Moscow, Russian Federation. E-mail: helenakoudriachova@yandex.ru*

[b] *V. N. Orekhovich Institute of Biomedical Chemistry, 119121 Moscow, Russian Federation*

DOI: 10.71267/mencom.7590

A structural model of the *Trypanosoma cruzi* L-galactonolactone dehydrogenase enzyme, a potential drug target for Chagas disease, is presented. Constructed using a hybrid approach, the model was validated *via* molecular dynamics, ensemble docking and experimental data (circular dichroism and IC_{50} values) with good agreement. The model holds promise for virtual screening of new inhibitors.



Keywords: *Trypanosoma cruzi*, Chagas disease, flavin adenine dinucleotide, AlphaFold, molecular dynamics, docking.

Trypanosoma cruzi is a protozoan parasite that causes Chagas disease, which affects millions of people in Latin America. Up to seven million people are infected, and 75 million are at risk. Unlike humans, the parasite cannot uptake vitamin C and must synthesize it.[1] The enzyme L-galactonolactone dehydrogenase from *T. cruzi* (TcGAL) is involved in the final step of vitamin C biosynthesis, making it a potential drug target.

TcGAL is a member of the aldonolactone oxidoreductase family,[2] a group of flavoenzymes that catalyze the final step in vitamin C biosynthesis across plants, fungi and animals. These enzymes oxidize sugar lactones to produce L-ascorbic acid, differing in their oxygen reactivity: plant GalDH enzymes act as dehydrogenases using cytochrome c as an electron acceptor, while fungal ALO and animal GULO enzymes are oxidases that directly utilize oxygen, producing hydrogen peroxide as a byproduct. Phylogenetic analysis indicates a common evolutionary origin for these enzymes,[3] with key mutations driving their functional divergence across different lineages.

Other therapeutic targets[4] under discussion include enzymes from the ergosterol biosynthesis pathway (sterol 14α-demethylase, squalene synthase, squalene monooxygenase, lanosterol synthase, sterol 24-C-methyltransferase and hydroxymethylglutaryl-CoA reductase), the antioxidant defense system (trypanothione reductase and iron superoxide dismutase) and virulence-associated factors (cruzipain and trans-sialidase). Despite their promise, these targets remain largely in the research and preclinical stages, requiring significant further development before clinical application can be realized.

TcGAL consists of 505 residues and includes a flavin adenine dinucleotide (FAD) cofactor. The membranotropic properties of TcGAL complicate folding into an active conformation *in vitro*.[5] However, TcGAL can be stabilized and studied in reverse micelles.[6,7] Experimental determination of the protein structure is challenging and it remains unknown. In this study, a structural model of TcGAL was designed using a hybrid methodology and supported by experimental data.

AlphaFold2[8] predicted the 3D structure of the apoenzyme (average pLDDT score 87.8) using the gene sequence (UniProt ID: Q4DPZ5). The unstructured fragment Ala277–Gly239 (average pLDDT score 43.0) was kept as is. The FAD cofactor structure taken from the homologous enzyme (PDB ID: 5OEP) was inserted into the apoenzyme cavity. Visual inspection confirmed a good fit.

The system was prepared for molecular dynamics (MD) simulations using AmberTools22. The ff14SB and GAFF force fields were used for the protein and FAD, respectively. The phosphate groups of the FAD cofactor were assigned as negatively charged and deprotonated. The atomic charges for FAD were estimated using the AM1-BCC method. The AMBER topology was converted to GROMACS format using ParmEd v3.4.3.

Constrained energy minimization to eliminate steric clashes between the apoenzyme and the cofactor was performed in GROMACS 2022.4. The system was solvated in a truncated octahedral box using the TIP3P water model. Na^+ and Cl^- ions were added to neutralize the charge, reaching 0.15 M. The system was subjected to minimization, heating to 300 K and equilibration in the NVT and NPT ensembles. This was followed by a 300 ns production MD simulation.

MD simulations showed acceptable protein stability (RMSD < 3.5 Å). A flexible loop formed by the Ala277–Gly239 sequence was detected (average RMSF of 2.56 Å). This loop

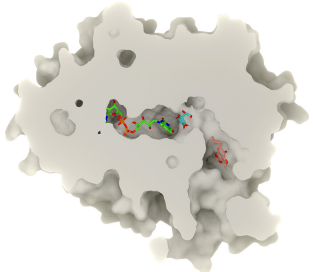


Figure 1 Molecular surface of TcGAL.

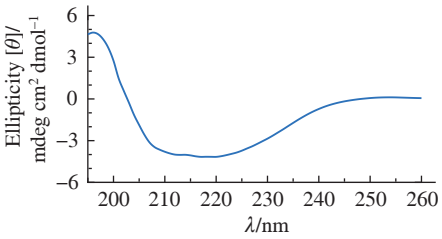


Figure 2 CD spectrum of TcGAL in reverse micelles. Conditions: 0.05 M enzyme, 0.04 M AOT (docusate sodium) lipid in isooctane, 10 mM sodium phosphate buffer (pH 7.2), 20 °C, degree of hydration [H_2O] : [AOT] = 30 : 1.

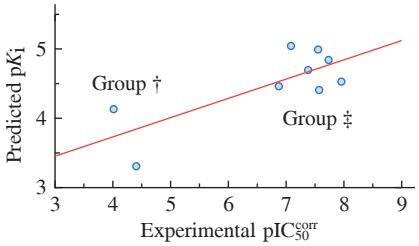


Figure 3 Relationship between corrected experimental pIC_{50} values and docking-predicted pK_i values at a standard concentration C of 1 M.

also had low AlphaFold2 pLDDT scores and is likely involved in regulation and activity by covering the entry to the active site. The protein surfaces were analyzed with PyMOL, revealing a tunnel to the active site (Figure 1).

Clustering of the trajectories in GROMACS using the single linkage method yielded six representative protein conformations. The secondary structure was assigned using DSSP and compared with circular dichroism (CD) spectroscopy data (Figure 2). The model showed 33% α-helices and 18% β-sheets, which are comparable to 35.5 and 17.5%, respectively, from the CD data† within experimental error.

One representative structure from each cluster was prepared using MGLTools v1.5.7 for blind ensemble docking with AutoDock Vina v1.2.3 engine.[9] The ligands included twenty-three known compounds, related to inhibitors, substrates, products, and analogs,[7] and four acid anions. The obtained scores ΔG_d were converted to dissociation constants $K_i = \exp(\Delta G_d/RT)$.

The experimental IC_{50}^{exp} values measured in the multiphase assay[7] were corrected to account for the partitioning of the inhibitor between the organic and aqueous phases:

$$pIC_{50}^{corr} = -\log IC_{50}^{corr} + \log(1 + P), \quad (1)$$

where P was calculated using Chemprop[10] v2.0.3 trained on experimental values of $\log P$.[11] Equation (1) was obtained by solving

$$\log P = \log \frac{[I]_o}{[I]_w}, \quad (2)$$
$$[I]_w + [I]_o = C_I,$$

where C_I, $[I]_w$ and $[I]_o$ are the inhibitor concentrations: total, in the aqueous and in the organic phases, respectively.

The corrected experimental pIC_{50}^{corr} values were compared with the docking pK_i constants. Inhibitors with a maximum % inhibition less than 50% were excluded, as were charged triphenylphosphonium-containing inhibitors, since reliable $\log P$ values cannot be predicted and the AutoDock Vina scoring function cannot account for arbitrary charges. The Pearson correlation for uncharged molecules was 0.77 (Figure 3), which is considered adequate,[12] indicating the usefulness of the model for virtual screening. However, the correlation should be interpreted with caution, since it may appear stronger due to the uneven distribution of values forming two distinct groups.

An analysis of the docked poses showed that substrates, products and sugar-like analogs [Figure 4(*a*)] bind inside the active site; small inhibitors like allylbenzene and apiole can potentially bind both inside and outside the active site [Figure 4(*b*)]; larger lycorine [see Figure 4(*b*)] and bifunctional inhibitors with a linker [Figure 4(*c*)] bind outside the active site.

AlphaFold employs MSA (multiple sequence alignment) and deep learning to predict protein 3D structures, providing valuable insights for solving crystal and Cryo-EM structures.[13]

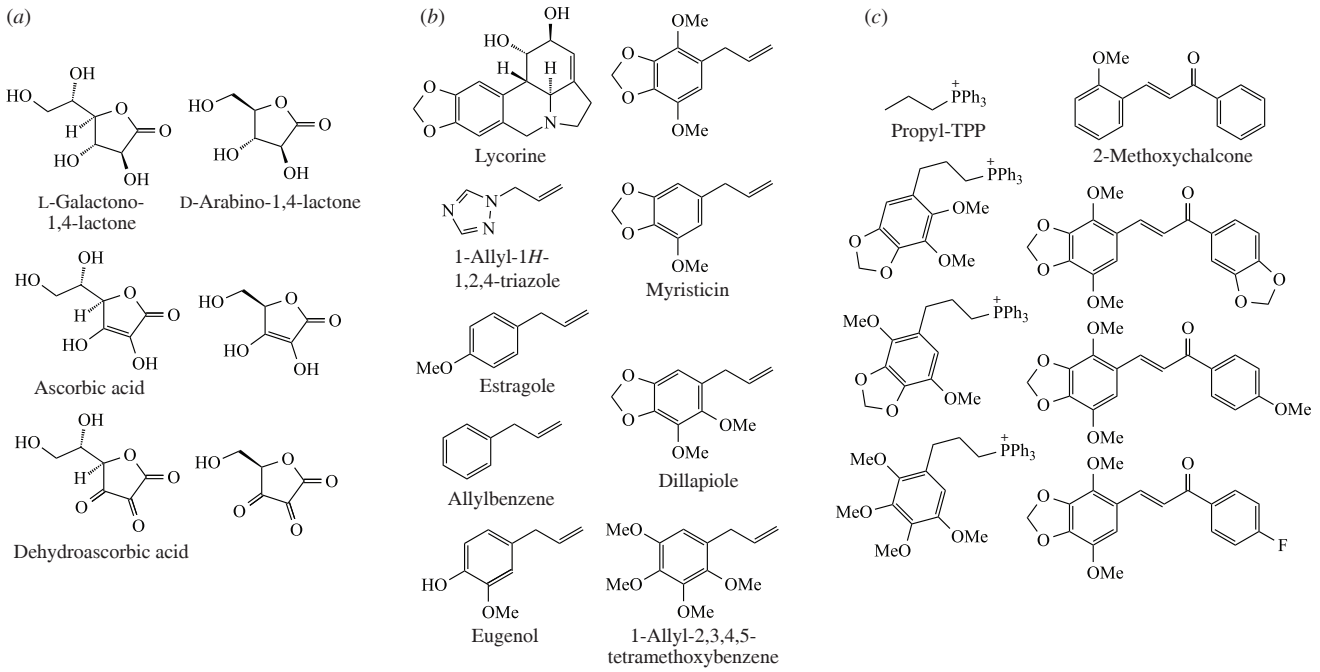


Figure 4 Docked ligands: (*a*) substrates, products and analogs (all ionized forms were considered); (*b*) known inhibitors used in the correlation analysis; (*c*) known inhibitors with a maximum % inhibition less than 50% or with uncompensated total charge.

† CD spectrum was measured using a Jasco J-815 circular dichroism spectrometer (Japan). Secondary structure analysis was performed by deconvolution using CDNN v2.1 software (Applied Photophysics Ltd., UK). The results revealed 35.6% α-helices, 17.5% β-sheets (both antiparallel and parallel), 16.1% β-turns and 32.0% disordered structures. The total deconvolution error is 1.2%.

However, its performance for drug targets is mixed due to limitations in accounting for protein flexibility and uncertainty in predictions for non-crystallized proteins. Additionally, publication bias should not be overlooked.

This research highlights the potential of TcGAL as a drug target in *T. cruzi* vitamin C biosynthesis. By combining AlphaFold2 and classical homology modeling, MD simulations and ensemble docking, we obtained a model representing the structure and binding sites of TcGAL. This model is promising for the virtual screening of new inhibitors and advances therapeutic research in Chagas disease.

This work was supported by the Non-Commercial Foundation for the Support of Science and Education 'INTELLECT'. The measurements were performed using the equipment (Jasco J-815 CD spectrometer) of the Moscow State University development program. The research was conducted using the equipment of the shared research facilities of the HPC computing resources at Lomonosov Moscow State University.

References

1 F. J. Logan, M. C. Taylor, S. R. Wilkinson, H. Kaur and J. M. Kelly, *Biochem. J.*, 2007, **407**, 419; https://doi.org/10.1042/bj20070766.
2 N. G. H. Leferink and W. J. H. van Berkel, in *Flavins and Flavoproteins: Methods and Protocols*, eds. S. Weber and E. Schleicher, Humana Press, New York, 2014, pp. 95–111; https://doi.org/10.1007/978-1-4939-0452-5_6.
3 A. Boverio, N. Jamil, B. Mannucci, M. L. Mascotti, M. W. Fraaije and A. Mattevi, *Nat. Commun.*, 2024, **15**, 4158; https://doi.org/10.1038/s41467-024-48410-1.
4 P. García-Huertas and N. Cardona-Castro, *Biomed. Pharmacother.*, 2021, **142**, 112020; https://doi.org/10.1016/j.biopha.2021.112020.
5 E. V. Kudryashova, N. G. H. Leferink, I. G. M. Slot and W. J. H. van Berkel, *Biochim. Biophys. Acta, Proteins Proteomics*, 2011, **1814**, 545; https://doi.org/10.1016/j.bbapap.2011.03.001.
6 A. A. Chudin and E. V. Kudryashova, *Analytica*, 2022, **3**, 36; https://doi.org/10.3390/analytica3010004.
7 A. A. Chudin, I. D. Zlotnikov, S. S. Krylov, V. V. Semenov and E. V. Kudryashova, *Biochemistry (Moscow)*, 2023, **88**, 131; https://doi.org/10.1134/s000629792301011x.
8 J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli and D. Hassabis, *Nature*, 2021, **596**, 583; https://doi.org/10.1038/s41586-021-03819-2.
9 J. Eberhardt, D. Santos-Martins, A. F. Tillack and S. Forli, *J. Chem. Inf. Model.*, 2021, **61**, 3891; https://doi.org/10.1021/acs.jcim.1c00203.
10 E. Heid, K. P. Greenman, Y. Chung, S.-C. Li, D. E. Graff, F. H. Vermeire, H. Wu, W. H. Green and C. J. McGill, *J. Chem. Inf. Model.*, 2024, **64**, 9; https://doi.org/10.1021/acs.jcim.3c01250.
11 M. Popova, O. Isayev and A. Tropsha, *Sci. Adv.*, 2018, **4**, eaap7885; https://doi.org/10.1126/sciadv.aap7885.
12 J. M. Carpio Arévalo and J. C. Amorim, *Sci. Rep.*, 2022, **12**, 4742; https://doi.org/10.1038/s41598-022-08359-x.
13 R. Nussinov, M. Zhang, Y. Liu and H. Jang, *Drug Discovery Today*, 2023, **28**, 103551; https://doi.org/10.1016/j.drudis.2023.103551.

Received: 12th August 2024; Com. 24/7590