

# Crithidia fasciculata gene of unknown function CFAC1\_070014400 encodes a putative alpha/beta hydrolase with a conserved active site

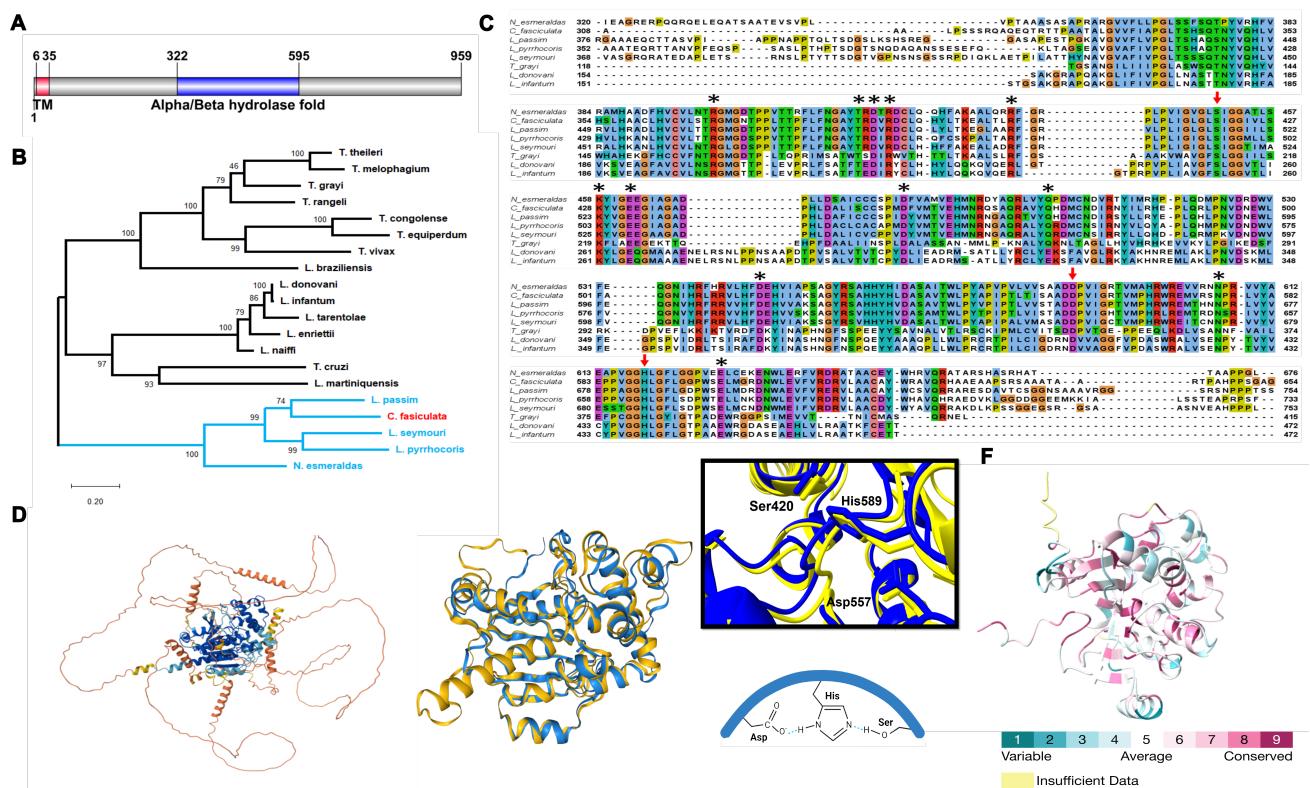
Justin Davis

Department of Biology, North Carolina Wesleyan University, Rocky Mount, North Carolina, United States

To whom correspondence should be addressed: JDavis@NCWU.edu

## Abstract

A gene of unknown function, CFAC1\_070014400 (Transcript ID: CFAC1\_070014400.1), identified in *Crithidia fasciculata* was analyzed using bioinformatic methods to study the sequence and structure of its encoded protein. Results from domain predictions, conservation analysis, and structural comparisons indicate the encoded protein is a member of the alpha/beta hydrolase superfamily and contains a conserved active site. Based on phylogenetic analysis, the encoded protein groups with other putative alpha/beta hydrolase homologs from related kinetoplastid parasites.



**Figure 1. Sequence and Structural Characterization of CFAC1\_070014400:**

A) Domain structure of CFAC1\_070014400 indicating the location of predicted sequence features. Two primary domains include a putative alpha/beta hydrolase fold and a predicted transmembrane domain. Created using Illustrator for Biological Sequences (Liu et al., 2015) based on predictions from InterPro (Blum et al., 2020). (B) Maximum likelihood phylogenetic tree calculated using the MEGA12 evolutionary analysis software using default settings (Kumar et al., 2024). Support values are shown from 500 bootstrap repetitions. (C) Multi-Sequence Alignment of the alpha/beta hydrolase fold domain showing sequence conservation across diverse homologs from related kinetoplastid parasites. The multi-sequence alignment was created using Clustal Omega then exported to Jalview for visualization (Sievers and Higgins 2017; Waterhouse et al., 2009). Red arrows indicate the residues of the predicted catalytic triad (Ser102, His271 and Asp239), which are highly conserved and predicted to be functional. ConSurf-identified highly-conserved functional residues in CFAC1\_070014400 indicated with asterisks (Ashkenazy et al., 2016). (D) AlphaFold3 (Abramson et al., 2024) predicted structural model of the full length CFAC1\_070014400 protein. Dark blue represents a very high model confidence (pLDDT >90), light blue represents medium confidence (90 >pLDDT >70), yellow represents low confidence (70 >pLDDT >50), and orange represents very low confidence (pLDDT <50). (E) Structural overlay of the CFAC1\_070014400 alpha/beta hydrolase domain (Blue) with alpha/beta hydrolase-1 domain-containing protein (yellow) from *Leishmania infantum* (*L. infantum*). Structure alignment was done using the Foldseek server and the AFDB-proteome database (Kempen et al., 2024). The side box shows the predicted active site containing the conserved catalytic

triad. The bottom portion shows a model representation of the catalytic triad. (F) AlphaFold3 predicted the structural model of the CFAC1\_070014400 alpha/beta hydrolase domain with coloring based on ConSurf conservation scores.

## Description

Several members of the family *Trypanosomatidae* are of importance to global health. The family *Trypanosomatidae* is composed entirely of protozoan parasites, which includes clinically relevant human parasites like *Trypanosoma brucei* (*T. brucei*) and *Leishmania* spp. *T. brucei* causes African sleeping sickness and members of the *Leishmania* genus causes Leishmaniasis in humans. The genomes of these parasites have been sequenced and provided novel insights into parasite biology (Berriman et al., 2005; Ivens et al., 2005). Understudied members of this family include parasites like *Crithidia fasciculata* (*C. fasciculata*) that complete their monoxenous life cycles in a single mosquito host. *C. fasciculata*'s genome has been sequenced and has been used as a model organism to study the related human pathogens, *Trypanosoma brucei* and *Leishmania* (Filosa et al., 2019). A vast majority of *C. fasciculata* genes are annotated as hypothetical proteins with unknown function. Here, we predict the encoded function of an unknown gene in *C. fasciculata* to advance our understanding of parasite genetics.

Based on our analysis, the uncharacterized *C. fasciculata* gene CFAC1\_070014400 (Shanmugasundram et al., 2023) is a member of the alpha/beta hydrolase superfamily, subfamily 4 (IPR050960). Members of this family carry out diverse sets of enzymatic functions associated with lipid metabolism (Lenfant et al., 2013). alpha/beta hydrolase domain-containing proteins are understudied in kinetoplastid parasites, but phospholipases have been recognized for their importance in pathogenesis in related parasites *T. brucei* and *Leishmania* spp. (Bordon et al., 2018; Monic et al., 2022). Based on our analysis, CFAC1\_070014400 also contains a single transmembrane helix, which suggests that it has a possible membrane localization. Other alpha/beta hydrolase domain-containing proteins, found in mammalian cells, are anchored to membranes by a single transmembrane helix and play key roles in lipid metabolism and signalling (Lord et al., 2013).

The InterPro web server (Blum et al., 2025) identified the 959-amino acid CFAC1\_070014400 protein as a member of the alpha/beta hydrolase superfamily, subfamily 4 (IPR050960). The sequence was found to contain an alpha/beta hydrolase fold (IPR029058) and a single transmembrane helix (**Figure 1A**).

The softwares Phobius (Käl et al., 2004) and TMHMM (Hallgren et al., 2022) both identified a single transmembrane helix at the N-terminus of CFAC1\_070014400 (6-35) (**Figure 1A**). A putative signal peptide was also identified using both TargetP (Armenteros et al., 2019) and DeepLoc (Thumuluri et al., 2022). DeepLoc predicted the membrane association of CFAC1\_070014400 to be peripheral (0.7540). The softwares BUSCA (Savojardo et al., 2018) and DeepLoc both predicted the CFAC1\_070014400 protein to be membrane associated, but no confident prediction of subcellular localization could be determined. BUSCA predicted CFAC1\_070014400 to localize to the plasma membrane, while DeepLoc predicted the localization to be the lysosome/vacuole.

We used the evolutionary analysis software MEGA12 (Kumar et al., 2024) to construct a maximum likelihood phylogenetic tree with the CFAC1\_070014400 protein and putative homologs identified by an NCBI BlastP search in related kinetoplastid parasites (**Figure 1B**). CFAC1\_070014400 appears to be most closely related with putative homologs from other monoxenous trypanosomatid parasites.

Using an NCBI BlastP search, homologs of the CFAC1\_070014400 protein were identified across a diverse set of kinetoplastid parasites. These included monoxenous trypanosomatids, as well as *Leishmania* spp. and *Trypanosoma*. Homologs were also identified in diverse species belonging to Plantae and Fungi. No homologs were identified in Animalia, Archaeobacteria, or Bacteria. We used ClustalOmega and Jalview (Sievers and Higgins 2017; Waterhouse et al., 2009) to construct a multiple sequence alignment (**Figure 1C**) in order to visualize sequence conservation across representative homologs. There appears to be significant sequence conservation within the alpha/beta hydrolase domain, which indicates these residues are likely important across multiple species. Additionally, many of these residues were identified by ConSurf (Ashkenazy et al., 2016) to be highly conserved across multiple species as well as functional and are marked with asterisks. These included residues of the putative catalytic triad (Ser102, His271 and Asp239). The full ConSurf results for CFAC1\_070014400 are available as extended data.

We used AlphaFold3 (Abramson et al., 2024) to predict a tertiary structure of the full length CFAC1\_070014400 protein (**Figure 1D**). The program predicted confident folding in the putative Alpha/Beta hydrolase domain (residues 322- 595). The core structural region consists of beta-sheets flanked by alpha helices. This structural core is conserved and is similar to what is observed in other alpha/beta hydrolases across diverse eukaryotes (Ozhelvaci and Steczkiewicz 2025). When the putative alpha/beta hydrolase domain of CFAC1\_070014400 was overlaid with its nearest structural homolog from *L. infantum*, we observed many features that were overlaid with each other including the putative catalytic triad (**Figure 1E**). The putative catalytic triad we identified in CFAC1\_070014400 consisted of the residues, Ser102, His271 and Asp239. This canonical catalytic triad is highly conserved in alpha/beta hydrolase family proteins and generally consists of a histidine, an acid (aspartic acid or glutamate) and a nucleophilic residue (Ozhelvaci and Steczkiewicz 2025). In mammalian alpha/beta hydrolase proteins, a serine residue in the active site generally serves as the nucleophile (Lord et al., 2013). In our identified active site, a histidine is present, and aspartic acid likely acts as the acidic residue while serine

would likely act as the nucleophile. Mutagenesis studies have previously shown that mutating even one of the amino acid residues in the catalytic triad, results in a complete loss of enzyme function (Navia-Paldanius et al., 2012). Our analysis also showed that residues within this putative alpha/beta hydrolase domain are highly conserved including our putative catalytic triad (**Figure 1F**). This highlights the key role these residues play in enzymatic activity and the likely reason they remain highly conserved across a variety of eukaryotes.

Based on sequence and structural features, we predict CFAC1\_070014400 is part of the alpha/beta hydrolase superfamily. Additionally, based on previously published studies of alpha/beta hydrolase domain-containing proteins as well as structural analysis of our protein, CFAC1\_070014400 might be involved in lipid metabolism and signaling. Lipid metabolism plays key roles in *C. fasciculata* parasites including, membrane homeostasis, metabolism and environmental stress resistance. Next steps for functional studies of this protein could involve expressing and purifying the protein in a heterologous system to perform esterase/lipase screens to confirm alpha/beta hydrolase activity and identify putative substrates. Site-directed mutagenesis could then be performed to mutate the predicted residues in the catalytic triad to confirm involvement in enzymatic activity. Taken together, these experiments would confirm the bioinformatics predictions presented here.

**Acknowledgements:** The author thanks Dr. Emily Weber and Dr. Daniel Elias for the critical reading and evaluation of this manuscript.

### Extended Data

Description: Full ConSurf results for the CFAC1\_070014400 protein. . Resource Type: Text. File: [consurf\\_colored\\_seq\\_Cf AB hydrolase\(1\).pdf](#). DOI: [10.22002/np0r4-pwt21](#)

### References

- Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al., Jumper JM. 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 630(8016): 493-500. PubMed ID: [38718835](#)
- Almagro Armenteros JJ, Salvatore M, Emanuelsson O, Winther O, von Heijne G, Elofsson A, Nielsen H. 2019. Detecting sequence signals in targeting peptides using deep learning. *Life Sci Alliance* 2(5): 10.26508/lsa.201900429. PubMed ID: [31570514](#)
- Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N. 2016. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* 44(W1): W344-50. PubMed ID: [27166375](#)
- Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renauld H, Bartholomeu DC, et al., El-Sayed NM. 2005. The genome of the African trypanosome *Trypanosoma brucei*. *Science* 309(5733): 416-22. PubMed ID: [16020726](#)
- Blum M, Andreeva A, Florentino LC, Chuguransky SR, Grego T, Hobbs E, et al., Bateman A. 2025. InterPro: the protein sequence classification resource in 2025. *Nucleic Acids Res* 53(D1): D444-D456. PubMed ID: [39565202](#)
- Bordon MLAC, Laurenti MD, Ribeiro SP, Toyama MH, Toyama DO, Passero LFD. 2018. Effect of phospholipase A(2) inhibitors during infection caused by *Leishmania (Leishmania) amazonensis*. *J Venom Anim Toxins Incl Trop Dis* 24: 21. PubMed ID: [30181736](#)
- Filosa JN, Berry CT, Ruthel G, Beverley SM, Warren WC, Tomlinson C, et al., Povelones M. 2019. Dramatic changes in gene expression in different forms of *Crithidia fasciculata* reveal potential mechanisms for insect-specific adhesion in kinetoplastid parasites. *PLoS Negl Trop Dis* 13(7): e0007570. PubMed ID: [31356610](#)
- Hallgren J, Tsirigos KD, Pedersen MD, Almagro Armenteros JJ, Marcatili P, Nielsen H, et al., Winther O. 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *bioRxiv*. <https://doi.org/10.1101/2022.04.08.487609>
- Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, et al., Myler PJ. 2005. The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309(5733): 436-42. PubMed ID: [16020728](#)
- Käll L, Krogh A, Sonnhammer EL. 2004. A combined transmembrane topology and signal peptide prediction method. *J Mol Biol* 338(5): 1027-36. PubMed ID: [15111065](#)
- Kumar S, Stecher G, Suleski M, Sanderford M, Sharma S, Tamura K. 2024. MEGA12: Molecular Evolutionary Genetic Analysis Version 12 for Adaptive and Green Computing. *Mol Biol Evol* 41(12): 10.1093/molbev/msae263. PubMed ID: [39708372](#)
- Lenfant N, Hotelier T, Bourne Y, Marchot P, Chatonnet A. 2013. Proteins with an alpha/beta hydrolase fold: Relationships between subfamilies in an ever-growing superfamily. *Chem Biol Interact* 203(1): 266-8. PubMed ID: [23010363](#)

Lord CC, Thomas G, Brown JM. 2013. Mammalian alpha beta hydrolase domain (ABHD) proteins: Lipid metabolizing enzymes at the interface of cell signaling and energy metabolism. *Biochim Biophys Acta* 1831(4): 792-802. PubMed ID: [23328280](#)

Monic SG, Lamy A, Thonnus M, Bizarra-Rebelo T, Bringaud F, Smith TK, Figueiredo LM, Rivière L. 2022. A novel lipase with dual localisation in *Trypanosoma brucei*. *Sci Rep* 12(1): 4766. PubMed ID: [35306507](#)

Navia-Paldanius D, Savinainen JR, Laitinen JT. 2012. Biochemical and pharmacological characterization of human  $\alpha/\beta$ -hydrolase domain containing 6 (ABHD6) and 12 (ABHD12). *J Lipid Res* 53(11): 2413-24. PubMed ID: [22969151](#)

Ozhelvaci F, Steczkiewicz K. 2025.  $\alpha/\beta$  Hydrolases: Toward Unraveling Entangled Classification. *Proteins* 93(4): 855-870. PubMed ID: [39623291](#)

Savojardo C, Martelli PL, Fariselli P, Profiti G, Casadio R. 2018. BUSCA: an integrative web server to predict subcellular localization of proteins. *Nucleic Acids Res* 46(W1): W459-W466. PubMed ID: [29718411](#)

Shanmugasundram A, Starns D, Böhme U, Amos B, Wilkinson PA, Harb OS, et al., Jones. 2023. TriTrypDB: An integrated functional genomics resource for kinetoplastida. *PLOS Neglected Tropical Diseases* 17: e0011058. DOI: [doi.org/10.1371/journal.pntd.0011058](https://doi.org/10.1371/journal.pntd.0011058)

Sievers F, Higgins DG. 2018. Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci* 27(1): 135-145. PubMed ID: [28884485](#)

Thumuluri V, Almagro Armenteros JJ, Johansen AR, Nielsen H, Winther O. 2022. DeepLoc 2.0: multi-label subcellular localization prediction using protein language models. *Nucleic Acids Res* 50(W1): W228-W234. PubMed ID: [35489069](#)

van Kempen M, Kim SS, Tumescheit C, Mirdita M, Lee J, Gilchrist CLM, Söding J, Steinegger M. 2024. Fast and accurate protein structure search with Foldseek. *Nat Biotechnol* 42(2): 243-246. PubMed ID: [37156916](#)

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25(9): 1189-91. PubMed ID: [19151095](#)

**Funding:** This research was funded in part by internal grants for professional development awarded by North Carolina Wesleyan University.

**Conflicts of Interest:** The authors declare that there are no conflicts of interest present.

**Author Contributions:** Justin Davis: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, resources, software, visualization, writing - original draft, writing - review editing.

**Reviewed By:** Anonymous

**History:** Received November 30, 2025 **Revision Received** February 23, 2026 **Accepted** April 1, 2026 **Published Online** April 3, 2026 **Indexed** April 17, 2026

**Copyright:** © 2026 by the authors. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Citation:** Davis J. 2026. *Crithidia fasciculata* gene of unknown function CFAC1\_070014400 encodes a putative alpha/beta hydrolase with a conserved active site. *microPublication Biology*. [10.17912/micropub.biology.001967](https://doi.org/10.17912/micropub.biology.001967)