The Predicted Structure of a Thermophilic Malate Synthase

Shaelee Nielsen, Jantzen Orton, & Bruce R. Howard*

Department of Chemistry and Physics, Southern Utah University, Cedar City, UT

Student: shaeleenielsen@gmail.com, jantzen.orton@gmail.com
Mentor: howard@suu.edu*

ABSTRACT

This project aims to solve the structure of the crenarchaeal *Sulfolobus acidocaldarius* enzyme malate synthase. Other known malate synthase enzymes have been found to require a magnesium ion in the active site to carry out catalytic activities, but a study reported that *S. acidocaldarius* malate synthase does not require magnesium. This suggests a novel mechanism for this enzyme. Additionally, the mature *S. acidocaldarius* protein is approximately 100 residues larger than any other structurally characterized malate synthase. It has also been reported to form a dimer, while previously solved structures have only displayed monomeric, trimeric, and hexameric arrangements. We plan to determine the structure experimentally. However, major advances in the accuracy of protein structure prediction were made recently by AlphaFold, an artificial intelligence system developed by DeepMind, which has revolutionized the field and has largely solved the protein folding problem. A similar AI system, RoseTTAFold, developed by David Baker's lab at the University of Washington, has been made publicly available. Here, we report our analysis of the structure of this protein, predicted using both of these algorithms and of a predicted structural model for the dimeric form of the enzyme using ClusPro. Our results strongly support a conserved catalytic mechanism requiring magnesium, which is common with all previously solved malate synthase isoforms.

KEYWORDS

Glyoxylate Cycle; Malate synthase; Protein Prediction; Thermophile; *Sulfolobus acidocaldarius;* Magnesium; AlphaFold; RoseTTAFold

INTRODUCTION

The glyoxylate cycle is a metabolic pathway that relies on several citric acid cycle enzymes along with two additional enzymes, isocitrate lyase and malate synthase, to synthesize citric acid cycle intermediates from two-carbon units.[1] Evidence of this pathway was first discovered in Bacteria (*E. coli*),[2] but the cycle's enzymes have since been identified in Eukarya and Archaea, as well.[3-5] Operation of this cycle is essential for many microorganisms to grow on a media where acetate is the sole carbon source.[6] Furthermore, the glyoxylate cycle contributes to survival in the pathogenic organisms *Candida albicans* and *Mycobacterium tuberculosis*, once they are engulfed by a host macrophage.[7, 8]

In organisms employing the glyoxylate cycle, an isocitrate molecule is cleaved to form succinate and the two-carbon compound glyoxylate by isocitrate lyase. In step two, malate synthase catalyzes the Claisen condensation of glyoxylate and an incoming acetyl-CoA to form a malyl-CoA intermediate that is then hydrolyzed to malate and CoA. In this way, the two decarboxylation steps of the citric acid cycle are bypassed, which generates additional precursors for amino acid and carbohydrate biosynthesis.

Previous efforts have identified four isoforms of malate synthase but only three isoforms have been structurally characterized.[9] Originally, two of these were distinguished by whether growth was observed on acetate (MSA) or glycolate (MSG) in *E. Coli*, which contains both.[10] All solved structures in bacteria are one of these two isoforms: MSA or MSG. Although both isoforms, A and G, exist as monomers in currently determined structures, their size and amino acid sequence vary significantly. Members of isoform A

consist of ~530 amino acid residues, whereas those from isoform G consist of ~730 residues; the structurally conserved regions of the sequences retain an 18% identity.[11] MSA is predicted to form oligomers in some cases.[12, 13]

The two remaining isoforms have only been identified in Archaea at present. The first reported example of an archaeal malate synthase to be purified came from _Haloferax volcanii_,[5] a halophile that grows in the bottom sediment of the Dead Sea at high salt concentrations.[14] Interestingly, this protein does not closely resemble MSA or MSG sequences and is only 433 amino acid residues.[15] Due to these differences and the discovery of other halophilic malate synthases through genome sequencing,[16] this class of enzyme was denoted as MSH.[17] Data from gel-filtration chromatography and crystal structures revealed that isoform H exists in a trimer/hexamer equilibrium, and in comparison to MSA and MSG, exhibits a deletion at the N-terminus and a truncated C-terminal domain. Despite these variations, the geometry of the active site and catalytic mechanism were conserved.[9]

The central protein fold that contains the active site in the aforementioned isoforms is a $(\beta/\alpha)_8$ (TIM) barrel. A crystal structure of the glyoxylate complex in _E. coli_ MSG revealed that Asp 631 donated from a C-terminal domain, and Arg 338 within the TIM barrel are both conserved to provide acid-base chemistry; Asp 631 deprotonates the methyl group of acetyl-CoA, which subsequently forms an enolate that is stabilized by the positive charge carried by Arg 338. Two other critical residues, Glu 427 and Asp 455, coordinate an essential Mg^{2+} ion into place, thereby polarizing the glyoxylate substrate for nucleophilic attack. The positive charge carried by Mg^{2+} (cooperating with Arg 338) also helps stabilize the oxyanion that is formed from the attack of the enolate intermediate.[18] **(Figure 1)**
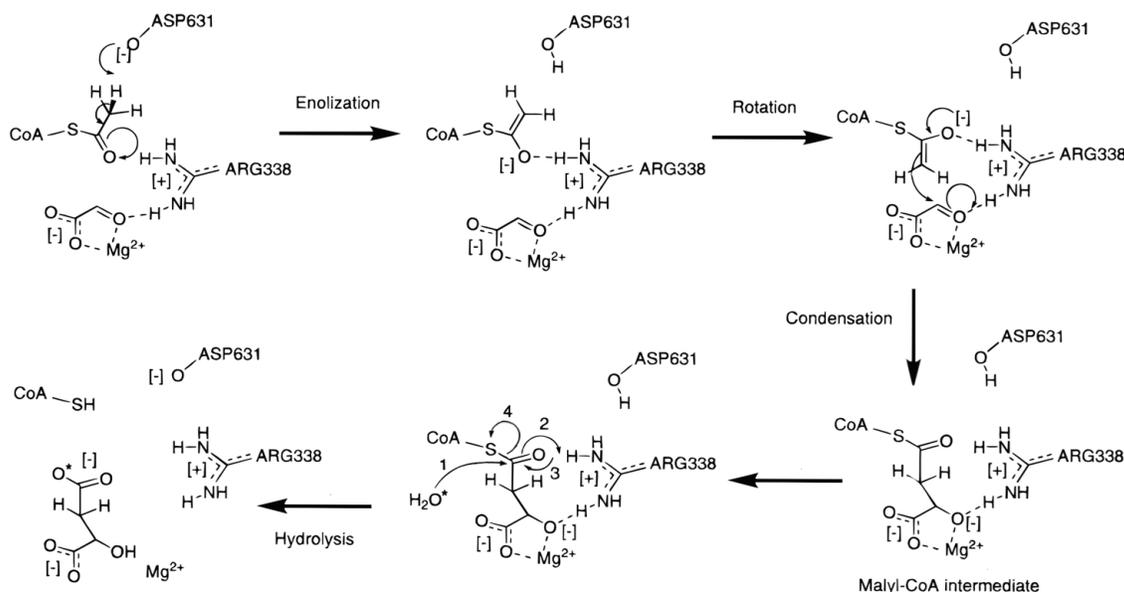

Figure 1. A currently proposed enzymatic mechanism for malate synthase with _E. coli_ MSG numbering.[18]

The fourth isoform of malate synthase, found in crenarchaeal species like _Sulfolobus acidocaldarius_, has yet to be structurally solved. This organism is a thermoacidophile that grows optimally at 75-80 °C and a pH of 2-3, typically found in acid thermal soils or acid hot springs.[19] Genome sequencing revealed that the malate synthase protein is 824 residues in length,[20] about 100 residues longer than MSG. In contrast to the other structures, it was reported to function without Mg^{2+} and to form a functional homodimer.[21] The significant difference in cofactor dependence suggests that the enzyme may employ a novel catalytic mechanism.

In order to investigate this unique type of malate synthase, its catalytic requirements, and its evolutionary relationship compared to previously characterized isoforms, we have undertaken the present study. Simultaneously, our study led us to explore the recent advancements in protein structure prediction through the application of artificial intelligence technology. AlphaFold, developed by DeepMind, has revolutionized structural biology by demonstrating accurate predictions of protein structures comparable in quality to experimentally-determined structures.[22] RoseTTAFold, is a similar system that was developed by David Baker's lab at the University of Washington.[23] Both systems were released in 2021. Furthermore, the ClusPro web server uses computational methods

to sample billions of potential protein-protein conformations.[24-27] These results are further refined through energy minimization calculations that produce the most likely models for a given interaction. These systems were implemented into the study design in order to explore their potential application to this enzyme.

METHODS AND PROCEDURES

Analysis of the Predicted Structures

The protein sequence was retrieved from the UniProt database (Identifier # Q4J6V5) and submitted to the RoseTTAFold public server found at https://robetta.bakerlab.org/ for structural predictions. The structure output by the server was downloaded as a PDB file and was viewed using PyMOL.[31] Analysis of the structure allowed us to identify a TIM barrel that appeared homologous to other barrels that contain the active site in previously determined structures of malate synthase.

In order to conduct a more detailed comparison, representative structures for each one of the known malate synthase isoforms were downloaded from the Protein Data Bank (PDB):[32] MSA from *Bacillus anthracis* (3CUX), MSG from *Escherichia coli* (1D8C), and MSH from *Haloferax volcanii* (3OYZ). The following four conserved catalytic residues were identified inside each TIM barrel: an arginine to act a bidentate cation which bridges the two substrates during the course of the reaction, an aspartate that acts as the catalytic base, and another aspartate cooperating with a glutamate residue to chelate the required cofactor. **(Figure 1)**

To compare the spatial positioning of the TIM barrel surrounding these key residues, protein alignment and overlays were performed. The MSG isoform from *Escherichia coli* (1D8C) served as the stationary molecule to which all overlays were made. We isolated the 3D coordinates of the C-alphas from each of the four key catalytic residues in each isoform and created PDB files containing solely those coordinates.

Using the PyMOL alignment tool, each of these sets of coordinates were aligned onto the fixed coordinates for these four C-alpha atoms of 1D8C. The LSQ algorithm in the COOT (Crystallographic Object-Oriented Toolkit) software was used to overlay each protein onto its corresponding four-coordinate aligned file.[28,29] These overlaid structures were saved and viewed simultaneously in PyMOL, allowing for direct structural comparison of the active sites.

For each isoform and the RoseTTAFold model, the distances were measured between the C-alphas of these residues as (A) aspartate base to arginine, (B) Mg^{2+}-chelating aspartate to arginine, (C) Mg^{2+}-chelating aspartate to glutamate, and (D) Mg^{2+}-chelating aspartate to the aspartate base **(Figure 3, Table 1)**.

The AlphaFold predicted structure was retrieved through the UniProt database (Identifier # A0A0U3GUU1). As with the other structures, the coordinates of the four catalytic residue C-alphas in the AlphaFold structure were aligned on those in 1D8C and the respective distances were measured. Additionally, overlays of the AlphaFold and RoseTTAFold structural models were compared and analyzed using PyMOL.

Prediction of a Dimeric Form

The ClusPro protein-protein docking server at https://cluspro.bu.edu/ was accessed using the "Dimer Classification" feature to predict potential dimeric forms of the RoseTTAFold model.[30] The server output 26 potential structures according to the "balanced" ranking option. PyMOL was used to view and draw conclusions from these dimeric models.

RESULTS AND DISCUSSION

Conserved Mechanism

The overlay of MSG, MSA, MSH, and the RoseTTAFold model indicates that the $(\beta/\alpha)_8$ (TIM) barrel domain is conserved across the various isoforms **(Figure 2)**.

At an atomic level, the distances between catalytic residues are also conserved. The distances between the key residues in the active site are very similar across each representative isoform, including both of the predicted structures for *S. acidocaldarius* **(Figure 3,**

Table 1). This further suggests that the functions of the catalytic arginine, base, and both Mg^{2+}–coordinating residues are maintained in the fourth isoform of malate synthase from *S. acidocaldarius*. Hence, we propose the following functions in *S. acidocaldarius:* Arg-214 acts as a bidentate cation, Asp-600 acts as the catalytic base, and both Asp-326 and Glu-298 chelate the suggested cofactor.
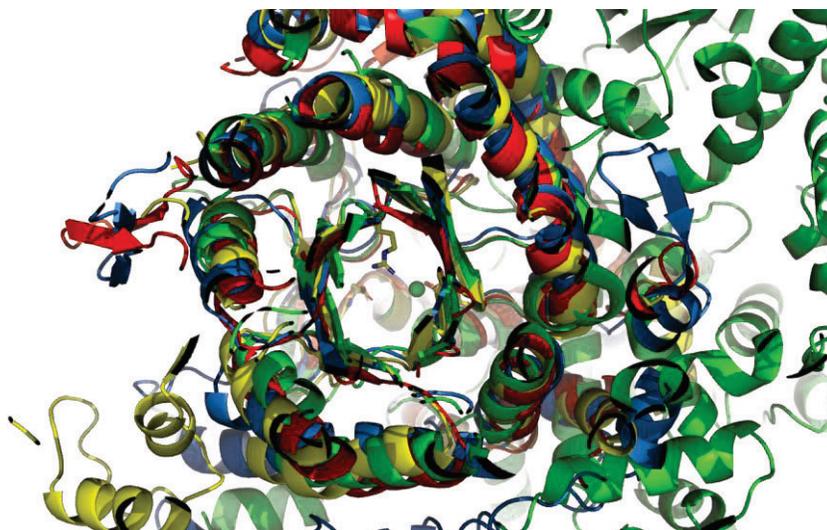

Figure 2. Overlay of MSA (red), MSH (green), and the RoseTTAFold model (blue) onto *E. coli* MSG (yellow).[31]
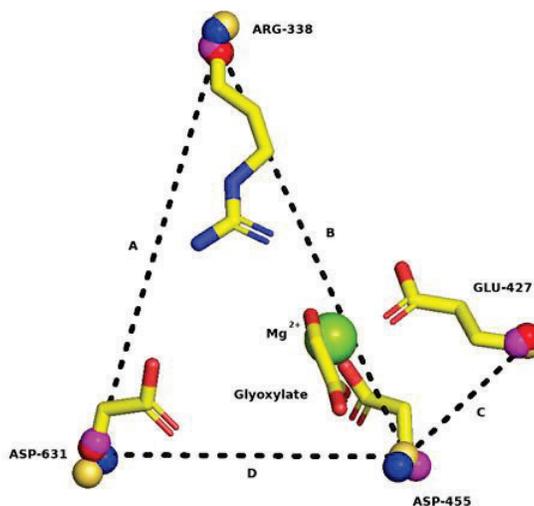

Figure 3. Relative position of Cα-atoms of key catalytic residues and relative distance measurements shown for the RoseTTAFold model (blue spheres). Also shown are the Cα-atoms for MSA (red spheres), MSH (yellow spheres), and the AlphaFold model (purple spheres) aligned onto those of MSG.[31]

Distance (Å)	MSG	MSA	MSH	RoseTTAFold	AlphaFold
A	13.9	13.4	15.1	14.3	13.3
B	14.7	14.5	15.1	15.2	15.0
C	5.5	5.8	5.6	5.8	5.4
D	10.1	10.1	10.2	9.6	10.4

Table 1. Relative distances between Cα-atoms of key catalytic residues shown in Angstroms (Å) for each isoform and the predicted model.[31]

The conserved $(\beta/\alpha)_8$ (TIM) barrel domain and the conserved positions of key catalytic residues within the barrel provide strong evidence for a conserved catalytic mechanism in the malate synthase of *S. acidocaldarius*. A proposed mechanism for malate synthase in previously solved structures of MSG, MSA, and MSH requires magnesium in the active site (**Figure 1**). The conserved TIM-barrel fold and positioning of these key catalytic residues include the magnesium-coordinating groups. This calls into question the report that this isoform did not require magnesium for activity.[21] It is important to remember that the structures used in the analysis are only predictions, so results cannot be confirmed until the structure of the protein is determined experimentally.
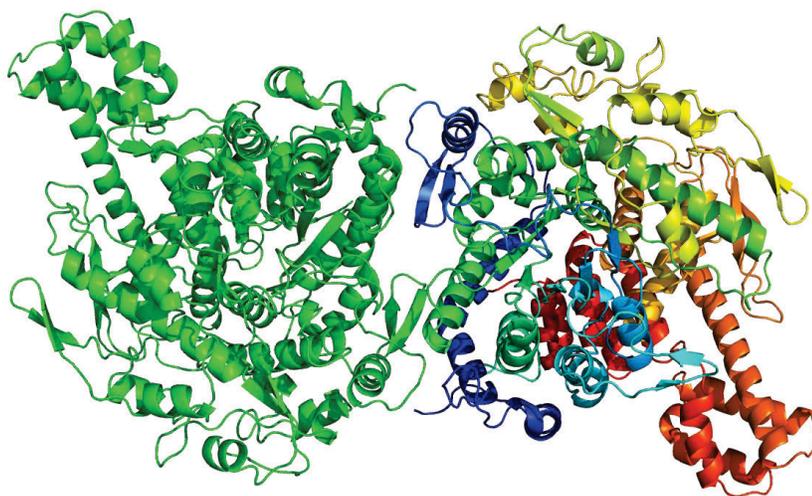
Predicted Dimer

The "Dimer Classification" feature from the ClusPro server using the balanced energy function produced the top 26 dimer predicted arrangements using the RoseTTAFold model.[30] Using PyMOL, all of these dimer predictions were viewed simultaneously by superimposing one subunit of each dimer **(Figure 4)**. This allowed us to compare the relative positions of the predicted interfaces. Each of the top 26 predicted dimer interfaces are found within one of 3 main locations on the RoseTTAFold model. The green-yellow-tan cluster contains 21 predictions, the blue cluster contains 3 predictions, and the red cluster contains 2 predictions.
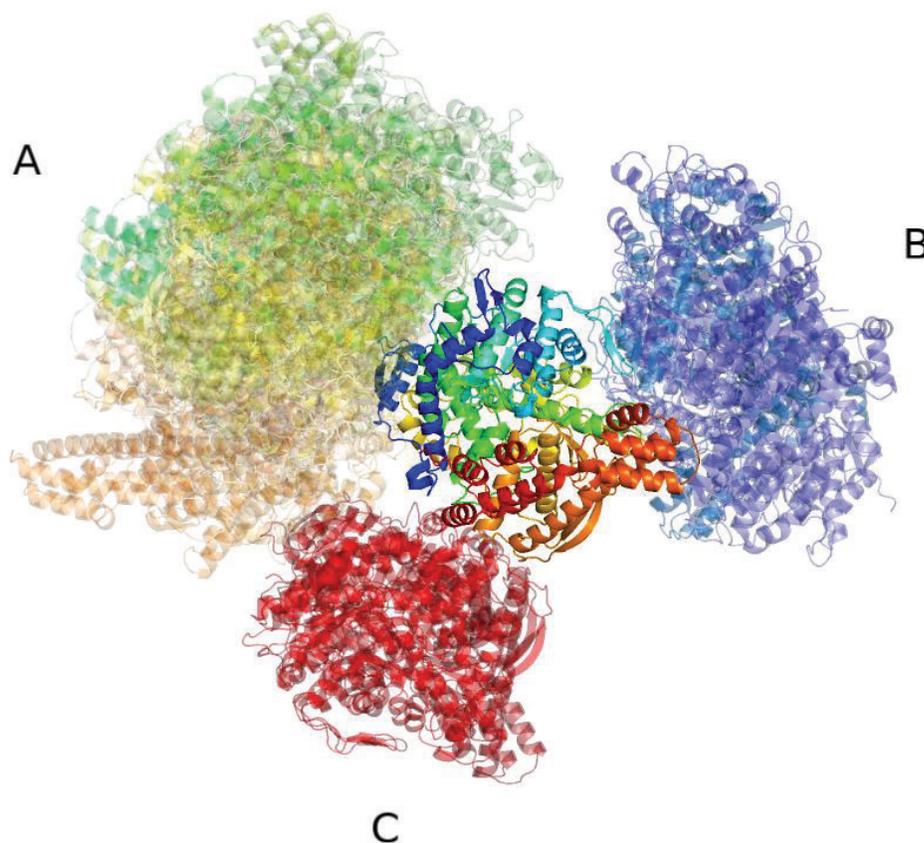
Figure 4. Three surfaces predicted as potential dimerization interfaces. Top 26 predicted dimer arrangements for the RoseTTAFold model using ClusPro are shown, with one subunit of each dimer prediction superimposed (centrally located and colored blue to red from the N- to the C-terminus). The green-yellow-tan cluster of structure shown at the top left comprises most of the predicted structures, and also includes the lowest-energy dimer prediction shown in figure 5. However, two other locations were identified as potential dimerization interfaces and are shown as a blue cluster on the right, or as a red cluster to the lower left of the superimposed subunit.[31]

The lowest-energy predicted dimer is located within the green-yellow-tan cluster of structures shown in Figure 4. The location and orientation of this dimer suggest that the extended N-terminus of *S. acidocaldarius* malate synthase is involved in forming the dimerization interface **(Figure 5)**. Interestingly, this interface also corresponds to the location of intersubunit interactions observed in the only experimentally determined structure of an oligomeric malate synthase reported. This is the malate synthase from *Haloferax*

volcanii, determined by X-ray crystallography, which was shown to form active trimers and hexamers using gel-filtration chromatography.[9] The hexameric form is composed of two back-to-back trimers as observed in the crystal structure. When one of the subunits in the *H. volcanii* hexamer is overlain onto one of the subunits in the predicted dimer shown in Figure 5, the locations of the respective subunit interfaces can be compared directly. The interface in this lowest-energy predicted dimer structure of *S. acidocaldarius* corresponds to a position on the *H. volcanii* subunit that interacts with two different subunits in the opposing trimer within the hexamer. Conversely, the locations of the red and the blue clusters in Figure 4 don't correspond to any of the subunit interfaces within the *H. volcanii* hexameric arrangement. These observations further support the possibility that the predicted dimer shown in Figure 5 may represent the true dimer interface. It is important to note that the presence and location of this dimer arrangement is only a prediction and cannot be confirmed without experimentally solving the protein structure.


Figure 5. Top predicted RoseTTAFold dimer arrangement, as predicted by ClusPro, viewed along an approximate two-fold rotation axis of symmetry, with one subunit colored blue to red from the N- to the C-terminus. N-terminal segments (shown in dark blue in the monomer on the right) come together to largely form the dimerization interface.[30, 31]

In addition to the RoseTTAFold server, we also utilized the AlphaFold structure prediction. Using PyMOL, we performed an overlay of the structures predicted by each algorithm **(Figure 6)**. When comparing the structures, we noticed differences in the spatial orientation of some secondary structures. While the structures are not identical, they are very similar, including the conserved structure of the TIM barrel forming the active site. Most of the protein is structured in generally the same way in both predictions. Most of these differences between the predictions can be explained by slight shifts and rotational movements of the various portions of the proteins. However, one helical subdomain, shown in the lower right side of **Figure 6**, is rotated differently relative to the main TIM barrel domain.

In conclusion, we found strong evidence for a conserved mechanism in the malate synthase of *S. acidocaldarius*, suggesting that magnesium is required in the active site. The ClusPro server prediction suggests that the extended N-terminus of *S. acidocaldarius* malate synthase is involved in forming a dimerization interface. In addition, we found slight differences between the predicted models from RoseTTAFold and AlphaFold. These findings are based on predicted structures, not confirmed protein structures, and results cannot be confirmed without an experimentally solved protein structure. But based on our detailed comparisons, we anticipate that Arg-214 acts as a bidentate cation to bridge the two substrates and stabilize the oxyanions formed during the catalytic cycle, Asp-600 acts as the catalytic base, and both Asp-326 and Glu-298 chelate the suggested Mg^{2+} cofactor.

The differences found between the RoseTTAFold and AlphaFold predictions underscore the need for an experimentally determined protein structure. This will allow verification of the predicted structure and mechanism involved in the malate synthase of *S.*

acidocaldarius. An experimentally solved structure will confirm the existence of a dimer and the details of the interface. Additionally, this would provide an opportunity to determine which algorithm (RoseTTAFold or AlphaFold) produced the more accurate prediction.
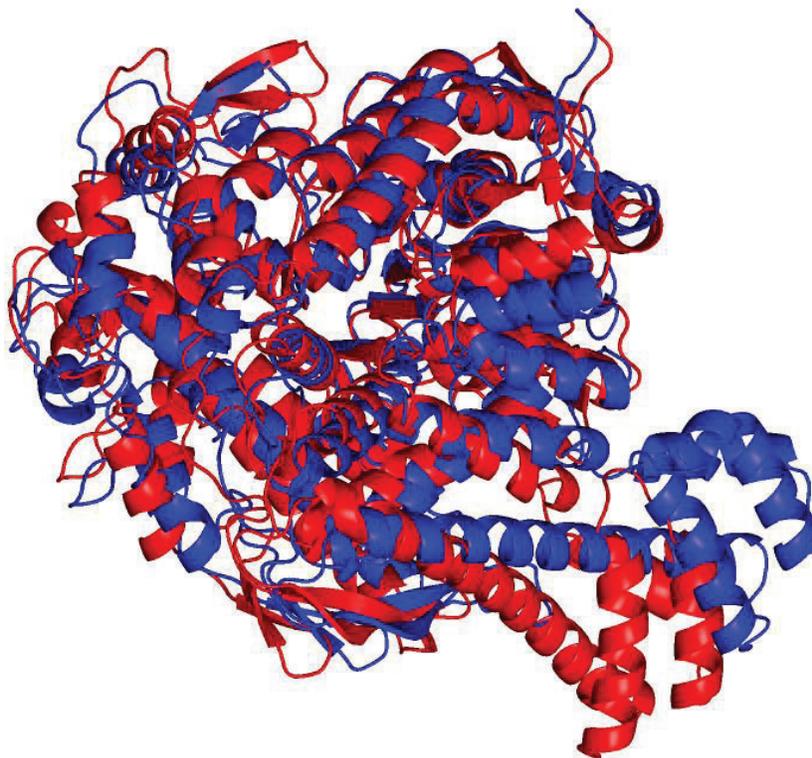
Figure 6. Overlay of the RoseTTAFold monomer model (blue) and the AlphaFold monomer model (red). The overlay was performed by superimposing the C-alpha atoms of four catalytic residues in each monomer as described in the methods section and shown in Figure 3. Both monomers are shown as backbone cartoon traces to clarify the structural variations between these two predicted structures.[31]

REFERENCES
1. Kornberg, H. L., and Krebs, H. A. (1957) Synthesis of cell constituents from C2-units by a modified tricarboxylic acid cycle. *Nature* 179, 988-991. *https://doi.org/10.1038/179988a0*
2. Ajl, S. J. (1956) Conversion of acetate and glyoxylate to malate. *J Am Chem Soc* 78(13), 3230-3231. *https://doi.org/10.1021/ja01594a079*
3. Kornberg, H. L., and Beevers, H. (1957) A mechanism of conversion of fat to carbohydrate in castor beans. *Nature* 180(4575), 35–36. *https://doi.org/10.1038/180035a0*
4. Nakazawa, M., Minami, T., Teramura, K., Kumamoto, S., Hanato, S., Takenaka, S., Ueda, M., Inui, H., Nakano, Y., and Miyatake, K. (2005) Molecular characterization of a bifunctional glyoxylate cycle enzyme, malate synthase/isocitrate lyase, in Euglena gracilis. *Comp Biochem Physiol B Biochem Mol Biol* 141, 445-452. *https://doi.org/10.1016/j.cbpc.2005.05.006*
5. Serrano, J. A., Camacho, M., and Bonete, M. J. (1998) Operation of glyoxylate cycle in halophilic archaea: presence of malate synthase and isocitrate lyase in Haloferax volcanii. *FEBS Lett* 434(1-2), 13-16. *https://doi.org/10.1016/s0014-5793(98)00911-9*

6. Kornberg, H. L. (1966) The role and control of the glyoxylate cycle in Escherichia coli. *Biochem J* 99(1), 1-11. *https://doi.org/10.1042/bj0990001*

7. McKinney, J. D., Höner zu Bentrup, K., Muñoz-Elías, E. J., Miczak, A., Chen, B., Chan, W. T., Swenson, D., Sacchettini, J. C., Jacobs, W. R., Jr., and Russell, D. G. (2000) Persistence of Mycobacterium tuberculosis in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase. *Nature* 406(6797), 735–738. *https://doi.org/10.1038/35021074*

8. Lorenz, M. C., and Fink, G. R. (2001) The glyoxylate cycle is required for fungal virulence. *Nature* 412, 83-86. *https://doi.org/10.1038/35083594*

9. Bracken, C. D., Neighbor, A. M., Lamlenn, K. K., Thomas, G. C., Schubert, H. L., Whitby, F. G., and Howard, B. R. (2011) Crystal structures of a halophilic archaeal malate synthase from Haloferax volcanii and comparisons with isoforms A and G. *BMC Struct Biol* 11(23), 1-19. *https://doi.org/10.1186/1472-6807-11-23*

10. Vanderwinkel, E., and De Vlieghere, M. (1968) Physiology and genetics of isocitritase and the malate synthases of Escherichia coli. *Eur J Biochem* 5(1), 81-90. *https://doi.org/10.1111/j.1432-1033.1968.tb00340*

11. Lohman, J. R., Olson, A. C., and Remington, S. J. (2008) Atomic resolution structures of Escherichia coli and Bacillus anthracis malate synthase A: Comparison with isoform G and implications for structure-based drug discovery. *Protein Sci* 17(11), 1935–1945. *https://doi.org/10.1110/ps.036269.108*

12. Durchschlag, H., Biedermann, G., and Eggerer, H. (1981) Large-scale purification and some properties of malate synthase from bakers-yeast. *Eur J Biochem* 114, 255–262.

13. Khan, A.S., Van Driessche, E., Kanarek, L., and Beeckmans, S. (1993) Purification of the glyoxylate cycle enzyme malate synthase from maize (Zea mays L.) and characterization of a proteolytic fragment. *Protein Expr Purif* 4(6), 519-28. *https://doi.org/10.1006/prep.1993.1068*

14. Mullakhanbhai, M.F., and Larsen, H. (1975) Halobacterium volcanii spec. nov., a Dead Sea halobacterium with a moderate salt requirement. *Arch Microbiol* 104, 207–214. *https://doi.org/10.1007/BF00447326*

15. Serrano, J. A., and Bonete, M. J. (2001) Sequencing, phylogenetic and transcriptional analysis of the glyoxylate bypass operon (ace) in the halophilic archaeon Haloferax volcanii. *Biochim Biophys Acta* 1520(2), 154–162. *https://doi.org/10.1016/s0167-4781(01)00263-9*

16. Baliga, N. S., Bonneau, R., Facciotti, M. T., Pan, M., Glusman, G., Deutsch, E. W., Shannon, P., Chiu, Y., Weng, R. S., Gan, R. R., Hung, P., Date, S. V., Marcotte, E., Hood, L., and Ng, W. V. (2004) Genome sequence of Haloarcula marismortui: a halophilic archaeon from the Dead Sea. *Genome Res* 14(11), 2221–2234. *https://doi.org/10.1101/gr.2700304*

17. Thomas, G., Lamlenn, K., and Howard, B. R. (2009) Crystallization and preliminary x-ray diffraction of a halophilic archaeal malate synthase. *AJUR* 8(2 & 3), 15–23.

18. Howard, B. R., Endrizzi, J. A., and Remington, S. J. (2000) Crystal structure of Escherichia coli malate synthase G complexed with magnesium and glyoxylate at 2.0 A resolution: mechanistic implications. *Biochemistry* 39(11), 3156–3168. *https://doi.org/10.1021/bi992519h*

19. Brock, T. D., Brock, K. M., Belly R. T., and Weiss, R. L. (1972) Sulfolobus: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch Microbiol* 84, 54-68. *https://doi.org/10.1007/BF00408082*

20. Chen, L., Brügger, K., Skovgaard, M., Redder, P., She, Q., Torarinsson, E., Greve, B., Awayez, M., Zibat, A., Klenk, H. P., and Garrett, R. A. (2005) The genome of Sulfolobus acidocaldarius, a model organism of the Crenarchaeota, *J Bacteriol* 187(14), 4992–4999. *https://doi.org/10.1128/JB.187.14.4992-4999.2005*

21. Uhrigshardt, H., Walden, M., John, H., Petersen, A., and Anemüller, S. (2002) Evidence for an operative glyoxylate cycle in the thermoacidophilic crenarchaeon Sulfolobus acidocaldarius. *FEBS Lett* 513(2-3), 223–229. *https://doi.org/10.1016/s0014-5793(02)02317-7*

22. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., and Hassabis, D. (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583-589. *https://doi.org/10.1038/s41586-021-03819-2*

23. Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., Wang, J., Cong, Q., Kinch L. N., Schaeffer, R. D., Millán, C., Park, H., Adams, C., Glassman, C. R., DeGiovanni, A., Pereira, J. H., Rodrigues A. V., van Dijk, A. A., Ebrecht, A. C., Opperman, D. J., Sagmeister, T., Buhlheller, C., Pavkov-Keller, T., Rathinaswamy, M. K., Dalwadi, U., Yip, C. K., Burke, J. E., Garcia, K. C., Grishin, N. V., Adams, P. D., Read, R. J., and Baker, D. (2021) Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 373:6557, 871-876. *https:doi.org/10.1126/science.abj8754*

24. Desta I. T., Porter K. A., Xia B., Kozakov D., and Vajda S. (2020) Performance and its limits in rigid body protein-protein docking. *Structure* 28(9), 1071-1081. *https://doi.org/10.1016/j.str.2020.06.006*

25. Vajda S., Yueh C., Beglov D., Bohnuud T., Mottarella S. E., Xia B., Hall D. R., and Kozakov D. (2017) New additions to the ClusPro server motivated by CAPRI. *Proteins* 85(3), 435-444. *https://doi.org/10.1002/prot.25219*

26. Kozakov D., Hall D. R., Xia B, Porter K. A., Padhorny D., Yueh C., Beglov D., and Vajda S. (2017) The ClusPro web server for protein-protein docking. *Nat Protoc* 12(2), 255-278. *https://doi.org/10.1038/nprot.2016.169*

27. Kozakov D., Beglov D., Bohnuud T., Mottarella S., Xia B, Hall D. R., Vajda, S. (2013) How good is automated protein docking? *Proteins* 81(12), 2159-66. *https://doi.org/10.1002/prot.24403*

28. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Cryst* 66(4), 486-501. *https://doi.org/10.1107/S0907444910007493*

29. Emsley P., and Cowtan K. (2004) Coot: Model-building tools for molecular graphics. *Acta Cryst* 60(Pt 12 Pt 1), 2126–2132.

30. Yueh C., Hall D. R., Xia B., Padhorny D., Kozakov D., and Vajda S. (2017) ClusPro-DC: Dimer classification by the Cluspro server for protein–protein docking. *J Mol Biol* 429(3), 372-381. *https://doi.org/10.1016/j.jmb.2016.10.019*

31. Schrödinger, L., and DeLano, W. (2020) *PyMOL*. Retrieved from *http://www.pymol.org/pymol*

32. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N, Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000) The RCSB Protein Data Bank, www.rcsb.org. Nucleic Acids Research 28, 235–242.

ABOUT THE STUDENT AUTHORS

Jantzen Orton graduated from Southern Utah University in April 2022, earning a Bachelor of Science in Biology and a minor in Chemistry. He is attending medical school at the Texas Tech University Health Sciences Center.

Shaelee Nielsen is currently attending Southern Utah University. She plans to graduate with a Bachelor of Science in Biology in Spring 2024 and then complete medical school.

PRESS SUMMARY

Our project focuses on how a malate synthase enzyme works in a microorganism called *Sulfolobus acidocaldarius*. Most malate synthase enzymes need magnesium to function properly, but one study reported that the malate synthase in this microorganism does not. This is interesting because it suggests a new way the enzyme could work. Recently, new methods have been developed that use artificial intelligence to predict protein structure. We used these new methods to predict the structure of the malate synthase enzyme and found the structure to be very similar to the structures of the other malate synthases. This suggests that the malate synthase in *Sulfolobus acidocaldarius* has a function similar to the other malate synthases, suggesting that magnesium is required for proper functioning.