Research in Molecular Medicine

Functional Annotation of Some Hypothetical Genes in the *Schistosoma* Parasite Based on Reciprocal Best Structural-hit Relationship

Arezou Askari Rad[1], Jamal Fayazi[1*], Mohammad Taghi Beigi Nassiri[1], Alireza Hasani Baferani[2]

1. Department of Animal Sciences, Faculty of Animal and Food Science, Agricultural Sciences and Natural Resources University of Khuzestan, Ahvaz, Iran.
2. Department of Animal Sciences, Agricultural Extension & Education Institute, Tehran, Iran.

Use your device to scan and read the article online

Citation Askari Rad A, Fayazi J, Beigi Nassiri MT. Functional Annotation of Some Hypothetical Genes in the *Schistosoma* Parasite Based on Reciprocal Best Structural-hit Relationship. Research in Molecular Medicine. 2022; 10(4):225-234. https://doi.org/10.32598/rmm.10.4.1076.4

https://doi.org/10.32598/rmm.10.4.1076.4

Article Type:
Research Paper

Article info:
Received: 7 Sep 2022
Revised: 28 Sep 2022
Accepted: 15 Oct 2022

Keywords:
Schistosoma, Structure, Functional annotation, AlphaFold

ABSTRACT

Background: The *Schistosoma* parasite is responsible for several overlooked tropical diseases, which cause significant economic losses in livestock. This parasite is increasingly found in the central areas of the northern provinces of Iran. To generate an effective treatment, it is crucial to understand thoroughly how the parasite's genes work. Currently, the roles of numerous genes in these parasites are unknown, so their identification and targeting of them are challenging. Conventional techniques for assigning functions to proteins depend on the similarity of their sequences. Yet, this method does not always recognize similarities between distantly related proteins. Research has shown that taking the protein's structure in the process of predicting its function can be helpful in pinpointing proteins whose functions are not known yet.

Materials and Methods: In our study, we utilized two advanced technologies, AlphaFold and Foldseek, to deduce the functions of theoretical proteins in the *Schistosoma* parasite. We accomplished this by contrasting the structure of *Schistosoma* proteins with those of *Caenorhabditis elegans*, a closely related model organism, using Foldseek to identify reciprocal best matches. Our research involved an in-depth examination of two specific predictions, evaluating evidence for functional resemblances, such as patterns of protein interactions and similarities in functional domains.

Results: Our results indicate that one of the analyzed genes is likely involved in embryogenesis, while the other might be connected to the egg-laying process of the *Schistosoma* parasite.

Conclusion: Function of some hypothetical proteins can be inferred bases on their structural similarities to annotated proteins, especially proteins with a low sequence similarity to annotated proteins.

*** Corresponding Author:**
Jamal Fayazi, PhD.
Address: *Department of Animal Sciences, Faculty of Animal and Food Science, Agricultural Sciences and Natural Resources University of Khuzestan, Ahvaz, Iran.*
Phone: *+98 (916) 6124162*
E-mail: *j_fayazi@asnrukh.ac.ir*

Introduction

Schistosoma genus comprises a broad parasitic worm with a significant impact on human health, livestock, and agricultural production, causing economic instability in various places, especially in developing countries. The 2021 World Health Organization (WHO) estimates show that at least 251 million people are at risk of disease exposure and need preventive measures whose implementation can reduce the complications and symptoms of this parasite over several years [1]. Transmission of *Schistosoma* has been reported in 78 countries worldwide, while effective treatment measures have been followed in just 51 countries [1].

Schistosoma has a complex life cycle that relies on several organisms. *Schistosoma* larvae are released into the water by infected snails that live in freshwater [2]. After host contact with contaminated water, the larvae penetrate the skin, mature inside the body, and live inside the vessels. The female *Schistosoma* starts spawning, some of which are excreted in the urine or feces, and some remain in the body tissues, leading to the host's immune system reaction and increased damage to the host [2]. If the urine or feces of an infected host gets into fresh water, the eggs are released into the water, penetrating the snails' bodies and starting the next stages of their development. Finally, infected snails release *Schistosoma* larvae into the water, completing the cycle [3].

Infectious diseases caused by *Schistosoma* are known as schistosomiasis and are relatively common in Asian and African cattle [2]. Although the clinical symptoms of this disease are not severe, its high prevalence causes a significant impact on the growth and efficiency of livestock production [2]. In addition, this parasitic infection increases the possibility of getting infected with other parasites or bacterial diseases. That is why cattle infected with this parasite face a price drop [2].

During 2012-13, after reports of heavy losses of some sheep in the central district of Mazandaran Province, the Iran Veterinary Organization identified *Schistosoma* parasites in animals with high losses [4]. In the mentioned study, 21% of the infected sheep died. The same study showed that female sheep's loss rate is higher than male sheep's [4].

Reports show that 19 different strains of *Schistosoma* can cause disease in animals, of which 8 are more important [2]. *Schistosoma bovis* has a relatively high prevalence among different strains in the Middle East [2]. This parasite has male and female genders, and its various strains can mate with each other. Reproduction of pathogenic strains with pathogenic strains in humans has also been observed in vitro [2].

The WHO considers schistosomiasis a neglected tropical disease, meaning that research on this disease is insufficient. Since this parasite is relatively far from model organisms in the phylogenetic tree, many unknowns exist. For non-model organisms, the most common ways of functional annotation of genes are exploring the similarity of the protein sequences of genes with annotated proteins in other organisms [5]. In particular, two common methods for predicting gene function are sequence signature prediction with the InterProScan tool [6] and identifying the orthologous relationship of a gene with the genes of model organisms [5]. Despite the usefulness of these tools, if the protein sequence in one organism differs significantly from its homologs in other organisms, these tools lose their power.

The tertiary structure of proteins is more conserved than their sequence, so in some cases, even though two homologous proteins lack significant sequence similarity, both have substantial structural similarity [7]. For this reason, the structure of a protein provides researchers with new clues about its function. Until recently, the use of structure for function prediction was not very popular, mainly because of the small number of experimentally resolved protein structures. Until 2020, the most common structure prediction methods were based on homology modeling with experimentally resolved protein structures. These methods are known as template-based modeling. Due to the nature of these methods, they are not accurate enough to predict the structure of proteins whose similar proteins' structures have not been experimentally resolved. Despite these limitations, it has been shown that the protein structures predicted by template-based modeling methods can be useful for functional annotation [8].

AlphaFold, developed by Google DeepMind, was the first program to predict the structure of proteins with the same accuracy as experimental methods [9]. AlphaFold has little reliance on sequence similarity to the experimentally resolved protein structures, and the program's GitHub claims that the accuracy of the predicted structures does not change significantly if it works independently of the template structures. To predict the structure, AlphaFold first finds the homologous sequences to the query protein sequence and does multiple sequence alignment of all hits [9]. In the next step, AlphaFold predicts the distance between amino acids based on the

covariation of amino acids in homologous sequences, assuming that strongly covarying amino acids are close to each other in the tertiary structure of the protein [9]. In other words, AlphaFold predicts the distance of amino acids to predict the structure. AlphaFold can predict how two proteins fold next to each other with the same rationale, which can help infer protein-protein interaction, a capability introduced in AlphaFold-multimer [10].

As mentioned earlier, AlphaFold relies on multiple sequence alignments for structure prediction, but AlphaFold was not the first method to use sequence alignment for structure prediction. Previously, models such as Raptor-X used sequence alignment to predict the structure [11]. Still, none were as accurate as AlphaFold, whose credit goes to the efficient use of artificial intelligence and its branches, such as deep learning in structure prediction. In AlphaFold, a new type of neural network named Evoformer was used, establishing a connection between sequence alignment and structure using the attention mechanism [9]. Attention-based mechanisms in machine learning displayed significant success [12].

AlphaFold first needs to find homologous sequences of the query sequence to predict the structure. The default AlphaFold method to find similar sequences is computationally heavy and time-consuming. Shortly after introducing the AlphaFold, scientists presented a modified version, known as ColabFold, that is faster than the original version; it uses more optimized methods of sequence searching [13].

If there are not enough homologous sequences of the query sequence in the databases, AlphaFold cannot predict the structure accurately. ESMFold [14] and OmegaFold [15] are programs based on large language models and can predict protein structure without finding homologous sequences. The accuracy of these programs in routine structure prediction tasks is slightly lower than the accuracy of AlphaFold. Still, they can be run at a lower computational cost, and their lack of reliance on homologous sequences increases their range of applications. For phylogenetically remote sequences with limited homologous sequences, ESMfold and Omegafold are more accurate than AlphaFold [14, 15].

To predict the function of the query sequences, it is necessary to find structurally similar annotated proteins after having the structure in hand. Foldseek is a structural alignment program whose speed is thousands of times faster than alternative tools [16]. A combination of AlphaFold and Foldseek can be used on a large scale for functional annotation based on structure. Ruperti et al. used the combination of Foldseek and AlphaFold to predict the function of proteins in a marine sponge [17]. First, they found the closest protein structures to sponge proteins in model organisms using Foldseek and then attributed the function of the proteins in the closest hit's orthology group. Their findings showed that functionally annotated genes can increase to 50% if protein structures are used for their workflow rather than protein sequences [17].

In another study, it was reported that using Foldseek instead of BLASTP to find reciprocal best-hit relationships between proteomes of two organisms can uncover new orthology relationships, which in turn can be employed for functional annotation of the proteins [18].

In this study, we aimed to predict the function of some hypothetical genes in the *Schistosoma* parasite based on the orthology of its proteins to the proteins of *Caenorhabditis elegans*, the closest model organism to Schistosoma. To this end, we predicted the orthology relationship by finding the reciprocal best structural hits between the structures of the organisms' proteins. Next, the function of the annotated proteins in *C. elegans* was attributed to their orthologs in Schistosoma. As a case study, we have provided evidence supporting the predicted functions.

Materials and Methods

Among different *Schistosoma* strains, the structure of the proteome of *Schistosoma mansoni* has been predicted by AlphaFold. Accordingly, it was used for functional annotation purposes in this study. The fourth version of protein structures was downloaded in tarred format from the AlphaFold website [19]. Tarred files included structures in both PDB and mmCIF formats. Structures in mmCIF format were removed, and PDB structures were tarred again. All processes were done in Ubuntu 18.04. Foldseek was run with the "rbh" option for reporting the reciprocal best hits between two proteomes. The scripts used for this study are available at github.

Gene ontology data for *S. mansoni* and *C. elegans* were retrieved from UniProt by specifying the taxonomy ID in the search bar. The taxonomy IDs of *S. mansoni* and *C. elegans* are 6183 and 6239, respectively. The retrieved data was downloaded in tsv format. For functional domain inspection, the domains predicted by Pfam that are accessible on the InterPro website [20] were used.

ColabFold was used to predict the interaction of two proteins. The number of recycles was set to 3, and the

latest version of AlphaFold parameters was utilized for the structure prediction. TM-align was used to compare the structures of two proteins, which are available on the RCSB website [20].

Results

AlphaFold contains 13865 predicted structures for *S. mansoni* and 19694 for *C. elegans*. AlphaFold reports its confidence estimate for each amino acid with pLDDT, which is a number between 0 and 100, and higher values indicate a higher confidence level. Figure 1A shows the distribution of the "mean of pLDDT for each protein" for both *S. mansoni* and *C. elegans*. Figure 1B displays the distribution of the fraction of amino acids in each protein with a pLDDT below 50.

Foldseek could establish the reciprocal best-hit relationship for 3341 protein pairs in *C. elegans* and *S. mansoni*. Among the pairs, 182 pairs had a bit score below 100. The distribution of sequence identity for the hits is shown in Figure 2.

Among the established relationships, 193 relationships were made by hypothetical proteins in *S. mansoni*. Of the hypothetical proteins, 36 lacked predicted molecular functions, even though their counterparts in the *C. elegans* worm had known functions. Similarly, for 60 proteins, the biological pathway was unknown, and for another 60, their final cellular location remained unidentified despite this information being available for their *C. elegans* counterparts.

As a case study, we looked deeper at the predictions made for two uncharacterized proteins of *S. mansoni*. The availability of orthogonal data for function verification has been the rationale for selecting proteins.

A0A5K4FEN7 is probably involved in embryogenesis

The protein A0A5K4FEN7, which is known as the "uncharacterized protein" in *S. mansoni*, is the reciprocal best match with protein Q21194, "described as a guanine nucleotide exchange factor rei-2" in *C. elegans*. The structure and predicted aligned error figure of A0A5K4FEN7 are shown in Figure 3. Figure 4 shows the structural alignment of A0A5K4FEN7 with Q21194. According to the UniProt database, both proteins have a Pfam family of SH3 domain-binding protein 5 (SH3BP5). The 4 alpha helices shown in Figure 4 are the domain-containing regions of both proteins.

Q21194 protein contributes to the embryogenesis and cell division in *C. elegans* by regulating the distribution of the Rab 11.1 gene (UniProt ID O01803). Our data show that the reciprocal best match of O01803 in *S. mansoni* is A0A5K4F920. Since the function of Q21194 is dependent on its interaction with O01803, to confirm the prediction further, the interaction of these two proteins in the *C. elegans* and the *S. mansoni* was predicted and compared using ColabFold. The results show that the ipTM values for two proteins in *C. elegans* and *S. mansoni* parasite are 0.728 and 0.802, respectively.

Four alpha helices found in both proteins are similar to each other. The same Pfam domain has been predicted for the same region of the proteins. The TM-score of the alignments is 0.51.

A0A5K4F0A7 is likely to be involved in a neurotransmitter synthesis process

A0A5K4F0A7 in *S. mansoni* and Q9XTQ6 in *C. elegans* are the best reciprocal hits. The AlphaFold predictions are shown in Figure 5. A0A5K4F0A7 protein is

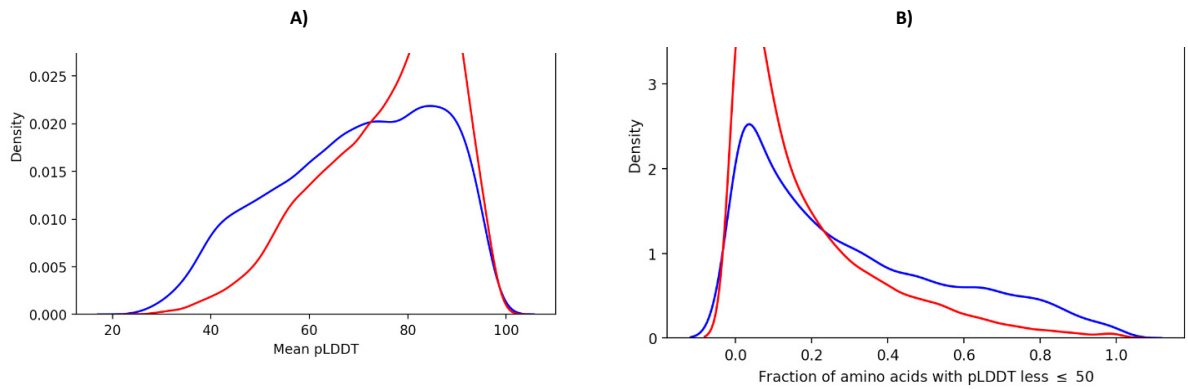

Figure 1. Comparing pLDDT values for *S. mansoni* and *C. elegans*

A) The distribution of the mean pLDDT for each protein, B) The distribution of the fraction of amino acids with pLDDT below 50

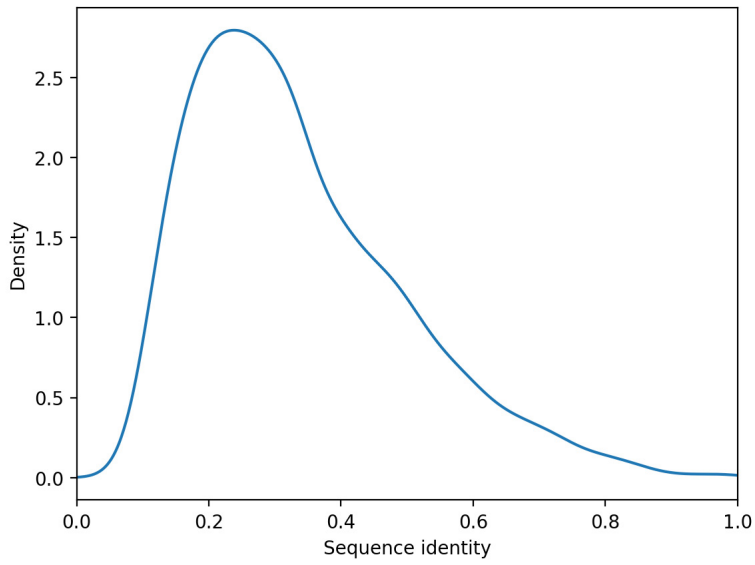

Figure 2. The distribution of sequence identity of Foldseek hits

described as an "uncharacterized protein" in *S. mansoni*, while Q9XTQ6 is described as "tyramine beta-hydroxylase" in *C. elegans*.

The structural alignment of A0A5K4F0A7 and Q9XTQ6 is shown in Figure 6. Studies have shown that this protein is needed to convert tyramine to octopamine, a neurotransmitter [21]. Q9XTQ6 has two domains for binding to the divalent copper ions. Similar domains are also present in A0A5K4F0A7. Furthermore, the PROSITE patterns (accessible through the InterProI website) [6] show that both A0A5K4F0A7 and Q9XTQ6 contain the PROSITE pattern (ID: PS00084) described as "copper type II, ascorbate-dependent monooxygenases signature 1."

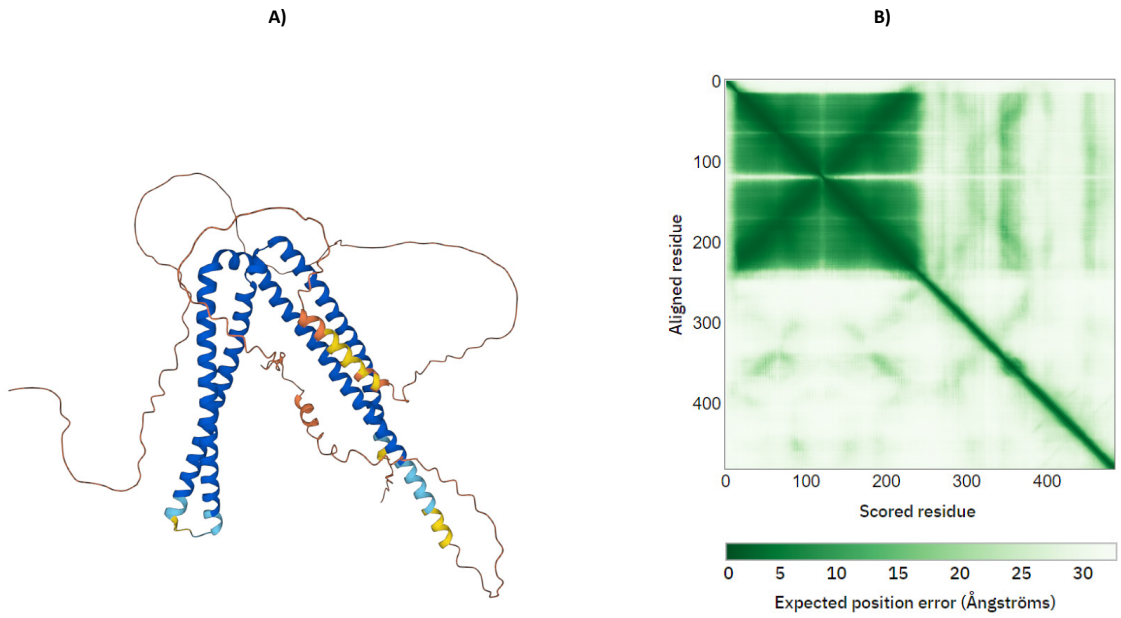

Figure 3. Alphafold predictions for A0A5K4FEN7

A) Predicted structure, B) pAE plot the residues are colored based on pLDDT

Note: Low values are shown in orange/yellow, and high values are displayed in blue. The photos were retrieved from the AlphaFold website.

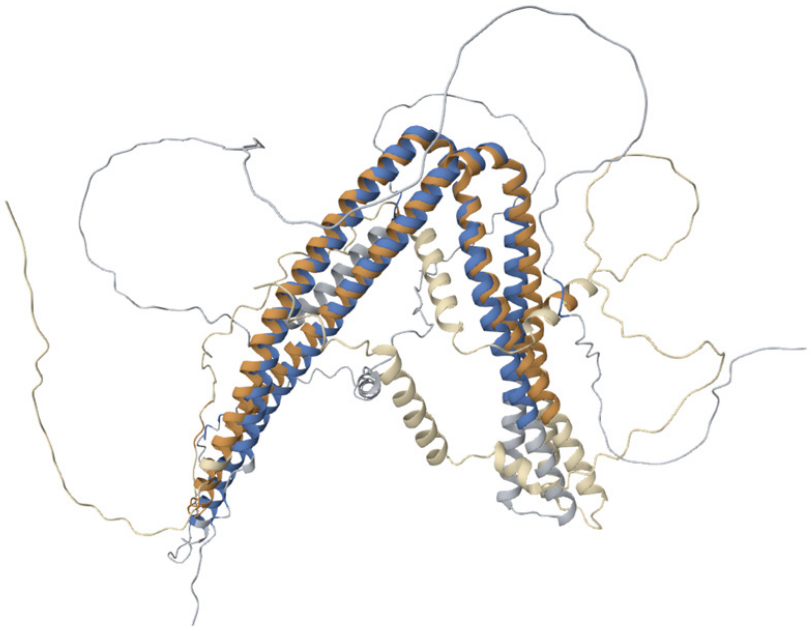

Figure 4. The structural alignment of Q21194 (orange) with A0A5K4FEN7 (blue) shows the structural alignment of A0A5K-4FEN7 with Q21194

Note: According to the UniProt database, both proteins have a Pfam family of SH3 domain-binding protein 5 (SH3BP5). The 4 alpha helices are the domain-containing regions of both proteins.

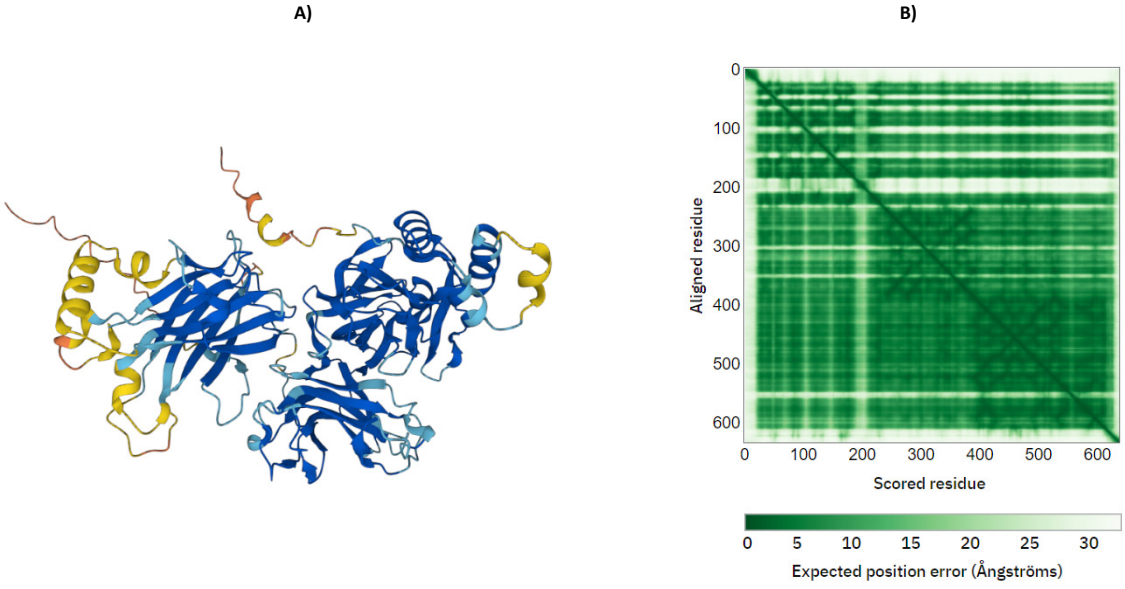

Figure 5. AlphaFold predictions for A0A5K4F0A7

A) The predicted 3D structure, B) The pAE Plot

Note: The photos were retrieved from the AlphaFold website.

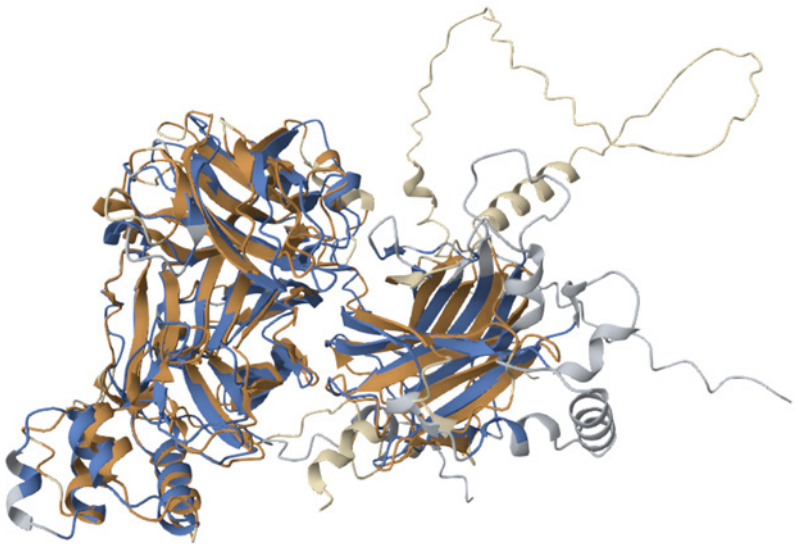

Figure 6. The structural alignment of Q9XTQ6 (orange) with A0A5K4F0A7 (blue)
Note: The TM-score of the alignment is 0.75.

Discussion

The direct comparison of pLDDT distribution for *S. mansoni* and *C. elegans* proteome (Figure 1) shows that the *C. elegans* structures have a higher confidence. Figure 1A illustrates that the mean confidence score is generally lower for *S. mansoni*. Figure 1B demonstrates that in *S. mansoni*, a higher fraction of each protein has a lower confidence score than *C. elegans*. As mentioned in the introduction, AlphaFold relies on sequence alignment of the query protein with its homologs for the structure prediction. As *S. mansoni* is a non-model organism, we expect a relatively smaller number of homologous proteins to be available than *C. elegans*, which might have lower confidence scores.

Figure 2 shows most established relationships have less than 30% sequence identity. If such relationships were established with sequence alignment tools, they would have been in a twilight zone and could not be used reliably [22]. On the other hand, according to the Foldseek reports, proteins with a bit score above 100 most likely belong to the same structural superfamilies, showing that most established relationships are highly confident. Therefore, this study exemplifies the higher performance of the structural alignment tools compared to sequence alignment tools for establishing the reciprocal best-hit relationship.

Considering the precise position of each amino acid, the pAE plot shows how well we can predict the location of other amino acids with respect to the query residue. If two amino acids are part of a rigid globular domain, their pAE would be zero, as shown by dark green color in pAE plots (Figures 3B and 5B). According to the definition, A0A5K4FEN7 contains a globular N-terminal domain and an unstructured turn, shown in a spaghetti-like stand in Figure 3B. The globular domain corresponds to the 4 helices shown in Figure 3A. The rest of the protein does not fold into a rigid domain. As mentioned, the Pfam predicted domain also coincides with the domain inferred from the predicted 3D structure. For A0A5K4F0A7, the pAE plot shows a big central domain and two tiny non-domain structures in N and C-terminals (Figure 5B). Both non-domain structures have a low confidence score, as shown by the orange and yellow colors in (Figure 5A).

The TM-scores of A0A5K4FEN7 and A0A5K4F0A7 with their corresponding hits are above 0.5, showing that both hits are highly similar. The ipTM shows the reliability of interaction prediction, and values higher than 0.5 are considered significant. Given the high structural similarity of A0A5K4FEN7 with Q21194, the similarity of the domains, the existence of its interacting protein in *S. mansoni*, and the high confidence interaction of A0A5K4FEN7 with its partner, A0A5K4F920, we proposed that A0A5K4FEN7 has the same function as its counterpart

in *C. elegans* and contributes to embryogenesis and cell division.

For A0A5K4F0A7, above domain similarity and high structural similarity (TM-score=0.75) with its best reciprocal hit (Q9XTQ6), both A0A5K4F0A7 and Q9XTQ6 contain the identical PROSITE pattern. PROSITE patterns are relatively short patterns summarizing the conserved patterns observed in the active site of enzymes. For A0A5K4F0A7, the PS00084 has been predicted whose pattern is "H-H-M-x(2)-F-x-C," meaning that the active site should contain "histidine-histidine-methionine-x-x-phenylalanine-x-cysteine," where x could be any amino acid. This pattern has been predicted on 294-301 of A0A5K4F0A7, suggesting that not only the overall domain sequence but also the active site pattern is conserved. Given all this information, it is proposed that A0A5K4F0A7 is probably involved in the egg-laying process of *Schistosoma* parasites. Considering the importance of the egg-laying process for the survival of its generation, knocking down A0A5K4F0A7 with RNAi is suggested to see if its knockdown could stop the parasite's reproduction.

Conclusion

Our results show that for 193 uncharacterized proteins in *S. mansoni*, the reciprocal best hit in *C. elegans* can be found using structural similarity. Two of such proteins were further studied. In both cases, proteins in *S. mansoni* and their counterparts in *C. elegans* were annotated with identical Pfam domains. Q9XTQ6 is involved in regulating the distribution of Rab11.1 in *C. elegans*. Our data show that A0A5K4F0A7, the counterpart of Q9XTQ6 in *S. mansoni*, also interacts with the ortholog of Rab11.1. Besides, our data suggest that according to structural similarity, domain similarity, and having identical sequence signatures with Q9XTQ6, A0A5K4F0A7 might be involved in the egg-laying process of Schistosoma.

In summary, our results show that protein structures can be used to predict the function of hypothetical proteins in less-studied organisms. Besides, the protein-protein interaction prediction using AlphaFold can help validate the predicted functions. In future studies, we would like to explore the knocking down of A0A5K4F0A7 and its effect on the egg-laying of the *S. mansoni*.

Ethical Considerations

Compliance with ethical guidelines

All ethical principles were considered in this article.

Funding

This work was financially supported by Department of Animal Sciences, Agricultural Sciences and Natural Resources University of Khuzestan.

Authors contribution's

Study design: Arezou Askari Rad, Jamal Fayazi and Alireza Hasani Baferani; Experiments: Arezou Askari Rad; Supervision: Jamal Fayazi, Mohammad Taghi Beigi Nassiri and Alireza Hasani Baferani; Final approval: All authors.

Conflict of interest

The authors declared no conflict of interest.

Acknowledgements

The authors acknowledge the support provided by the Research Deputy of the Agricultural Sciences and Natural Resources University of Khuzestan, whose assistance was invaluable.

References

[1] WHO. Schistosomiasis. Geneva: WHO; 2023. [Link]

[2] Merck, Co I. The merck veterinary manual. New Jersey: Merck & Company; 2008. [Link]

[3] Centers for Disease Control and Prevention. Disease. Atlanta: Centers for Disease Control and Prevention; 2023. [Link]

[4] Hosseini SH, Rohani Shahrestani S, Tabari Pour R, Yousefi MR. [Report of sheep ornithobilharzia in central of Mazandaran province 2102-13 (Persian)]. J Vet Res. 2016; 71(3):277-81. [Link]

[5] No Author. Gene Ontology Consortium [Internet]. 2023 [Updated 9 September 2023]. Available from: [Link]

[6] Paysan-Lafosse T, Blum M, Chuguransky S, Grego T, Pinto BL, Salazar GA, et al. InterPro in 2022. Nucleic Acids Res. 2023; 51(D1):D418-27. [DOI:10.1093/nar/gkac993] [PMID]

[7] Illergård K, Ardell DH, Elofsson A. Structure is three to ten times more conserved than sequence-a study of structural response in protein cores. Protein. 2009; 77(3):499-508. [DOI:10.1002/prot.22458] [PMID]

[8] Gligorijević V, Renfrew PD, Kosciolek T, Leman JK, Berenberg D, Vatanen T, et al. Structure-based protein function prediction using graph convolutional networks. Nat Commun. 2021; 12(1):3168. [DOI:10.1038/s41467-021-23303-9] [PMID]

[9] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021; 596(7873):583-9. [DOI:10.1038/s41586-021-03819-2] [PMID]

[10] Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-Multimer. biorxiv. 2021. [Preprint]. [DOI:10.1101/2021.10.04.463034]

[11] Xu J. Distance-based protein folding powered by deep learning. Proc Natl Acad Sci U S A. 2019; 116(34):16856-65. [DOI:10.1073/pnas.1821309116] [PMID]

[12] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. Advances in Neural Information Processing Systems. 2017. [Link]

[13] Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: Making protein folding accessible to all. Nat Methods. 2022; 19(6):679-82. [DOI:10.1038/s41592-022-01488-1] [PMID]

[14] Lin Z, Akin H, Rao R, Hie B, Zhu Z, Lu W, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. Science. 2023; 379(6637):1123-30. [DOI:10.1126/science.ade2574] [PMID]

[15] Wu R, Ding F, Wang R, Shen R, Zhang X, Luo S, et al. High-resolution de novo structure prediction from primary sequence. BioRxiv. 2022. [DOI:10.1101/2022.07.21.500999]

[16] van Kempen M, Kim SS, Tumescheit C, Mirdita M, Lee J, Gilchrist CLM, et al. Fast and accurate protein structure search with Foldseek. Nat Biotechnol. 2024; 42(2):243-6. [DOI:10.1038/s41587-023-01773-0] [PMID]

[17] Ruperti F, Papadopoulos N, Musser JM, Mirdita M, Steinegger M, Arendt D. Cross-phyla protein annotation by structural prediction and alignment. Genome Biol. 2023; 24(1):113. [DOI:10.1186/s13059-023-02942-9] [PMID]

[18] Monzon V, Paysan-Lafosse T, Wood V, Bateman A. Reciprocal best structure hits: Using AlphaFold models to discover distant homologues. Bioinform Adv. 2022; 2(1):vbac072. [DOI:10.1093/bioadv/vbac072] [PMID]

[19] No Author. AlphaFold website [Internet]. 2023 [Updated 9 September 2023]. Available from: [Link]

[20] Zhang Y, Skolnick J. TM-align: A protein structure alignment algorithm based on the TM-score. Nucleic Acids Res. 2005; 33(7):2302-9. [DOI:10.1093/nar/gki524] [PMID]

[21] Alkema MJ, Hunter-Ensor M, Ringstad N, Horvitz HR. Tyramine functions independently of octopamine in the Caenorhabditis elegans nervous system. Neuron. 2005; 46(2):247-60. [DOI:10.1016/j.neuron.2005.02.024] [PMID]

[22] Rost B. Twilight zone of protein sequence alignments. Protein Eng. 1999; 12(2):85-94. [DOI:10.1093/protein/12.2.85] [PMID]

This Page Intentionally Left Blank