Science Letters ISSN 2345-5463 DOI: https://doi.org/10.47262/SL/14.1.132026120

Research article
2026 | Volume 14 | Issue 1 | Article ID 1326120sl

ARTICLE INFO

Received
January 05, 2026
Revised
March 08, 2026
Accepted
March 11, 2026
Published
April 13, 2026

***Corresponding author**
Muhammad Ehsan
E-mail
muhammad.ehsan@iub.edu.pk

Keywords
Fasciola gigantica
In silico analysis
Superantigen-like protein
Structural bioinformatics
AlphaFold modeling

How to Cite
Yousaf A, Ehsan M, Shahid K, Shahzad MI, Javed A. In silico structural and functional profiling of FGIG_11154, a novel immune-interacting protein of *Fasciola gigantica*. Science Letters 2026; 14(1):1326120sl

SCAN ME

Open Access

In Silico Structural and Functional Profiling of FGIG_11154, a Novel Immune-Interacting Protein of *Fasciola gigantica*

Areeba Yousaf[1,2], Muhammad Ehsan[1]*, Kinza Shahid[1,2], Mirza Imran Shahzad[2], Amir Javed[3]

[1]Department of Parasitology, Faculty of Veterinary and Animal Sciences, The Islamia University of Bahawalpur, 63100, Punjab, Pakistan
[2]Department of Biochemistry & Molecular Biology, Institute of Biochemistry, Biotechnology & Bioinformatics, The Islamia University of Bahawalpur, 63100 Punjab, Pakistan
[3]Department of Physiology and Biochemistry, Cholistan University of Veterinary and Animal Sciences, Bahawalpur, 63100, Punjab, Pakistan

Abstract

Fasciola gigantica is a major trematode parasite of veterinary and public health importance, yet a substantial proportion of its proteome remains functionally uncharacterized. In this study, *in silico* characterization of the hypothetical protein FGIG_11154 (226 amino acids) was performed to infer its structural and functional properties. Physicochemical analysis revealed a hydrophilic, instability-prone protein with a negative GRAVY value and a high proportion of charged residues, suggesting potential involvement in molecular interactions. Sequence similarity analysis demonstrated moderate conservation among *Fasciola* species, while limited homology with bacterial proteins indicated lineage-specific functionality. Conserved domain analysis identified a significant superantigen-like protein SSL4 domain, implicating a possible role in host immune modulation. Subcellular localization predictions yielded contrasting results, indicating potential extracellular as well as nuclear localization, suggesting multifunctional or context-dependent behavior. Signal peptide and transmembrane helix analyses confirmed that FGIG_11154 is a non-membrane, soluble protein likely secreted via a non-classical pathway. Secondary structure prediction revealed a coil-rich architecture, consistent with intrinsic disorder. Three-dimensional structure modeling using AlphaFold indicated low per-residue confidence, further supporting the presence of disordered regions; however, structural validation via ProSA-web and Ramachandran plot analysis confirmed overall model plausibility. Collectively, these findings suggest that FGIG_11154 may function as a flexible, immune-interacting protein potentially involved in host–parasite interactions.

Introduction

Fasciola gigantica is a helminthic trematode that causes fascioliasis, a zoonotic disease often overlooked but with a significant global impact on livestock and humans, particularly in tropical and subtropical regions [1]. The parasite causes significant economic losses to the livestock industry due to low productivity, liver condemnation, and high treatment costs, as well as an increasing concern for the general human population [2]. Although it is of significance, the molecular pathways of the survival of parasites, adaptation to their hosts, and immune evasion are not fully understood.

Recent genome sequencing studies have shown that a very large proportion of the *F. gigantica* proteome is made up of hypothetical proteins that are not functionally annotated [3]. These unknown proteins are a knowledge gap, because they can be important in the development of the parasites, their interactions with the host, and pathogenicity. Such proteins should be elucidated functionally to identify new drug targets, vaccine candidates, and diagnostic markers. *In silico* methods have become strong tools for the initial characterization of the hypothetical proteins, especially in organisms that are difficult to validate in practice [4]. Computational studies combining physicochemical profiling, sequence homology, conserved domain detection, subcellular localization prediction, and structural modeling can be very useful in the understanding of protein functionality, stability, and biological relevance. This has been effectively used in the case of parasitic helminths to select proteins for downstream experimental studies [4, 5]. Of particular interest in the biology, trematodes is the group of proteins involved in host immune modulation because they need to circumvent or manipulate host defense systems in order to allow chronic infections [6, 7]. The intrinsically disordered regions of such proteins also indicate the presence of molecular recognition, signaling, and interaction with various targets of the host [8].

This paper provides an *in silico* characterization of a hypothetical *F. gigantica* protein, FGIG_11154, by using a set of sequence-based prediction tools, structural prediction tools, and localization prediction tools. This study aims to determine the possible functional role and biological importance of it. This article offers novel details on a protein that had not been previously annotated and is a part of the emerging literature on the *F. gigantica* molecular biology, which would form the basis of subsequent experimental confirmation and therapeutic investigation.

Materials and Methods

Retrieval of protein sequence

The hypothetical protein FGIG_11154 of *F. gigantica* was used to extract the amino acid sequence of the protein using the National Center of Biotechnology Information (NCBI) database of proteins (https://www.ncbi.nlm.nih.gov/protein). This sequence was retrieved in the form of FASTA with the accession number TPP64944.1 in GenBank. The protein has 226 amino acids, and it became the reference sequence of all future computational analyses [9].

Physicochemical property analysis

The physicochemical characterization of FGIG_11154 was conducted under the ProtParam tool that is provided in the ExPASy server (https://web.expasy.org/protparam/). The analysis comprised molecular composition, total number of positively and negatively charged residues, extinction coefficient, estimated half-life, instability index, aliphatic index, and grand average of hydropathicity (GRAVY). These parameters were determined from the primary amino acid sequence and were utilized in studying the protein stability, solubility, and biochemical behavior [10].

Sequence similarity analysis

The sequence homology was performed using BLASTp on the NCBI non-redundant (nr) protein database (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Standard settings such as percent identity, query coverage, E-values, and functional annotations of top hits were used to evaluate evolutionary conservation and imply possible functional relevance [11].

Conserved domain and motif identification

The protein sequence was initially searched against InterProScan (https://www.ebi.ac.uk/interpro/search/sequence/), which consists of a combination of protein signature databases, to identify conserved domains. Additional domain validation was done in the NCBI Conserved Domain Database (CDD) (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Domains identified were also tested depending on domain boundaries, accession numbers, and E-values to inform functional significance [12, 13].

Subcellular localization prediction

Subcellular localization was predicted using two independent computational tools. CELLO v2.5 (https://cello.life.nctu.edu.tw/) was employed to predict localization based on support vector machine (SVM) classification [14]. PSORT II (https://psort.hgc.jp/form2.html) was used to estimate localization probabilities based on k-nearest neighbor (k-NN) analysis and motif recognition [15]. Predictions from both tools were compared to assess potential multifunctional or context-dependent localization.

Signal peptide and transmembrane helix prediction

The presence of a classical N-terminal signal peptide was evaluated using SignalP version 6.0 (https://services.healthtech.dtu.dk/service.php?SignalP-6.0). Prediction of transmembrane helices was performed using TMHMM version 2.0 (https://services.healthtech.dtu.dk/service.php?TMHMM-2.0). These analyses were conducted to determine secretion pathways and membrane association [16, 17].

Secondary structure prediction

Secondary structure elements were predicted using GOR IV via the NPS@ server (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_gor4.html) and PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/). The proportions of α-helices, β-strands, and random coils were calculated, and qualitative agreement between both methods was assessed to infer folding tendencies and structural flexibility [18, 19].

Tertiary structure prediction

The three-dimensional structure of FGIG_11154 was predicted using the AlphaFold Protein Structure Database (https://alphafold.ebi.ac.uk/). The predicted PDB model corresponding to the full-length sequence was retrieved and used for further validation. Model confidence was assessed using predicted Local Distance Difference Test (pLDDT) scores [20].

Structural validation and quality assessment

Global and local model quality was tested using the ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php) to conduct the structural validation of the model predicted. Z-score values and energy plots of the residues were compared. The quality of stereochemicals was also evaluated with the help of PROCHECK, and the statistics with Ramachandran plots were calculated to determine the distribution of the backbone dihedral angles and the overall reliability of the structures [21, 22].

Assessment of structural disorder

Structural predictions were understood together with the secondary structure composition and AlphaFold confidence scores to assess the occurrence of intrinsically disordered regions [20]. Data interpretation took into consideration the relationship between the predicted disorder, low-confidence structural areas, and the possible functional flexibility.

Results

Sequence retrieval and primary structure analysis

The hypothetical protein FGIG_11154 of *F. gigantica* was successfully obtained in the NCBI database of proteins with an accession number of TPP64944.1. The protein is composed of 226 amino acids and the N-terminal methionine residue. The main sequence was highly polar and charged, which implied the possibility of polar and charged residues being involved in the molecular interactions and not in the rigid structural function.

Physicochemical characterization

The ProtParam tool was used to determine the physicochemical properties of FGIG 11154 (Table 1). The protein had 29 negatively charged residues (Asp + Glu) and 19 positively charged residues (ARG + Lys), which gives the protein an overall negative charge. The instability index calculated was 67.64, and this was found to indicate an unstable protein; the aliphatic index was found to be 57.52, which indicated moderate thermostability. The negative GRAVY value (-0.581) indicated that the protein was a hydrophilic one, which is in line with the fact that it was a soluble non-membrane-associated protein. It was predicted to have a half-life of around 30 hours in mammalian reticulocytes, more than 20 hours in yeast, and more than 10 hours in Escherichia coli, which is moderate stability in a variety of biological systems.

Sequence similarity and homology Analysis

Analysis of the BLASTp data with the NCBI non-redundant protein database showed that FGIG_11154 is conserved in the *Fasciola* species (Table 2). The most similar protein was the hypothetical proteins of

Table 1 Physicochemical properties of *FGIG_11154* predicted using ProtParam.

Atomic composition	
Carbon	1076
Hydrogen	1687
Nitrogen	297
Oxygen	373
Sulfur	6
Formula	$C_{1076}H_{1687}N_{297}O_{373}S_6$
Total number of atoms	3439
Extinction coefficients	
Extinction coefficients are in units of $M^{-1}cm^{-1}$, at 280 nm in water.	
Ext. coefficient	12615
Abs 0.1% (=1g/l), (assuming all pairs of Cys residues form cystines)	0.506
Ext. coefficient	12490
Abs 0.1% (=1 g/l), (assuming all Cys residues are reduced)	0.501
Estimated half-life	
Mammalian reticulocytes *(in vitro)*	30 hours
Yeast *(in vivo)*	>20 hours
Escherichia coli *(in vivo)*	>10 hours
Instability index	
Instability index (II)	67.64
Aliphatic index	57.52
Grand average of hydropathicity (GRAVY)	-0.581

F. hepatica with 55-56% sequence identity, with about 48% query coverage and statistically significant E-values. Conversely, low similarity (<31% identity) was identified with bacterial proteins, such as representatives of the OmpA family, which implies that FGIG_11154 is a trematode-specific or lineage-specific protein. Such observations indicate evolutionary conservation in species of *Fasciola* as well as in favor of the lack of well-characterized homologs.

Conserved domain identification

InterProScan of the conserved domains did not find any known protein families or domains. But further examination on the NCBI Conserved Domain Database (CDD) revealed that it contained a superantigen-like protein SSL4 domain (PRK13042) between amino acid residues 38 and 115 with an E-value of 4.18 × 10^-6 (Table 3), indicating statistical significance.

The discovery of this domain indicates that it may be involved in immune-related mechanisms, especially in host-parasite interactions. The difference between InterProscan and CDD findings underscores the need to have a combination of functional inference tools (Fig. 1).

Subcellular localization prediction

Subcellular localization predictions yielded contrasting results (Table 4). CELLO predicted FGIG_11154 to be extracellular with the highest support vector machine score (1.880), suggesting possible secretion or surface association. In contrast, PSORT II predicted predominant nuclear localization with 94.1% reliability, supported by the presence of putative nuclear localization signals (KRHR and RHRK motifs). These divergent predictions indicate that FGIG_11154 may exhibit multifunctional or context-dependent localization, a feature commonly observed in regulatory or immune-interacting proteins.

Signal peptide and transmembrane helix analysis

SignalP 6.0 analysis revealed the absence of a classical N-terminal signal peptide, suggesting that FGIG_11154 is unlikely to be secreted via the conventional endoplasmic reticulum-Golgi pathway and may instead follow a non-classical secretion mechanism. Furthermore, TMHMM analysis predicted no transmembrane helices, indicating that the protein is not membrane-associated and is likely soluble in nature. These results support the possibility of non-classical secretion mechanisms or intracellular functional roles.

Secondary structure prediction

The analysis of the secondary structure by GOR IV showed that FGIG_11154 is mainly composed of random coils (64.60%), then there are alpha-helices (21.24%), and lastly the 2-strands (14.16) (Table 5). No 3_{10} helices, π helices, or β bridges were identified. The GOR IV results were reproducible with PSIPRED predictions that showed that there was a coil-rich architecture, which is a signature of structural flexibility and possible intrinsic disorder.

Table 2 BLASTp results showing homologous proteins of FGIG_11154.

Accession	Organism	Protein / Annotation	Identity (%)	Query Coverage	E-value
TPP64944.1	*Fasciola gigantica*	Hypothetical protein FGIG_11154	100	100	1e-161
CAM0512007.1	*Fasciola hepatica*	Unnamed protein product	55.05	48	2e-19
THD20126.1	*Fasciola hepatica*	Hypothetical protein D915_009213	56.52	48	6e-05
WP_140735178.1	*Mesorhizobium sp. B2-3-4*	OmpA family protein	30.69	45	3e-05

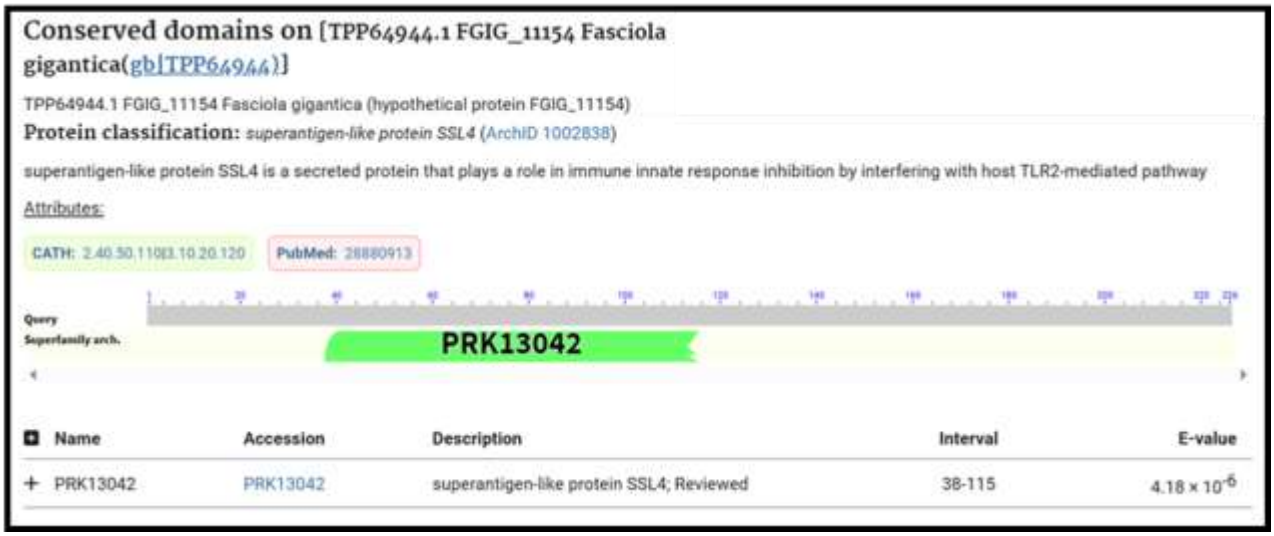

Fig. 1 Conserved domain architecture of FGIG_11154 showing the SSL4 domain (residues 38–115) identified using NCBI-CDD.

Table 3 Conserved domain analysis of FGIG_11154.

Tool used	Domain detected	Accession	Position (aa)	E-value	Functional insight
InterProScan	None detected	–	–	–	Protein may be poorly characterized
NCBI CDD	Superantigen-like protein SSL4	PRK13042	38–115	4.18×10^{-6}	Possible role in immune modulation

Table 4 Subcellular localization prediction of FGIG_11154 using CELLO and PSORT II.

Tool	Predicted localization	Score / Reliability	Key evidence	Biological interpretation
CELLO	Extracellular	Score = 1.880	Highest SVM score among all compartments	May be secreted or surface-associated; possible role in host–parasite interaction
PSORT II	Nuclear	94.1% reliability	Presence of putative NLS motifs (KRHR, RHRK); k-NN prediction (82.6% nuclear)	Suggests potential regulatory or DNA-associated function

Table 5 Predicted secondary structure composition of FGIG_11154.

Secondary structure element	GOR IV (Residues)	GOR IV (%)	PSIPRED (Prediction trend)*
α-Helix (H)	48	21.24	Moderately abundant
β-Strand (E)	32	14.16	Present in multiple regions
Random coil (C)	146	64.60	Predominant
3_{10} Helix (G)	0	0.00	Not detected
π-Helix (I)	0	0.00	Not detected
β-Bridge (B)	0	0.00	Not detected
Turns / bends	0	0.00	Minimal
Ambiguous states	0	0.00	None

Tertiary structure prediction

AlphaFold Protein Structure Database was used to predict the three-dimensional structure of FGIG_11154. The model had a low average pLDDT value (35.81), which means that it has low per-residue confidence and is, therefore, mostly an intrinsically disordered protein. The predicted structure was used as a framework for further validation and assessment of disorder, even in the case of low confidence scores (Fig. 2).

Structural validation and quality assessment

The structural validation with ProSA-web gave a Z-score of 0.11, a range that is acceptable with native

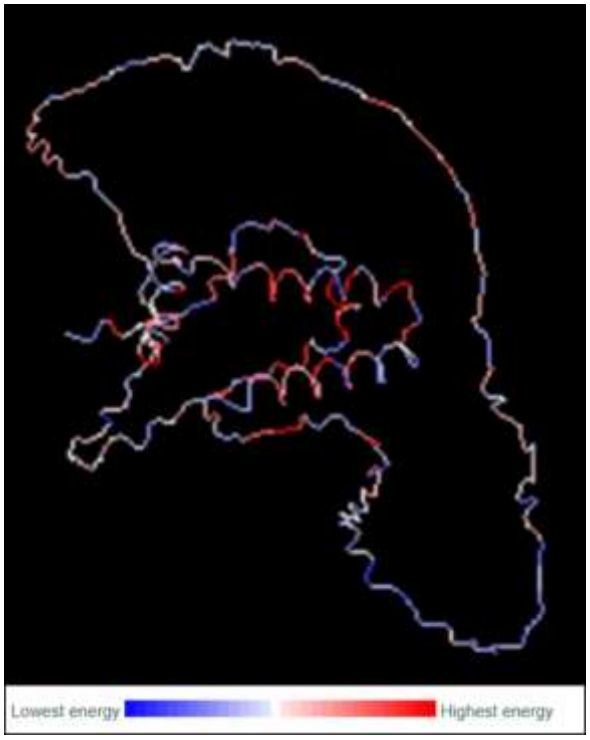

Fig. 2 AlphaFold-predicted tertiary structure of FGIG_11154 colored by pLDDT confidence scores. Regions with low confidence (blue) indicate intrinsic disorder, while higher-confidence regions (red) represent comparatively structured segments.

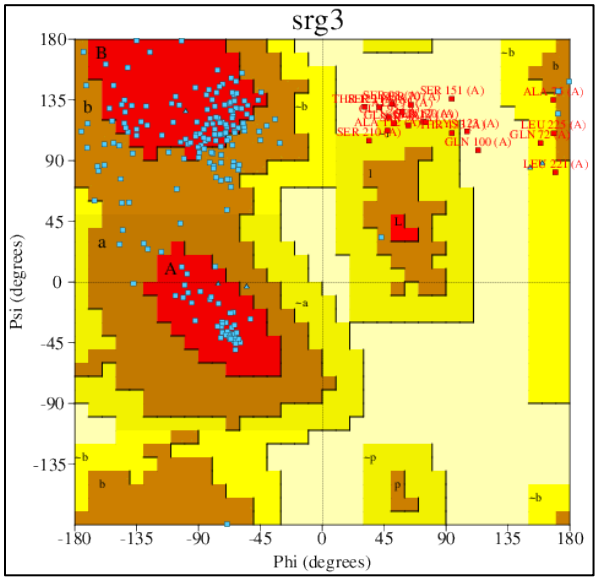

Fig. 3 Ramachandran plot analysis of the predicted FGIG_11154 structure generated using PROCHECK.

proteins of the same size, and goes to show that the global model is of reasonable quality. The energy plot in the form of a curve made of residues did not have any extreme energy values, indicating that there were no significant structural discrepancies. The Ramachandran plot analysis showed that the proportion of residues in the most preferred regions was 59.3%, the proportion in the additional allowed regions was 31.1%, the proportion in the generously allowed regions was 5.3%, and that of disallowed regions was only 4.3% (Table 6). The total G-factor value of -0.67 showed good stereochemical geometry. The intermediate statistics of residues in preferred regions are in line with the occurrence of intrinsic disordered sections (Fig. 3).

Discussion

This study presents the detailed *in silico* characterization of the hypothetical protein FGIG_11154 of *F. gigantica* and gives renewed information about the possible biological role of this protein in the survival of the parasite and its interrelation with the host. The percentage of hypothetical proteins in trematode genomes is significant, and functional annotation of hypothetical proteins is a significant problem. This study enhances the comprehension of a hitherto unknown protein that could have immunomodulatory implications through the incorporation of physicochemical, structural, and functional prediction analyses.

Physicochemical analysis indicated that FGIG_11154 is a hydrophilic protein having a high percentage of charged residues and a negative GRAVY value, which indicates a soluble property and possible involvement of the protein in protein-protein interactions or protein-host interactions [23]. The instability index of the protein indicated that it was unstable, although these properties are often found in regulatory proteins or proteins that interact to mediate an effect, such as the immune modulators or the signaling proteins. Instability profiles have also been found to be similar in parasite effector proteins that depend on conformational flexibility and not on rigid structural stability to do their job [24]. Analyses of sequence similarity showed that FGIG 11154 was moderately conserved across Fasciola species, but had little homology with other taxa' proteins. This conservation by lineage implies that it is possible that FGIG_11154 is capable of doing something specific to parasites, which is not completely conserved throughout organisms [25]. These proteins have been suggested to be involved in host adaptation and immune evasion, with many of them being good targets in therapeutic or diagnostic studies. The small degree of similarity to bacterial proteins further increases the chances of functional redundancy with host proteins, which is a key factor to consider in drug

Table 6 Ramachandran plot statistics of FGIG_11154.

Region Type	Residues	Percentage (%)
Most favored regions (A, B, L)	124	59.3
Additional allowed regions (a, b, l, p)	65	31.1
Generously allowed regions (~a, ~b, ~l, ~p)	11	5.3
Disallowed regions (XX)	9	4.3
Total residues analyzed	209	100

target prioritization [26, 27]. The most important finding of this research is that the superantigen-like protein SSL4 domain is found in FGIG_11154. The development of superantigen-like domains can be most effectively described in bacterial systems, which contribute to immune modulation through interaction with host immune receptors and the modulation of immune signaling pathways. This domain appearing in a protein of a trematode is an indication of possible evolutionary intersections into immune-interfering strategies. The immune modulation of long-term infections is a requirement in the helminths, and the interaction between proteins and the host immune factors is an important part of immune modulation. The identification of an SSL4-like domain thus provides an opportunity that FGIG_11154 might be involved in immune evasion or immune regulation in *F. gigantica* infection [28, 29]. The further predictions of the subcellular localization also endorse the possible multifunctionality of FGIG_11154. CELLO was able to predict extracellular localization, whereas PSORT II showed strong nuclear localization with high confidence [30].

This type of contrasting prediction does not occur in isolation from proteins that are involved in host-parasite interactions, especially those that can be shuttled between cellular compartments or that can have an intracellular as well as extracellular effect. The existence of putative nuclear localization signals indicated that FGIG_11154 could be involved in interaction with nucleic acids or nuclear proteins, with the possibility that it affects transcriptional or regulatory activities. Alternatively, extracellular localization can be indicative of secretion via non-classical pathways, which is often used by helminths to deliver immunomodulatory molecules into the host environment [31, 32]. It is also evidenced by the lack of a classical signal peptide and transmembrane helices that FGIG_11154 is a soluble protein that can employ unusual secretion pathways. The non-classical secretion is more and more characterized as a typical feature of parasite effector proteins that enable them to circumvent conventional secretory routes but retain the ability to have immunological effects on the host. This characteristic is consistent with the extracellular activity of FGIG_11154 as predicted and increases its suitability as a host-interacting molecule [33, 34].

The results of the analyses of the secondary and tertiary structures demonstrated an intrinsic disorder protein architecture characterized by random coils and low AlphaFold confidence scores. The intrinsically disordered proteins (IDPs) are known to interact with a variety of binding partners, as well as to change structure to suit various functional situations. IDPs are often related to immune evasion, signaling, and regulation in parasitic organisms. The chaotic appearance of FGIG_11154 is thus in agreement with a role in dynamic interaction between hosts and parasites, and not enzyme or structural activity [35, 36]. Although there was low structural confidence, validation analysis proved the overall plausibility of the model of the predicted three-dimensional. The reasonable statistics of acceptable ProSA-web Z-score and reasonable values of Ramachandran plot indicate that the model can be used to infer the functional aspect, especially when analyzed in terms of intrinsic disorder. Disordered regions can also be the cause of non-optimal stereochemistry, a typical characteristic of flexible and interaction-prone proteins [37].

Together, the findings of the current paper indicate that FGIG_11154 is a versatile protein and lineage-specific, which possesses the characteristics of immunomodulatory effectors in parasitic helminths. Its predicted SSL4 domain, hydrophilic character, non-classical secretion capacity, and intrinsic disorganization are all indications of a host immune interaction. Although these predictions could not be confirmed by experimental validation, the current *in silico* analysis offers a solid ground for future functional, immunological, and vaccine-related investigations. Notably, the discovery of these proteins increases our knowledge of *F. gigantica* pathobiology and is part of the overall endeavor of discovering new targets to regulate fascioliasis [38, 39].

In future research, experimental evidence about the functions of FGIG_11154 predicted should be validated. Profiles of expression during different developmental stages, localization experiments with immunofluorescence, and interaction experiments with host immune components would provide important functional information. Also, its immunogenicity and epitope mapping may be used to determine its vaccine candidacy. Experimental

methods like circular dichroism or nuclear magnetic resonance spectroscopy can use structural refinement to further understand the role of intrinsic disorder in its functionality. Overall, this study provides a solid computational basis for the functional annotation of hypothetical proteins in *F. gigantica* and highlights the significance of combining in silico methods to give priority to biologically important targets. The lessons of this study would help to gain deeper insights into the biology of parasites and would assist in further developing effective mechanisms of controlling and managing fascioliasis.

Conclusion

This research paper provides an *in silico* characterization of the hypothetical protein FGIG 11154 of *F. gigantica* and offers new data on the possible structural and functional functions of this protein. Combining sequence-based analysis, physicochemical profiling, identification of domains, prediction of subcellular localization, and structural modeling, this work will fill a critical gap in the functional annotation of *F. gigantica* hypothetical proteins. These results indicated that FGIG_11154 is a hydrophilic, soluble, and largely disordered protein with an intermediate level of evolutionary conservation across species of Fasciola. The structure of a superantigen-like domain of protein SSL4 has suggested the potential role in host immune modulation, subcellular localization predictions, and the lack of a classical secretion signal suggested multifunctionality and possible non-classical secretion. Intrinsically disordered regions supplement the hypothesis that FGIG_11154 is a flexible interaction protein, as opposed to a rigid enzyme or structural part, that is engaged in host-parasite communication.

Conflict of Interest

The authors had no conflicts of interest to disclose.

References

[1] Lalor R, Cwiklinski K, Calvani NED, Dorey A, Hamon S, Corrales JL, et al. Pathogenicity and virulence of the liver flukes *Fasciola hepatica* and *Fasciola gigantica* that cause the zoonosis fasciolosis. Virulence 2021; 12(1):2839-67.

[2] Ogbaje L, Malann Y, Azare B, Njoku M, Jegede O. Economic impact of high burden of fascioliasis in FCT caused by liver damage. Direct Res J Agric Food Sci 2025; 13(2):37-41.

[3] Castañeda S, Ramírez JD. From genes to worms: a deep dive into helminth omics. In: Recent advances in parasitomics: implications for parasite and vector research. 2025; p. 207-46.

[4] Konhar R, Das KC, Nongrum A, Samal RR, Sarangi SK, Biswal DK. *In silico* design of an epitope-based vaccine ensemble for fasciolopsiasis. Front Genet 2025; 15:1451853.

[5] Abdallah MM, Suliman RAI, Ahmed YT, Yahia M. In silico analysis of hypothetical proteins in *Pseudomonas aeruginosa* PAC1: structural and functional insights. J Genet Eng Biotechnol 2025; 23(3):100515.

[6] Graczyk TK. Immunobiology of trematodes in vertebrate hosts. In: Advances in trematode biology. CRC Press; 2024. p. 383-404.

[7] Adukpo S. The dual nature of host–parasite interactions: exploring protozoans and helminths in symbiosis and pathogenesis; 2025.

[8] Noli Truant S, Redolfi DM, Sarratea MB, Malchiodi EL, Fernández MM. Superantigens, a paradox of the immune response. Toxins 2022; 14(11):800.

[9] Hasan MR, Tithi SR, Khatun B, Chowdhury UR, Mahmud A. *In silico* characterization of hypothetical protein AAFM48_25015 from *Burkholderia pseudomallei*: structural and functional analysis. J Biosci Public Health 2025;1(3):28-44.

[10] Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy server. In: The proteomics protocols handbook. Springer; 2005; p. 571-607.

[11] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990; 215(3):403-10.

[12] Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics 2014; 30(9):1236-40.

[13] Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al. CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Res 2019; 48(D1):D265-D8.

[14] Yu CS, Chen YC, Lu CH, Hwang JK. Prediction of protein subcellular localization. Proteins 2006; 64(3):643-51.

[15] Nakai K, Kanehisa M. A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics 1992; 14(4):897-911.

[16] Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol 2022; 40(7):1023-5.

[17] Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 2001; 305(3):567-80.

[18] Garnier J, Osguthorpe DJ, Robson B. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J Mol Biol 1978; 120(1):97-120.

[19] Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 1999; 292(2):195-202.

[20] 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.
[21] Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 2007; 35(Suppl 2):W407-W10.
[22] Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993; 26(2):283-91.
[23] Asad MA, Shorna SA, Mizan M, Nath RD, Saikat ASM, Uddin ME. Bioinformatics approaches for the molecular characterization and structural elucidation of a hypothetical protein of *Aedes albopictus*. Eng Proc 2024; 67(1):14.
[24] Suleman M, Yaseen AR, Naeem A, Alzahrani FM, Alzahrani KJ, Alsharif KF, et al. From sequence to solution: computational design of a multi-epitope vaccine candidate against *Francisella tularensis*. Probiotics Antimicrob Proteins 2025; 1-29, 10.1007/s12602-025-10862-2.
[25] Nagarajan G. Unraveling evolutionary insights into AVT peptide conservation and antimicrobial motif prediction across taxa. Int J Mol Sci 2025; 26(16):8026.
[26] Cwiklinski K, Dalton JP. Exploiting comparative omics to understand the pathogenic and virulence-associated protease: anti-protease relationships in *Fasciola hepatica* and *Fasciola gigantica*. Genes 2022; 13(10):1854.
[27] Farfán-López M, Espinoza-Culupú A, García-de-la-Guarda R, Serral F, Sosa E, Palomino MM, et al. Prioritisation of potential drug targets against *Bartonella bacilliformis* by an integrative *in silico* approach. Mem Inst Oswaldo Cruz 2020; 115:e200184.
[28] Naujack P. Insights into the molecular interaction between Schistosoma mansoni egg-secreted glycoprotein IPSE/alpha-1 and the host cell. Justus-Liebig-Universität Gießen; 2024.
[29] Gazzinelli-Guimaraes PH, Nutman TB. Helminth parasites and immune regulation. F1000Res 2018; 7:F1000 Faculty Rev-685.
[30] Xiong E, Zheng C, Wu X, Wang W. Protein subcellular location: the gap between prediction and experimentation. Plant Mol Biol Rep 2016; 34(1):52-61.
[31] Cuesta-Astroz Y, Santos A, Oliveira G, Jensen LJ. Analysis of predicted host–parasite interactomes reveals commonalities and specificities related to parasitic lifestyle and tissue tropism. Front Immunol 2019; 10:212.
[32] Soleymani N, Sadr S, Santucciu C, Dianaty S, Lotfalizadeh N, Hajjafari A, et al. Unveiling novel insights in helminth proteomics: advancements, applications and implications. Biologics 2024; 4(3):314-44.
[33] Khan MAA, Ami JQ, Faisal K, Chowdhury R, Ghosh P, Hossain F, et al. An immunoinformatic approach driven by experimental proteomics: *in silico* design of a subunit vaccine targeting secretory proteins of *Leishmania donovani*. Parasit Vectors 2020;13(1):196.
[34] Mladineo I, Rončević T, Gerdol M, Tossi A. Helminthic host defense peptides: using the parasite to defend the host. Trends Parasitol 2023; 39(5):345-57.
[35] Zanotti G. Intrinsic disorder and flexibility in proteins: a challenge for structural biology and drug design. Crystallogr Rev 2023; 29(2):48-75.
[36] Tompa P, Schad E, Tantos A, Kalmar L. Intrinsically disordered proteins: emerging interaction specialists. Curr Opin Struct Biol 2015; 35:49-59.
[37] Patra P, Ghosh P, Patra BC, Bhattacharya M. Biocomputational analysis and in silico characterization of an angiogenic protein (RNase5) in zebrafish (Danio rerio). Int J Pept Res Ther 2020; 26(4):1687-97.
[38] Hernandez Alvarez L, Naranjo Feliciano D, Hernandez Gonzalez JE, de Oliveira Soares R, Barreto Gomes DE, Pascutti PG. Insights into the interactions of Fasciola hepatica cathepsin L3 with substrates and inhibitors through *in silico* approaches. PLoS Negl Trop Dis 2015; 9(5):e0003759.
[39] Akıl M, Aykur M, Karakavuk M, Can H, Döşkaya M. Construction of a multiepitope vaccine candidate against *Fasciola hepatica* using immunogenic excretory/secretory antigens. Expert Rev Vaccines 2022; 21(7):993-1006.