ARTICLE

Computational docking analysis of *Schistosoma haematobium* egg derived proteins: Implications of IPSE/α-1 and serpin in bladder cancer development

Arthur F. de Vera, Leila Sophia A. Casano, Alyssa Nicole M. Elemia, Kelsey Beatrice F. Faderanga, Adrian Delfin L. De Jesus, A-Jay M. Florentino, Jennuel Herbert B. Frias, and Mary Rose F. Lirio*

Institute of Health Sciences and Nursing, Far Eastern University – Manila, Nicanor Reyes Street, Sampaloc, Manila, 1008 Metro Manila

ABSTRACT

Schistosoma haematobium is the primary cause of urogenital schistosomiasis worldwide. If left untreated, the disease can lead to chronic inflammation and *Schistosoma*-induced bladder cancer. This study explores the carcinogenic potential of two proteins secreted by *S. haematobium* eggs, IPSE/α-1 and serpin. In silico methods were used to examine their physicochemical properties and interactions with human host proteins. Results show that both proteins exhibit high stability, extended half-lives, and binding affinities with cancer-related host proteins. HADDOCK analysis revealed a negative binding score for IPSE/α-1 with EGFR, suggesting it could potentially induce tissue hyperplasia, while serpin demonstrated stronger binding to p53, potentially inhibiting its tetramerization and rendering the protein nonfunctional. These findings provide valuable insights into the molecular mechanisms driving *Schistosoma*-induced squamous cell carcinoma (SCC) of the bladder and highlight potential targets for further research in parasite-associated cancer development. Moreover, these findings offer new opportunities for cancer prevention and lay the groundwork for developing targeted therapies and immunotherapies against parasite-associated cancers. Integrating these insights into therapeutic strategies could significantly enhance efforts to prevent and treat *Schistosoma*-induced malignancies.

INTRODUCTION

Schistosoma haematobium is a blood fluke parasite commonly associated with urogenital schistosomiasis, kidney damage, and bladder fibrosis (International Agency for Research on Cancer, 2012). People acquire infection through direct contact with freshwater containing the parasite's infectious stage, the cercariae. After penetrating human skin, the larvae migrate through the bloodstream and mature into adult flukes, which reside in the venous plexus of the bladder. There, adult females deposit eggs, some of which become trapped in the bladder wall, triggering chronic inflammation and fibrosis. Over time, this pathological process contributes to severe complications, including bladder cancer.

*Corresponding author
Email Address: mlirio@feu.edu.ph
Date received: 23 September 2025
Dates revised: 08 November 2025, 24 November 2025
Date accepted: 08 December 2025
DOI: https://doi.org/10.54645/2025182GZL-51

KEYWORDS

bladder cancer, *in silico*, IPSE/α-1, oncology, parasitology, *Schistosoma haematobium*, serpin

Schistosomiasis remains a major public health issue that affects nearly 240 million people worldwide, with about 90% of those at risk living in Sub-Saharan Africa (Nazareth et al., 2022), and it causes approximately 200,000 deaths annually (van Hoogstraten et al., 2023). Egypt once had the highest prevalence, with more than 40% of the population affected until the 1980s. During that time, bladder cancer became the most common cancer-related cause of death in men aged 20–44 and the second most common in women. Historical data showed that schistosomiasis increases the risk of bladder cancer development by 1.72 times (95% CI 1.0-2.9), and 16% of bladder cancer cases were linked to *S. haematobium* infection. Mass treatment with praziquantel and improvements in water sanitation reduced the prevalence of schistosomiasis to 1-2%, leading to a corresponding decline in cancer incidence (Salem et al., 2011).

To better understand how schistosomiasis progresses to bladder cancer, researchers have focused on the molecular mechanisms by which the parasite induces inflammation and tissue remodeling in the bladder. Proteomic studies have identified several egg-secreted proteins that not only help the parasite evade the host immune response but also promote pathological changes such as hyperplasia and angiogenesis. Among these secreted proteins are the IL-4-inducing principle of the *Schistosome* egg (IPSE/α-1) and serine protease inhibitor (serpin).

Knuhr et al. (2018) reported that IPSE/α-1 stimulates basophils to release interleukin-4 (IL-4), directing naïve T helper cells toward a T helper 2 (Th2) response. This promotes an anti-inflammatory mechanism that is less effective in clearing the parasite than the pro-inflammatory response of Th1. In contrast, Farling (2017) emphasized that chronic inflammation promotes squamous cell carcinoma (SCC), highlighting a paradox in schistosomal infection leading to malignancy. Furthermore, Mbanefo et al. (2020) demonstrated that IPSE/α-1, when injected into mouse bladders, could induce angiogenesis and urothelial hyperplasia. This study suggests that IPSE/α-1 may directly contribute to carcinogenesis beyond its immunomodulatory role.

On the other hand, serpin acts as a "molecular mousetrap" that undergoes a conformational change to irreversibly bind and inactivate host serine proteases (Huntington, 2001). "Trapping" these proteases disrupts the proteolytic cascades necessary for normal immune function and tissue remodeling (Chan et al., 2024). Cagnin et al. (2024) reported that overexpression of SerpinB3, a homologous protein, inhibits apoptotic pathways in the human bladder. Although the precise molecular mechanisms remain unclear, this dysregulation of immune surveillance and apoptosis may contribute to tumor progression.

These proteins are of particular interest in oncologic studies due to their ability to manipulate the host microenvironment. IPSE/α-1 promotes angiogenesis and urothelial hyperplasia, while serpin interferes with apoptosis. Their combined effects suggest a direct role in shaping a microenvironment conducive to carcinogenesis (Kontomanolis et al., 2020). While previous research has begun to explore the molecular activities of these proteins, the exact mechanisms by which they modulate host cellular pathways and potentially contribute to carcinogenesis remain to be clarified.

This study addresses this gap through an *in silico* investigation of the structural and functional properties of IPSE/α-1 and serpin. Computational analysis was used to predict physicochemical properties and structural regions that may participate in host–parasite interactions promoting urothelial carcinogenesis. Amino acid sequences of the parasite proteins were retrieved from the National Center for Biotechnology Information (NCBI) – GenBank and UniProt databases, and 3D models were constructed through homology modeling and refinement to identify regions that may mediate these carcinogenic effects.

MATERIALS AND METHODS

To characterize the carcinogenic potential of the parasite proteins, different *in silico* tools were utilized. Fig. 1 outlines how the study characterized *S. haematobium* egg proteins and assessed their potential roles in SCC of the bladder development.

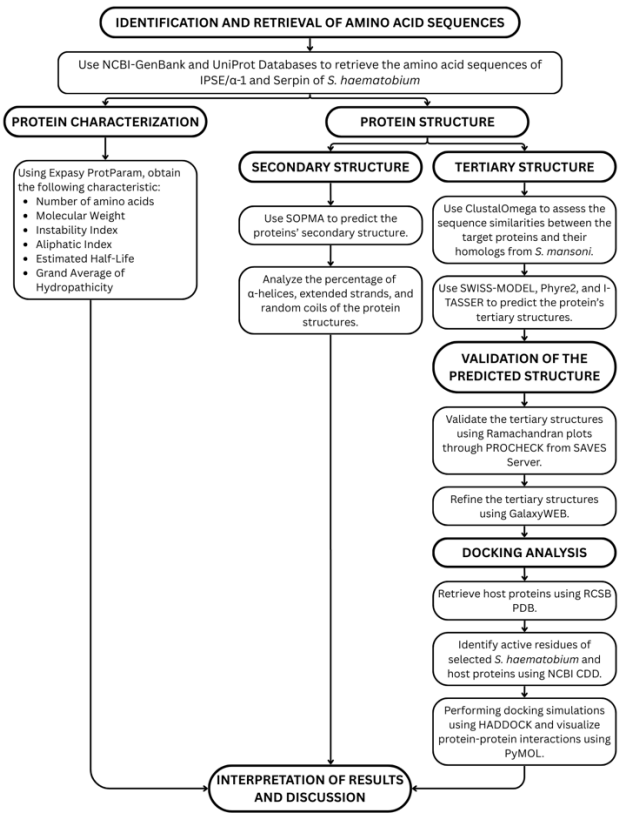


Figure 1: Flowchart of the study. Workflow of the *in silico* analysis conducted to characterize *Schistosoma haematobium*–secreted proteins IPSE/α-1 and serpin and evaluate their interactions with human host proteins. The workflow outlines the retrieval of amino acid sequences, protein characterization, secondary and tertiary structure prediction, model validation, docking analysis, and interpretation of results.

Amino Acid Sequence Retrieval

The amino acid sequences of IPSE/α-1 and serpin were retrieved from NCBI-GenBank Release 257 (https://www.ncbi.nlm.nih.gov/genbank/) (Sayers et al., 2021). The keywords used to select the targeted sequences were "IPSE/α-1", "Serpin", and "*Schistosoma haematobium*". To ensure the accuracy and reliability of these sequences, they were cross-validated using UniProt Release 2024_06 (https://www.uniprot.org/) (Ahmad et al., 2025).

Previous studies aimed to characterize the physicochemical properties of proteins and their possible interactions with other proteins, including those by Alam et al. (2023) and Cunanan et al. (2023), utilized NCBI-GenBank to retrieve protein sequences and were cross-validated using UniProt database.

Determination of Physicochemical Properties

The physicochemical properties of IPSE/α-1 and serpin were analyzed through Expasy ProtParam (https://web.expasy.org/protparam/) (Gasteiger et al., 2005), which generated molecular parameters relevant to predicting their carcinogenic potential. This tool has been widely used to

characterize parasite-secreted proteins associated with carcinogenesis, such as the 31 secretory proteins from *Opistorchis viverrini* and 22 from *Clonorchis sinensis* linked to cholangiocarcinoma (Machicado et al., 2021), and the *Helicobacter pylori* protein HPF63-1454 in the development of gastric adenocarcinoma (Alam et al., 2023).

Determination of Protein Structure

For secondary structure determination, the sequences were uploaded to SOPMA v.2.16.0. (https://npsa-prabi.ibcp.fr/cgibin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) (Geourjon & Deléage, 1995) to analyze α-helices, extended strands, and random coils of the proteins.

For homology modeling, target protein sequences and their homologs from *Schistosoma mansoni* were subjected to sequence alignment using ClustalOmega (https://www.ebi.ac.uk/jdispatcher/msa/clustalo) (Sievers et al., 2011). After confirming the sequence identity of the proteins between the two different species, these sequences were uploaded to SWISSMODEL (https://swissmodel.expasy.org) (Guex et al., 2009; Waterhouse et al., 2018) to generate tertiary structures. Homologous structures from *S. mansoni* were used as templates in SWISS-MODEL because of their high overall sequence similarity to *S. haematobium*, ensuring accurate homology-based predictions (Wu et al., 2021). Alam, Saikat, and Uddin (2023) similarly used SOPMA and SWISSMODEL for predicting *H. pylori* protein structures.

To further enhance modeling reliability, two additional homology modeling tools were used: Phyre v.2.2. (www.sbg.bio.ic.ac.uk/~phyre2) (Powell et al., 2025) and I-TASSER v.5.2. (https://zhanggroup.org/I-TASSER/) (Yang Zhang, 2008; Roy et al., 2010). Each tool employs different algorithms and scoring methods, allowing cross-validation and reducing the risk of tool-specific bias or error.

The generated structures were downloaded as PDB files and validated using the PROCHECK tool on SAVES v.6.1. (https://saves.mbi.ucla.edu) (Laskowski et al., 1993). Models labeled as "good quality", with 85–90% of their residues in the favored regions of the Ramachandran plot (Park et al., 2023), were further refined using GalaxyWEB (https://galaxy.seoklab.org) (Seok et al., 2021). The top model was uploaded and refined with default settings, which included sidechain optimization and full model relaxation. This refinement improved both global and local structural quality, enhancing the model's physical accuracy and its ability to represent the actual protein (Heo et al., 2013).

Alam et al. (2023) also used PROCHECK on SAVES to validate *H. pylori* protein models, while Bhargavi et al. (2025) applied GalaxyWEB to refine Histone Deacetylase 1 (HDAC1) modeled by Modeller 10.6 software on their prediction of HDAC1 protein-ligand interactions linked to cancer.

Docking Analysis

Active residues of the modeled proteins were identified using the NCBI Conserved Domains Database (CDD) (https://www.ncbi.nlm.nih.gov/Structure/cdd) (Wang et al., 2023). Only domains involved in ligand binding and signal transduction were used to ensure biologically meaningful interactions.

Docking simulations were performed using HADDOCK v.2.4. (https://rascar.science.uu.nl/haddock2.4/) (Honorato et al., 2021; Honorato et al., 2024), where the parasite protein was tagged as the "ligand" and the host protein as the "receptor". Default clustering parameters were applied, and only the top-scoring model was analyzed. Its binding energy was used to estimate the stability of the predicted interactions. Key regions related to immune function or cell cycle control were visualized using PyMOL v.3.1.5.1. (https://www.pymol.org) (Schrodinger, 2015) to better understand their potential effects on host pathways.

The prediction of protein-protein interactions (PPIs) using HADDOCK has already been applied to investigate the potential mechanism of *H. pylori* in infecting host cells (Akcelik-Deveci, 2024). In this study, the outer inflammatory protein A (OipA) of *H. pylori* was sequenced from cultured *H. pylori* G27 strain through DNA extraction, then amplification, purification, and sequencing. The sequence was used to predict the interaction between the OipA protein and host cells. Using the HADDOCK tool, multiple cell membrane receptors were identified, explaining the pathogenicity of *H. pylori* OipA protein.

Statistical Analysis

To analyze the predictions from the various computational tools used in this study, the Grand Average of Hydropathicity (GRAVY), Ramachandran plotting, and HADDOCK scoring were applied. These statistical analyses were selected because they each assess key aspects of protein structure and interaction relevant to docking-based carcinogenesis research.

GRAVY measures the overall hydrophilic or hydrophobic nature of a protein. Proteins with negative GRAVY scores are more hydrophilic and may interact readily with host proteins in aqueous environments, potentially influencing carcinogenic pathways (Abdollahi et al., 2020). In this study, GRAVY was used to determine whether host proteins from aqueous or lipophilic compartments should be selected for analysis.

GalaxyWEB utilizes Ramachandran plotting to assess and refine the tertiary structures of the proteins that were generated by SWISS-MODEL, Phyre2, and I-TASSER. This validation step ensures the structural integrity of the models before docking (Park et al., 2023).

The HADDOCK score combines electrostatics, desolvation energy, Van der Waals, and restraint violations into a single energy-based metric. More negative HADDOCK scores indicate stronger and more stable binding conformations between parasite and host proteins (Yan et al., 2020).

This scoring provides insight into possible mechanisms of immune modulation and tumor promotion.

RESULTS

The amino acid sequences for both proteins were retrieved from NCBI-GenBank, where IPSE/α1 had a GenBank ID ATJ03502.1, while serpin had the GenBank ID AAA19730. In addition, the UniProt protein database was used to conduct cross-validation of these protein sequences. This ensured that the retrieved sequences are reliable and accurate for *in silico* analysis.

Protein Characterization

Table 1: Predicted Physicochemical Properties of IPSE/α-1 and Serpin. The amino acid sequences of IPSE/α-1 and serpin were retrieved from NCBI-GenBank, and their physicochemical properties were determined using the Expasy ProtParam tool.

Properties	IPSE/α-1	Serpin
Number of Amino Acids	134	406
Molecular Weight	15,072.42 Da	45,841.92 Da
Instability Index	35.54 (Stable)	34.56 (Stable)
Aliphatic Index	83.58	85.39
Estimated Half-Life	30 hours (mammalian reticulocytes, in vitro). >20 hours (yeast, in vivo). >10 hours (*Escherichia coli,* in vivo).	30 hours (mammalian reticulocytes, in vitro). >20 hours (yeast, in vivo). >10 hours (*Escherichia coli,* in vivo).
Grand Average of Hydropathicity (GRAVY)	-0.165 (hydrophilic)	-0.215 (hydrophilic)
IPSE = Interleukin-4-Inducing Principle of the *Schistosome* Egg		

As presented in Table 1, IPSE/α-1 is a smaller protein, consisting of 134 amino acids and weighing 15,072.42 Da, whereas serpin consists of 406 amino acids and has a molecular weight of 45,841.92 Da. The instability indices of IPSE/α-1 (35.54) and serpin (34.56) indicate that both proteins are stable under physiological conditions, and their aliphatic indices (83.58 and 85.39, respectively) suggest high thermostability. Both proteins exhibit estimated half-lives of approximately 30 hours in mammalian reticulocytes, over 20 hours in yeast, and over 10 hours in Escherichia coli, indicating that both are likely stable across different biological systems. Both proteins exhibit a negative Grand Average of Hydropathicity (GRAVY) score, indicating their hydrophilic nature. Serpin has a more hydrophilic characteristic, scoring -0.215, while IPSE/α-1 scored -0.165.

Table 2: Predicted Structural Characteristics of IPSE/α-1 and Serpin. Using the amino acid sequences retrieved from NCBI GenBank, the secondary structural elements of the target proteins were predicted using the SOPMA tool.

Secondary Structure	IPSE/α-1	Serpin
α-Helices	18.65%	39.41%
Extended Strands	30.60%	15.52%
Random Coils	50.75%	45.07%
IPSE = Interleukin-4-Inducing Principle of the *Schistosome* Egg		

SOPMA secondary structure analysis revealed that IPSE/α-1 comprises 18.65% α-helices, 30.60% extended strands, and 50.75% random coils. In contrast, serpin exhibits 39.41% α-helices, 15.52% extended strands, and 45.07% random coils.

Homology Modeling

Sequence alignment using ClustalOmega showed that IPSE/α-1 homologs from *S. haematobium* and *S. mansoni* share 85.83% sequence identity (Figure 2.a), while their serpin homologs share 76.11% (Figure 2.b). Although the two species exhibit high proteomic similarity (82–92% identity on average) (Wu et al., 2021), minor differences in amino acid sequences were still observed in several regions. These variations may influence species-specific folding, surface charge distribution, or protein–protein interaction affinity (E Lohning et al., 2017). Such differences could potentially change the binding affinity of *S. haematobium* proteins to certain host targets with which *S. mansoni* may interact weakly or strongly, thereby contributing to species-specific pathogenic mechanisms.

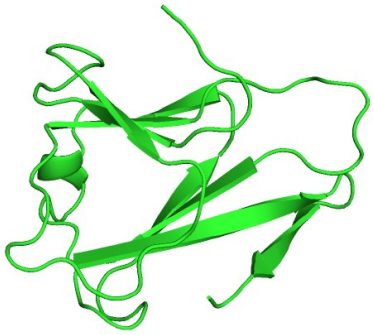
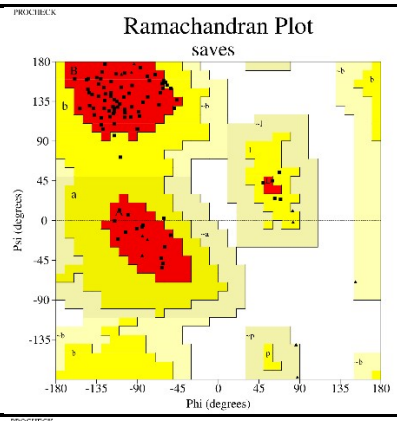
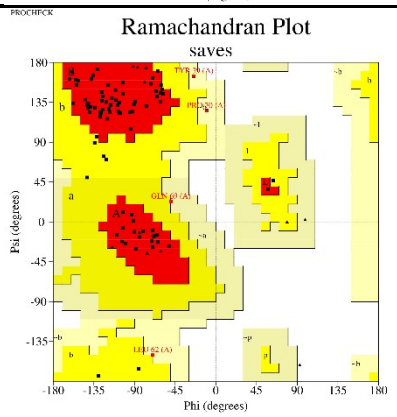
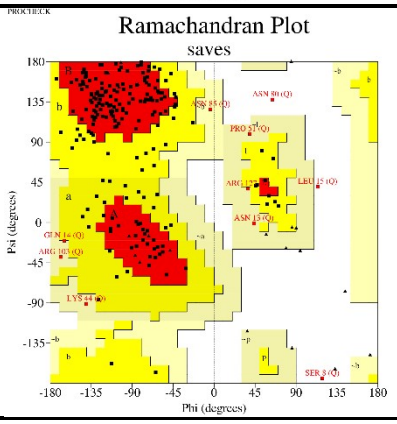
Despite these minor differences, the conserved residues forming the protein active-site motifs were retained, indicating that the overall tertiary structure and functional domains are likely preserved across species (Meyer et al., 2011; Ambadapadi et al., 2016). Therefore, *S. mansoni* homologs serve as reliable structural templates for *S. haematobium* proteins.

However, to ensure that the subtle conformational features unique to *S. haematobium* were accurately represented, homology modeling was performed. This approach allowed refinement of the predicted 3D structures to capture species-specific folding patterns, side-chain orientations, and potential local flexibility. The resulting models thus provide a more physiologically relevant representation of *S. haematobium* IPSE/α-1 and serpin, suitable for subsequent docking and interaction analyses with host proteins.

Figure 2: Pairwise sequence alignment of *S. haematobium* and *S. mansoni* IPSE/α-1 (a) and serpin (b) homologs generated using ClustalOmega. Identical residues are marked with "*", strongly similar residues with ":", weakly similar residues with ".", and non-similar residues are unmarked. The alignments show 85.83% sequence identity for IPSE/α-1 and 76.11% for serpin.

As protein models generated through homology modeling are prone to residual errors, validation is essential to assess their reliability in studying protein-protein interactions. PROCHECK generates a Ramachandran plot that quantifies the percentage of residues in the most favored regions, which represent sterically favorable conformations.

Table 3: Raw Protein Structures of IPSE/α-1 from Homology Modeling. Using the amino acid sequence of IPSE/α-1 retrieved from NCBI GenBank, three structural models were generated using SWISS-MODEL, Phyre2, and I-TASSER. The models were evaluated with PROCHECK to determine the percentage of residues in *favored regions*. In the Ramachandran plot, red areas indicate *most favored* dihedral angles, light yellow areas represent *generously allowed* angles, and white areas correspond to *disallowed regions*. Amino acids shown in red font lie outside favored regions and reduce the overall *favored region score*. The best model generated will be refined and used for subsequent analyses.

IPSE/α-1			
HOMOLOGY MODELLING TOOL	**Tertiary Structure**	**Ramachandran Plot**	**Favored Region**
SWISS-MODEL			89.4%*
Phyre2			86.4%
I-TASSER			68.5%

IPSE = Interleukin-4-Inducing Principle of the *Schistosome* Egg

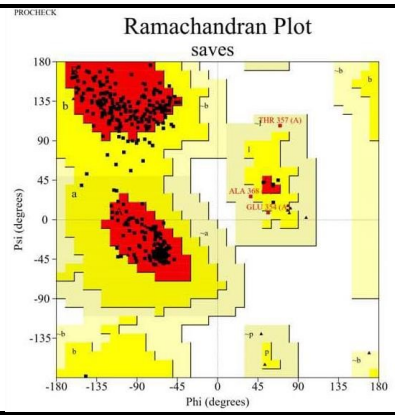
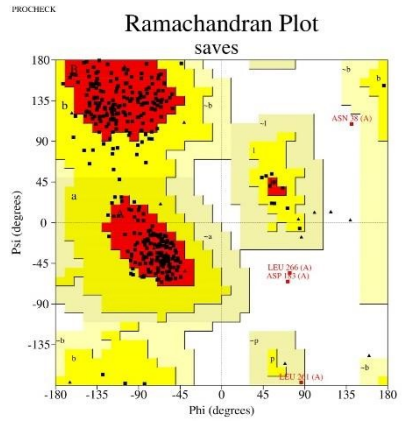
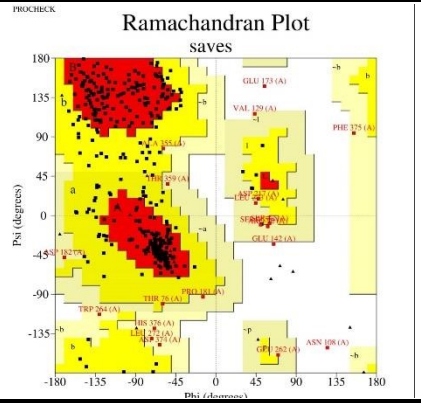
Red = most favored, Yellow = allowed, Light yellow = generously allowed, White = disallowed regions of φ–ψ angles.

Amino acids shown in red font lie outside favored regions and decrease the model's favored region score.

** Best model predicted and will be refined through GalaxyWEB.*

The tertiary structure IPSE/α-1 predicted by SWISS-MODEL garnered 89.4% favored region, the highest score among the three tools, indicating the best structure predicted. Phyre2 and I-TASSER had favored regions of 86.4% and 68.5%, respectively (Table 3).

Table 4: Raw Protein Structures of Serpin from Homology Modeling. Using the amino acid sequence of serpin retrieved from NCBI GenBank, three structural models were generated using SWISS-MODEL, Phyre2, and I-TASSER. The models were evaluated with PROCHECK to determine the percentage of residues in *favored regions*. In the Ramachandran plot, red areas indicate *most favored* dihedral angles, light yellow areas represent *generously allowed* angles, and white areas correspond to *disallowed regions*. Amino acids shown in red font lie outside favored regions and reduce the overall *favored region score*. The best model generated will be refined and used for subsequent analyses.

Serpin			
HOMOLOGY MODELING TOOL	**Tertiary Structure**	**Ramachandran Plot**	**Favored Region**
SWISS-MODEL			90.9%*
Phyre2			86.5%
I-TASSER			71.5%

Red = most favored, Yellow = allowed, Light yellow = generously allowed, White = disallowed regions of φ–ψ angles.

Amino acids shown in red font lie outside favored regions and decrease the model's favored region score.

** Best model predicted and will be refined through GalaxyWEB.*

After generating the tertiary structure of IPSE/α-1, the same methodology was done to model serpin. PROCHECK assessed that the model generated by SWISS-MODEL has a 90.9% favored region, followed by Phyre2 (86.5%) and I-TASSER (71.5%) (Table 4).

The models for both proteins generated by SWISS-MODEL were subjected to further refinement through GalaxyWEB (Table 5). By refining their conformations, GalaxyWEB helps ensure that the models used more closely reflect the proteins' conformations in the body and are reliable for docking analyses and predictions. According to Cunanan et al. (2023), a good model should have at least 90% favored region.

GalaxyWEB created five refined models, and the best model was selected based on their favored region (reported as Rama Favored) and MolProbity Scores, with priority given to the former.

For IPSE/α-1, model 1 scored the highest favored region of 99.0%, while also achieving the lowest MolProbity score of 1.519. Meanwhile, model 1 of serpin also had the highest favored region of 99.0% but had the second lowest MolProbity score of 1.695 (Table 6).

Table 5: Summary of GalaxyWEB Refinement. The models generated by SWISS-MODEL were uploaded to GalaxyWEB for structural quality refinement. The tool assessed the model's *MolProbity score*, which indicates overall structure quality, where lower values denote better geometry. It also evaluated the *Favored Region (reported as Rama favored)*, representing the percentage of residues in favored Ramachandran regions. The favored region percentage is prioritized over the MolProbity score, as it better reflects stereochemical reliability.

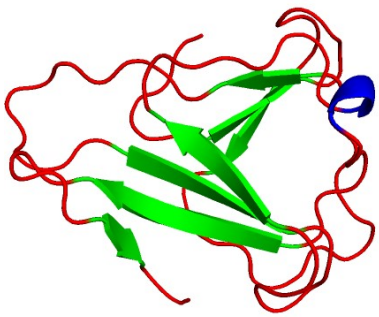
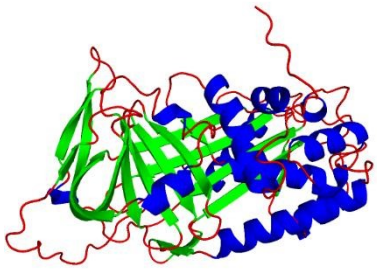
	IPSE/α-1		**Serpin**	
Model	MolProbity	Favored Region (Rama Favored)	MolProbity	Favored Region (Rama Favored)
Initial	2.241	95.1	1.240	95.0
MODEL 1	1.519*	99.0*	1.695*	99.0*
MODEL 2	1.713	99.0	1.646	98.5
MODEL 3	1.620	99.0	1.693	98.8
MODEL 4	1.696	99.0	1.641	98.8
MODEL 5	1.565	98.0	1.654	98.5

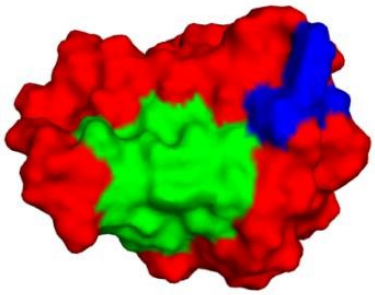
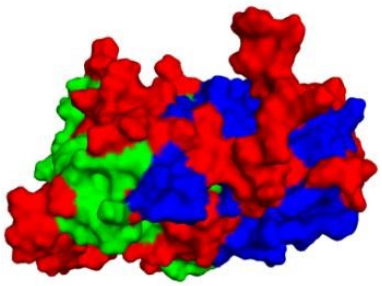
IPSE = Interleukin-4-Inducing Principle of the *Schistosome* Egg

MolProbity = overall structure quality score (lower values denote better geometry).
Favored Region (Rama Favored) = number of residues in favored Ramachandran regions in percentage (%).

** Best-refined model used for subsequent analysis.*

Table 6: Refined Models of the Proteins. The top models generated by GalaxyWEB, refined from SWISS-MODEL outputs, are presented. The ribbon diagrams depict the secondary structures, where *blue* indicates α-helices, *green* indicates extended strands, and *red* represents random coils. The surface models show the proteins' three-dimensional conformation in a biological context. The *favored region* indicates the percentage of residues in the most favorable conformations on a Ramachandran plot, while the *MolProbity score* reflects the stereochemical quality of the protein structure by assessing steric clashes.

Structural Parameters	**IPSE/α-1**	**Serpin**
Ribbon		

Surface		
Favored Region (Rama Favored)	99.0%	99.0%
MolProbity	1.519	1.695

IPSE = Interleukin-4-Inducing Principle of the *Schistosome* Egg

Blue = α-helices; Green = extended-strands; Red = random coils.

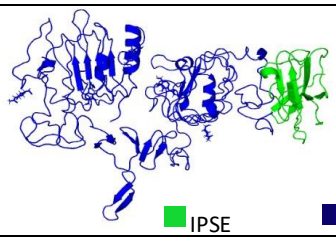
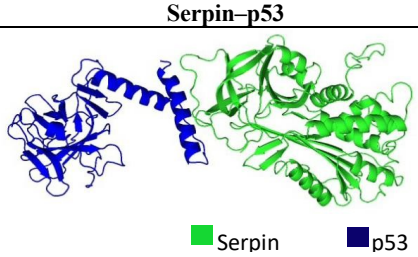
Favored Region (Rama Favored) = corresponds to the percentage of residues in the most favorable φ–ψ angle conformations in a Ramachandran plot (Park et al., 2023).
MolProbity = measures stereochemical quality by evaluating steric clashes, bond geometry, and side-chain conformations, where lower scores indicate more favorable overall structure (Williams et al., 2017; Alhumaid & Tawfik, 2024).

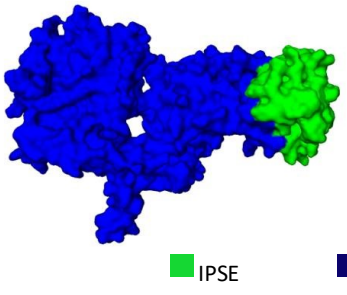
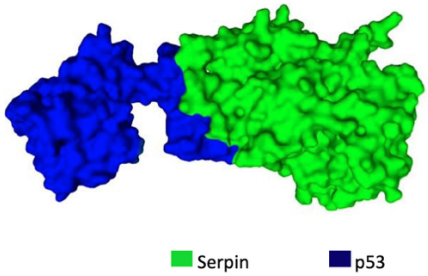
Docking Analysis

IPSE/α-1 contains active residues at positions 26–112 amino acids, corresponding to the IL-4-inducing immunoglobulin-binding domain. While serpin protein exhibits active residues at positions 23–402 amino acids, corresponding to the reactive center loop (RCL) of serpins.

Host proteins selection was guided by the physicochemical properties of the parasite's proteins and by supporting literature. Both IPSE/α-1 and serpin exhibit hydrophilic properties, suggesting their preference for aqueous environments such as the extracellular space and nucleoplasm. This is consistent with previous findings of Pennington et al. (2017) stating that homologs of IPSE/α-1 from *S. mansoni* are localized at the extracellular space of the host, while Janciauskiene et al. (2024) reported nuclear localization for certain serpins, including SerpinB5. Since protein homologs often retain conserved localization and functional patterns across species, these data collectively support the predicted cellular distribution of IPSE/α-1 and serpin from *S. haematobium*. These are used as the basis of the rational selection of the epidermal growth factor receptor (EGFR) domain IV located between residues 505–637 (PDB ID: 1IVO) to simulate interaction with IPSE/α-1, while using p53 and its tetramerization domain located at residues 319–357 as binding partner with serpin. Selection of these domains will further be discussed in the succeeding section.

Table 7: Docking Analysis Result. Docking analysis between the parasite and host proteins was performed using HADDOCK. The refined parasite protein models from GalaxyWEB were docked with host protein structures retrieved from the RCSB PDB. Active residues were identified using NCBI CDD, and the resulting complexes were visualized using PyMOL. The *HADDOCK score* indicates the overall binding efficiency, where more negative values denote stronger interactions. *Electrostatic, desolvation,* and *van der Waals energies* represent the individual energetic contributions to complex formation, while the *buried surface area (Å²)* corresponds to the solvent-inaccessible region upon binding, with larger values indicating stronger interactions.

	IPSE-EGFR	**Serpin–p53**
Ribbon		

Surface		
HADDOCK SCORE	-49.1 ± 10.5	- 143 ± 5.2
Electrostatic Energy	-229.0 ± 63.4kcal/mol	-404.6 ± 43.8 kcal/mol
Desolvation Energy	0.0 ± 2.9kcal/mol	14.1 ± 4.9 kcal/mol
Van der Waals Energy	-66.2 ± 11.7kcal/mol	-67.6 ± 6.8 kcal/mol
Buried Surface Area	1834.2 ± 57.4 Å^2	2348.6 ± 32.7 Å^2
Potential Molecular Mechanism	EGFR's Intramolecular Tether Disruptor	P53 Tetramerization Inhibitor

IPSE-EGFR = Interleukin-4-Inducing Principle of the *Schistosome* Egg – Epidermal Growth Factor Receptor complex

Serpin–p53 = Serpin – p53 complex

HADDOCK Score – A composite score derived by combining various energy terms that describe the interaction between the host and parasite proteins. It is automatically calculated by HADDOCK based on its docking algorithm. A more negative score indicates a stronger and more efficient predicted binding.

Electrostatic Energy (kcal/mol) – Represents the energy contribution from interactions between oppositely charged amino acid side chains at the protein interface. More negative values indicate stronger electrostatic attraction and improved binding stability.

Desolvation Energy (kcal/mol) – Reflects the energy change associated with removing water molecules from the protein surfaces to enable closer contact between the binding partners. Negative values contribute favorably to the interaction.

Van der Waals Energy (kcal/mol) – Accounts for non-covalent attractive and repulsive forces between atoms of the interacting proteins. More negative values suggest better steric complementarity and physical compatibility at the binding interface.

Buried Surface Area (Å²) – Refers to the portion of the protein surfaces that becomes inaccessible to solvent after complex formation. A larger buried surface area generally corresponds to stronger and more stable protein–protein interactions.

The docking analysis between IPSE/α-1 and the EGF receptor domain of EGFR, as well as serpin and the tetramerization domain of p53, revealed potential interactions inside the host, as reflected in the HADDOCK results (Table 7). The IPSE/α-1-EGFR complex yielded a HADDOCK score of -49.1 ± 10.5, with electrostatic energy of -229.0 ± 63.4 kcal/mol, desolvation energy of 0.0 ± 2.9 kcal/mol, van der Waals energy of -66.2 ± 11.7 kcal/mol, and a buried surface area (BSA) of 1834.2 ± 57.4 Å². While serpin–p53 complex had a HADDOCK score of -143 ± 5.2, alongside an electrostatic energy of -404.6 ± 43.8 kcal/mol, desolvation energy of -14.1 ± 4.9 kcal/mol, van der Waals energy of -67.6 ± 6.8 kcal/mol, and a BSA of 2348.6 ± 32.7 Å².

DISCUSSION

Immunogenic Potential and Persistence of *Schistosomal* Proteins

The physicochemical properties of IPSE/α-1 (MW: 15,072.42 Da) and serpin (MW: 45,841.92 Da) indicate that both proteins are large enough to interact with host proteins and elicit a response. Proteins with molecular weight greater than 10,000 Da are more likely to elicit a host response (Miller & Stevens, 2021) due to their structural complexity and ability to present multiple epitopes for host proteins to interact with (Pedroza-Escobar et al., 2023).

Furthermore, the instability indices of IPSE/α-1 (35.53) and serpin (34.56) indicate that both proteins are predicted to be stable under physiological conditions, as values below 40 generally denote protein stability (Ocampo et al., 2024). Their high aliphatic indices (83.58 for IPSE/α-1 and 85.39 for serpin) are greater than 80, which indicates strong thermostability. Higher proportion of aliphatic side chains: alanine, valine, leucine, and isoleucine, enhances hydrophobic interactions within the protein core, helping maintain structural integrity over a wide temperature range (Ahmed et al., 2022). Both proteins also exhibit similar half-lives, lasting about 30 hours in mammalian reticulocytes, over 20 hours in yeast, and more than 10 hours in *Escherichia coli*. These characteristics suggest that IPSE/α-1 and serpin may persist within the host by resisting degradation and evading immune clearance, possibly allowing prolonged interaction with host tissues and promoting sustained inflammation and tissue damage.

Docking Analysis

Both IPSE/α-1 and serpin have negative GRAVY scores, suggesting their hydrophilicity and likely solubility in aqueous environments. This indicates that these proteins function in the extracellular space and nucleoplasm, enhancing their potential to interact with host proteins in the same compartments. Specifically, the extracellular domains of EGFR and the nucleoplasmic localization of p53 are both exposed to aqueous environments. This supports the rationale for selecting these host proteins as docking partners for IPSE/α-1 and serpin, respectively (Jernigan, 2022; Darmawati et al., 2022).

Beyond solubility and localization, the functional domains and secondary structure of these proteins provide additional insights into their host-interacting potential. The IL4-inducing immunoglobulin-binding domain (residue 26–112) is a region conserved across *Schistosoma* species which adopts a βγ-crystallin

fold stabilized by disulfide bonds (Cys23–Cys26, Cys59–Cys93, and Cys111–Cys121). This domain mediates the binding of IPSE/α-1 with host protein, inducing cytokine production (Meyer et al., 2011). Given the protein's signaling capability, it may also initiate a receptor-mediated signaling cascade through interacting with the EGFR domain IV.

Meanwhile, the RCL domain of serpin (residue 23–402) is the structural feature responsible for trapping target proteases and modulating cell-regulatory processes (Ambadapadi et al., 2016). Although p53 is not a serine protease, this domain is still relevant in this analysis because several serpins, such as ovalbumin, possess this domain and demonstrate non-canonical and non-inhibitory functions that involve direct binding with non-protease proteins. Such interactions are consistent with the regulatory and anti-apoptotic roles observed in nuclear serpins (Cagnin et al., 2024; Janciauskiene et al., 2024).

The HADDOCK analysis revealed that IPSE/α-1 interacts with EGFR at lower affinity (-49 ± 10.5), whereas serpin forms a more stable complex with p53 (-143 ± 5.2). HADDOCK score combines multiple energy components: electrostatic, van der Waals, desolvation, and restrain energies, to estimate the overall binding favorability, with more negative scores reflecting stronger and more stable complexes (Saponaro et al., 2020).

A breakdown of the HADDOCK score provides further insight into the interaction profiles of these complexes. Electrostatic energy, which reflects charged complementarity, is less negative for the IPSE/α-1–EGFR complex (-229 ± 63.4 kcal/mol) but is strongly negative for serpin–p53 (-404.6 ± 4.9 kcal/mol), indicating greater charge-based stabilization in the latter (Almeida, 2021).

Desolvation energy, the energetic benefit of displacing water molecules during complex formation, is near neutral for IPSE/α-1–EGFR complex (0.0 ± 2.9 kcal/mol) but favorable for serpin–p53 complex formation (-14 ± 4.9 kcal/mol), suggesting water displacement benefits the serpin–p53 interaction more (Pathak et al., 2021).

Van der Waals energy represents weak interactions between atoms that are close together. These forces depend on the ability of protein surfaces to fit snugly, where more negative values indicate a better physical fit between protein surfaces (Alrosan et al., 2022). Both complexes show almost similar van der Waals energies (IPSE/α-1–EGFR: -66.2 ± 11.7 kcal/mol; serpin–p53: -67.6 ± 6.8 kcal/mol), suggesting that both protein pairs have decent surface complementarity and close packing.

BSA indicates how much surface area becomes inaccessible to solvent upon complex formation. According to Ran and Gestwicki (2018), a larger BSA usually means stronger and more stable binding, with more contact points between the proteins. The serpin–p53 complex (2348.6 ± 32.7 Å²) has a higher BSA than the IPSE/α-1–EGFR complex (1834.2 ± 57.4 Å²). This supports the idea that serpin and p53 form a tighter and more extensive interface than IPSE/α-1 and EGFR.

Secondary structure differences help explain these docking outcomes. IPSE/α-1 has a relatively low α-helix content (18.65%) and higher proportions of extended strands (30.60%) and random coils (50.75%), indicating a flexible conformation capable of adapting to diverse binding interfaces (ZanettiDomingues, 2020). This flexibility may allow IPSE/α-1 to compete at the domain II/IV interface of EGFR, destabilizing the inhibitory tether and favoring ligand-independent activation.

In contrast, serpin exhibits a higher proportion of α-helices (39.41%) and a substantial number of random coils (45.07%), resulting in a more rigid and compact yet adaptable structure. According to Rehman et al. (2022), α-helices provide structural stability critical in protein–protein interactions, while random coils introduce flexibility that allows conformational fitting to binding partners. This structural combination likely enhances serpin's ability to form a stable complex with p53, consistent with its more negative HADDOCK score.

Carcinogenic Pathways

The HADDOCK results and physicochemical characteristics of the target proteins suggest that protein-protein interactions are highly likely to occur between IPSE/α-1 and EGFR, and between serpin and p53. The energy differences between these complexes reflect variations in their binding strength and nature.

Epidermal Growth Factor Receptor (EGFR) Carcinogenic Pathway

EGFR is a transmembrane tyrosine kinase receptor that regulates cell proliferation, differentiation, and survival (Fig. 3). The domain II/IV tether is necessary to maintain EGFR in its inactive conformation and inhibit uncontrolled downstream signaling leading to hyperplasia. IPSE/α-1 may disrupt this intramolecular tether by outcompeting domain II for binding with domain IV, shifting the receptor into an extended conformation. This exposes the EGFR's ectodomain and induces ligand-independent activation (Sugiyama et al., 2023). This finding may help address the unresolved observation of angiogenesis and urothelial hyperplasia associated with IPSE/α-1 (Mbanefo et al., 2020) as chronic EGFR activation is known to promote angiogenesis, alongside cell proliferation and survival. These processes are implicated in schistosomiasis-induced carcinogenesis.

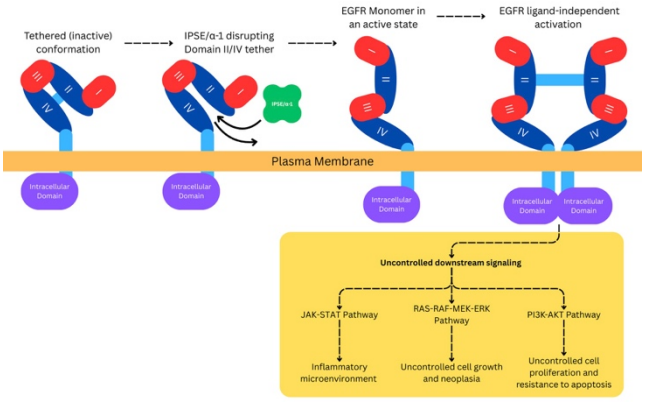


Figure 3: Predicted molecular pathways illustrating the proposed carcinogenic mechanisms of *Schistosoma haematobium*–secreted protein IPSE/α-1 in the development of squamous cell carcinoma of the bladder. IPSE/α-1 is shown to interact with the epidermal growth factor receptor (EGFR) on the urothelial cell surface, potentially disrupting the domain II/IV tether and inducing receptor phosphorylation. This interaction may lead to ligand-independent EGFR activation and stimulation of downstream signaling cascades, including the JAK–STAT, RAS–RAF–MEK–ERK, and PI3K–AKT pathways. These signaling events promote an inflammatory microenvironment, uncontrolled cell proliferation, angiogenesis, and inhibition of apoptosis. The pathway illustrates how IPSE/α-1 may facilitate tumorigenic processes through EGFR-mediated signaling based on computational interaction analyses.

Upon activation, EGFR triggers the JAK/STAT cascade, in which JAK phosphorylation recruits and dimerizes STAT proteins. Chronic activation of this pathway promotes an inflammatory microenvironment that favors tumor growth (Philips et al., 2022).

EGFR also activates the RAF-MEK-ERK cascade, a MAPK pathway involved in cell growth. It begins with RTK and RAS activation, followed by RAF, MEK1/2, and ERK1/2. This cascade drives proliferation and survival, increasing DNA

stress and mutation risk. Such disruptions bypass cell cycle checkpoints, allowing uncontrolled growth and neoplasia (Ullah et al., 2022).

Another key target is the PI3K/AKT pathway, where activated EGFR recruits PI3K, which then activates AKT and mTOR. Persistent mTOR activation blocks apoptosis and promotes cell cycle progression (Rascio et al., 2021).

These findings align with AlHariry et al. (2024), who found increased EGFR expression in *Schistosoma*-induced SCC of the bladder. Their observations show that chronic schistosomiasis is associated with higher EGFR levels compared to non-*Schistosoma* SCC.

Although they did not identify the specific mechanism, the data supports the hypothesis that IPSE/α-1 may interact with EGFR and promote its activation in these cancers.

P53 Carcinogenic Pathway

Figure 4: Predicted molecular pathway illustrating the proposed carcinogenic mechanism of *Schistosoma haematobium*–secreted serpin in the modulation of p53-mediated tumor suppression. Serpin is shown to interact with the tumor suppressor protein p53, binding to its tetramerization domain and preventing the formation of the functional p53 tetramer. This disruption impairs p53's DNA-binding and pro-apoptotic functions, allowing damaged DNA to continue replicating and promoting abnormal cell proliferation. The pathway illustrates how serpin may contribute to schistosomiasis-associated bladder carcinogenesis through inhibition of p53-mediated signaling, as supported by computational docking and interaction analyses.

Serpin may disrupt p53-mediated tumor suppression by binding to its tetramerization domain (Fig. 4). Normally, p53 safeguards cells by responding to DNA damage and cellular stress, inducing cell cycle arrest, DNA repair, or apoptosis. These functions depend on the formation of a functional p53 tetramer that binds DNA and activates regulatory genes (Chen et al., 2022).

If serpin binds tightly to this domain, it may block p53 from assembling into a tetramer. This can leave p53 in a non-functional monomeric form. As a result, p53 cannot bind DNA, stop the cell cycle, and repair DNA damage. This leads to unchecked cell division and the accumulation of mutations.

This hypothesis helps explain the findings of Santos et al. (2021), who observed abnormally high but largely non-functional p53 in SCC tissues of schistosomiasis patients. In contrast, normal bladder tissues express p53 at very low levels, which rapidly degrade after doing their job (Salomao et al., 2021). Serpin's potential inhibition of p53 tetramerization provides a molecular mechanism for this accumulation of inactive p53.

Chronic schistosomiasis may amplify this effect, as continuous secretion of serpin increases its local concentration in host tissues. Sustained inhibition of p53 function could therefore contribute to carcinogenesis, particularly in the bladder, where *S. haematobium* eggs accumulate.

Potential Therapeutic Methods

For decades, schistosomiasis control and the prevention of its progression to bladder cancer have primarily relied on mass drug administration of praziquantel, the only antischistosomal drug with high efficacy (Molehin, 2020). While it remains the primary treatment, praziquantel is only effective against adult parasites, and a single treatment may not fully eradicate the infection (Secor & Montgomery, 2015). Additionally, Montgomery and Richards (2018) reported that praziquantel administration reduces the egg burden by 95–99% in patients who are not cured of schistosomiasis. However, this reduction is attributed to the host's immune response to praziquantel-damaged adult worms, rather than a direct effect on the eggs.

The drug acts by disrupting the parasite's tegument, leading to increased calcium ion influx, muscle paralysis, and parasite death, with a reduction in egg output. Given its mechanism, it cannot target *Schistosoma* eggs or their secreted proteins. Consequently, it cannot prevent or reverse the carcinogenic effects of egg-derived molecules such as IPSE/α-1 and serpin, which contribute to immune modulation and tumorigenesis (Sabra et al., 2025).

Furthermore, praziquantel's side effects, such as vomiting, abdominal pain, headache, and nausea (Kuevi et al., 2023), could further burden cancer patients who already experience fatigue, immunosuppression, and gastrointestinal distress. The added toxicity of praziquantel may further compromise their overall health and quality of life, making the search for alternative strategies even more critical.

In line with these, alternative strategies must be explored to develop a targeted therapeutic approach for *Schistosoma*-induced SCC of the bladder. One potential strategy is the development of small-molecule inhibitors that specifically disrupt the interactions between schistosome and host proteins, particularly IPSE/α-1–EGFR and serpin–p53. By disrupting these interactions, the modulation of host cellular pathways that lead to tumor development may be prevented.

In this study, it was formulated that IPSE/α-1 could potentially disrupt the intramolecular tether between EGFR domain II and IV. Weng et al. (2018) mentioned that over-expression and activation of EGFR can lead to uncontrolled cell proliferation, inhibition of apoptosis, and increased vascular permeability. Using TKIs could block EGFR autophosphorylation triggered by the proposed mechanism in which the parasite protein untethers domains II and IV (Shah & Lester, 2019). TKIs are primarily used for the treatment of non-small-cell lung cancer, including squamous cell carcinoma. Considering the inhibitory capability of TKIs, their potential use could be explored in SCC of the bladder, particularly in schistosomiasis-induced malignancies.

Moreover, serpin may play a potential role in inhibiting the tetramerization of p53, leading to the accumulation of mutated p53 and the loss of its proper function. A study by Zatloukalová et al. (2018) emphasized that treatment with the methylated form of PRIMA-1 (APR-246) can restore p53 tetramerization, thereby recovering its tumor suppressor function. APR-246 is commonly used in anticancer therapy, specifically for cancers with p53 mutations.

These computational findings elucidate how *S. haematobium*-secreted proteins may contribute to bladder carcinogenesis through distinct host–pathogen interactions. While previous research has largely focused on chronic inflammation as the primary driver of

schistosomiasis-associated carcinogenesis, the present computational findings address a key research gap by providing mechanistic insight into how specific parasite-derived proteins may directly influence tumorigenic pathways in the host. The molecular docking and interaction analyses revealed possible affinities between IPSE/α-1 and EGFR, as well as between serpin and p53, suggesting that these secreted proteins can modulate host cellular signaling involved in proliferation and apoptosis. These findings expand the current understanding of schistosome-mediated oncogenesis by demonstrating plausible protein–protein interactions that link infection to cancer initiation at the molecular level.

Future work should involve experimental validation through both *in vitro* assays and *in vivo* studies to confirm the predicted binding sites and interaction stability, as well as to verify these mechanisms within a physiological context. The potential applicability of tyrosine kinase inhibitors (TKIs) to disrupt IPSE/α-1–EGFR signaling, and the p53-reactivating compound APR-246 to counteract serpin-induced p53 dysfunction, serves as a translational link between computational predictions and biological validation. These approaches collectively advance both the molecular understanding and therapeutic management of schistosomiasis-associated bladder cancer.

Study Limitations

Although these findings shed light on the molecular interactions contributing to bladder carcinogenesis, it is important to recognize the limitations of this study. Computational modeling offers a powerful and cost-effective means to explore molecular mechanisms and generate hypotheses that guide further research. However, these predictions have inherent limitations. The results are based solely on computational simulations and thus require validation through experimental studies such as *in vitro* assays and *in vivo* studies. These experiments could be used to confirm the biological relevance of the findings in this study. Molecular docking relies on simplified assumptions about protein flexibility and solvent environment, which can result in false positives or overestimated binding affinities. Moreover, this study did not account for post-translational modifications like glycosylation in IPSE/α-1, which may significantly influence its interaction with glycan-sensitive receptors such as EGFR. The omission of such modifications could limit the accuracy of interaction predictions. Additionally, this study was limited to the analysis of pathways involving EGFR and p53, which are upregulated in schistosomiasisassociated SCC of the bladder. Other molecular interactions and signaling pathways beyond this scope were not explored and may be considered in future studies to provide a more comprehensive understanding of parasite-induced oncogenesis. Finally, focusing on only two parasite proteins restricts the broader understanding of *Schistosoma*'s complex molecular interplay with the host.

CONCLUSION

Schistosoma haematobium has been linked to SCC of the bladder, primarily due to chronic inflammation and affected pathways caused by the interaction of the parasite-derived proteins with host proteins. Using molecular characterization and homology modeling, this study explored the potential role of IPSE/α-1 and serpin and their mechanism in influencing a microenvironment conducive to carcinogenesis.

Molecular characterization determined the clinical significance of the proteins' estimated half-life, instability, aliphatic indices, and GRAVY scores. The results showed that IPSE/α-1 and serpin are hydrophilic and structurally stable under physiological conditions, which may allow them to persist in the host environment for longer periods. The predicted secondary structures also revealed that serpin has a higher proportion of α-helices, suggesting a more stable structure than IPSE/α-1. Meanwhile, IPSE/α-1 has more extended strands and random coils, allowing it to be more flexible to perform specialized functions.

Through homology modeling and docking analyses, the study determined the potential interaction of IPSE/α-1 with EGFR and serpin with the p53 protein. The findings suggest that IPSE/α1 may disrupt the intramolecular tether between EGFR domain II and IV, activating EGFR signaling and inducing carcinogenesis by uncontrollably stimulating cellular pathways involved in cell proliferation and survival. On the other hand, serpin may bind strongly to the tetramerization domain of p53, rendering p53 non-functional and inhibiting DNA repair pathways, which promotes cancer cell survival.

Through *in silico* methods, the results of the study established the physicochemical properties and structures of IPSE/α-1 and serpin along with their interactions with the host. These computational approaches can be used as an initial screening tool for understanding biomolecules and physiological processes without the need for live models. When *in silico* methods yield promising findings, experimental validation through *in vitro* and *in vivo* studies should follow.

Experimental techniques such as co-immunoprecipitation assays can be used to confirm direct binding between the parasite and host proteins. Cell-based functional assays may reveal their downstream effects on cell survival, proliferation, and apoptosis. *In vivo* studies using animal models are essential to assess the biological consequences of these interactions in a whole-organism context.

Furthermore, future research may employ X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy to expand the findings generated by *in silico* analysis. These techniques may provide high-resolution, three-dimensional structural data that confirm protein folding, binding interfaces, and molecular interactions. This will provide further understanding of the precise mechanisms involved in IPSE/α-1—EGFR and serpin—p53 interactions.

The study also highlights the therapeutic potential of targeting these PPIs. Inhibitors could be developed to disrupt these interactions, possibly restoring normal host cellular function and preventing *Schistosoma*-induced carcinogenesis. TKIs used for treating EGFR-driven cancers may help block signaling pathways activated by IPSE/α-1 binding. Similarly, compounds like APR-246 may counteract the inhibitory effects of serpin on the tetramerization of p53. These targeted approaches could reduce tumor progression in patients with schistosomiasis-associated bladder cancer.

Additionally, if *in silico* methods are to be used, future studies should explore IPSE/α-1 glycosylation, as it may improve docking analysis with EGFR and other host proteins that rely on glycans for binding. Other parasite-derived proteins can also be studied, as this may uncover additional novel PPIs that could further improve knowledge of the molecular mechanisms of *Schistosoma*-induced bladder cancer.

By integrating computational, biochemical, and structural biology strategies, future research can contribute to the development of targeted interventions against *Schistosoma*-induced bladder cancer, ultimately advancing knowledge in parasite-associated carcinogenesis and developing potential treatment options.

ACKNOWLEDGMENTS

The researchers extend their heartfelt gratitude to all who supported them throughout the completion of this research paper. They are especially thankful to their parents, loved ones, and friends for their unwavering support and encouragement, which serves as a constant source of motivation. Deep appreciation goes to their research adviser, Ms. Mary Rose Lirio, for her invaluable guidance, insight, and inspiration. They also appreciate Ms. Melisa Mondoy, Ms. Laarni Lacorte, and Mr. Edrick Dianzon for their constructive feedback during the proposal defense, which helped strengthen the study. Above all, they are grateful to Almighty God for His guidance and blessings, which enabled them to overcome challenges and complete this work.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

CONTRIBUTIONS OF INDIVIDUAL AUTHORS

The original conception of the study was formulated by Arthur F. de Vera. All authors contributed to the refinement of the study's concepts and design, collected data on the proteins' physicochemical properties and structures, and conducted the docking simulations. Arthur F. de Vera, Leila Sophia A. Casano, Alyssa Nicole M. Elemia, and Kelsey Beatrice F. Faderanga interpreted and analyzed the study's findings. All authors have read, critically reviewed, and approved of this study.

REFERENCES

Abdollahi, S., Morowvat, M. H., Savardashtaki, A., Irajie, C., Najafipour, S., Zarei, M., & Ghasemi, Y. (2020). Amino acids sequence-based analysis of arginine deiminase from different prokaryotic organisms: An in silico approach. *Recent Patents on Biotechnology, 14*(3), 235-246.

Ahmad, S., Jose da Costa Gonzales, L., Bowler-Barnett, E. H., Rice, D. L., Kim, M., Wijerathne, S., ... & Martin, M. J. (2025). The UniProt website API: facilitating programmatic access to protein knowledge. *Nucleic Acids Research*, gkaf394.

Ahmed, Z., Zulfiqar, H., Tang, L., & Lin, H. (2022). A statistical analysis of the sequence and structure of thermophilic and non-thermophilic proteins. *International Journal of Molecular Sciences, 23*(17), 10116. https://doi.org/10.3390/ijms231710116

Akcelik-Deveci, S., Kılıç, E., Mansur-Ozen, N., Timucin, E., Buyukcolak, Y., & Oktem-Okullu, S. (2024). Identification of interaction partners of outer inflammatory protein A: Computational and experimental insights into how Helicobacter pylori infects host cells. *Plos one, 19*(10), e0300557.

Alam, M. M., Saikat, A. S. M., & Uddin, M. E. (2023). Bioinformatics Approaches for Structural and Functional Annotation of an Uncharacterized Protein of Helicobacter pylori. *Engineering Proceedings, 37*(1), 61.

AlHariry, N. S., El Saftawy, E. A., Aboulhoda, B. E., Abozamel, A. H., Alghamdi, M. A., Hamoud, A. E., & Ghanam, W. A. E. K. (2024). Comparison of tissue biomarkers between Non-Schistosoma and Schistosoma-Associated Urothelial Carcinoma. *Tissue and Cell*, 102416.

Alhumaid, N. K., & Tawfik, E. A. (2024). Reliability of alphafold2 models in virtual drug screening: A focus on selected class A GPCRs. International Journal of Molecular Sciences, 25(18), 10139.

Almeida, F. C. L., Sanches, K., Pinheiro-Aguiar, R., Almeida, V. S., & Caruso, I. P. (2021). Protein surface interactions—theoretical and experimental studies. *Frontiers in Molecular Biosciences*, 8. https://doi.org/10.3389/fmolb.2021.706002

Alrosan, M., Tan, T. C., Easa, A. M., Gammoh, S., & Alu'datt, M. H. (2022). Molecular forces governing protein-protein interaction: Structure-function relationship of complexes protein in the food industry. *Critical reviews in food science and nutrition, 62*(15), 4036-4052.

Ambadapadi, S., Munuswamy-Ramanujam, G., Zheng, D., Sullivan, C., Dai, E., Morshed, S., ... & Lucas, A. (2016). Reactive center loop (RCL) peptides derived from serpins display independent coagulation and immune modulating activities. *Journal of Biological Chemistry, 291*(6), 2874-2887.

Bhargavi, M., Bhargavi, S. G., & Sripada, E. (2025). Computational Approach for HDAC1 Predicting Protein-Ligand Interactions for Cancer through Homology Modelling, Virtual Screening and Molecular Docking. *Convergence of Technology & Biology─ Transforming Life Sciences,* 8-21.

Cagnin, S., Pontisso, P., & Martini, A. (2024). SerpinB3: A multifaceted player in health and disease —Review and Future Perspectives. *Cancers, 16*(14), 2579.

Chan, E. D., King, P. T., Bai, X., Schoffstall, A. M., Sandhaus, R. A., & Buckle, A. M. (2024). The inhibition of serine proteases by serpins is augmented by negatively charged heparin: A concise review of some clinically relevant interactions. In *International Journal of Molecular Sciences* (Vol. 25, Issue 3). https://doi.org/10.3390/ijms25031804

Chen, C., Fu, G., Guo, Q., Xue, S., & Luo, S. Z. (2022). Phase separation of p53 induced by its unstructured basic region and prevented by oncogenic mutations in tetramerization domain. *International Journal of Biological Macromolecules, 222*, 207-216.

Cunanan, D. J., Gonzalvo, P. C. N., Gochuico, M. L. C., Falzado, C. D. V., Cruz, A. E., Dejapa, E. P., Godin, C. R. M. Y., Fenix, T. N. T., & Lirio, M. R. F. (2023). Molecular Characterization and homology modeling of intracellular adhesion regulatory (ICAR) proteins in Biofilm-Producing Staphylococcus species. *Asian Journal of Biological and Life Sciences, 12*(2), 301-309. https://doi.org/10.5530/ajbls.2023.12.41

Darmawati, S., Ethica, S. N., Prastiyanto, M. E., Depamede, S. N., Putri, E. O., & Kamaruddin, M. (2022). Molecular characterization of a 42 kDa subunit pili protein of Salmonella typhi causes typhoid fever. *Biodiversitas Journal of Biological Diversity, 23*(2).

E Lohning, A., M Levonis, S., Williams-Noonan, B., & S Schweiker, S. (2017). A practical guide to molecular docking and homology modelling for medicinal chemists. *Current topics in medicinal chemistry, 17*(18), 2023-2040.

Farling, K. B. (2017). Bladder cancer. *The Nurse Practitioner, 42*(3), 26-33. https://doi.org/10.1097/01.pr.0000512251.61454.5d

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. E., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In *The proteomics protocols handbook* (pp. 571-607). Totowa, NJ: Humana press.

Geourjon, C. and Deléage, G. 1995. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments, C*omput Appl Biosci*., 11:681684.

Guex, N., Peitsch, M. C., & Schwede, T. (2009). Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*, 30(S1), S162-S173.

Heo, L., Park, H., & Seok, C. (2013). GalaxyRefine: Protein structure refinement driven by side-chain repacking. *Nucleic Acids Research, 41*(W1), W384-W388. https://doi.org/10.1093/nar/gkt458

Honorato, R. V., Koukos, P. I., Jiménez-García, B., Tsaregorodtsev, A., Verlato, M., Giachetti, A., ... & Bonvin, A. M. (2021). Structural biology in the clouds: the WeNMR-EOSC ecosystem. *Frontiers in molecular biosciences*, 8, 729513.

Honorato, R. V., Trellet, M. E., Jiménez-García, B., Schaarschmidt, J. J., Giulini, M., Reys, V., ... & Bonvin, A. M. (2024). The HADDOCK2. 4 web server for integrative modeling of biomolecular complexes. *Nature protocols*, 19(11), 3219-3241.

Huntington, J. A., & Carrell, R. W. (2001). The serpins: Nature's molecular mousetraps. *Science Progress*, *84*(2), 125–136. https://doi.org/10.3184/003685001783239032 International Agency for Research on Cancer. (2012). *Biological agents*. In www.ncbi.nlm.nih.gov.

International Agency for Research on Cancer. https://www.ncbi.nlm.nih.gov/books/NBK304343/

Janciauskiene, S., Lechowicz, U., Pelc, M., Olejnicka, B., & Chorostowska-Wynimko, J. (2024). Diagnostic and therapeutic value of human serpin family proteins. *Biomedicine & Pharmacotherapy, 175*, 116618.

Jernigan, R. L., Khade, P., Kumar, A., & Kloczkowski, A. (2022). Using Surface Hydrophobicity Together with Empirical Potentials to Identify Protein-Protein Binding Sites: Application to the Interactions of E-cadherins. *Methods in molecular biology (Clifton, N.J.), 2340*, 41–50. 580 https://doi.org/10.1007/978-1-0716-1546-1_3

Knuhr, K., Langhans, K., Nyenhuis, S., Viertmann, K., Kildemoes, A. M., Doenhoff, M. J., Haas, H., & Schramm, G. (2018). Schistosoma mansoni egg-released ipse/alpha-1 dampens inflammatory cytokine responses via Basophil interleukin (IL)-4 and IL-13. *Frontiers in Immunology, 9*. https://doi.org/10.3389/fimmu.2018.02293

Kontomanolis, E. N., Koutras, A., Syllaios, A., Schizas, D., Mastoraki, A., Garmpis, N., ... & Fasoulakis, Z. (2020). Role of oncogenes and tumor-suppressor genes in carcinogenesis: a review. *Anticancer research, 40*(11), 6009-6015.

Kuevi, D. N. O., Acquah, F. A., Amuquandoh, A., & Abbey, A. P. (2023). Challenges and proven recommendations of praziquantel formulation. *Journal of Clinical Pharmacy and Therapeutics, 2023*(1), 3976392.

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Applied Crystallography*, 26(2), 283291.

Machicado, C., Soto, M. P., La Chira, L. F., Torres, J., Mendoza, C., & Marcos, L. A. (2021). In silico prediction of secretory proteins of Opisthorchis viverrini, Clonorchis sinensis and Fasciola hepatica that target the host cell nucleus. *Heliyon, 7*(7).

Mbanefo, E. C., Agbo, C. T., Zhao, Y., Lamanna, O. K., Thai, K. H., Karinshak, S. E., Khan, M. A., Fu, C. L., Odegaard, J. I., Saltikova, I. v., Smout, M. J., Pennington, L. F., Nicolls, M. R., Jardetzky, T. S., Loukas, A., Brindley, P. J., Falcone, F. H., & Hsieh, M. H. (2020). IPSE, an abundant egg secreted protein of the carcinogenic helminth Schistosoma haematobium, promotes proliferation of bladder cancer cells and angiogenesis. *Infectious Agents and Cancer, 15*(1). https://doi.org/10.1186/s13027- 020-00331-6

Meyer, N. H., Schramm, G., & Sattler, M. (2011). 1H, 13C and 15N chemical shift assignments of IPSEΔNLS. *Biomolecular NMR assignments, 5*(2), 225-227.

Miller, L. E., & Stevens, C. D. (2021). *Clinical immunology and serology: A laboratory perspective* (5th Ed.). F.A. Davis Company.

Molehin, A. J. (2020). Schistosomiasis Vaccine Development: Update on human clinical trials. *Journal of Biomedical Science, 27*(1). https://doi.org/10.1186/s12929-020-0621-y

Montgomery, S. P., & Richards Jr, F. O. *285 Blood Trematodes: Schistosomiasis*.

Nazareth, L. C., Lupenza, E. T., Zacharia, A., & Ngasala, B. E. (2022). Urogenital schistosomiasis prevalence, knowledge, practices and compliance to MDA among school-age children in an endemic district, southern East Tanzania. *Parasite Epidemiology and Control, 18*, e00257.

Ocampo, D. N., Olorosisimo, H. E., Palting, E. J., Pascual, L. A., Ramos, E. S., Ruela, J. A., Salonga, S., Sanchez, A. B. F., & Lirio, M. R. (2024). Proteomic analysis of AeD7L1 and AeD7L2 in Aedes aegypti to the biogenic amines in host hemostasis. *International Journal of Mosquito Research, 11*(2), 37–48.

Park, S. W., Lee, B. H., & Kim, M. K. (2023). Revisiting the Ramachandran plot based on statistical analysis of static and dynamic characteristics of protein structures. *Journal of Structural Biology, 215*(1), 107939. https://doi.org/10.1016/j.jsb.2023.107939

Pathak, Y., Mehta, S., & Priyakumar, U. D. (2021). Learning atomic interactions through solvation free energy prediction using graph neural networks. *Journal of Chemical Information and Modeling, 61*(2), 689-698.

Pedroza-Escobar, D., Castillo-Maldonado, I., González-Cortés, T., Delgadillo-Guzman, D., RuizFlores, P., Cruz, J. H., Espino-Silva, P.-K., Flores-Loyola, E., Ramirez-Moreno, A., Avalos-

Soto, J., Téllez-López, M.-Á., Velázquez-Gauna, S.-E., Garcia-Garza, R., Vertti, R. D., & Torres-León, C. (2023). Molecular bases of protein antigenicity and determinants of immunogenicity, Anergy, and Mitogenicity. *Protein Peptide Letters, 30*(9), 719-733. https://doi.org/10.2174/0929866530666230907093339

Pennington, L. F., Alouffi, A., Mbanefo, E. C., Ray, D., Heery, D. M., Jardetzky, T. S., ... & Falcone, F. H. (2017). H-IPSE is a pathogen-secreted host nucleus-infiltrating protein (infiltrin)

expressed exclusively by the Schistosoma haematobium egg stage. *Infection and Immunity, 85*(12), 10- 1128.

Philips, R. L., Wang, Y., Cheon, H., Kanno, Y., Gadina, M., Sartorelli, V., ... & O'Shea, J. J. (2022). The JAK-STAT pathway at 30: Much learned, much more to do. *Cell, 185*(21), 3857-3876.

Powell, H. R., Islam, S. A., David, A., & Sternberg, M. J. (2025). Phyre2. 2: a community resource for template-based protein structure prediction. *Journal of molecular biology*, 168960.

Ran, X., & Gestwicki, J. E. (2018). Inhibitors of protein-protein interactions (PP|s): An analysis of scaffold choices and buried surface area. *Current opinion in chemical biology*, 44, 75-86

Rascio, F., Spadaccino, F., Rocchetti, M. T., Castellano, G., Stallone, G., Netti, G. S., & Ranieri, E. (2021). The pathogenic role of PI3K/AKT pathway in cancer onset and drug resistance: an updated review. *Cancers, 13*(16), 3949.

Rehman, I., Farooq, M., & Botelho, S. (2022, December 11). Biochemistry, Secondary protein structure. *StatPearls - NCBI Bookshelf*. https://www.ncbi.nlm.nih.gov/books/NBK470235/

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. Nature protocols, 5(4), 725-738.

Sabra, A. N. A., Salem, M. B., William, S., Hammam, O. A., & El-Lakkany, N. M. (2025). Optimizing the therapeutic benefits of synriam combined with praziquantel in mice harbouring juvenile and mature Schistosoma mansoni. *Scientific Reports, 15*(1), 28867.

Salem, S., Mitchell, R. E., El-Alim El-Dorey, A., Smith, J. A. & Barocas, D. A. Successful control of schistosomiasis and the changing epidemiology of bladder cancer in Egypt. *BJU Int.* 107, 206– 211 (2011).

Salomao, N., Karakostis, K., Hupp, T., Vollrath, F., Vojtesek, B., & Fahraeus, R. (2021). What do we need to know and understand about p53 to improve its clinical value?. *The Journal of Pathology, 254*(4), 443-453.

Santos, L. L., Santos, J., Gouveia, M. J., Bernardo, C., Lopes, C., Rinaldi, G., Brindley, P. J., & da Costa, J. M. C. (2021). Review urogenital schistosomiasis—history, pathogenesis, and bladder cancer. In *Journal of Clinical Medicine* (Vol. 10, Issue 2). https://doi.org/10.3390/jcm10020205

Saponaro, A., Maione, V., Bonvin, A. M., & Cantini, F. (2020). Understanding docking complexes of macromolecules using HADDOCK: the synergy between experimental data and computations. *Bio-protocol, 10*(20), e3793-e3793.

Sayers, E. W., Cavanaugh, M., Clark, K., Pruitt, K. D., Schoch, C. L., Sherry, S. T., & Karsch-Mizrachi, I. (2021). GenBank. *Nucleic acids research*, 49(D1), D92-D96.

Schrodinger, L. L. C. (2015). The PyMOL molecular graphics system. Version, 1, 8. *CiNii Research*. https://cir.nii.ac.jp/crid/1370294643858081026

Secor, W. E., & Montgomery, S. P. (2015). Something old, something new: is praziquantel enough for schistosomiasis control? *Future MedicinalChemistry, 7*(6), 681–684. https://doi.org/10.4155/fmc.15.9

Seok, C., Baek, M., Steinegger, M., Park, H., Lee, G. R., & Won, J. (2021). Accurate protein structure prediction: what comes next. *Biodesign*, 9(3), 47-50.

Shah, R., & Lester, J. F. (2019). Tyrosine kinase inhibitors for the treatment of EGFR Mutation-Positive Non–Small-Cell lung Cancer: a clash of the generations. *Clinical Lung Cancer, 21*(3), e216– e228. https://doi.org/10.1016/j.cllc.2019.12.003

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., ... & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*, 7(1), 539.

Sugiyama, M. G., Brown, A. I., Vega-Lugo, J., Borges, J. P., Scott, A. M., Jaqaman, K., ... & Antonescu, C. N. (2023). Confinement of unliganded EGFR by tetraspanin nanodomains gates EGFR ligand binding and signaling. *Nature communications, 14*(1), 2681.

Ullah, R., Yin, Q., Snell, A. H., & Wan, L. (2022, October). RAF-MEK-ERK pathway in cancer evolution and treatment. In *Seminars in cancer biology* (Vol. 85, pp. 123-154). Academic Press.

van Hoogstraten, L. M., Vrieling, A., van der Heijden, A. G., Kogevinas, M., Richters, A., & Kiemeney, L. A. (2023). Global trends in the epidemiology of bladder cancer: challenges for public health and clinical practice. *Nature reviews Clinical oncology, 20*(5), 287-304.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., ... & Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research*, 46(W1), W296-W303.

Wang, J., Chitsaz, F., Derbyshire, M. K., Gonzales, N. R., Gwadz, M., Lu, S., ... & Marchler-Bauer, A. (2023). The conserved domain database in 2023. Nucleic acids research, 51(D1), D384D388.Weng, M., Chang, J., Hung, W., Yang, Y., & Chien, M. (2018). The interplay of reactive oxygen species and the epidermal growth factor receptor in tumor progression and drug resistance. *Journal of Experimental & Clinical Cancer Research, 37*(1). https://doi.org/10.1186/s13046-018-0728-0

Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S., & Richardson, D. C. (2017). Molprobity: More and better reference data for improved all-atom structure validation. *Protein Science, 27* (1), 293-315. https:// doi.org/10.1002/pro.3330

Wu, K., Zhai, X., Huang, S., Jiang, L., Yu, Z., & Huang, J. (2021). Protein kinases: potential drug targets against Schistosoma japonicum. *Frontiers in Cellular and Infection Microbiology, 11*, 691757

Yan, Y., Tao, H., He, J., & Huang, S. Y. (2020). The HDOCK server for integrated protein–protein docking. *Nature protocols, 15*(5), 1829-1852.

Yang Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. BMC bioinformatics, 9(1), 40.

Zanetti-Domingues, L. C., Bonner, S. E., Martin-Fernandez, M. L., & Huber, V. (2020). Mechanisms of action of EGFR tyrosine kinase receptor incorporated in extracellular vesicles. *Cells, 9*(11), 2505.

Zatloukalová, P., Galoczová, M., & Vojtěšek, B. (2018). Prima-1 and APR-246 in cancer therapy. *Klinicka Onkologie, 31*(Suppl2). https://doi.org/10.14735/amko20182s71