IOSR Journal of Research & Method in Education (IOSR-JRME)
e-ISSN: 2320–7388, p- ISSN: 2320-737x Volume 16, Issue 4 Ser. 3 (July. – August. 2026) 48-53
www.iosrjournals.org

Decoding The Egg's Achilles' Heel: Immunoinformatics Unveils A Stable C-Terminal Epitope For Next-Generation Schistosomiasis Diagnosis

Leonardo Ferreira Oliveira[1], Nayara Gonçalves Pereira[2], Thiago Alves Xavier Dos Santos[2], Aderson Vasconcelos Santos[3], Danilo Dos Reis Santos[4], Joao Bosco De Souza Júnior[5], Lara Thalita Ramos Freire[6], Tamara Alves De Souza[6], Agueda Maria De França Tavares[7], Mateus Guedes Araújo[7], Ana Elisa Lelis De Oliveira[8]

[1]*Master's Degree In Animal Science - Universidade Federal De Minas Gerais, Brazil*
[2]*Ph.D. Candidate In Biotechnology - Universidade Estadual De Montes Claros, Brazil*
[3]*Graduate Certificate In Hematology And Blood Banking - Academia De Ciência E Tecnologia, Brazil*
[4]*Graduate Certificate In Microbiology - Faculdades Unidas Do Norte De Minas, Brazil*
[5]*Bachelor's Degree In Chemical Engineering - Faculdade De Ciência E Tecnologia De Montes Claros, Brazil*
[6]*Bachelor's Degree In Biomedical Sciences - Faculdades Unidas Do Norte De Minas, Brazil*
[7]*Master's Candidate In Animal Science - Universidade Federal De Minas Gerais, Brazil*
[8]*Bachelor's Student In Pharmacy - Afya Centro Universitário Montes Claros, Brazil*

Abstract:
Background*: Schistosomiasis, caused by Schistosoma mansoni, remains a significant global health burden, affecting over 240 million people. Current diagnostics rely heavily on crude Soluble Egg Antigen (SEA), which faces challenges in standardization, high production costs, and cross-reactivity. Immunoinformatics offers a transformative approach for the rational design of precise molecular targets. This study aims to identify and evaluate B-cell epitopes from the immunodominant P40 egg protein using an integrated in silico approach to develop robust candidates for rapid point-of-care (POC) diagnostic tools.*
Materials and Methods*: In this computational study, the primary sequence and three-dimensional structure of the S. mansoni P40 protein (UniProt P12812) were analyzed using a consensus of four bioinformatic tools: BepiPred-3.0 for linear epitopes, and DiscoTope-3.0 and ElliPro for conformational epitopes, integrated with AlphaFold structural modeling. Antigenicity was assessed using VaxiJen v2.0 (parasite model, threshold 0.5). Structural exposure and accessibility were validated through PyMOL visualization. Epitopes were selected based on the convergence of sequence, surface geometry, and structural context.*
Results*: Integration of the predictive tools revealed a non-uniform distribution of antigenic signals, with a distinct consensus hotspot located in the C-terminal portion of the molecule. The region comprising residues 298-316 was the only segment concordantly identified by all four tools as a high-performance antigen. This region exhibited high structural reliability (average pLDDT 92.6) and was part of a larger conformational surface (ElliPro Cluster 2, score 0.747) that spatially approximates residues 235-267 and 298-316. BepiPred-3.0 showed the highest confirmation rate (100%), while linear ElliPro was the most conservative (50%).*
Conclusion: *The integrated immunoinformatics approach identified the C-terminal region (residues 298-316) of the P40 protein as the most promising target for the development of a specific and scalable S. mansoni immunodiagnostic. This synthetic epitope candidate offers a robust alternative to crude antigens, potentially enhancing the sensitivity and specificity of point-of-care tests in endemic areas.*
Key Word: *Schistosoma mansoni; Immunoinformatics; P40 Protein; B-cell Epitopes; Point-of-care Diagnostics.*

Date of Submission: 04-08-2026 Date of Acceptance: 14-08-2026

I. Introduction

Schistosomiasis, caused by the trematode helminth *Schistosoma mansoni*, remains one of the neglected tropical diseases with the highest global socioeconomic impact, affecting more than 240 million people in tropical and subtropical communities[1]. Despite international control efforts, the morbidity associated with chronic infection-including liver fibrosis and portal hypertension-continues to burden healthcare systems in

endemic areas[2]. Currently, transmission control and clinical management rely almost exclusively on mass drug administration with praziquantel[3]. However, dependence on a single drug raises critical concerns regarding the selection of resistant strains and underscores the urgency for developing new diagnostic and therapeutic tools to monitor intervention efficacy and advance toward disease elimination[1,3].

Historically, the Soluble Egg Antigen (SEA) has been established as one of the most relevant diagnostic targets for detecting active and chronic infections, serving as the basis for enzyme-linked immunosorbent assays (ELISA) for antibody detection[4]. Although SEA exhibits high sensitivity, its large-scale use is severely limited by technical and logistical challenges[5]. The production of the native antigen requires complex maintenance of the parasite's life cycle in the laboratory, resulting in high production costs and inherent difficulties in standardization between different batches[6]. Such restrictions prevent the routine implementation of these tests in remote and low-income areas, where laboratory infrastructure is scarce and the diagnostic need is most pressing[4, 6].

In this scenario, immunoinformatics emerges as a transformative approach, enabling the transition from crude and heterogeneous antigens to the rational design of precise molecular targets[7]. Through advanced machine learning algorithms and structural modeling, it is possible to perform systematic prediction of B-cell epitopes derived from immunodominant SEA proteins, such as the P40 protein[8]. This methodology allows for the identification of peptide segments that maintain high specificity and immunogenicity while eliminating cross-reactivity regions with other parasites[9]. Furthermore, predicted epitopes can be chemically synthesized at a low cost with high reproducibility, overcoming the standardization barriers of traditional biological antigens[7,9].

The core strategy of this study is based on the application of immunoinformatics tools for the design of B-cell mimetic epitopes with superior diagnostic potential. By focusing on specific synthetic epitopes, it is possible to mitigate the limitations of native SEA, providing a robust platform for the development of rapid point-of-care (POC) tests[10]. These lateral flow assays, based on defined synthetic peptides, offer the thermal stability and ease of use necessary for field diagnosis, eliminating the need for sophisticated equipment and highly trained technicians[11]. Thus, the replacement of complex antigens with rational epitopes represents a qualitative leap in the precision and accessibility of schistosomiasis diagnosis[10,11].

The present study aims to fill the critical gap between the need for precise diagnosis and the limitations of current technologies. The central hypothesis maintains that the identification of consensus epitopes in the C-terminal portion of the *S. mansoni* P40 protein, through the integration of linear and conformational predictions, will result in high-performance diagnostic candidates. By aligning the technological innovation of immunoinformatics with the World Health Organization's global efforts for schistosomiasis elimination, this work proposes a scalable and sustainable pathway for strengthening epidemiological surveillance in endemic areas.

II. Material And Methods

Protein Sequence Acquisition

The amino acid sequence of the *S. mansoni* P40 protein was retrieved from the UniProt database (https://www.uniprot.org/) under the accession code P12812 (P40_SCHMA). This database was selected for its curated functional and structural annotations, ensuring traceability and reproducibility for subsequent analyses.

Three-Dimensional Structure Acquisition

The three-dimensional structural model of the protein was obtained from the AlphaFold Protein Structure Database (https://alphafold.ebi.ac.uk/) under the identifier AF-P12812-F1, generated via deep learning-based prediction. The reliability of the model was assessed using the per-residue confidence score (pLDDT), which served as the structural quality criterion for the subsequent structure-based prediction stages.

Linear B-Cell Epitope Prediction

Linear B-cell epitopes were predicted using BepiPred-3.0, available on the Immune Epitope Database platform (IEDB; http://tools.iedb.org/bcell/, "B Cell Prediction - Sequence Based" module). This tool employs protein language models and exhibits superior performance compared to its previous version (BepiPred-2.0) in identifying antigenic regions solely from the primary sequence. Regions with scores above the threshold recommended by the tool were selected, prioritizing contiguous sequences with higher lengths and scores.

Conformational Epitope Prediction

Since more than 90% of antibody-recognized B-cell epitopes are conformational (discontinuous), the AlphaFold-derived three-dimensional structure was submitted to two complementary structural prediction tools available on the IEDB platform. DiscoTope-3.0 ("B Cell Prediction – Structure Based"; http://tools.iedb.org/discotope/), based on machine learning, estimates the propensity of each surface residue to

form a conformational epitope based on properties such as solvent accessibility and local structural microenvironment. ElliPro (http://tools.iedb.org/ellipro/), in turn, is based on protein surface geometry, calculating the Protrusion Index (PI) for each residue or region, under the premise that segments that are more exposed and projected outward from the globular core of the protein have a higher probability of antibody recognition. The residues and regions predicted by both tools were compared, and those identified concordantly by both methods were prioritized in subsequent analysis stages, serving as an additional criterion for prediction reliability.

Antigenicity Assessment

The peptide sequences selected in the linear and conformational prediction stages were submitted to the VaxiJen v2.0 server (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html), using the "parasite" model specific for parasitic organism antigens, with a threshold of 0.5. Sequences with scores equal to or higher than this value were classified as potential protective antigens, with a theoretical capacity to induce a humoral immune response.

Structural Visualization

The spatial localization of the predicted epitopes was analyzed on the protein's three-dimensional structure using PyMOL software (https://pymol.org/). The regions corresponding to the selected linear and conformational epitopes were highlighted on the structure, allowing for the verification of their exposure on the molecule's surface-an essential criterion for immune system accessibility-as well as their distribution throughout the protein architecture.

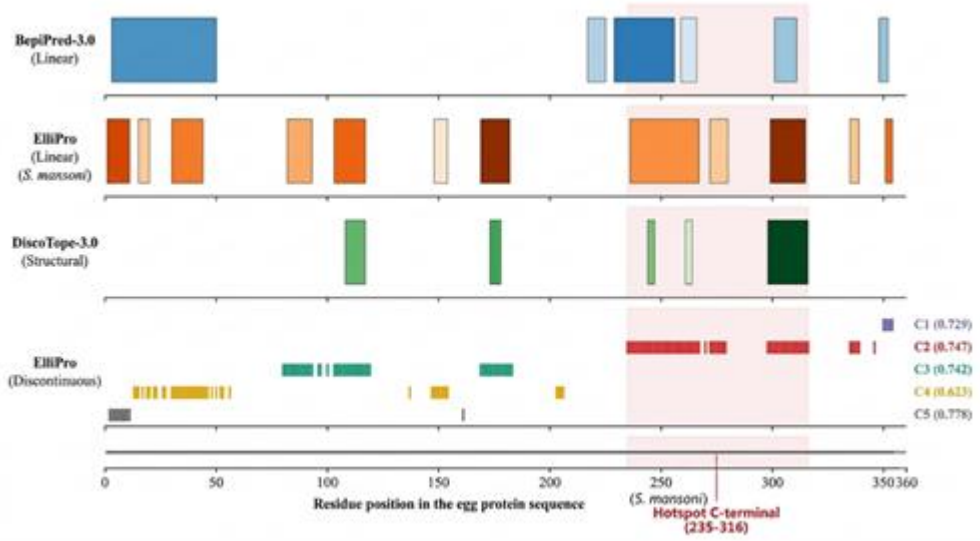
Integration and Final Sequence Selection

Candidate epitope sequences were selected based on the convergence of three criteria: (i) positive prediction by BepiPred-3.0 and/or agreement between DiscoTope-3.0 and ElliPro; (ii) accessible localization on the three-dimensional structure, confirmed by visual inspection in PyMOL; and (iii) an antigenicity score equal to or higher than 0.5 in VaxiJen v2.0 (parasite model).

III. Result

The integration of B-cell epitope predictions obtained with BepiPred-3.0, ElliPro (linear and discontinuous), and DiscoTope-3.0 revealed a non-uniform distribution pattern along the *S. mansoni* egg protein sequence, with a clear concentration of antigenic signal in the C-terminal portion of the molecule (Figure no 1).

Figure no 1. Consensus map of predicted B-cell epitopes along the egg protein sequence.



Caption description: Each track represents a tool; color intensity is proportional to the score reported by that tool. In the bottom track, vertical bars indicate individual residues composing each ElliPro conformational epitope (cluster). The red-shaded region (residues 235-316) corresponds to the C-terminal consensus hotspot.

Spatial overlap of the predictions allowed for the delimitation of eight regions of interest. Comparison of these regions with VaxiJen v2.0 antigenicity scores (Figure no 2) showed that the 298-316 region was the

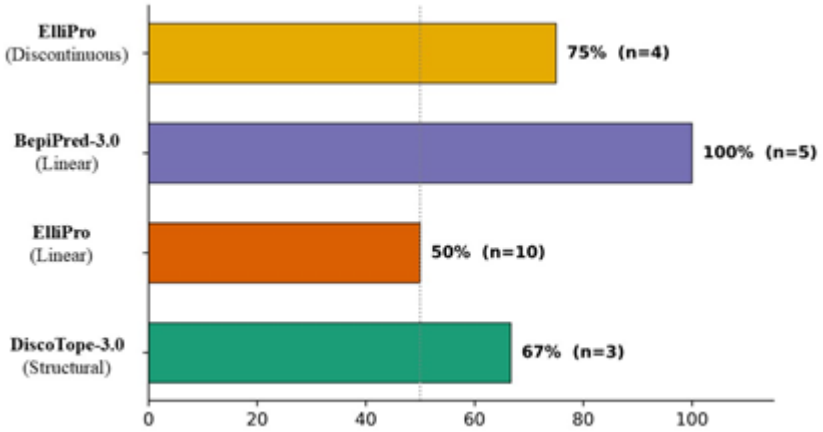
only one classified as an ANTIGEN concordantly by all four tools simultaneously, while the 217-267 region received positive support from three of them. The N-terminal (1-50) and central (82-182) regions showed heterogeneous signals-positive in some tools and negative in others-consistent with epitopes of lower predictive robustness.

Figure no 2. Antigenicity convergence heatmap by region and tool.

Caption description: Each cell corresponds to the average VaxiJen v2.0 score (parasite model) for the tool's peptides overlapping the region; gray cells indicate the absence of tool prediction in the region. Blue outlines highlight the C-terminal hotspot regions (R5 and R7).

Regarding the relative performance of each tool (Figure no3), BepiPred-3.0 showed the highest confirmation rate - 100% of predicted peptides (5/5) were classified as ANTIGEN - followed by discontinuous ElliPro (75%; 3/4) and DiscoTope-3.0 (67%; 2/3). Linear ElliPro was the most conservative, with only 50% confirmation (5/10), reflecting greater sensitivity to local conformational variations.

Figure no 3. Fraction of peptides predicted by each tool confirmed as ANTIGEN by VaxiJen v2.0.

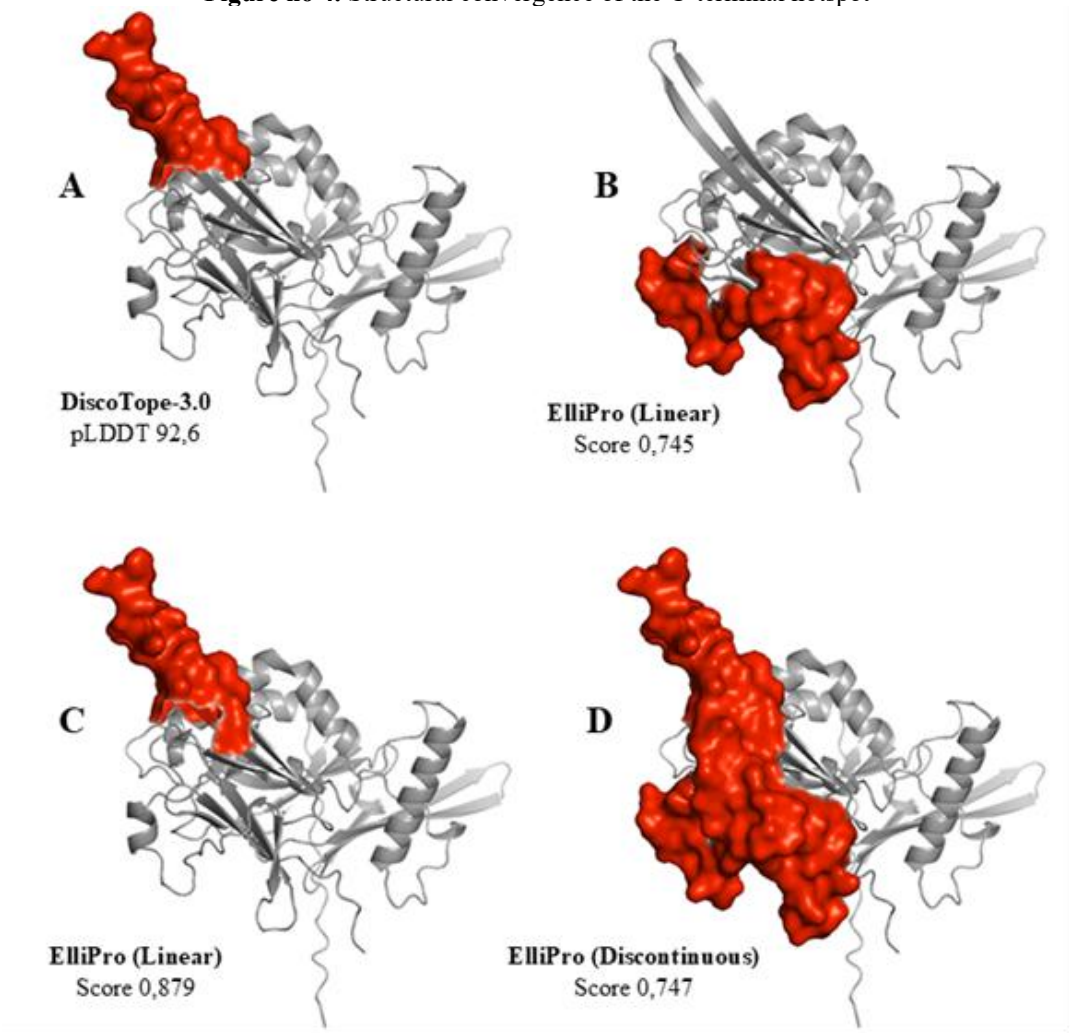


Caption description: Proportion of epitopes predicted by BepiPred-3.0, ElliPro (linear and discontinuous), and DiscoTope-3.0 that were classified as antigens by VaxiJen v2.0. Values in parentheses indicate the number of sequences analyzed (n).

IV. Discussion

The convergence of the four bioinformatic tools in the 298-316 region constitutes the central finding of this study, highlighting it as an immunodominant segment of high diagnostic relevance. This is the only segment simultaneously signaled by methods of distinct methodological natures-sequence-based prediction (BepiPred-3.0), surface geometry (ElliPro), and three-dimensional structural context (DiscoTope-3.0). The robustness of this prediction is corroborated by the high structural reliability of the AlphaFold model in the region, with an average pLDDT of 92.6, ensuring that the prediction rests on a stable protein architecture rather than zones of structural uncertainty[12]. Agreement between independent methods is an established strategy in immunoinformatics to substantially reduce the probability of false positives and identify targets with greater potential for in vitro experimental validation[7, 13].

Figure no 4. Structural convergence of the C-terminal hotspot



Caption description: Structural localization of the C-terminal hotspot predicted by independent methods on the same egg protein structure (red surface = epitope; gray cartoon = rest of the protein). (A) DiscoTope-3.0, residues 298–316. (B) Linear ElliPro, residues 236-267. (C) Linear ElliPro, residues 299-315. (D) Discontinuous ElliPro, Cluster 2 (67 residues), joining regions A-C into a single conformational surface. The near-complete overlap between panels A and C - obtained by independent methods (3D structure vs. surface geometry) - is, in itself, visual evidence of the robustness of this epitope; panel D shows how the native protein folding projects this same region, added to the 236-267 segment (panel B), into a single contiguous and exposed face of the molecule

The conformational epitope identified in ElliPro Cluster 2 (score 0.747) reinforces this conclusion by demonstrating that the native protein folding spatially approximates residues from regions 235-267, 298-316, and 335-339, which are distant in the primary linear sequence. This finding is biologically significant, as more than 90% of B-cell epitopes recognized during natural infection are conformational[14]. The formation of a

continuous antigenic surface in the C-terminal portion suggests that the immunogenicity of the P40 protein does not arise solely from isolated linear motifs but from an exposed three-dimensional architecture, which is crucial for recognition by polyclonal antibodies in infected patients[8,14]. This characteristic guides strategic biotechnology decisions, suggesting that the use of truncated recombinant domains or multi-epitope peptides may be superior to the use of short linear peptides to maximize diagnostic sensitivity[7,15].

In contrast, the N-terminal region (residues 1-50), although signaled by three tools, did not show equivalent structural support in DiscoTope-3.0. This pattern is consistent with regions of high intrinsic flexibility or partially hidden segments in the native globular conformation, which may limit antibody accessibility[16]. In the literature, regions with low structural prediction density are frequently associated with lower efficacy in serological diagnostic assays, qualifying this segment as a second-line candidate compared to the C-terminal hotspot[12,16].

From a biotechnological application perspective, the linear peptide corresponding to the 298-316 region meets ideal characteristics for implementation in rapid diagnostic platforms: high theoretical antigenicity, high structural confidence, and feasibility of direct chemical synthesis [10]. The use of defined synthetic epitopes has proven to be an effective strategy to mitigate cross-reactivity with other helminths (such as Ascaris lumbricoides and Ancylostoma spp.), a chronic problem associated with the use of crude antigens like SEA[4,11]. Furthermore, the chemical stability of synthetic peptides favors their incorporation into lateral flow tests (Point-of-Care), which are essential for epidemiological surveillance in remote areas where Kato-Katz microscopy exhibits low sensitivity[10,17].

Altogether, the results indicate that the consensus strategy between tools of complementary natures - sequence, surface geometry, and 3D structure - is significantly more informative than the isolated application of individual predictors. The identification of the C-terminal region of the P40 protein as the most promising target for the rational development of an *S. mansoni* immunodiagnostic aligns with current trends in reverse vaccinology and precision diagnostics, offering a solid scientific basis for the advancement of new schistosomiasis control tools[7,9,13].

V. Conclusion

This study demonstrates the disruptive potential of immunoinformatics in overcoming the diagnostic bottlenecks of schistosomiasis mansoni. The rational prediction of B-cell epitopes, especially the hotspot identified in the C-terminal region of the P40 protein, offers a robust and scalable alternative to the use of crude antigens. The applicability of these synthetic epitopes as a basis for the development of rapid diagnostic tests is fundamental for transforming surveillance in low-income areas, where conventional microscopy often fails due to low parasite burden. In summary, this innovative approach not only accelerates the development of high-precision diagnostic tools but also provides the technological support necessary for schistosomiasis control and elimination strategies on a global scale.

References

[1]. World Health Organization. WHO Guideline On Control And Elimination Of Human Schistosomiasis. Geneva: WHO; 2022.
[2]. Colley DG, Bustinduy AL, Secor WE, King CH. Human Schistosomiasis. The Lancet. 2014;383(9936):2253–2264.
[3]. Cioli D, Pica-Mattoccia L, Basso A, Guidi A. Praziquantel: 40 Years And Still Going Strong. International Journal For Parasitology: Drugs And Drug Resistance. 2014;4(3):255–261.
[4]. Doenhoff MJ, Chiodini PL, Hamilton JV. Diagnosis Of Schistosomiasis By Antibody Detection. Expert Review Of Molecular Diagnostics. 2004;4(2):173–181.
[5]. Okoye IC, Et Al. Limitations Of Current Diagnostic Techniques For Schistosomiasis. Trends In Parasitology. 2024;30(1):1–2.
[6]. Bergquist R, Utzinger J, Keiser J. Schistosomiasis: From Risk Assessment To Control. Trends In Parasitology. 2015;31(1):16–24.
[7]. Sanches RCO, Et Al. Immunoinformatics Design Of Multi-Epitope Peptide-Based Antigens For Schistosomiasis Diagnosis. Frontiers In Immunology. 2021;12:621706.
[8]. Nene V, Dunne DW, Johnson KS, Rodriguez V, Darby B, Stadecker MJ. Sequence And Expression Of A Major Egg Antigen From Schistosoma Mansoni. Molecular And Biochemical Parasitology. 1986;21(2):179–188.
[9]. De Gregorio E, Rappuoli R. The Next Generation Of Vaccines: From Informatics To Mice To Humans. Science. 2019;365(6452):451–454.
[10]. Corstjens PLAM, Et Al. Point-Of-Care Tests For Schistosomiasis: The Need For Precision. PLOS Neglected Tropical Diseases. 2020;14(2):E0008036.
[11]. Cardoso DT, Et Al. Challenges To Schistosomiasis Control Program In Brazil: Setbacks And Critical Analysis. Revista Da Sociedade Brasileira De Medicina Tropical. 2024;57:E00409.
[12]. Jumper J, Evans R, Pritzel A, Et Al. Highly Accurate Protein Structure Prediction With Alphafold. Nature. 2021;596(7873):583–589.
[13]. Rappuoli R, Bottomley MJ, D'Oro U, Finco O, De Gregorio E. Reverse Vaccinology 2.0: Human Immunology Instructs Vaccine Design. Journal Of Experimental Medicine. 2016;213(4):469–481.
[14]. Potocnakova L, Bhide M, Pulzova LB. B-Cell Epitope Prediction: Tools And Applications In Infectious Diseases. Briefings In Bioinformatics. 2021;22(3):Bbaa152.
[15]. Reis R, Et Al. Synthetic Peptides For Schistosomiasis Diagnosis: A Systematic Review. Expert Review Of Molecular Diagnostics. 2023;23(5):401–412.
[16]. Kandasamy S, Et Al. Accessibility And Flexibility Of B-Cell Epitopes: Implications For Diagnostic Design. Proteomics. 2022;22(8):2100145.