



## Bioinformatics-based prediction and screening of immunogenic epitopes of *Toxoplasma gondii* rhostry proteins 7, 21 and 22 as candidate vaccine target

Fariha Ayub<sup>a</sup>, Haroon Ahmed<sup>a,\*</sup>, Tehreem Sohail<sup>a</sup>, Khuram Shahzad<sup>a</sup>, Figen Celik<sup>b</sup>, Xu Wang<sup>c</sup>, Sami Simsek<sup>b</sup>, Jianping Cao<sup>c,d,\*\*</sup>

<sup>a</sup> Department of Biosciences, COMSATS University Islamabad (CUI), Park Road, Chakh Shahzad, Islamabad, Pakistan

<sup>b</sup> Department of Parasitology, Faculty of Veterinary Medicine, Firat University, Elazig, Turkey

<sup>c</sup> National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention (Chinese Center for Tropical Diseases Research), Key Laboratory of Parasite and Vector Biology, National Health Commission of the People's Republic of China, WHO Collaborating Center for Tropical Diseases, Shanghai, China

<sup>d</sup> The School of Global Health, Chinese Center for Tropical Diseases Research, Shanghai Jiao Tong University School of Medicine, Shanghai, China

### ARTICLE INFO

#### Keywords:

*Toxoplasma gondii*  
Rhoptry proteins  
ROP7  
ROP21  
ROP22  
Bioinformatics  
Vaccine  
Epitope

### ABSTRACT

**Introduction:** Toxoplasmosis is a well-known zoonotic disease caused by *Toxoplasma gondii*. The main causes of the disease range from eating undercooked or contaminated meat and shellfish to cleaning litter trays into which cats that excreted toxoplasma via faeces. This pathogen can live for a very long time, possibly a lifetime, within the bodies of humans and other animals.

**Aims and objectives:** This study aimed to predict and analyse candidate immunogenic epitopes for vaccine development by evaluating the physio-chemical properties, multiple sequence alignment, secondary and tertiary structures, phosphorylation sites, transmembrane domains, and signal peptides, of *T. gondii* rhostry proteins ROP7, ROP21, and ROP22 using bioinformatics tools.

**Methods:** To find immunogenic epitopes of rhostry proteins, numerous bioinformatics web servers were used containing multiple sequence alignment, physiochemical properties, antigenicity and allergenicity, post-translational modification sites (PTMs), signal peptides, transmembrane domains, secondary and tertiary structures, and screening of predicted epitopes. We evaluated immunogenic linear B-cell epitopes as candidate proteins for vaccine development.

**Results:** Nine epitopes were identified for each protein, and analysis of immunogenicity, revealed three candidate epitopes for ROP7, one for ROP21, and four for ROP22. Among all candidate epitopes, ROP22 contained the most immunogenic epitopes with immunogenicity score of 0.50575.

**Conclusion:** We acquired detailed information on predicted immunogenic epitopes using in-silico methods. The results provide a foundation for further experimental analysis of toxoplasmosis, and potential vaccine development.

\* Corresponding author.

\*\* Corresponding author. National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention (Chinese Center for Tropical Diseases Research), Key Laboratory of Parasite and Vector Biology, National Health Commission of the People's Republic of China, WHO Collaborating Center for Tropical Diseases, Shanghai, China.

E-mail addresses: [haroonahmad12@yahoo.com](mailto:haroonahmad12@yahoo.com) (H. Ahmed), [caojp@chinacdc.cn](mailto:caojp@chinacdc.cn) (J. Cao).

<https://doi.org/10.1016/j.heliyon.2023.e18176>

Received 28 December 2022; Received in revised form 4 July 2023; Accepted 10 July 2023

Available online 11 July 2023

2405-8440/© 2023 Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

Toxoplasmosis, caused by *T. gondii*, is a serious global public health issue [1]. The apicomplexan parasite *T. gondii* is unusual in its ability to infect a wide range of intermediate hosts, including all mammals and birds. Infection mainly occurs via consumption of living tissue-cysts present in undercooked meat or oocysts in cat faeces. Congenital transmission may occur when a primary infection arises during pregnancy [2]. Rhoptries, dense granules, and micronemes are three types of apical secretory organelle in *T. gondii*. Rhoptries have a club-like shape and two different halves comprising; an anterior duct (neck) and a posterior bulb. *T. gondii* is the only apicomplexan parasite species in which this particular apical secretory organelle has been found [3,4]. Rhoptries help *T. gondii* to actively enter the host cell, and they play a role in the development of the parasitophorous vacuole, a unique intracellular compartment in which the parasite grows rapidly and escapes intracellular elimination [4].

Rhoptry proteins (ROPs) are essential for parasite survival inside host cells, and they contribute to the various stages of parasite invasion [3–6]. Rhoptries comprise between 1% and 30% of the total volume of toxoplasma cells. Many ROPs including ROP11, ROP16, ROP17, ROP19, ROP29, ROP38, ROP20, ROP21, ROP27, ROP25, ROP28, ROP30, ROP31, ROP32, ROP33, ROP34, ROP35, ROP39, ROP41, and ROP46, are active kinases. Others such as ROP2, ROP8, ROP5, ROP4 and, ROP7 are anticipated to be pseudo-kinases [7].

ROP7 is a member of the ROP2 family. It is found in rhoptries at the apical end of permeable tachyzoites after synthesis and maturation, and it colocalises with ROP1 and translocates into the parasitophorous vacuole following invasion [8,9]. ROP21 contributes to host cell invasion, and it has a protective effect against additional strains of *T. gondii*, but this needs further investigation [10]. ROP22 also contributes to host cell invasion, and it could be useful for developing anti-*T. gondii* vaccines with several components because it partially protects the mice against acute and chronic *T. gondii* infections [1].

Many *T. gondii* antigens have undergone immunogenic effectiveness testing to guard against both acute and chronic toxoplasmosis [1]. Toxoplasmosis vaccine halt disease spreading. In the past 15 years, significant advancements have been made in the fields of immunological techniques, parasite mutant creation, gene cloning, and antigen isolation and characterisation [11].

Commonly given toxoplasma medications only work against the tachyzoite stage; while agents are ineffective against encysted bradyzoites, which are very long lived. They may also cause adverse reactions including hypersensitivity and bone marrow suppression [12]. Consequently, establishing vaccine strategies and developing vaccine pipelines are of utmost importance for both cattle and humans due to the global burden of latent illness. It is strongly advised to detect parasite antigens with high immunogenicity for this purpose [13]. Genes encoding ROPs essential in protozoan pathogenicity and virulence, have gained much attention [14].

Bioinformatics is a powerful and rapidly expanding field [15]. Bioinformatics is mostly used for evaluation of gene and protein expression and to evaluate properties such as immunogenicity, protein structures, and physiochemical characteristics. This knowledge advances our comprehension of protein molecules, and helps analysts choose the ideal proteins or particular epitopes for vaccine development and diagnosis studies [16].

The aim of the present study was to predict and analyse candidate immunogenic epitopes for vaccine development by evaluating physio-chemical properties, multiple sequence alignment (MSA), secondary and tertiary structure, phosphorylation sites, transmembrane domains, and signal peptides, of *T. gondii* proteins (ROP7, ROP21, ROP22) using bioinformatics tools.

## 2. Materials and methods

### 2.1. Protein sequence retrieval

The complete amino acid sequences of ROP7, ROP21, and ROP22 were accessed from the GenBank database (<https://www.ncbi.nlm.nih.gov>). The accession numbers of proteins are XP\_018638552.1, EPT29717.1, and EPT32442.1, for ROP7, ROP21 and ROP22, respectively. Sequences were retrieved in FASTA format [17]. The NCBI Nucleotide database, which links to relevant data including taxonomy, genomes, protein sequences, protein structures, and articles in PubMed, provides access to GenBank [18].

### 2.2. Assessing physiochemical properties

ExPasy ProtParam (<https://web.expasy.org/protparam/>), was used, to evaluate the number of amino acids, molecular weight (MW), aliphatic index, instability index, estimated half-life, extinction coefficients, total number of positively and negatively charged residues and grand average of hydropathicity (GRAVY), and thereby assess the physiochemical features of ROPs [19].

### 2.3. Multiple sequence alignment

The MSA of the three ROPs was performed by using Clustal Omega. (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). Many sequence analysis techniques start with MSA, and this progressive alignment heuristic is used to compute most alignments [20].

### 2.4. Antigenicity and allergenicity

Antigenicity and allergenicity of proteins were calculated using Antigenpro (<https://scratch.proteomics.ics.uci.edu/>) and Vaxijen 2.0 ([http://www.ddg-pharmfac.net/Vaxijen/Vaxijen/Vaxijen\\_help.html](http://www.ddg-pharmfac.net/Vaxijen/Vaxijen/Vaxijen_help.html)) web servers. To predict antigenicity, Antigenpro was used

alone for microarray analysis [21], while Vaxijen was used to estimate allergenicity. Vaxijen was created to enable the classification of antigens without using sequence alignment, which is only based on physiochemical properties of proteins. Vaxijen can be used independently or alongside alignment-based prediction techniques [22]. To assess the allergenicity of proteins, we used the Allertop webserver (<https://www.ddg-pharmfac.net/Allertop/>). Allertop is a webtool used to predict allergens based on primary physiochemical characteristics of proteins [23].

### 2.5. Prediction of post-translational modification sites

Almost all proteins undergo post-translational modification (PTM). Knowing the potential PTMs of template proteins may help us better understand the molecular processes in which they participate because the functions of altered proteins are frequently adversely affected by these changes [24]. The phosphorylation sites of toxoplasma ROP7, ROP21 and ROP22 were evaluated using NetPhos 3.1 (<https://services.healthtech.dtu.dk/service.php?NetPhos-3.1>).

### 2.6. Prediction of signal peptides, transmembrane domains, and isoelectric point

Signal peptides in amino acid sequences were identified using the webserver SignalP 6.0 (<https://services.healthtech.dtu.dk/service.php?SignalP>). This uses a transformer protein language model for structural prediction. Proteins have short amino acid sequences known as signal peptides that direct them into or across membranes. Various bioinformatics tools are available for prediction of signal peptides, but many are unable to differentiate between signal types, as proteins can have different types of signal peptides including Sec/SPI ("standard" secretory signal peptides transported by the Sec translocon and cleaved by Signal Peptidase I), Sec/SPII (lipoprotein signal peptides transported by the Sec translocon and cleaved by Signal Peptidase II), Tat/SPI (Tat signal peptides transported by the Tat translocon and cleaved by Signal Peptidase I), and no signal peptide [25]. Using the TMHMM 2.0 webserver (<https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>) probable transmembrane domains were predicted. TMHMM can correctly predict 97–98% of transmembrane helices accurately. Moreover, TMHMM can distinguish between membrane and soluble proteins along with sensitivity and specificity, but the precision can be affected in the presence of signal peptides [26]. The isoelectric point (pI) of ROP7, ROP21, and ROP22 was calculated using the isoelectric point calculator (IPC) online server (<http://isoelectric.org/>). IPC is a webserver used to precisely calculate the pI of a protein using dissociation constant (pKa) values [27].

### 2.7. Prediction of secondary structure

The position-specific iterated prediction (PSIPRED) online server (<http://bioinf.cs.ucl.ac.uk/psipred>) and the PSI-BLAST results were employed to elucidate the secondary structure of ROP7, ROP21 and ROP22. To predict the secondary structure using PSIPRED, users can select a protein structure prediction server, execute a program of their choice, and receive prediction email and graphically over the web [28].

### 2.8. Prediction of tertiary structure

To predict tertiary structure we used, PHYRE2 (<http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index>). PHYRE2 builds 3D models, examines ligand-binding sites, and assesses the impact of amino acid variation such as nonsynonymous single nucleotide polymorphisms (SNPs) for a given input protein sequence [29].

### 2.9. Prediction of linear B-cell epitopes

Based on the 3D structure of a protein and using default settings, the immune Epitope Database (IEDB) Ellipro webserver (<http://tools.iedb.org/ellipro/>) was used to predict the linear B-cell epitopes. After submission of protein ID, the default threshold for epitope minimum prediction was 0.5 (minimum score) and 6 (maximum distance). Ellipro uses Thornton's method along with residue clustering, Jmol viewer, and MODELLER programs to predict, evaluate and visualise antibody epitopes of a protein of interest [30].

### 2.10. Screening of linear B-cell epitopes

Linear B-cell epitopes toxicity, antigenicity and, allergenicity were investigated by using various webserver. To assess the toxicity of epitopes, we used the Toxinpred webserver with a threshold of 0.5 ([https://webs.iitd.edu.in/raghava/toxinpred/multi\\_submit.php](https://webs.iitd.edu.in/raghava/toxinpred/multi_submit.php)). ToxinPred is an online server that predicts the toxicity of peptides or proteins. It is a useful tool for the construction of vaccines using non-toxic peptides [31]. Non-toxic linear epitopes were examined for antigenicity using Vaxijen (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) with a threshold of 0.5. After exploring toxicity and antigenicity, selected epitopes were further assessed for their allergenicity using the online server Allertop with a threshold of (<https://www.ddg-pharmfac.net/AllerTOP/>).

### 2.11. Prediction of immunogenicity of probable epitopes

To analyse immunogenicity of linear epitopes, we used the IEDB server class I immunogenicity (<http://tools.iedb.org/>)



immunogenicity/) [32].

### 3. Results

#### 3.1. Prediction of physicochemical properties

The GenBank database was used to retrieve the amino acid sequences of ROP7 (XP\_018638552.1), ROP21 (EPT29717.1), and ROP22 (EPT32442.1) in FASTA Format. ROP7 is a 575 amino acid polypeptide sequence with a pI of 6.53 and a molecular mass of 63,395.25 Da, according to the ProtParam server results. A total of 64 positively and 67 negatively charged residues were found in the protein. At 280 nm, the 8926 atoms have an extinction coefficient of 74,370. The half-life was calculated 30 h *in vitro* (mammalian reticulocytes), >20 h *in vivo* (yeast), and >10 h *in vivo* (*Escherichia coli*). The instability score was calculated as 42.43, indicating that it is unstable. The aliphatic and GRAVY scores were calculated to be 86.16 and -0.243, respectively.

ROP21 is a 749 amino acid polypeptide with a pI of 6.27 and a molecular mass of 82,280.42 Da. A total of 87 positively and 95 negatively charged residues were found in the protein. At 280 nm, the 11,550 have an extinction coefficient of 86,860. The half-life was calculated as 30 h *in vitro* (mammalian reticulocytes), >20 h *in vivo* (yeast), and >10 h *in vivo* (*E. coli*). The instability score was calculated as 48.53, indicating that it is unstable. The aliphatic and GRAVY scores were calculated to be 76.30 and -0.425, respectively.

ROP22 is a 611 amino acid polypeptide with pI of 7.71 and a molecular mass of 68,931.40 Da. A total of 79 positively and 78 negatively charged residues were found in the protein. At 280 nm, the 9585 atoms have an extinction coefficient of 75,400. The half-life was calculated as 30 h *in vitro* (mammalian reticulocytes), >20 h *in vivo* (yeast), and >10 h *in vivo* (*E. coli*). The instability score was calculated as 48.39, indicating that it is unstable. The aliphatic and GRAVY scores were calculated to be 80.64 and -0.383, respectively.

#### 3.2. Multiple sequence alignment

For MSA of ROP7, ROP21 and, ROP22 we used Clustal Omega. In the MSA asterisk in which “\*” means all sequences have the same residue (i.e.; fully conserved), no dots mean no conservation, “.” means some conservation, “:” means partially conserved (i.e.; amino acids with somewhat different properties, and “-” means a gap inserted for better alignment). Results for MSA are shown in Fig. 1.

#### 3.3. Antigenicity and allergenicity

The antigenicity of ROP7, ROP21, and ROP22 was predicted, and the scores are 0.316442, 0.728490 and 0.600087 for Antigenpro and 0.5852, 0.5966 and 0.5859 for Vaxijen. Thus, both web servers showed that ROP7, ROP21 and ROP22 are not allergenic.

#### 3.4. Prediction of post-translational modification sites

The results of the Netphos 3.1 webserver predicted that ROP7 has 47 phosphorylation sites; including 18 threonines, 4 tyrosines, and 25 serines. ROP21 has 79 phosphorylation sites; including 31 threonines, 2 tyrosines, and 46 serines, and ROP22 has 64

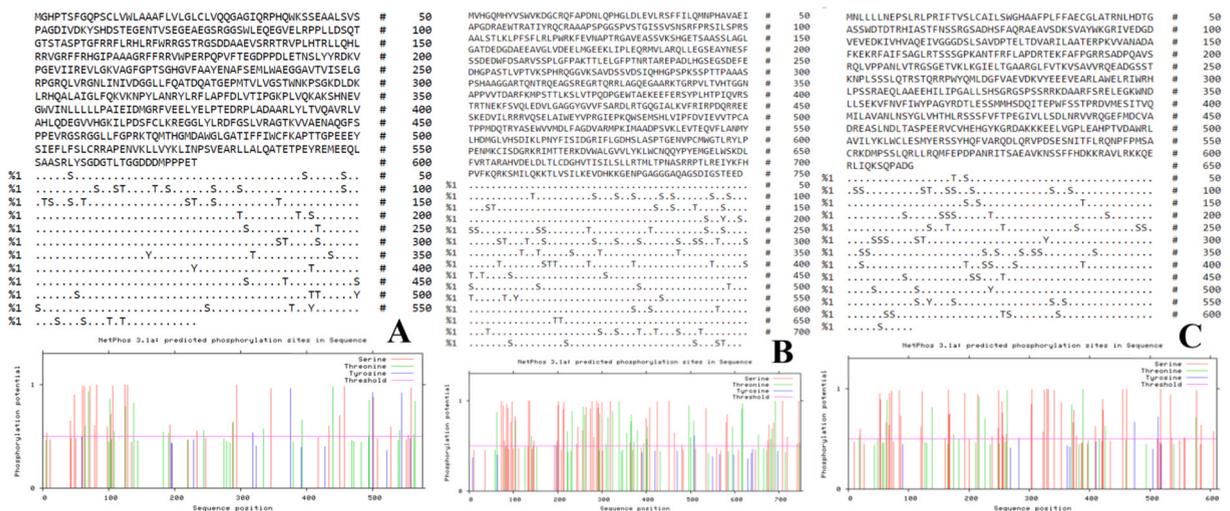


Fig. 2. (a–c). Prediction of phosphorylation sites using the Netphos 2.0 webserver. ROP7 has 47 phosphorylation sites (a). ROP21 has 79 phosphorylation sites (b). ROP22 has 64 phosphorylation sites (c). The predicted sites were for S (serine), T (threonine) and Y (tyrosine).

phosphorylation sites; including 11 threonines, 50 serines, and 3 tyrosines. Results are shown in Fig. 2(a–c).

### 3.5. Prediction of signal peptides, transmembrane domains and isoelectric point

Based on the SignalP-6.0 and TMHMM results, ROP7 and ROP22 were predicted to include a signal peptide at the cleavage site between residues 30–34 and 30–31, with a probability of 0.864365 and 0.591464, respectively. Meanwhile, ROP21 was not predicted to have a signal peptides, hence there is no predicted cleavage site. Results are summarized in Table 1 and Fig. 3(a–c). In addition, one transmembrane domain was predicted in ROP7, but none were predicted for ROP21 or ROP22 as shown in Fig. 4(a–c). The predicted pI values for ROP7, ROP21 and ROP22 were 6.25, 5.99 and 7.05 plotted in Fig. 5(a–c).

### 3.6. Prediction of secondary structure

The secondary structures of ROP7, ROP21 and ROP22 were predicted and visualised using the PSIPRED webserver. ROP7 consists of 56% (322/575) random coil, 33.7% (194/575) alpha helix and 10% (59/575) beta sheet. ROP21 consists of 61.28% (459/749) random coil, 32.44% (243/749) alpha helix, and 6.27% (47/749) beta sheet, ROP22 consists of 49.42% (302/611) random coil, 41.73% (255/611) alpha helix and 8.83% (54/611) beta sheet. Results are shown in Fig. 6(a–c).

Amino acid types can be classified into four types: small nonpolar, hydrophobic, polar, and aromatic plus cysteine. ROP7 contains 36% (211/575) small nonpolar, 22.9% (132/575) hydrophobic, 30.7% (177/575) polar, and 9.56% (55/575) aromatic plus cysteine, ROP21 contains 37.38% (280/749) small nonpolar, 21.89% (164/749) hydrophobic, 32.84% (246/749) polar, and 7.87% (59/749) aromatic plus cysteine, ROP22 contains 34.53% (211/611) small nonpolar, 21.76% (133/749) hydrophobic, 34.69% (212/611) polar, and 9% (55/611) aromatic plus cysteine. Results are shown in Fig. 7(a–c).

### 3.7. Prediction of tertiary structure

PHYRE2 was used for tertiary structure prediction based on the most similar templates. For ROP7, 344 residues were modelled with 100% confidence and 60% coverage, and 35% identity with the most similar template. ROP21, 239 residues were modelled with 100% confidence and 39% coverage, and 24% identity with the most similar template, for ROP22, 258 residues were modelled 100% confidence and 42% coverage, and 21% identity with the most similar template. Results are shown in Fig. 8(a–c).

### 3.8. Prediction of linear B-cell epitopes

The Ellipro web tool was used to predict linear epitopes. ROP7, ROP21, and ROP22 all contain nine linear epitopes. The predicted epitopes are listed in Table 2 with number of residues and score present in every epitope.

### 3.9. Screening of linear B-cell epitopes

ToxinPred was used to predict toxic and non-toxic epitopes. All nine epitopes of ROP7 and ROP21 were found to be non-toxic, while one epitope of ROP22 was toxic. Results are summarized in Table 3.

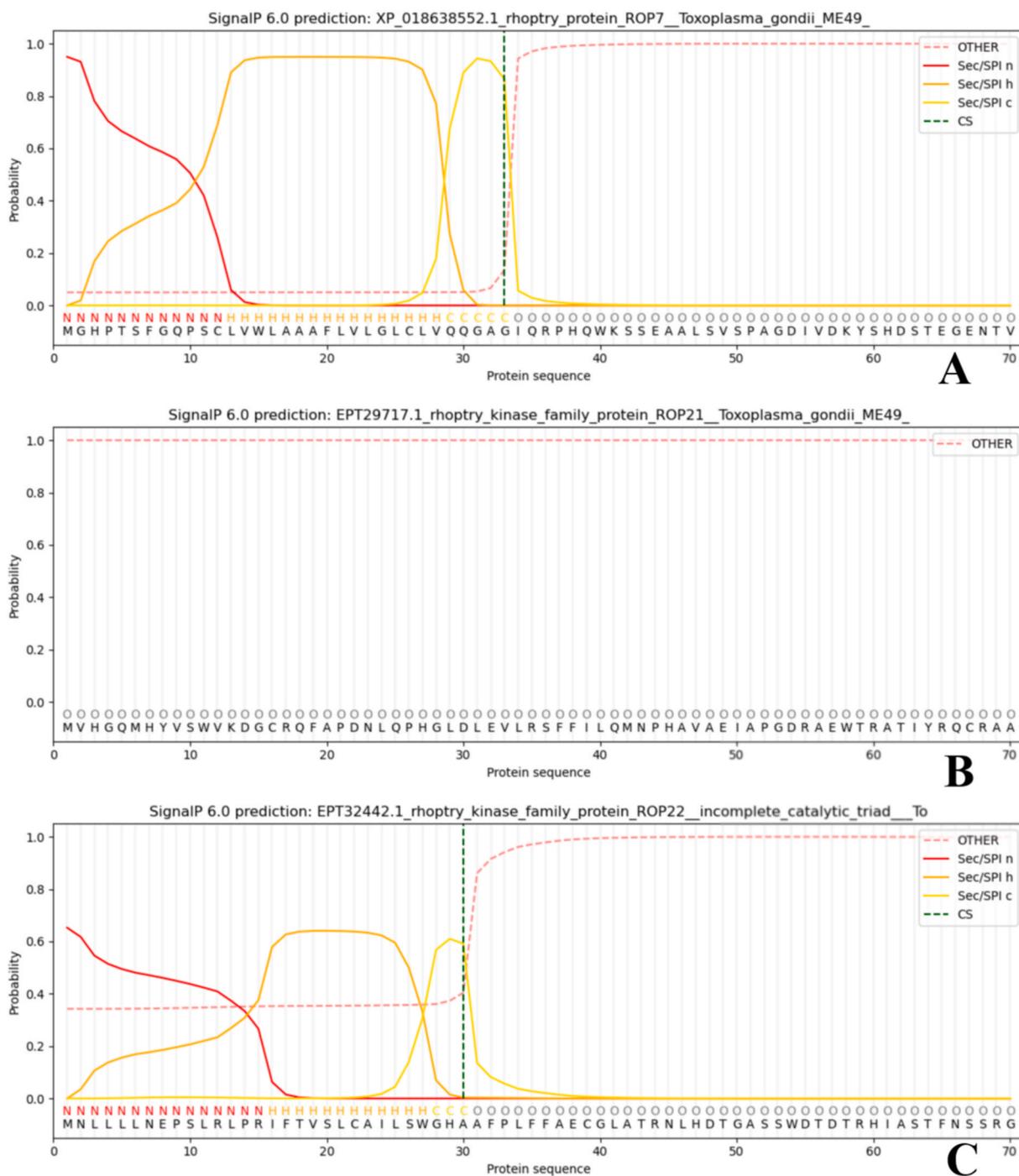
The online tool Vaxijen was then used to check the antigenicity of non-toxic epitopes. ROP7 contains seven probable antigens, ROP21 contains four probable antigens, and ROP22 contain six probable antigens as shown in Table 4. All probable antigens were checked for allergenicity using the Allertop webserver, which showed that ROP7 contains three probable non-allergen epitopes, ROP21 contains one probable non-allergen epitope, and ROP22 contains four non-allergen epitopes Results are summarized in Table 4.

### 3.10. Prediction of immunogenicity of probable epitopes

The immunogenicity of non-allergen linear epitopes was checked using the IEDB webserver, which calculated immunogenic score for every epitope according to immunogenicity, predicting that all the epitopes given in Table 5 can be considered as probable epitopes.

**Table 1**  
Prediction of signal peptides using the SignalP webserver.

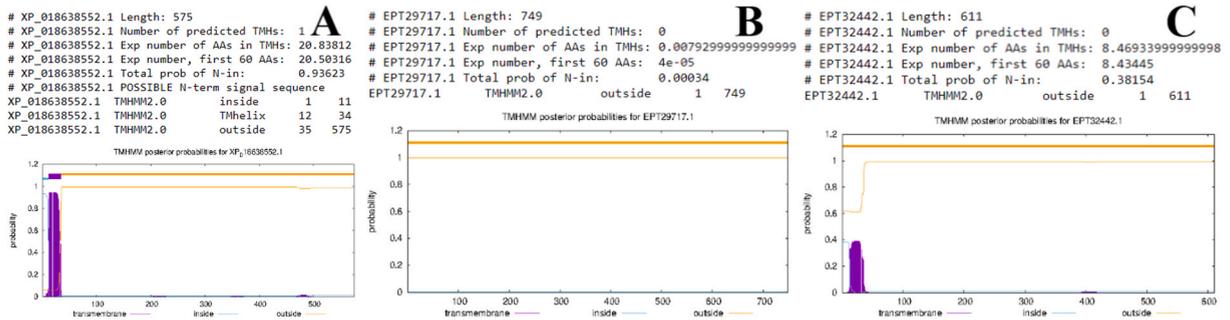
Protein	Cleavage site	Probability	Other	Signal peptide (Sec/SPI)	Lipoprotein signal peptide (Sec/SPII)	TAT signal peptide (Tat/SPI)	TAT lipo-protein signal peptide (Tat/SPII)	Pilin-like signal peptide (Sec/SPIII)
<b>ROP7</b>	30 and 34	0.864365	0.0554	0.9431	0.0007	0.0003	0.0002	0.0003
<b>ROP21</b>	–	–	1.0001	0	0	0	0	0
<b>ROP22</b>	30 and 31	0.591464	0.361	0.6322	0.0056	0.0004	0.0003	0.0005



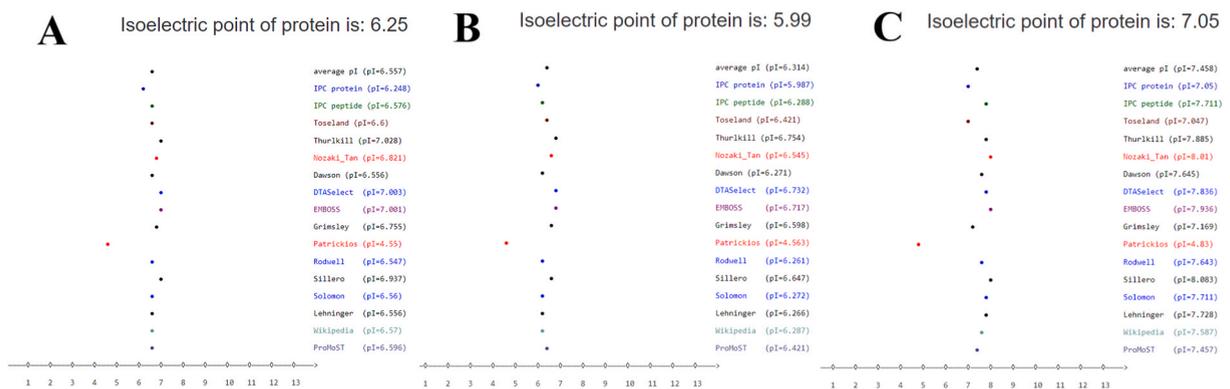
**Fig. 3.** (a–c). Prediction of signal peptides using the Signalp 6.0 webserver. ROP7 has a signal peptide at the cleavage site between residues 30–34 (a). ROP21 has no signal peptides (b). ROP22 has a signal peptide at the cleavage site between residues 30–31. Sec/SPI, Sec/SPII, Tat/SPI, Tat/SPII, Sec/SPIII (c).

#### 4. Discussion

The *T. gondii* parasite affects ~30% of humans, and the best control strategies are immunisation and vaccination. However, most infected humans do not show clinical symptoms, and those with underlying immune deficits may experience severe effects. For instance, toxoplasma encephalitis occurs in ~45% of AIDS (acquired immune deficiency syndrome)-related immunocompromised individuals [33]. According to immunological research on antigens, immune responses are mediated by certain immunogenic epitopes



**Fig. 4.** (a–c). Prediction of transmembrane domains using by TMHMM-2.0 webserver. ROP7 contains one transmembrane domain (a). ROP21 contains no transmembrane domain (b). ROP22 contains no transmembrane domain (c). TMHMM predicts the length of the protein sequence, the number of transmembrane helices (TMHs), the number of amino acids in TMHs, the first 60 amino acids, the probability of N in TMHs, and possible N-terminal signal sequences.



**Fig. 5.** (a–c). Prediction of the isoelectric point (pI) using the isoelectric point calculator (IPC). The pI of ROP7 is 6.25 (a). The pI of ROP21 is 5.99 (b). The pI of ROP22 is 7.05 (c).

rather than the full. Protein antigens can have epitope structures that are detrimental to protective immunity in addition to the structures of epitope used by B, T, cytotoxic T lymphocyte (CTL), and natural killers (NK) cells to mediate immune responses [34].

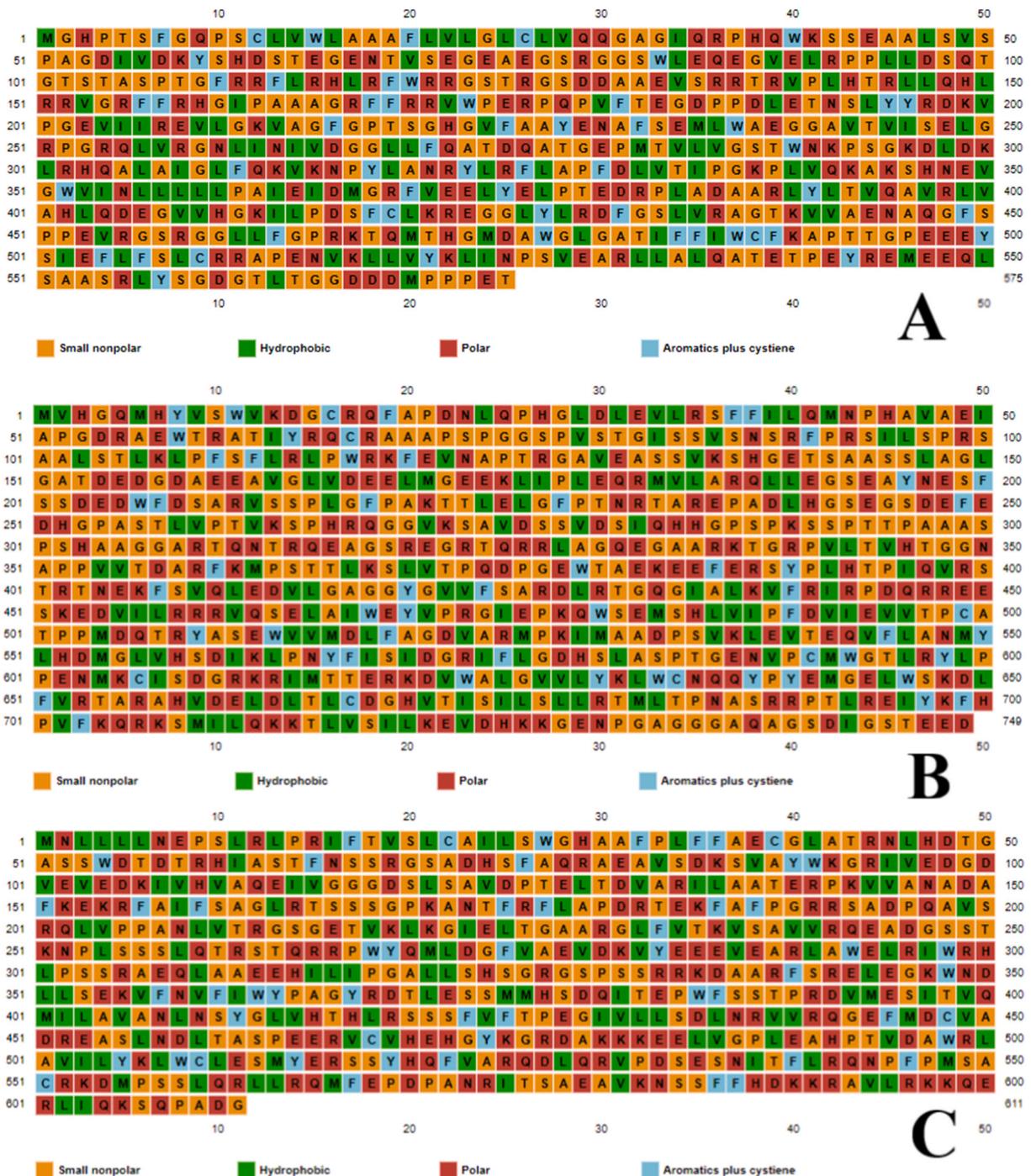
Vaccines are important tools for preserving worldwide health. Traditional vaccine technologies have been used across a wide range of bacterial and viral pathogens. There are numerous instances where traditional vaccine technologies have failed, including infections, viral antigens, and emerging pathogens [35]. The usage of vaccines has improved the quality of life for humans and other animals and saved a numerous lives. Immunoinformatics is a powerful computational method in the study of infectious diseases for investigating features relevant to host-pathogen association. Assigning tasks to uncharacterised genes that can be targeted as candidates in vaccine design using bioinformatics techniques can be effective for pregnant women in vaccine trials and programmes [36]. Studying the antigenic and immunogenic epitopes of *T. gondii* has improved our knowledge of the composition and function of antigens, the interactions between antigens and antibodies, and many other immunological concepts. It has also contributed to the development of novel diagnostic tools and vaccines [34].

ROPs are distinct secretory molecules present in all apicomplexan invasive stages [37]. Toxoplasma ROPs have been subjected to proteomic investigation, revealing many serine-threonine kinases and protein phosphatases [5]. In the present study, several bioinformatics webserver were used to predict and evaluate the various properties and immunogenic epitopes of three proteins (ROP7, ROP21 and, ROP22) of *T. gondii*. ExPasy was used to predict the physicochemical properties of proteins. The protein sequence of ROP7, ROP21, and ROP22 comprise 575, 749 and 611 residues, with MW values of 63395.25, 82580.42, and 68931.40 Da, implying good antigenic nature because antigens with MW of <5–10 kDa are considered weakly immunogenic [38]. The instability scores of ROP7, ROP21 and ROP22 was calculated as 42.43, 48.53 and 48.39, indicating that all three are unstable. The instability index of a protein provides an estimate of the stability in a test tube. A protein whose instability index is smaller than 40 is predicted as stable, a value above 40 predicts that the protein may be unstable (ProtParam, n.d.). The calculated aliphatic indices and GRAVY scores of ROP7, ROP21, ROP22 were 86.16 and -0.243, 76.30 and -0.425, and 80.64 and -0.383, respectively. Negative GRAVY scores indicate a hydrophilic nature, consistent with strong interactions with the water molecules surrounding membranes. Target proteins are more stable over a higher temperature range if the aliphatic index is higher [39].

Many sequence analysis techniques start with MSA [20]. In the present study, MSA was performed by using Clustal Omega. This revealed high sequence identity and similarity between all three proteins shown in Fig. 1. A favourable vaccine possesses probable

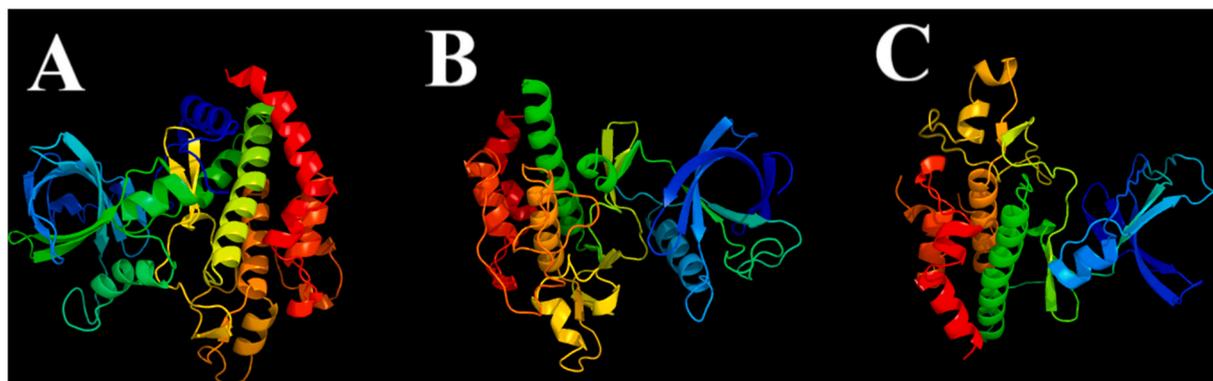


**Fig. 6.** (a–c). Secondary structure prediction using PSIPRED. ROP7 consists of 56% random coil, 33.7% alpha helix, and 10% beta sheet (a). ROP21 consists of 61.28% random coil, 32.44% alpha helix, and 6.27% beta sheet (b). ROP22 contains 49.42% random coil, 41.73% alpha helix, and 8.83% beta sheet (c). Graphical illustrations show strands, helices and coils.



**Fig. 7.** (a–c). Prediction of amino acid type by PSIPRED. ROP7 contains 36% small nonpolar, 22.9% hydrophobic, 30.7% polar, and 9.56% aromatic plus cysteine amino (a). ROP21 contains 37.38% small nonpolar, 21.89% hydrophobic, 32.84% polar, and 7.87% aromatic plus cysteine amino acid residues (b). ROP22 contains 34.53% small nonpolar, 21.76% hydrophobic, 34.69% polar, and 9% aromatic plus cysteine (c). Graphical illustrations show small nonpolar, hydrophobic, polar, and aromatics plus cysteine amino acids.

antigenicity, while lacking allergenicity. In this study, Antigenpro, Vaxijen and Allertop webserver were used, and they predicted that ROP7, ROP21, ROP22 are probable immunogenic and non-allergenic proteins. Antigenpro predicted antigenicity probability values for ROP7, ROP21 and ROP22 of 0.316442, 0.728490, and 0.600087. In immunology, virology, and bacteriology, the identification of antigen proteins that can cause a large humoral immune response is crucial. This is also important for a variety of practical applications, including vaccine design and diagnostics [21].



**Fig. 8.** (a–c). Tertiary structure prediction using PHYRE2. 3D structure of ROP7 (a). 3D structure of ROP21 (b). 3D structure of ROP22 (c). 3D models show helices, loops and strands.

**Table 2**

Prediction of linear epitopes using the Ellipro IEDB webserver.

Protein	Chain	Start	End	Peptide	No. of residues	Score
ROP7	A	453	470	PETTAERMLPFGQHHPTL	18	0.812
ROP7	A	223	260	EYAADSLVSTSLWNTGQ PFRVESELGERPRTLVRG	35	0.788
ROP7	A	275	281	TDQETGE	7	0.763
ROP7	A	334	356	VKDPQVLS	8	0.763
ROP7	A	289	300	YFTEAIK	8	0.748
ROP7	A	313	317	RGIKN	5	0.707
ROP7	A	493	518	PNTDDAALGGSEWIF RSCKNIPQVVR	26	0.674
ROP7	A	190	201	NLYFQGFRGTDP	12	0.627
ROP7	A	437	444	RDGASAVS	8	0.594
ROP21	A	252	258	RDVERLE	7	0.805
ROP21	A	212	235	ATVVPQNAETTVDLSLSQ GERCLK	24	0.776
ROP21	A	311	329	AVQSQPPFAQLSPGYA	16	0.772
ROP21	A	413	418	KVGTRG	6	0.77
ROP21	A	508	514	NFDRRRR	7	0.76
ROP21	A	353	364	FVYVFRGDEGIL	12	0.746
ROP21	A	430	440	REFLNASTATF	11	0.719
ROP21	A	267	274	MGAENSRS	8	0.687
ROP21	A	470	501	IKGSWKRPRLRVPGTDSLAFGSCPTLPDFVKT	32	0.66
ROP22	A	462	481	PELEARRATISYHRDRRT LM	20	0.822
ROP22	A	344	366	KDPQKKMIRVMWVLSR	17	0.774
ROP22	A	284	290	TDETGE	6	0.77
ROP22	A	322	326	RGIKN	5	0.734
ROP22	A	232	269	EMQSAADSLVSTSLWNGQPFRVESELGERPRTLVRG	36	0.721
ROP22	A	385	397	LSHSSTHKSLVHH	13	0.714
ROP22	A	300	309	TEPPSNAIK	9	0.695
ROP22	A	503	529	PIGGSEWIFRSCKNIPQVRA	21	0.667
ROP22	A	446	453	RDGARVVS	8	0.621

PTM of proteins is a crucial element of cellular regulatory mechanisms that modify the physical and chemical characteristics of proteins including folding, stability, conformation and activity, which in turn affects how they function [6]. In this study, the Netphos 2.0 webserver was used to identify PTMs of ROP7, ROP21, and ROP22. ROP7 has 47 phosphorylation sites, ROP7 has 79 phosphorylation sites, and ROP7 has 64 phosphorylation sites shown in Fig. 2(a–c). SignalP 6.0 and TMHMM webservers were used to predict signal peptides and transmembrane domains. ROP7 and ROP22 include a signal peptide at cleavage sites between residues 30–34 and 30–31, with probability values of 0.864365 and 0.591464, respectively. ROP21 does not have a signal peptide. ROPs have an N-terminal signal sequence and C-terminal hydrophobic region assumed to be transmembrane domains [21]. In this study, one transmembrane domain was predicted for ROP7, and no transmembrane domains were predicted for ROP21 and ROP22 shown in Fig. 4(a–c).

Numerous proteomic processes and biochemical phenomena can be predicted with reasonable accuracy using pI values estimated based on the amino acid sequence. Under physiological conditions, the pI value indicates whether a protein will carry a net positive and negative charge. Proteins with pI > 7.0 are basic, and proteins with pI < 7.0 are acidic [40]. In this study, the pI of ROP7, ROP21

**Table 3**  
Prediction of non-toxic epitopes by the ToxinPred webserver.

Protein Name	Peptide sequence	Svm score	Prediction	Hydro-phobicity	Hydro-pathicity	Hydro-philicity	Charge	Mol weight
ROP7	PETTAERMLP	-1.47	Non-Toxic	-0.16	-0.81	-0.05	0.00	2062.60
	FGQHHPTL							
ROP7	EYAADSLVST	-1.92	Non-Toxic	-0.22	-0.58	0.19	-1.00	3922.84
	SLWNTGQPFR							
	VESELGERPRT							
	LVRG							
ROP7	TDQETGE	-0.97	Non-Toxic	-0.41	-2.26	1.20	-3.00	778.82
ROP7	VKDPQVLS	-1.09	Non-Toxic	-0.15	-0.14	0.21	0.00	885.14
ROP7	YFTERAIK	-1.28	Non-Toxic	-0.26	-0.60	0.19	1.00	1027.29
ROP7	RGIKN	-0.80	Non-Toxic	-0.52	-1.56	0.88	2.00	586.76
ROP7	PNTDDAALGG	-0.99	Non-Toxic	-0.20	-0.65	0.15	0.00	2872.58
	SEWIFRSCKNIP							
	QPVR							
ROP7	NLYFQGFRGTDP	-0.97	Non-Toxic	-0.16	-0.83	-0.26	0.00	1414.71
ROP7	RDGASAVS	-0.100	Non-Toxic	-0.23	-0.28	0.51	0.00	761.88
ROP21	RDVERLE	-1.15	Non-Toxic	-0.63	-1.64	1.67	-1.00	916.09
ROP21	ATVWPQNAETT	-1.36	Non-Toxic	-0.22	-0.69	0.22	0.00	2672.35
	VDSLSSQGERKL							
	K							
ROP21	AVQSQPFAQLS	-1.21	Non-Toxic	-0.01	-0.15	-0.52	0.00	1661.07
	PGYA							
ROP21	KVGTRG	-0.70	Non-Toxic	-0.36	-0.95	0.68	2.00	616.80
ROP21	NFDRRRR	-1.01	Non-Toxic	-1.11	-3.17	1.81	3.00	1019.21
ROP21	FVYVFRGDEGIL	-1.28	Non-Toxic	0.07	0.72	-0.41	-1.00	1414.80
ROP21	REFLNASTATF	-0.69	Non-Toxic	-0.13	-0.06	-0.19	0.00	1256.52
ROP21	MGAENSRS	-0.44	Non-Toxic	-0.36	-1.23	0.62	0.00	851.00
ROP21	IKGSWKRPRLRV	-1.23	Non-Toxic	-0.14	-0.23	0.03	3.00	3462.49
	PGTDSLAFGSCT							
	PLPDFVKT							
ROP22	PELEARRATISYH	-1.60	Non-Toxic	-0.46	-1.19	0.65	2.50	2472.10
	RDRRTLML							
ROP22	KDPQKKMIRVM	-0.67	Non-Toxic	-0.38	-0.85	0.52	5.00	2143.94
	WVLSR							
ROP22	TDETGE	-0.95	Non-Toxic	-0.36	-0.25	1.37	-3.00	650.67
ROP22	EMQSAADSLVSTS	-1.90	Non-Toxic	-0.26	-0.69	0.31	1.00	4573.74
	LWNGQPFRVESEL							
	GERPRTLVRG							
ROP22	LSHSSTHKSIVHH	-1.30	Non-Toxic	-0.18	-0.68	-0.25	3.00	1469.82
ROP22	TEPPSNAIK	-0.80	Non-Toxic	-0.22	-1.03	0.42	0.00	956.18
ROP22	PIGGSEWIFRSCKN	-1.18	Non-Toxic	-0.14	-0.32	-0.06	2.00	2356.05
	IPQPVRA							
ROP22	RDGARVVS	-1.12	Non-Toxic	-0.38	-0.44	0.72	1.00	859.05

and ROP22 was 6.25, 5.99 and 7.05, respectively.

Secondary structure and amino acid types were predicted by the PSIPRED webserver. In a polypeptide chain, the pattern of hydrogen bonding between carboxyl oxygen and amino hydrogen atoms determines the secondary structure, with alpha helices and beta sheets the most common types [41]. In this study, the three major secondary structural elements in the ROP7 protein are 56% random coil, 33.7% alpha helix, and 10% beta sheet. The ROP21 protein consist of 61.28% random coil, 32.44% alpha helix, and 6.27% beta sheet. The ROP22 protein contains 49.42% random coil, 41.73% alpha helix, and 8.83% extended sheet. Fig. 6. Amino acid type prediction is important to understand the nature of a protein. In this study, four amino acid types (small nonpolar, Hydrophobic, Polar, and Aromatic plus cysteine) were explored for ROP7, ROP21 and ROP22, and differences were noted.

It is well known that the function of a protein depend on its structure. In this study, tertiary structures of ROP7, ROP21 and ROP22 were predicted by PHYRE2 webserver. For ROP7, 344 residues were modelled with 100% confidence and 60% coverage, and 35% identity with the most similar template. For ROP21, 239 residues were modelled with 100% confidence and 39% coverage, and 24% similarity with the most similar template protein. For ROP22, 258 residues were modelled with 100% confidence and 42% coverage, and 21% identity with the most similar template.

A critical step in the development of epitope-based vaccines, therapeutic antibodies, and diagnostic tools is the identification of B-cell epitopes. Currently, epitope-based antibodies are the most promising class of biopharmaceuticals [42]. In this study B-cell linear epitopes were predicted using IEDB and Ellipro webserver. Ellipro, derived from Ellipsoid and Protrusion, uses a modified method for predicting continuous B-cell epitopes in protein regions projecting from the globular surface of antigens combined with a residue clustering algorithm for predicting discontinuous epitopes from amino acid sequence [30].

All pathogenic alterations caused by the accumulation, mis localisation, and multimerisation of disease-specific proteins are collectively referred to as protein toxicity, one of the main pathogenic pathways for most neurodegenerative illnesses, but the specifics

**Table 4**  
Prediction of antigenicity and allergenicity of non-toxic epitopes.

Protein	Peptide	Antigenicity	Allergenicity
ROP7	EYAADSLVSTSLWNTGQPFR	0.6194	Probable allergen
	VESELGERPRTLVRG		
ROP7	YFTERAIK	1.2032	Non-allergen
ROP7	NLYFQGFRGTDP	0.8924	Non-allergen
ROP7	PNTDDAALGGSEWIFRSCKN	0.6111	Probable allergen
	IPQPVR		
ROP7	TDQETGE	0.9401	Probable allergen
ROP7	RDGASAVS	0.5750	Probable allergen
ROP7	VKDPQVLS	0.6342	Non-allergen
ROP21	KVGTRG	3.4472	Non-allergen
ROP21	MGAENSRS	0.7091	Probable allergen
ROP21	ATVWPQNAETTVDLSLQSG	0.5915	Probable allergen
	ERKLK		
ROP21	IKGSWKRPSLRVPGTDSLAF	0.8577	Probable allergen
	GSCTPLPDFVKT		
ROP22	PELEARRATISYHRDRRTLM	0.6916	Non-allergen
ROP22	RDGARVVS	0.8556	Probable allergen
ROP22	EMQSAADSLVSTSLWNGQP	0.8758	Probable allergen
	FRVESELGERPRTLVRG		
ROP22	TDETGE	1.0113	Probable allergen
ROP22	TEPPSNAIK	1.0141	Non-allergen
ROP22	KDPQKKMIRVMWVLSR	0.7415	Non-allergen

**Table 5**  
Prediction of immunogenicity score of probable epitopes using the immunogenicity IEDB webserver.

Protein	Peptide	Length	Score
ROP7	YFTERAIK	8	0.3129
ROP7	NLYFQGFRGTDP	12	0.27882
ROP7	VKDPQVLS	8	-0.08726
ROP21	KVGTRG	6	0.10046
ROP22	PELEARRATISYHRDRRTLM	20	0.50575
ROP22	TDETGE	6	0.10456
ROP22	TEPPSNAIK	9	-0.07117
ROP22	KDPQKKMIRVMWVLSR	17	-0.72342

of protein toxicity remain unknown [43]. Thus, it is important to choose non-toxic protein epitopes for vaccine and drug development. In this study, the ToxinPred webserver was used to assess the toxicity of epitopes, which showed that all nine epitopes of ROP7 and ROP21 were non-toxic, while one epitope of ROP22 was toxic. To choose the best candidate epitopes of ROP proteins, antigenicity and allergenicity were checked using online Vaxijen and Allertop. Vaxijen predicted that ROP7 contains seven probable antigens, ROP21 contains four probable antigens, and ROP22 contain six probable antigens. Allertop predicted that ROP7 contains three probable non-allergen epitopes, ROP21 contains one probable non-allergen epitope, and ROP22 contains four non-allergen epitopes.

In this study, the final step was to explore the immunogenicity of probable epitopes and to calculate Class I immunogenicity scores using the IEDB web tool. The immunogenicity of a peptide major histocompatibility complex (MHC) is predicted by this tool using the properties of the amino acids and their positions within the peptide. This server gave scores for every probable epitope of ROP7, ROP21 and ROP22 according to their immunogenicity as summarized in Table 5. It is well known that a higher score indicates a greater probability of eliciting an immune response. Based on evaluation by all tools, three immunogenic epitopes of ROP7, one immunogenic epitope of ROP21, and four immunogenic epitopes of ROP22 were predicted.

## 5. Conclusion

The goal of this study was to use bioinformatics to analyse ROP proteins (ROP7, ROP21, ROP22) of *T. gondii*, and identify candidate immunogenic and antigenic epitopes. Several candidate epitopes of ROP proteins were identified using in silico approaches, indicating that these methods can be useful for immunogenic epitope prediction. All identified epitopes had good immunogenic scores. The protein with the highest immunogenic score (0.50575) was ROP22. ROP7 had two high-scoring epitopes with scores of 0.3129 and 0.27882. ROP21 had one immunogenic epitope with a score of 0.10046. The predicted immunogenic epitopes for ROP7, ROP21 and ROP22 suggests they may serve as candidate proteins for vaccine development for toxoplasmosis.

## Declarations

### Author contribution statement

Haroon Ahmed; Jianping Cao: conceived and designed the experiments; Fariha Ayub; Xu Wang collected, analysed and interpreted the data; Performed the experiments; Fariha Ayub wrote the paper; Haroon Ahmed; Tehreem Sohail; Figen Celik; Khuram Shahzad; Xu Wang; Sami Simsek; Jianping Cao: revised the paper.

### Funding statement

This work was supported by grants from the National Natural Science Foundation of China [grant numbers 81971969 and 82272369 to J.C.]. The funders had no role in the study design, the data collection and analysis, the decision to publish, or the preparation of the manuscript.

### Data availability statement

Not applicable.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

## References

- [1] Z. Zhang, Y. Li, Q. Xie, P. Li, X. Nan, L. Kong, D. Zeng, Z. Ding, S. Wang, The molecular characterization and immunity identification of rhoptry protein 22 of *Toxoplasma gondii* as a DNA vaccine candidate against toxoplasmosis, *J. Eukaryot. Microbiol.* 66 (1) (2019) 147–157, <https://doi.org/10.1111/jeu.12639>.
- [2] R.E. Lyons, R. McLeod, C.W. Roberts, *Toxoplasma gondii* tachyzoite-bradyzoite interconversion, *Trends Parasitol.* 18 (5) (2002) 198–201, [https://doi.org/10.1016/s1471-4922\(02\)02248-1](https://doi.org/10.1016/s1471-4922(02)02248-1).
- [3] A. Camejo, D.A. Gold, D. Lu, K. McFetridge, L. Julien, N. Yang, K.D.C. Jensen, J.P.J. Saeij, Identification of three novel *Toxoplasma gondii* rhoptry proteins, *Int. J. Parasitol.* 44 (2) (2014) 147–160, <https://doi.org/10.1016/j.ijpara.2013.08.002>.
- [4] H. Długoska, *Toxoplasma* rhoptries: unique secretory organelles and source of promising vaccine proteins for immunoprevention of toxoplasmosis, *J. Biomed. Biotechnol.* 2008 (2008), 632424, <https://doi.org/10.1155/2008/632424>.
- [5] P.J. Bradley, L.D. Sibley, Rhoptries: an arsenal of secreted virulence factors, *Curr. Opin. Microbiol.* 10 (6) (2007) 582–587, <https://doi.org/10.1016/j.mib.2007.09.013>.
- [6] H. El Hajj, E. Demy, J. Poncet, M. Lebrun, B. Wu, N. Galéotti, M.N. Fourmaux, O. Mercereau-Puijalon, H. Vial, G. Labesse, J.F. Dubremetz, The ROP2 family of *Toxoplasma gondii* rhoptry proteins: proteomic and genomic characterization and molecular modeling, *Proteomics* 6 (21) (2006) 5773–5784, <https://doi.org/10.1002/pmic.200600187>.
- [7] E. Talevich, N. Kannan, Structural and evolutionary adaptation of rhoptry kinases and pseudokinases, a family of coccidian virulence factors, *BMC Evol. Biol.* 13 (2013) 117, <https://doi.org/10.1186/1471-2148-13-117>.
- [8] L. Wang, G. Lu, A. Zhou, Y. Han, J. Guo, H. Zhou, H. Cong, S. He, Evaluation of immune responses induced by rhoptry protein 5 and rhoptry protein 7 DNA vaccines against *Toxoplasma gondii*, *Parasite Immunol.* 38 (4) (2016) 209–217, <https://doi.org/10.1111/pim.12306>.
- [9] H.E. Hajj, M. Lebrun, M.N. Fourmaux, H. Vial, J.F. Dubremetz, Characterization, biosynthesis and fate of ROP7, a ROP2 related rhoptry protein of *Toxoplasma gondii*, *Mol. Biochem. Parasitol.* 146 (1) (2006) 98–100, <https://doi.org/10.1016/j.molbiopara.2005.10.011>.
- [10] Z. Zhang, Y. Li, M. Wang, Q. Xie, P. Li, S. Zuo, L. Kong, C. Wang, S. Wang, Immune protection of rhoptry protein 21 (ROP21) of *Toxoplasma gondii* as a DNA vaccine against toxoplasmosis, *Front. Microbiol.* 9 (2018), <https://doi.org/10.3389/fmicb.2018.00909>.
- [11] J. Kur, L. Holec-Gasior, E. Hiszczyńska-Sawicka, Current status of toxoplasmosis vaccine development, *Expert Rev. Vaccine* 8 (6) (2009) 791–808, <https://doi.org/10.1586/erv.09.27>.
- [12] M. Antczak, K. Dzitko, H. Długoska, Human toxoplasmosis—Searching for novel chemotherapeutics, *Biomedicine & Pharmacotherapy [Biomed. Pharmacother.* 82 (2016) 677–684, <https://doi.org/10.1016/j.biopha.2016.05.041>.
- [13] G. Saadatinia, M. Golkar, A review on human toxoplasmosis, *Scand. J. Infect. Dis.* 44 (11) (2012) 805–814, <https://doi.org/10.3109/00365548.2012.693197>.
- [14] J.C. Boothroyd, J.-F. Dubremetz, Kiss and spit: the dual roles of *Toxoplasma* rhoptries, *Nat. Rev. Microbiol.* 6 (1) (2008) 79–88, <https://doi.org/10.1038/nrmicro1800>.
- [15] P. Romano, R. Giugno, A. Pulvirenti, Tools and collaborative environments for bioinformatics research, *Briefings Bioinf.* 12 (6) (2011) 549–561, <https://doi.org/10.1093/bib/bbr055>.
- [16] R.M. Fereig, H.H. Abdelbaky, Comparative study on *Toxoplasma gondii* dense granule protein 7, peroxiredoxin 1 and 3 based on bioinformatic analysis tools, *Ger. J. Microbiol.* 2 (1) (2022) 30–38, <https://doi.org/10.51585/gjm.2022.1.0013>.
- [17] P.-A. Binz, J. Shofstahl, J.A. Vizcaino, H. Barsnes, R.J. Chalkley, G. Menschaert, E. Alpi, K. Clauser, J.K. Eng, L. Lane, S.L. Seymour, L.F.H. Sánchez, G. Mayer, M. Eisenacher, Y. Perez-Riverol, E.A. Kapp, L. Mendoza, P.R. Baker, A. Collins, et al., Proteomics standards initiative extended FASTA format, *J. Proteome Res.* 18 (6) (2019) 2686–2692, <https://doi.org/10.1021/acs.jproteome.9b00064>.
- [18] D.A. Benson, M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, J. Ostell, K.D. Pruitt, E.W. Sayers, GenBank, *Nucleic Acids Res.* 46 (D1) (2018) D41–D47, <https://doi.org/10.1093/nar/gkx1094>.
- [19] M.R. Wilkins, E. Gasteiger, A. Bairoch, J.-C. Sanchez, K.L. Williams, R.D. Appel, D.F. Hochstrasser, Protein identification and analysis tools in the ExPASy server, in: *2-D Proteome Analysis Protocols*, vol. 112, Humana Press, 2003, pp. 531–552.
- [20] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, *Mol. Syst. Biol.* 7 (1) (2011) 539, <https://doi.org/10.1038/msb.2011.75>.
- [21] C.N. Magnan, M. Zeller, M.A. Kayala, A. Vigil, A. Randall, P.L. Felgner, P. Baldi, High-throughput prediction of protein antigenicity using protein microarray data, *Bioinformatics (Oxford, England)* 26 (23) (2010) 2936–2943, <https://doi.org/10.1093/bioinformatics/btq551>.
- [22] I.A. Doytchinova, D.R. Flower, Vaxijen: a server for prediction of protective antigens, tumour antigens and subunit vaccines, *BMC Bioinf.* 8 (1) (2007) 4, <https://doi.org/10.1186/1471-2105-8-4>.
- [23] I. Dimitrov, D.R. Flower, I. Doytchinova, Allertop—a server for in silico prediction of allergens, *BMC Bioinf.* 14 (Suppl 6) (2013) S4, <https://doi.org/10.1186/1471-2105-14-S6-S4>.

- [24] N. Blom, T. Sicheritz-Pontén, R. Gupta, S. Gammeltoft, S. Brunak, Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence, *Proteomics* 4 (6) (2004) 1633–1649, <https://doi.org/10.1002/pmic.200300771>.
- [25] J.J. Almagro Armenteros, K.D. Tsirigos, C.K. Sønderby, T.N. Petersen, O. Winther, S. Brunak, G. von Heijne, H. Nielsen, SignalP 5.0 improves signal peptide predictions using deep neural networks, *Nat. Biotechnol.* 37 (4) (2019) 420–423, <https://doi.org/10.1038/s41587-019-0036-z>.
- [26] A. Krogh, B. Larsson, G. von Heijne, E.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, *J. Mol. Biol.* 305 (3) (2001) 567–580, <https://doi.org/10.1006/jmbi.2000.4315>.
- [27] L.P. Kozłowski, IPC - isoelectric point calculator, *Biol. Direct* 11 (1) (2016) 55, <https://doi.org/10.1186/s13062-016-0159-9>.
- [28] L.J. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server, *Bioinformatics (Oxford, England)* 16 (4) (2000) 404–405, <https://doi.org/10.1093/bioinformatics/16.4.404>.
- [29] L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J.E. Sternberg, The PHYRE2 web portal for protein modeling, prediction and analysis, *Nat. Protoc.* 10 (6) (2015) 845–858, <https://doi.org/10.1038/nprot.2015.053>.
- [30] J. Ponomarenko, H.-H. Bui, W. Li, N. Fusseder, P.E. Bourne, A. Sette, B. Peters, Ellipro: a new structure-based tool for the prediction of antibody epitopes, *BMC Bioinf.* 9 (1) (2008) 514, <https://doi.org/10.1186/1471-2105-9-514>.
- [31] S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, Open Source Drug Discovery Consortium, G.P.S. Raghava, In silico approach for predicting toxicity of peptides and proteins, *PLoS One* 8 (9) (2013), e73957, <https://doi.org/10.1371/journal.pone.0073957>.
- [32] J.J.A. Calis, M. Maybeno, J.A. Greenbaum, D. Weiskopf, A.D. De Silva, A. Sette, C. Keşmir, B. Peters, Properties of MHC class I presented peptides that enhance immunogenicity, *PLoS Comput. Biol.* 9 (10) (2013), e1003266, <https://doi.org/10.1371/journal.pcbi.1003266>.
- [33] M.E. Grigg, J. Ganatra, J.C. Boothroyd, T.P. Margolis, Unusual abundance of atypical strains associated with human ocular toxoplasmosis, *J. Infect. Dis.* 184 (5) (2001) 633–639, <https://doi.org/10.1086/322800>.
- [34] Y. Wang, G. Wang, J. Cai, H. Yin, Review on the identification and role of *Toxoplasma gondii* antigenic epitopes, *Parasitol. Res.* 115 (2) (2016) 459–468, <https://doi.org/10.1007/s00436-015-4824-1>.
- [35] M.S. Gebre, L.A. Brito, L.H. Tostanoski, D.K. Edwards, A. Carfi, D.H. Barouch, Novel approaches for vaccine development, *Cell* 184 (6) (2021) 1589–1603, <https://doi.org/10.1016/j.cell.2021.02.030>.
- [36] A.N. Oli, W.O. Obialor, M.O. Ifeanyi-chukwu, D.C. Odimegwu, J.N. Okoyeh, G.O. Emechebe, S.A. Adejumo, G.C. Ibeanu, Immunoinformatics and vaccine development: an overview, *ImmunoTargets Ther.* 9 (2020) 13–30, <https://doi.org/10.2147/ITT.S241064>.
- [37] J.F. Dubremetz, Rhoptries are major players in *Toxoplasma gondii* invasion and host cell interaction, *Cell Microbiol.* 9 (4) (2007) 841–848, <https://doi.org/10.1111/j.1462-5822.2007.00909.x>.
- [38] J.A. Berzofsky, in: Immunogenicity and Antigen Structure. Fundamental Immunology, third ed., 1993, pp. 235–282. <https://cir.nii.ac.jp/crid/1573950399094337792>.
- [39] H. Majidani, A. Dalimi, F. Ghaffarifar, M. Pirestani, A.D. Ghaffari, Computational probing of *Toxoplasma gondii* major surface antigen 1 (SAG1) for enhanced vaccine design against toxoplasmosis, *Microb. Pathog.* 147 (104386) (2020), 104386, <https://doi.org/10.1016/j.micpath.2020.104386>.
- [40] B.D. Halligan, ProMoST: a tool for calculating the pI and molecular mass of phosphorylated and modified proteins on two-dimensional gels, *Methods Mol. Biol. (Clifton, N.J.)* 527 (2009) 283–298, [https://doi.org/10.1007/978-1-60327-834-8\\_21](https://doi.org/10.1007/978-1-60327-834-8_21), ix.
- [41] R.Y. Yada, R.L. Jackman, S. Nakai, Secondary structure prediction and determination of proteins—a review, *Int. J. Pept. Protein Res.* 31 (1) (1988) 98–108, <https://doi.org/10.1111/j.1399-3011.1988.tb00011.x>.
- [42] L. Potocnakova, M. Bhide, L.B. Pulzova, An introduction to B-cell Epitope mapping and in silico Epitope prediction, *J. Immunol. Res.* 2016 (2016), 6760830, <https://doi.org/10.1155/2016/6760830>.
- [43] C.G. Chung, H. Lee, S.B. Lee, Mechanisms of protein toxicity in neurodegenerative diseases, *Cell. Mol. Life Sci.: CMLS* 75 (17) (2018) 3159–3180, <https://doi.org/10.1007/s00018-018-2854-4>.