CLINICAL AND EXPERIMENTAL VACCINE RESEARCH

Clin Exp Vaccine Res 2020;9:146-158
https://doi.org/10.7774/cevr.2020.9.2.146
pISSN 2287-3651 • eISSN 2287-366X

In-depth computational analysis of calcium-dependent protein kinase 3 of *Toxoplasma gondii* provides promising targets for vaccination

Hamidreza Majidiani[1]*, Shahrzad Soltani[2]*, Ali Dalir Ghaffari[3], Mohamad Sabaghan[4], Ali Taghipour[3], Masoud Foroutan[2]

[1]Zoonotic Diseases Research Center, Ilam University of Medical Sciences, Ilam; [2]Abadan Faculty of Medical Sciences, Abadan; [3]Department of Parasitology, Faculty of Medical Sciences, Tarbiat Modares University, Tehran; [4]Behbahan Faculty of Medical Sciences, Behbahan, Iran

Received: May 30, 2020
Revised: June 23, 2020
Accepted: July 28, 2020

Corresponding author: Masoud Foroutan, PhD
Abadan Faculty of Medical Sciences, Abadan, 6313833177, Iran
Tel: +98-61-53265361, Fax: +98-61-53384008
E-mail: masoud_foroutan_rad@yahoo.com

No potential conflict of interest relevant to this article was reported.

This study was financially supported by the Behbahan Faculty of Medical Sciences, Behbahan, Iran (Grant no., 99013).

*These authors contributed equally to this work.

Purpose: The *Toxoplasma gondii* calcium-dependent protein kinase-3 (CDPK3) is a key enzyme for parasite egress, control of calcium-dependent permeabilization in parasitophorous vacuole membrane and tissue cyst formation. In this study, we comprehensively explored the bioinformatics features of this protein to improve vaccine design against *T. gondii.*

Materials and Methods: Various web servers were employed for the analysis of physicochemical properties, post-translational modifications, localization in the subcellular milieu, secondary and tertiary structures, as well as B-cell, major histocompatibility complex (MHC)-binding and cytotoxic T-lymphocyte (CTL) epitopes.

Results: This protein was a 537 amino acid antigenic and non-allergenic molecule with a molecular weight of 60.42 kDa, a grand average of hydropathicity score of -0.508, and aliphatic index of 79.50. There exists 46.74% alpha helix, 12.48% extended strand, and 40.78% random coil in the secondary structure. Ramachandran plot of the refined model demonstrated 99.3%, 0.7%, and 0.0% of residues in the favored, allowed and outlier areas, respectively. Besides, various potential B-cell (continuous and conformational), MHC-binding and CTL epitopes were predicted for *Toxoplasma* CDPK3 protein.

Conclusion: This article provides a foundation for further investigations, and laid a theoretical basis for the development of an appropriate vaccine against *T. gondii* infection.

Keywords: *Toxoplasma*, Calcium-dependent protein kinase-3, Bioinformatics, *In silico*, Vaccine

Introduction

The intracellular ubiquitous apicomplexan, *Toxoplasma gondii* (*T. gondii*), is the causative agent of toxoplasmosis, a significant zoonosis with harsh outcomes in livestock and human being [1]. The feline species are the only definitive hosts, which shed unsporulated oocysts in their fecal matter and contaminate the surroundings [2]. Approximately, one-third of the world population is seropositive regarding *T. gondii* infection [3]. The protozoan employs different transmission routes, including: food/water sources contaminated with sporulated oocysts, ingestion of meat products containing tissue cysts, transplacental infection, organ transplantation, and infected blood donors [3-5]. Toxoplasmosis is asymptomatic in healthy subjects, while the clinical manifestations would emerge in immunocompromised individuals with poor prognosis. Various parameters influence the level of morbidity and mortality of *Toxoplasma* infection, comprising host's age, gender, immune status, close contact with cats, occu-

KVS KOREAN VACCINE SOCIETY

© Korean Vaccine Society.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

pation, cultural and feeding behaviors as well as the parasite genotype [3].

Anti-malaria and antibacterial agents constitute the commonly prescribed drugs for toxoplasmosis treatment; however, they may entail side effects such as suppression of bone marrow, hypersensitivity, and teratogenicity. Additionally, they are only pivotal on tachyzoite stages, while there is no efficacy on chronic tissue cysts, which in turn, could be a source for disease recrudescence upon suppressed immunity [6]. Immunoprophylaxis is a preventive tactic through vaccination approaches, enabling the immune machinery to properly recognize, isolate, and eliminate the pathogenic agent upon exposure. Initial attempts were based on live/attenuated parasites or crude lysates [7]. Later, sophisticated progress in *Toxoplasma* molecular biology revealed its antigenic repertoires, comprising major surface antigens, micronemes, rhoptries, and dense granule antigens, with potential to be used as DNA and protein vaccines or prime-boost strategies. In addition to the main *T. gondii* antigens, there exists a wide array of crucial enzymes in the metabolic, transcription, and signaling pathways [7-10]. The distinct family of calcium-dependent protein kinases (CDPKs) of *Toxoplasma* plays a substantial role in the cellular invasion and egress as well as protist gliding motility. The CDPKs are exclusively expressed in the plants, apicomplexans and ciliates [10,11]. The CDPK3 is an essential enzyme for parasite egress from the host cell, calcium-based permeabilization control of parasitophorous vacuole membrane as well as the tissue cyst formation in mouse brain [12,13]. Hence, CDPK3 possesses the extensive potential to be applied for immunization strategies as previously approved [13].

Prediction of immunodominant epitopes of a particular molecule utilizing web-based predictive algorithms have opened new doors towards improved vaccine design and better immunization outcome [14]. In the current study, we have exploited a wide array of bioinformatics tools for in-depth excavation of the CDPK3 protein to find the candidate epitopes.

Materials and Methods

Amino acid sequence

The whole amino acid sequence of CDPK3 was retrieved from the public database of ToxoDB (https://toxodb.org/toxo/) server for bioinformatics analysis.

Prediction of physico-chemical functions

The physico-chemical properties of CDPK3 were predicted by using the Expasy ProtParam online server (https://web.expasy.org/protparam/), which demonstrates the number of amino acids, protein isoelectric point (pI) and molecular weight (MW), instability index, aliphatic index (AI), calculated half-life *in vitro* and *in vivo*, extinction coefficients, total number of residues with positive and negative charges as well as the grand average of hydropathicity (GRAVY) [15].

Post-translational modification sites of calcium-dependent protein kinase-3

The post-translational modification (PTM) sites, including phosphorylation and acylation regions of the CDPK3 protein were predicted by NetPhos 3.1 (http://www.cbs.dtu.dk/services/NetPhos/) and CSS-Palm (http://csspalm.biocuckoo.org/online.php) online tools, respectively [16].

Transmembrane domains and subcellular localization of calcium-dependent protein kinase-3

The subcellular localization and potential transmembrane domains of *T. gondii* CDPK3 protein were predicted utilizing PSORT II (http://psort.hgc.jp/form2.html) and TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM-2.0/) web servers, respectively [16].

Prediction of secondary and tertiary structures

The Garnier-Osguthorpe-Robson (GOR) server (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html) predicted the secondary structure of CDPK3 protein [17]. Subsequently, the three-dimensional (3D) models of protein sequence were constructed by SWISS-MODEL employing a homology-modeling method (https://swissmodel.expasy.org/) [16,18].

Tertiary structure refinement and validation

The GalaxyRefine web server at http://galaxy.seoklab.org/ was employed to refine the most suitable 3D model [19], on the basis of CASP10-tested refinement technique [20]. This bioinformatics tool rehashes structure disturbance followed by total structural relaxation via dynamics simulation [21]. Further, the overall quality of the refined structure was confirmed in the ProSA-web at https://prosa.services.came.sbg.ac.at/prosa.php. This server gives a total score for each specific structure, so out of range scores indicate probable errors in the predicted protein [22]. In the following, Ramachandran

plots of the initial and refined models were created using RAMPAGE server at (http://mordred.bioc.cam.ac.uk/-rapper/rampage.php) [23]. According to Laskowski et al. [24], "this server validates protein structure based on energetically allowed and disallowed dihedral angle psi (ψ) and phi (ϕ) of amino acid residues."

Prediction of linear and conformational B-cell epitopes

The continuous 20-mer B-cell epitopes were predicted by ABCpred server with a threshold of 0.75% (http://crdd.osdd.net/raghava/abcpred/) [25]. Moreover, B-cell epitopes were mapped based on physico-chemical characteristics, including accessibility, flexibility, polarity, hydrophilicity, turns, exposed surface, and antigenic propensity using Bcepred web tool available at http://crdd.osdd.net/raghava/bcepred/bcepred_submission.html [26]. Another B-cell epitope prediction tool was ProtScale, which was used to graphically evaluate epitopes based on alpha helix, beta-turn, hydrophobicity, average flexibility, and percent of accessible residues (https://web.expasy.org/protscale/) [15]. Prediction of conformational B-cell epitopes was done using the ElliPro tool of the Immune Epitope Database (IEDB; http://tools.iedb.org/ellipro) by default options, i.e., 0.5 min-score and 6 Å max distance. This server evaluates the epitopes through protein shape, neighbor residue clustering and residual protrusion index [27].

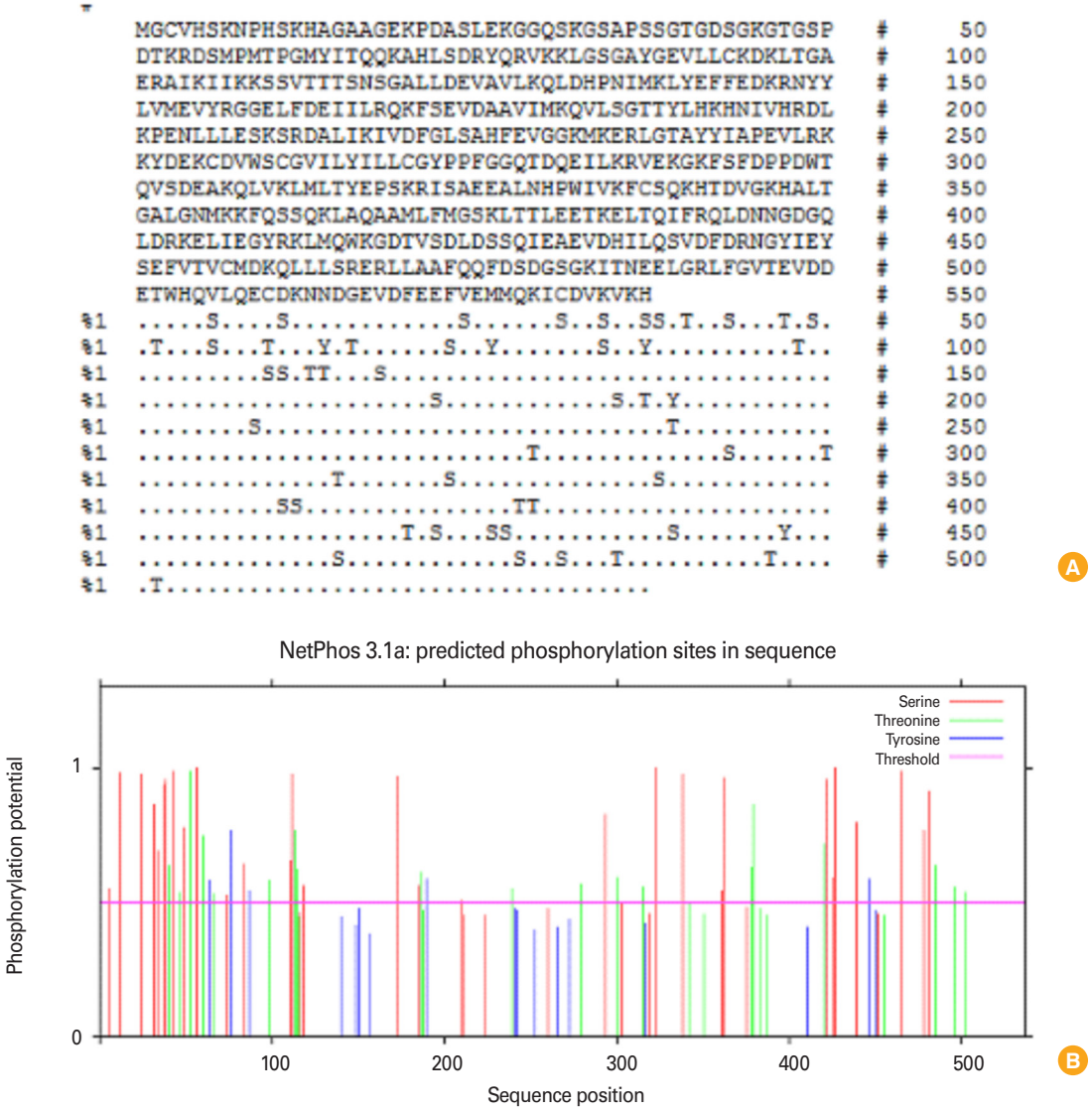

Fig. 1. NetPhos server output for CDPK3 phosphorylation sites. (A) The number of predicted sites, based on S (serine), T (threonine), and Y (tyrosine). (B) Prediction diagram of CDPK3 phosphorylation sites (http://www.cbs.dtu.dk/services/NetPhos-2.0/output.php). CDPK3, calcium-dependent protein kinase-3.

 https://doi.org/10.7774/cevr.2020.9.2.146

Prediction of major histocompatibility complex-specific epitopes

Those peptides from CDPK3 having affinity to major histocompatibility complex (MHC)-I (http://tools.iedb.org/mhci/) and MHC-II (http://tools.immuneepitope.org/mhcii) molecules were predicted by the IEDB server (recommended method 2.22), based on the half-maximal inhibitory concentration (IC_{50}) score. The 10-mer MHC-I epitopes were predicted for H2-Db, H2-Dd, H2-Kb, H2-Kd, H2-Kk, and H2-Ld mouse alleles, whereas prediction of 15-mer MHC-II epitopes was performed for H2-IAb, H2-IAd, and H2-IEd mouse alleles.

Table 1. The acylation sites of CDPK3 sequence

ID	Position	Peptide	Score
TGME49_305860 CDPK3 (*T. gondii*)	3	*****MGCVHSKNPH	17.659
TGME49_305860 CDPK3 (*T. gondii*)	93	AYGEVLLCKDKLTGA	2.928
TGME49_305860 CDPK3 (*T. gondii*)	256	RKKYDEKCDVWSCGV	6.868
TGME49_305860 CDPK3 (*T. gondii*)	261	EKCDVWSCGVILYIL	9.703
TGME49_305860 CDPK3 (*T. gondii*)	270	VILYILLCGYPPFGG	3.091
TGME49_305860 CDPK3 (*T. gondii*)	337	HPWIVKFCSQKHTDV	1.44
TGME49_305860 CDPK3 (*T. gondii*)	457	YSEFVTVCMDKQLLL	1.777
TGME49_305860 CDPK3 (*T. gondii*)	510	WHQVLQECDKNNDGE	0.756
TGME49_305860 CDPK3 (*T. gondii*)	531	VEMMQKICDVKVKH*	1.269

CDPK3, calcium-dependent protein kinase-3; *T. gondii, Toxoplasma gondii.*

Prediction of cytotoxic T-lymphocyte epitopes

The cytotoxic T-lymphocyte (CTL) epitopes specific to CDPK3 of *T. gondii* were analyzed and predicted by using CTLpred online server according to 75.8% accuracy and combined approach (http://www.imtech.res.in/raghava/ctlpred/index.html). The default settings for the prediction were artificial neural network of 0.51 and support vector machine of 0.36 [28].

Evaluation of antigenic and allergenic profiles

Protein antigenicity was estimated by two web servers: ANTIGENpro (http://scratch.proteomics.ics.uci.edu/) [29] and VaxiJen ver. 2.0 (http://www.ddg-pharmfac.net/vaxijen/) [30]. ANTIGENpro prediction, mainly relies on microarray analysis data, without dependence on alignment and pathogen. The novel alignment-free prediction of VaxiJen is exerted through an auto cross covariance-mediated sequence transformation into uniform vectors of primary amino acid properties, with prediction accuracy of 70% to 89% (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen_help.html). Also, the allergenic profile of CDPK3 was predicted by AllergenFP

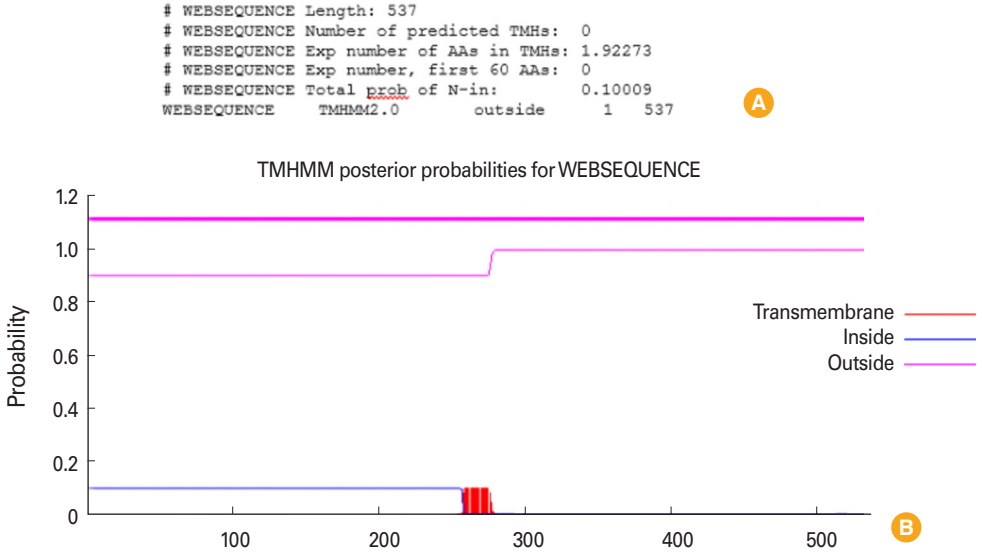

Fig. 2. Bioinformatics analysis of the transmembrane domain of CDPK3 sequence (http://www.cbs.dtu.dk/services/TMHMM-2.0/). (A) Number of predicted TMHs: the number of predicted transmembrane helices; exp number of AAs in TMHs: the expected number of amino acids in transmembrane helices. If this number is larger than 18 it is very likely to be a transmembrane protein (*OR* have a signal peptide); Exp number–first 60 AAs: the expected number of amino acids in transmembrane helices in the first 60 amino acids of the protein. If this number more than a few, you should be warned that a predicted transmembrane helix in the N-term could be a signal peptide; total prob. of N-in: the total probability that the N-term is on the cytoplasmic side of the membrane; POSSIBLE N-term signal sequence: a warning that is produced when "Exp number–first 60 AAs" is larger than 10 (http://www.cbs.dtu.dk/services/TMHMM-2.0/TMHMM2.0.guide.html#output). (B) Graphical illustration of transmembrane domain analysis of CDPK3. CDPK3, calcium-dependent protein kinase-3.

ver. 1.0 server with a descriptor-based fingerprint method to provide 88.9% accuracy of prediction (http://ddg-pharmfac.net/AllergenFP/) [31].

Ethical statement

This study received the approval from the Behbahan Faculty of Medical Sciences Ethical Committee (IR.BHN.REC.1399.034). Ethical issues (including plagiarism, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

Results

General features of *Toxoplasma* calcium-dependent protein kinase-3 gene

The ToxoDB server was applied for *T. gondii* CDPK3 protein sequence (accession ID: TGME49_305860). This protein encompassed 537 amino acid residues with a hypothesized pI of 5.98 and MW of 60,429.82. The total number of residues with positive (Arg+Lys) and negative (Asp+Glu) charges was 72 and 81, respectively. There exist a total number of 8,472 atoms in the sequence with the extinction coefficient of 53,330 M^{-1} cm^{-1} in the water at 280 nm wavelength. The half-life of the CDPK3 was estimated at 30 hours (mammalian reticulo-

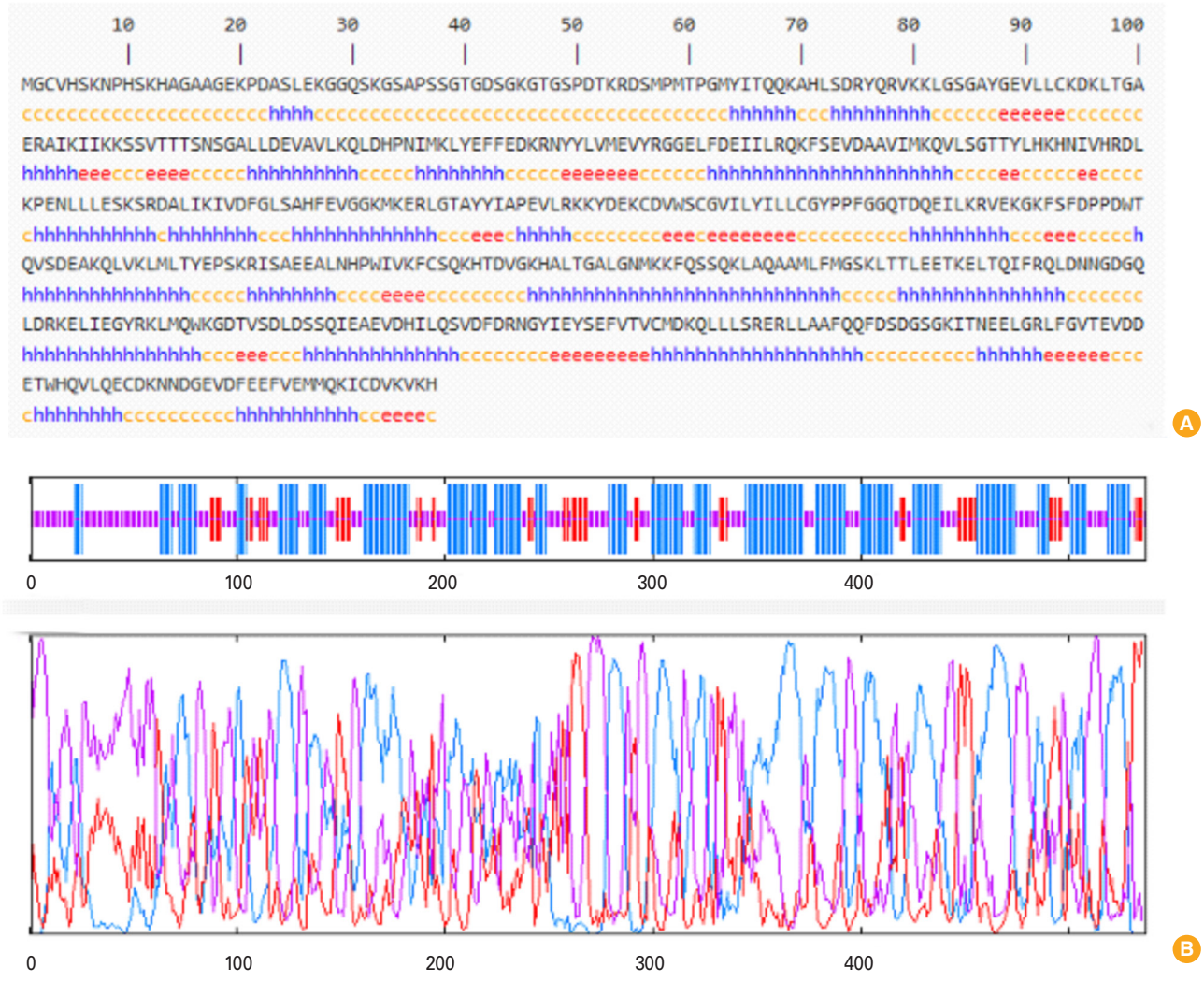

Fig. 3. (A) GOR IV server results suggested that CDPK3 encompasses 40.78% random coil, 12.48% extended strand, and 46.74% α-helix in secondary structure; (B) graphical result of the secondary structure prediction of CDPK3 using GOR IV online server (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html). GOR, Garnier-Osguthorpe-Robson; CDPK3, calcium-dependent protein kinase-3.

 https://doi.org/10.7774/cevr.2020.9.2.146

cytes, *in vitro*), >20 hours (yeast, *in vivo*), and >10 hours (*Escherichia coli*, *in vivo*). According to instability calculation, the protein was classified as stable with a score of 32.03. Also, the AI and GRAVY of the protein were 79.50 and -0.508, respectively.

Prediction of post-translational modification sites of calcium-dependent protein kinase-3

According to NetPhos 3.1 and CSS-Palm analysis, there observed nine acylation sites and 54 phosphorylation regions (serine, 30; threonine, 19; tyrosine, 5) in the sequence, rendering a total number of 63 PTM sites (Fig. 1, Table 1).

Transmembrane domains and subcellular localization

Based on the TMHMM server, no transmembrane domain was detected in the CDPK3 protein (Fig. 2). Furthermore, subcellular localization results were as follows: 39.1% cytoplasmic, 34.8% nuclear, 8.7% Golgi, 4.3% peroxisomal, 4.3% plasma membrane, 4.3% vesicles of secretory system, and 4.3% mitochondrial.

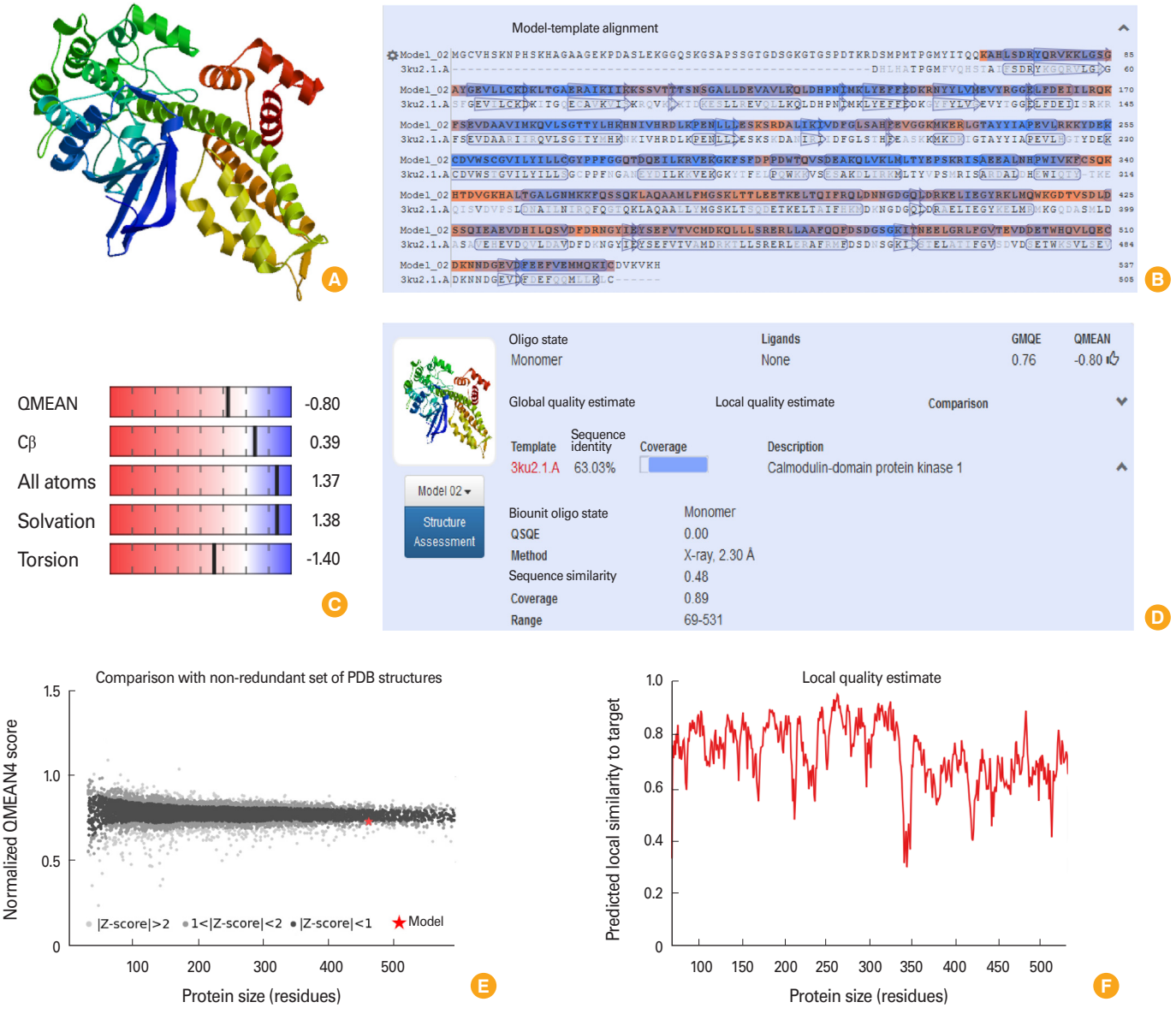

Fig. 4. SWISS-MODEL server output (https://swissmodel.expasy.org/). (A) Computed three-dimensional model; (B) model-template alignment; (C) global quality estimate; (D) sequence identity and coverage; (E) comparison with non-redundant set of PDB structures; and (F) local quality estimate.

Secondary and tertiary structure assessment

Elements of the secondary structure of CDPK3 including alpha helix, extended strand, and random coil were predicted using GOR IV web server, implying 251 (46.74%) alpha helix, 67 (12.48%) extended strand, and 219 (40.78%) random coil (Fig. 3). The 3D construct of the simulated models of the protein was drawn by SWISS-MODEL. Totally, four models were predicted, among which the model with high coverage and 63.03% sequence identity was selected as the most suitable model. The whole details of the SWISS-MODEL analysis are

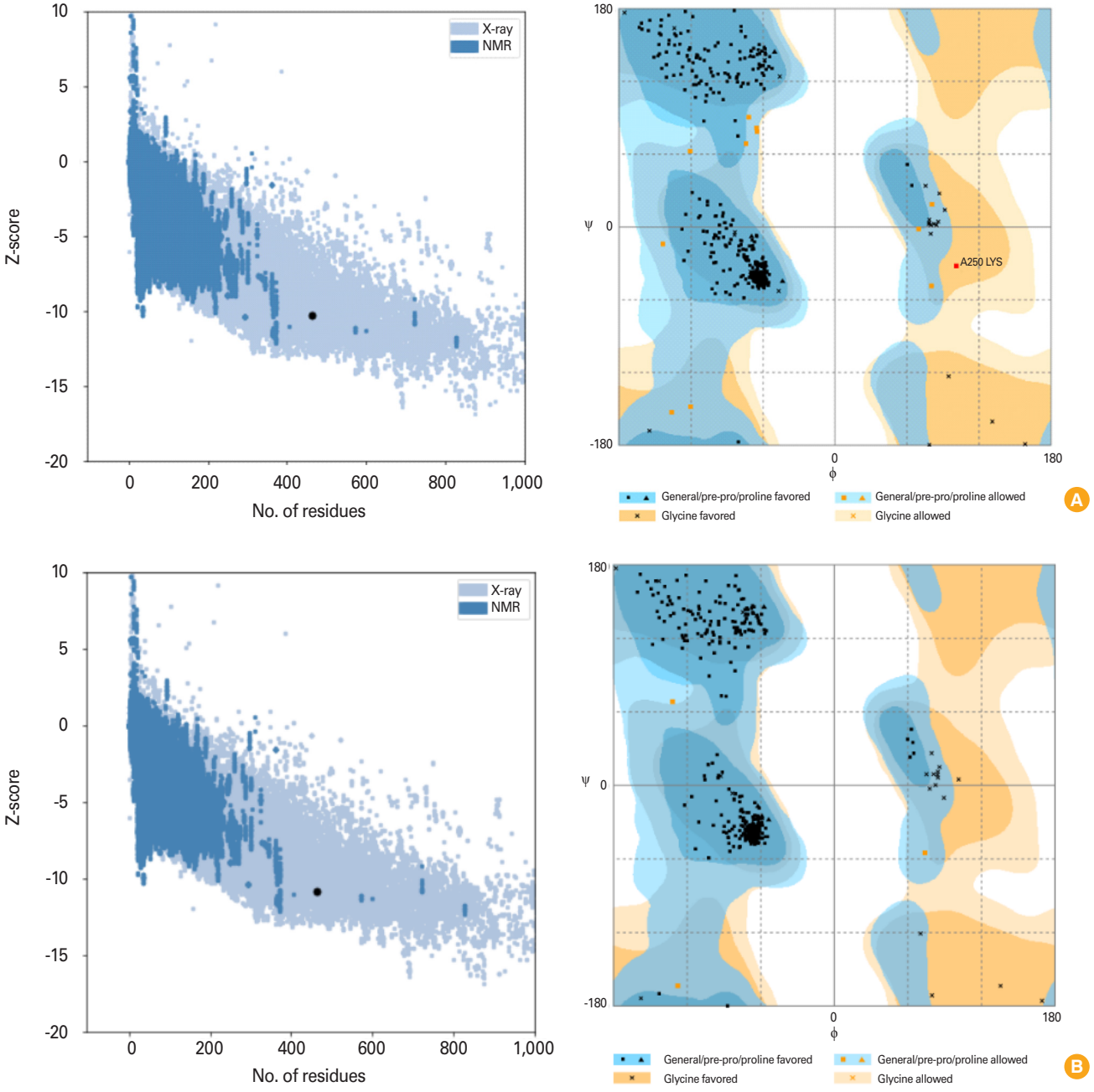

Fig. 5. Confirmation of the three-dimensional structure of CDPK3 by Ramachandran plots (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) and ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php). (A) Crude model: the Z-score was estimated as -10.31, with Ramachandran plot analysis of 97.4% of amino acid residues in favored region, 2.4% and 0.2% in allowed and outlier regions, respectively. (B) Refined model: the Z-score was increased to -10.83 indicating closer quality to protein structure defined by X-ray crystallography, with change in RAMPAGE results as follow: 99.3% of residues in favored regions, 0.7% in allowed regions, and 0.0% in outlier regions. CDPK3, calcium-dependent protein kinase-3; NMR, nuclear magnetic resonance.

 https://doi.org/10.7774/cevr.2020.9.2.146

Table 2. Linear B-cell epitopes from full-length calcium-dependent protein kinase-3 protein using ABCpred server

Rank	Sequence	Start position	Score
1	EEALNHPWIVKFCSQKHTDV	325	0.91
2	SEFVTVCMDKQLLLSRERLL	451	0.87
2	SQKHTDVGKHALTGALGNMK	338	0.87
3	SSQIEAEVDHILQSVDFDRN	426	0.86
3	QIFRQLDNNGDGQLDRKELI	388	0.86
3	SGTGDSGKGTGSPDTKRDSM	38	0.86
4	PDASLEKGGQSKGSAPSSGT	21	0.85
5	GMYITQQKAHLSDRYQRVKK	62	0.82
6	LHKHNIVHRDLKPENLLLES	190	0.81
6	LDEVAVLKQLDHPNIMKLYE	122	0.81
7	RVKKLGSGAYGEVLLCKDKL	78	0.79
7	SQKLAQAAMLFMGSKLTTLE	362	0.79
8	GEVDFEEFVEMMQKICDVKV	516	0.78
8	QSKGSAPSSGTGDSGKGTGS	30	0.78
8	VGGKMKERLGTAYYIAPEVL	229	0.78
8	LLESKSRDALIKIVDFGLSA	206	0.78
9	YGEVLLCKDKLTGAERAIKI	87	0.77
9	IKKSSVTTTSNSGALLDEVA	107	0.77
10	GTGSPDTKRDSMPMTPGMYI	46	0.76
11	QVLQECDKNNDGEVDFEEFV	505	0.75
11	GRLFGVTEVDDETWHQVLQE	490	0.75
11	VHSKNPHSKHAGAAGEKPDA	4	0.75
11	KQLVKLMLTYEPSKRISAEE	307	0.75
11	WTQVSDEAKQLVKLMLTYEP	299	0.75
11	LKRVEKGKFSFDPPDWTQVS	284	0.75
11	DFGLSAHFEVGGKMKERLGT	220	0.75

Table 4. IC_{50} values for CDPK3 binding to MHC class I molecules obtained using the IEDB[a]

MHC-I allele[b]	CDPK3 start–stop[c]	Peptide sequence	CDPK3 percentile rank[d]
H2-Db	56–65	SMPMTPGMYI	0.37
	181–190	KQVLSGTTYL	0.79
	427–436	SQIEAEVDHI	1.45
H2-Dd	98–107	TGAERAIKII	0.28
	156–165	YRGGELFDEI	0.76
	407–416	IEGYRKLMQW	1.6
H2-Kb	368–377	AAMLFMGSKL	1.07
	218–227	IVDFGLSAHF	1.6
	239–248	TAYYIAPEVL	2.75
H2-Kd	63–72	MYITQQKAHL	1.15
	187–196	TTYLHKHNIV	3.9
	75–84	RYQRVKKLGS	4.55
H2-Kk	234–243	KERLGTAYYI	0.93
	143–152	FEDKRNYYLV	1.65
	521–530	EEFVEMMQKI	1.9
H2-Ld	56–65	SMPMTPGMYI	1.8
	57–66	MPMTPGMYIT	1.8
	133–142	HPNIMKLYEF	2.9

IC_{50}, the half-maximal inhibitory concentration; CDPK3, calcium-dependent protein kinase-3; MHC, major histocompatibility complex.
[a]The Immune Epitope Database (http://tools.iedb.org/mhci/).
[b]H2-Db, H2-Dd, H2-Kb, H2-Kd, H2-Kk, and H2-Ld alleles are mouse MHC class I molecules.
[c]Ten amino acids for analysis was used each time.
[d]Low percentile rank=high level binding; high percentile rank=low level binding; IC_{50} values=percentile rank.

Table 3. B-cell linear epitopes based on different physico-chemical parameters using Bcepred server

Prediction parameter	Epitope sequence
Hydrophilicity	HSKNPHSK; AAGEKPDAS; EKGGQSKGSAPSSGTGDSGKGTGSPDTKRDSM; GSGAYGE; SSVTTTSNSGA; LESKSRDA; RKKYDEKCD; GGQTDQEI; TQVSDEAKQ; CSQKHTDV; QLDNNGDGQLDRK; KGDTVSDLDSSQ; QQFDSDGSGKI; GVTEVDDET; QECDKNNDGEVD.
Flexibility	CVHSKNPH; PDASLEKGGQSKGSAPSSGTGDSGKGTGSPDTKRDS; QRVKKLG; KIIKKSSVTTTSNS; IILRQKF; NLLLESKSR; FEVGGKMK; EILKRVEKGK; TQVSDEA; TYEPSKR; KFCSQKH; GNMKKFQSSQ; RQLDNNGD; DTVSDLD; QQFDSDGS; QECDKNND.
Accessibility	VHSKNPHSKHA; AAGEKPDASLEKGGQSKGSAP; KGTGSPDTKRDSMPMT; MYITQQKAHLSDRYQRVKKLGS; KDKLTGAER; KIIKKSSVTT; LKQLDHPNIMK; YEFFEDKRNYYLV; LRQKFSEVD; TTYLHKHN; VHRDLKPENLLLESKSRDAL; GGKMKERLGTA; PEVLRKKYDEKCD; GGQTDQEILKRVEKGKF; DPPDWTQVSDEAKQLVK; LTYEPSKRISAEEALNHP; CSQKHTDVGK; GNMKKFQSSQKLAQ; KLTTLEETKELTQIFRQLDNN GDGQLDRKELIEGYRKLMQWK; QSVDFDRNGYIEY; QQFDSDGSGKITNEELGR; TEVDDETWHQ; LQECDKNNDGEVD.
Turns	HSKNPHSKH; TTSNSGA; QLDHPNI; QLDNNGDGQ; CDKNNDGE.
Exposed surface	SKNPHSKH; SPDTKRDSM; DRYQRVKKL; KIIKKSS; FEDKRNYYL; HRDLKPEN; LESKSRDA; GKMKERLG; PEVLRKKYDEKCD; DQEILKRVEKGK; TYEPSKRI; NMKKFQSSQK; GQLDRKELIE; RKLMQWK; QECDKNNDGE; KVKH.
Polarity	VHSKNPHSKHAGA; EKGGQSK; SPDTKRDSM; KAHLSDRYQRVKKLGS; KLTGAER; YEFFEDKRNYY; EVYRGGE; EIILRQKFSEV; TTYLHKHNIVHRDLKPEN; LLESKSRDA; GGKMKERLGTA; APEVLRKKYDEKCDVW; DQEILKRVEKGKF; EAKQLVK; TYEPSKRISAEEALNHP; KFCSQKHTDVGKH; KLTTLEETKELTQ; DGQLDRKELIEGYRKLMQWK; IEAEVDHI; LLLSRERLLAA; GKITNEELGRLF; VTEVDDETWHQ; ECDKNNDGEVDFEEFVEMM; CDVKVKH.
Antigenic propensity	CVHSKNP; YGEVLLCKDKL; VLKQLDH; YYLVMEVYR; LFDEIIL; VIMKQVL; LHKHNIV; LIKIVDFGLS; KCDVWSCGVILYILLCGYPPF; QLVKLMLTY; LNHPWIVKFCSQK; EVDHILQSVDF; YSEFVTVCMDKQL; LFGVTEV; HQVLQECDK; QKICDVKVKH.

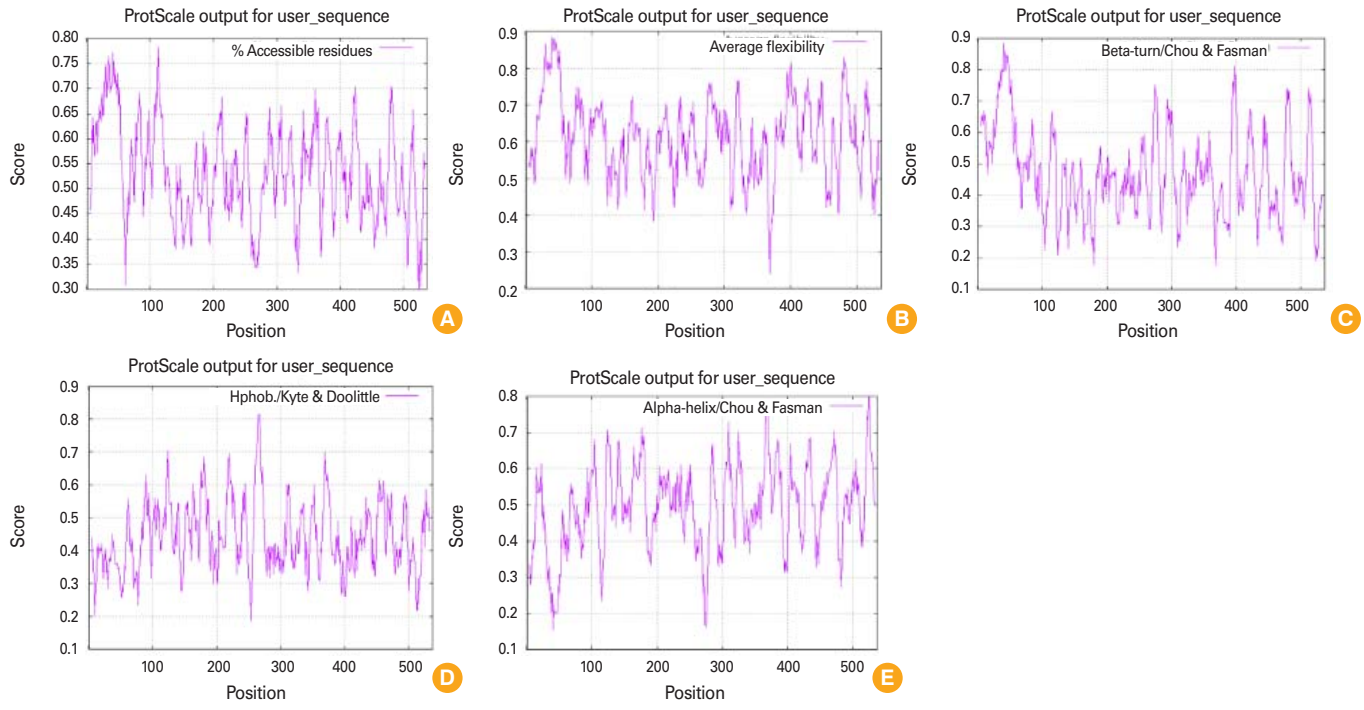

Fig. 6. Linear B-cell epitopes of CDPK3 protein sequence predicted by ProtScale server (https://web.expasy.org/protscale/), based on percent of accessible residues (A), average flexibility (B), beta turn (C), hydrophobicity (D), and alpha helix (E). CDPK3, calcium-dependent protein kinase-3.

Table 5. IC_{50} values for CDPK3 binding to MHC class II molecules obtained using the IEDB[a)]

MHC-II allele[b)]	CDPK3 start–stop[c)]	Peptide sequence	CDPK3 percentile rank[d)]
H2-IAb	236–250	RLGTAYYIAPEVLRK	2.40
	237–251	LGTAYYIAPEVLRKK	2.50
	235–249	ERLGTAYYIAPEVLR	3.10
H2-IAd	461–475	QLLLSRERLLAAFQQ	3.80
	462–476	LLLSRERLLAAFQQF	5.00
	460–474	KQLLLSRERLLAAFQ	5.35
H2-IEd	311–325	KLMLTYEPSKRISAE	3.85
	137–151	MKLYEFFEDKRNYYL	3.85
	70–84	AHLSDRYQRVKKLGS	4.30

IC_{50}, the half-maximal inhibitory concentration; CDPK3, calcium-dependent protein kinase-3; MHC, major histocompatibility complex.
[a)]The Immune Epitope Database (http://tools.immuneepitope.org/mhcii).
[b)]H2-IAb, H2-IAd, and H2-IEd alleles are mouse MHC class II molecules.
[c)]Fifteen amino acids for analysis was used each time.
[d)]Low percentile rank=high level binding; high percentile rank=low level binding; IC_{50} values=percentile rank.

Table 6. Predicted CDPK3 epitopes by CTLpred[a)]

Peptide rank	Start position[b)]	Sequence	Score (ANN/SVM)[c)]
1	71	HLSDRYQRV	0.68/1.5769321
2	235	ERLGTAYYI	0.95/1.1198854
3	143	FEDKRNYYL	0.65/1.0252087
4	304	DEAKQLVKL	0.80/0.71549234
5	259	WSCGVILYI	0.64/0.86811315
6	123	DEVAVLKQL	0.76/0.7451107
7	232	KMKERLGTA	0.92/0.58396977
8	316	YEPSKRISA	0.95/0.5085333
9	327	ALNHPWIVK	0.34/1.1046382
10	385	ELTQIFRQL	0.93/0.50219928

CDPK3, calcium-dependent protein kinase-3; ANN, artificial neural network; SVM, support vector machine.
[a)]CTLpred, available online at http://www.imtech.res.in/raghava/ctlpred/index.html.
[b)]Nine amino acids for analysis was used.
[c)]The default ANN and SVM cut-off scores were set 0.51 and 0.36, respectively.

illustrated in Fig. 4.

Refinement and validation of tertiary structure

Comparison of analysis criteria of various refined models in the GalaxyRefine server demonstrated that model number 1 as the best refined structure, having GDT-HA (0.9865), RMSD (0.298), MolProbity (1.438), clash score (8.1), poor rotamers (0.0), and Rama favored (99.3), compared to other models. Subsequently, the quality of the refined model was appraised using the ProSA-web tool, indicating a Z-score of -10.31 in the crude model compared to -10.83 in the refined model. Based on Ramachandran plot analysis, there were 458 (99.3%) resi-

 https://doi.org/10.7774/cevr.2020.9.2.146

dues in the favored region with 3 (0.7%) and 0 (0.0%) residues in the allowed and outlier regions of the refined model, respectively. On the other hand, 449 (97.4%), 11 (2.4%), and 1 (0.2%) residues were found in favored, allowed and outlier regions of the crude model, respectively (Fig. 5).

Continuous and conformational B-cell epitopes

The high-score 20-mer linear B-cell epitopes were predicted using the ABCpred web tool (Table 2). Also, Table 3 and Fig. 6 demonstrate continuous B-cell epitopes based on some physico-chemical properties, analyzed by Bcepred and ProtScale web servers. The results of the ElliPro analysis of conformational B-cell epitopes showed that there were six epitopes encompassing (1) 54 residues (score=0.815), (2) 40 residues (score=0.778), (3) 17 residues (score=0.763), (4) 3 residues (score=0.728), (5) 94 residues (score=0.677), and (6) 6 residues (score=0.588).

Major histocompatibility complex-binding epitopes

The predicted MHC-I (10-mer) and MHC-II (15-mer) epitopes were recognized based on calculated IC_{50} values for peptide-binding to mouse alleles. It is worth mentioning that the lower percentile ranks (or IC_{50} values) indicate the higher level affinity, which represents a better T-cell epitopes and vice versa (Tables 4, 5).

Cytotoxic T-lymphocyte epitope prediction

The CTLpred server was utilized to analyze CTL-specific epitopes. Totally, 10 high-ranked 9-mer CTL epitopes were predicted in the CDPK3 protein, being embedded in Table 6.

Antigenic and allergenic profiles

The antigenic profile of CDPK3 was predicted by ANTIGENpro and VaxiJen web servers with scores of 0.821125 and 0.5967 (threshold: 0.5). Based on AllergenFP analysis, the CDPK3 protein was evaluated as probable non-allergen.

Discussion

Over a century has passed from *T. gondii* discovery, a zoonotic widespread protozoan with a special interest in pregnant women and immunosuppressed individuals [3,32]. The weak immune status of at-risk people causes tachyzoite invasion to virtually all nucleated host cells, leading to the clinical disease. On the other hand, the parasites may hide as bradyzoite stages inside tissue cysts in hosts, suggesting the risk of opportunistic infection upon suppressed immune responses [3]. Thus, implementing immunoprophylactic strategies is highly recommended to prevent acute and/or chronic infections. Accordingly, identification and recognition of the precise immune-mediated processes are pivotal to battle *Toxoplasma*. In this sense, primitive attempts date back to the 1940s, which later prompted Dye test development as the first serodiagnosis for *T. gondii* infection [33].

In decades later, various investigations corroborated the protective efficacy of interleukin-12, interferon-γ (IFN-γ) as well as primary T lymphocyte subsets (CD_4^+ and CD_8^+), indicating the key role of acquired, cell-mediated immunity to limit the *Toxoplasma* infection [34,35]. Although, design and application of a vaccine candidate for the aim of vaccination are not such feasible as it may pretend; a great deal of research on vaccination during the last decades is good evidence for this allegation [7-9]. The only success of *Toxoplasma* live immunization was the development of "Toxovax", which protects sheep from congenital infection, while such live vaccine is unsafe for human use [36]. Hence, attempts in the human section mostly relied on vector-based, DNA and protein vaccines, each with its own benefits and drawbacks [7]. Besides, employing only one antigenic compound as a vaccine candidate probably do not provide sufficient immunity; thus, multi-antigenic and/or multi-epitope vaccines would be more immunoreactive, with strong excitation of IFN-γ-producing $CD8^+$ T cells and the subsequent parasite elimination [7,9].

The advent of computer science made the *in-silico* sensing of candidate epitopes of a specific sequence possible, decreasing experimental costs and facilitating high-quality vaccine design [37-39]. Designing *Neisseria meningitidis* multi-epitope-based vaccine is a leader in this novel field of research, being extended for other pathogens, including parasitic protozoa [40-42].

In *T. gondii*, a successive calcium-related molecular events exist, being mediated by a distinct family of protein kinases, called the CDPKs. Due to a lack of CDPKs expression in fungal and mammalian cells, they merit further excavation as a potential target for immunization against toxoplasmosis [10]. Current *in silico* study was done for bioinformatics excavation of *Toxoplasma* CDPK3 protein using a wide range of web-based tools. Based on the ProtParam output for physico-chemical properties, the 60.42 kDa CDPK3 molecule was a good antigen (poor immunogens are below 5–10 kDa) [43] and a stable protein with an instability index of 32.03. Regarding AI

(79.50), higher scores demonstrate increased uniformity of the target protein in a wide extent of temperatures. A negative GRAVY index was obtained from the CDPK3 protein (-0.508), rendering its hydrophilic nature and better interaction in the water-based micro-molecular milieu. Particular features of a vaccine candidate would be its immunogenic and non-allergenic profiles, as shown for *T. gondii* CDPK3 protein. The PTMs are known as a set of enzymatic functions in many eukaryotes, which occur following protein biosynthesis and determines protein activity [44]. Herein, we predicted the phosphorylation and acylation sites of the protein using NetPhos 3.1 and CSS-Palm web servers, respectively. Accordingly, there was 54 phosphorylation and nine acylation regions in the CDPK3 sequence, with a total of 63 PTM sites which inevitably affect protein activity. The absence of transmembrane domains improves presenting antigen for humoral and cellular priming of immunity and rapid responses [45].

Formally, the secondary structure is determined by the hydrogen bonds pattern between carboxyl oxygen and amino hydrogen atoms in a polypeptide chain, with α-helices and β-structures as the most occurring forms [46]. Secondary structure analysis of the GOR IV server showed three fundamental elements, including 12.48% extended strand, 40.78% random coil, and 46.74% α-helix. The tertiary structure is defined on the basis of the bonds and interactions of side chains present in a target protein [47]. In this study, we used the SWISS-MODEL for the CDPK3 3D structure prediction. Subsequently, the best-fit model (sequence identity=63.03) was rehashed by GalaxyRefine server, then RAMPAGE was applied for validation of refined model number 1, with quality scores, including GDT-HA (0.9865), RMSD (0.298), MolProbity (1.438), clash score (8.1), poor rotamers (0.0), and Rama favored (99.3). The RAMPAGE output confirmed the improvements in the tertiary structure of the refined model comparable to the crude model, with 99.3% versus 97.4%, 0.7% versus 2.4%, and 0.0% versus 0.2% of residue percentages in the favored, allowed and outlier areas, respectively.

Similar to the case of most other pathogens, innate immunity to *T. gondii* paves the way for the activation of acquired immune responses, i.e., humoral and cell-mediated immunity [48]. In the humoral phase, anti-*Toxoplasma* antibodies, particularly immunoglobulin G are a key element to limit parasite adhesion to surface receptors on host cells. They also play as opsonizing agent for macrophages to facilitate phagocytosis. Cellular immunity, particularly IFN-γ-secreting T cells and T-CD_8^+ cells, are pivotal to confine parasite multiplication and reactivation of bradyzoites within tissue cysts [49,50]. Although, in a protein sequence of interest, there may be affinity to different molecular/cellular types such as B-cells, CTL and/or MHC molecules. Thus, we exploited various web-based algorithms to analyze these specific epitopes. The ABCpred analysis revealed linear B-cell epitopes 65.93% accuracy and a threshold of >0.75. Also, B-cell continuous epitopes based on some physico-chemical options were predicted by Bcepred and ProtScale websites. ElliPro analysis demonstrated six conformational B-cell epitopes in the protein sequence, which are crucial for optimized antigen-antibody matching. Enhanced priming of T-cells would be done through better presentation of antigenic particles by antigen presenting cells. Hence, the MHC-binding affinity (IC_{50}) of epitopes was evaluated by the IEDB server. As shown in our results, low percentile ranks in a specific epitope indicate higher affinity for MHC-binding. Prediction of CTL epitopes is of utmost significance for vaccine design against toxoplasmosis. According to the CTLpred server (combined approach, 75.8% accuracy), ten 9-mer epitopes were predicted to be in the protein sequence with the potential ability for CTL induction.

In conclusion, prediction of epitopes with immunoprotective characteristics through immunosensing methods is still in its infancy, though it must be considered for future works for optimizations in the vaccine design and better immunization outcomes. For the next step, it is highly recommended to construct a multi-epitope vaccine and devise *in vivo* experiments, based on the findings of current investigation using various adjuvants and in the context of different immunization protocols.

ORCID

Hamidreza Majidiani *https://orcid.org/0000-0001-5568-1366*
Shahrzad Soltani *https://orcid.org/0000-0001-9898-1297*
Ali Dalir Ghaffari *https://orcid.org/0000-0001-9635-2876*
Mohamad Sabaghan *https://orcid.org/0000-0003-4670-767X*
Ali Taghipour *https://orcid.org/0000-0002-6876-5902*
Masoud Foroutan *https://orcid.org/0000-0002-8661-7217*

References

1. Dubey JP. Toxoplasmosis: a waterborne zoonosis. Vet Parasitol 2004;126:57-72.
2. Rahimi MT, Daryani A, Sarvi S, et al. Cats and Toxoplasma gondii: a systematic review and meta-analysis in Iran. On-

derstepoort J Vet Res 2015;82:e1-10.
3. Wang ZD, Liu HH, Ma ZX, et al. Toxoplasma gondii infection in immunocompromised patients: a systematic review and meta-analysis. Front Microbiol 2017;8:389.
4. Fallahi S, Rostami A, Nourollahpour Shiadeh M, Behniafar H, Paktinat S. An updated literature review on maternal-fetal and reproductive disorders of Toxoplasma gondii infection. J Gynecol Obstet Hum Reprod 2018;47:133-40.
5. Foroutan-Rad M, Majidiani H, Dalvand S, et al. Toxoplasmosis in blood donors: a systematic review and meta-analysis. Transfus Med Rev 2016;30:116-22.
6. Antczak M, Dzitko K, Dlugonska H. Human toxoplasmosis: searching for novel chemotherapeutics. Biomed Pharmacother 2016;82:677-84.
7. Zhang NZ, Wang M, Xu Y, Petersen E, Zhu XQ. Recent advances in developing vaccines against Toxoplasma gondii: an update. Expert Rev Vaccines 2015;14:1609-21.
8. Foroutan M, Zaki L, Tavakoli S, Soltani S, Taghipour A, Ghaffarifar F. Rhomboid antigens are promising targets in the vaccine development against Toxoplasma gondii. EXCLI J 2019;18:259-72.
9. Foroutan M, Ghaffarifar F, Sharifi Z, Dalimi A, Jorjani O. Rhoptry antigens as Toxoplasma gondii vaccine target. Clin Exp Vaccine Res 2019;8:4-26.
10. Foroutan M, Ghaffarifar F. Calcium-dependent protein kinases are potential targets for Toxoplasma gondii vaccine. Clin Exp Vaccine Res 2018;7:24-36.
11. Billker O, Lourido S, Sibley LD. Calcium-dependent signaling and kinases in apicomplexan parasites. Cell Host Microbe 2009;5:612-22.
12. McCoy JM, Whitehead L, van Dooren GG, Tonkin CJ. TgCDPK3 regulates calcium-dependent egress of Toxoplasma gondii from host cells. PLoS Pathog 2012;8:e1003066.
13. Zhang NZ, Huang SY, Zhou DH, et al. Protective immunity against Toxoplasma gondii induced by DNA immunization with the gene encoding a novel vaccine candidate: calcium-dependent protein kinase 3. BMC Infect Dis 2013;13:512.
14. Wang Y, Wang G, Cai J, Yin H. Review on the identification and role of Toxoplasma gondii antigenic epitopes. Parasitol Res 2016;115:459-68.
15. Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. The proteomics protocols handbook. Totowa, NJ: Humana Press; 2005. p.571-607.
16. Zhou J, Wang L, Zhou A, et al. Bioinformatics analysis and expression of a novel protein ROP48 in Toxoplasma gondii. Acta Parasitol 2016;61:319-28.
17. Garnier J, Gibrat JF, Robson B. GOR method for predicting protein secondary structure from amino acid sequence. Methods Enzymol 1996;266:540-53.
18. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. Electrophoresis 2009;30 Suppl 1:S162-73.
19. Ko J, Park H, Heo L, Seok C. GalaxyWEB server for protein structure prediction and refinement. Nucleic Acids Res 2012;40:W294-7.
20. Nugent T, Cozzetto D, Jones DT. Evaluation of predictions in the CASP10 model refinement category. Proteins 2014;82(Suppl 2):98-111.
21. Nain Z, Karim MM, Sen MK, Adhikari UK. Structural basis and designing of peptide vaccine using PE-PGRS family protein of Mycobacterium ulcerans: an integrated vaccinomics approach. Mol Immunol 2020;120:146-63.
22. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 2007;35:W407-10.
23. Lovell SC, Davis IW, Arendall WB 3rd, et al. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins 2003;50:437-50.
24. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283-91.
25. Saha S, Raghava GP. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins 2006;65:40-8.
26. Saha S, Raghava GP. BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. Proceedings of the Third International Conference on Artificial Immune Systems; 2004 Sep 13-16; Catania, Italy. Berlin: Springer; 2004. p.197-204.
27. Ponomarenko J, Bui HH, Li W, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinformatics 2008;9:514.
28. Bhasin M, Raghava GP. Prediction of CTL epitopes using QM, SVM and ANN techniques. Vaccine 2004;22:3195-204.
29. Magnan CN, Zeller M, Kayala MA, et al. High-throughput prediction of protein antigenicity using protein microar-

ray data. Bioinformatics 2010;26:2936-43.
30. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 2007;8:4.
31. Dimitrov I, Naneva L, Doytchinova I, Bangov I. AllergenFP: allergenicity prediction by descriptor fingerprints. Bioinformatics 2014;30:846-51.
32. Rostami A, Riahi SM, Contopoulos-Ioannidis DG, et al. Acute Toxoplasma infection in pregnant women worldwide: a systematic review and meta-analysis. PLoS Negl Trop Dis 2019;13:e0007807.
33. Coombes JL, Hunter CA. Immunity to Toxoplasma gondii: into the 21st century. Parasite Immunol 2015;37:105-7.
34. Gazzinelli RT, Hakim FT, Hieny S, Shearer GM, Sher A. Synergistic role of CD4+ and CD8+ T lymphocytes in IFN-gamma production and protective immunity induced by an attenuated Toxoplasma gondii vaccine. J Immunol 1991;146:286-92.
35. Gazzinelli RT, Hieny S, Wynn TA, Wolf S, Sher A. Interleukin 12 is required for the T-lymphocyte-independent induction of interferon gamma by an intracellular parasite and induces resistance in T-cell-deficient hosts. Proc Natl Acad Sci U S A 1993;90:6115-9.
36. Wastling JM, Harkins D, Buxton D. Western blot analysis of the IgG response of sheep vaccinated with S48 Toxoplasma gondii (Toxovax). Res Vet Sci 1994;57:384-6.
37. Hajissa K, Zakaria R, Suppian R, Mohamed Z. Epitope-based vaccine as a universal vaccination strategy against Toxoplasma gondii infection: a mini-review. J Adv Vet Anim Res 2019;6:174-82.
38. Foroutan M, Ghaffarifar F, Sharifi Z, Dalimi A, Pirestani M. Bioinformatics analysis of ROP8 protein to improve vaccine design against Toxoplasma gondii. Infect Genet Evol 2018;62:193-204.
39. Foroutan M, Ghaffarifar F, Sharifi Z, Dalimi A. Vaccination with a novel multi-epitope ROP8 DNA vaccine against acute Toxoplasma gondii infection induces strong B and T cell responses in mice. Comp Immunol Microbiol Infect Dis 2020;69:101413.
40. Adu-Bobie J, Capecchi B, Serruto D, Rappuoli R, Pizza M. Two years into reverse vaccinology. Vaccine 2003;21:605-10.
41. Caro-Gomez E, Gazi M, Goez Y, Valbuena G. Discovery of novel cross-protective Rickettsia prowazekii T-cell antigens using a combined reverse vaccinology and in vivo screening approach. Vaccine 2014;32:4968-76.
42. Ding J, Qian W, Liu Q, Liu Q. Multi-epitope recombinant vaccine induces immunoprotection against mixed infection of Eimeria spp. Parasitol Res 2012;110:2297-306.
43. Berzofsky JA, Cease KB, Cornette JL, et al. Protein antigenic structures recognized by T cells: potential applications to vaccine design. Immunol Rev 1987;98:9-52.
44. Walsh C. Posttranslational modification of proteins: expanding nature's inventory. Englewood, CO: Roberts and Company Publishers; 2006.
45. Ghaffari AD, Dalimi A, Ghaffarifar F, Pirestani M. Structural predication and antigenic analysis of ROP16 protein utilizing immunoinformatics methods in order to identification of a vaccine against Toxoplasma gondii: an in silico approach. Microb Pathog 2020;142:104079.
46. Yada RY, Jackman RL, Nakai S. Secondary structure prediction and determination of proteins: a review. Int J Pept Protein Res 1988;31:98-108.
47. Schueler-Furman O, Wang C, Bradley P, Misura K, Baker D. Progress in modeling of protein structures and interactions. Science 2005;310:638-42.
48. Sasai M, Yamamoto M. Innate, adaptive, and cell-autonomous immunity against Toxoplasma gondii infection. Exp Mol Med 2019;51:1-10.
49. El-Kady IM. T-cell immunity in human chronic toxoplasmosis. J Egypt Soc Parasitol 2011;41:17-28.
50. Suzuki Y, Orellana MA, Schreiber RD, Remington JS. Interferon-gamma: the major mediator of resistance against Toxoplasma gondii. Science 1988;240:516-8.