J Appl Biotechnol Rep. 2024 March;11(1):1207-1219

Journal of
Applied Biotechnology Reports

doi 10.30491/JABR.2023.388522.1612

Original Article

In Silico Characterization of Epitopes for Vaccine Design against *Escherichia coli* based on Heme-Utilization Protein

Fateme Sefid [1,2], Zahra Payandeh [3], Saeed Khalili [4], Zahra Sadat Hashemi [5], Alireza Zakeri [4], Armina Alagheband Bahrami [6], Seyed Mehdi Kalantar [1], Hakimeh Moghaddas Sani [7], Navid Pourzardosht [8*]

[1] Department of Medical Genetics, Shahid Sadoughi University of Medical Science, Yazd, Iran
[2] Department of Biology, Science and Art University, Yazd, Iran
[3] Immunology Research Center, Biomedicine Institute, Tabriz University of Medical Sciences, Tabriz, Iran
[4] Department of Biology Sciences, Shahid Rajaee Teacher Training University, Tehran, Iran
[5] ATMP Department, Breast Cancer Research Center, Motamed Cancer Institute, ACECR, Tehran, Iran
[6] Department of Biotechnology, School of Advanced Technologies in Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[7] Kidney Research Center, Tabriz University of Medical Sciences, Tabriz, Iran
[8] Biochemistry Department, Guilan University of Medical Sciences, Rasht, Iran

Corresponding Author: Navid Pourzardosht, PhD, Assistant Professor, Biochemistry Department, Guilan University of Medical Sciences, Rasht, Iran. Tel: +98-9113377533, E-mail: pourzardosht@gums.ac.ir

Received March 5, 2023; **Accepted** September 5, 2023; **Online Published** March 15, 2024

Abstract

Introduction: *E. coli* heme-utilization (ChuA) protein is an outer membrane protein, which has been shown as an amenable target for vaccine design studies. In the present study, we aimed to identify and characterize the most potent B and T cell epitopes of ChuA protein to unveil its most immunogenic regions.

Materials and Methods: In the present study, homology modeling was invoked to determine the three-dimensional (3D) structure of *E. coli* heme-utilization protein (ChuA). Thereafter, membrane topology, ligand binding site, surface accessibility, and clefts were assigned for ChuA, Linear and conformational B cell epitopes and T cell epitopes were predicted for ChuA. The 2D and 3D interaction plots between the most potent T cell epitopes and HLA-A020 and HLA-DRB0101 structures were drawn following the molecular docking analyses.

Results: Our results indicated that ChuA is heme ligand transporter protein, which forms a common beta-barrel structure. It is located in the membrane via 22 membrane-spanning regions. Residue–based pockets and clefts were identified on the ChuA protein. Immunological analyses revealed 9 highly potent B cell epitopes. Among predicted T cell epitopes 2 most potent epitopes were analyzed for HLA binding via molecular docking. The YSKQPGYG and FAAATTMSY epitopes showed stable interactions with HLA-A020 and HLA-DRB0101.

Conclusions: Our immunological, biochemical, and functional analysis highlighted the region of the ChuA protein, which harbors the highest immunogenic properties for vaccination purposes. Our strategy to employ 3D structure prediction and epitope prediction results could be deemed as an amenable approach for efficient vaccine design in various platforms.

Keywords: Urinary Tract Infections, Vaccine, Iron Receptor, Bioinformatics, OMP

Citation: Sefid F, Payandeh Z, Khalili S, Hashemi ZS, Zakeri A, Alagheband Bahrami A, et al. *In Silico* Characterization of Epitopes for Vaccine Design against *Escherichia coli* based on Heme-Utilization Protein. J Appl Biotechnol Rep. 2024;11(1):1207-1219. doi: 10.30491/JABR.2023.388522.1612

Introduction

Enterohemorrhagic *Escherichia coli* O157:H7 is an important zoonotic pathogen serotype in human and animals worldwide and induces bacterial diarrhea and hemolytic uremic syndrome and can cause acute renal failure and death.[1] The main identified source of bacteria is cattle, therefore, to reduce the incidence of infection cattle vaccination can be a suitable strategy.[2] Recent investigations have shown the effectiveness of different vaccine formulations in calves and mice to decrease the bacteria and production of neutralizing antibodies.[3] Different proteins such as shiga toxin 2B subunit,[4] Tir, Intimin-531 or EspA,[5] and outer membrane protein F, have been used to induce immunogenicity as vaccines. Epitope vaccines have been effectively used against a variety of bacterial infections.[6] Epitopes are recognized by major histocompatibility complex I and II (MHC I and II) molecules and trigger immunogenic responses to defeat infection.

Iron is a vital nutrient element for bacterial growth; however, its accessibility is restricted in mammalian hosts.[7] Due to the limited accessibility of pathogens to iron in the host, *E. coli* has evolved membrane systems to transport sufficient amounts of iron into the cell. Siderophores are low-molecular-weight compounds and transport chelated iron to the cell through membrane receptors. An alternative system is using outer membrane transport systems to obtain host compounds harboring iron in their structures. Various

Gram-negative bacteria including *Haemophilus influenza* type b,[8] *Vibrio cholera*, and *Shigella dysenteriae*[9] use heme protein and its different compositions as an iron supply by expressing membrane proteins. *E. coli* heme-utilization (ChuA) protein is an outer membrane protein that is synthesized in response to lack of iron and as an iron transport system, which imports heme or hemoglobin as iron source.[10] After targeting a single component of the iron uptake system using conventional vaccines, no other iron uptake mechanism can compensate the generated iron deficiency.

The coding sequence of ChuA is highly homologous to the sequence of the *ShuA* gene, a heme receptor coding sequence identified in *Shigella dysenteriae* type 1.[11] ChuA coding gene is located in the heme transport locus, which has a widespread distribution among pathogenic *E. coli* strains. Heme transport locus harbors eight open reading frames and is localized within the region equivalent to 78.7 min in *E. coli* K-12 chromosome.[12] Pathogenic bacteria secrete cytotoxins to effectively achieve the intracellular heme and/or hemoglobin sources. This feature alongside the ability of these bacteria to invade the host tissue cells has been considered as an efficient way to obtain the required iron and extend the bacterial infection[12] The virulence of these pathogens is negatively affected by the lack of iron and probably targeting ChuA protein and restricting heme transport will inhibit the onset of infection. Therefore, ChuA can be considered as a suitable candidate to design and develop new vaccines to defeat *E. coli* infections.

Since protein structure is associated with its function, having information about the conformation of the proteins makes it possible to, in part, predict protein function and also identify their binding ligands[13] Furthermore, unraveling the three-dimensional (3D) structure of the proteins will pave the way to properly modify the proteins, design efficient drugs, and also to predict useful immunogen vaccine epitopes.[14,15] However, the 3D structure of ChuA protein has not experimentally been resolved, yet.

The number of proteins with identified structures is far less than the number of proteins whose structures have been determined which accentuates the importance of solving the structure of these proteins. Determination of high-resolution 3D structure of proteins is still a challenge and many experiments fail to solve the correct structure.[16] Moreover, the purification and crystallization processes of outer membrane proteins are complicated.[16] Since the experimental methods are cost-effective and cumbersome, extensive efforts have been made to develop predictive *in silico* procedures.[17] In this regard, multiple algorithms and methods such as homology modeling have been developed. Homology modeling is a computational template-based procedure to predict protein structures based on template proteins with known structure.[18]

The purpose of the current study is to predict the 3D structure and also antigenic B- and T-cell epitopes of the ChuA protein. Moreover, the antigenicity of epitopes and their binding ability to MHC class-I (MHC I) and class-II (MHC II) by *in silico* molecular docking approach are analyzed.

Materials and Methods

Sequence Availability, Homology Alignment and Template Search

The ChuA protein sequence with NCBI (http://www.ncbi.nlm.nih.gov/protein) accession number of AAC44857.1 and Uniprot (https://www.uniprot.org/) ID of P77069 was obtained in FASTA format. The protein sequence was used as a query for BLAST against non-redundant protein database (http://blast.ncbi.nlm.nih.gov/Blast.cgi). Also, in the same server we searched for the probable putative conserved domains of the protein query. PSI-BLAST against protein data bank (PDB) was used to identify homologous structures of the protein.[19,20]

Protein Antigenicity Prediction

The Vaxijen (v.2.0) server[21] (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) was used to predict the efficient immunogenic protein sequence. The threshold value for being a probable antigen was set at 0.4.

Protein Structure Evaluation

Protparam server[23] (http://expasy.org/tools/protparam.html) was employed to estimate and predict physiochemical properties of the protein. Secondary structure features of ChuA protein were predicted using iterative threading assembly refinement (I-TASSER)[24] (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) and phyre2 method (http://www.sbg.bio.ic.ac.uk/phyre2). These are improved *in silico* protein secondary structure prediction servers which use multiple alignments to predict consensus structure.

Topology and Signal Peptide Prediction

Topology prediction of integral transmembrane (TM) proteins provides valuable information about their function. Furthermore, the retrieved structure can be used as a useful template for further experimental studies. PRED-TMBB[25] (http://biophysics.biol.uoa.gr/PRED-TMBB/) predicts the TM strands and the topology of β-barrel outer membrane proteins of gram-negative bacteria and also predicts the topology of the loops in protein structure. SignalP 4.1[26] server (http://www.cbs.dtu.dk/services/SignalP/) localizes signal peptide cleavage sites in a variety of organisms based on several combined artificial neural networks.

3D structure Prediction and Evaluation

The 3D structure of the protein was predicted using different homology modeling software. The protein sequence in FASTA format was used to retrieve the predicted structure

from Phyre2 and SWISS-MODEL servers. In parallel I-TASSER was also used to predict protein structure. SWISS-MODEL[27] (https://swissmodel.expasy.org/), predicts the 3D structure of proteins through comparative modeling using experimentally solved protein structures as template. Protein homology/analogy recognition engine V2.0 (Phyre2)[28] (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) generates 3D protein models. Moreover, Phyre2 predicts ligand binding sites of the protein and the effects of amino acid mutations on protein features. I-TASSER generates a model for proteins following template-based fragment assembly simulation strategy. The function of the protein is inferred from similar structures in protein databases.[24] The accuracy of constructed 3D models was qualitatively evaluated by verify3D, Anolea,[29] QMEAN,[30] and Procheck.[31]

Orientation of the Protein 3D Structure in Membrane

The Positioning of Proteins in Membrane (PPM) server (http://opm.phar.umich.edu/server.php) was applied for spatial positioning of the ChuA as a transmembrane protein using its predicted 3D model structure as an input data. The related orientations of proteins in membranes (OPM) database provides extensive information about structure and location of membrane-associated proteins.[32]

Ligand Binding Site

COFACTOR (http://zhanglab.ccmb.med.umich.edu/COFACTOR/) predicts the biological function of the proteins based on their structure, sequence, and protein-protein interactions (PPI). COFACTOR is a protein threading algorithm, which uses the BioLiP protein function database to find the best structure matches to predict functional sites and homologies. Retrieved homologue templates will give some information about the function, gene ontology (GO), enzyme commission (EC), and ligand-binding sites. For GO, further information can be gained from UniProt-GOA by sequence and sequence -profile alignments and from a search tool for the retrieval of interacting genes/proteins (STRING).[33]

Surface Accessible Pockets and Clefts Analyses

Pocket regions were defined using different online servers. Structural pockets and cavities are related to the binding and functional regions of the proteins and nucleic acids.[34] GHECOM (Grid-based HECOMi finder) server (http://strcomp.protein.osaka-u.ac.jp/ghecom/) is a program for finding multi-scale pockets on protein surface using mathematical morphology. The computed atlas of surface topography of proteins (CastP) server (http://sts.bioe.uic.edu/castp/) uses the weighted Delaunay triangulation and the alpha complex for shape measurements. It provides information to recognize and measure accessible and inaccessible pockets for proteins and other molecules.[35] DEPTH (http://mspc.bii.a-star.edu.sg/tankp/help.html) is a server for computing or predicting depth, cavity sizes, ligand binding sites and PKa of ionizable residues in proteins. Depth measures the closest distance of a residue or atoms to bulk solvent.[36]

Prediction Antigenic B-cell Epitopes

Recognition of B-cell epitopes is vital to develop an efficient vaccine, immunodiagnostic tests, and also to produce antibody. Therefore, *in silico* methods for reliable and accurate prediction of linear B-cell epitopes are highly desirable. B cell epitopes were identified using immune epitope database (IEDB)[37] server (http://tools.iedb.org/bcell/). IEDB server predicts linear epitopes, beta-turn structures, surface accessibility, flexibility, antigenicity, and hydrophilicity of the epitopes through BepiPred, Chou–Fasman, Emini, Karplus–Schulz, Kolaskar–Tongaonkar, and Parker prediction algorithms. DiscoTope (http://www.cbs.dtu.dk/services/DiscoTope/) predicts discontinuous B cell epitopes by measuring the surface accessibility and a novel epitope propensity amino acid score. The final scores are calculated by combining the propensity scores of residues in spatial proximity and the contact numbers.[38-40] Ellipro server (http://tools.iedb.org/ellipro/) is a new structure-based tool that predicts continuous epitopes in the accessible protein regions on the surface of the proteins. ElliPro uses three different algorithms to approximate the protein shape as an ellipsoid, measure the residue protrusion index (PI) and cluster neighboring residues based on calculated PI values.[41]

Prediction of T-cell Epitopes

To predict T-cell epitopes, selected B-cell epitopes were utilized. To identify antigenic T-cell epitopes that bind both MHC I and MHC II classes two Propred-1 (http://www.imtech.res.in/raghava/propred1/)[42] and Propred (http://www.imtech.res.in/raghava/propred/)[43] servers were used. The epitopes interacted at least with 15 MHC alleles and the antigenic ones based on VaxiJen score were chosen as probable vaccine antigen candidates. The interaction of selected antigenic epitopes with two frequent MHC alleles, HLA-A0201 and HLADRB1-0101 was assessed using T-epitope designer (http://www.bioinformation.net/ted/designer.htm)[44] and MHCpred (http://www.ddg-pharmfac.net/mhcpred/MHCPred/)[45] servers. MHCpred calculates the half maximal (50%) inhibitory concentration (IC50) of epitopes and T-epitope designer predicts if the epitopes bind with high affinity to HLA-A0201 allele and also if they interact with more than 75% of total HLA alleles. Epitopes with high affinity to HLA-A0201 and DRB1*0101 (IC50 <100 nmol/L) were selected.

Molecular Docking of Predicted T-cell Epitopes with HLA Structures

The RCSB Protein data bank (PDB) (http://www.rcsb.org/pdb/home/home.do) was searched to find 3D structures

of Hla-A0201 and HLA-DRB0101. The PDB editor software was invoked to remove all non-protein moieties and any accompanying epitopes. The 3D structures of the epitopes were built by PEP-FOLD2.0 (http://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD/). Then, the *.pdb files were converted to in *.pdbqt format using AutoDock Tools v.1.5.6. All epitopes were docked against their corresponding HLA by standalone AutoDock Vina v. 1.1.2 software. The center grid box was set to be IN maximum size. This would consider the whole HLA molecule for docking search and the results would be more reliable. All other parameters were set to be as default. The binding energy of the interactions between the epitopes and the HLA molecules were calculated by AutoDock Vina v. 1.1.2 software. The initially docked protein complex files were subjected to restricted sidechain optimization followed by soft rigid-body minimization using FireDock. The 2D interaction plot between the docked HLA and epitope structures were drawn using the LigPlot c.2.5.5 software.

Results

Sequence Availability, Homology Alignment and Template Search

The ChuA protein sequence with 660 amino acids was saved in FASTA format and was used as a query for BLAST to find sequences with high similarity. BLAST and PSI-BLAST searches revealed several homologous hits to the ChuA protein sequence. Retrieved sequences mostly belonged to CirA superfamily of proteins in TonB dependent/ligand-gated channels. These are outer membrane receptor proteins that transport substrates including siderophores (ferric chelates), vitamin B12, nickel complexes, and carbohydrates. The conserved structure of these channels shares a 22-stranded beta-barrel (C-terminal) folded around a plug domain (N-terminal). The first hit with the highest identity score (98%) and coverage (95%) was outer membrane heme receptor ShuA from *Shigella dysenteriae* (PDB code: 3FHH-A). Supplementary Figure1 shows the alignment of ChuA and ShuA proteins using Clustal W program.

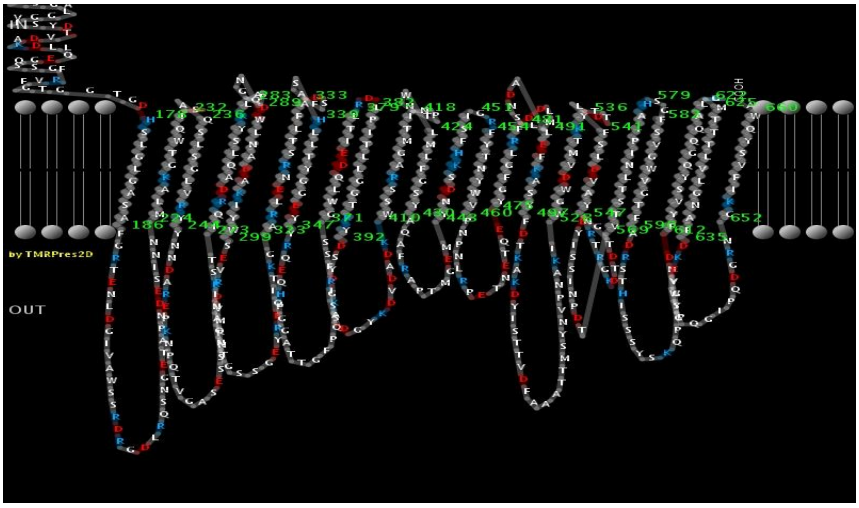

Figure 1. A 2D Topology Model of ChuA Protein Predicted by PRED-TMBB.

Antigenicity Prediction of ChuA Protein

The immunogenic and antigenic efficiency of ChuA protein was predicted using VaxiJen (cutoff ≥0.4). The calculated VaxiJen score of ChuA protein was 0.6266 and suggests the antigenic propensity of the protein.

Primary and Secondary Structure Evaluation

The molecular weight of ChuA (660 amino acids) was estimated to be 72429.05 Da with theoretical isoelectric point (pI) of 5.17. This indicates total negative charge of the protein which accords to the calculated number of 72 negative (Asp + Glu) and 57 positive (Arg + Lys) residues. The grand average of hydropathicity (GRAVY) and instability index factors were calculated to be -0.474 and 28.00, respectively. These suggest the protein is hydrophilic and also stable in test tube. Moreover, aliphatic index of 67.58 implies thermo-stability of the ChuA protein. Coils, helices and strands are components constitute the secondary structure of proteins and prediction of the proportion of these components in protein structure is used to validate its tertiary structure. The phyr2 prediction estimated the content of secondary structure components in ChuA protein as alpha helix (7%), extended strand (57.00%), and random coil (23%). The result of secondary structure prediction by I-TASSER was different and calculated alpha helix, beta strand and coil content of ChuA was predicted to be 6.96%, 47.5%, and 45.45%, respectively.

Topology and Signal Peptide Prediction

A two-dimensional (2D) topology model of ChuA was built using PRED-TMBB server and the predicted inside, transmembrane and outside regions of the protein are

indicated in Figure 1. The model suggests that the protein has β-barrel structure in native form and harbors 22 transmembrane anti-parallel β-strands. Strands forming β–barrel are linked together through 11 loops at the outer or through 10 turns at the inner sides of the membrane. SignalP4.1 identified the first 28 residues at the N terminal of the protein sequence as signal peptide.

ChuA Protein Modeling

SWISS-MODEL and phyr2 servers recruited for homology modeling and protein templates were built based on target-template alignments. The best identified template by both SWISS-MODEL and phyr2 was heme/hemoglobin outer membrane transporter ShuA from *Shigella dysenteriae* (PDB: 3FHH). Based on information retrieved from SWISS-MODEL, ShuA protein has a sequence identity of 99.53%, and coverage of residues (29-660) of 95%. Coverage indicates the fraction of the query sequence structure that can be predicted from the template, and the plausibility of the resulting model. Phyr2 modeled 620 residues (94%) of input sequence with 100.0% confidence based on the highest scored template protein, ShuA (PDB: 3FHH). I-TASSER was the other online server employed to predict the appropriate structural model for ChuA protein. According to I-TASSER program, ShuA was the only protein suggested by all ten programs implemented by I-TASSER server as a suitable model for ChuA protein. Sequence identity and coverage of residues between ChuA model and ShuA were 99% and 94%, respectively. The predicted models were evaluated through a series of tests including verify3D, Anolea, QMEAN, and procheck (Table 1). Moreover, the secondary structure of the models was assigned using DSSP algorithm (Table 1). According to the result of evaluation among suggested models from different servers, the model retrieved from SWISS-MODEL with the highest quality was selected for further assessments.

Table 1. Evaluation Results for the Obtained Models from Three Different Servers for ChuA Protein

Models	QMEAN Z-Score	Verify 3-D (% of aa with score ≥ 0.2)	Procheck		QMEAN score6	DSSP		
			% in most favoured	% in allowed regions		alpha	beta	coil
Swiss-model	-3.311	83.54%	90.9	8.8	0.464	3.8	55.4	40.8
I-TASSER	-4.317	76.52%	75.1	23.6	0.373	5	53.2	41.8
Phyr2	-3.456	76.70%	85.7%	14.2	0.45	4.35	55.8	39.9

QMEAN score6: Calculated score of the whole model that indicates the reliability of predicted structure ranging from 0 to 1. QMEAN Z-score: Demonstrates the quality of a model. Positive Z-scores are associated with models of poor quality. Verify 3D: Shows degree of the consistency between protein sequence and structure atomic coordination. Procheck: Determines the geometry of bulk protein and each residue to be within a given range.

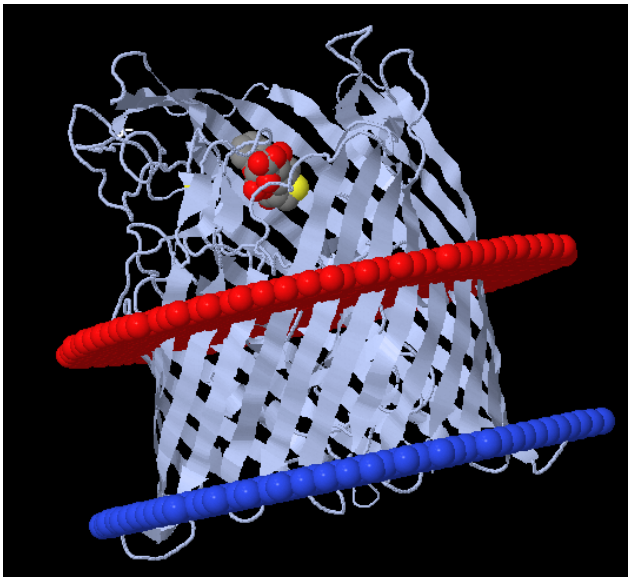

Figure 2. Orientation of the ChuA Protein Model in the Cell Membrane. Embedded_residue numbers include 1-46, 48, 63-64, 72-80, 90-112, 115, 118-132, 138-147, 151-160, 165-174, 193-204, 207-217, 247-256, 261-271, 294-303, 309-319, 350-359, 362-372, 390-412, 452-461, 470-478, 510-518, 520-528, 548-557, 561-570, 589-598, 602, 604-615, 631-640.

Orientation of the Protein 3D Structure in Membrane

The energetically preferred orientation of the ChuA protein model within the membrane was predicted using online PPM server. PPM server is able to predict the position and spatial arrangement of peripheral and integral proteins in the membrane. Based on obtained information from this server, depth of peptide protrusion into the nonpolar membrane region calculated to be 24.6 ± 0.3 Å. Measured angle between peptide axis and the bilayer (tilt angle) determined to be 1.0 ± 1.0°. The residues embedded into the membrane hydrocarbon core are mentioned in Figure 2.

Ligand Binding Site Prediction

Ligand binding sites of modeled protein was determined based on homologous known proteins using COFACTOR software. Among suggested proteins hemophore receptor from *Serratia marcescens* as a heme binding protein had the highest C-score. The result indicates the involvement of conserved residues from the plug (residues ASP68, LEU70, THR118) in iron binding site (Supplementary Figure 2). Moreover, COFACTOR predicted the function of ChuA protein by identifying 24 GO terms. High consistency was observed among the suggested functions (GO terms). Taken altogether, results suggested that ChuA is a transmembrane protein and functions as a heme ligand transporter protein. Involvement of the protein in metabolic and cellular processes has also been suggested.

Surface Accessible Pockets and Clefts on the ChuA Protein

Binding sites or pockets of ChuA protein surface were predicted using GHECOM server. The produced graph

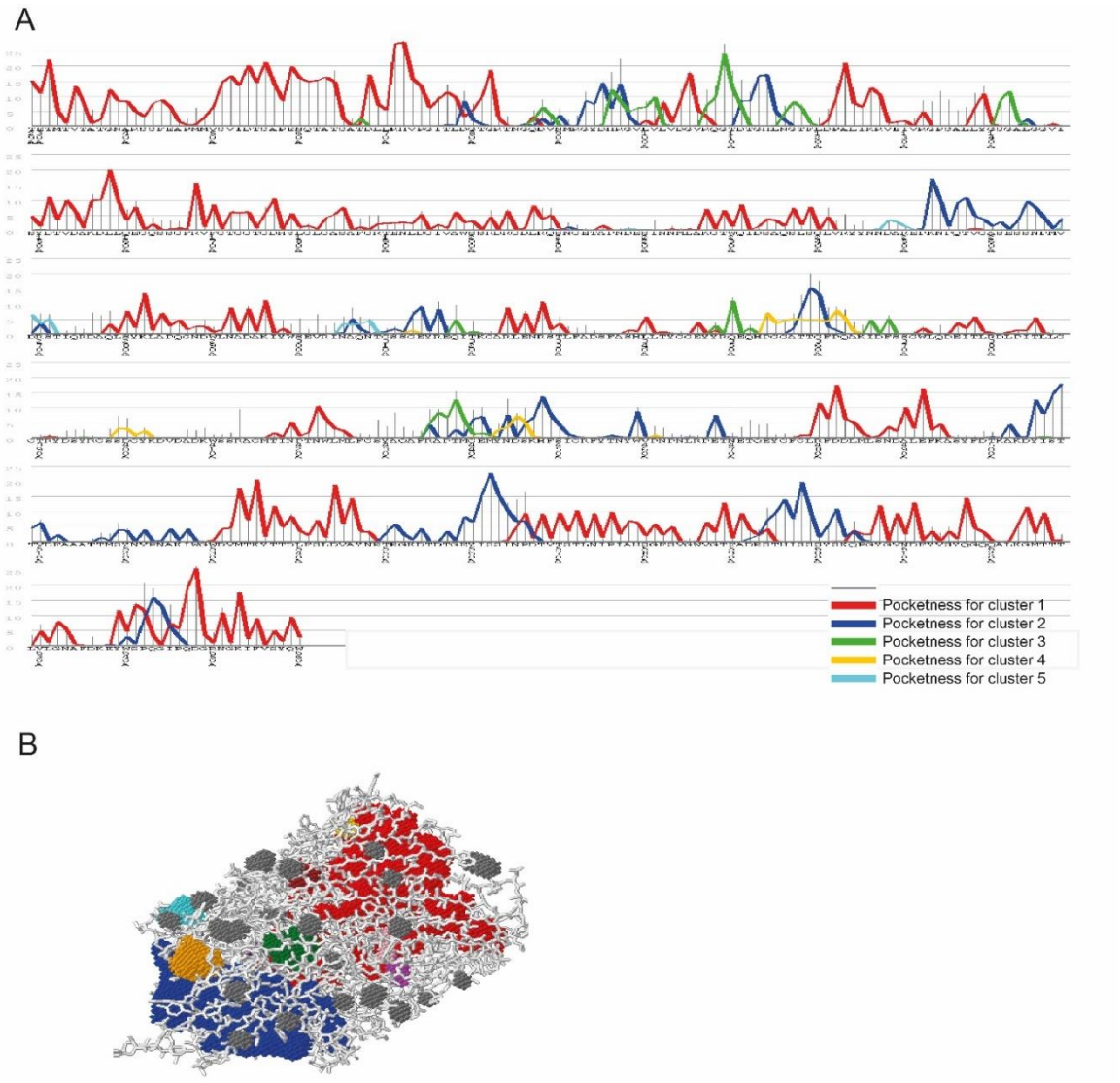

Figure 3. GHECOM Results Showing Graph Residue-Based Pocketness and Jmol View of ChuA Protein Pocket Structure. **A)** Contribution of ChuA residues to cluster pocketness. Lines indicate the value of pocketness [%] for each residue. Residues located in a deeper and larger pocket have a larger value of pocketness. The cluster number of pockets has been indicated by the color of the pocketness bar. **B)** Grid-based pocketness cluster of ChuA protein.

shows 5 identified residue–based pockets on the protein (Figure 3A). The pockets are often associated with binding sites of proteins. The result of analyses of pocketness indicates the residues that create the cluster with the highest degree of pocketness. The degree in pocketness of cluster 1 is higher compared to other clusters and is in consistence with the result of COFACTOR as a probable binding site for heme (Figure 3B). CastP server predicts pockets and cavity regions of 3D structure of the protein which can be related to the binding activity of the protein. This server analytically measures the area and the volume of each pocket and cavity, both in a solvent accessible surface and molecular surface. Figure 4 shows predicted pockets and cavities that reside in the target protein. DEPTH server calculates the distance of buried atom/residue from bulk solvent and provides information to define physical and chemical characteristics of protein structures and predict protein-protein interaction regions, phosphorylation and ligand binding sites.[36] Residue depth and solvent-accessible area values of all residues are calculated based on crystal structures of proteins and have been utilized to estimate the probability of each residue to be involved in binding activity. Figure 5 shows the probability of residues forming a binding site and residue depth plot and a 3D rendition of the cavity prediction.

B Cell Epitope Identification

To identify linear B-cell epitopes IEDB B-cell epitope prediction server was used to obtain information about epitope antigenic propensity, surface accessibility, beta turn structure and flexibility of the protein. The regions of the ChuA protein with probable ability of B cell activation were

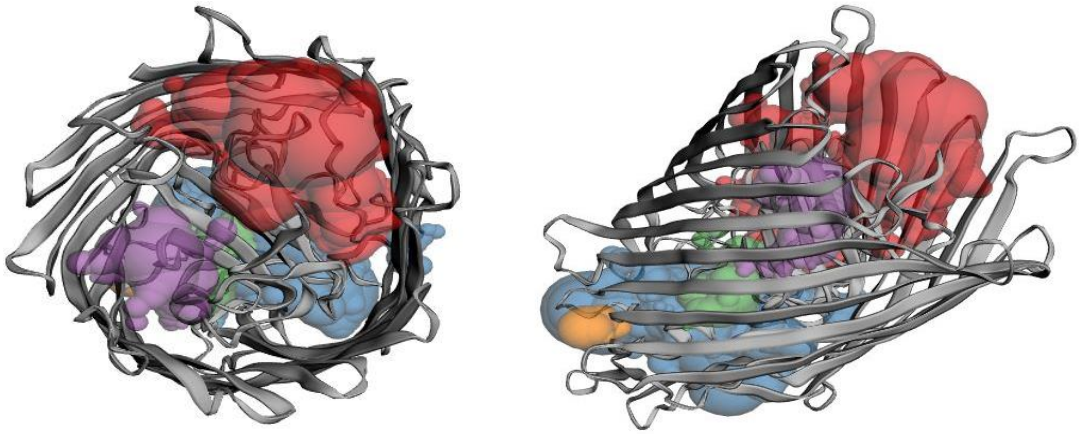

Figure 4. CastP Results Showing Surface Accessible Pockets and also Interior Inaccessible Cavities. Pockets are colored based on area and volume size and are shown from top (left) and lateral (right) view. The most important one is illustrated in red and others are shown in blue, green, purple, and orange, respectively.

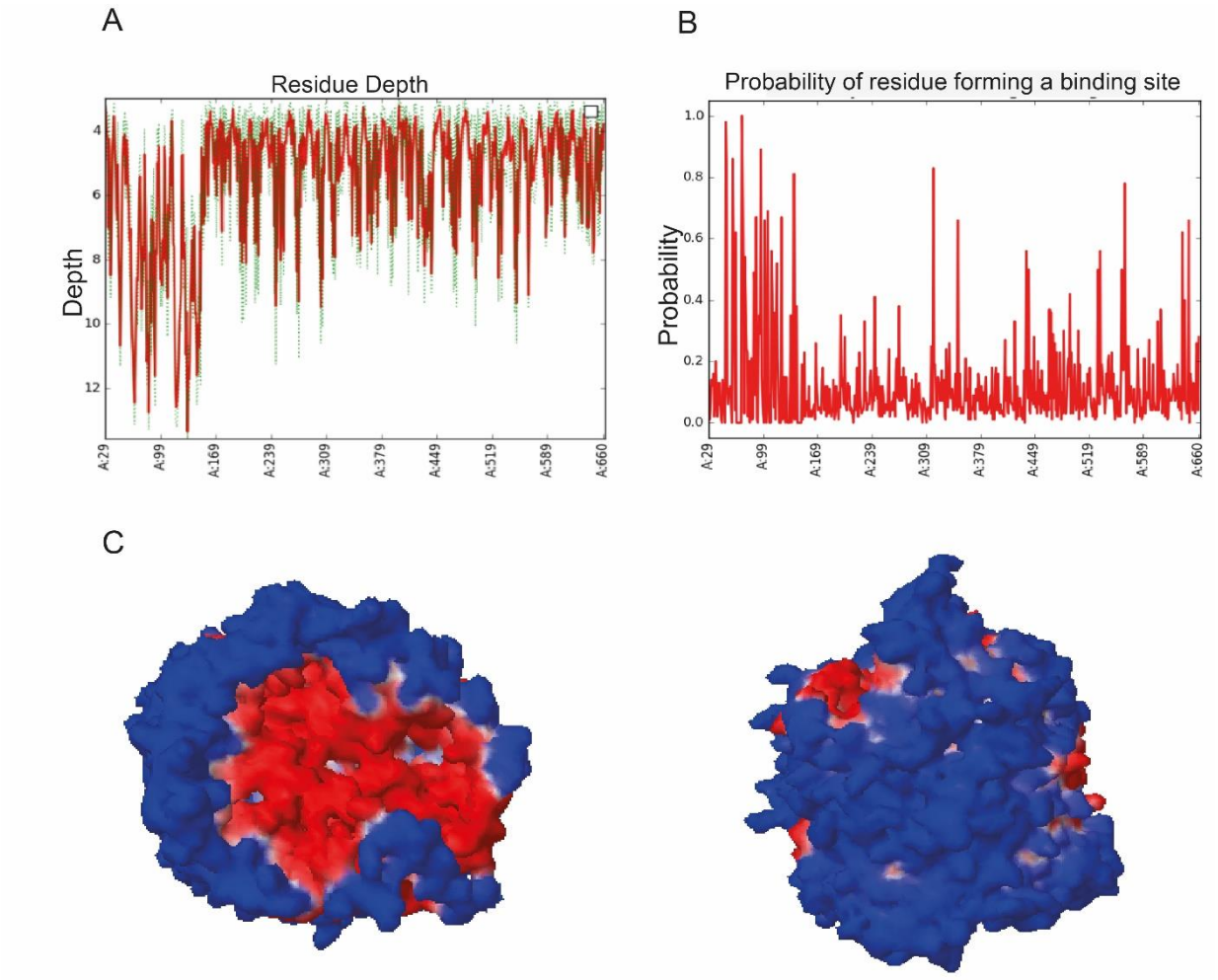

Figure 5. Probability of Residues in Creating a Binding Site and 3D Views of the Cavity Prediction. **A)** Residue depth plot predicted by Depth server, and **B)** 2D plot indicating the probability of participation of each residue in forming a binding site. The predicted binding site residues are listed below the plot. **C)** 3D renditions of the cavity prediction is shown from top (left) and lateral (right) view. Residues of the predicted binding cavity are colored in red and the rest of the protein is in blue.

identified through each tool. Supplementary Figure 3 shows the result of each used tool in IEDB server. In the first step, the antigenic epitopes recognized at least by 4 out of 6 IEDB server different tools considered as common epitopes. Sequences 205-218 and 249-267 were among residues with highest probability. ABCpred was the other server used to

predict B cell epitopes of target protein (Table 1S). Further investigations were performed using ElliPro (Table 2S) and Disco Tope (Table 3S) servers by loading PDB structure of the protein (Figure 6). The predicted epitopes by these four servers are highly varied and are overlapped. Considering VaxiJen cutoff >0.5 nine sequences (20mer) were selected as common suitable B cell epitopes from ChuA protein (Table 2).

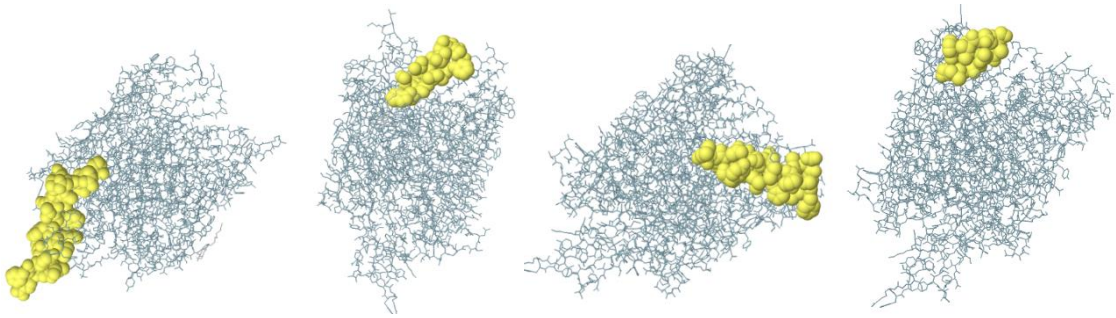

Figure 6. 4 Best Linear and Discontinuous Epitopes with the Highest Score Predicted by Ellipro Server.

Table 2. Antigenic B-Cell Epitopes of E.coli ChuA are Identified Using IEDB, ABCPred, Disco Tope, and ElliPro Servers

Sequence Position	Sequence	Vaxijen Score
307-326	QNTGSSGEYREQITKGARLE	1.348
340-360	LTYGGEYYRQEQHPGGATTG	1.224
204-223	RGDLRQSNGETAPNDESINN	1.099
551-570	RGKDTDTGEYISSINPDTVT	1.069
591-620	ADRSTHISSSYSKQPGYGVN	1.01
390-410	RYDSYRGSSDGYKDVDADKW	0.986
451-471	IGRFYTNYWVPNPNLRPETN	0.804
252-271	PKNPQTVGASESSNPMVDRS	0.602
504-523	DYISTTVDFAAATTMSYNVP	0.5217

Table 3. Predicted B- and T-Cell Epitopes. Number of MHC binding alleles and Vaxigen score of each epitope in indicated. Underlined T-cell epitopes were used for binding ability determination

B-cell Epitopes	T-cell Epitopes	Position	No. of MHCI Binding Alleles	No. of MHCII Binding Alleles	Vaxigen Score
QNTGSSGEYREQITKGARLE	**YREQITKGA**	315-323	12	10	0.5801
LTYGGEYYRQEQHPGGATTG	YYRQEQHPG	346-354	7	2	0.6024
ADRSTHISSSYSKQPGYGVN	**YSKQPGYGV**	611-619	32	1	1.1979
RGDLRQSNGETAPNDESINN	**LRQSNGETA**	207-215	18	4	2.5536
IGRFYTNYWVPNPNLRPETN	YTNYWVPNP	455-463	8	3	0.7038
	IGRFYTNYW	451-459	15	7	-0.1990
DYISTTVDFAAATTMSYNVP	**FAAATTMSY**	512-520	20	9	0.6707
RYDSYRGSSDGYKDVDADKW	YRGSSDGYK	394-402	11	2	2.161
RGKDTDTGEYISSINPDTVT	YISSINPDT	560-568	12	1	0.727

Table 4. Binding Features of Selected T-Cell Epitopes Predicted by MHCPred v.2

T-cell Epitopes	T-Epitope Designer Score of T-Cell Epitopes for HLA- A0201	IC_{50} Value of T-Cell Epitopes	Confidence of Prediction (Max = 1)
YREQITKGA	658.25	38.37	0.89
YSKQPGYGV	256.37	51.88	1
LRQSNGETA	-351.28	1194	1
FAAATTMSY	443.67	10	1

T Cell Epitope Identification

To recognize MHC I and MHC II binding T cell epitopes all identified antigenic B cell epitopes with VaxiJen cutoff >0.5 were further assessed using ProPred 1 and ProPred servers with default settings. T-cell epitopes that create complex with both MHC I and MHC II were identified and epitopes recognized by more than 15 MHC classes were selected for further investigation (Table 3). Among identified 9mer epitopes four epitopes were chosen and their antigenic features were confirmed using VaxiJen cutoff >0.5. Binding ability of each epitope was predicted using MHCPred v.2 (Table 4). According to MHCPred analysis two 9mer sequences with IC50 value under 100nM (high affinity) and confidence prediction of 1 were chosen as suitable for vaccine candidates. The YSKQPGYG (called Y) sequence is recognized by both HLA-A0201 MHC I class and HLA-DRB0101 MHC II. These two MHC alleles have high frequency hence, the interaction of selected peptides was

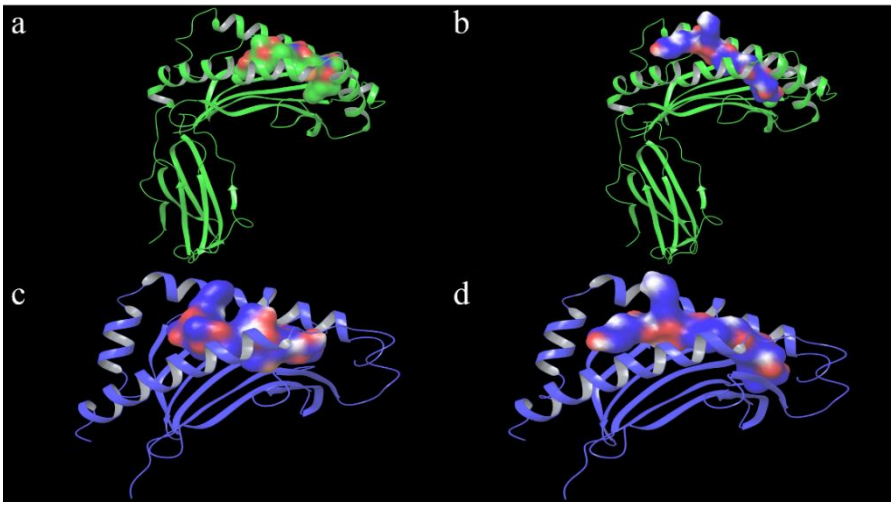

Figure 7. The Docking Results between the Predicted F and Y T Cell Epitopes and the Hla-A020 and HLA-DRB0101 Molecules. The complexes of Hla-A020/F **(a)**, Hla-A020/Y **(b)**, HLA-DRB0101/F **(c)**, and HLA-DRB0101/Y **(d)** are presented. The HLA molecules are in ribbon presentation while the epitopes are in surface presentation.

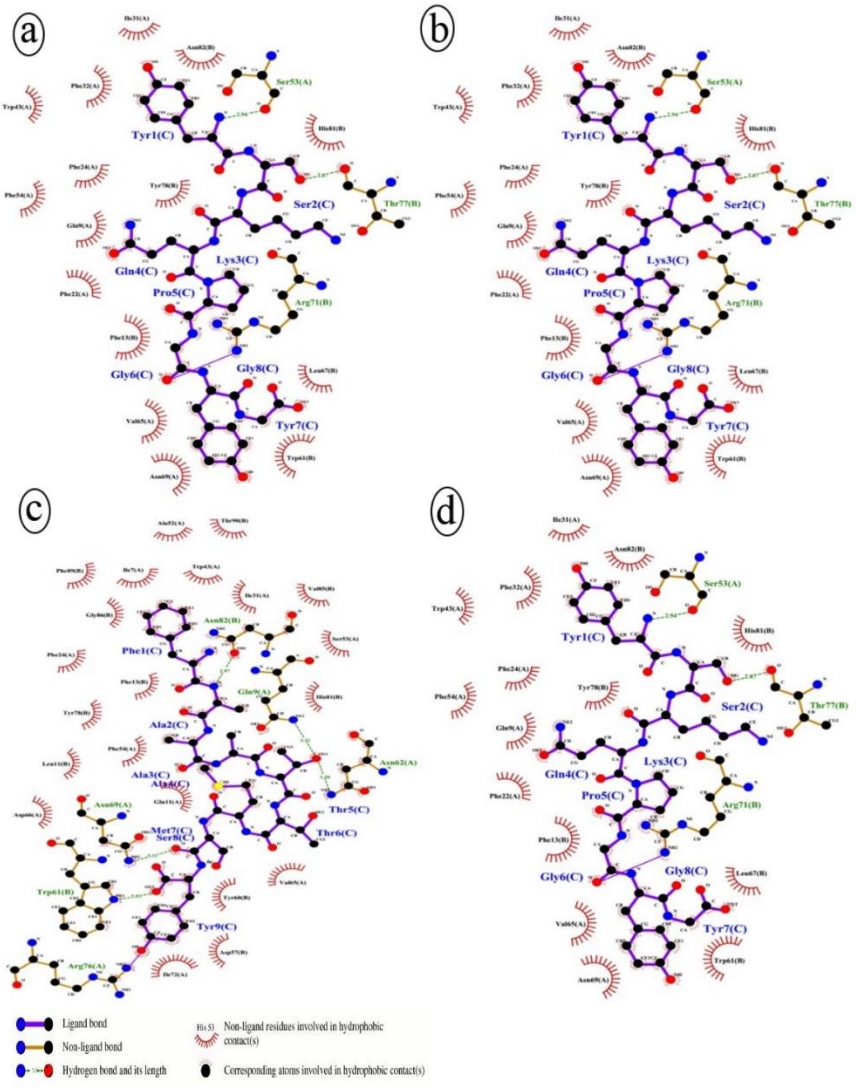

Figure 8. The 2D Interaction Results between the Predicted F and Y T Cell Epitopes and the Hla-A020 and HLA-DRB0101 Molecules. The complexes of Hla-A020/F **(a)**, Hla-A020/Y **(b)**, HLA-DRB0101/F **(c)**, and HLA-DRB0101/Y **(d)** are presented.

analyzed using MHCPred. The other FAAATTMSY (called F) epitope bind to 20 MHC I and 9 MHC II alleles. According to the result of topology prediction using PRED-TMBB server in Figure 2, both selected epitopes are located on 8th and 10th loops of ChuA protein in outer membrane region which further proves the suitability of epitopes. The structure under the PDB ID of 4UQ3 was used as the 3D structure for HLA-A020 and the structure under the PDB ID of 1AQD was used as the 3D structure for HLA-DRB0101. Molecular docking analyses indicated that predicted epitope structures are capable of binding to cavities within the grooves of HLA molecules. Moreover, the binding energies between the epitopes and HLA molecules were calculated to be -8.2, -8.7, -9.4, and -9.1 for HLA-A020/F, HLA-A020/Y, HLA-DRB0101/F and HLA-DRB0101/Y, respectively. The structures of built complexes were refined by the FireDock server (Figure 7). The 2D interaction plots of the built complexes were also drawn using the LigPlot software. Given the interaction between the peptide and HLA molecules, the results of the binding energies from docking analyses were confirmed (Figure 8). The HLA-DRB0101/F complex had the highest number of hydrogen bonds (5 hydrogen bonds) compared to the other complexes (3 hydrogen bonds). Thus, this complex is expected to have the most stable interaction with its docked peptide.

Discussion

The development of immunogenic agents for the practical activation of the adaptive immune system is one of the captivating and also challenging fields to achieve an efficient vaccine.[46,47] *E. coli* is known as a lethal hospital-acquired infection and to design an effective vaccine it has been subjected to extensive studies. The widespread occurrence of *E. coli* infections in hospitals particularly in intensive care units (ICUs), emphasizes the necessity of developing potential vaccines against this pathogen.[48] Accurate vaccine design could be facilitated by acquiring information about the protein 3D structure. The development of experimental or computational procedures has made it possible to unravel the structural conformation of proteins. These procedures provide some information about the biochemical features and function of target proteins. To overcome the disadvantages of using costly, labor-intensive, and error-prone experimental procedures, computational methods are increasingly drawing the researcher's attention to be used in the determination of the 3D structure of proteins.[14,49,50]

The current study aimed to predict the 3D structure and also antigenic B- and T-cell epitopes of the ChuA protein. Moreover, the antigenicity of epitopes and their binding ability to MHC class-I (MHC I) and class-II (MHC II) by *in silico* molecular docking approach are analyzed. To achieve reliable and accurate models, the prediction process was performed using homology modeling.[51]

Several homologous proteins were found to ChuA protein using the BLAST homology search tool. These homologous proteins mostly belong to the outer membrane-channels superfamily, TonB-dependent/Ligand-Gated channels, and ligand-gated-channel protein family.[52] These protein families are considered as virulence factors and play essential roles in bacterial pathogenicity. Amongst, the 3D structure for one of the proteins called heme/hemoglobin outer membrane transporter ShuA from *Shigella Dysenteriae* has been already resolved (PDB: 3FHH_A). ShuA belongs to the TonB-dependent transporter family (TBDT) and its resolved structure represents valuable information about the architecture of the TBDT family.[53] The predicted 3D structures for unresolved proteins could provide some information about the immunogenic characteristics of proteins and were used to design probable vaccine candidates. Homology modeling is a major method with the highest accuracy in computational prediction of protein conformation.[54,55] To have a successful homology modeling, a similar and reliable template is required which could be acquired based on the similarity between query and template protein using sequence alignment methods. A reliable template is defined as a sequence with low e-value, high query identity (>35%), and coverage against the target protein. Therefore, the protein with top scoring can be considered as the best template. The improvement of the quality of predicted structures could be achieved by a refinement run. Model refinement aims to refinement is to adequately draw the predicted model toward its native structure regarding hydrogen bonding, the position of side chain groups, and protein backbone. Assessment of the quality of full-atom refined constructed models was performed based on two sets of measure parameters. The first set of parameters is based on the overall conformational similarity between the predicted model and experimental structure including the root-mean-square deviation (RMSD). Template modeling (TM)-score is considered as the second set. Lower value of RMSD and higher TM-score/GDT-TS correspond to more accurate predicted models.[56,57] Furthermore, the overall secondary structure of the ChuA protein was predicted based on the template. The results of this study indicate that the structure of ChuA protein is mainly consisted of multiple trans-membrane β-strands in an antiparallel arrangement. The predicted models suggest that β-strands assembled together and created a β-barrel structure in the membrane.[58] In β–barrel structure adjacent β strands are connected by loops on the surface side and turns on the periplasmic side of the protein. This protein harbors more than 11 external loops. It is assumed that side chains of surface-exposed residues in the loops are involved in binding to Fe-siderophore complex.

Previous data on various pathogenic antigens indicates that the epitope density of an antigen is associated with antigenicity and immunogenicity features.[59] Therefore, gaining information about the 3D structure of B-cell epitopes is essential to

developing vaccines, designing diagnostic immune tests, and antibody production. *In silico* methods provide reliable tools in the prediction of tertiary and linear epitopes recognized by B- and T-cells. Linear epitopes are short stretches of contiguous residues belonging to a fragment of a protein, while, in the primary sequence of the tertiary epitopes no contiguity is observed, however, protein folding will bring the residues to the vicinity of each other.[60] Epitomic data can be used to find the promising ChuA fragments with higher epitope density. The identified fragments can be utilized to develop protective immunogens which can efficiently elicit humoral responses and induce the production of high avidity epitope-specific neutralizing monoclonal and polyclonal antibodies. The most efficient Linear B cell fragments of ChuA protein as a vaccine candidate are predicted to be located on the largest surface-exposed loops. Immunogens presented on the surface-exposed loops are consisted of continuous B cell epitopes which are spatially in close vicinity. The predicted epitopes in this study were in agreement with the results of previously recognized epitopes whose efficiency has been shown using approved antibodies. This agreement verified the accuracy of the computational methods used to predict the structure of immunogenic epitopes. Inclusion of T cell epitopes within the sequence of the final antigen would guarantee high immunogenicity of the vaccine candidate. In this regard, we have predicted the T cell epitope content for the ChuA protein. Two predicted T cell epitopes were restricted to Hla-A020 and HLA-DRB0101. Our docking analyses have confirmed the epitope prediction results. The interaction between the epitopes and HLA molecules were in correct orientation within the major groove of the HLA molecule. Moreover, the binding energies indicated that the interactions are strong enough to make an effective interaction and consequent activation of T cell epitopes. It could be expected that a vaccine candidate containing T cell epitopes would be more successful in eliciting humoral immunity due to the activated T helper cells.

Conclusion

In summary, it should be considered that computational methods are amenable tools to fill the gap between the tremendous number of identified protein sequences and their tertiary structures. Data retrieved from computational studies on immunogenic structures of pathogens can be considered as a suitable source to develop efficient vaccines. Predicting the tertiary structure of immunogenic agents and epitopes could pave the way to unravel the conformation, function, and subsequently therapeutic effects of the antigen and will facilitate further vaccine development studies.

Authors' Contributions

Conceptualization by FS, ZP, SK, ZSH, AAB, SMK, HMS, NP; Methodology, Validation, and Investigation by FS, ZP, SK, YA, ZSH, NP; validation and Formal analyses by FS, ZP, YA, SK, ZSH, AAB, SMK, HMS, and NP; Writing - original draft preparation by FS, ZP, SK, Writing - review and editing by FS, ZP, YA, SK, ZSH, AAB, SMK, HMS, NP; Supervision by FS, ZP, SK.

Conflict of Interest Disclosures

The authors declare that they have no conflicts of interest.

Acknowledgment

The authors thank Yazd University of Medical Sciences and Zanjan University of Medical Sciences for support to conduct this work.

References

1. Nguyen Y, Sperandio V. Enterohemorrhagic *E. coli* (EHEC) pathogenesis. Front Cell Infect Microbiol. 2012;2:90. doi:10.3389/fcimb.2012.00090
2. McNeilly TN, Naylor SW, Mahajan A, Mitchell MC, McAteer S, Deane D, et al. *Escherichia coli* O157: H7 colonization in cattle following systemic and mucosal immunization with purified H7 flagellin. Infect Immun. 2008;76(6):2594-602. doi:10.1128/iai.01452-07
3. Fingermann M, Avila L, De Marco MB, Vázquez L, Di Biase DN, Müller AV, et al. OMV-based vaccine formulations against Shiga toxin producing *Escherichia coli* strains are both protective in mice and immunogenic in calves. Hum Vaccines. 2018;14(9):2208-13. doi:10.1080/21645515.2018.1490381
4. Martorelli L, Garbaccio S, Vilte DA, Albanese AA, Mejías MP, Palermo MS, et al. Immune response in calves vaccinated with type three secretion system antigens and Shiga toxin 2B subunit of *Escherichia coli* O157: H7. PloS One. 2017;12(1):e0169422. doi:10.1371/journal.pone.0169422
5. McNeilly TN, Mitchell MC, Corbishley A, Nath M, Simmonds H, McAteer SP, et al. Optimizing the protection of cattle against *Escherichia coli* O157: H7 colonization through immunization with different combinations of H7 flagellin, Tir, intimin-531 or EspA. PLoS One. 2015;10(5):e0128391. doi:10.1371/journal.pone.0128391
6. Liu X, Chen C, Chen C, Ding R, Marslin G. Construction of a recombinant OmpC dominant epitope-based vaccine against *escherichia coli* and evaluation of its immunogenicity and protective immunity. Jundishapur J Microbiol. 2017;10(11):e55652. doi:10.5812/jjm.55652
7. Guerinot ML. Microbial iron transport. Annu Rev Microbiol. 1994;48(1):743-72. doi:10.1146/annurev.mi.48.100194.003523
8. Cope LD, Yogev RA, Muller-Eberhard U, Hansen EJ. A gene cluster involved in the utilization of both free heme and heme: hemopexin by *Haemophilus influenzae* type b. J Bacteriol. 1995;177(10):2644-53. doi:10.1128/jb.177.10.2644-2653.1995
9. O'malley SM, Mouton SL, Occhino DA, Deanda MT, Rashidi JR, Fuson KL, et al. Comparison of the heme iron utilization systems of pathogenic vibrios. J Bacteriol. 1999;181(11):3594-8. doi:10.1128/jb.181.11.3594-3598.1999
10. Torres AG, Payne SM. Haem iron-transport system in enterohaemorrhagic *Escherichia coli* O157: H7. Mol

Microbiol. 1997;23(4):825-33. doi:10.1046/j.1365-2958.1997.2641628.x
11. Vial PA, Robins-Browne R, Lior H, Prado V, Kaper JB, Nataro JP, et al. Characterization of enteroadherent-aggregative *Escherichia coli*, a putative agent of diarrheal disease. J Infect Dis. 1988;158(1):70-9. doi:10.1093/infdis/158.1.70
12. Wyckoff EE, Duncan D, Torres AG, Mills M, Maase K, Payne SM. Structure of the *Shigella dysenteriae* haem transport locus and its phylogenetic distribution in enteric bacteria. Mol Microbiol. 1998;28(6):1139-52. doi:10.1046/j.1365-2958.1998.00873.x
13. Floudas CA, Fung HK, McAllister SR, Munnigmann M, Rajgaria R. Advances in protein structure prediction and de novo protein design: A review. Chem Eng Sci. 2006;61(3):966-88. doi:10.1016/j.ces.2005.04.009
14. Payandeh Z, Rajabibazl M, Mortazavi Y, Rahimpour A. *In silico* analysis for determination and validation of human CD20 Antigen 3D Structure. Int J Pept Res Ther. 2019;25:123-35. doi:10.1007/s10989-017-9654-9
15. Khalili S, Rasaee MJ, Bamdad T. 3D structure of DKK1 indicates its involvement in both canonical and non-canonical Wnt pathways. Mol Biol. 2017;51:155-66. doi:10.1134/S0026893317010095
16. Blundell T, Carney D, Gardner S, Hayes F, Howlin B, Hubbard T, et al. Knowledge-based protein modelling and design. Eur J Biochem. 1988;172(3):513-20. doi:10.1111/j.1432-1033.1988.tb13917.x
17. Haste Andersen P, Nielsen M, Lund OL. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. Protein Sci. 2006;15(11):2558-67. doi:10.1110/ps.062405906
18. Rahman A, Zomaya AY. An overview of protein-folding techniques: issues and perspectives. Int J Bioinform Res Appl. 2005;1(1):121-43. doi:10.1504/IJBRA.2005.006911
19. Fiser A. Protein structure modeling in the proteomics era. Expert Rev Proteom. 2004;1(1):97-110. doi:10.1586/14789450.1.1.97
20. Gish W, States DJ. Identification of protein coding regions by database similarity search. Nat Genet. 1993;3(3):266-72. doi:10.1038/ng0393-266
21. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC bioinformatics. 2007;8(1):4. doi:10.1186/1471-2105-8-4
22. Yu CS, Cheng CW, Su WC, Chang KC, Huang SW, Hwang JK, Lu CH. CELLO2GO: a web server for protein subCELlular LOcalization prediction with functional gene ontology annotation. PloS One. 2014;9(6):e99368. doi:10.1371/journal.pone.0099368
23. Gasteiger E, Hoogland C, Gattiker A, Duvaud SE, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy server. Humana press; 2005.
24. Yang J, Zhang Y. Protein structure and function prediction using I-TASSER. Curr Protoc Bioinformatics. 2015;52(1):5-8. doi:10.1002/0471250953.bi0508s52
25. Bagos PG, Liakopoulos TD, Spyropoulos IC, Hamodrakas SJ. PRED-TMBB: a web server for predicting the topology of β-barrel outer membrane proteins. Nucleic Acids Res. 2004;32(suppl_2):W400-4. doi:10.1093/nar/gkh417
26. Petersen TN, Brunak S, Von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. 2011;8(10):785-6. doi:10.1038/nmeth.1701
27. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 2003;31(13):3381-5. doi:10.1093/nar/gkg520
28. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015;10(6):845-58. doi:10.1038/nprot.2015.053
29. Melo F, Feytmans E. Assessing protein structures with a non-local atomic interaction energy. J Mol Biol. 1998;277(5):1141-52. doi:10.1006/jmbi.1998.1665
30. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics. 2011;27(3):343-50. doi:10.1093/bioinformatics/btq662
31. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr. 1993;26(2):283-91. doi:10.1107/S0021889892009944
32. Lomize MA, Pogozheva ID, Joo H, Mosberg HI, Lomize AL. OPM database and PPM web server: resources for positioning of proteins in membranes. Nucleic Acids Res. 2012;40(D1):D370-6. doi:10.1093/nar/gkr703
33. Roy A, Yang J, Zhang Y. COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. Nucleic Acids Res. 2012;40(W1):W471-7. doi:10.1093/nar/gks372
34. Kawabata T. Detection of multiscale pockets on protein surfaces using mathematical morphology. Proteins: Struct Funct Bioinformatics. 2010;78(5):1195-211. doi:10.1002/prot.22639
35. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acids Res. 2006;34(suppl_2):W116-8. doi:10.1093/nar/gkl282
36. Tan KP, Nguyen TB, Patel S, Varadarajan R, Madhusudhan MS. Depth: a web server to compute depth, cavity sizes, detect potential small-molecule ligand-binding cavities and predict the pKa of ionizable residues in proteins. Nucleic Acids Res. 2013;41(W1):W314-21. doi:10.1093/nar/gkt503
37. Vita R, Overton JA, Greenbaum JA, Ponomarenko J, Clark JD, Cantrell JR, et al. The immune epitope database (IEDB) 3.0. Nucleic Acids Res. 2015;43(D1):D405-12. doi:10.1093/nar/gku938
38. Davies MN, Flower DR. Harnessing bioinformatics to discover new vaccines. Drug Discov Today. 2007;12(9-10):389-95. doi:10.1016/j.drudis.2007.03.010
39. Reimer U. Prediction of linear B-cell epitopes. Epitope Mapping Protocols: Second Edition. 2009:335-44. doi:10.1007/978-1-59745-450-6_24
40. Ansari HR, Raghava GP. Identification of conformational B-cell Epitopes in an antigen from its primary sequence. Immunome Res. 2010;6(1):6. doi:10.1186/1745-7580-6-6
41. Ponomarenko J, Bui HH, Li W, Fusseder N, Bourne PE, Sette A, Peters B. ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinformatics. 2008;9:514. doi:10.1186/1471-2105-9-514
42. Singh H, Raghava GP. ProPred1: prediction of promiscuous MHC Class-I binding sites. Bioinformatics. 2003;19(8):1009-14. doi:10.1093/bioinformatics/btg108
43. Singh H, Raghava GP. ProPred: prediction of HLA-DR binding sites. Bioinformatics. 2001;17(12):1236-7. doi:10.1093/bioinformatics/17.12.1236
44. Kangueane P, Sakharkar MK. T-Epitope Designer: A HLA-peptide binding prediction server. Bioinformation. 2005;1(1):21-4. doi:10.6026/97320630001021
45. Guan P, Doytchinova IA, Zygouri C, Flower DR. MHCPred: a server for quantitative prediction of peptide–MHC binding. Nucleic Acids Res. 2003;31(13):3621-4. doi:10.1093/nar/gkg510
46. Bauer HW, Alloussi S, Egger G, Blümlein HM, Cozma G, Schulman CC, et al. A long-term, multicenter, double-

blind study of an *Escherichia coli* extract (OM-89) in female patients with recurrent urinary tract infections. Eur Urol. 2005;47(4):542-8. doi:10.1016/j.eururo.2004.12.009
47. Uehling DT, Hopkins WJ, Elkahwaji JE, Schmidt DM, Leverson GE. Phase 2 clinical trial of a vaginal mucosal vaccine for urinary tract infections. J Urol. 2003; 170(3):867-9. doi:10.1097/01.ju.0000075094.54767.6e
48. Goluszko P, Goluszko E, Nowicki B, Nowicki S, Popov V, Wang HQ. Vaccination with purified Dr Fimbriae reduces mortality associated with chronic urinary tract infection due to *Escherichia coli* bearing Dr adhesin. Infect Immun. 2005;73(1):627-31. doi:10.1128/iai.73.1.627-631.2005
49. Jahangiri A, Rasooli I, Gargari SL, Owlia P, Rahbar MR, Amani J, et al. An *in silico* DNA vaccine against *Listeria monocytogenes*. Vaccine. 2011;29(40):6948-58. doi:10.1016/j.vaccine.2011.07.040
50. Sefid F, Rasooli I, Jahangiri A. In silico determination and validation of baumannii acinetobactin utilization a structure and ligand binding site. BioMed Res Int. 2013; 2013:172784. doi:10.1155/2013/172784
51. Kleywegt GJ, Jones TA. Databases in protein crystallography. Acta Crystallogr D Biol Crystallogr. 1998;54(6):1119-31. doi:10.1107/S0907444998007100
52. Brillet K, Reimmann C, Mislin GL, Noël S, Rognan D, Schalk IJ, et al. Pyochelin enantiomers and their outer-membrane siderophore transporters in fluorescent pseudomonads: structural bases for unique enantiospecific recognition. J Am Chem Soc. 2011;133(41):16503-9. doi:10.1021/ja205504z
53. Cobessi D, Meksem A, Brillet K. Structure of the heme/hemoglobin outer membrane receptor ShuA from Shigella dysenteriae: heme binding by an induced fit mechanism. Proteins: Struct Funct Bioinformatics. 2010;78(2):286-94. doi:10.1002/prot.22539
54. Oakhill JS, Sutton BJ, Gorringe AR, Evans RW. Homology modelling of transferrin-binding protein A from *Neisseria meningitidis*. Protein Eng Des Sel. 2005;18(5):221-8. doi:10.1093/protein/gzi024
55. Sefid F, Rasooli I, Payandeh Z. Homology modeling of a Camelid antibody fragment against a conserved region of *Acinetobacter baumannii* biofilm associated protein (Bap). J Theor Biol. 2016;397:43-51. doi:10.1016/j.jtbi.2016.02.015
56. Payandeh Z, Rajabibazl M, Mortazavi Y, Rahimpour A, Taromchi AH. Ofatumumab monoclonal antibody affinity maturation through *in silico* modeling. Iran Biomed J. 2018;22(3):180-92. doi:10.22034/ibj.22.3.180
57. Maiorov VN, Crippen GM. Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins. J Mol Biol. 1994;235(2):625-34. doi:10.1006/jmbi.1994.1017
58. Bagos PG, Liakopoulos TD, Hamodrakas SJ. Evaluation of methods for predicting the topology of β-barrel outer membrane proteins and a consensus prediction method. BMC Bioinformatics. 2005;6:7. doi:10.1186/1471-2105-6-7
59. Liu W, Chen YH. High epitope density in a single protein molecule significantly enhances antigenicity as well as immunogenicity: a novel strategy for modern vaccine development and a preliminary investigation about B cell discrimination of monomeric proteins. Eur J Immunol. 2005;35(2):505-14. doi:10.1002/eji.200425749
60. Chen J, Liu H, Yang J, Chou KC. Prediction of linear B-cell epitopes using amino acid pair antigenicity scale. Amino Acids. 2007;33:423-8. doi:10.1007/s00726-006-0485-9
61. Chen X, Zaro JL, Shen WC. Fusion protein linkers: property, design and functionality. Adv Drug Deliv Rev. 2013;65(10):1357-69. doi:10.1016/j.addr.2012.09.039
62. Gokhale RS, Khosla C. Role of linkers in communication between protein modules. Curr Opin Chem Biol. 2000; 4(1):22-7. doi:10.1016/S1367-5931(99)00046-0
63. Haddad J, Whitehead GF, Katsoulidis AP, Rosseinsky MJ. In-MOFs based on amide functionalised flexible linkers. Faraday Discuss. 2017;201:327-35. doi:10.1039/C7FD00085E
64. Argos P. An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. J Mol Biol. 1990; 211(4):943-58. doi:10.1016/0022-2836(90)90085-Z