HOSTED BY
ELSEVIER

Contents lists available at ScienceDirect

Journal of Acute Disease

journal homepage: www.jadweb.org

Original article http://dx.doi.org/10.1016/j.joad.2015.11.006

Immunological study on integrated PilQ and disulphide loop region of PilA against acute *Pseudomonas aeruginosa* infection: In silico analysis and *in vitro* production

Alireza Salimi Chirani[1*], Robabeh Majidzadeh[2], Hossein Dabiri[1], Javad Rezaei[3], Ali Esmaili[1], Yasamin Abdanan Kord[4], Narges Khabazzadeh Tehrani[5], Negin Attaran[4]

[1]*Department of Medical Microbiology, Shahid Beheshti University of Medical Sciences, Tehran, Iran*

[2]*Department of Chemistry, Pharmaceutical Sciences Branch, Islamic Azad University, Tehran, Iran*

[3]*Department of Clinical Biochemistry, Shahid Beheshti University of Medical Science, Tehran, Iran*

[4]*Department of Microbiology, Science and Research Branch, Islamic Azad University, Fars, Iran*

[5]*Biology Department, Science and Research Branch, Islamic Azad University, Tehran, Iran*

ARTICLE INFO

Article history:
Received 27 Aug 2015
Received in revised form 27 Oct 2015
Accepted 9 Nov 2015
Available online 8 Jan 2016

Keywords:
Pseudomonas aeruginosa
In silico analysis
Type IV pili
PilQ
Vaccine
Immunogen

ABSTRACT

Objective: Nowadays, *Pseudomonas aeruginosa* (*P. aeruginosa*), the highly regarded opportunistic pathogen, is the leading cause of morbidity and mortality worldwide. The *P. aeruginosa* type IV pili (T4P) as a multiple functional surface organelle in the development of acute *P. aeruginosa* infections have been well documented. Today, in silico analysis is a quick, and cost-effective tool for vaccine development.

Methods: In present study, several turns' motifs along with the chimeric protein were predicted. Based on the hydropathy analysis, numerous antibody-accessible hydrophilic regions were characterized in the chimeric protein. A synthetic chimeric gene, encoding integrated PilQ and disulphide loop region of PilA, was designed. Modeling was done to predict the 3D structure of protein. The model was validated by using Ramachandran plot statistics and by ProSA server. Identification of B-cell and T-cell corresponding epitopes was done by using appropriate servers.

Results: The closer 3D model to the native form of the chimeric protein was achieved. Validation results showed that 95.1% residues were in favor region and 3.6% of amino acid residues were in the allowed region. The B-cell epitope mappings showed that almost all the epitopes had irregular enriched structures. The major histocompatibility complex binding sequence prediction identified several human major histocompatibility complex class I and II restricted T-cell epitopes. The integrated PilQ and PilA disulphide loop encoding regions in the frame of pET28a(+) vector were expressed and purified efficiently.

Conclusions: We expect that the two recognized antigenic determinants from our chimeric protein, "AYHKGNWSGYGKDGNIGIKDEDGMNCGPIAGSCTFPTTGTS-KSPSPFVDLGAKDATSG" and "GPIAGSCTFPTTGTSKSPSP", can be able to evoke strong both humoral and cell-mediated immune responses in mouse models.

1. Introduction

Pseudomonas aeruginosa (*P. aeruginosa*), as an important opportunistic human pathogen, can cause various types of infectious diseases such as chronic obstructive pulmonary disease, cystic fibrosis and ventilator-associated pneumonia. Today, *P. aeruginosa* infections in extensive burn and immune-compromised patients have high morbidity and mortality rates[1]. Once *P. aeruginosa* colonizes in human host, the bacterium often spreads rapidly and can cause tissue destruction due to multiple virulence factors[2]. Type IV pili (T4P) and its secretion apparatus in *P. aeruginosa* serves as a key component for twitching motility, colonization and

*Corresponding author: Alireza Salimi Chirani, Department of Medical Microbiology, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
Tel: +98 21 22432517
E-mail: Alireza.SalimiChirani@sbmu.ac.ir
Peer review under responsibility of Hainan Medical College.

adhesion to various receptors[3]. The C-terminal disulphide loop (DSL) region of PilA serves as a functional binding domain of T4P in *P. aeruginosa*. Previous studies confirmed that the LD_{50} of piliated *P. aeruginosa* strains was 10-fold higher than non-piliated mutants[4,5]. Earlier studies showed that specific humoral immunity against the DSL region of pilA can be protective[6,7]. The presence of the protective continuous epitope, two β-turns within the DSL region and its unique curved shape conformation led to production of good-quality, cross-reactive antibodies against this region[8]. In *P. aeruginosa*, the outer membrane PilQ, as a gated channel, facilitates extrusion and retraction of T4P through across the outer membrane. The PilQ molecule is extremely stable T4P biogenesis machine which belongs to the GspD secretin super family[9]. In *P. aeruginosa*, only about 200–300 residues in the C-terminal region of the PilQ monomer make up conserved domain. The PilQ multimeric channel is abundant in *P. aeruginosa* cell membrane[10]. The *P. aeruginosa* PilQ protein was known as an efficient Th17 cell activator and elicited strong interleukin-17 inducer[11]. Up to now, numerous methods were experienced to produce a therapeutic and effective antiserum against *P. aeruginosa*. Despite long period efforts, some of them were not cost-effective and had several disadvantages. However, there is no commercial vaccine against *P. aeruginosa* available on the market[2]. In this study, we decided to exploit bioinformatic tools and modeling approach to better evaluate and characterize the PilQ monomer protein integrated with major pilin DSL region of *P. aeruginosa* PAO1 strain. These immunoinformatics tools help us to choose appropriate antigenic and immunogenic areas of the recombinant integrated PilQ380-706 and DSL region (QD). Our results indicate essential characteristics of QD as an appropriate vaccine candidate. Finally, we expressed and purified QD by bacterial expression system. The results of our study are discussed in details in the following paragraphs.

2. Materials and methods

2.1. Primary sequence analysis and construct design

The sequence 380-706 of PilQ and DSL region of PilA molecule from *P. aeruginosa* PAO1 strain were retrieved from gene Bank (NCBI, http://www.ncbi.nlm.nih.gov) in FASTA format with accession numbers NP_253727.1 and NP_253215.1 respectively. Selected sequence of PilQ was subjected to blastp against non-redundant protein sequence database (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The resulted sequences with a similarity > 90%, coverage > 90%, and E-value < 10^{-4} were selected and aligned with multiple sequence alignment tools from European Bioinformatics Institute server (http://www.ebi.ac.uk)[12]. In the next step, we used the helix-forming sequence (Glu-Ala-Ala-Ala-Lys) [$(EAAAK)_4$] as an integrating linker between PilQ and PilA DSL region[13]. In actual fact, in the study, the longest, conserved and well solvent-exposed sequence of PilQ from *P. aeruginosa* was selected which attached to the DSL by the linker sequence. Schematic representation of chimeric QD domains was designed by Domain Graph 2.0[14]. The primary structure and the several basic physicochemical properties of the QD sequences were evaluated by using ProtParam tool (http://web.expasy.org/protparam/). The protein sub-cellular localization of QD among Gram-negative bacteria was predicted by PSLpred server (http://www.imtech.res.in/raghava/pslpred/). PSLpred server uses a hybrid approach-based method within overall high accuracy of 91.2%[15]. Presence of signal sequence in our chimeric sequence was checked by SignalP (http://www.cbs.dtu.dk/services/SignalP-3.0/)[16], and Signal-3L software (http://www.csbio.sjtu.edu.cn/bioinf/Signal-3L/)[17].

The solubility of QD was analyzed by Recombinant Protein Solubility Prediction (http://biotech.ou.edu/#rt) with accuracy of 94%[18]. The QD surface accessibility was calculated by Immune Epitope Database (IEDB) surface accessibility prediction (http://tools.immuneepitope.org) at default threshold. The QD encoding sequence was optimized by ExPASy tools (http://www.expasy.org/) based on 15% cut-off for codon efficiency and except for positions with strong secondary structures. The chimeric QD encoding sequence in frame of expression vector pET28a(+) plasmid was constructed by Biomatik Corporation, Canada. The *Escherichia coli* (*E. coli*) BL21 (DE3) was used as host expression strain.

2.2. Antigenicity prediction

The potential hydrophilic regions and antigenicity value prediction were performed based on Kyte-Doolittle technique [19], Hopp/Woods hydrophilicity scale[20], Kolaskar and Tongaonkar method by IEDB B-cell antigenic prediction site [21]. We used VaxiJen v2.0 server (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html) to ensure accuracy of antigenic areas[22]. VaxiJen is the first server for alignment-independent prediction of protective antigens solely based on the physicochemical properties of proteins for vaccine development.

2.3. Prediction of protein secondary and 3D structures

The secondary structure of QD was studied by the MINNOU server (http://minnou.cchmc.org/) by using default parameters. The α, β, and γ turn prediction was analyzed by the AlphaPred, BetaTPred2, and GammaPred servers, respectively[23–25]. The IMTECH server also predicted weakly polar interactions between the side-chain aromatic rings and hydrogen's of backbone amides (Ar-HN interactions)[26].

The 3D structure of QD was predicted by online iterative threading assembly refinement (I-TASSER) software at http://zhanglab.ccmb.med.umich.edu/I-TASSER/[27]. The predicted result was validated by ProSA server (https://www.came.sbg.ac.at/prosa.php)[28]. In proteomics studies, structure refinement model is a closer form of molecule to the native structure.

The refinement process was done in terms of H-hydrogen bonding networks, backbone topology, chain residue standing, and atomic-level energy minimization. The selected 3D structure of QD was corrected by minimization of atomic-level energy by using ModReiner online server (http://zhanglab.ccmb.med.umich.edu/ModRefiner/)[29]. We used the refined model at predictions of discontinuous B-cell epitopes. After refinement process, the 3D modeling of QD was visualized by ChemBio office software. The 3D hydrophobicity modeling of QD was predicted by Discovery Studio (Accelrys Co. Ltd) software.

2.4. Identification of antigenic B-cell epitopes

B-cell epitope mapping to discovery of protective B-cell epitope plays a critical role in vaccine development process. For prediction of continuous B-cell epitopes (20 mers), the full-length sequence of QD was subjected to BCPreds (http://ailab.ist.psu.edu/bcpreds/) (cutoff score > 0.8)[30]. BCPreds utilizes both BCPred and amino acid pair (AAP) prediction methods with specificity of 75%. The predicted epitopes by BCPreds server were confirmed by two other web servers, ABCpred (http://www.imtech.res.in/raghava/abcpred/) and VaxiJen servers[31]. The ones with cutoff score > 0.8 were selected for ABCpred and VaxiJen server. In the next level, we used IEDB BepiPred tool (http://www.cbs.dtu.dk/services/BepiPred/) to predict flexible length linear B-cell epitopes with threshold of 0.50 using a combination of a hidden Markov model and a propensity scale method[32]. For prediction of discontinuous B-cell epitopes, we used CBTOPE (http://www.imtech.res.in/raghava/cbtope/) and DiscoTope servers (http://www.cbs.dtu.dk/services/DiscoTope/)[33]. The CBTOPE server uses the primary structure of a protein for prediction with the overall accuracy of 85%. DiscoTope server uses 3D structure of proteins[34]. The default threshold of the DiscoTope server was −3.7 to facilitate specifying the epitope residues.

2.5. Prediction of T-cell epitopes

Prediction of human leukocyte antigen (HLA) restricted peptides within QD as an antigen was detected by ProPred-I server (http://www.imtech.res.in/raghava/propred1) for prediction of major histocompatibility complex (MHC) class-I binding sites (47 alleles)[35], and ProPred server (http://www.imtech.res.in/raghava/propred) for MHC class-II HLA-DR binding site prediction (51 alleles)[36], by using matrix-based TEPITOPE algorithm. In addition, in the case of PropPred, 4% default threshold level and 5% proteasome filter were selected. The IC_{50} of common epitopes from both servers was calculated by using quantitative prediction of binding affinity of peptide-MHC binding site in MHCPred (http://www.ddg-pharmfac.net/mhcpred/MHCPred/)[37]. Epitopes with IC_{50} value < 1000 nmol/L for allele DRB1*0101 were selected as a good-quality epitope.

2.6. Overexpression and purification

The constructed plasmid pET28a(+)/QD was transformed into the *E. coli* BL21 (DH3) competence cells according to invitrogen™ protocol. The transformed cells were incubated at 37 °C in Luria-Bertani (LB) agar containing kanamycin (50 μg/mL) for overnight. After incubation, the plasmid was extracted from single colony by using the commercial extraction kit (GeneJET plasmid miniprep kit, Fermentas, Germany). The accuracy of the transformation was verified by PCR reaction and enzymatic double digestion by using the Bam HI and Hind III restriction enzymes. In order to detect the integrated QD encoding sequence in the frame of pET28a(+) expression vector, the forward: F-5′-CGTTGGCTACGACAAAAGCC-3′ and reverse: R-5′-GTACGGCAGTTCACCCAGAA-3′ primers were designed in the current study by Primer-3 software[38].

The QD protein was expressed by incubation of transformed BL21 cells in LB broth containing kanamycin (50 μg/mL) medium supplemented with 2.5 mmol/L isopropyl-β-D-thiogalactopyranoside (IPTG). The chimeric QD gene expression was studied at 37 °C at different post-induction intervals (4, 6 and 8 h). The expressed QD protein was purified by nickel-nitrilotriacetic acid magnetic agarose beads according to a protocol previously described with slight modification[39]. The expression of the recombinant protein was monitored in each elute fraction by sodium dodecyl sulfate-polyacrylamide gel electrophoresis analysis. Immune blotting analysis was performed as previous protocol[40]. The QD protein quantification was measured by the Bradford method.

3. Results

3.1. Sequence analysis and antigenic segments

Multiple sequence alignment of selected PilQ sequence showed highly conserved amino acids among the selected important *P. aeruginosa* strains: PAO1, B136, LESB58, M18, NCGM2, PA7, DK2, SCV20265 and PA14 (Figure 1). Schematic depiction of chimeric QD domains was shown in Figure 2. The physical and chemical parameters of the chimeric QD protein were shown in Table 1. The number of negatively charged residues was slightly higher than positively charged residues.

The PSLpred server identified QD as a periplasmic protein with the reliability index of 1. No potential signal peptide was detected in the QD protein by Signal-3L and SignalP servers. The QD sequence had a 95% chance of solubility when overexpressed in *E. coli*. The average score of surface accessibility was 1.000. The minimum and maximum scores and applied threshold were 0.059, 5.865 and 1.000 respectively. The Emini server detected some surface-accessible patches in different lengths along with full-length QD protein (Table 2).

3.2. Prediction of secondary structure

The MINNOU server defined each residue into α-helix, β-sheet, and random coil structures. The MINNOU server showed 8 α-helix, 18 β-sheet and 26 random coil segments along with the QD (Figure 3). Turn elements in protein structure were locations where the protein chain turned around its direction. The prediction showed 6 α turns, 18 β turns, and 18 γ turns identified in the chimeric antigen (Table 3). NHPred server predicted just four weakly polar Ar (i)-HN (i+1) interactions in QD protein. The donor and acceptor residues and their respective positions were Tyr7-Ala8, Phe17-Gln18, Tyr 48-Gln 49, and Tyr 84-Asp 85. The 3D modeling was completed based on multiple-threading alignments, sequence homology comparison, and comparative modeling methods by I-TASSER. Three 3D model structures for QD protein were predicted. The top confidence score (C-score) of the models was −0.97. The expected template modeling (TM)-score for this model was 0.48 ± 0.15. The estimated root-mean-square deviation was 10.4 ± 4.6 Å. This model was used for evaluation and refinement. The ProSA-web server validated the QD protein structure quality. In the ProSA system, an overall quality score for a specific input structure was calculated. The Z-score, −3.49 of QD was in the range of native conformations (Figure 4).

```
PAO1        ELIQVNYAKAADIAKLFQSVTSDGGQEGKEGGRGSITVDDRTNSIIAYQPQERLDELRRIVSQLDIPVRQVMIEARIVEANVGYDKSLGVRWGGAYHKGNW
B136-33     ELIQVNYAKAADIAKLFQSVTSDGGQEGKEGGRGSITVDDRTNSIIAYQPQERLDELRRIVSQLDIPVRQVMIEARIVEANVGYDKSLGVRWGGAYHKGNW
LESB58      ELIQVNYAKAADIAKLFQSVTSDGGQEGKEGGRGSITVDDRTNSIIAYQPQERLDELRRIVSQLDIPVRQVMIEARIVEANVGYDKSLGVRWGGAYHKGNW
M18         ELIQVNYAKAADIAKLFQSVTSDGGQEGKEGGRGSITVDDRTNSIIAYQPQERLDELRRIVSQLDIPVRQVMIEARIVEANVGYDKSLGVRWGGAYHKGNW
NCGM2.S1    ELIQVNYAKAADIAKLFQSVTSDGGQEGKEGGRGSITVDDRTNSIIAYQPQERLDELRRIVSQLDIPVRQVMIEARIVEANVGYDKSLGVRWGGAYHKGNW
PA7         ELIQVNYAKAADIAKLFQSVTSDGGQEGKEGGRGSITVDDRTNSIIAYQPQERLDELRRIVSQLDIPVRQVMIEARIVEANVGYDKSLGVRWGGAYHKGNW
DK2         ELIQVNYAKAADIAKLFQSVTSSG—EEKENSRGSITVDDRTNSIIAYQPQERLDELRRIVSQLDIPVRQVMIEARIVEANVGYDKSLGVRWGGAYHKGNW
SCV20265    ELIQVNYAKAADIAKLFQSVTSSG—EEKENSRGSITVDDRTNSIIAYQPQERLDELRRIVSQLDIPVRQVMIEARIVEANVGYDKSLGVRWGGAYHKGNW
UCBPP-PA14  ELIQVNYAKAADIAKLFQSVTSSG—EEKENSRGSITVDDRTNSIIAYQPQERLDELRRIVSQLDIPVRQVMIEARIVEANVGYDKSLGVRWGGAYHKGNW
            ************************.*  * **..******************:**********************************************:***

PAO1        SGYGKDGNIGIKDEDGMNCGPIAGSCTFPTTGTSKSPSPFVDLGAKDATSGIGIGFITDNIILDLQLSAMEKTGNGEIVSQPKVVTSDKETAKILKGSEVP
B1          SGYGKDGNIGIKDEDGMNCGPIAGSCTFPTTGTSKSPSPFVDLGAKDATSGIGIGFITDNIILDLQLSAMEKTGNGEIVSQPKVVTSDKETAKILKGSEVP
LESB58      RGYGKDGNIGIKDEDGMNCGPIAGNCTFPTTGTSKSPSPFVDLGAKDATSGIGIGFITDNIILDLQLSAMEKTGNGEIVSQPKVVTSDKETAKILKGSEVP
M18         NGYGKDGNIGIKDEDGMNCGPIAGNCTFPTTGTSKSPSPFVDLGAKDATSGIGIGFITDNIILDLQLSAMEKTGNGEIVSQPKVVTSDKETAKILKGSEVP
NCGM2.S1    NGYGKDGNIGIKDEDGMNCGPIAGNCTFPTTGTSKSPSPFVDLGAKDATSGIGIGFITDNIILDLQLSAMEKTGNGEIVSQPKVVTSDKETAKILKGSEVP
PA7         NAYGKDGNMGIKDEEGLNCGPIAGNCSFPTVGTNKSPSPFVDLGAKDATSGIGIGFITDNIILDLQLSAMEKTGNGEIISQPKVVTSDKETAKILKGSEVP
DK2         NSYGKNGNMGIKDKEGYNCGPFQGQCTFPTR—DNSPTPFVDMGAKDATSGIGIGFITDNIILDLQLSAMEKTGNGEVVSQPKVVTSDKETAKILKGSEIP
SCV20265    NSYGKNGNMGIKDKEGYNCGPFQGQCTFPTR—DNSPTPFVDMGAKDATSGIGIGFITDNIILDLQLSAMEKTGNGEVVSQPKVVTSDKETAKILKGSEIP
UCBPP-PA14  NSYGKNGNMGIKDKEGYNCGPFQGQCTFPTR—DNSPTPFVDMGAKDATSGIGIGFITDNIILDLQLSAMEKTGNGEVVSQPKVVTSDKETAKILKGSEIP
            .***:**:****::* ****: *.*:***  .:**:*****:*****:**********************************::*****************.*:*

PAO1        YQEASSSGATSTSFKEAALSLEVTPQITPDNRIIVEVKVTKDAPDYQNMLNGVPPINKNEVNAKILVNDGETIVIGGVFSNEQSKSVEKVPFLGELPYLGR
B1          YQEASSSGATSTSFKEAALSLEVTPQITPDNRIIVEVKVTKDAPDYQNMLNGVPPINKNEVNAKILVNDGETIVIGGVFSNTQSKSVDKVPFLGELPYLGR
LESB58      YQEASSSGATSTSFKEAALSLEVTPQITPDNRIIVEVKVTKDAPDFDRALNGVPPINKNEVNAKILVNDGETIVIGGVFSNTQSKSVDKVPFLGELPYLGR
M18         YQEASSSGATSTSFKEAALSLEVTPQITPDNRIIVEVKVTKDAPDFDRALNGVPPINKNEVNAKILVNDGETIVIGGVFSNTQSKSVDKVPFLGELPYLGR
NCGM2.S1    YQEASSSGATSTSFKEAALSLEVTPQITPDNRIIVEVKVTKDAPDFDRALNGVPPINKNEVNAKILVNDGETIVIGGVFSNEQTKAVDKVPFLGELPYLGR
PA7         YQEASSSGATSTSFKEAALSLEVTPQITPDNRIIVEVKVTKDAPDYQNMLNGVPPINKNEVNAKILVNDGETIVIGGVFSNTQSKAVDKVPFLGELPYLGR
DK2         YQEASSSGATSTSFKEAALSLEVTPQITPDNRIIVEVKVTKDAPDFQNALNGVPPINKNEVNAKILVNDGETIVIGGVFSNTQSKAVDKVPFLGELPYLGR
SCV20265    YQEASSSGATSTSFKEAALSLEVTPQITPDNRIIVEVKVTKDAPDFQNALNGVPPINKNEVNAKILVNDGETIVIGGVFSNTQSKAVDKVPFLGELPYLGR
UCBPP-PA14  YQEASSSGATSTSFKEAALSLEVTPQITPDNRIIVEVKVTKDAPDFQNALNGVPPINKNEVNAKILVNDGETIVIGGVFSNTQSKAVDKVPFLGELPYLGR
            ************************************************:::*****************************  *:*::*::*******:******

PAO1        LFRRDTVTDRKNELLVFLTPRIMNN
B1          LFRRDTVTDRKNELLVFLTPRIMNN
LESB58      LFRRDTVSDVKNELLVFLTPRIMNN
M18         LFRRDTVSDVKNELLVFLTPRIMNN
NCGM2.S1    LFRRDTVSDVKNELLVFLTPRIMNN
PA7         LFRRDTVVDRKNELLVFLAPRIMNN
DK2         LFRRDTVSDTKNELLVFLTPRIMNN
SCV20265    LFRRDTVSDTKNELLVFLTPRIMNN
UCBPP-PA14  LFRRDTVSDTKNELLVFLTPRIMNN
            ******* * ********:******
```

Figure 1. The multiple sequence alignment of PilQ380-706 among *P. aeruginosa* strains: PAO1, B136, LESB58, M18, NCGM2, PA7, DK2, SCV20265 and PA14.

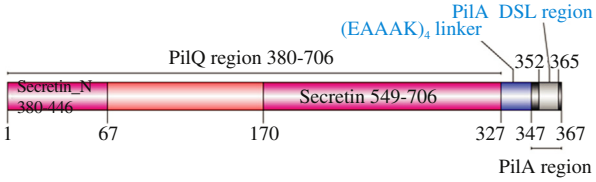

Figure 2. Schematic representation of *P. aeruginosa* PAO1 antigenic construct consisting of PilQ regions and the C-terminal DSL region of PilA integrated by helix-forming linker $(EAAAK)_4$.

Table 1

Physicochemical properties of QD shown by ProtParam server.

Characteristics	Property
Number of amino acids	367
Molecular weight (dalton)	39361.4
Total number of negatively charged residues (Asp + Glu)	49
Total number of positively charged residues (Arg + Lys)	42
Theoretical isoelectric point (pI)	5.14
Estimated half-life	1 h (mammalian reticulocytes, *in vitro*); 30 min (yeast, *in vivo*); > 10 h (*E. coli*, *in vivo*)
Instability index	32.09 (classified as stable protein)
Extinction coefficient value at 280 nm	27550 mg/mL
Aliphatic index	82.92
Grand average hydropathy	–0.342

3.3. Evaluation and refinement of QD 3D model

The QD protein structure refinement showed several structural changes in the structure which scattered throughout the length of the protein (Figure 5). After refinement process, the most abundant structural conversion throughout full-length recombinant QD protein was conversion of random-coil into turn element features. The comparison of conformational change of QD structure before and after refinement process was shown in detail in Table 4.

Table 2

Prediction of surface-accessible sequences of QD by Emini server.

Peptide	Start position	End position	Peptide length
DGGQEGKEG	23	31	9
YQPQERLDEL	48	57	10
TGTSKSP	132	138	7
VPYQEA	201	206	6
KVTKDAPDYQN	240	250	11
INKNEV	258	263	6
NEQSKSV	283	289	7
LFRRDTVTDRKN	304	315	12
STQDPMF	354	360	7

3.4. Model stability assessment

Structure validation and assessment of model stability were calculated by using Ramachandran (φ/ψ) plot provided by RAMPAGE server. The Ramachandran plot analysis of the model prior to refinement showed that 293 residues (80.3%) were in favor regions (A, B and L) and 46 (12.6%) of amino acid residues were in the allowed regions (a, b, l, and p) of the plot. The number of residues in outlier region was 26 (7.1%) (Figure 6A).

In the next step, after refinement, corresponding plot revealed that 347 (95.1%) residues from improved model were in favor regions (A, B and L) and 13 (3.6%) of residues were in the allowed regions (a, b, l, and p) of the plot. Only 5 (1.4%) residues of QD were in outlier region. The five residues that covered the unflavored position (Ser87, Pro202, 227, 231, Pro256) were demonstrated by red dots in Figure 6B.

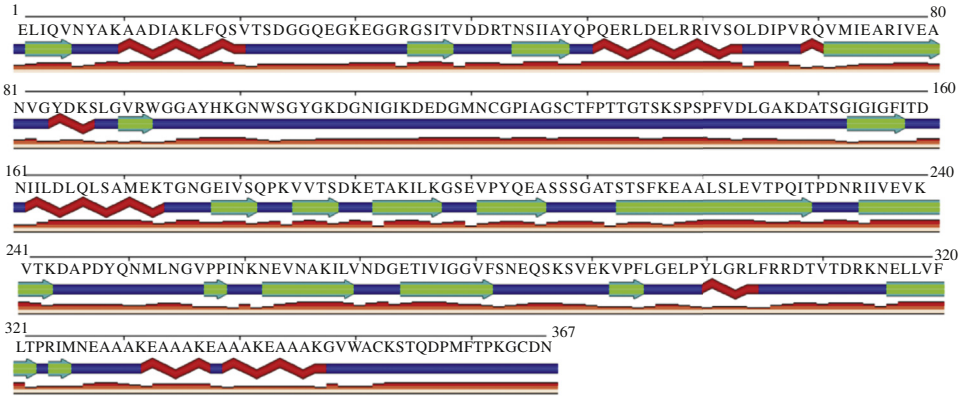

Figure 3. Analysis of QD protein secondary structure by MINNOU server.
Green and red colors represent β-sheet and α-helix structures respectively while the blue line represents random coil. The orange columns show confidence of prediction for each position.

3.5. Hydropathy analysis

Based on the Kyte-Doolittle method, there were three hydrophobic regions in the full-length QD protein. The major hydrophobic area was extended from the amino acid 150 to 165. Based on the Hopp-Woods scale, the regions of 31-35, 48-50, 181-200, 230-242, 306-318, 332-341 contained the potential hydrophilic region. After refinement process, 3D hydrophobicity modeling of QD represented several antibody-accessible hydrophilic surfaces (Figure 7).

3.6. Antigenicity prediction

IEDB B-cell analysis resource revealed several antigenic sites listed in Table 5 by Kolaskar and Tongaonkar antigenicity

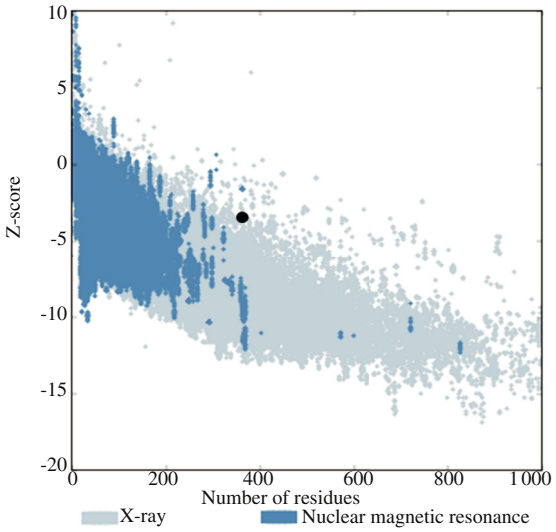

Figure 4. The ProSA-web Z-score plot.
The Z-scores of all recorded protein chains in Protein Data Bank determined by X-ray crystallography (light blue) or nuclear magnetic resonance spectroscopy (dark blue) with respect to their number of residues. The Z-score plot for QD (≤ 10) is shown as a black dot.

Table 3
The number and position of turns along with full QD sequence predicted from primary structure of QD protein.

Conformation	Number	Position
α-helix	9	1-17, 52-61, 69-81, 163-173, 189-196, 216-224, 262-267, 315-319, 325-352
β-strand	14	17-21, 33-39, 43-47, 88-92, 127-129, 141-143, 152-160, 179-181, 185-188, 224-226, 235-242, 267-270, 274-281, 319-322
Random coil	18	21-33, 39-43, 47-52, 61-69, 81-88, 92-127, 129-141, 143-152, 160-163, 173-179, 181-185, 196-216, 226-235, 242-262, 270-274, 281-315, 322-325, 352-367
α turns	6	21-30, 111-140, 246-255, 286-295, 305-310, 346-365
β turns	18	19-35, 39-43, 63-71, 91-94, 97-100, 105-152, 158-163, 172-184, 188-192, 196-200, 206-212, 218-224, 229-233, 242-256, 269-273, 282-299, 304-313, 320-367
γ turns	18	21-33, 40-43, 65-67, 97-100, 108-143, 147-152, 159-162, 174-177, 189-191, 207-209, 230-233, 243-256, 271-273, 283-299, 306-313, 327-332, 340-343, 345-367

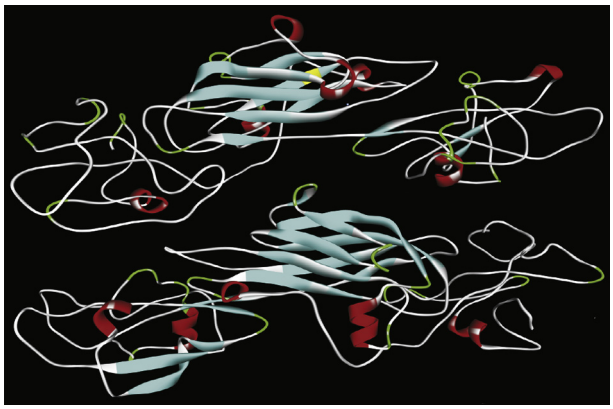

Figure 5. The 3D structure of QD protein.
Before (up) and after (down) refinement process as viewed by ChemBio office software.

Table 4
Comparison of conformational change of QD structure before and after refinement process.

Position	Before refinement	After refinement
42-43, 74-75-76, 80-81, 88-89, 114-115, 123, 134, 158, 178-179, 188, 194-195-196, 201-202-203, 228-229, 238-239	Coil	Sheet
164-165, 247-248, 328-327, 333-332	Coil	Turn
254	Coil	Helix
33-34, 125-126, 205, 302-303-304	Turn	Coil
61-62-63-64, 172	Turn	Helix
54-55-56, 189-190-191-192	Helix	Coil
197-198-199, 250	Helix	Turn

prediction method. The prediction showed the shortest sequence consisting of 6 amino acids with the start position of 265 and the longest one consisting of 25 residues with the start position of 57. Overall prediction for the QD by VaxiJen server classified QD as probable antigens with a value of 0.63, above the normal threshold value of 0.40.

3.7. Identification of B-cell and T-cell epitopes

A good antigenic and immunogenic vaccine candidate should be hydrophobic and can provoke both B-cell and T-cell mediated immunity. Therefore, QD protein was first subjected to BCpreds server. The server generated 18 linear B-cell epitopes with specificity of 75%.

In general, antigenic epitopes having BCpreds cutoff value of > 0.6 were selected. The sequences were further checked by using ABCpred and VaxiJen servers based on former criteria as mentioned (Table 6).

The prediction results by three servers showed the five topmost scoring patches centered on residues Q26 (score: 2.136), G117 (score: 2.120), E172 (score: 1.872), S208 (score: 1.713) and T174 (score: 1.175). The patches Q26 and G117 were located in the N-terminal while the patches centered at T174, E172 and S208 were located at the middle region of the QD protein. The IEDB BepiPred tool predicted 11 continuous B-cell epitopes in variable length. The longest one was composed of 58 amino acids in contrast to the shortest one of 7 amino acids. Most of them got high VaxiJen scores (Table 7). Furthermore, several surface-exposed conformational epitopes were discovered by CBTOPE (Table 8) and DiscoTope server (Table 9).

The QD full-length sequences were analyzed for T-cell restricted epitopes. For screening T-cell restricted antigenic epitopes, ProPred (HLA-DR restricted patches) (Table 10) and ProPred-I (47 alleles for MHC class-I binding sites) (Table 11) identified several common T-cell restricted epitopes that shared continuous B-cell epitope sequences.

These antigenic epitopes can provoke both humoral and cellular mediated immunity (Table 10).

3.8. In vitro production

By specific primers and PCR reaction, double digestion and sequencing method, we identified the transformed construct. The

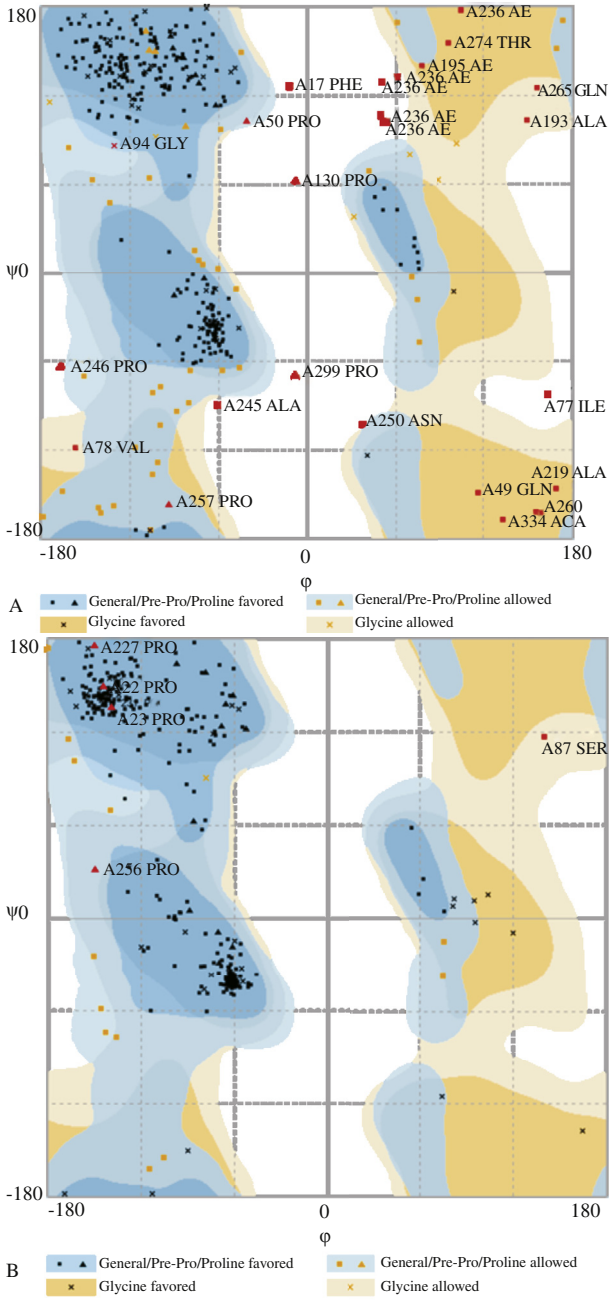

Figure 6. The QD protein structure validation before (A) and after refinement (B) by using Ramachandran plot.

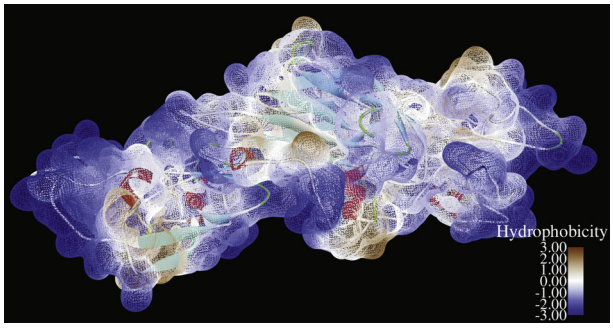

Figure 7. The 3D hydrophobicity modeling of QD after refinement process.
The brown color represents the most hydrophobic region while the blue color indicates antibody-accessible and high-score hydrophilic residues.

Table 5
The highest scores of antigenic sequences in the full-length QD by using Kolaskar and Tangaonkar prediction method by IEDB B-cell analysis resource.

Start position	End position	Peptide	Length
10	21	AADIAKLFQSVT	12
45	51	IIAYQPQ	7
57	81	LRRIVSQLDIPVRQVMIEARIVEAN	25
121	130	GPIAGSCTFP	10
137	145	SPSPFVDLG	9
163	169	ILDLQLS	7
179	188	IVSQPKVVTS	10
198	206	GSEVPYQEA	9
218	228	EAALSLEVTPQ	11
234	244	RIIVEVKVTKD	11
265	270	AKILVN	6
274	280	TIVIGGV	7
287	304	KSVEKVPFLGELPYLGRL	18
316	323	ELLVFLTP	8
346	355	AKGVWACKST	10

expected band of the PCR product (658 bp) was detected (Figure 8A).

The integrity of pET28a(+)/QD plasmid was confirmed by double digestion method (Figure 8B).

The higher amount of recombinant protein was achieved by incubation in LB broth containing kanamycin (50 μg/mL) medium and the addition of 2.5 mmol/L IPTG for 6 h and 8 h induction by IPTG at 37 °C and with shaking at 250 rRPM (Figure 9).

The reasonable yield of pure QD protein was obtained in this study (Figure 10). By the Bradford method, up to 2 mg of the recombinant QD from a 500 mL IPTG-inoculated culture was detected.

Finally, the purified QD was confirmed by western immunoblotting method (Figure 11).

Table 6
The QD protein epitopes identified by BCpreds (BCPred + AAP), ABCpred and VaxiJen modules.

Position	Epitopes	BCpreds score	ABCpred score	VaxiJen score
169	SAMEKTGNGEIVSQPKVVTS	1.000	0.84	0.9800
49	QPQERLDELRRIVSQLDIPV	1.000	0.65	0.0755
121	GPIAGSCTFPTTGTSKSPSP	1.000	0.84	0.5900
190	KETAKILKGSEVPYQEASSS	1.000	0.76	0.3096
249	QNMLNGVPPINKNEVNAKIL	1.000	0.68	0.3111
25	GQEGKEGGRGSITVDDRTNS	1.000	0.93	1.8956
196	LKGSEVPYQEASSSGATSTS	0.996	0.82	1.4226
98	KGNWSGYGKDGNIGIKDEDG	0.994	0.79	1.1962
324	RIMNEAAAKEAAAKEAAAKE	0.973	0.68	1.0548
172	EKTGNGEIVSQPKVVTSDKE	0.970	0.82	1.0544
93	GGAYHKGNWSGYGKDGNIGI	0.914	0.87	0.5406
224	EVTPQITPDNRIIVEVKVTK	0.887	0.87	0.7278
246	PDYQNMLNGVPPINKNEVNA	0.816	0.71	–0.0414
348	GVWACKSTQDPMFTPKGCDN	0.777	0.88	0.6156
346	AKGVWACKSTQDPMFTPKGC	0.688	0.92	0.1984
227	PQITPDNRIIVEVKVTKDAP	0.475	0.69	0.5948

Table 7
The prediction of the significant liner B-cell epitopes from QD protein by using IEDB analysis resource.

Start position	End position	Peptide	Peptide length	VaxiJen score
20	41	VTSDGGQEGKEG GRGSITVDDR	22	2.3713
47	55	AYQPQERLD	9	0.8772
82	85	VGYDKSLGVR	4	1.3896
95	152	AYHKGNWSGYGKD GNIGIKDEDGMNCGPI AGSCTFPTTGTSKSPS PFVDLGAKDATSG	58	0.8271
171	181	MEKTGNGEIVS	11	2.0123
184	193	KVVTSDKETA	10	0.8268
200	216	EVPYQEASSSGATSTSF	17	1.2107
226	232	TPQITPD	7	2.6047
241	249	VTKDAPDYQ	11	0.6501
308	316	DTVTDRKNE	9	1.1910

Table 8
Prediction of QD discontinuous B-cell epitopes by using CBTOPE server.

Amino acid	Position	Probability scale
DIAKL	12-16	4
TSDG	21-24	4
TVDD	36-40	4
A	47	4
DIP	65-67	4
V	78	4
GYD	83-85	4
SLGVRW	87-92	4
KG	98-99	4
S	102	4
KDGNIGIKDEDGM	106-118	4
GP	121-122	4
CTF	127-129	4
L	144	4

(*continued on next page*)

Table 8 (*continued*)

Amino acid	Position	Probability scale
IGI	153-155	4
FI	157-158	4
L	168	4
KTG	173-175	4
EIV	178-180	4
PK	183-184	4
V	201	4
Y	203	4
EA	205-206	4
SSGATS	208-213	4
KE	217-218	4
EV	224-225	4
VT	241-242	4
DA	244-245	4
QNMLN	249-253	5
GVP	254-256	4
A	265	4
G	279	4
EQS	284-286	4
VP	292-293	4
FL	294-295	5
GE	296-297	4
RRDT	306-309	4
VTD	310-312	5
RKN	313-315	4
NE	327-328	4
C	352	4
QD	356-357	4
PMFT	358-361	5
PKGCDN	362-367	4

Table 10

Common T-cell restricted epitopes from QD protein (51 human MHC class-II alleles with threshold value = 4) that may induce both the B-cell and T-cell immunity.

Position	HLA-DR alleles restricted epitopes	Number of MHC class II binding alleles	VaxiJen score	IC_{50} value
36-44	ITVDDRTNS	8	1.7011	5.45
82-90	VGYDKSLGV	15	0.5413	217.27
104-112	YGKDGNIGI	10	2.0035	109.14
117-126	MNCGPIAGS	9	0.7500	45.08
129-137	FPTTGTSKS	6	1.8216	5.20
141-149	FVDLGAKDA	11	1.2758	2.24
185-193	VVTSDKETA	11	1.2436	602.56
229-237	ITPDNRIIV	8	0.3716	98.40
237-245	VEVKVTKDA	14	1.2626	3 118.89

Table 11

Common MHC I (47) alleles restricted epitopes from recombinant QD protein with 4% threshold that may induce both the B-cell and T-cell immunity.

Position	Predicted MHC class-I alleles restricted epitopes	Number of MHC class I binding alleles	VaxiJen score	IC_{50} value
37-45	TVDDRTNSI	12	0.7172	4.95
96-104	YHKGNWSGY	14	2.0035	21.48
98-106	KGNWSGYGK	5	0.7500	1 931.97
121-129	GPIAGSCTF	10	1.8216	1 918.67
127-136	CTFPTTGTSK	3	1.2758	81.66
133-141	GTSKSPSPF	8	−0.8234	198.61
324-332	RIMNEAAAK	8	0.5258	5.74

Table 9

Prediction of discontinuous B-cell epitopes from QD by using DiscoTope server.

Amino acid	Position	Contact number	Propensity score	DiscoTope score
Y	7	14	−2.290	−3.636
S	22	7	−2.998	−3.458
G	24	4	−1.933	−2.171
QEGKEG	26-31	10-22-7-8-26-0	0.459, −0.313, 0.815, 1.007, 0.456, 1.125	−0.744, −2.807, −0.084, −0.029, −2.587, 0.996
RG	33-34	1-13	−0.986, −2.462	−0.988, −3.674
D	39	0	−4.156	−3.678
YQPQERLDELRRI	48-60	19-6-1-11-7-2-3 -0-8-16-20-4-8	−1.403, −0.731, 1.235, 2.345, 1.305, 1.660, 2.948, 2.694, 1.891, 1.060, 0.250, −1.221, −2.966	−3.426, −1.337, 0.978, 0.810, 0.350, 1.239, 2.264, 2.384, 0.754, −0.902, −2.079, −1.541, −3.545
SQLDI PVRQ V	62-71	3-14-18-4-5-15 -13-10-0-8	−3.243, −1.783, −0.927, −1.955, −0.891, −2.208, −2.414, −1.784, −2.247, −2.390	−3.215, −3.188, −2.891, −2.191, −1.364, −3.679, −3.631, −2.729, −1.989, −3.035
GAYHKGN WSGYGKD	94-107	20-16-11-17-5-0-1 -13-1-1-24-0-8-20	−0.899, 0.265, 2.291, 3.065, 4.598, 5.072, 5.465, 4.935, 5.282, 4.256, 1.919, 1.220, −0.256, −0.418	−3.095, −1.605, 0.762, 0.757, 3.495, 4.489, 4.722, 2.872, 4.560, 3.652, −1.062, 1.080, −1.146, −2.670
AS	206-207	2-5	−1.490, −2.991	−1.549, −3.222
NGVPPINKNE	253-262	7-11-6-14-4-1 -8-14-3	−0.789, 0.008, −0.893, −1.301, 0.586, 0.035, −0.365, −1.164, −1.670	−1.504, −1.258, −1.480, −2.762, 0.059, −0.084, −1.243, −2.640, −1.823
EQ	284-285	5-0	−2.710, −2.378	−2.974, −2.104
V TDR	310-313	12-0-8-0	−2.610, −0.544, −0.226, −1.914	−3.690, −0.481, −1.120, −1.694
T	355	0	−3.961	−3.505
KGC	363-366	3-7-6	−0.506, −0.822, −2.355	−0.792, −1.533, −2.774

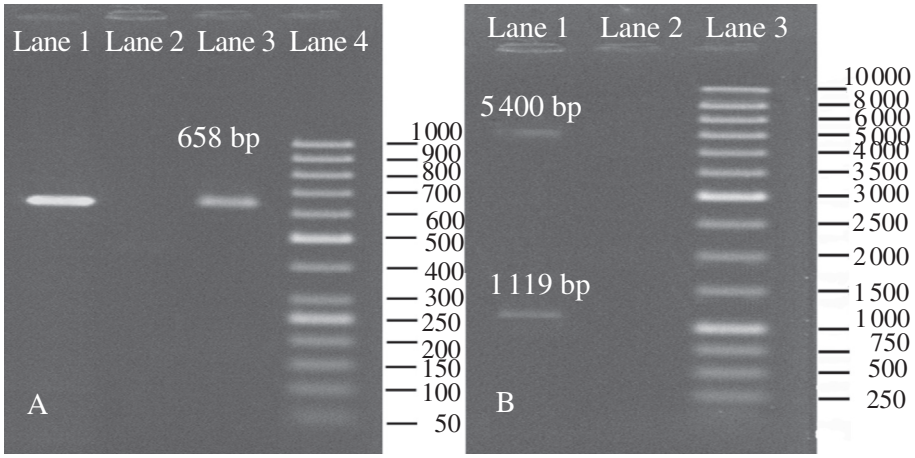

Figure 8. Identification of pET28a(+)/QD by PCR and restriction enzyme digestion.
A: Electrophoresis analysis (1%) and expected band of chimeric QD gene after PCR; Lane 1: 50 bp DNA ladder; Lane 2: Single expected band of QD gene (658 bp) isolated from positive recombinant *E. coli* BL21 (DE3) strains; Lane 3: Negative control; Lane 4: PCR product of QD gene in frame of synthetic pET28a(+)/QD as positive control; B: Double digestion of recombinant pET28a(+)/QD with Bam HI, Hind III restriction enzymes; Lane 1: Double digestion result of the pET28a(+)/QD (pET28a: 5 400 bp and chimeric QD: 1 119 bp); Lane 2: Negative control; Lane 3: 1 kb DNA marker.

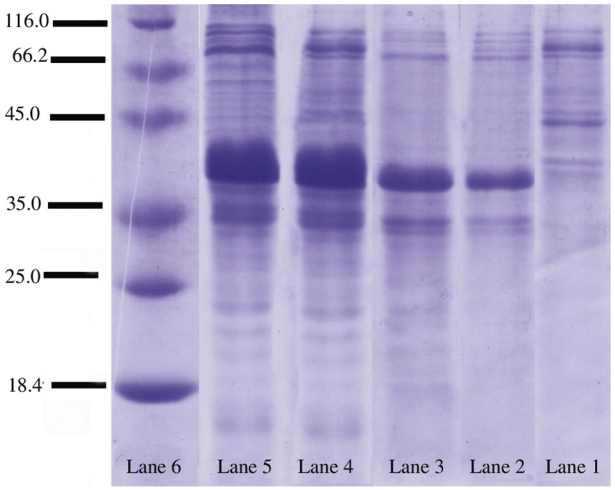

Figure 9. Analysis of chimeric QD protein expression at different harvesting time in 2.5 mmol/L IPTG concentration.
Lane 1: No-induced; Lane 2: 2 h after induction; Lane 3: 4 h after induction; Lane 4: 6 h after induction; Lane 5: 8 h after induction; Lane 6: marker.

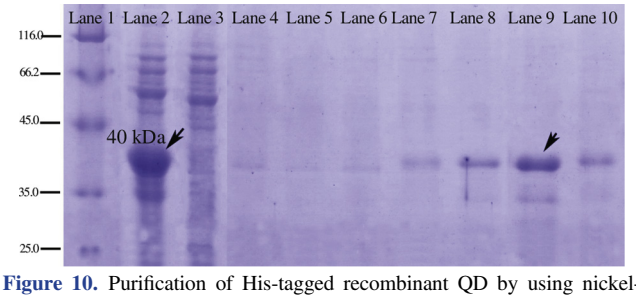

Figure 10. Purification of His-tagged recombinant QD by using nickel-nitrilotriacetic acid agarose. The fractions were analyzed by SDS-PAGE.
Lane 1: Marker proteins (Fermentas# SM0431); Lane 2: Cleared lysate containing induced recombinant QD; Lane 3: Flow through; Lane 4: Washed with the buffer containing 20 mmol/L imidazole; Lane 5: Washed with the buffer containing 50 mmol/L imidazole; Lanes 6-7: Washed with the buffer containing 150 mmol/L imidazole; Lane 8: Washed with the buffer containing 250 mmol/L imidazole; Lane 9: Eluted recombinant protein with modified method and buffer; Lane 10: Eluted recombinant protein with standard method.

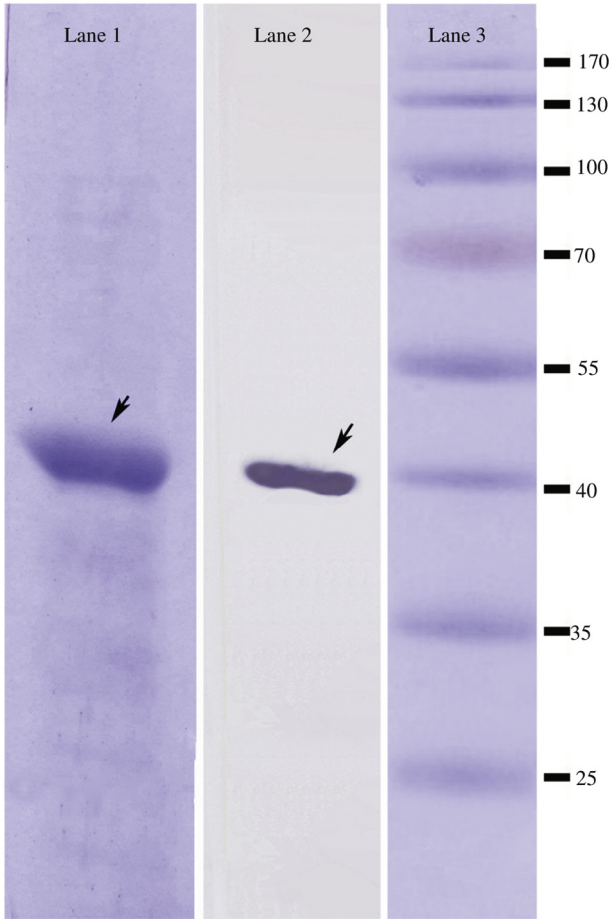

Figure 11. Western blotting analysis of expressed recombinant QD by using His-Tag antibody.
Lane 1: Coomassie blue-stained sodium dodecyl sulfate-polyacrylamide gel electrophoresis; Lane 2: Western blots with mouse monoclonal anti-His-Tag antibody; Lane 3: M-pre-stained protein weight marker (Thermo Scientific Fermentas).

4. Discussion

Host's innate and adaptive immunity plays a significant role in prevention, recognition, clearance and controlling of acute *P. aeruginosa* infection. However, *P. aeruginosa* may develop several strategies to circumvent the host's immune response and also antibiotic tolerance. Despite all efforts, the future of immunization against *P. aeruginosa* remains controversial and there still is no a truly effective vaccine available[2]. Former studies showed that efficiency of immunization with outer membrane proteins and T4P components induced enhanced clearance of *P. aeruginosa*[2]. In this study, the highly conserved segments of PilQ combined to DSL domain as an adhesive part of T4P major subunit is assumed to be a strong and protective immunogen against most of *P. aeruginosa* strains[3]. The roles of linkers to improve immunogenicity and reduce reactogenicity of fusion protein have been confirmed. In this study, the two domains of recombinant QD (the chimeric PilQ and the DSL region of PilA) were separated by four repeated Glu-Ala-Ala-Ala-Lys peptide sequences as a helix-forming linker. The three-repeated alanine between Glu^- and Lys^+ as solvent-exposed salt bridge stabilized α-helix structure formation[41]. Small-angle X-ray scattering based studies showed that the longer linkers (n = 4, 5) in comparison to the short linkers (n = 2, 3) can efficiently solvate chimeric proteins. In addition, α-helix-forming linkers such as four repeated EAAAK sequence are more elongated than flexible linkers, and can separate different domains efficiently[42].

Estimated pI value (pI < 7) showed acidic property of QD which facilitated the efficient purification process. The low quantity coefficient at 280 nm represents the low number of tryptophan, tyrosine, and disulfide bonds. The high aliphatic index indicates the thermal stability of QD for a wide range of temperature. The low rate of grand average of hydropathicity index of recombinant QD (−0.342) indicates that our chimeric protein is hydrophilic molecule. ProtParam server recognized QD as a stable protein (instability index: 32.09). The estimated half-life of greater than 10 h in *E. coli* is attractive characteristic for QD as a vaccine candidate. The PilQ as a member of the secretin family of outer membrane proteins is involved in secretion of T4P in *P. aeruginosa* in acute phase of infection[43].

Up to date, several vaccine candidates have been analyzed against *P. aeruginosa* infections[2,44], as regards, QD was not listed as an effective immunogen and vaccine candidate. However, native PilQ was just introduced as good Th17-stimulating protein[11]. The high solubility (99.5%) implies that QD could be purified under native state when expressed in *E. coli*. Random coil in QD (53.13%) was the frequent structure. In spite of the fact that a protein with its native 3D structure could be an excellent candidate for immunogenic studies, QD should be purified with denaturants like guanidine hydrochloride or urea. Unlike majority of the proteins, the QD secondary structures contain higher percentage of random coils and β-sheet. Random coil structures allow for better recognition of conformational epitope by antibodies[45]. The turn motifs are located on the surface of proteins where they could be accessed easily to immune system. The numbers of β, and γ turns in compare to α motif in the QD were high. Structural and functional characterization of a protein is one frequent dilemma in biology.

In this study, both comparative and *ab initio* methods were considered for prediction of 3D structure of QD protein. The appropriate accuracy and reliability assessments revealed that the predicted 3D structure of I-TASSER could be accurate and reliable. The structure of QD was refined in atomic level and the initial model closer to the native structure was obtained.

The accuracy of the predicted QD 3D structure is clarified by TM-score and C-score. The TM score or TM-score as a measure of similarity between two protein structures is more accurate than the often used root-mean-square deviation. The TM-score between (0, 1], indicates the difference between two structures where 1 indicates a perfect match between two structures. The estimated TM-score (0.48 ± 0.15) confirmed the precision of the predicted model. A TM-score > 0.5 generally points to the correct topology. C-score is used for estimating the quality of predicted models by I-TASSER in the range of [−5, 2], where a C-score of higher value implies a model with a high confidence. These clarify accuracy of the predicted QD 3D structure. Z-sore as a symbol of the deviation of total energy of molecular structure categorizes QD in the range of native conformations. The Ramachandran plot revealed desirable stability in the chimeric QD protein. The ability to induce adaptive immune response is another main feature of an efficient vaccine candidate. B-cell and T-cell epitope mapping is critical strategy in vaccine design possess and antibody production. The BCPreds was the main modules to predict continuous B-cell epitopes. Several high score 20 mers B-cell restricted epitopes were achieved by BCPreds server and most of them obtained the high ABCpred and VaxiJen scores.

Polar amino acids are those with side chains that prefer to reside in an aqueous (*i.e.* water) environment. All the predicted epitopes were rich of polar amino acids. The amino acids Glu and Asn were the most abundant clearly polar amino acids in epitope sequences. Specific AAPs of epitope sequences Y: G and P: D were presented in the predicted sequences[46]. Further comparative analysis of the secondary structure results and predicted linear epitope sequences showed almost of the continuous B-cell epitope patches in antibody-accessible areas which significantly enriched with (irregular) random coil and turn structures compared to non-epitopic parts. These flexible secondary structures involve in conformational change of epitopes upon antibody binding. It has been observed that almost 90% of all B-cell epitopes are conformational epitopes[33]. CBTOPE and DiscoTope servers discovered a wide range of discontinuous B-cell epitope protein which scattered throughout the QD protein. Since the DiscoTope calculates surface accessibility and predicts discontinuous B-cell epitopes from protein 3D structures, it seems to be more reliable. If enough quantities of peptide from an antigen bind to a MHC on the surface of antigen presenting cells, the T-cell can prompt the cellular immunity[47].

Despite of progress in various computational methods on evaluating the success of various MHC-peptide binding prediction, there is no consensus on a perfect method[48]. However, in this study, ProPred and ProPred-I servers discovered a lot of HLA-A, HLA-B, HLA-C (MHC class I proteins) and HLA-DR (MHC class II) restricted epitopic fractions from full-length QD protein. ProPred or ProPred-I recognized that QD protein could bind to several different types of HLA alleles which could trigger immune response. In the next step, comparative analysis of the continuous B-cell epitope results and predicted T-cell restricted epitopes showed several T-cell epitopes

derived from B-cell binding sequences from QD protein which possibly will evoke strong both humoral and cell-mediated immune responses. The two peptide sequences of QD antigen, "AYHKGNWSGYGKDGNIGIKDEDGMNCGPIAGSCT FPTTGTSKSPSPFVDLGAKDATSG" and "GPIAGSCTFPTT GTSKSPSP" possessed these properties. In this study, *E. coli* BL21 (DE3) strain as suitable prokaryotic expression system could express heterologous QD protein. In this study, the *E. coli* BL21 (DE3) strain showed significant expression after 6–8 h of induction at 37 °C. There was no toxic effect of QD protein on the *E. coli* BL21 cells. The poly-histidine-tag sequence efficiently helped the purification of QD by Ni^{2+}-affinity chromatography method. Our antigen remains to be tested for the immunological efficacy.

Several regions of QD recombinant protein from *P. aeruginosa* were found to be efficient antigens. Almost all conserved QD antigenic patches were enriched with irregular coil and turn motif. Our predictions showed that the QD chimeric protein with several T-cell epitopes derived from B-cell epitopes could be expressed efficiently in *E. coli* BL21 (DE3) strain and could serve as a good subunit of vaccine candidate against *P. aeruginosa*.

Conflict of interest statement

The authors report no conflict of interest.

Acknowledgments

The authors acknowledge Department of Medial Microbiology, Shahid Beheshti University of Medical Sciences, Tehran, Iran and useful suggestions made by Dr. Mojgan Bandepour from Cellular and Molecular Biology Research Center, School of Medicine, Shahid Beheshti University of Medical Sciences.

We declare that Shahid Beheshti University of Medical Sciences supported this work and that there is no conflict of interests about publication of this paper.

References

[1] Staudinger BJ, Muller JF, Halldórsson S, Boles B, Angermeyer A, Nguyen D, et al. Conditions associated with the cystic fibrosis defect promote chronic *Pseudomonas aeruginosa* infection. *Am J Respir Crit Care Med* 2014; **189**(7): 812-24.

[2] Priebe GP, Goldberg JB. Vaccines for *Pseudomonas aeruginosa*: a long and winding road. *Expert Rev Vaccines* 2014; **13**(4): 507-19.

[3] Giltner CL, Nguyen Y, Burrows LL. Type IV pilin proteins: versatile molecular modules. *Microbiol Mol Biol Rev* 2012; **76**(4): 740-72.

[4] Sato H, Okinaga K, Saito H. Role of pili in the pathogenesis of *Pseudomonas aeruginosa* burn infection. *Microbiol Immunol* 1988; **32**(2): 131-9.

[5] Bucior I, Pielage JF, Engel JN. *Pseudomonas aeruginosa* pili and flagella mediate distinct binding and signaling events at the apical and basolateral surface of airway epithelium. *PLoS Pathog* 2012; **8**(4): e1002616.

[6] Cachia PJ, Hodges RS. Synthetic peptide vaccine and antibody therapeutic development: prevention and treatment of *Pseudomonas aeruginosa*. *Biopolymers* 2003; **71**(2): 141-68.

[7] Heiniger RW, Winther-Larsen HC, Pickles RJ, Koomey M, Wolfgang MC. Infection of human mucosal tissue by *Pseudomonas aeruginosa* requires sequential and mutually dependent virulence factors and a novel pilus-associated adhesin. *Cell Microbiol* 2010; **12**(8): 1158-73.

[8] Sheth HB, Glasier LM, Ellert NW, Cachia P, Kohn W, Lee KK, et al. Development of an anti-adhesive vaccine for *Pseudomonas aeruginosa* targeting the C-terminal region of the pilin structural protein. *Biomed Pept Proteins Nucleic Acids* 1995; **1**(3): 141-8.

[9] Choi DS, Kim DK, Choi SJ, Lee J, Choi JP, Rho S, et al. Proteomic analysis of outer membrane vesicles derived from *Pseudomonas aeruginosa*. *Proteomics* 2011; **11**(16): 3424-9.

[10] Koo J, Tang T, Harvey H, Tammam S, Sampaleanu L, Burrows LL, et al. Functional mapping of PilF and PilQ in the *Pseudomonas aeruginosa* type IV pilus system. *Biochemistry* 2013; **52**(17): 2914-23.

[11] Wu W, Huang J, Duan B, Traficante DC, Hong H, Risech M, et al. Th17-stimulating protein vaccines confer protection against *Pseudomonas aeruginosa* pneumonia. *Am J Respir Crit Care Med* 2012; **186**(5): 420-7.

[12] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994; **22**(22): 4673-80.

[13] Chen X, Zaro JL, Shen WC. Fusion protein linkers: property, design and functionality. *Adv Drug Deliv Rev* 2013; **65**(10): 1357-69.

[14] Ren J, Wen L, Gao X, Jin C, Xue Y, Yao X. DOG 1.0: illustrator of protein domain structures. *Cell Res* 2009; **19**(2): 271-3.

[15] Bhasin M, Garg A, Raghava GP. PSLpred: prediction of subcellular localization of bacterial proteins. *Bioinformatics* 2005; **21**(10): 2522-4.

[16] Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 2004; **340**(4): 783-95.

[17] Shen HB, Chou KC. Signal-3L: a 3-layer approach for predicting signal peptides. *Biochem Biophys Res Commun* 2007; **363**(2): 297-303.

[18] Diaz AA, Tomba E, Lennarson R, Richard R, Bagajewicz MJ, Harrison RG. Prediction of protein solubility in *Escherichia coli* using logistic regression. *Biotechnol Bioeng* 2010; **105**(2): 374-83.

[19] Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 1982; **157**(1): 105-32.

[20] Hopp TP, Woods KR. Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad Sci U S A* 1981; **78**(6): 3824-8.

[21] Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett* 1990; **276**(1–2): 172-4.

[22] Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* 2007; **8**: 4.

[23] Kaur H, Raghava GP. Prediction of alpha-turns in proteins using PSI-BLAST profiles and secondary structure information. *Proteins* 2004; **55**(1): 83-90.

[24] Kaur H, Raghava GP. Prediction of beta-turns in proteins from multiple alignment using neural network. *Protein Sci* 2003; **12**(3): 627-34.

[25] Kaur H, Raghava GP. A neural-network based method for prediction of gamma-turns in proteins from multiple sequence alignment. *Protein Sci* 2003; **12**(5): 923-9.

[26] Kaur H, Raghava GP. Role of evolutionary information in prediction of aromatic-backbone NH interactions in proteins. *FEBS Lett* 2004; **564**(1-2): 47-57.

[27] Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 2010; **5**(4): 725-38.

[28] 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**(Web Server issue): W407-10.

[29] Xu D, Zhang Y. Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophys J* 2011; **101**(10): 2525-34.

[30] El-Manzalawy Y, Dobbs D, Honavar V. Predicting linear B-cell epitopes using string kernels. *J Mol Recognit* 2008; **21**(4): 243-55.

[31] Saha S, Raghava GP. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* 2006; **65**(1): 40-8.
[32] Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Res* 2006; **2**: 2.
[33] Ansari HR, Raghava GP. Identification of conformational B-cell epitopes in an antigen from its primary sequence. *Immunome Res* 2010; **6**: 6.
[34] Kringelum JV, Lundegaard C, Lund O, Nielsen M. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS Comput Biol* 2012; **8**(12): e1002829.
[35] Singh H, Raghava GP. ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics* 2003; **19**(8): 1009-14.
[36] Singh H, Raghava GP. ProPred: prediction of HLA-DR binding sites. *Bioinformatics* 2001; **17**(12): 1236-7.
[37] Guan P, Doytchinova IA, Zygouri C, Flower DR. MHCPred: bringing a quantitative dimension to the online prediction of MHC binding. *Appl Bioinformatics* 2003; **2**(1): 63-6.
[38] Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 2012; **13**: 134.
[39] Reichelt P, Schwarz C, Donzeau M. Single step protocol to purify recombinant proteins with low endotoxin contents. *Protein Expr Purif* 2006; **46**(2): 483-8.
[40] Legaree BA, Daniels K, Weadge JT, Cockburn D, Clarke AJ. Function of penicillin-binding protein 2 in viability and morphology of *Pseudomonas aeruginosa*. *J Antimicrob Chemother* 2007; **59**(3): 411-24.
[41] Arai R, Ueda H, Kitayama A, Kamiya N, Nagamune T. Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng* 2001; **14**(8): 529-32.
[42] Arai R, Wriggers W, Nishikawa Y, Nagamune T, Fujisawa T. Conformations of variably linked chimeric proteins evaluated by synchrotron X-ray small-angle scattering. *Proteins* 2004; **57**(4): 829-38.
[43] Gellatly SL, Hancock RE. *Pseudomonas aeruginosa*: new insights into pathogenesis and host defenses. *Pathog Dis* 2013; **67**(3): 159-73.
[44] Döring G, Pier GB. Vaccines and immunotherapy against *Pseudomonas aeruginosa*. *Vaccine* 2008; **26**(8): 1011-24.
[45] Saad B, Corradin G, Bosshard HR. Monoclonal antibody recognizes a conformational epitope in a random coil protein. *Eur J Biochem* 1988; **178**(1): 219-24.
[46] Rubinstein ND, Mayrose I, Halperin D, Yekutieli D, Gershoni JM, Pupko T. Computational characterization of B-cell epitopes. *Mol Immunol* 2008; **45**(12): 3477-89.
[47] Davies MN, Flower DR. Harnessing bioinformatics to discover new vaccines. *Drug Discov Today* 2007; **12**(9-10): 389-95.
[48] Lafuente EM, Reche PA. Prediction of MHC-peptide binding: a systematic and comprehensive overview. *Curr Pharm Des* 2009; **15**(28): 3209-20.