



# A potent multivalent vaccine for modulation of immune system in atherosclerosis: an *in silico* approach

**Ahmad Karkhah<sup>1</sup>, Jafar Amani<sup>2</sup>**

<sup>1</sup>Cellular and Molecular Biology Research Center, Student Research Committee, School of Medicine, Babol University of Medical Sciences, Babol; <sup>2</sup>Applied Microbiology Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran

Received: September 8, 2015

Revised: October 22, 2015

Accepted: November 5, 2015

Corresponding author: Jafar Amani, PhD  
 Applied Microbiology Research Center, Baqiyatallah University of Medical Sciences, Vanak Sq. Molasadra St., Tehran, P.O. Box 19395-5487, Iran  
 Tel: +98-21-82482568, Fax: +98-21-88068924.  
 E-mail: Jafar.amani@gmail.com

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

**Purpose:** Atherosclerosis is classically defined as an immune-mediated disease characterized by accumulation of low-density lipoprotein cholesterol over intima in medium sized and large arteries. Recent studies have demonstrated that both innate and adaptive immune responses are involved in atherosclerosis. In addition, experimental and human models have recognized many autoantigens in pathophysiology of this disease. Oxidized low-density lipoproteins,  $\beta$ 2 glycoprotein I ( $\beta$ -2-GPI), and heat shock protein 60 (HSP60) are the best studied of them which can represent promising approach to design worthwhile vaccines for modulation of atherosclerosis.

**Materials and Methods:** *In silico* approaches are the best tools for design and evaluation of the vaccines before initiating the experimental study. In this study, we identified immunogenic epitopes of HSP60, ApoB-100, and  $\beta$ -2-GPI as major antigens to construct a chimeric protein through bioinformatics tools. Additionally, we have evaluated physico-chemical properties, structures, stability, MHC binding properties, humoral and cellular immune responses, and allergenicity of this chimeric protein by means of bioinformatics tools and servers.

**Results:** Validation results indicated that 89.1% residues locate in favorite or additional allowed region of Ramachandran plot. Also, based on Ramachandran plot analysis this protein could be classified as a stable fusion protein. In addition, the epitopes in the chimeric protein had strong potential to induce both the B-cell and T-cell mediated immune responses.

**Conclusion:** Our results supported that this chimeric vaccine could be effectively utilized as a multivalent vaccine for prevention and modulation of atherosclerosis.

**Keywords:** Atherosclerosis, Vaccine, HSP60, Apolipoprotein B-100, Beta 2-glycoprotein I



© Korean Vaccine Society.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Introduction

Atherosclerosis is classically defined as a chronic inflammatory state in which low-density lipoprotein cholesterol (LDL-c) accumulate over intima in medium sized and large arteries [1]. According to the World Health Organization (WHO) reports, cardiovascular diseases are the number one cause of death globally. In particular, cardiovascular diseases account for approximately 40% of all deaths in the world [2]. Hyperlipidemia, smoking, hyperhomocysteinemia, diabetes, and some infectious microorganisms such as cytomegalovirus and *Chlamydia pneumonia* are known as common risk factors which predispose people for this chronic disorder [3].

Previously, atherosclerosis was assumed merely as a passive aggregation of LDL-c

in arterial walls, but recent studies have made this hypothesis more complex and atherosclerosis is now considered as a chronic inflammatory disease [4]. Over the past decade, many investigations have demonstrated that both innate and adaptive immune systems are involved in pathogenesis of atherosclerosis [5]. Many immune cells and molecules such as dendritic cells, B and T lymphocytes, macrophages, neutrophils, and inflammatory cytokines are implicated in initiation and development of atherosclerotic lesions [6].

At the first stages of this disease, high low-density lipoprotein particles in peripheral blood circulation trap in the extracellular matrix of arterial intima. Accumulated LDL-c in sub endothelial layer is converted to oxidized low-density lipoprotein (ox-LDL) through reaction with reactive oxygen species. *In vitro* and *in vivo* experiments have shown that classical atherosclerosis risk factors lead to simultaneous expression of adhesion molecules on the surface of endothelial cells. This over expression of adhesion molecules acts as a “danger signal” which activates both the innate immune and the adaptive immune systems. Many immune cells such as monocytes and T cells recruit into the arterial intima and respond to many endogenous and exogenous antigens [7,8].

Based on above-mentioned evidence, modulation of immune system could represent a useful approach to prevent or treat this disease. Although extensive usage of various drugs such as statin family improve atherosclerosis, these drugs are associated with several side effects [9]. More studies are needed to introduce more effective strategies which confer favorable improvement in atherosclerosis. During last decade, many reports have described the idea of an immune-modulatory approach such as active immunization for improvement of atherosclerosis [10]. The concept of active immunotherapy against atherosclerosis goes back to mid-twenty century when Gero et al. [11] observed that immunization with beta-lipoproteins inhibit progression of atherosclerosis.

One of the most important steps in immunization is to know which antigens are implicated in the disease. As yet, many autoantigens including ox-LDL,  $\beta$ 2 glycoprotein I ( $\beta$ -2-GPI), and heat shock protein (HSP) 60 have recognized that play important roles in development of atherosclerosis [12].

ApoB-100, the major protein permanently associated with ox-LDL, has attracted attentions in vaccine researches. Some studies indicated that active immunization using a number of different ApoB-100 peptide sequences such as p2, p143, and p210 led to 40%-70% reduction in development of atherosclerosis [13-15].

HSPs, a super family of stress proteins classified based on their molecular weight, have been implicated in atherogenesis [16]. Active immunization with various kinds of HSPs have revealed different effects on atherosclerosis. Immunization with HSP65 and their reactive T-cell epitopes was reported to reduce atherosclerotic lesions in different models of experimental atherosclerosis [17,18]. Also, several studies have stated that immunization with ApoB-100 and HSP60 peptides together exerts synergetic atheroprotective effect [19].

A number of studies demonstrated that  $\beta$ -2-GPI have been involved in the pathogenesis of atherosclerosis [20]. Human  $\beta$ -2-GPI or apolipoprotein H is a plasma protein belongs to the complement control protein superfamily which circulates in the blood at high concentration. This protein has a pivotal role in the binding of anti-phospholipid antibodies (aPL) to negatively charged phospholipids and is identified as an important autoantigen for the development of thromboembolic complications in both arteries and the veins in some autoimmune diseases such as atherosclerosis [21,22].

Taken the above reports together, both pro-atherogenic and atheroprotective roles for the immune system in atherosclerosis have interested many researchers to design immune-modulatory strategies for active immunization against many of these autoantigens which could probably decrease progression of the atherosclerosis. Since it has been proven that *in silico* approaches are the best tools for design and evaluation of the vaccines before initiating the experimental study, in this study a new structural model containing three effective autoantigens with antigenic determinants of HSP60, ApoB-100, and  $\beta$ -2-GPI was designed. This chimeric gene was fused together by an appropriate linker and codon optimization for expression in *Escherichia coli* was carried out. The chimeric protein structure and its ability to induce CD4+ and CD8+ immune responses against these antigens were predicted by *in silico* approaches.

## Materials and Methods

### Protein retrieval and sequence analysis

The protein sequences of hHSP60, ApoB-100 and  $\beta$ -2-GPI have retrieved in FASTA format from Uniprot Knowledgebase data, have also used accession Nos. P10809, P04114, and P02749, respectively. Multiple sequence alignments were carried out using National Centre for Biotechnology Information (NCBI) at <http://www.ncbi.nlm.nih/Blast>.

### Construct design and optimization of chimeric gene

Recombinant hHSP60–ApoB-100 have constructed by fusing the C-terminal of hHSP60 and known atheroprotective sequences of ApoB-100 including p2, p143, and p210. Then, the N-terminal of  $\beta$ -2-GPI including I, II domains was fused to early structure. Final chimeric structure was constructed by adding the domain V of  $\beta$ -2-GPI in the C-terminal of constructed protein using hydrophobic GGGGS amino acid linker. *In silico* analysis has performed to optimize the chimeric gene by using Jcat software [23]. The chimeric gene was designed for cloning and expression in *E. coli*. VaxiJen server was used to predict the immunogenicity of the constructed protein and each protein alone [24].

### mRNA structure prediction

The mRNA secondary structure of the chimeric gene was analyzed by the program mfold at <http://www.bioinfo.rpi.edu/applications/mfold> [25].

### The physico-chemical parameters evaluation

The physico-chemical parameters such as theoretical isoelectric point (pI), molecular weight, total number of positive and negative residues, extinction coefficient, half-life, instability index, aliphatic index, and grand average hydropathy (GRAVY) were computed using the ExPASy's ProtParam (<http://us.expasy.org/tools/protparam.html>) [26].

### Secondary and tertiary structure prediction and analysis

The secondary protein structure prediction and analysis were carried out by Garnier-Osguthorpe-Robson secondary structure prediction server [27]. The 3D structure of the recombinant protein has generated using the I-TASSER online server which generates three-dimensional models with their confidence score (C-score) [28]. In addition, the quality of the resulting stereochemistry of structure was evaluated by Ramachandran plot in PROCHECK software (<http://www.ebi.ac.uk/thornton-srv/software/PROCHECK>).

### Prediction of B-cell epitopes

The amino acid sequence of chimeric protein was subjected to predict continuous and discontinuous B-cell epitopes employing the online servers. The first, analysis of chimeric protein for continuous B-cell epitopes was performed using Bcepred (<http://www.imtech.res.in/raghava/bcepred/>) [29]. Then, Discotope server was used (<http://www.cbs.dtu.dk/services/DiscoTope>) to predict discontinuous B-cell epitopes

from three-dimensional protein structures [30].

### Prediction of T-cell epitopes

Prediction of T-cell epitopes was carried out by CTLPred at <http://www.imtech.res.in/raghava/ctlpred/>. This server provide a direct method for prediction of cytotoxic T lymphocyte (CTL) epitopes crucially based on Quantitative Matrix (QM), Support Vector Machine (SVM), and Artificial Neural Network (ANN) in subunit vaccine design [31].

### Prediction of MHC binding peptides

Prediction of the promiscuous MHC class-I binding peptides was carried out applying ProPred-I at <http://www.imtech.res.in/raghava/propred1/> [32]. Also, this online server predicts MHC-I binding peptides for 47 various alleles. Prediction of binding peptides to class II MHC molecules was performed by means of RANKPEP online server [33].

### IgE epitopes and allergenic sites prediction

Possible allergenic sites was predicted by AlgPred. AlgPred determines allergenic sites based on similarity of known epitope with any region of protein [34].

### Validation methods

Several servers were applied to predict MHC binding peptides, B-cell and T-cell epitopes. These servers generally use SYFPEYITHI, MHCPEP, MHCBN, and jenPep databases. SYFPEYITHI and MHCPEP provide approximately 7,000 and 13,000 peptide sequences of various organisms including human, chicken, mouse, and apes, respectively that are known to bind to MHC-I and MHC-II [35,36]. Spatial proximity and the contact number are the most important difference between these databases. In addition default threshold for identification of epitopes was (-3.7) which provide the sensitivity and specificity of 0.47 and 0.75, respectively.

## Results

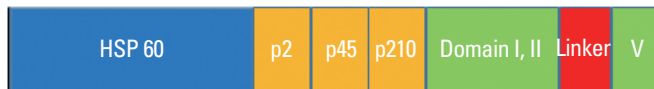
### Construct design and optimization of chimeric gene

The residues 61-361 in the N-terminal region of HSP60 have been reported to be effective in some immunization studies [4]. Also, experimental results indicated that the ApoB-100 peptide sequences are responsible for modulating of atherogenesis in active immunization of animal models (Table 1) [13,14]. For the third fragment, the amino acid residues of domain I, II, and V of  $\beta$ -2-GPI protein were selected. Based on

these sequences compared by NCBI, N-terminal region of HSP60 (300 amino acids), ApoB-100 (60 amino acids), and  $\beta$ -2-GPI (222 amino acids), it was shown that these three parts were highly conserved. Linker consisting of GGGGS amino acid was used to separate these three domains from each other. Restriction sites interfering with cloning, instability elements, and all the cis-acting sites were removed. Schematic diagram of protein domain structures with linker's sites designed is shown in Fig. 1. Codon bias and GC content were calculated in both wild type and chimeric genes. The codon adaptation index (CAI) of chimeric gene was 0.66, while that of wild type gene was only 0.63. The GC content was reduced from 50% to 49.65%, so it is expected an increase in the mRNA stability of the chimeric gene (Fig. 2). Antigen index was evaluated by VaxiJen server for HSP60, ApoB-100 and  $\beta$ -2-GPI,

**Table 1.** ApoB-100 peptide sequences used for modulating lipoprotein autoimmunity in atherosclerosis

ApoB-100 peptide	Amino acid No.	Sequence	Ref
p2	16-35	ATRFK HLRKY TYNYE AESSS	[15]
p45	661-680	IEIGL EGKGF EPTLE ALFGK	[15]
p210	3,136-3,155	KTTKQ SFDLS VKAQY KKNKH	[15]



**Fig. 1.** Schematic representation of chimeric construct consists of heat shock protein 60 (HSP 60), ApoB-100, and  $\beta$ 2 glycoprotein I ( $\beta$ -2-GPI) genes bound together by appropriate linkers for expression in *Escherichia coli*. It is important to emphasize that ApoB-100 peptides were applied as an appropriate linker as well as immunogenic epitopes in our chimeric construct.

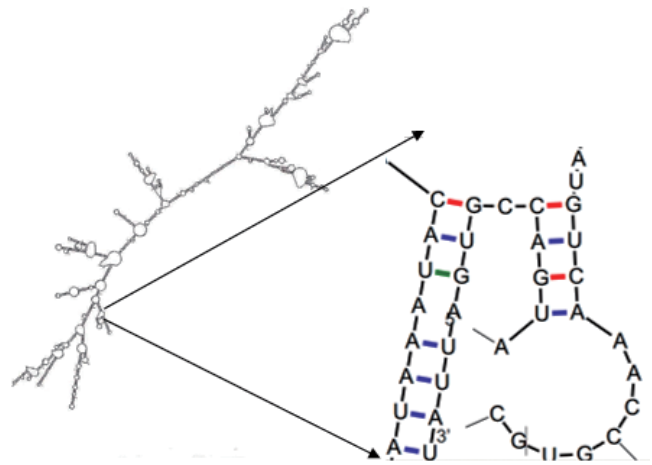
and chimeric gene which are 0.46, 1.10, 0.51, and 0.53, respectively.

**mRNA structure prediction**

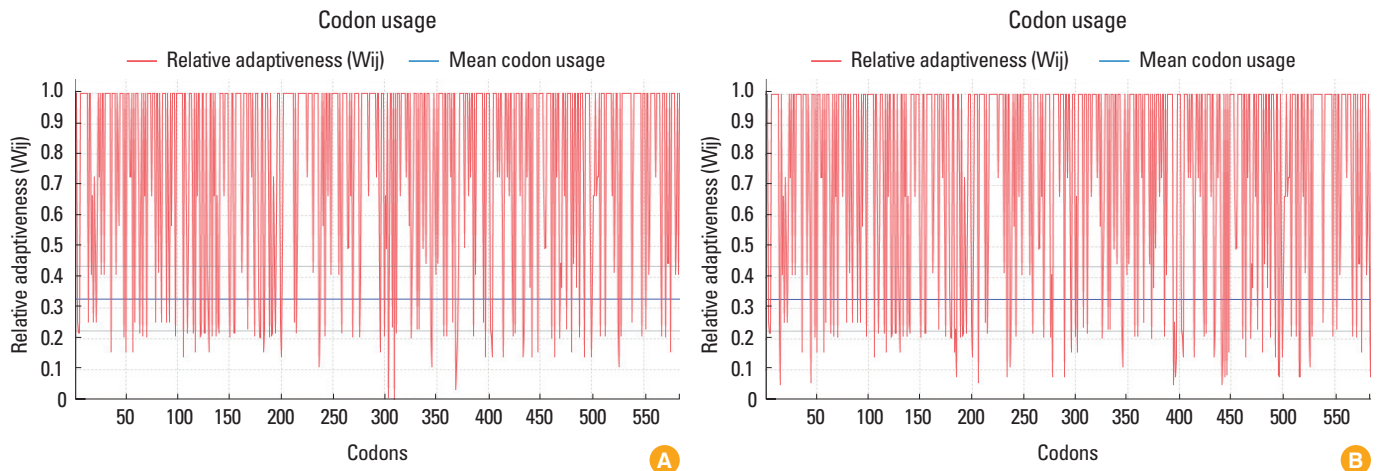
At this stage of research, the minimum free energy for secondary structures formed by RNA molecules was employed.

**Table 2.** Free energy details related to 5' end of chimeric gene mRNA structure by mfold web server

Structural element	Free energy (kcal/mol)	Base pair
External loop	-0.6	9 ss bases and 1 closing helices
Stack	-2.1	External closing pair: U2-A1756
Stack	-2.4	External closing pair: G3-C1755
Stack	-2.2	External closing pair: A4-U1754
Helix	-6.7	4 Base pairs



**Fig. 3.** Prediction of RNA secondary structure of chimeric gene by mfold server. Predicted structure has no hairpin and pseudo knot at 5' site of mRNA.



**Fig. 2.** Graphical view of codon usage in optimized chimeric gene (A) and wild type (B).

The best predicted structure had  $\Delta G = -582.10$  kcal/mol and at the site of first nucleotides at 5' was observed not a hairpin nor pseudoknot (Fig. 3). Also, Free energy details related to 5' end of chimeric gene mRNA structure is shown in Table 2.

**The physico-chemical parameters evaluation**

The calculated molecular weight and pI of chimeric protein were 64.1 kDa and 8.63, respectively. Extinction coefficient of chimeric protein at 280 nm was 46,840/M/cm. Half-life was estimated to be 30 hours in mammalian reticulocytes, more than 20 in yeast and *E. coli*. Expsy ProtParam classifies the chimeric protein as stable protein with instability index, 35.59. Aliphatic index and GRAVY of chimeric protein were defined 83.38 and -0.243, respectively.

**Secondary and tertiary structure prediction and analysis**

The secondary structure of chimeric protein was predicted by online servers are given in Fig 4. First the HSP60, ApoB-100, and  $\beta$ -2-GPI were applied as test sequences to validate our

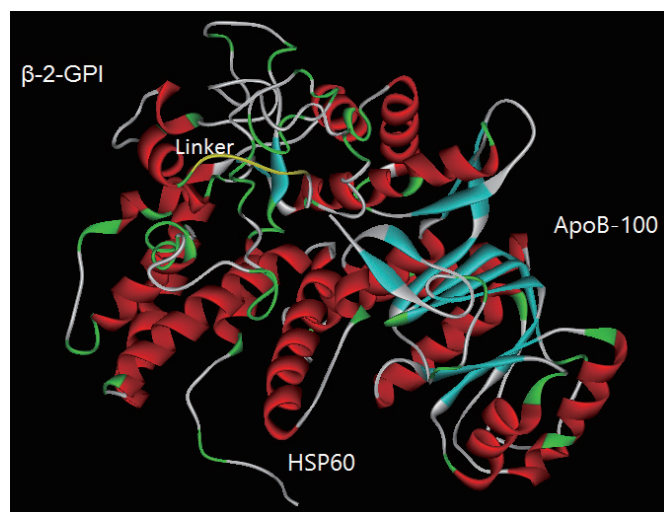
**Table 3.** Percentage of secondary structure elements of chimeric and single proteins

Sequence	Alpha helix	Extended strand	Random coil
HSP60	41.86	14.29	43.85
ApoB-100	30.00	20.00	50.00
$\beta$ -2-GPI	3.52	26.87	69.60
Chimeric gene	29.08	17.18	53.74

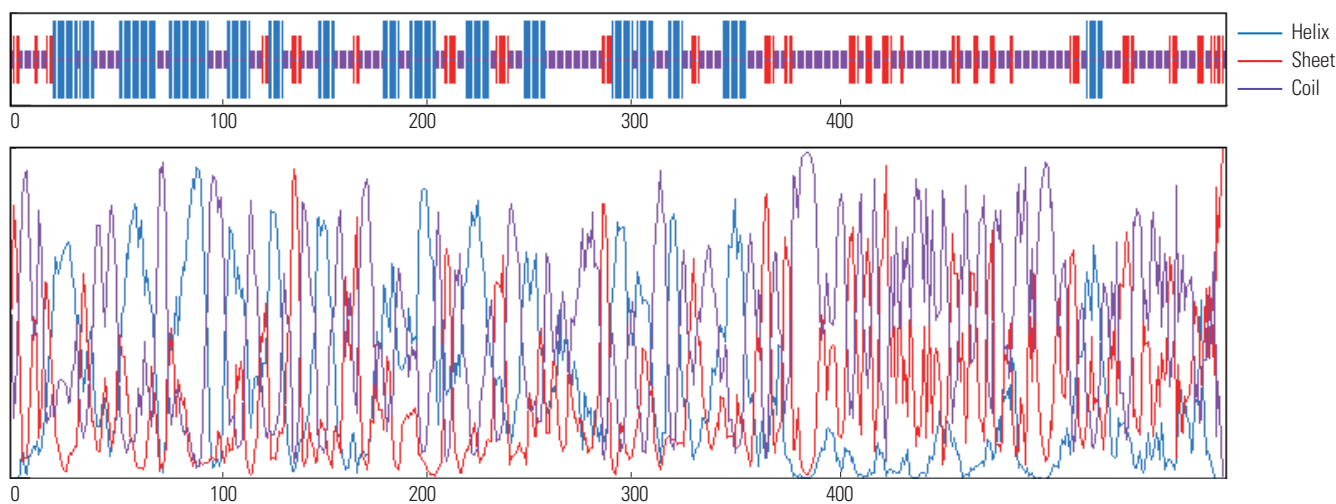
Values are presented as percentage.  
HSP60, heat shock protein 60;  $\beta$ -2-GPI,  $\beta$ 2 glycoprotein I.

secondary structure prediction method. The obtained results indicated that structural contents of protein include extended strand, random coil, and alpha helix. Additionally, predicted secondary structure for chimeric protein had 29.08% alpha helix, 17.18% extended strand and 53.74% random coil (Table 3).

The three-dimensional structure of chimeric protein was constructed by I-TASSER software. Tertiary structure of the generated protein using I-TASSER revealed a protein with three main domains linked together with linkers (Fig. 5). The C-score of models predicted by I-TASSER was -1.33. C-score



**Fig. 5.** Predicted structure of constructed protein using I-TASSER software. The three-dimensional structure generated by I-TASSER software showed a protein with three main domains linked together with appropriated linkers. HSP60, heat shock protein 60;  $\beta$ -2-GPI,  $\beta$ 2 glycoprotein I.



**Fig. 4.** Graphical results for secondary structure prediction of chimeric protein. Purple, red, and blue colors indicate extended strand, coil, and helix, respectively.

is usually in the range of -5 to and the higher the value of C-score, the higher the model confidence. In addition, the template modeling (TM)-score and root-mean-square deviation (RMSD) for this model were  $0.55 \pm 0.15$  and  $10.8 \pm 4.6$ , respectively. Chimeric protein analysis by the Ramachandran plot showed that 74.9% of amino acid residues from the structure modeled by I-TASSER were incorporated into the favored regions of the plot. Apart from that, 14.2% of residues were in allowed regions of the plot and 10.9% in outlier region (Fig. 6).

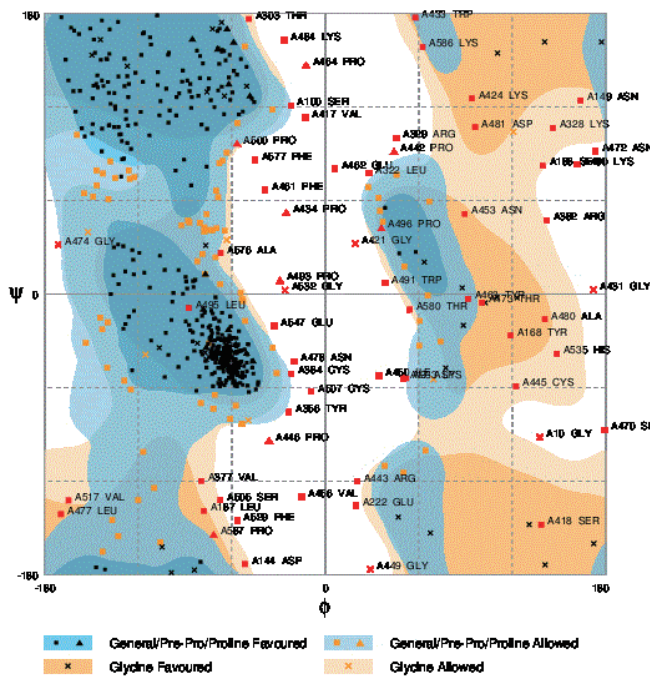


Fig. 6. Validation of protein structure by Ramachandran plot.

**Prediction of B-cell epitopes**

The continuous B-cell epitopes of constructed protein were predicted through Bcepred software. All 16-mers with Bcepreds cutoff score >0.9 in B-cell epitopes were selected and shown in Table 4. Also, continuous B-cell epitopes of chimeric protein were selected based on single characters including hydrophilicity, antigenicity, flexibility, accessibility, polarity, and exposed surface as shown in Table 5. Furthermore, obtained results indicated that did not locate any epitope at linker sites between different domains (amino acid, 501 to 505). The Discotope server was applied for prediction of discontinuous B-cell epitopes. Discotope server identified 52 B-cell epitopes residues out of 588 all residues (Table 6).

**Prediction of T-cell epitopes**

CTLpred is known as a useful server for prediction of CTL epitopes regarding to MHC class I based on QM, SVM, and

Table 4. B-Cell epitopes from full length proteins using Bcepred

Rank	Sequence	Start position	Score
1	YTTFEYPNTISFSCNT	458	0.95
2	DGTIEVPKCFKEHSSL	560	0.94
2	VWVAKAPGFGDNRKNO	238	0.94
3	GGGSSCKVPVKATV	501	0.93
4	MTVIIQSWGSPKVTK	1	0.92
5	AKSIDLKDQKYNIGAK	22	0.91
6	HVAIAGRTPKPPDLP	376	0.90
7	YKNIGAKLVQDVANNT	31	0.89
8	ELEIIEGMKFDGRYIS	151	0.89

Table 5. Epitopes predicted in chimeric protein by different parameters based on Bcepred software

Prediction parameter	Epitope sequence
Hydrophilicity	SPKVTKD, QDVANNTNEEAGDGTTTAT, SANGDKEIGN, TVKDGKTLNDE, NTSKGQKCE, AEDVDGEA, GDNRKNQ, EDVQPHD, NYEAESSKTTKQS, TCPKDD, NGADSAKCTEEGK, APGGGGSSCK, CKNKEKKCSYTEDAQ, KTDASDVKP
Flexibility	DVANNTN, EGFEKISKGA, VIAELKKQS, ATISANG, DAMKKVGR, YFINTSKGQK, VLLSEKKI, PGFGDNRKNO, YNYEAESSKTTK, GYVSRGGM, VCAPGGGGSS, KIQEFK, SFFCKNKEKK
Accessibility	QSWGSPKVTKD, KSIDLKDQKYNIGAK, QDVANNTNEEAGD, AKEGFKEI, NPVEIRR, IAELEKQSKPVTTPPEIAQ, SANGDKE, MKKVGRKG, TVKDGKTLNDEL, EGMKFDGRY, INTSKGQKCEFOD, LSEKISS, ANAHRKPL, GFGDNRKNQLKDM, NLEDVQPHDLGK, KGATRFK, EGKGFEP, LFGKRKYTYNYEAESSKTTKQSF, SVKAQYKKNKHMIS, GRTCPKDDLP, PLKTFYEPGEEIT, KPGYVSRGGMRK, RYTTFEYPNTIS, KCTEEGKWSPELP, KVPVKA, YQGERVKIQEFKNGM, FCKNKEKKCSYTEDAQ, PKCFKEH, WKTDASDVKP
Turns	DVANNTNEE
Exposed surface	SPKVTKD, IDLKDQKYN, AELKKQSKPVT, MKKVGRKG, NTSKGQK, GDNRKNQLKDM, GKRKYTYNYE, SSKTTKQS, VKAQYKKNKHM, PKPDDLP, KVPVKA, ERVKIQEFKNG, FCKNKEKKCSY
Polarity	KSIDLKDQKYN, ARSIAKEGFKEI, NPVEIRRGVM, IAELEKQSKP, AMKKVGRKG, DELEIIEGMKFD, SKGQKCEF, LSEKISSI, EIANAHRKPLVI, AEDVDGEA, GDNRKNQLKD, EDVQPHD, KGATRFKHLIE, LFGKRKYTYN, EAESSK, VKAQYKKNKHMIS, FYEPGEEIT, SRGGMRKF, SAKCTEEGKWSPE, YQGERVKIQEFKNGM, FCKNKEKKCSYTE, EVPKCFKEHSSLA
Antigenic propensity	YISPYFI, YVLLSEKK, SSIQSIVP, HRKPLVII, LSTLVNLRLKVLQV, KVGEVIVT, HLIEIGL, QSFDSL, HMISPVLLIFSSFLCHV, LPFSTVPLKTFY, ITYSCKPGYVS, RKFIPLTGLW, LKCTPRVCPF, SPELPVC, GSSCKVPVKK, DKVSFFCKN, TIEVPKCFKE

**Table 6.** Conformational B-cell epitopes from full length protein using DiscoTope server

Epitope	Position
GLU, GLN, SER, TRP, GLY, SER, PRO	3, 14, 1, 0, 5, 2, 7
GLY	1
LYS, THR, LEU, ASN	8, 0, 2, 2
GLY	4
LYS	0
LEU, GLU, ALA, GLN, TYR, LYS, LYS	6, 14, 13, 7, 8, 4, 5
GLN	1
GLU, ARG	1, 5
LYS, GLY, PHE, GLU, PRO, THR, LEU, GLU, ALA, LEU, PHE, GLY, LYS, ARG, LYS	13, 6, 22, 17, 6, 6, 19, 1, 3, 0, 16, 17, 0, 0, 10
TYR, THR, TYR, ASN, TYR, GLU, ALA	8, 15, 9, 9, 20, 20, 15
GLU	9
ALA	1
ASP	4
LYS, PRO, CYS	1, 0, 4

**Table 7.** Prediction of T-cell epitope by CTLPred

Peptide rank	Start position	Sequence	Score
1	58	TVLARSIAK	1.000
2	380	AGRTC PKPD	1.000
3	410	YSCKPGYVS	0.990

ANN in subunit vaccine design (Table 7).

**Prediction of MHC binding peptides**

Predicted peptides for binding to MHC-I over multiple alleles including MHC-2Kb, MHC-Db, MHC-2Dd, MHC-2Kd, MHC-2Kk, and MHC-2Ld were selected and showed in Table 8. RANKPEP online server was applied to predict binding peptides to class II MHC molecules as shown in Table 9.

**IgE epitopes and allergenic sites prediction**

Allergenic site prediction using AlgPred showed that the protein sequence does not contain experimentally proven IgE epitope.

**Discussion**

Atherosclerosis is classically known as a complex inflammatory disorder of arterial intima. Over the past decade, many studies have demonstrated that both protective and pathogenic immune responses are major factors responsible for development of atherosclerosis [5]. Given the significant roles

**Table 8.** The result of prediction for MHC-I epitopes

MHC type	Start and end position
MHC-Db	17-25, 73-81, 85-93, 121-129, 140-148, 150-158, 196-204, 226-234, 263-278, 355-374, 424-440
MHC-Dd	17-42, 52-60, 79-87, 104-112, 118-131, 134-142, 144-152, 162-170, 179-202, 206-214, 220-228, 246-259, 266-274, 278-291, 300-308, 314-334, 359-367, 382-390, 419-427, 430-438, 448-461, 478-486, 491-499, 501-516, 526-539, 560-568
MHC-Kb	26-34, 123-131, 140-148, 162-170, 178-187, 194-202, 221-229, 300-308, 354-362, 390-398, 400-408, 458-466, 567-575
MHC-Kd	17-27, 30-38, 85-93, 101-109, 140-154, 163-171, 179-192, 196-204, 263-273, 317-325, 355-377, 400-418, 424-438, 462-470, 475-483, 517-525, 551-559
MHC-Kk	17-34, 47-77, 79-93, 107-127, 146-165, 178-204, 206-229, 267-290, 294-309, 313-326, 335-343, 348-363, 371-379, 388-396, 400-410, 427-435, 451-469, 487-495, 517-534, 551-563, 570-578
MHC-Ld	11-19, 23-31, 77-85, 101-113, 152-160, 172-180, 199-207, 224-232, 238-246, 258-266, 297-305, 310-326, 340-348, 362-372, 384-401, 403-411, 417-425, 427-477, 499-507, 509-517, 533-542, 561-573

**Table 9.** Prediction of MHC-II epitopes

Rank	Position	Sequence	Score
1	31	YKNIGAKLV	18.129
2	109	IAQVATISA	15.593
3	305	FKHLIEIGL	15.277
4	359	NKHMISPVL	15.012
5	155	IEGMKFDRG	14.484
6	376	HVAIAGRTC	13.579
7	254	LKDMAIATG	12.327
8	178	CEFQDAYVL	12.032
9	410	YSCKPGYVS	11.36
10	184	YVLLSEKKI	10.601
11	518	YQGERVKIQ	10.224
12	158	MKFDRGYIS	9.578
13	463	YPNTISFSC	8.695

of immune system in development of this vascular disorder, many reports suggested an immune-modulatory strategy such as active immunization to modulate atherosclerosis [37]. Many modalities of vaccine have been used in preclinical studies which represent promising result in modulation of atherosclerosis via active immunization [38]. One of the most important steps in active immunization is recognition of appropriate antigens implicated in the disease. Recent studies have identified many autoantigens such as ox-LDL,  $\beta$ -2-GPI, and HSP60 which play important roles in development of atherosclerosis [12].

Considering that atherosclerosis is a multifactorial disease, it is important to note that blocking of an autoantigen neces-

sarily not to be able modulate atherosclerosis effectively. But this problem could be promoted by using multi-component vaccines which are more suitable in preventing and modulating of atherosclerosis in comparison with a mono-subunit vaccine.

In the present study, we selected HSP60, ApoB-100, and  $\beta$ -2-GPI proteins that play pivotal role in development of atherosclerosis. For designing of our multivalent vaccine, we selected the residues 61-361 in the N-terminal region of HSP60 reported to be effective in some immunization studies [4]. Also, the ApoB-100 peptide sequences including p2, p143, and p210 were selected. Recent studies indicated that active immunization using these ApoB-100 peptide sequences leads to 40%-70% reduction in development of atherosclerosis [13,14]. For the third fragment, the amino acid residues of domain I, II, and V of  $\beta$ -2-GPI protein were selected. This protein is a central antigenic target for the aPL in the pathophysiology of several autoimmune diseases specially atherosclerosis [20,21]. A logical explanation for this selection is that domains I and II are the best binding sites for the aPL. Also, the structural data suggest a simple membrane binding mechanism in which the positive charges on domain V interact with the negatively charged phospholipids [39,40]. Since these domains play key roles in pathogenesis of this disease, an immune-modulatory strategy for active immunization against domain I, II, and V could probably be useful for prevention and treatment of this disease by production of specific blocker antibodies against these domains.

The constructed fusion proteins need to the appropriate linkers to bind various parts of these proteins. For efficient separation of the different domains of our chimeric protein, a linker consisting of GGGGs repeats was applied to form a monomeric hydrophobic  $\alpha$ -helix and random coil. Our successful experience in using this linker in chimeric gene has shown that it could lead to logically acceptable results. Our selected linker could significantly control the distance and reduce the interference between the domains.

One of the most important steps in designing synthetic genes is codon optimization. We optimized our synthetic gene to enhance recombinant protein production in *E. coli*. CAI on the optimized gene sequence was 0.66, while wild type gene CAI was only 0.63. This chimeric gene also showed well immunogenic properties in VaxiJen server analysis.

Expression of proteins is significantly dependent on mRNA secondary structure which confer beneficial roles such as regulation of gene expression [41]. Mfold is a common online

server usually used to predict RNA secondary structure. The results from mRNA prediction indicated that the mRNA had enough stability for effective translation in the host. Our results showed that the first nucleotides at 5' of mRNAs did not have a hairpin or pseudoknot and the folding free energies of 5'-untranscribed regions were higher than other genomic regions. Therefore, based mentioned results we can expect to increase the translation rates, half-lives, and transcript number of chimeric protein in the host.

Physico-chemical parameters of our chimeric gene were analyzed by ProtParam software. The pI value ( $pI > 7.75$ ) showed basic nature of the protein. Extinction coefficient of chimeric protein at 280 nm was high. ExPASy ProtParam classifies the chimeric protein as stable based on the instability index (instability index, 35.59).

*Ab initio* I-TASSER software was used for predicting three-dimensional structure of the chimeric protein. Our results suggested that *ab initio* I-TASSER software could predict the folds and good resolution model for our multivalent protein. RMSD and TM score were used to evaluate the reliability and accuracy of predicted models. Expected TM score of  $0.55 \pm 0.15$  validates the accuracy of model. A TM-score more than 0.5 usually indicate the correct topology model and Z-score and C-score were used for showing its confidence. Analysis of chimeric structure based on Ramachandran plot showed a protein with full stability. In Ramachandran plot analysis, approximately an inconsiderable 10.9% of the residues were located in outlier region that could probably be due to the presence of chimeric junctions.

A crucial step in the development of multi-component vaccine based on recombinant protein is to identify B-cell and T-cell epitopes which induce strong B-cell and T-cell responses [42,43]. In this study, we identified continuous B-cell epitopes of the chimeric protein based on several different features such as antigenicity, accessibility, hydrophobicity, secondary structure, and flexibility analysis. The analysis of all continuous B-cell epitopes was performed completely to obtain good results, without any missing. The identified epitopes on protein surface could interact easily with antibodies, and they were generally flexible [44].

Another important step in computational studies is determination of discontinuous epitopes that are essential for antibody-antigen interaction. DiscoTope which seems to be more reliable than other software recognized 52 conformational epitopes for B cells.

CTLpred server was used for identifying potential T-cell



epitopes of chimeric proteins. The selected T-cell epitopes were antigenic with potential to interact with human leukocyte antigen alleles.

In summary, we have applied several bioinformatics methods for immunogenicity prediction of chimeric protein. Our results demonstrated that epitopes of our constructed protein, designed from HSP60, ApoB-100, and  $\beta$ -2-GPI could induce B cell and T cell-mediated immune responses successfully. Therefore, our multivalent vaccine is suggested as a potential candidate for modulation and reduction of atherosclerosis. As the next steps to complete our study on vaccine development, we recommend *in vitro* synthesis and *in vivo* experimental studies to evaluate the efficacy of this chimeric vaccine.

## ORCID

Ahmad Karkhah <http://orcid.org/0000-0002-9511-7839>

Jafar Amani <http://orcid.org/0000-0002-5155-4738>

## References

1. Tuttolomondo A, Di Raimondo D, Pecoraro R, Arnao V, Pinto A, Licata G. Atherosclerosis as an inflammatory disease. *Curr Pharm Des* 2012;18:4266-88.
2. Hansson GK. Inflammation, atherosclerosis, and coronary artery disease. *N Engl J Med* 2005;352:1685-95.
3. Mizuno Y, Jacob RF, Mason RP. Inflammation and the development of atherosclerosis. *J Atheroscler Thromb* 2011; 18:351-8.
4. Wick G, Jakic B, Buszko M, Wick MC, Grundtman C. The role of heat shock proteins in atherosclerosis. *Nat Rev Cardiol* 2014;11:516-29.
5. Hansson GK, Libby P. The immune response in atherosclerosis: a double-edged sword. *Nat Rev Immunol* 2006;6:508-19.
6. Businaro R, Tagliani A, Buttari B, et al. Cellular and molecular players in the atherosclerotic plaque progression. *Ann N Y Acad Sci* 2012;1262:134-41.
7. Tabas I, Williams KJ, Boren J. Subendothelial lipoprotein retention as the initiating process in atherosclerosis: update and therapeutic implications. *Circulation* 2007;116:1832-44.
8. Nilsson J, Wigren M, Shah PK. Vaccines against atherosclerosis. *Expert Rev Vaccines* 2013;12:311-21.
9. Pichardo-Almarza C, Metcalf L, Finkelstein A, Diaz-Zuccarini V. Using a systems pharmacology approach to study the effect of statins on the early stage of atherosclerosis in humans. *CPT Pharmacometrics Syst Pharmacol* 2015;4:e00007.
10. Chyu KY, Shah PK. Advances in immune-modulating therapies to treat atherosclerotic cardiovascular diseases. *Ther Adv Vaccines* 2014;2:56-66.
11. Gero S, Gergely J, Jakab L, et al. Inhibition of cholesterol atherosclerosis by immunisation with beta-lipoprotein. *Lancet* 1959;2:6-7.
12. Chyu KY, Shah PK. Can we vaccinate against atherosclerosis? *J Cardiovasc Pharmacol Ther* 2014;19:77-82.
13. Chyu KY, Zhao X, Reyes OS, et al. Immunization using an Apo B-100 related epitope reduces atherosclerosis and plaque inflammation in hypercholesterolemic apo E (-/-) mice. *Biochem Biophys Res Commun* 2005;338:1982-9.
14. Fredrikson GN, Soderberg I, Lindholm M, et al. Inhibition of atherosclerosis in apoE-null mice by immunization with apoB-100 peptide sequences. *Arterioscler Thromb Vasc Biol* 2003;23:879-84.
15. Nilsson J, Fredrikson GN, Bjorkbacka H, Chyu KY, Shah PK. Vaccines modulating lipoprotein autoimmunity as a possible future therapy for cardiovascular disease. *J Intern Med* 2009;266:221-31.
16. Kilic A, Mandal K. Heat shock proteins: pathogenic role in atherosclerosis and potential therapeutic implications. *Autoimmune Dis* 2012;2012:502813.
17. Long J, Lin J, Yang X, et al. Nasal immunization with different forms of heat shock protein-65 reduced high-cholesterol-diet-driven rabbit atherosclerosis. *Int Immunopharmacol* 2012;13:82-7.
18. Klingenberg R, Ketelhuth DF, Strodthoff D, Gregori S, Hansson GK. Subcutaneous immunization with heat shock protein-65 reduces atherosclerosis in Apoe(-)/(-) mice. *Immunobiology* 2012;217:540-7.
19. Lu X, Chen D, Endresz V, et al. Immunization with a combination of ApoB and HSP60 epitopes significantly reduces early atherosclerotic lesion in Apobtm2SgyLdlrtm-1Her/J mice. *Atherosclerosis* 2010;212:472-80.
20. Serrano A, Garcia E, Serrano M, et al. IgA antibodies against beta2 glycoprotein I in hemodialysis patients are an independent risk factor for mortality. *Kidney Int* 2012;81:1239-44.
21. De Groot PG, Meijers JC, Urbanus RT. Recent developments in our understanding of the antiphospholipid syndrome. *Int J Lab Hematol* 2012;34:223-31.
22. Banzato A, Pengo V. Clinical relevance of beta(2)-glyco-

- protein-I plasma levels in antiphospholipid syndrome (APS). *Curr Rheumatol Rep* 2014;16:424.
23. Grote A, Hiller K, Scheer M, et al. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res* 2005;33:W526-31.
  24. Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* 2007;8:4.
  25. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 2003;31:3406-15.
  26. Wilkins MR, Gasteiger E, Bairoch A, et al. Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol* 1999;112:531-52.
  27. Sen TZ, Jernigan RL, Garnier J, Kloczkowski A. GOR V server for protein secondary structure prediction. *Bioinformatics* 2005;21:2787-8.
  28. Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 2008;9:40.
  29. El-Manzalawy Y, Dobbs D, Honavar V. Predicting flexible length linear B-cell epitopes. *Comput Syst Bioinformatics Conf* 2008;7:121-32.
  30. 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:e1002829.
  31. Bhasin M, Raghava GP. Prediction of CTL epitopes using QM, SVM and ANN techniques. *Vaccine* 2004;22:3195-204.
  32. Singh H, Raghava GP. ProPred1: prediction of promiscuous MHC class-I binding sites. *Bioinformatics* 2003;19:1009-14.
  33. Reche PA, Glutting JP, Reinherz EL. Prediction of MHC class I binding peptides using profile motifs. *Hum Immunol* 2002;63:701-9.
  34. Saha S, Raghava GP. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res* 2006;34:W202-9.
  35. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* 1999;50:213-9.
  36. Brusic V, Rudy G, Harrison LC. MHCPEP, a database of MHC-binding peptides: update 1997. *Nucleic Acids Res* 1998;26:368-71.
  37. de Jager SC, Kuiper J. Vaccination strategies in atherosclerosis. *Thromb Haemost* 2011;106:796-803.
  38. Kimura T, Tse K, Sette A, Ley K. Vaccination to modulate atherosclerosis. *Autoimmunity* 2015;48:152-60.
  39. de Groot PG, Meijers JC. beta(2) -Glycoprotein I: evolution, structure and function. *J Thromb Haemost* 2011;9:1275-84.
  40. Willis R, Pierangeli SS. Anti-beta2-glycoprotein I antibodies. *Ann N Y Acad Sci* 2013;1285:44-58.
  41. Gaspar P, Moura G, Santos MA, Oliveira JL. mRNA secondary structure optimization using a correlated stem-loop prediction. *Nucleic Acids Res* 2013;41:e73.
  42. Jahangiri A, Rasooli I, Reza Rahbar M, Khalili S, Amani J, Ahmadi Zanoos K. Precise detection of *L. monocytogenes* hitting its highly conserved region possessing several specific antibody binding sites. *J Theor Biol* 2012;305:15-23.
  43. Rahbar MR, Rasooli I, Gargari SL, et al. A potential in silico antibody-antigen based diagnostic test for precise identification of *Acinetobacter baumannii*. *J Theor Biol* 2012;294:29-39.
  44. Delfani S, Imani Fooladi AA, Mobarez AM, Emaneini M, Amani J, Sedighian H. In silico analysis for identifying potential vaccine candidates against *Staphylococcus aureus*. *Clin Exp Vaccine Res* 2015;4:99-106.