Research Article **Bioinformatics**

International Journal of Pharma and Bio Sciences **ISSN 0975-6299**

3D STRUCTURAL MODELING OF NEUTRALIZING SCFV AGAINST GLYCOPROTEIN-D OF HSV-1 AND EVALUATION OF ANTIGEN-ANTIBODY INTERACTIONS BY BIOINFORMATIC METHODS

MOZAFAR MOHAMMADI[1,3] AND FOROOGH NEJATOLLAHI*[1,2,3]

[1]Recombinant antibody laboratory, Department of Immunology, Shiraz University of Medical Sciences, Shiraz, Iran
[2]Shiraz HIV/AIDS Research Center, Shiraz University of Medical Sciences, Shiraz, Iran
3School of Advanced medical Sciences and Technologies, Shiraz University of Medical Sciences, Shiraz, Iran

ABSTRACT

Structural studies on the interactions between antigens and antibodies are absolutely necessary and beneficial due to the understanding the principles that influence the interactions. In this work, we modeled 3D structure of specific neutralizing single chain variable fragment (scFv) recombinant antibody against glycoprotein-D (gD) of Herpes Simplex Virus-1 (HSV-1), using Modeller 9.11. To get the best alignment result, the ClustalW2 program was also used. The modeled molecule was evaluated by various servers such as ModEval and SAVES. The scFv binding sites were assessed by Pocket Finder and CASTp programs. The interactions between glycoprotein-D antigen and modeled antibody were thoroughly studied using ClusPro 2.0, a protein docking server. The docking results indicated that the neutralizing anti-gD scFv antibody has high affinity for its ligand due to different types of interactions including hydrogen bonds, hydrophobic forces, and salt bridges that taking place between the two molecules.

KEYWORDS: Ag-Ab interaction, GlycoproteinD of HSV, Modeling, Molecular docking, Neutralizing scFv, Pocket Finding.

FOROOGH NEJATOLLAHI
Recombinant antibody laboratory, Department of Immunology, Shiraz University of Medical Sciences, Shiraz, Iran Shiraz HIV/AIDS Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

INTRODUCTION

Antibodies (Abs) are distinct from each other in their structures and functions which serve as paradigms for ligand recognition [1]. Specific binding of Abs to different types of antigens (Ags) is due to various interactions between Ab and Ag which occurs between residues present at the Ab binding site and the antigen [2]. Also, shape complementarity plays an important role in antigen recognition and can be used as the main criterion for docking studies on antibody-antigen complexes [3]. Single chain variable fragments (scFv) recombinant antibodies are small antibodies that composed of variable light (V_L) and variable heavy (V_H) domains that are linked by a polypeptide linker [4]. These recombinant antibodies lack the Fc region, leading to low immunogenicity and making them better therapeutic agents compared to intact monoclonal Abs [5,6]. ScFvs have been found to be applicable agents in clinical use such as therapeutic gene delivery, immunotoxin based cancer therapy [7] and treatment of infectious diseases [8]. The high affinity property of scFvs that provides a tight interaction with the antigen, has made them effective targeting reagents. In this study, we modeled 3D structure of a specific neutralizing anti-gD scFv recombinant antibody selected and characterized in our previous study[9] and investigated its interactions with glycoprotein-D antigen of HSV-1 using bioinformatic methods.

METHODS

The 3D model structure for neutralizing anti-gD scFv was made by employing homology modeling from the available crystal structure of single chain antibodies using Modeller 9.11 [10,11] which is an automated homology modeling software. The 3D crystal structures of anti-HIV-1 gp120 single-chain variable fragment antibody variant (PDB code 3JUY) [12] and Mimicry-recognizing Native 2d10- scFv (PDB code 4H0G) [13] were directly retrieved from the protein data bank, as templates. To get the best alignment result the ClustalW2 program (http://www.ebi.ac.uk/Tools/msa/clustalw2/) was also used to align the target protein sequence against the template sequences. The simulated structure was evaluated by the ModEval (http://modbase.compbio.ucsf.edu/evaluation) and SAVES (http://nihserver.mbi.ucla.edu/SAVES) programs. In order to find the cavities and pockets on the antibody (i.e. the potential antibody-antigen binding sites), we used the Pocket Finder (http://www.modelling.leeds.ac.uk/pocketfinder) and Castp[14] (http://sts.bioengr.uic.edu/castp/about.php) programs. The Herpes simplex virus-1 glycoprotein D crystal structure (PDB code 1JMA) [15] was retrieved directly from protein data bank. Docking of HSV-1 gD antigen to scFv antibody was carried out using ClusPro 2.0 server[16-20] (http://cluspro.bu.edu/login.php), an automated protein–protein docking server. The best docking result with the lowest energy level was selected and further investigations on the model were performed by the Chimera [21], MolSurfer[22] and YASARA [23] softwares.

RESULTS AND DISCUSSION

(i) Sequence analysis of VH and VL regions

Table 1
Frameworks and CDRs of neutralizing anti-gD scFv molecule and comparing them to human V- gene family[9].

Region	VH1 Family	scFv-gD$_2$ VH
FR1	QMQLVQSGAEVKKTGSSVKVSCKASGYTFT	RGAAGESGAEVKKTGSSVKVSCKAS
CDR1	GYTFTYRYLH	GYTFTYRYLH
FR2	WVRQAPGQALEWMG	WVRQAPGQALEWMG
CDR2	WITPFNGNTNYAQKFQD	WITPFNGNTNYAQKFQD
FR3	RVTITRDRSMSTAYMELSSLRSEDTAMYYCAR	RVTITRDRSMSTAYMELSSLRSEDTAMYYCAR
CDR3	- - - - - - - - AFDI	GADTAMAGAFDI
JH3	WGQGTMVTVSS	WGQGTMVTVSS
Region	VL1 Family	scFv-gD$_2$ VL
FR1	QSVLTQPPSVSAAPGQKVTISC	QSVLTQPASVSGSPGQSITISC
CDR1	SGSSSNIGNNYVS	SGSSSNIGNNYVS
FR 2	WYQQLPGTAPKLLIY	WYQQLPGTAPKLLIY
CDR2	DNNKRPS	DNNKRPS
FR3	GIPDRFSGSKSGTSATLGITGLQTGDEADYYC	GIPDRFSGSKSGTSATLGITGLQTGDEADYYC
CDR3	GTWDSSLSAVV	GTWDSSLSAGV
JL3	FGGGTKLTVL	FGGGTKVTVL
	Linker	GGGGSGGGGSGGGGS

As a result of significant changes showed in table 1, considerable specificity of the antibody molecule is performed [24] and it shows a specific antibody selection against the peptide. The CDR3 region of VH in the antibody molecule plays an important role in the recognition of antigen [25]. Changes in the Frameworks are also effective in the formation of proper spatial structure of the whole antibody molecule and the creation of antigen-binding pocket cavity. Amino acids that are near the antibody binding site and outside CDRs are highly effective in the creating of antibody's tendency to antigen and specificity for its ligand [26]. The Residue of amino acids that either do not interact with the antigen or are outside the binding site, at the molecular level, can produce conformational differences in the whole body of the antibody molecule, inducing subtle changes in the binding activity [27].

(ii) Modeling and assessment of modeled structure

Modeller is a software that uses homology modeling and builds three-dimensional structure of proteins. It acts through a series of geometrical criterions for creating a probability density function for each atom position in the protein[11]. The method relies on an input sequence alignment between the target amino acid sequence to be modeled and a template protein whose structure has been solved. The best alignment result was obtained using ClustalW2 program (Fig.1).

Result of alignment

```
3JUY          -QVQLVQSGAEVKKPGASVKVSCQASGYRFSHFTVHWVRQAPGQRFEWMGWINPYNGNKE 59
4H0G          MEIQLQQSGPELVKPGASVKISCKASGYSFTDYIMLWVKQSHGKSLEWIGNINPYYGSTS 60
scFv-gD       -RGAAGESGAEVKKTGSSVKVSCKASGYTFTYRYLHWVRQAPGQALEWMGWITPFNGNTN 59
                .    :**.*: *.*:***:**:**** *:   : **:*: *: :**:* *.*: *...

3JUY          FSAKFQDRVTFTADTSANTAYMELRSLRSADTAVYYCARVGEWGWDDSPYDNYYMDVWGK 119
4H0G          YNLKFKGKATLTVDKSSSTAYMQLNSLTSEDSAVYYCARKNYYGSS--------LDYWGQ 112
scFv-gD       YAQKFQDRVTITRDRSMSTAYMELSSLRSEDTAMYYCARGADTAMAG------AFDIWGQ 113
              :  **:.:.*:* * * .****:* ** * *:*:*****     .          :* **:

3JUY          GTTVIVSS----------GGGGSEIVLTQAPGTLSLSPGERATFSCRSS----HSIRSRR 165
4H0G          GTTLTVSSAKTTG-----GGGGSDVVMTQTPFSLPVSLGDQASISCRSSQSLVHSNGNTY 167
scFv-gD       GTMVTVSSGGGGSGGGGSGGGGSQSVLTQ-PASVSGSPGQSITISCSGSS---SNIGNNY 169
              ** : ***          *****: *:** * ::. * *:  ::** .*     .  .

3JUY          VAWYQHKPGQAPRLVIHGVSNRASGISDRFSGSGSGTDFTLTITRVEPEDFALYYCQVYG 225
4H0G          LHWYLQKPGQSPKLLIYKVSNRFSGVPDRFSGSGSGTDFTLKISRVEAEDLGVYFCSQST 227
scFv-gD       VSWYQQLPGTAPKLLIYDNNKRPSGIPDRFSGSKSGTSATLGITGLQTGDEADYYCGTWD 229
              : ** : ** :*:*:*:  .:* **:.****** ***. ** *: ::. * . *:*

3JUY          ASSYT--FGQGTKLERK 240
4H0G          HVPYT--FGGGTKLEIK 242
scFv-gD       SSLSAGVFGGGTKLTVL 246
                  :  ** ****
```

Figure 1

The alignment result of ClustalW2 program. Stars and colons indicate exact similarity and partial similarity between amino acids of the three molecules, respectively and dashes indicate the gaps between amino acids.

Modeller, offered 10 models for our target protein, the best one among which, was chosen with the lowest DOPE (Discreted Optimized Protein Energy) score, -22286.881. This energy level indicates the stability and proper folding of the created model comparing to other derived models (Fig. 2) .Regions of the model that are built without a template, are much less accurate than the rest of the model [28]. The loop refinement was applied and except the linker region, the other loops were modified.

3D structure of scFv-gD

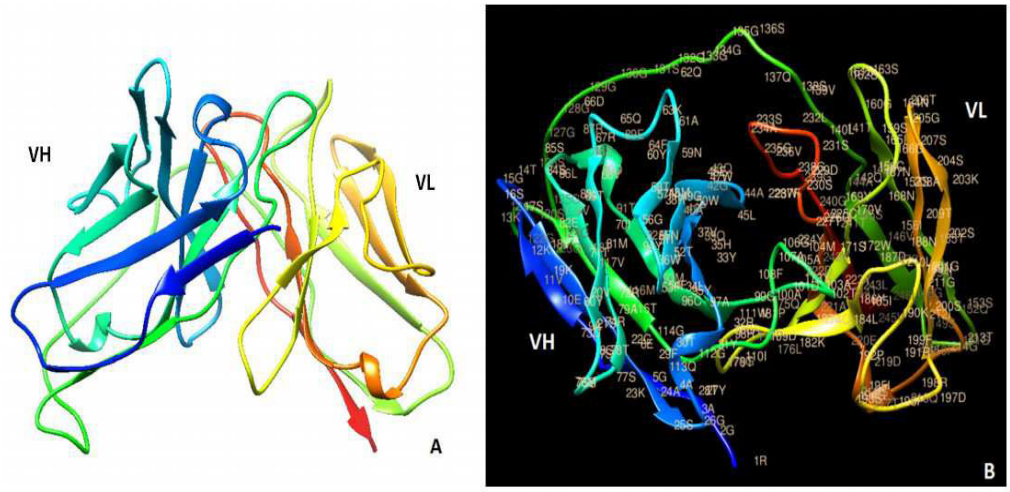


Figure 2

The 3D scenes obtained from molecular modeling of scFv-gD. A. depth-cued, rounded ribbon style created by Chimera shows domains and second structures of molecule by different colors. B. The top view of the same molecule with names and numbers of residues.

The obtained PDB format of model was evaluated through ModEval (Table 2) and SAVES (Table3) programs.

Table 2
Results obtained from ModEval.

	Model scores
Predicted RMSD	2.612
Predicted Native Overlap (3.5 Å)	0.889
GA341	1.000

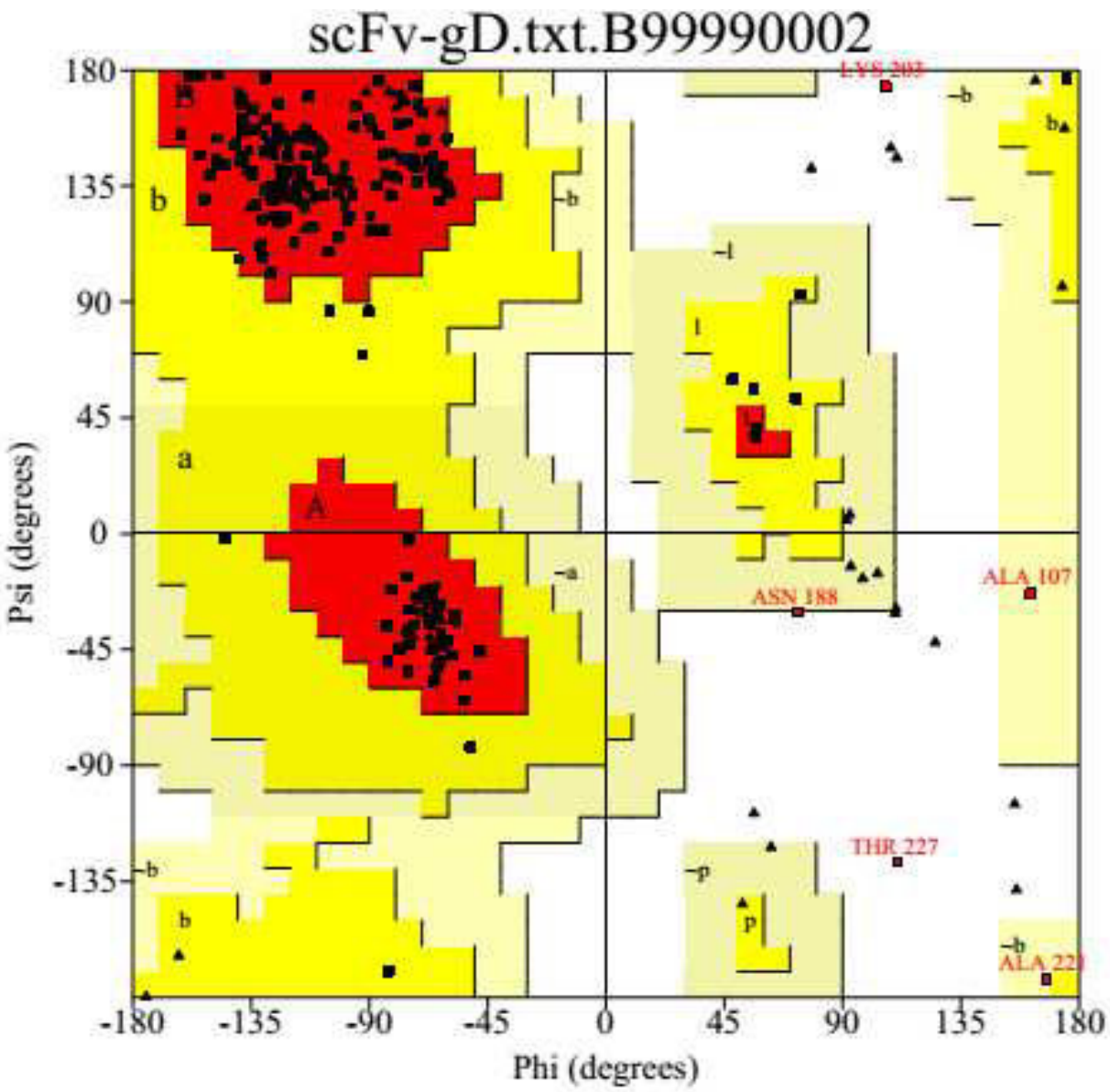
Due to the presence of three repeats of GGGGS sequence as a linker in our protein against one repeat in the homologous proteins (3JUY, 4H0G), the score of RMSD (Root-mean-square deviation) was 2.612 (Table 2), which in spite of loop refinement, no significant change was observed. There was 55% identity between anti-gD scFv sequence and the homologous protein sequences. A predicted model could be reliable when the GA341 score is higher than a cut-off 0.7(i.e. a probability of the correct fold larger than 95%), the method uses the percentage of sequence identity between the template and the model as a parameter. On the other hand, protein folding is reliable when at least 30% of its C alpha atoms superpose within 3.5 Å of their correct positions (Table 2) [29]. This was about 89% for our modeled scFv structure. According to the obtained scores by ModEval, it can be claimed that the predicted model of our protein is acceptable ≥ 95%[30,31]. To assess protein structures, the SAVES server uses different tools, such as the Verify_3D that determines the compatibility of an atomic model (3D) with its own amino acid sequence (1D) by assigning a structural class based on its location and environment (alpha, beta, loop, polar, nonpolar, etc) and comparing the results to good structures[32]. Another tool is the ProcCheck that Checks the stereochemical quality of a protein structure by analyzing residue-by-residue geometry and overall structure geometry[33,34]. According to the verify- 3D criterions, the results provided by SAVES show that, the modeled structure is 99.19 percent acceptable. Torsion angles are among the most important local structural parameters that control protein folding. The obtained data from ramachandran plot in the ProCheck, indicated that over 90% of amino acids are located in low energy areas, according to their φ and ψ torsion angles 91.8% residues were in most favored region and 5.6% residues were in additional allowed region. This indicated a proper 3D structure model for anti-gD scFv. However, due to the presence of 12 glycines in the linker that usually create loop, it caused slight errors in the ramachandran plot; 1% and 1.5% residues were in generousely allowed and disallowed regions, respectively (Fig. 3 & Table 3). The ERRAT2 program gave overall quality factor 69.912 which indicates that the model is acceptable.

Table 3
The SAVES server results.

Program	score
Verify_3D	**99.19% Pass**
ProCheck	**Ramachandran plot: 88.8% core, 8.2% allow, 1.5% gener, 1.5% disall**
	M/c bond lengths: 97.4% within limits
	M/c bond angles: 88.4% within limits
	Planar groups: 100.0% within limits
ERRAT2	**Overall quality factor: 69.912**

Core, this region shows the most favorable conformations of atoms and displays the sterically allowed conformations for beta strands. Allow, this is where rare left-handed alpha helices lie. Gener (Generously allowed), this is the place right-handed alpha helices lie. Disall (Disallowed regions), it has almost no outlined region and this conformation is disfavored due to steric clash. M/c bond lengths, main-chain bond lengths in the structure. M/c bond angles, main-chain bond angles in the structure. Planar groups, aromatic rings (Phe, Tyr, Trp, His) and planar end-groups (Arg, Asn, Asp, Gln, Glu).

Ramachandran plot



Plot statistics

Residues in most favoured regions [A,B,L]	180	91.8%
Residues in additional allowed regions [a,b,l,p]	11	5.6%
Residues in generously allowed regions [~a,~b,~l,~p]	2	1.0%
Residues in disallowed regions	3	1.5%
Number of non-glycine and non-proline residues	196	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	40	
Number of proline residues	8	
Total number of residues	246	

Figure 3

According to φ and ψ torsion angles, 91.8% residues are in most favored region and 5.6% residues are in additional allowed region. 1% and 1.5% residues are in generously allowed and disallowed regions, respectively.

CDRs of scFv-gD

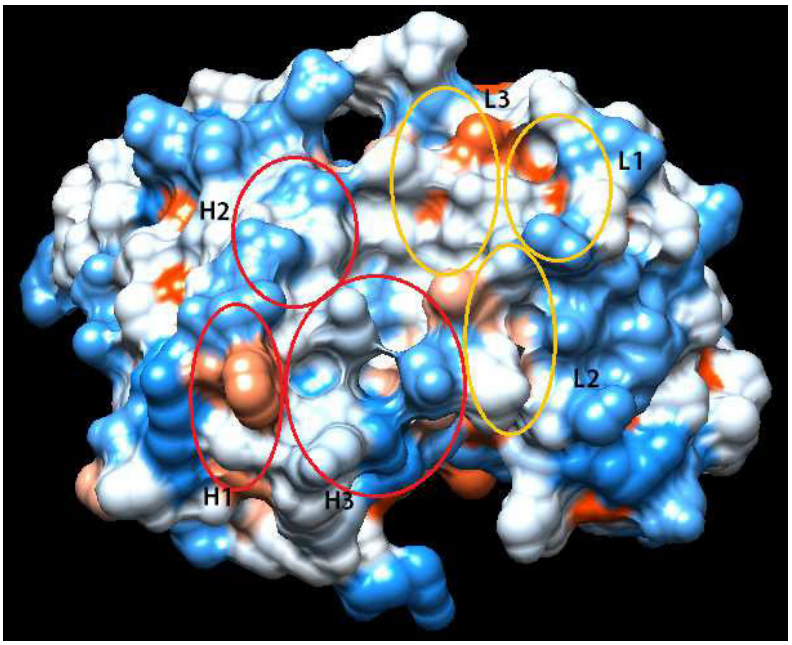


Figure 4
Top view of the scFv-gD (hydrophobic surfaces) provided by Chimera software showing the VH and VL regions.

(iii) Pockets and active binding sites finding

The function of a protein molecule takes place through interaction with other molecules, such as DNA, substrate, ligand, and other domains of proteins. The 3-D conformation of a protein provides the necessary shape and physicochemical texture to facilitate its interactions with the other molecules. Structural information about the surface areas of a protein molecule can help in the study of relationship between function and structure of the protein. Features of protein surface areas are helpful for determining the binding specificity [35]. Each of the programs, Castp and Pocket Finder considering their defined algorithms, introduced several areas as ligand-binding pocket or cavity. Through comparison the areas using these two programs, it was identified that the amino acids Y 27, Y 31, R 32, Y 33, L 34 of VH and G 99, A 100, D 101, T 102, A 103, M 104, A 105, G 106, A 107, F 108, D 109, I 110 of VL, are common among areas. Among these, amino acids Y 27, R 32, Y 33, L 34 and G 99, A 100, D 101, I 110 were able to create about 260 Cubic Angstroms ($Å^3$) volume, as main antigen binding domain (Fig.5). The hydrophobic and aromatic amino acids seem to have a tendency for being in the CDRs and antigen recognition activity. It seems that the aromatic side-chains (Tyr) are solvent-exposed more than water-soluble proteins and provide strong participation in the interaction with the ligand [36].

Pocket binding sites of scFv-gD

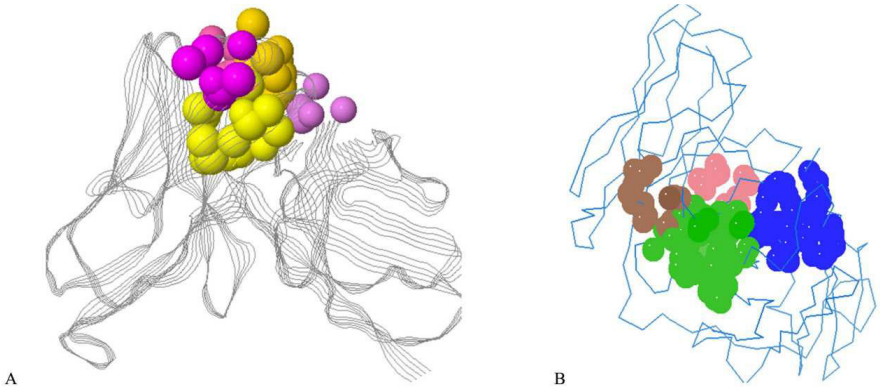


Figure 5
Active binding sites (colored balls) in anti gD-scFv provided by A. Castp (front view) and B. Pocket Finder (top view) servers

(iv) Antibody – Antigen Docking

We used ClusPro 2.0 for the study of Antibody – Antigen interaction which is the first fully automated web-based molecular docking program. The server uses a supercomputer for ligand docking at Boston university. The ClusPro 2.0 server which is developed by the groups of Dima Kozakov and Sandor Vajda, was the best in the server category in the latest rounds of CAPRI experiment [37]. Billions of putative complexes assessed by the molecular docking algorithms, retaining a preset number with favorable surface complementarities. A filtering method is then applied to this set of complexes and those with good electrostatic and desolvation free energies are selected for further clustering [16-20]. We selected the best results based on the lowest energy calculated ($E=0.50E_{rep}+-0.20E_{att}+600E_{elec}+0.25E_{DARS}$ [19]); which was -250.1 Kcal/mol (Fig.6). This energy lowering in the antibody–antigen complex was mainly due to stabilizing by hydrogen bonding, salt bridges and hydrophobic interactions.

2D & 3D viewpoint of scFv-gD and HSV-1gD interaction

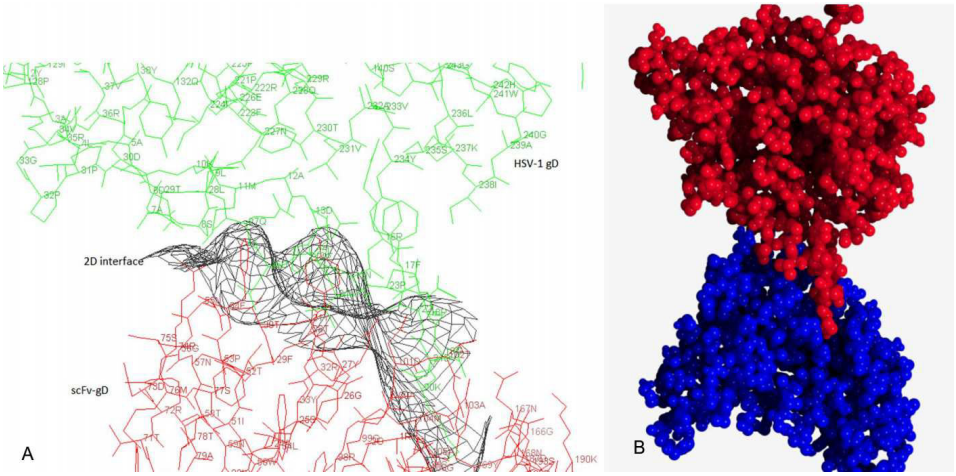


Figure 6
A. 2D viewpoint of interaction between anti gD- scFv and HSV-1 gD obtained from MolSurfer software. B. The ball style created by YASARA software from the best docking result, HSV-1 gD (red) and anti gD-scFv (blue).

Table 4
Interaction of antibody amino acids with gD epitope peptide at interatomic distance less than 5 Å.

Inter molecular interactions	scFv - gD		HSV-1 Glycoprotein D		
	Amino acid	Position	Amino acid	Position	Distances (Å)
Hydrogen bonds	Asp OD2	109	Lys HZ	20	1.75
	Asp OD1	109	Lys HZ	20	1.79
	Arg HH1	98	Gly O	19	1.85
	Thr OG1	28	Lue O	25	1.94
	Tyr HH	31	Lue O	25	1.83
	Tyr OH	31	Gln HE2	27	3.58
	Thr H	30	Asp OD2	26	2.16
	Arg HH2	74	Asp OD2	26	1.76
	Arg HH2	74	Gln O	27	3.15
	Arg HH1	74	Thr OG1	29	1.77
	Arg HH2	32	Asn OD1	15	2.81
	Arg HH2	32	Phe O	17	1.74
	Asp OD1	101	Asn OD1	15	3.47
	Asp OD1	101	Arg HH2	18	3.49
Hydrophobic interaction	Phe	54	Met	11	
	Phe	54	Pro	14	
	Tyr	31	Pro	14	
	Tyr	31	Phe	17	
	Tyr	31	Tyr	234	
	Tyr	31	Lue	25	
pi-pi interaction	Tyr	31	Phe	17	5.0
Salt bridge	Asp OD1	109	Lys NZ	20	2.61
	Tyr OH	31	Gln NE2	27	2.94
	Arg NH2	74	Asp OD1	26	2.73
	Asp OD1	101	Arg NH2	18	3.05

OD is a standard abbreviation in the PDB database for "Oxygen Delta ", OG for "Oxygen Gamma", "NH" for a hydrogen on the N of an Arg residue, NE for "nitrogen of ε carbon", HE for "hydrogen of ε carbon" and Z is used for "ammonium (nitrogen; NZ and hydrogen; HZ)" on Lys.

Not only these bonds but also van der Waals' interactions, and pi-pi interactions Lead to bond strength in the interface area between two protein molecules (Table 4). Multiple bonding between the antigen and the antibody ensures that the antigen will be tightly bound to the antibody. All bonds were between amino acids located in the CDR regions and amino acids that formed the intended epitope on the antigen. Although two arginines, 98 and 74, are not in the CDR regions, they participated in the binding to the antigen due to close arisen space provided by other connections. This indicates that the effect of other amino acids outside the CDR regions is needed for proper folding and spatial structure of antibody binding to antigen and the amino acids of framework regions have effective roles in this regard. Among the amino acids of the antigen that interact with paratope, the amino acids Pro 14, Asn 15, Phe 17, Arg 18, Gly 19, and Lys 20 were in the epitope region, but amino acids Met 11, Lue 25, Asp 26, Gln 27, Thr 29, and Tyr 234 were not in this region. Nonetheless, due to the spatial structure created around the epitope and approaching of non-epitope amino acids to the interface region between the two macromolecules, they were able to participate in the interactions and caused the decline in the binding free energy and greater stability of binding activity between scFv and HSV-1 gD molecules. The aromatic rings of the two adjacent amino acids, tyrosine 31 and phenylalanine 17, led to pi-pi interaction between them. This class of interaction involves direct attraction between arene rings. It is clear that the rings of these two amino acids form the parallel displaced configuration, that is, by it self, one of the most stable conformations in pi-pi interaction (Fig.7). The hydrophobic interactions (Fig.7) that normally play effective roles in the antibody – antigen bindings, are detected between Tyr 31 of

antibody and Pro 14, Phe 17, Tyr 234 and Lue 25 of antigen, and also were between Phe 54 of antibody and Met 11 and Pro 14 of antigen[38]. The salt bridges were formed between charged amino acids, Asp 109, Tyr 31, Arg 74, Asp 101 on scFv and amino acids Lys 20, Gln 27, Asp 26, Arg 18 on antigen, respectively. The groups of polar and charged residues that are located on protein-protein interfaces enhance the stability of the complex [39]. It seems that the hydrogen bonds along with the hydrophobic forces and salt bridges have the main role in the interactions between the gD and scFv-gD recombinant antibody (Fig.7). Lysine 20 and phenylalanine 17 of epitope, tyrosine 31, arginine 74 and aspartic acid 101 & 109 of the antibody play an important role in connecting the two molecules. As shown in the table 4, these amino acids are able to form several interactions alone. It is noteworthy that tyrosine 31 (as an aromatic partially hydrophobic amino acid) of antibody is the only amino acid which has been found in all interactions. This is described by its large size (hydrophobic effect), large polarization (van der Waals interactions), ability to form hydrogen bonds, and rigidity (less loss of conformational entropy upon complexation). Thus, the amount of aromatic rings would give 'stickiness' to the CDRs and give diversity to antibody specificity. Antigen specificity would obtain from the interacting surfaces, complementarity shapes, formed by the proper positioning of the aromatic rings and the polar or charged groups located in correct position [36]. In addition, lysine 20 of antigen, because of its spatial structure and location on the epitope, has been able to enter into the pocket conveniently created on the antibody molecule (antigen binding pocket), and it stables the binding of these two molecules not only chemically but also physically. This form has been created as a result of docking with the least expenditure of energy. It also shows the high-affinity of scFv-gD to the antigen.

Different interactions between scFv-gD and HSV-1 gD

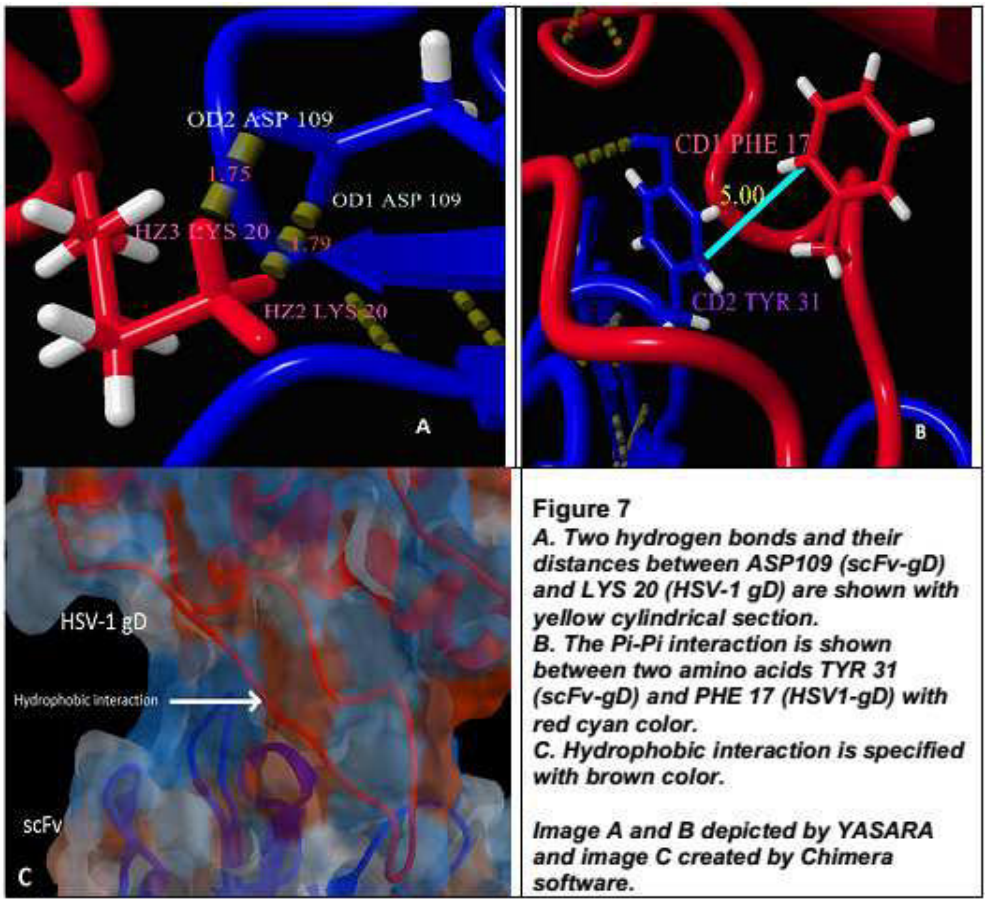


Figure 7
A. Two hydrogen bonds and their distances between ASP109 (scFv-gD) and LYS 20 (HSV-1 gD) are shown with yellow cylindrical section.
B. The Pi-Pi interaction is shown between two amino acids TYR 31 (scFv-gD) and PHE 17 (HSV1-gD) with red cyan color.
C. Hydrophobic interaction is specified with brown color.

Image A and B depicted by YASARA and image C created by Chimera software.

CONCLUSION

Anti gD - scFv recombinant antibody was well modeled with respect to the assessments made on its structure. Moreover, antibody–antigen docking result revealed the high affinity of anti gD- scFv antibody for HSV-1 glycoprotein-D antigen. Hydrogen bonds, hydrophobic forces, and salt bridges were found to have important roles in the interaction between the antibody and gD. These data confirmed the 76% neutralization effect of anti-gD scFv antibody obtained in our previous study [9].

ACKNOWLEDGEMENT

The authors acknowledge Shiraz University of Medical Sciences for financial support, Grant no 91-6153. Some molecular graphics and analyses were performed with the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).

Conflict of Interest

Conflict of interest declared none.

REFERENCES

1. Yuriev, E., Ramsland, P. A. & Edmundson, A. B. Docking of combinatorial peptide libraries into a broadly cross-reactive human IgM. *Journal of molecular recognition : JMR* 14, 172-184, doi:10.1002/jmr.533 (2001).
2. Valjakka, J. *et al.* Crystal structure of an in vitro affinity- and specificity-matured anti-testosterone Fab in complex with testosterone. Improved affinity results from small structural changes within the variable domains. *The Journal of biological chemistry* 277, 44021-44027, doi:10.1074/jbc.M208392200 (2002).
3. Feng, J., Li, Y., Zhang, W. & Shen, B. Rational design of potent mimic peptide derived from monoclonal antibody: antibody mimic design. *Immunology letters* 98, 311-316, doi:10.1016/j.imlet.2004.12.006 (2005).
4. Deckert, P. M. Current constructs and targets in clinical development for antibody-based cancer therapy. *Current drug targets* 10, 158-175 (2009).
5. Monnier, P. P., Vigouroux, R. J. & Tassew, a. N. G. In Vivo Applications of Single Chain Fv (Variable Domain) (scFv) Fragments. *Antibodies* 2, 193-208 (2013).
6. Nejatollahi, F., Jaberipour, M. & Asgharpour, M. Triple blockade of HER2 by a cocktail of anti-HER2 scFv antibodies induces high antiproliferative effects in breast cancer cells. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*, doi:10.1007/s13277-014-1854-8 (2014).
7. Chowdhury, P. S., Viner, J. L., Beers, R. & Pastan, I. Isolation of a high-affinity stable single-chain Fv specific for mesothelin from DNA-immunizedmice by phage display and construction of a recombinant immunotoxin with anti-tumor activity. *Proc. National. Acad. Sci. USA* 95, 669–674 (1998).
8. Xu, L. M. *et al.* scFv antibodies against infectious bursal disease virus isolated from a combinatorial antibody library by flow cytometry. *Biotechnology letters* 36, 1029-1035, doi:10.1007/s10529-014-1463-z (2014).
9. Nejatollahi, F., Bagheri, V., Motamedifar, M. & Moazen, B. Neutralizing human single-chain antibodies against Herpes Simplex Virus type 1 glycoprotein D from a phage display library. *International Journal of Microbiology and Immunology Research* 2(3), 033-040 (2014).
10. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *Journal of molecular biology* 234, 779-815, doi:10.1006/jmbi.1993.1626 (1993).
11. Fiser, A. & Sali, A. Modeller: generation and refinement of homology-based protein structure models. *Methods in enzymology* 374, 461-491, doi:10.1016/S0076-6879(03)74020-8 (2003).
12. Clark, K. R. & Walsh, S. T. Crystal structure of a 3B3 variant--a broadly

neutralizing HIV-1 scFv antibody. *Protein science : a publication of the Protein Society* 18, 2429-2441, doi:10.1002/pro.255 (2009).
13. Tapryal, S., Gaur, V., Kaur, K. J. & Salunke, D. M. Structural evaluation of a mimicry-recognizing paratope: plasticity in antigen-antibody interactions manifests in molecular mimicry. *Journal of immunology* 191, 456-463, doi:10.4049/jimmunol.1203260 (2013).
14. Dundas, J. *et al.* CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic acids research* 34, W116-118, doi:10.1093/nar/gkl282 (2006).
15. Carfi, A. *et al.* Herpes simplex virus glycoprotein D bound to the human receptor HveA. *Molecular cell* 8, 169-179 (2001).
16. Brenke, R. *et al.* Application of asymmetric statistical potentials to antibody-protein docking. *Bioinformatics* 28, 2608-2614, doi:10.1093/bioinformatics/bts493 (2012).
17. Comeau, S. R., Gatchell, D. W., Vajda, S. & Camacho, C. J. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics* 20, 45-50 (2004).
18. Comeau, S. R., Gatchell, D. W., Vajda, S. & Camacho, C. J. ClusPro: a fully automated algorithm for protein-protein docking. *Nucleic acids research* 32, W96-99, doi:10.1093/nar/gkh354 (2004).
19. Kozakov, D., Brenke, R., Comeau, S. R. & Vajda, S. PIPER: an FFT-based protein docking program with pairwise potentials. *Proteins* 65, 392-406, doi:10.1002/prot.21117 (2006).
20. Kozakov, D. *et al.* How good is automated protein docking? *Proteins* 81, 2159-2166, doi:10.1002/prot.24403 (2013).
21. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* 25, 1605-1612, doi:10.1002/jcc.20084 (2004).
22. Gabdoulline, R. R., Wade, R. C. & Walther, D. MolSurfer: A macromolecular interface navigator. *Nucleic acids research* 31, 3349-3351 (2003).
23. Krieger, E., Koraimann, G. & Vriend, G. Increasing the precision of comparative models with YASARA NOVA--a self-parameterizing force field. *Proteins* 47, 393-402 (2002).
24. Kabat, E. A. & Wu, T. T. Identical V region amino acid sequences and segments of sequences in antibodies of different specificities. Relative contributions of VH and VL genes, minigenes, and complementarity-determining regions to binding of antibody-combining sites. *Journal of immunology* 147, 1709-1719 (1991).
25. Xu, J. L. & Davis, M. M. Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities. *Immunity* 13, 37-45 (2000).
26. Chothia, C. *et al.* Conformations of immunoglobulin hypervariable regions. *Nature* 342, 877-883, doi:10.1038/342877a0 (1989).
27. Greenspan, N. S. Dimensions of antigen recognition and levels of immunological specificity. *Advances in cancer research* 80, 147-187 (2001).
28. Chung, S. Y. & Subbiah, S. A structural explanation for the twilight zone of protein sequence homology. *Structure* 4, 1123-1127 (1996).
29. Melo, F., Sanchez, R. & Sali, A. Statistical potentials for fold assessment. *Protein science : a publication of the Protein Society* 11, 430-448, doi:10.1002/pro.110430 (2002).
30. Shen, M. Y. & Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein science : a publication of the Protein Society* 15, 2507-2524, doi:10.1110/ps.062416606 (2006).
31. Eramian, D., Eswar, N., Shen, M. Y. & Sali, A. How well can the accuracy of comparative protein structure models be predicted? *Protein science : a publication of the Protein Society* 17, 1881-1893, doi:10.1110/ps.036061.108 (2008).
32. Bowie, J. U., Luthy, R. & Eisenberg, D. A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 253, 164-170 (1991).
33. Laskowski, R. A., Moss, D. S. & Thornton, J. M. Main-chain bond lengths and bond angles in protein structures. *Journal of*

molecular biology 231, 1049-1067, doi:10.1006/jmbi.1993.1351 (1993).
34. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of biomolecular NMR* 8, 477-486 (1996).
35. Binkowski, T. A., Naghibzadeh, S. & Liang, J. CASTp: Computed Atlas of Surface Topography of proteins. *Nucleic acids research* 31, 3352-3355 (2003).
36. Kumagai, I. & Tsumoto, K. Antigen–Antibody Binding. *Nature Publishing Group* (2001).
37. Lensink, M. F. & Wodak, S. J. Docking, scoring, and affinity prediction in CAPRI. *Proteins: Structure, Function, and Bioinformatics* 81, 2089-2095 (2013).
38. Tina, K. G., Bhadra, R. & Srinivasan, N. PIC: Protein Interactions Calculator. *Nucleic acids research* 35, W473-476, doi:10.1093/nar/gkm423 (2007).
39. Sheinerman, F. B., Norel, R. & Honig, B. Electrostatic aspects of protein-protein interactions. *Current opinion in structural biology* 10, 153-159 (2000).