

Designing a new bi-specific tandem single-chain variable fragment antibody against tumor necrosis factor- α and interleukin-23 via in silico studies for the treatment of Rheumatoid arthritis

Asgar Barkhordaria

Isfahan University of Medical Sciences

Karim Mahnam (✉ mahnam.karim@sku.ac.ir)

Shahrekord University <https://orcid.org/0000-0002-9906-4957>

Hamid Mir Mohammad Sadeghi

Isfahan University of Medical Sciences

Research

Keywords: Molecular dynamics simulation, Bispecific antibody, Rheumatoid arthritis, Tumor necrosis factor- α , Interleukin-23, Docking

Posted Date: February 11th, 2020

DOI: <https://doi.org/10.21203/rs.2.23111/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Rheumatoid arthritis disease is a chronic autoimmune inflammatory disease that mainly causes synovial joint inflammation and cartilage destruction. Tumor necrosis factor- α (TNF- α) is a pivotal cytokine that plays an important role in rheumatoid arthritis. The treatments focusing on a single cytokine' inhibition, are able to clinically produce meaningful responses in only about half of the treated patients due to multiple cytokines involved in this disease. In the present study, a bi-specific tandem single-chain variable fragment was designed in order to suppress both human tumor necrosis factor- α and interleukin-23 (IL23) as a potential therapeutic drug candidate for this disease. To do so, at first, eight bi-specific tandem single-chain variable fragment models were built against tumor necrosis factor- α and interleukin-23 cytokines with different domain orders and then 50 ns molecular dynamics simulation was performed for each one and thereafter structural properties were exploited. Results: MD simulation results indicate that the domains' order strongly affects tandem single-chain variable fragment properties and in overall, the fragment VLIL23+Linker+VHIL23+linker+VLTNF+Linker+VHTNF +His6 (VL is variable light and VH is variable heavy fragments and His6 is six histidine) not only separated antibody domains but also had better stability and solvation energy. Conclusions: Hence, this structure can be considered as a potential drug for rheumatoid arthritis. It is hoped that this research could shed a light for the treatment of Rheumatoid arthritis disease.

Introduction

Rheumatoid arthritis (RA) is a chronic autoimmune disease of the joints that is characterized by synovial inflammation as well as cartilage and bone destruction[1]. This disease is associated with undeniable negative consequences for the individual such as losing his/her function and as a result work disability, which poses significant socio-economic challenges to the society[2]. Many clinical studies have validated the pivotal contribution of cytokines in the pathogenesis of RA [3]. Tumor necrosis factor (TNF- α) antagonist is assumed as one of the most efficient agents available for rheumatoid arthritis[4]. Randomized clinical trials and extension studies have revealed that all anti-TNF drugs are highly effective in both early and established RA. Large-scale observational and registry studies have evaluated their use and long-term efficacy in clinical preparation. It is proved that all anti-TNF drugs have not only a validated and rapid efficacy, but could also inhibit the progression of structural joint damage as due to their rapid and prolonged inflammation' suppression [5]. However, according to the results of clinical trials, almost 50% of RA patients treated with TNF- α inhibitors do not achieve efficient response. Moreover, some RA patients treated with cytokine inhibitors lose their sensitivity over the time and need to switch to other agents[6]. Therefore, there are still a high medical requirement and also alternative strategies to fight against inflammatory and autoimmune illnesses and to further improve patients' lives.

Applying a combined inhibitor for different cytokine targets may overcome this challenge via permitting more effective inflammation' suppression. Given the inflammatory process complexity and the heterogeneity among the patients with RA, it is very likely that the best therapeutic efficacy is obtained through using more than one single mechanism [1].

IL-23 and IL-17 plays important roles in both joint inflammation and the association of destructed bone, indicating that their inhibition would ameliorate both processes [7]. IL-23 is considered as a vital checkpoint in the generation, maintenance, and also activation determinant of pathogenic Th17 cells. This is largely attributable to the fact that only pre inflammatory Th17 cells express the IL-23 receptor (IL-23R) [8]. It could be noted that IL23 stabilizes IL17 expression by Th17 cells and seems to give 'license' to Th17 cells in order to become pathogenic. By inducing the expression of its own receptor, IL23 also induces an autocrine positive-feedback loop [3]. Proliferation and stabilization of Th17 cells also require IL-23[9]. Therefore, therapeutic strategies that specifically block Th17 cell development (i.e., IL-23) can be used as a putative treatment method for RA. IL-23 up regulates receptor activator' expression of nuclear factor- κ B ligand (RANKL) on fibroblast-like synoviocytes, myeloid precursors and on CD4 cells and eventually on driving osteoclast formation. The chief observation fueling this idea is that IL-23p19 $^{-/-}$ mice are fully resistant to the development of collagen-induced arthritis (an animal model of RA)[10]. In male DBA1 mice, collagen-induced arthritis is associated with increased expression levels of IL-23 and IL-23R by the synovial membrane. The intra-articular injection of an antibody to IL-23R significantly improved the arthritis and decreased the synovial expression of IL-23, IL-23R, IL-17, IL-1, IL-6, and TNF- α [11]. The intra-articular injection of IL-23 into the mice resulted in the local migration of neutrophils by increasing IL-17 production and decreasing IL-12 and IFN γ production. In vivo neutralization of IL-23 was uncovered to be the cause of diminishing IL-17 induction; so, it can be concluded that Th17 cells cannot be produced in IL-23 deficient mice [9]. These data confirm the fact that targeting IL-23p19 through a vaccination strategy is protective in collagen-induced arthritis. This specific targeting of IL-23 might constitute a promising therapeutic approach to investigate rheumatoid arthritis[12]. The serum level of IL-23 was significantly elevated in the patients with RA compared to the healthy controls. In addition, there was a significant positive relationship between the serum level of IL-23 in RA patients and individual disease activity' parameters[13–15].

Abovementioned data, suggest that IL-23 may potentially play a role in RA pathogenesis and therefore targeting the IL-23, may provide a new promising therapeutic approach in the treatment of RA. It seems that the combination of IL-23 with TNF- α inhibitor, may represent a very promising novel approach for the treatment of rheumatoid arthritis.

Bi-specific antibodies (BsAbs) are the second antibodies' generation used for therapeutic purposes in immunotherapy [16]. At least 30 different BsAb and related molecules are currently applied in clinical development and mainly oncology, autoimmune or chronic inflammatory indications[17]. The most successful clinical application of BsAbs up to now has been its usage in oncology by redirecting cytotoxic T cells to kill tumor cells. Simultaneous blockade of several different disease mediators may lead to a greater therapeutic efficacy and/or benefit for the patients than targeting individual disease' mediators[18].

Small recombinant bi-specific antibodies, devoid of all constant immunoglobulin parts, are assembled from the variable light (VL) and heavy chain (VH) domains of two antibodies. ScFv fragments of different specificity can be easily included into a bi-specific molecule by joining genetically the two scFv moieties with a more or less flexible peptide linker (VL + Linker + VH). In these tandem scFv (taFv) molecules, each antigen-binding site forms a separate folding unit. Consequently, tandem scFv molecules could be considered in different arrangements of scFv and variable domains (VL or VH)[19]. The most commonly used flexible linkers have the sequences which primarily consist of Gly and Ser residues ("GS" linker). An example of the most widely used flexible

linker is the sequence of (GGGGS)_n. The length of this GS linker could be optimized via adjusting the copy number “n” in order to achieve appropriate separation of the functional domains, or to maintain necessary inter-domain interactions. The incorporation of Ser or Thr would maintain the linker stability in aqueous solutions by hydrogen bonds' formation with water molecules and therefore reduces unfavorable interaction between the linker and protein moieties. The selection of (Gly)₈ linker was due to the reports accentuated that flexible linkers can increase the accessibility of an epitope to antibodies or improve protein folding. In sum, flexible linkers are generally rich in small or polar amino acids such as Gly and Ser and are able to provide good flexibility and solubility. In addition, flexible linkers can serve as passive linkers to keep a distance between the functional domains. The length of flexible linkers might be adjusted to permit proper folding or to achieve optimal biological activity of the fusion proteins [20]. Thus, tandem scFv molecules were constructed possessing short Ala-Ala-Ala or Gly-Gly-Gly-Gly-Ser linkers, or longer linkers, which can adopt secondary structures [19]. In this study, the flanking linkers (FL) within the scFv units and middle linkers (ML) connecting the two scFv (ATNF and AIL23 domains) molecules were (GGGGS)₃.

The association between an antibody and an antigen involves myriad of non-covalent interactions between the epitope (binding site on the Ag), and the paratopes (binding site on the Ab). Six hyper variable loops within the variable domains of Abs, commonly termed complementarity determining regions (CDRs), are widely assumed to be responsible for Ag recognition; while, the constant domains are believed to mediate effector activation. A set of CDRs constitutes a paratope; three CDRs are found in the heavy chain (H1, H2, H3) and three are revealed in the light chain (L1, L2, L3) [21].

There is a considerable interest in developing more effective treatments of the diseases by simultaneously targeting multiple components of the disease pathways.

There are currently five FDA-approved TNF inhibitors using for the treatment of RA, including etanercept, adalimumab, certolizumab, infliximab, and golimumab [22]. However, almost 50% of RA patients treated with TNF- α inhibitors do not show any notable positive response; therefore, it is tempting for the researchers to speculate a fusion protein for TNF- α that participate in the pro-inflammatory cytokine cascade [28] and the treatment of RA. Previous studies also reported that the bispecific fusion protein has curative effectiveness for RA [23, 24] or other inflammatory diseases[25–27].

Presently, the use of computational techniques in chemistry and biology, from the quantum mechanics of molecules to the dynamics of large complex molecular aggregates, is becoming quite popular. Molecular interactions steer chemical reactions, phase transitions and other physical phenomena and can be studied via molecular dynamics (MD) simulation that show detailed motion of molecules or atoms as a function of time. The properties of a model system, minima geometries of proteins and DNA, and binding drugs' free energy can be studied via MD simulation[28].

In this study, we aim at designing a new bispecific taFv against TNF- α and IL-23 using in silico simulation as a new treatment for Rheumatoid arthritis. At first, eight bispecific taFv were designed against TNF- α and IL-23 by molecular modeling and the impact of domains' order on structure and dynamics of taFv was investigated.

In the present study, we designed a new bispecific tandem scFv antibody, which is capable of simultaneously recognizing TNF- α and IL-23. Eight arrangements of AIL23 and ATNF with one linker were considered for this protein. These structures were examined via theoretical study because it was somewhat a less expensive and faster method. It should be noted that inappropriate design of taFv leads to unsuitable antibody' generation [29].

In the past two decades, numerous different types of bispecific taFv have been scrutinized and designed including: VH-VL—VH-VL [30, 31], VH-VL—VL-VH [32], VL-VH—VL-VH [33–35] and VL-VH—VH-VL [36, 37]. A bispecific taFv is composed of four variable domains; thus, there are eight different domain orders for building a bispecific taFv. Although, experimental methods are more precise, they obviously require extra time and cost. Consequently, it could be said that this research is the first attempt to design a bispecific taFv using in silico methods.

Methods

Molecular modeling

Figure 1 shows schematic view of designed bispecific taFv constructs. Tandem scFv molecules (taFv) are generated by connecting two scFv molecules (ATNF- α or AIL23) with an additional middle linker (ML) sequence. One scFv was consisted of variable light (VL) and variable heavy (VH) segments of IL23 or TNF- α domain (which are joined together by a flexible peptide linker). The sequence of VH-TNF was: “EVQLVESGGGLVQPGRSLRLSCAASGFTFDYAMHWVRQAPGKLEWVSAITWNSGHIDYADSVGRFTISRDNKNSLYLQMNSLRAEDTAVYYCAKVSYLSTASSLDYV”. Also, the sequence of VL-TNF was: “DIQMTQSPSSLSASVGRVITITCRASQGIIRNYLAWYQQKPGKAPKLLIYAASLTQSGVPSRFSGSGSGTDFTLTISLQPEDVATYYCQRYNRAPYTFGGQGTKVEIK”. The sequence of VH-IL23 was: “EVQLVQSGAEVKKPGESLKISCKGSGYSFNSYWIGWVRQMPGKLEWV GIIDPS NSYTRYSPSFQGVITISADKISITAYLQWSSLKASDTAMYYCARWYKPFQVWVGQGLTVTVSS”.

Furthermore, the sequence of VL-IL23 was: “QSVLTQPPSVSGAPGQRVTISCTGSSSNIGSGY DVHWYQQLPGTAPKLLIYGNSKRPSGVPDRFSGSKSGTSASLAITGLQSEADYYCASWTDGLSL VVFGGKTKLTVL”. Thus taFv are composed of four variable domains (two VH and two VL). All linkers have (GGGGS)₃ sequence. Eight domain orders were considered for bispecific taFv.

<Figure 1 near here >

The Fab (antigen-binding) fragment of one antibody is divided into the Fv (fragment variable) (Fv = VH + VL) and other fragments. Fv is the smallest fragment that binds to antigen via the connection of both heavy and light chains. CDR is on Fv fragment and the Fv of the Fab fragment of anti-TNF- α antibody has PDB code 4nyl and anti-IL-23 has PDB code 4 m6n. These structures were used as the model building' templates. Homology models of the models were

constructed using Modeller 9.16 software[38]. For each taFv structure, 1000 models were generated and sorted by their molecular probability density function (molpdf) values [39, 40]. Ten models with lower molpdf energy value for each taFv structure were selected out of 1000 generated models and used for further analysis.

The best model was selected based on stereochemistry among these ten models using Procheck software[41]. The amount of allowed regions' residues of the Ramachandran plots was higher than 91% in all taFv models, indicating that they are reliable homology modeled structures. For contraction, the Ramachandran plots were not exposed. In addition, the quality of final models were assessed using ProSA-web (Protein Structure Analysis)[42] and it was confirmed that all of them were appropriate energies (data not shown). Thereafter, best selected taFv models were used as starting structure for MD simulations. To do so, Pymol software was used to visualize the models [43].

Molecular Dynamics Simulation

MD simulations were carried out using the Gromacs 5 [44] under G43A1 force field. Each of these models was put in a center of cubic box filled with SPC216 water model and 10 Å far from the box edges. Anionic strength of 0.14 M (Na^+ and Cl^- ions) was also considered to simulate the physiological conditions. To avoid edge effects and better describe the condition of full hydration, periodic boundary conditions were applied for each hydrated protein system. Subsequently, the systems were energy minimized using steepest descent algorithm followed by conjugate gradients algorithm with a tolerance of 60 kJ/mol/nm. After energy minimization, the systems were equilibrated at the constant volume (NVT) for 1000 ps and constant pressure (NPT) for 2000 ps at 100 K in which, initial configuration of the structures were kept fixed. Finally, a 50 ns production of MD simulation at 300 K with a time step of 0.002 ps was performed for each system using isothermal coupling ensemble (NPT). Electrostatic interactions were computed using particle mesh Ewald (PME) method and the cut off electrostatic and van der Waals' interactions were 10 Å. All covalent bonds involving hydrogen atoms were constrained by LINCS algorithm during the production phase. Then, in total 400 ns MD of the simulation for all models were executed and analyzed via Gromacs analysis tools. Afterward, 10 snapshots were taken with 500 ps intervals during the last 5 ns of the simulations of each taFv structure. These structures were used to generate PQR files by the PDB2PQR web server (http://nbc-222.ucsd.edu/pdb2pqr_2.0.0/). Eventually, APBS software (Adaptive Poisson-Boltzmann Solver software) was applied to calculate solvation energies [45] and the average of 10 solvation free energies was calculated for each model. The dielectric constant was set to four for the protein and 78 for the solvent.

Results And Discussion

At first, the structures were designed with modeler software and then molecular dynamics simulation was implemented for each structure. Thereafter, several factors were analyzed including: RMSD, Rg, hydrogen bond numbers, distances between domains, SASA, free energies of solvation and electrostatic and van der Waals interactions, during the last 40 ns of the simulation. These parameters demonstrate protein stability and ability to bind to antigen as well as aberrant folding and aggregation of fusion proteins[46]. The amount of residues belonging to the most favored/additional allowed regions of the Ramachandran plots was higher than 98.4% in all taFv models, which indicates the reliability of homology modeled' structures. The ProSA website showed that Z-scores (overall model quality) of all models were negative and within the range of -9.37 to -10.02 which underlines its acceptable quality.

Rmsd, Potential Energy, Temperature:

To obtain the equilibrium time of each simulated protein model during MD simulation, the backbones root mean square deviation (RMSD) were calculated (Fig. 2). RMSD plot is usually used to evaluate the time required for a system to reach structural equilibrium and to estimate the duration of running a simulation. As it is illustrated in Fig. 2, all models represented stable dynamics and reached a plateau during the last 40 ns and then they were all analyzed and performed during the last 40 ns.

< Fig. 2 near here >

Table 1 shows the results of average potential energy, temperature, root mean square deviation (RMSD), radius of gyration (Rg), ratio of the total energy drift to the average of total energy during the last 40 ns of the simulation. The system temperature reached the plateau at 300 K after 10 ns (the temperature plot shows that all molecular systems became stabilized and remained stable throughout 50 ns of the molecular dynamics simulations) and fluctuated around 1 K during the last 40 ns of the simulation. Next, all systems reached thermal equilibration. The potential energy of all systems was measured to be negative and stable during the last 40 ns (throughout 50 ns) of the simulation. Also, the total energy drift was trivial in all systems; so that the simulation times seemed to be adequate. The results of backbone RMSD, potential and the temperature showed that the equilibrium and simulation times were sufficient in all systems.

Table 1

The average of the potential energy, temperature, root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), ratio of the total energy drift to average of total energy during the last 40 ns of simulation.

	taFv1	taFv2	taFv3	taFv4	taFv5	taFv6	taFv7	taFv8
RMSD (nm)	2.21 ± 0.14	1.23 ± 0.12	0.77 ± 0.04	1.04 ± 0.11	0.86 ± 0.07	0.89 ± 0.13	2.17 ± 0.18	1.58 ± 0.10
RG (nm)	2.82 ± 0.13	2.82 ± 0.08	2.38 ± 0.03	3.04 ± 0.15	2.77 ± 0.04	2.67 ± 0.10	2.90 ± 0.04	3.01 ± 0.03
Potential energy (Kj/mol)	-1079918 ± 1123	-1097259 ± 1090	-779838 ± 957	-1008838 ± 1050	-1495120 ± 1330	-1036047 ± 1089	-1810751 ± 1392	-1326445 ± 1232
Temperature (K)	300.02 ± 1.21	299.97 ± 1.19	299.99 ± 1.49	299.99 ± 1.28	299.99 ± 0.99	300.03 ± 1.24	300.02 ± 0.94	300.02 ± 1.06
Ratio	0.000117	0.000047	0.000442	0.000023	0.000164	0.000255	0.000122	0.000067

< Table 1 near here >

Radius Of Gyration:

The radius of gyration (Rg), which is a property linked to the molecule volume and compactness, was found to fluctuate between 2.38 nm and 3.04 nm. The gyration radius of eight taFvs models were presented in Fig. 3. The Rg of all models reached the stable state after 5 ns; while, the Rg of taFv4 and 8 were greater than other models.

< Fig. 3 near here >

In protein engineering, the functional domains' separation to reduce interference, is considered as a very important factor. As it was discovered by Arai et al [47], a higher Rg is associated with increased distance between the domains of a fusion protein. In our study, Rg was consistent with the distance between ATNF scFv and AIL23 scFv domains. These distances in taFv8 and taFv4 models were more than other models (Table 4); therefore, in this structure these domains have the least interference with each other. The absence of changes in Rg, confirms the fact that the protein properly folds; while, the changes of Rg value imply that protein structure unfolds over the time. The Rg value fluctuation of taFv4 is more than taFv8 (Table 1). Hence, it can be conclude that taFv8, remains more stable during the simulation time.

Table 4

The average of free energy of solvation ($\times 10000$ kj/mol) of all models, electrostatic and Van der Waals interactions energy (kj/mol) between protein and water and the average of short range electrostatic and Van der Waals interactions energy (kj/mol) between VL and VH segment of ATNF or AIL23 domain and distance between center of mass of two scFv (VL + VH) during the last 40 ns MD simulation

	taFv1	taFv2	taFv3	taFv4	taFv5	taFv6	taFv7	taFv8
ΔG_{sol}	-1.96	-2.25	-1.80	-1.91	-1.89	-1.88	-2.23	-2.38
Coul + LJ: Protein- water	-26164	-25693	-24757	-26499	-25739	-25918	-25871	-26174
Coul + LJ: VL with VH in ATNF domain	-551.7	-632.01	-764.36	-611.36	-616.72	-561.33	-691.91	-676.04
Coul + LJ: VH with VL in AIL23 domain	-524.9	-646.64	-633.84	-601.07	-635.12	-569.74	-593.05	-608.19
Distance AIL23 scFv domain with ATNF scFv domain	4.42 ± 0.32	4.51 ± 0.19	3.27 ± 0.07	5.11 ± 0.36	4.29 ± 0.09	3.96 ± 0.27	4.65 ± 0.10	4.93 ± 0.08

Solvent Accessible Surface Area (sasa):

Solvent accessibility surface area (SASA) of all models and CDR residues (in VH and VL parts) were computed during the 50 ns of MD simulation (Fig. 4) and the average of the last 40 ns MD of the simulation were exhibited in Table 2. Although many researchers take advantage of the SASA to monitor folding process, in current research, SASA was utilized to determine whether the selected model has sufficient area for binding to IL-23 and TNF antigen. The solvent accessible area plays important role in interaction between antigen and antibody. The higher surface area could mediate high affinity of the antigen-antibody complex. As it is displayed in this table, taFv1, taFv4, taFv7 and taFv8 have similar and higher total surface area than other models. The highest SASA value of CDR residues of ATNF scFv, was observed in taFv2 (27.59 nm²). Also, CDR residues of AIL-23 scFv were observed in taFv4 (25.19 nm²). Moreover, the ASAS of taFv8 in both ATNF and AIL23 was remarkable (about 23 nm²). The total SASA value of taFv8, taFv7, taFv4 and taFv1 were similar and greater than other constructs. The SASA of CDR in these models was more than other models (Table 2). So, it is concluded that these models have more ability for binding to the antigens.

Table 2

The average of total accessible surface area of all models and CDR residues during the last 40 ns of MD simulation.

Area (nm ²)	taFv1	taFv2	taFv3	taFv4	taFv5	taFv6	taFv7	taFv8
Total area	211.88 ± 0.39	206.59 ± 0.39	192.20 ± 0.38	212.85 ± 0.39	204.07 ± 0.38	207.01 ± 0.38	210.67 ± 0.38	210.31 ± 0.38
CDR surface in ATNF	26.85 ± 0.44	27.59 ± 0.47	26.65 ± 0.45	26.46 ± 0.43	25.87 ± 0.44	21.55 ± 0.37	23.97 ± 0.42	23.82 ± 0.41
CDR surface in AIL23	22.77 ± 0.38	20.73 ± 0.37	14.74 ± 0.32	25.19 ± 0.39	24.00 ± 0.38	24.23 ± 0.41	23.93 ± 0.39	23.76 ± 0.39

<Figure 4 near here >

< Table 2 near here >

Protein solubility, and consequently aggregation, is a critical issue in several areas of biochemical and biopharmaceutical research such as the production and maintenance of protein pharmaceuticals or industrial enzymes[48]. In the present study, interactions between the protein and water was analyzed by a number of H-bond, electrostatic and van der Waals' interactions. The energies between protein and water and free energies of solvation were also investigated.

In the present study, the interaction between the VH and VL domain in each scFv was analyzed by number of H-bonds, center of mass (COM) distance, and van der Waals plus electrostatic energies between domains. The pattern of hydrogen bonding was evaluated with a donor-acceptor distance cut-off of 0.35 nm and donor-hydrogen-acceptor angle cut-off of 30°. The average of hydrogen bonds within the models and between the protein and water are articulated in Table 3.

Table 3

The average of number of hydrogen bonds between VH and VL in AIL23 and ATNF domain and average of number of protein-protein hydrogen bonds and protein-water during the last 40 ns of MD simulation.

	taFv1	taFv2	taFv3	taFv4	taFv5	taFv6	taFv7	taFv8
Average of number of H- bonds within proteins	332.2 ± 10.1	334.0 ± 10.7	342.2 ± 10.6	333.1 ± 10.2	341.7 ± 11.3	334.3 ± 10.3	333.9 ± 10.4	329.2 ± 10.5
Average of number of H bonds between protein and water	824.6 ± 18.6	811.8 ± 18.8	785.1 ± 19.8	825.3 ± 22.2	807.7 ± 20.6	815.9 ± 20.7	818.4 ± 19.5	824.1 ± 18.3
The number of hydrogen bonds between VH and VL in AIL-23 domain	4.26 ± 1.61	5.40 ± 2.02	6.98 ± 1.60	6.55 ± 1.79	7.25 ± 1.57	5.36 ± 1.79	5.96 ± 1.64	4.94 ± 2.04
The number of hydrogen bonds between VH with VL in ATNF domain	5.09 ± 1.97	7.74 ± 1.95	11.33 ± 2.47	6.16 ± 1.81	6.85 ± 1.73	6.09 ± 1.62	8.21 ± 1.81	7.36 ± 1.89

Hydrogen Bond:

The hydrogen bonds are responsible for a major factor of stable conformation' maintenance of the protein [49]. H-bond analysis was performed to examine the H-bond for eight models during the 50 ns of the simulation period. The average number of H-bonds within the proteins (intramolecular) is given in Table 3. The number of intramolecular H-bonds was almost similar in all eight models. The taFv3 and 5 showed a slightly larger number of H-bonds formed during the last 40 ns of the simulation than other models.

The solubility of the models was evaluated by the measurement of intermolecular H-bond developed between taFv and water during simulation. The greatest numbers of intramolecular hydrogen bonds within the proteins were in taFv5 and taFv3 models. Also, the highest numbers of intermolecular hydrogen bonds between the protein and solvent were in taFv 4, 1 and 8 (825, 824 and 824, respectively).

The changes in the VH/VL association can modify relative position of the CDR loops which, in turn, can alter general shape of the antigen-binding site, as well as the arrangement of the side-chains that interact with the antigen [39]. Also, insufficient VH-VL interface stability of scFvs has often been suggested as the main cause of their aggregation and non-functionality[50]. In particular, specific H-bonds are supposed to be crucial for the conformation and stabilization of the CDRs and VH/VL interface' packing [39]. The average number of hydrogen bonds between VH and VL domains of scFvs were calculated and compared with each other. The results uncovered that in ATNF scFv domains the maximum number of hydrogen bonds was in the model 3 (11.3); while, in AIL23 scFv domain the greatest number of hydrogen bonds between VH and VL was in the model 5 (7.2) (Table 3). All models introduced an average of 4.2–11.3 hydrogen bonds throughout the last 40 ns of the simulation period. In addition, the numbers of hydrogen bonds were considerable in taFv8 (about 5 and 7.3 in AIL23 and ATNF respectively).

< Table 3 near here >

Free Energy Of Salvation:

The average of solvation free energy in all models is illustrated in Table 4. The least solvation free energy belongs to the model 8 and is -23800 kJ/mol. This energy is compatible with the number of hydrogen bonds between the models and waters (Table 4).

< Table 4 near here >

Electrostatic And Van Der Waals Interactions' Energy:

Attractive interactions between the protein and water are a dominant factor in ensuring high solubility. In addition, the short range van der Waals (LJ-SR) interaction, and short range coulomb (Coul-SR) interaction (electrostatic) energies were calculated between the protein and water. The lowest electrostatic + Van der Waals energies between them were obtained for taFv4, taFv8 and taFv1 (Table 4). These results are consistent with the number of hydrogen bonds between the protein and water. In total, considering the number of hydrogen bonds, the free energy of solvation, and van der Waals and electrostatic energies, the taFv8 exhibits the best results in terms of the interactions between the protein and water.

Van der Waals plus coulombic energy between the protein and water in taFv4 and taFv8 and taFv1 were more negative than other models and it is consistent with the number of hydrogen bonds between the protein and water in these models and solvation free energy in the model 8. Subsequently, the models four and eight and one and particularly model eight have more solubility than others.

Also the average of electrostatic plus van der Waals interactions' energies between VL and VH domain of ATNF and AIL23 are exposed in Table 4. The strong electrostatic plus van der Waals' interaction between them showed the protein stability and was in taFv3, taFv7 for ATNF domain and also in taFv2, taFv3 and taFv5 in AIL23 domain. A tight relationship was detected among electrostatic plus van der Waals' interaction energies and the number of hydrogen bonds between VH and VL segments.

Distance:

Spatial separation of the domains in hetero-functional chimeric proteins is important for independent work of domains [47]. Figure 5 and Table 4 represent the distance between the center of mass in AIL23 scFv and ATNF scFv domain or during 50 ns MD of the simulation. Figure 5 shows that among eight models, the maximum center of mass distance between anti-TNF- α scFv and anti-IL23 scFv, was belong to taFv4 and taFv8. Thus, in these models scFvs fragment can be separated from each other better than in other models. These results are consistent with their gyration radius. The distance between the center of mass VH and VL parts were calculated in both ATNF and AIL23 domains and were between 2.08 to 2.21 nm in the last 40 ns of the simulation. Then FL linker ((GGGS)₃) can separate VH and VL parts.

Although the number of hydrogen bonds between VH and VL and also the sum of van der Waals and electrostatic energies between VH and VL in ATNF and AIL23 is not better than other models, it seems that these parameters are not necessary for the stability, solubility and binding biphasic antibody to the antigens. Consequently, according to all calculated parameters taFv8 is selected as the best model.

It is worthy to say that the properties of the taFv8 can be improved by the application of middle linker modifications between the two scFv molecules.

<Figure 5 near here >

Conclusion

We designed a new biphasic (bispecific) antibody by using MD of the simulation to neutralize TNF- α and IL-23 cytokines. In silico analysis showed that taFv8 (VL_{IL23}+FL + VH_{IL23}+ML + VL_{TNF}+FL + VH_{TNF}+His6) has better structural properties comparing other taFvs. The reason is that almost all its structural properties such as Rg and solvation energy, hydrogen bonds, the distance between the domains and accessible surface area were better than other models. Overall, the methodology adopted in this study is fruitful in the drug discovery process and would improve drug efficacy and safety profiles in RA treatment. It should be noted that this is the first drug candidate with the combination of anti-IL23 and anti-TNF- α via molecular dynamics simulation which was used for the treatment of RA and it is hoped that this drug can be effective in ongoing experimental studies in our Lab.

Abbreviations

scFv:Single-chain variable fragment (VL + Linker + VH); taFv:tandem scFv (scFv + ML + scFv); ML:middle linker connecting the two scFv; RA:Rheumatoid arthritis; TNF- α :Tumor necrosis factor- α ; ATNF:anti-Tumor necrosis factor- α ; IL-23:Interleukin-23; AIL23:antiinterleukin-23; MDS:Molecular dynamics simulation; RMSD:root mean square deviation; Rg:radius of gyration; SASA:Solvent accessible surface area; CDR:complementarity determining region; BsAbs:Bispecific antibodies

Declarations

Ethics approval and consent to participate:

All authors have consent for publishing this manuscript in Theoretical Biology and Medical Modelling journal.

Reporting studies involving human participants, human or animal data and tissue:

Not applicable, this work is only a bioinformatics study.

Consent for publication of individual person's data:

Not applicable, this work was done only via authors data.

Availability of data and material:

All data and materials of this manuscript are available. Data sharing is applicable if somebody request.

Funding:

Design of this study and collection and analysis were performed by funding the Isfahan University, research department of medical sciences, Isfahan, Iran and interpretation of data and writing the manuscript were done by funding of Shahrekord university of Iran.

Authors' contributions:

All authors contributed in this manuscript. A. Barkhordari was contributed in collection and analysis data and K. Mahnam was contributed in interpretation of data and writing of the manuscript and H. Mirmohammad-Sadeghi was contributed in design of this work.

Acknowledgments

Not applicable

The content of this paper is extracted from the Ph.D thesis of A. Barkhordari. We thank the Isfahan University, Research Department of Medical Sciences, Isfahan, Iran and University of Shahrekord of Iran for their financial support of this project.

Competing Interests:

The authors declare that they have no competing interests.

References

- [1] Fischer JA, Hueber AJ, Wilson S, Galm M, Baum W, Kitson C, et al. Combined inhibition of tumor necrosis factor alpha and interleukin-17 as a therapeutic opportunity in rheumatoid arthritis: development and characterization of a novel bispecific antibody. *Arthritis & rheumatology* (Hoboken, NJ). 2015 Jan;67(1):51-62.
- [2] Silacci M, Lembke W, Woods R, Attinger-Toller I, Baenziger-Tobler N, Batey S, et al. Discovery and characterization of COVA322, a clinical-stage bispecific TNF/IL-17A inhibitor for the treatment of inflammatory diseases. *mAbs*. 2015 Jan 2;8(1):141-9.
- [3] McInnes IB, Buckley CD, Isaacs JD. Cytokines in rheumatoid arthritis - shaping the immunological landscape. *Nat Rev Rheumatol*. 2015 Jan;12(1):63-8.
- [4] Liu M, Xie M, Jiang S, Liu G, Li L, Liu D, et al. A novel bispecific antibody targeting tumor necrosis factor alpha and ED-B fibronectin effectively inhibits the progression of established collagen-induced arthritis. *Journal of biotechnology*. 2014 Sep 30;186:1-12.
- [5] Rubbert-Roth A, Atzeni F, Masala IF, Caporali R, Montecucco C, Sarzi-Puttini P. TNF inhibitors in rheumatoid arthritis and spondyloarthritis: Are they the same? *Autoimmunity reviews*. Jan;17(1):24-8.
- [6] Liu Z, Song L, Wang Y, Xu P, Guo X, Yang J, et al. A novel fusion protein attenuates collagen-induced arthritis by targeting interleukin 17A and tumor necrosis factor alpha. *International journal of pharmaceutics*. 2018 Aug 25;547(1-2):72-82.
- [7] Sherlock JP, Taylor PC, Buckley CD. The biology of IL-23 and IL-17 and their therapeutic targeting in rheumatic diseases. *LWW* 2015:71-5.
- [8] Yang J, Sundrud MS, Skepner J, Yamagata T. Targeting Th17 cells in autoimmune diseases. *Trends in pharmacological sciences*. 2014 Oct;35(10):493-500.
- [9] Mateen S, Zafar A, Moin S, Khan AQ, Zubair S. Understanding the role of cytokines in the pathogenesis of rheumatoid arthritis. *Clinica chimica acta; international journal of clinical chemistry*. 2016 Apr 1;455:161-71.
- [10] Croxford AL, Kulig P, Becher B. IL-12 and IL-23 in health and disease. *Cytokine & growth factor reviews*. 2014 Aug;25(4):415-21.
- [11] Wendling D. Interleukin 23: a key cytokine in chronic inflammatory disease. *Joint Bone Spine*. 2008 Oct;75(5):517-9.
- [12] Ratsimandresy RA, Duvallet E, Assier E, Semerano L, Delavallee L, Bessis N, et al. Active immunization against IL-23p19 improves experimental arthritis. *Vaccine*. 2011 Nov 21;29(50):9329-36.
- [13] Guo YY, Wang NZ, Zhao S, Hou LX, Xu YB, Zhang N. Increased interleukin-23 is associated with increased disease activity in patients with rheumatoid arthritis. 2013:850-4.
- [14] Dalila AS, Mohd Said MS, Shaharir SS, Asrul AW, Low SF, Shamsul AS, et al. Interleukin-23 and its correlation with disease activity, joint damage, and functional disability in rheumatoid arthritis. *The Kaohsiung journal of medical sciences*. 2014 Jul;30(7):337-42.

- [15] Al Fadl EMA, Fattouh M, Allam AA. High IL-23 level is a marker of disease activity in rheumatoid arthritis. 2013;85-92.
- [16] Thakur A, Huang M, Lum LG. Bispecific antibody based therapeutics: Strengths and challenges. *Blood reviews*. 2018 Jul;32(4):339-47.
- [17] Spiess C, Zhai Q, Carter PJ. Alternative molecular formats and therapeutic applications for bispecific antibodies. *Molecular immunology*. 2015 Oct;67(2 Pt A):95-106.
- [18] Chan AC, Carter PJ. Therapeutic antibodies for autoimmunity and inflammation. *Nat Rev Immunol*. 2010 May;10(5):301-16.
- [19] Kontermann RE. *Bispecific Antibodies*: Springer 2011.
- [20] Chen X, Zaro JL, Shen WC. Fusion protein linkers: property, design and functionality. *Advanced drug delivery reviews*. 2013 Oct;65(10):1357-69.
- [21] Sela-Culang I, Kunik V, Ofra Y. The structural basis of antibody-antigen recognition. *Frontiers in immunology*. 2013 Oct 8;4:302.
- [22] Brzustewicz E, Bryl E. The role of cytokines in the pathogenesis of rheumatoid arthritis—Practical and potential application of cytokines as biomarkers and targets of personalized therapy. *Cytokine*. 2015 Dec;76(2):527-36.
- [23] Li Q, Ren G, Xu L, Wang Q, Qi J, Wang W, et al. Therapeutic efficacy of three bispecific antibodies on collagen-induced arthritis mouse model. *International immunopharmacology*. 2014 Jul;21(1):119-27.
- [24] Qi J, Kan F, Ye X, Guo M, Zhang Y, Ren G, et al. A bispecific antibody against IL-1beta and IL-17A is beneficial for experimental rheumatoid arthritis. *International immunopharmacology*. 2012 Dec;14(4):770-8.
- [25] Robert R, Juglair L, Lim EX, Ang C, Wang CJH, Ebert G, et al. A fully humanized IgG-like bispecific antibody for effective dual targeting of CXCR3 and CCR6. *PloS one*.12(9):e0184278.
- [26] Yang Y, Zhang T, Cao H, Yu D, Zhang T, Zhao S, et al. The pharmacological efficacy of the anti-IL17 scFv and sTNFR1 bispecific fusion protein in inflammation mouse stimulated by LPS. *Biomedicine & pharmacotherapy = Biomedicine & pharmacotherapie*. 2017 Aug;92:905-12.
- [27] Jung K, Lee D, Lim HS, Lee SI, Kim YJ, Lee GM, et al. Double anti-angiogenic and anti-inflammatory protein Valpha targeting VEGF-A and TNF-alpha in retinopathy and psoriasis. *The Journal of biological chemistry*. 2010 Apr 22;286(16):14410-8.
- [28] Soleimani M, Mahnam K, Mirmohammad-Sadeghi H, Sadeghi-Aliabadi H, Jahanian-Najafabadi A. Theoretical design of a new chimeric protein for the treatment of breast cancer. *Research in pharmaceutical sciences*. May-Jun;11(3):187-99.
- [29] Stamova S, Cartellieri M, Feldmann A, Arndt C, Koristka S, Bartsch H, et al. Unexpected recombinations in single chain bispecific anti-CD3-anti-CD33 antibodies can be avoided by a novel linker module. *Molecular immunology*. 2011 Dec;49(3):474-82.
- [30] Choi BD, Kuan CT, Cai M, Archer GE, Mitchell DA, Gedeon PC, et al. Systemic administration of a bispecific antibody targeting EGFRvIII successfully treats intracerebral glioma. *Proceedings of the National Academy of Sciences of the United States of America*. 2013 Jan 2;110(1):270-5.
- [31] Dan Z, Tan Z, Xia H, Wu G. Construction and expression of D-dimer and GPIIb/IIIa single-chain bispecific antibody. *Experimental and therapeutic medicine*. 2013 Aug;6(2):552-6.
- [32] Arndt C, Feldmann A, Koristka S, Cartellieri M, Dimmel M, Ehninger A, et al. Simultaneous targeting of prostate stem cell antigen and prostate-specific membrane antigen improves the killing of prostate cancer cells using a novel modular T cell-retargeting system. *The Prostate*. 2014 Sep;74(13):1335-46.
- [33] Taki S, Kamada H, Inoue M, Nagano K, Mukai Y, Higashisaka K, et al. A Novel Bispecific Antibody against Human CD3 and Ephrin Receptor A10 for Breast Cancer Therapy. *PloS one*. 2015;10(12):e0144712.
- [34] Guettinger Y, Barbin K, Peipp M, Bruenke J, Dechant M, Horner H, et al. A recombinant bispecific single-chain fragment variable specific for HLA class II and Fc alpha RI (CD89) recruits polymorphonuclear neutrophils for efficient lysis of malignant B lymphoid cells. *J Immunol*. Feb 1;184(3):1210-7.
- [35] Bruenke J, Fischer B, Barbin K, Schreiter K, Wachter Y, Mahr K, et al. A recombinant bispecific single-chain Fv antibody against HLA class II and Fc gamma RIII (CD16) triggers effective lysis of lymphoma cells. *British journal of haematology*. 2004 Apr;125(2):167-79.
- [36] Cheng M, Santich BH, Xu H, Ahmed M, Huse M, Cheung NK. Successful engineering of a highly potent single-chain variable-fragment (scFv) bispecific antibody to target disialoganglioside (GD2) positive tumors. *Oncoimmunology*. 2016 Jun;5(6):e1168557.
- [37] Yamamoto K, Trad A, Baumgart A, Huske L, Lorenzen I, Chalaris A, et al. A novel bispecific single-chain antibody for ADAM17 and CD3 induces T-cell-mediated lysis of prostate cancer cells. *The Biochemical journal*. 2012 Jul 1;445(1):135-44.
- [38] Webb B, Sali A. Comparative Protein Structure Modeling Using MODELLER. *Current protocols in bioinformatics*. 2016 Jun 20;54:5 6 1-5 6 37.
- [39] Arcangeli C, Cantale C, Galeffi P, Rosato V. Structure and dynamics of the anti-AMCV scFv(F8): effects of selected mutations on the antigen combining site. *Journal of structural biology*. 2008 Oct;164(1):119-33.

- [40] Almagro JC, Beavers MP, Hernandez-Guzman F, Maier J, Shaulsky J, Butenhof K, et al. Antibody modeling assessment. *Proteins*. 2011 Nov;79(11):3050-66.
- [41] Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *Wiley Online Library* 1993:283-91.
- [42] Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic acids research*. 2007 Jul;35(Web Server issue):W407-10.
- [43] DeLano, W.L. The PyMOL molecular graphics system. 2002.
- [44] Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *Elsevier* 2015:19-25.
- [45] Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America*. 2001 Aug 28;98(18):10037-41.
- [46] N N, Zhu H, Liu J, V K, C GP, Chakraborty C, et al. Analysing the Effect of Mutation on Protein Function and Discovering Potential Inhibitors of CDK4: Molecular Modelling and Dynamics Studies. *PLoS one*. 2015;10(8):e0133969.
- [47] 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 Dec 1;57(4):829-38.
- [48] Kuroda Y, Suenaga A, Sato Y, Kosuda S, Taiji M. All-atom molecular dynamics analysis of multi-peptide systems reproduces peptide solubility in line with experimental observations. *Scientific reports*. 2016 Jan 28;6:19479.
- [49] Kamaraj B, Bogaerts A. Structure and Function of p53-DNA Complexes with Inactivation and Rescue Mutations: A Molecular Dynamics Simulation Study. *PLoS one*. 2015;10(8):e0134638.
- [50] Mohammadi M, Nejatollahi F, Sakhteman A, Zarei N. Insilico analysis of three different tag polypeptides with dual roles in scFv antibodies. *Journal of theoretical biology*. 2016 Aug 7;402:100-6.

Figures

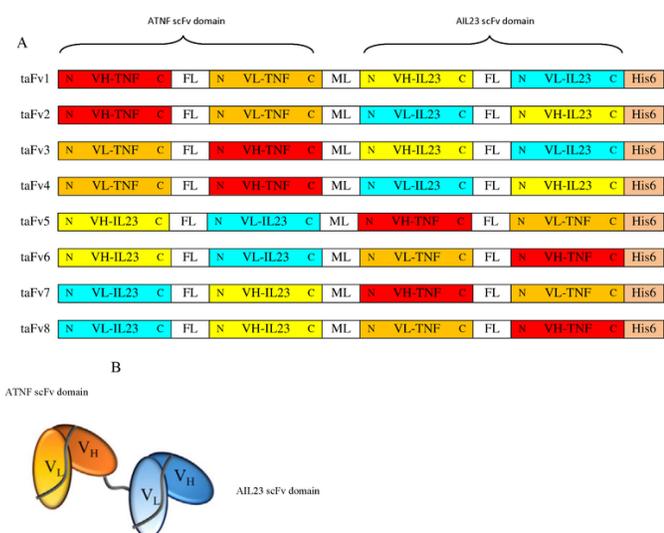


Fig. 1. Schematic view of designed bispecific taFv constructs (A) and one bispecific taFv (B). (VL-TNF: Variable light of ATNF, VH-TNF: Variable heavy of ATNF, VH-IL23: Variable heavy of AIL23, VL-IL23: Variable light of AIL23, ML(middle linker)= FL(flanking linkers)=(GGGS)₃, His6: Hexa histidine tag).

Figure 1

Schematic view of designed bispecific taFv constructs (A) and one bispecific taFv (B). (VL-TNF: Variable light of ATNF, VH-TNF: Variable heavy of ATNF, VH-IL23: Variable heavy of AIL23, VL-IL23: Variable light of AIL23, ML(middle linker)= FL(flanking linkers)=(GGGS)₃, His6: Hexa histidine tag).

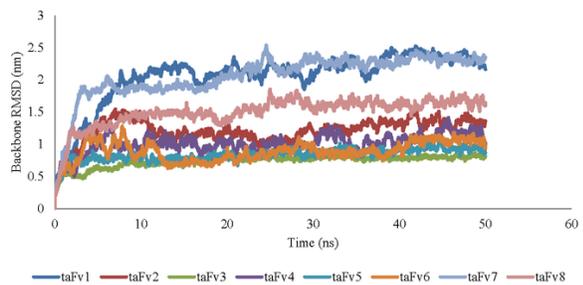


Figure 2. Backbone RMSD of all models during 50 ns MD simulation.

Figure 2

Backbone RMSD of all models during 50 ns MD simulation.

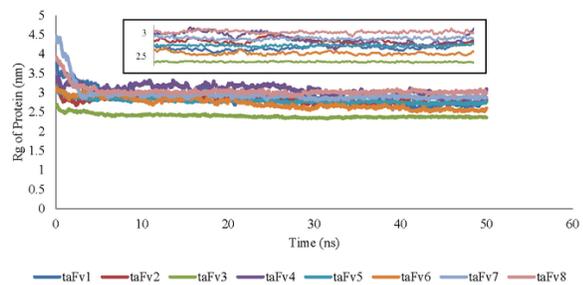


Figure 3. Rg of models during 50 ns MD simulation. The inset figure is related to the last 10 ns.

Figure 3

Rg of models during 50 ns MD simulation. The inset figure is related to the last 10 ns.

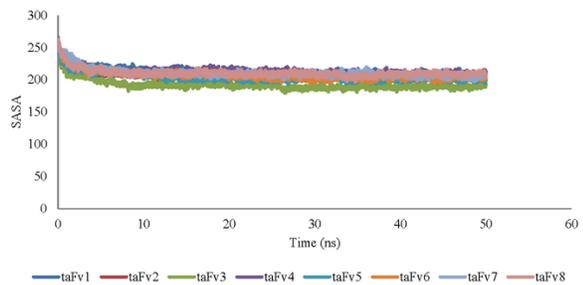


Figure 4. Total solvent accessible surface area (SASA) of models during 50 ns MD simulation.

Figure 4

Total solvent accessible surface area (SASA) of models during 50 ns MD simulation.

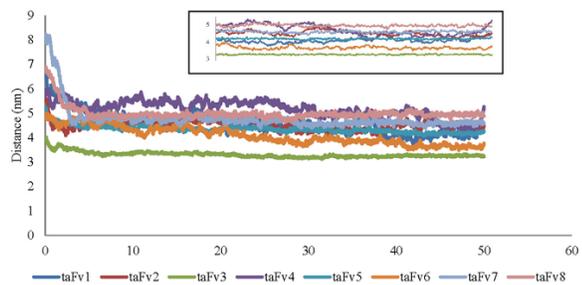


Figure 5. The center of mass distance between two scFv during 50 ns MD simulation

Figure 5

The center of mass distance between two scFv during 50 ns MD simulation.