Original Article

Theoretical design of a new chimeric protein for the treatment of breast cancer

Meysam Soleimani[1], Karim Mahnam[2], Hamid Mirmohammad-Sadeghi[1,3], Hojjat Sadeghi-Aliabadi[3,4], and Ali Jahanian-Najafabadi[1,3,*]

[1]Department of Pharmaceutical Biotechnology, School of Pharmacy and Pharmaceutical Sciences, Isfahan University of Medical Sciences, Isfahan, I.R. Iran.
[2]Biology Department, Faculty of Sciences, Shahrekord University, Shahrekord, I.R. Iran.
[3]Isfahan Pharmaceutical Sciences Research Center, School of Pharmacy and Pharmaceutical Sciences, Isfahan University of Medical Sciences, Isfahan, I.R. Iran.
[4]Department of Medicinal Chemistry, School of Pharmacy and Pharmaceutical Sciences, Isfahan University of Medical Sciences, Isfahan, I.R. Iran.

Abstract

p28 and NRC peptides are two anticancer peptides with various mechanisms have shown to be effective against breast cancer. Therefore, it seems that construction of a chimeric protein containing the two peptides might cause synergistic cytotoxic effects. However, since the two peptides bear opposite charges, production of a chimeric protein in which the two moieties do not intervene each other is difficult. In this study, our goal was to find a suitable peptide linker for the new chimeric protein in a manner that none of the peptides intervene the other's function. We selected some linkers with different characteristics and lengths and created a small library of the chimeric proteins harboring these linkers. Homology modeling and molecular dynamic simulation revealed that $(PA)_5P$ and $(EAAAK)_3$ linkers can separate the p28 and NRC peptides effectively. Thus, the chimeric protein linked with $(PA)_5P$ or $(EAAAK)_3$ linkers might show synergistic and stronger anticancer effects than the separate peptide moieties because they could exert their cytotoxic effects freely which is not influenced by the other part.

Keywords: NRC; p28; Linker; Salt Bridge; Homology modeling; Molecular Dynamic Simulation

INTRODUCTION

The global burden of cancer continues to increase due to an aging and lifestyle changes (1). Breast cancer is considered as the most common type of cancer after skin cancer and the second reason of death in the United States after lung cancer (2). Although chemotherapy and hormone therapy play important roles in treating cancer, the results of prospective study display that considerable number of patients did not respond to these remedies. The aim of pharmacotherapy of cancer is to kill malignant cells while minimizing the side effects on normal cells. Current therapies such as chemotherapy are not specialized for cancer cells and have side effects on other cells in the body. Furthermore, they are not able to destroy the slow-growing and drug resistant cancer cells and have the potential to create secondary malignancies (3). One of the new approaches for pharmacotherapy is to use chimeric peptides and proteins.

Recent studies demonstrated that some peptides can be used for their various functions such as tumor homing, cell penetrating and anticancer. Peptides became appropriate for drug design due to their low toxicity and immunogenicity, in expensiveness and rich chemical diversity (4,5).

Various studies reported that a segment of azurin protein, called p28, has the ability of selective entrance and activating the apoptosis in breast cancer cells (6,7). Caveolae-mediated endocytosis is involved in preferential entrance of p28 in breast cancer cells. Unlike endocytosis mediated by clathrin, endocytosis mediated by caveolae and lipid raft occurs at

*Corresponding author: A. Jahanian-Najafabadi
Tel: 0098 31 37927056, Fax: 0098 31 36680011
Email: jahanian@pharm.mui.ac.ir

neutral pH and prevents the breakdown of peptide by intracellular protease or nuclease (8). So it seems that this peptide is suitable for preferential transitions of other peptides to cancer cells.NRC peptide belongs to cationic antimicrobial peptide (CAPs) families and its anti-cancer features on breast cancer cells has been reported (9). Mitochondrial membrane damage is a mechanism that has been identified for destroying cancer cells by this peptide (10).

Because of the restrictions in medications used in the treatment of cancer there is a requirement to develop new drugs, which are able to destroy cancer cells selectively. One of the approaches which seem to be helpful is using chimeric peptides and proteins (5). In fact, by combining different domains of proteins or different peptides together, chimeric molecules could be designed, which are able to destroy cancer cells selectively. Nowadays, chimeric proteins are promising approaches in designing new remedies for breast cancer treatment. The main challenge in this field is to design these molecules in a way that different domains do not intervene each other's function (11).

Most of proteins in nature consist of distinctive functional domains, which have been linked to each other through linker peptides (12). According to the information obtained from natural linkers, researchers have designed different empirical linkers, which can be divided into three categories of flexible, rigid and cleavable linkers.

Flexible linkers generally consist of small amino acids, distinctively glycine and serine (GS linkers) (13). Rigid linkers consist of two categories: α helix forming linkers and linkers containing proline (14). $(PA)_n$ linker is one of the rigid linkers containing proline and alanine that has been used for separation of functional domains of various fusion proteins (15). (EAAAK) linker is an alpha-helix forming linker structure that is able to prevent the interference between two green fluorescent proteins (16).

Currently, the use of computational techniques in chemistry and biology, from the quantum mechanics of molecules to the dynamics of large complex molecular aggregates, is popular. Molecular interactions steer chemical reactions, phase transitions and other physical phenomena and can be studied via molecular dynamics (MD) simulation that show the 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 free energy of drugs to them can be studied via MD simulation (17). MD simulation can be used for chimeric protein design and checking the location and wrapping, distance, interaction of residues and possibly intervening of the moieties (18). In this study, our aim was to select the most appropriate amino acid linker for designing a chimeric protein composed of p28 and NRC peptides, so that no intervening interactions like salt bridges is permitted. In this study, $(PA)_nP$ (n = 1-5), $(EAAAK)_n$ (n = 1-3), and $(G)_nS$ (n = 1-8) linkers were used to create a small library of the chimeric protein and then the library was analyzed by various bioinformatics tools.

MATERIALS AND METHODS

Peptide sequences

In the present study we designed a chimeric protein containing the p28 (LSTAADMQGVVTDGMASGLDKDYLKPDD) and NRC (GRRKRKWLRRIGKGVKIIGGAALDHL) peptide sequences. Since p28 peptide has 7 aspartate and glutamate residues with negative charges, and NRC has 9 lysine and arginine residues with positive charges, the probability of salt bridge formation is high. This, consequently, may affect their function, and therefore, a suitable linker must be used for separation of the two moieties. Hence, a small library of linker peptides with different characteristics and lengths, i.e. $((PA)_nP$ (n = 1-5), $(EAAAK)_n$ (n = 1-3), and $(G)_nS$ (n = 1-8)) were prepared (Table 1).

Table 1. Sequence, length and type of the linkers used in the study.

Linker sequence	Linker type
$(PA)_{1-5}P$	Rigid linker
$(EAAAK)_{1-3}$	Rigid linker
$(G)_{1-8}S$	Flexible linker

Homology modeling

Three dimentional structures of the chimeric proteins with different linkers were created by I-TASSER (http://zhang.bioinformatics.ku.edu/I-TASSER) web server (19). Afterward, visual inspection was performed with VMD 1.9.2 software (20) to investigate any formation of salt bridges between functional groups of the two moieties.

Then, MODELLER software 9.13(21) was used to obtain more accurate structures of models. Multiple template modeling was performed using the crystal structure of the selected templates. One hundred models have been made for each one of the chimeric proteins by homology modeling. The best model which had the least molpdf energy was chosen (22). Quality of the various models was assessed using Rampage (23) and ProSA-web (24). For this purpose, protein model coordinates were submitted to Rampage site and the quality of the created protein structures were evaluated by Ramachandran plots. In these plots, the number of residues on allowed or disallowed areas determine the quality of the created protein model (25). Also energy plots were created by ProSA-web site (23) as a function of residue position.

Molecular dynamic simulation

The best chimeric models that obtained from homology modeling were subjected to MD simulation. MD simulation and molecular mechanic minimization were performed using GROMACS 4.5.6 package (26) under Gromos force field (G43A1) (27). The system was neutralized by adding 4 Na^+ ions. MD simulations were carried out at the constant temperature, constant pressure ensemble (NPT) and periodic boundary condition. Van der Waals forces were treated using a cut-off of 12 Å. The electrostatic interactions were calculated using the Particle-Mesh Ewald model with a 14 Å cut-off (28). The chimeric proteins were solvated by a layer of at least 12 Å in all directions. The frequency to update the neighbor list was 10 ps. MD simulation was accomplished in four steps for each system. In the first step, the entire system was minimized using the steepest descent followed by conjugate gradients algorithms. In the second step, the solvent and Na^+ ions were allowed to evolve using minimization and molecular dynamics at theconstanttemperature, constant volume ensemble (NVT) for 500 ps and at NPT ensemble for 1000 ps at 100 K in which initial configuration of the structures was kept fixed. In the third step, in order to obtain equilibrium geometry at 300 K and 1 atm, the system was heated at a weak temperature coupling ($\tau = 0.1$ ps) and pressure coupling ($\tau = 0.5$ ps). Berendsen algorithm was chosen for thermostat and barostat in equilibration phase (29).

LINCS algorithm was used to constrain the lengths of hydrogen-containing bonds (30). The temperature of the system was then increased from 100 K to 300 K and the velocities at each step were reassigned according to the Maxwell-Boltzmann distribution at that temperature and equilibrated for 200 ps. In the final step or production step, 95 ns MD simulations at 300 K with a time step of 2 fs was performed for each chimeric protein and final structures were obtained. The thermostat and barostat for production step were Nosé-Hoover thermostat and Parrinello-Rahman barostat (29). The production phase was performed at 300 K with 2 fs time step.

RESULTS

Homology modeling

Three dimensional structures of the chimeric protein were predicted with $(PA)_{1-5}P$, $(EAAAK)_{1-3}$, and $(G)_{1-8}S$ linkers using I-TASSER. I-TASSER server predicted five models each time, in which the model by the least confidence scores was selected. After selection of best predicted models, visual inspection of salt bridges in 3D structure of these models was performed using VMD 1.9.2 software. As it was expected, the $(G)_nS$ (n = 1-8) could not separate the p28 from NRC peptide and salt bridge was formed between their charged groups (data not shown). On the other hand, the study of different lengths of other linkers showed that the $(PA)_5P$ and $(EAAAK)_3$ can prevent the formation of salt

bridges between charged residues of p28 and NRC peptides. MODELLER software was used to produce more accurate prediction of 3D structure and to improve the quality of models. There was not a suitable template for homology modeling of the chimeric proteins in database. Thus, NCBI database was searched by using BLAST algorithm against protein data bank (PDB) to identify homologous structures of the p28 peptide, NRC peptide, $(PA)_5P$ linker, and $(EAAAK)_3$ linker, separately.

Based on the highest sequence similarity, PDB entries 2FT6, 2MF9, and 2QAR were used as templates for p28 peptide, $(PA)_5P$, and $(EAAAK)_3$ linkers, respectively. Various templates were used for NRC peptide. However, the best result was achieved when the model predicted by I-TASSER was used as a template for NRC peptide. Then, the template chimera using alignment of the selected templates was created (Fig. 1). This chimera template was used for homology modeling by MODELLER.

The Ramachandran plot of the chimeric protein with $(EAAAK)_3$ linker showed 76.1%, 89.6% residues in favored region, 22.4%, 4.5% in allowed region, and 1.5%, 4.5% in outlier region in the obtained model from I-TASSER and MODELLER, respectively (Fig. 2. A and B). Similarly, these percentages for the chimeric protein with $(PA)_5P$ linker were 82.5%, 100.0% residues in favored region, 9.5%, 0.0% in allowed region, 7.9%, 0.0% in outlier region in the obtained model from I-TASSER and MODELLER, respectively (Fig. 2. C and D).

A

```
>P1;2FT6
structureX:2FT6:50:A:77:A:azurin:::
LSTAADMQGVVTDGMASGLDKDYLKPDD------------------------------------------------*
>P1;2QAR
structureX:2QAR:95:B:107:B::::
-------------------------EAEAAAAEAAAKA----------------
--------*
>P1;NRC TASSER
structureX:NRC TASSER:1::26:::::
------------------------------------------GRRKRKWLRRIGKGVKI
IGGAALDHL*

>P1;chimeric peptide
sequence:chimeric:::69:::::
LSTAADMQGVVTDGMASGLDKDYLKPDDEAAAKEAAAKEAAAKGRRKRKWLRRIGKGVKI
IGGAALDHL*
```

B

```
>P1;2FT6
structureX:2FT6:50:A:77:A:azurin:::
LSTAADMQGVVTDGMASGLDKDYLKPDD-------------------------------
-----*
>P1;2MF9
structureX:2MF9:25:A:33:A::::
-------------------------PEPAPAPAP------------------------
-----*
>P1;NRC Tasser
structureX:NRC Tasser:1::26:::::
-----------------------------------------GRRKRKWLRRIGKGVKIIGGA
ALDHL*
>P1;chimeric peptide
sequence:chimeric:::65:::::
LSTAADMQGVVTDGMASGLDKDYLKPDDPAPAPAPAPAPGRRKRKWLRRIGKGVKIIGGA
ALDHL*
```

Fig. 1. The alignment of chimeric proteins for creation of template chimera. A; the alignment for the chimeric protein with $(EAAAK)_3$ linker and B; the alignment for the chimeric protein with $(PA)_5P$ linker.
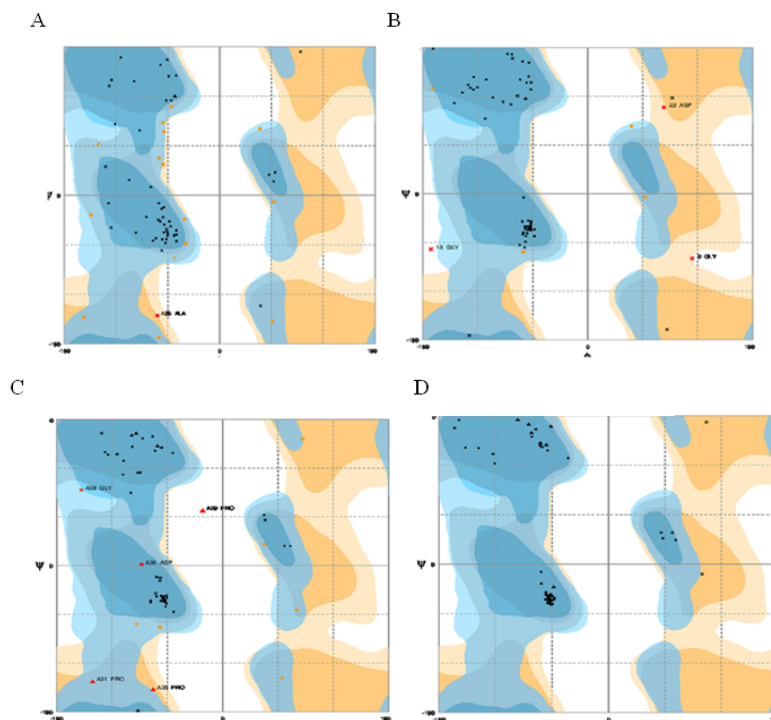
Fig. 2. The Ramachandran plot of the chimeric protein with (EAAAK)$_3$ and (PA)$_5$P linker. The chimeric protein with (EAAAK)$_3$ linker. A; model obtained from I-TASSER and B; model obtained from MODELLER. The chimeric protein with (PA)$_5$P linker. C; model obtained from I-TASSER and D; model obtained from MODELLER.
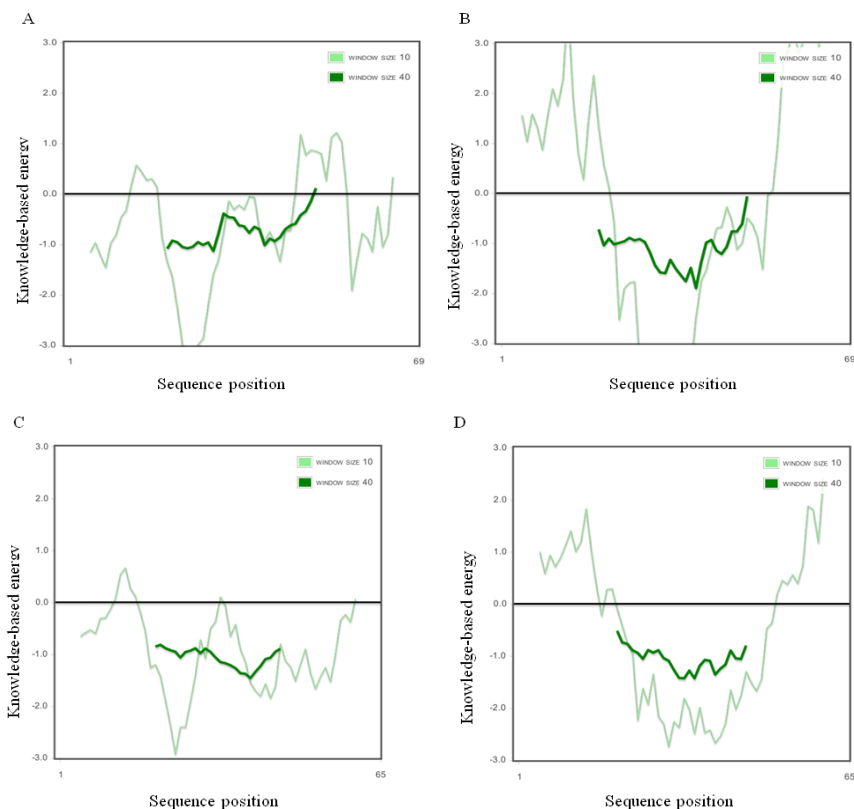

Fig. 3. Energy plot of two chimeric protein structures using the ProSA-web service. Energy resualt of the chimeric protein with (EAAAK)$_3$ linker. A; model obtained from I-TASSER and B; model obtained from MODELLER. Energy resault of the chimeric protein with (PA)$_5$P linker. C; model obtained from I-TASSER. D; model obtained from MODELLER.

We also applied the ProSA-web site for analysis of the chimeric protein. The ProSA-web site showed the energy plot of models in which positives correspond to problematic parts of the protein (24). The energies of residues in all models were negative with the exception of the N-terminal part of I-TASSER model for the chimeric protein with $(PA)_5P$ linker. Analysis of the ProSA-web results is depicted in Fig. 3.

Molecular dynamic simulation
Ninety five MD simulation was performed to evaluate the dynamic behavior of the chimeric protein.

Table 2 shows the results of average of the potential, kinetic energies, temperature, root mean square deviation (RMSD) of backbone relative to initial positions, Cα atoms root mean square fluctuation (RMSF), radius of gyration (Rg), ratio of the total energy drift to average of total energy, and distance between center of mass of two peptides (p28 and NRC) during the last 45 ns of the chimeric protein with the $(PA)_5P$ linker and 85 ns of the chimeric protein with the $(EAAAK)_3$ linker.

The RMSD value of the chimeric protein backbone atoms during MD simulation are shown in Fig. 4. The chimeric protein with $(PA)_5P$ and $(EAAAK)_3$ linkers reached a stable state after 40 ns and 10 ns of the MD simulation and became 1.2 ± 0.019 nm and 0.57 ± 0.016, respectively.

Also Rg after 40 ns for the chimeric protein with $(PA)_5P$ linker and after 10 ns for the chimeric protein with $(EAAAK)_3$ linker reached the Plateau at 1.12 ± 0.015 nm and 1.07 ± 0.011 nm, respectively (Fig. 5). These results suggest that 95 ns simulation was sufficient for stabilizing the chimeric proteins.

Table 2. Results of the last 45 ns and 85 ns of 95 ns MD simulation for the chimeric protein with $(PA)_5P$ and $(EAAAK)_3$ linkers.

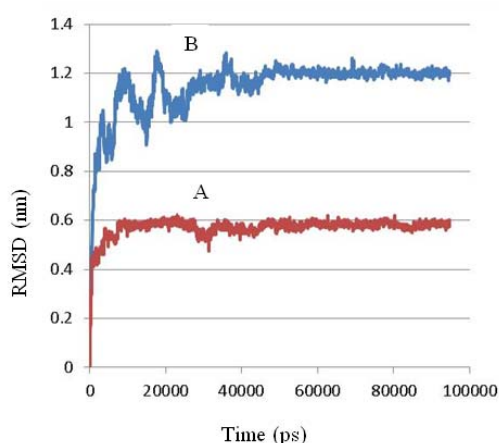| Parameters	Protein with $(PA)_5P$ linker	Protein with $(EAAAK)_3$ linker
Protein backbone RMSD (nm)	1.2 ± 0.019	0.57 ± 0.06
Protein backbone RMSF (nm)	0.47 ± 0.14	0.18 ± 0.06
Rg(nm)	1.12 ± 0.015	1.07 ± 0.011
Distance between center mass of two peptides (P28 and NRC)	1.15 ± 0.04	1.02 ± 0.03
Temperature (K°)	300 ± 2.35	300 ± 2.23
Potential (KJ/mol)	-247462 ± 603	-275541 ± 574
Total energy (KJ/mol)	-202427 ± 715	-225752 ± 709
Kinetic (KJ/mol)	45035 ± 353	49778 ± 369

Fig. 4. RMSD of the chimeric protein during 95 ns MD simulation. A; the RMSD for the chimeric protein with $(EAAAK)_3$ linker and B; the RMSD for the chimeric protein with $(PA)_5P$ linker.
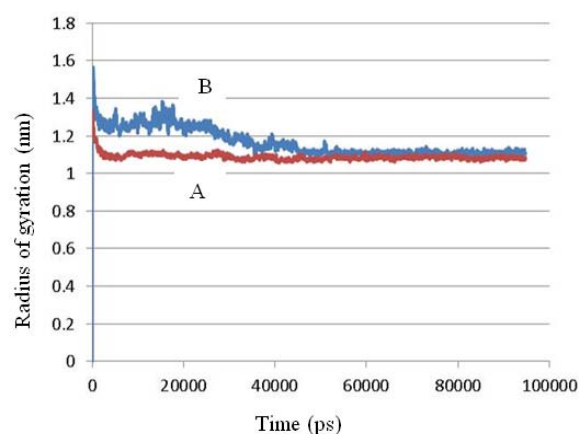

Fig. 5. Rg of the chimeric protein during 95 ns MD simulation. A; the Rg for the chimeric protein with $(EAAAK)_3$ linker and B; the Rg for the chimeric protein with $(PA)_5P$ linker.

We investigated the distance of six aspartate residues in p28 peptide and the lysine and arginine residues in NRC peptide using visual inspection in final 3D structure of the chimeric proteins by VMD 1.9.2 software. Then, the closest opposite charged residues in final 3D structures were selected to investigate the distance between them in MD simulation as a function of time (Figs. 6 and 7). In both chimeric proteins, the average distance and time of distance more than 4 Å between two carboxyl oxygen aspartate 6 with amide nitrogen atom depicted in Table 3.

We also calculated the distance between the geometrical centers of mass of p28 peptide from NRC peptide. As seen in Fig. 8 the average of this distance were 1.15 ± 0.04 nm for the chimeric protein with the $(PA)_5P$ linker and 1.02 ± 0.03 nm for the chimeric protein with the $(EAAAK)_3$ linker.
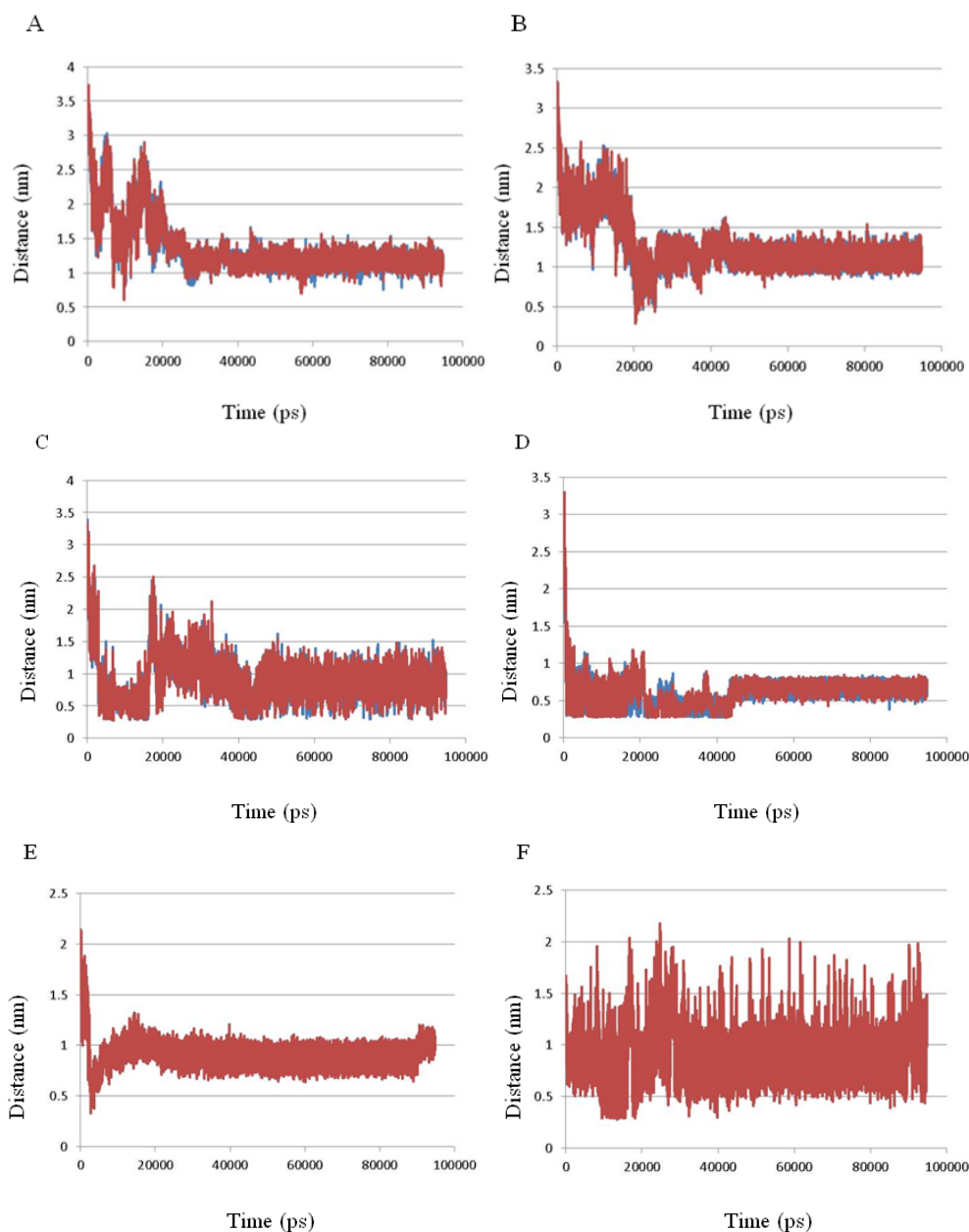

Fig. 6. The distances plotted between the N-O atoms in functional residue of the chimeric protein with $(PA)_3P$ linker. A; aspartate6-lysinee55, B; aspartate13-lysine55, C; aspartate20-argenine42, D; aspartate22-lysine55, E; aspartate27-lysine52 and F; aspartate28-argenine49. The blue line represents the distance between OD1-N and red line represents the distance between OD2-N.
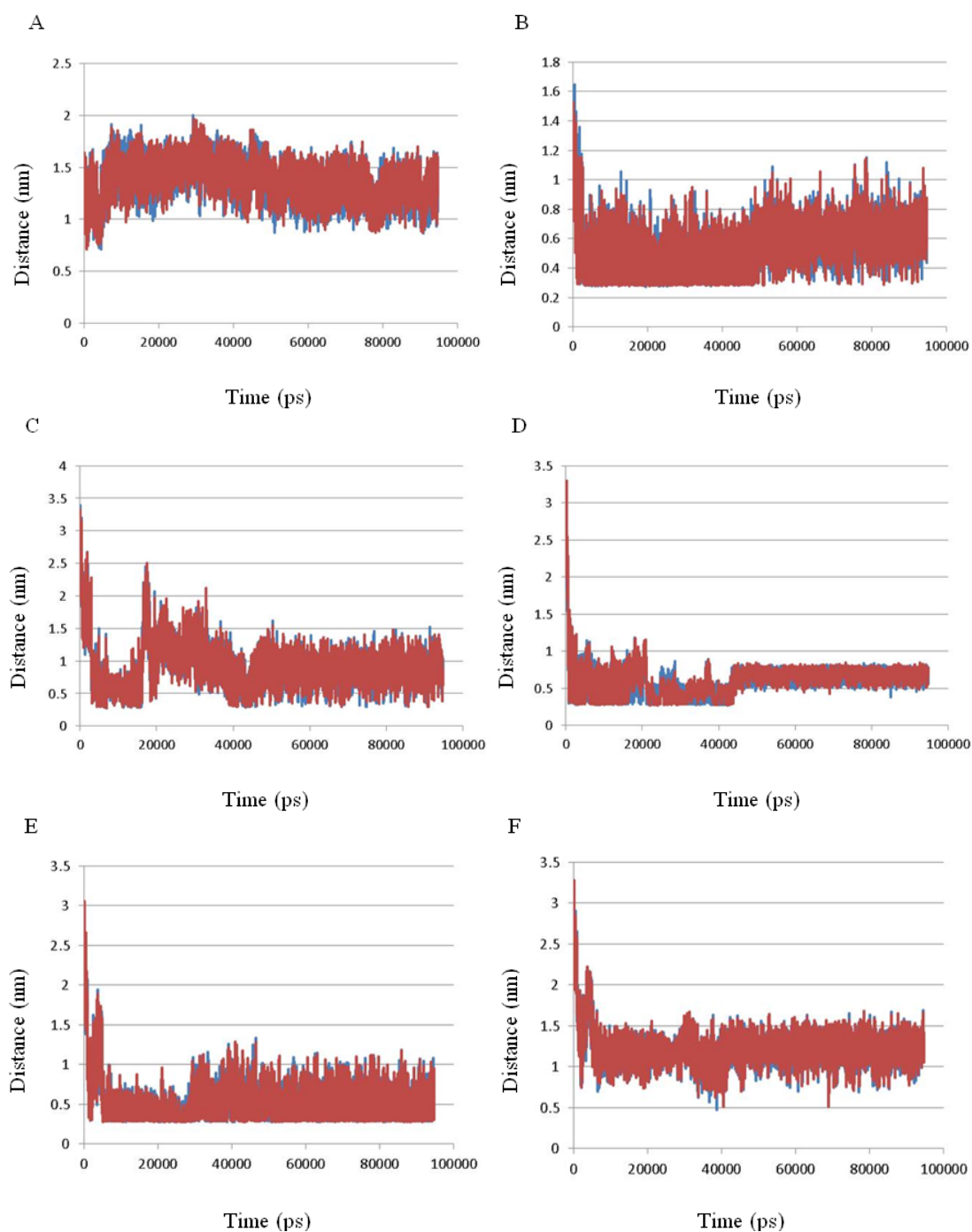
193

Fig. 7. The distances plotted between the N-O atoms in functional residue of the chimeric protein with $(EAAAK)_3$ linker. A; aspartate6-lysine56, B; aspartate13-lysine49, C; aspartate20-argenine53, D; aspartate22-lysine56, E; aspartate27-lysine56 and F; aspartate28-lysine56. The blue line represents the distance between OD1-N and red line represents the distance between OD2-N.

The flexibility of the chimeric protein was assessed by Cα RMSF (Fig. 9). It was observed that the chimeric protein with the $(PA)_5p$ linker have more fluctuations than the chimeric protein with the $(EAAAK)_3$ linker.

The average Cα RMSF for all residues in the chimeric protein with the $(PA)_5P$ linker was 0.47 ± 0.14 nm and it was 0.18 ± 0.06 nm for the chimeric protein with the $(EAAAK)_3$ linker.

Table 3. Average distance, standard deviation and time of distance more than 4 Å in the chimeric proteins with $(PA)_5P$ linker in the last 45 s of MD simulation or with $(EAAAK)_3$ linker in the last 85 s of MD simulation.

The chimeric protein with $(PA)_5P$ linker	**Opposite charged residue**	OD1D6-K55 OD2D6-K55	OD1D13-K55 OD2D13-K55	OD1D20-R42 OD2D20-R42	OD1D22-K55 OD2D20-K55	OD1D27-K52 OD2D27-K52	OD1D28-R49 OD2D28-R49
	Average distance in the last 45 s of MD simulation	1.19 1.18	1.14 1.14	0.85 0.85	0.66 0.66	0.87 0.87	0.91 0.90
	Standard deviation	0.09 0.09	0.09 0.09	0.19 0.19	0.08 0.08	0.07 0.07	0.20 0.20
	Time of distance more than 4 Å in the last 45 s of MD simulation	100% 100%	100% 100%	99.29% 99.31%	98.08% 98.38%	100% 100%	99.99% 99.99%
The chimeric protein with $(EAAAK)_3$ linker	**Opposite charged residue**	OD1D6-K56 OD2D6-K56	OD1D13-K49 OD2D13-K49	OD1D20-R53 OD2D20-R53	OD1D20-K56 OD2D20-K56	OD1D20-K56 OD2D20-K56	OD1D20-K56 OD2D20-K56
	Average distance in the last 85 s MD simulation	1.39 1.39	0.52 0.52	0.84 0.84	1.36 1.36	0.48 0.48	1.19 1.19
	Standard deviation	0.16 0.15	0.13 0.13	0.20 0.20	0.14 0.14	0.15 0.15	0.12 0.12
	Time of distance more than 4 Å in the last 85 s of MD simulation	100% 100%	81.05% 80.31%	98.58% 98.58%	100% 100%	70.25 67.90	100% 100%

D; Aspartic acid, K; Lysine, R; Arginine
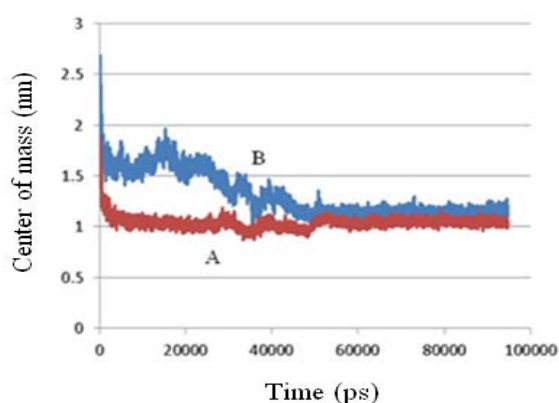
Fig. 8. Center of mass p28 from NRC peptides during 95 ns MD simulation of the chimeric proteins with A; center of mass in the chimeric protein with $(EAAAK)_3$ linker and B; center of mass in the chimeric protein with $(PA)_5P$ linker.
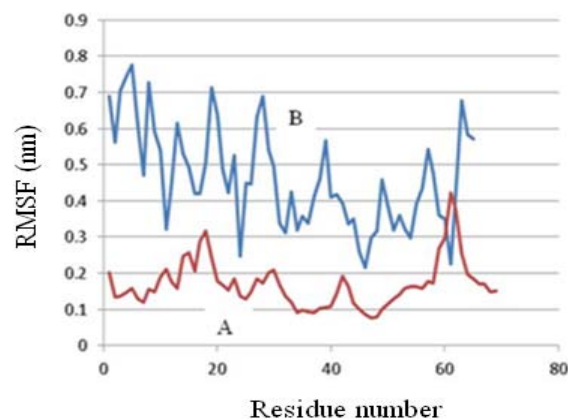
Fig. 9. RMSF value of Cα atoms during 95 ns MD simulation. A; RMSF value for the chimeric protein with $(EAAAK)_3$ and B; RMSF value the chimeric protein with $(PA)_5P$ linker.

DISCUSSION

At first, 3D structure of the chimeric proteins was predicted by I-TASSER web server. There was not a single suitable template for homology modeling of the chimeric protein with $(PA)_5P$ or $(EAAAK)_3$ linkers in databases. It has been demonstrated previously that using structural complementarity of the templates or segments matching in homology modeling improves the quality of models (31). In addition, the advantages of using multiple templates than single template have shown in other studies (32). We created a template chimera using different segment of templates (segment matching) that was structurally closer to the chimeric protein than each of the template separately. Then Clustal W (33) was used for the alignment of the chimeric proteins with different template sequences and creating the chimera template (Fig. 1). The quality of these models was evaluated by Rampage site and ProSA-web site and compared with I-TASSER result (Fig. 2 and 3). The Ramachandran plot of the chimeric protein with $(EAAAK)_3$ linker obtained from MODELLER showed that 89.6 percent of residues were placed in allowed regions that is higher than the percent of residues in allowed regions for structure obtained from I-TASSER (71.6%) (Figs. 2A and B). These percentage of residues in allowed regions for the chimeric protein with $(PA)_5P$ linker was 100% and 82.5% for MODELLER and I-TASSER, respectively (Figs. 2C and 2D). In both cases, the quality of the model obtained by the Modeller using template chimera was better than the quality model predicted by I-TASSER as seen in Fig. 2 and Fig. 3. The current study displayed that using template chimera, increases the accuracy of homology modeling. Obtaining better quality of the models from the MODELLER than the other software and web servers has been reported by other studies (32), hence the results produced should be reliable for molecular dynamic evaluation.

The models created by homology modeling cannot provide detailed information on the dynamic behavior of the chimeric proteins. The dynamic properties of a protein can be investigated by MD simulation (34). The MD simulation was used to gain a better comprehension of the salt bridge formation between charge residues in the chimeric proteins during MD simulation time (35). Small variations in potential and kinetic energies, temperature and RMSD of chimeric proteins and the very low ratio of the total energy drift to average of total energy showed that the simulation times were sufficient and those simulations were stable under simulation conditions and thermal equilibrium in the systems. We investigated the distance of six aspartate residues in p28 peptide and the lysine and arginine residues in NRC peptide using visual inspection in final 3D structure of the chimeric proteins by VMD 1.9.2 software. Then, the closest opposite charged residues in final 3D structures were selected to investigate the distance between them in MD simulation as a function of time. First, we had to select a distance criterion for the definition and formation of salt bridges in MD simulation. Kumar and coworkers defined salt bridge as ''an ion pair is defined as a salt bridge if the centroids of the side chain of charged group atoms in the residues lie within 4.0 Å of each other, and at least one pair of Asp or Glu side chain carbonyl oxygen and side chain nitrogen atoms of Arg, Lys, or His are also within this distance (36). Also the salt bridge was defined by Donald and colleagues as "an interaction between two groups of opposite charge in which at least one pair of heavy atoms is within hydrogen bonding distance" (37). As well, other studies define distance criterion ≤ 4 Å (38) or ≤ 3 Å (39) or 3.3 Å (40) for defining dehydrated salt bridge formation. According to distance criterion that defined in previous studies, we used 4 Å as a cut-off distance between the two carboxyl oxygen (OD1 or OD2) atoms of six aspartate in p28 peptide with the closest amide nitrogen atom of nine lysine and arginine in NRC peptide during the MD simulation to investigate the possibility of salt bridge formation between them. In the chimeric protein with $(PA)_5P$ linker, the nearest residue for aspartate 6, 13, 20, 22, 27, and 28 was lysine 55, lysine 55, arginine 42, lysine 55, lysine 52, and arginine 49, respectively. These residues for the chimeric

protein with $(EAAAK)_3$ linker, was lysine 56, lysine 49, arginine 53, lysine 56, lysine 56, and lysine 56, respectively. According to the 4Å criterion for salt bridge formation, as seen in Figs. 6, 7 and Table 2, these linkers are capable of separating the p28 and NRC peptide effectively so that the stable salt bridge did not form between their opposite charged residues. Centers of mass of p28 peptide from NRC peptide confirm these results (Fig. 8). Also, the p28 peptide has a lysine residue and NRC peptide has an aspartate residue that their distance in the final structure in both chimeric proteins is more than 2.2 nm.

Although the chimeric protein with the $(PA)_5P$ linker is four amino acid shorter than the chimeric protein with the $(EAAAK)3$ linker, the Rg of the one with the $(PA)_5P$ linker is bigger than that with the $(EAAAK)_3$(Fig. 5). Thus, the chimeric protein with the $(PA)_5P$ linker shows a more elongated conformation than the chimeric protein constructed with the $(EAAAK)_3$ linker. These results suggested that $(PA)_5P$ linker provides significant separation between the p28 and NRC peptide in the chimeric protein. Proline-rich linker forms a rigid and highly extended conformation and its open structure provides a large binding surface (41,42). Thus this linker provides a rigid spacer which leads to appropriate spatial arrangement of the p28 and NRC peptides. This arrangement might result in efficient prevention of the interaction between the two peptide moieties.

The flexibility of the chimeric protein was assessed by Cα RMSF (Fig. 9). It is observed that the chimeric protein with the $(PA)_5p$ linker have more fluctuations than the chimeric protein with the $(EAAAK)_3$ linker. This difference is likely due to the fact that $(PA)_5p$ linker create a less defined structure and the $(EAAAK)_3$ linker creates a rigid helix structure. Usually, well-structured regions in the protein show low Cα atoms RMSF value, whereas loop regions show high Cα RMSF values (43). The RMSD, Rg, and Cα atom RMSF parameters during the MD simulation suggested that the chimeric protein with the $(PA)_5P$ linker is more flexible. Despite of more flexibility, comparison of N-O distance of functional groups and center of mass of p28

and NRC peptide as a function of time during MD simulation indicates that the $(PA)_5P$ linker is more successful in spatial separation of functional groups and therefore prevention of salt bridge formation.

CONCLUSION

The present study revealed that the 3D structure of the chimeric protein obtained from the MODELLER using template chimera have better quality than I-TASSER. The molecular dynamic simulation carried out in this study suggests that the $(PA)_5P$ and $(EAAAK)_3$ linkers can effectively separate p28 peptide from NRC peptide. The observation suggests that the chimeric proteins constructed via $(PA)_5P$ and $(EAAAK)_3$ linkers could be successfully evaluated for their synergistic and specific breast cancer cells killing activities. In addition, this study showed that homology modeling and molecular dynamic simulation can be used in the design of chimeric proteins with special features. Further *in vitro* studies on expression and purification of the structures obtained in this study are undergoing.

ACKNOWLEDGMENTS

REFERENCES

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics. CA Cancer J Clin. 2014;64:9-29.
2. DeSantis C, Ma J, Bryan L, Jemal A. Breast cancer statistics, 2013. CA Cancer J Clin. 2014;64:52-62.
3. Tan M, Lan KH, Yao J, Lu CH, Sun M, Neal CL, *et al*. Selective inhibition of ErbB2-overexpressing breast cancer *in vivo* by a novel TAT-based ErbB2-targeting signal transducers and activators of transcription 3-blocking peptide. Cancer Res. 2006;66:3764-3772.
4. Craik DJ, Fairlie DP, Liras S, Price D. The future of peptide-based drugs. Chem Biol Drug Des. 2013;81:136-147.
5. Thundimadathil J. Cancer treatment using peptides: current therapies and future prospects. J Amino Acids.2012,2012:ID 967347.
6. Punj V, Bhattacharyya S, Saint-Dic D, Vasu C, Cunningham EA, Graves J, *et al*. Bacterial cupredoxin azurin as an inducer of apoptosis and

regression in human breast cancer. Oncogene. 2004;23:2367-2378.

7. Yamada T, Fialho AM, Punj V, Bratescu L, Gupta TK, Chakrabarty AM. Internalization of bacterial redox protein azurin in mammalian cells: entry domain and specificity. Cell Microbiol. 2005;7:1418-1431.

8. Taylor BN, Mehta RR, Yamada T, Lekmine F, Christov K, Chakrabarty AM, *et al*. Noncationic peptides obtained from azurin preferentially enter cancer cells. Cancer Res. 2009;69:537-546.

9. Hilchie AL, Doucette CD, Pinto DM, Patrzykat A, Douglas S, Hoskin DW. Pleurocidin-family cationic antimicrobial peptides are cytolytic for breast carcinoma cells and prevent growth of tumor xenografts. Breast Cancer Res. 2011;13:R102.

10. Morash MG, Douglas SE, Robotham A, Ridley CM, Gallant JW, Soanes KH. The zebrafish embryo as a tool for screening and characterizing pleurocidin host-defense peptides as anti-cancer agents. Dis Model Mech. 2011;4:622-633.

11. Chen X, Zaro JL, Shen WC. Fusion protein linkers: property, design and functionality. Adv Drug Deliv Rev. 2013;65:1357-1369.

12. Wriggers W, Chakravarty S, Jennings PA. Control of protein functional dynamics by peptide linkers. Biopolymers. 2005;80:736-746.

13. Klein JS, Jiang S, Galimidi RP, Keeffe JR, Bjorkman PJ. Design and characterization of structured protein linkers with differing flexibilities. Protein Eng Des Sel. 2014;27:325-330.

14. George RA, Heringa J. An analysis of protein domain linkers: their classification and role in protein folding. Protein Eng. 2002;15:871-879.

15. Zhao HL, Yao XQ, Xue C, Wang Y, Xiong XH, Liu ZM. Increasing the homogeneity, stability and activity of human serum albumin and interferon-alpha2b fusion protein by linker engineering. Protein Expr Purif. 2008;61:73-77.

16. Arai R, Ueda H, Kitayama A, Kamiya N, Nagamune T. Design of the linkers which effectively separate domains of a bifunctional fusion protein. Protein Eng. 2001;14:529-532.

17. Karplus M, Kuriyan J. Molecular dynamics and protein function. Proc Natl Acad Sci U S A.2005;102:6679-6685.

18. Mahnam K, Saffar B, Mobini-Dehkordi M, Fassihi A, Mohammadi A. Design of a novel metal bindingpeptide by molecular dynamics simulation to sequester Cu and Zn ions. Res Pharm Sci. 2014;9:69-82.

19. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc. 2010;5:725-738.

20. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. J Mol Graph. 1996;14:33-38.

21. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, *et al*. Comparative protein structure modeling using MODELLER. Curr Protoc Protein Sci. 2007;Chapter 2:Unit 29.

22. Eswar N, Eramian D, Webb B, Shen MY, Sali A. Protein structure modeling with MODELLER. Methods Mol Biol. 2008;426:145-159.

23. Lovell SC, Davis IW, Arendall WB3rd, de Bakker PI, Word JM, Prisant MG, *et al*. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins. 2003;50:437-450.

24. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2007;35:407-410.

25. Hooft RWW, Sander C, Vriend G. Objectively judging the quality of a protein structure from a Ramachandran plot. Comput Appl Biosci. 1997;13:425-430.

26. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, *et al*. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics. 2013;29:845-854.

27. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. J Comput Chem. 2005;26:1701-1718.

28. Darden T, York D, Pedersen L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. J Chem Phys. 1993;98:10089-10092.

29. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys. 1984;81:3684-3690.

30. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A linear constraint solver for molecular simulations. J Comput Chem. 1997;18:1463-1472.

31. Chakravarty S, Godbole S, Zhang B, Berger S, Sanchez R. Systematic analysis of the effect of multiple templates on the accuracy of comparative models of protein structure. BMC Struct Biol. 2008;8:31.

32. Larsson P, Wallner B, Lindahl E, Elofsson A. Using multiple templates to improve quality of homology models in automated homology modeling. Protein Sci. 2008;17:990-1002.

33. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, *et al*. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23:2947-2948.

34. Karplus M, Kuriyan J. Molecular dynamics and protein function. Proc Natl Acad Sci U S A. 2005;102:6679-6685.

35. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. Nat Struct Biol. 2002;9:646-652.

36. Kumar S, Nussinov R. Close-range electrostatic interactions in proteins. Chembiochem. 2002;3:604-617.

37. Donald JE, Kulp DW, DeGrado WF. Salt bridges: geometrically specific, designable interactions. Proteins. 2011;79:898-915.

38. Barlow DJ, Thornton JM. Ion-pairs in proteins. J Mol Biol. 1983;168:867-885.

39. Holbrook JA, Tsodikov OV, Saecker RM, Record MT, Jr. Specific and non-specific interactions of

integration host factor with DNA: thermodynamic evidence for disruption of multiple IHF surface salt-bridges coupled to DNA binding. J Mol Biol. 2001;310:379-401.

40. Ma L, Sundlass NK, Raines RT, Cui Q. Disruption and formation of surface salt bridges are coupled to DNA binding by the integration host factor: a computational analysis. Biochemistry. 2011;50: 266-275.

41. Williamson MP. The structure and function of proline-rich regions in proteins. Biochem J. 1994;297:249-260.

42. Boze H, Marlin T, Durand D, Perez J, Vernhet A, Canon F, *et al*. Proline-rich salivary proteins have extended conformations. Biophys J. 2010;99:656-665.

43. Saladino AC, Xu Y, Tang P. Homology modeling and molecular dynamics simulations of transmembrane domain structure of human neuronal nicotinic acetylcholine receptor. Biophys J. 2005;88:1009-1017.