RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

Exploring the binding properties of agonists interacting with human TGR5 using structural modeling, molecular docking and dynamics simulations

Thangaraj Sindhu and Pappu Srinivasan*

Molecular Biology Lab, Department of Bioinformatics, Alagappa University, Karaikudi - 630004, Tamilnadu, India

***Corresponding author**

Dr. P. Srinivasan, M.Sc., Ph.D.

Assistant Professor

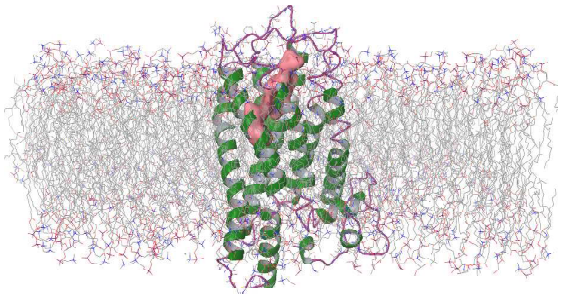
Tel: +91-4565-230725

Mobile: +91-9444482814

E-mail: sri.bioinformatics@gmail.com

Table of contents entry

TGR5**,** a membrane-bound receptor for bile acids act as a potential pharmacological target in the treatment of type II diabetes. In the computational study, structural modeling and binding site prediction of TGR5 receptor was performed. Two well-known agonists of TGR5 used to investigate the mode and mechanism of binding.



Abstract

TGR5, G-protein coupled receptor acts as a promising target for the treatment of diabetes, obesity and metabolic syndromes. Understanding the activation of TGR5, structural conformation and the mode of mechanism upon binding with agonists is crucial for the development of new drugs. In the absence of experimental data, homology modeling was performed to predict the structure of TGR5. Molecular dynamics simulation of 100 ns was performed to investigate the stability of the constructed model embedded in lipid bilayer. A combined method consisting of molecular docking and binding free energy calculations was performed to understand the binding mechanism of two experimentally proved selective TGR5 agonists. Additionally, 30 ns of protein-ligand complex dynamics were performed to reproduce the mechanism of interaction. Both agonists shared similar binding mode and showed four common hydrogen bonding interactions with TGR5. Thus, the results could provide more knowledge on the activation of TGR5 by agonists and helpful in the development of novel potent agonists.

Keywords: TGR5, homology modeling, agonists, membrane-based molecular dynamics, essential dynamics

1. Introduction

Diabetes mellitus, as one of the most prevalent and multifactorial metabolic disorders with abnormalities of carbohydrate, lipid, protein, triglycerides and closely associated with many other complications. It is characterized by the imperfection of insulin secretion in the pancreas and/or insulin action.[1-3] The prevalence of diabetes rapidly increasing and considered to be one of the major and growing public health problems worldwide. According to International Diabetic Federation, approximately 246 million people are affected by diabetes and anticipated 380 million in 2025.[4] Takeda G-protein coupled receptor 5 (TGR5), a member of G-protein coupled receptor family that composed of seven transmembrane helices. It is also known as membrane-type receptor for bile acids (M-BAR) and G-protein coupled bile acid receptor 1 (GPBAR1).[5] TGR5 receptor is expressed predominantly in various tissues including intestine, spleen, gall bladder, placenta, ovary, lung, brown adipose tissue, macrophages and monocytes.[6, 7]

Two categories of agonists (natural and synthetic) have been used for the activation of TGR5 receptor. Bile acids (BA) such as cholic acid (CA) and lithocholic acid (LCA) are considered as naturally occurring TGR5 agonists. BAs act as potent signal molecules that can activate TGR5 receptor.[8] Regulation of gene expression and activation of adenylyl cyclase cAMP signaling pathways can be achieved through the bile acid binding TGR5.[9] Previous reports suggest that activated TGR5 by BAs improved insulin sensitivity and glucose tolerance.[10] Synthetically derived small molecules are another category of TGR5 agonists that acts more potent than natural agonists. Apart from these two categories of agonists, semi synthetic agonists were also used to activate TGR5 receptor.

Recently, two potent semi synthetic derivatives were identified as agonists of TGR5 including 6α-ethyl-23(S)-methyl- 3α,7α,12α-trihydroxy-5β-cholan-24-oic acid (INT-777), a specific TGR5 agonist and 6α- ethyl-3α ,7α ,23-trihydroxy-24-nor-5β-cholan-23-sulfate sodium salt (INT-767), dual FXR/TGR5 agonist.[11, 12] Among the two potent agonists of TGR5, INT-777 exhibited excellent agonistic properties *in vivo* and *in vitro* were considered as a potent drug candidate for the treatment of type II diabetes.[5] The activation of TGR5 by INT-777 results in the regulation of insulin secretion from pancreas, stimulation of glucagon-like peptide-1 (GLP-1) release in intestine enteroendocrine cells, inhibition of glucagon release and reduction of appetite.[13] The dual agonist INT-767 has been reported to be a potential drug in the treatment of non-alcoholic fatty liver disease in obese diabetic mice.[14]

The present work mainly focused on the analysis of the structure and function of TGR5 membrane protein. Crystal structure of human TGR5 has not been solved in the experiment till now. Hence, the mechanistic study was initiated with the development of a homology model of human TGR5. Molecular dynamics (MD) simulation was carried out in order to investigate the structural conformation and stability of the protein embedded in the POPC lipid bilayer. The molecular mechanism of the binding characteristics of TGR5 remains unknown. Thus, the determination of interaction characteristics of TGR5 with agonists is essential. In the present study, two potent agonists such as INT-777 and INT-767 were used to investigate the mode and mechanism of binding using combined molecular docking, molecular dynamics and binding free energy calculations.

2. Results and Discussion

Homology model construction and validation

A total of 5 models were generated using I-Tasser server that combines the methods of threading, *ab initio* modeling and structural refinement. Among the five similar models, model 1 suggested as a good model based on the C-score generated in I-Tasser. However, all 5 models were validated using Procheck server in order to evaluate the stability and reliability of the generated models. Ramachandran plot was created in Procheck to check whether the constructed model was folded correctly relative to the solved crystal structure of proteins available in protein data bank.[15] A model having more than 80% residues in the most favoured region is considered to be a good quality model.[16] From the Procheck results, only model 4 was found to have highest residues (88.3%) in the favoured region of Ramachandran plot. Ramachandran plot obtained for model 4 represented in Supplementary figure 1 **(S1)** shows that psi/phi angle of a majority of the residues are present in most favoured region. A total of 241 residues (88.3%) were distributed in the most favored region, 9.2% including 25 residues in the allowed region, 1.8% including 5 residues in the generously allowed region, 0.7% including two residues such as Pro69 and Asn4 present in the disallowed region, suggesting that the predicted structure of TGR5 (model 4) is satisfactory. The percentage was calculated only for 273 non-glycine and non-proline residues distributed in the different four regions of the Ramachandran plot. Side chain of Glycine has only one hydrogen atom and therefore can adopt phi and psi angles in four quadrants of the Ramachandran plot. Proline has 5-membered-ring side chain connecting cα to backbone N and thereby shows a limited combination of phi and psi angles. Due to these reasons, Glycine (27)

and Proline (28) residues and end residues (2) are not included for calculating the percentage. All GPCR family proteins shares common structural features of seven hydrophobic transmembrane helices with an extracellular amino (N) terminus and an intracellular carboxyl (C) terminus.[17] The predicted three dimensional structure of TGR5 based on Procheck results displayed in **Figure 1**, reveals that the model consisting of seven transmembrane α- helices and disulphide bridges.

The Z-score generated using ProSA analysis is standardized statistically-derived structure quality assessment scales that include chi-1/chi-2 rotamer normality, packing quality, backbone conformation and ramachandran plot appearance.[18] ProSA analysis in **S2** showed Z-score of -5.6 kcal/mol, suggesting that the overall quality of the model is good and it is near to that of experimental protein structures that solved by Xray and NMR. The calculated interaction energy using ProSA shows that the maximum number of residues has negative interaction energy, whereas very few residues have positive interaction energy that confirms the reliability of the model. Overall, the Procheck and ProSA analysis suggest the model 4 as stable and suitable conformation for further work. Furthermore, the detailed quality of the model 4 was analyzed using exhaustive docking studies.

MD simulation analysis-Phase I

Stability profile analysis by Root mean square deviation

Membrane protein accounts for 70% of drug targets to treat various diseases. About only 2% of crystal structure of membrane proteins have been solved and deposited in Protein Data Bank. Due to the limited high resolution crystal structure information, application of MD simulations among membrane proteins plays an important role in finding stability and flexibility of the modeled structure embedded in lipid bilayer.[19] In the present study, MD simulation of 100 ns was carried out to explore the most recent homology model of TGR5 within the hydrated lipid environment using POPC bilayer. Conformational and geometrical properties obtained from the MD trajectories were examined to understand the dynamic changes and structural stability of $TGR5_{apo}$. Furthermore, MD simulation of modeled $TGR5_{apo}$ was carried out in order to obtain a refined structure with least potential energy. The POPC membrane is typically divided into two general regions including hydrocarbon core region and interfacial region. The hydrocarbon core in the center is dominated by the aliphatic side chain and the interfacial region comprises lipid head groups. The interfacial region of lipid bilayer is considered to be exposing to solvents.[20]

Two snapshots of protein depicted in **Figure 2**, clearly explain the structural changes observed in $TGR5_{apo}$ during the simulation time of 0 ns and 100 ns. From the figure 2, it was observed that the C-terminal and N-terminal end of the protein was buried into the interfacial region of the bilayer whereas the centre part of TGR5 was buried into the hydrocarbon core of the membrane.

Dynamic stability of $TGR5_{apo}$ over the simulation of 100 ns was analyzed using backbone root mean square deviation (RMSD) of cα atoms. The plot, describing the backbone RMSD of the $TGR5_{apo}$ state during the simulation time of 100 ns is shown in **Figure 3**. From this figure, it can be seen that the protein RMSD became stable after initial deviation. Further, it maintained stable conformation throughout the simulation time of 100 ns. An initial rise of RMSD over the first 2 ns was due to the relaxation of $TGR5_{apo}$ in the lipid bilayer membrane environment. The RMSD of backbone cα atoms of $TGR5_{apo}$ was observed from 0.2 nm to 0.46 nm that indicates the stability and equilibration of protein. The exhibited average (mean), RMSD of backbone cα atoms of $TGR5_{apo}$ was 0.41 nm with calculated means of standard deviation (SD) of 0.03 nm. The less RMSD of backbone atoms strongly suggests the stable dynamic behavior of TGR5 model and can be used for docking studies. Furthermore, Gromacs inbuilt tool g_energy was used to calculate the potential energy of protein at all frames and the protein conformation with least potential energy was taken for protein-ligand interaction studies.

Root mean square fluctuation

About 330 residues of $TGR5_{apo}$ were calculated by means of root mean square fluctuation (RMSF) to explore the flexibility of each residue. RMSF values calculate the magnitudes of each residual fluctuation. The magnitudes of the fluctuation of each residue in a protein are represented by the heights of peaks in **Figure 4**. Few amino acids of TGR5 showed minor fluctuations, including Asn4 (0.47 nm), Ser5 (0.41 nm), Thr6 (0.42 nm), Leu202 (0.50 nm), Glu203 (0.41 nm), Arg204 (0.59 nm), Ala205 (0.43 nm), Val206 (0.44 nm) and Cys208 (0.46 nm). Among these amino acids, Arg204 showed increased flexibility than other residues of TGR5. From the results of RMSF, it was concluded that all the large fluctuations of the residues were more pronounced in loop regions and relatively very small fluctuations were occurred in the α-helical regions of TGR5.

Solvent accessible surface area (SASA)

Biomolecular surface area that assessable to solvent molecules was calculated using g_sas and plotted in **Figure 5a**. Non-polar solvation free energy was estimated by means of SASA. The free energy of the non-polar solvation of each atom in a molecule is proportional to the SASA. The term non polar accounts for the rearrangement of solvent molecules around the solute and interaction of van der Waals contact between the solute and solvent molecules.[21] Conformational changes of modeled TGR5 over the simulation period can be estimated using SASA calculations. The calculated average SASA value for $TGR5_{apo}$ was 802 nm. SASA values obtained for $TGR5_{apo}$ during the simulation of 100 ns were relatively stable, indicating that no significant changes in the structure of TGR5. From the SASA graph, it was observed that the accessibility was retained at around 10 ns. The results confirmed that the residues of TGR5 were well exposed and accessible to the solvent.

Radiation of gyration

Radiation of gyration (Rg) is another probe to identify structural stability. Rg was calculated to measure the compactness of modeled structure. The plot illustrated in **Figure 5b** shows that Rg values were stabilized at 30 ns, indicating the equilibration of protein. The average Rg value of TGR5 at 100 ns was found to be 2.23 nm with the SD of 0.01 nm. The average value of Rg was calculated from the trajectory file generated from MD simulations. Relatively steady Rg value was maintained over the course of 100 ns indicating the stable folding of modeled protein TGR5. Rg of protein were well correlated with the backbone RMSD of cα atoms. Further, Dictionary of secondary structure of proteins (DSSP) analysis was compared and discussed below with two agonists of TGR5.

Binding mode analysis of INT-777

Molecular docking studies were carried out using the TGR5 structure with least potential energy (average structure) obtained from MD simulation. The conformation with least potential energy was observed at 8.2 ns of MD simulation as shown in **Figure 6**. Induced fit docking studies were performed to analyze the binding mode of two agonists of TGR5. Glide and IFD score was generated based on the binding affinity of each complex. According to the docking results, the first pose corresponds to the lowest docked energy was taken for analysis and further studies. Docking analysis of INT-777 with TGR5 showed the docking score of -16.689 kcal/mol and glide energy of -71.414 kcal/mol. The highest docking score obtained from IFD indicates the better interaction of INT-777 agonist with the active site of TGR5. Binding mode analysis of

INT-777 showed hydrogen bond interactions with active site residues of TGR5 such as Ser21, Asn76, Tyr89, Asn93, Ser157, Tyr240 and Ser270, forming a total of six hydrogen bonds. Docking results revealed that hydroxyl groups present in INT-777 formed hydrogen bonding interactions with Ser21 (HO···OH, bond length=1.73Å), Asn76 (NH_2···OH, bond length=1.77Å), Ser157 (C=O···HO, bond length=1.82Å) and Ser270 (OH.··OH, bond length=2.02Å). Further, the carboxyl group of INT-777 formed hydrogen bond interaction with the amine group of Asn93 (NH_2···O=C, bond length=1.80Å) whereas a hydroxyl group of active site residues including Tyr89 (OH···O-C, bond length=1.97Å) and Tyr240 (OH···O-C, bond length=2.02Å) interacted with the oxygen atom of agonist. In addition, active site residues including Ala17, Leu18, Leu24, Leu68, Pro69, Pro72, Trp75, Pro92, Phe96, Leu97, Leu263 and Leu266 were involved in hydrophobic contact with INT-777. The IFD score of the compound INT-777 was -599.485. From the analysis, it was observed that only polar amino acids (Ser21, Ser157, Ser270, Asn76, Asn93, Tyr89 and Tyr240) play a crucial role in the formation of hydrogen bonds with INT-777. Hydrogen bonding and hydrophobic interactions between INT-777 and active site residues of TGR5 play an important role in the stabilization of ligand conformation at the binding pocket.

Binding mode analysis of INT-767

Docking score of INT-767 was -12.207 kcal/mol whereas, the glide energy and IFD score was -64.534 kcal/mol and -599.685 kcal/mol, respectively. Binding mode analysis of INT-767 showed hydrogen bonding interactions with five active site residues of TGR5 such as Ser21 (HO···HO, bond length =2.03Å), Tyr89 (OH···O-S-O, bond length=1.78Å), Trp237 (NH···O=S, bond length=2.45Å), Tyr240 (OH···O-S-O, bond length=1.76Å) and Ser270 (HO···HO, bond length=2.06Å). From the interaction pattern, it was observed that sulfate ion present in INT-767 established hydrogen bonds with Tyr89, Trp237 and Tyr240. The hydroxyl group of INT-767 was interacting with Ser21 and Ser270, forming hydrogen bonds. The compound INT-767 was also adjacent to some hydrophobic residues of binding site, including Ala17, Leu18, Leu24, Ala25, Leu68, Pro69, Pro72, Pro92, Phe96, Leu97, Trp75, Leu263 and Leu266. Formation of hydrogen bonds between TGR5 and INT-767 was associated with only polar amino acids. The interactions of INT-777 and INT-767 with active site residues of TGR5 are shown in Figure 6a and 6b**,** respectively.

Docking procedures are usually used to predict the affinity between the ligand and protein through the identification of correct conformation of ligands in the binding pocket of a

protein.[22] In the present study, the correct conformation of two agonists into the binding site of TGR5 was predicted. The affinity between the two agonists and TGR5 was identified using the interaction pattern and docking scores. The highest docking score obtained for both agonists suggests that the ligands strongly bound with the TGR5 receptor. The binding site predicted for TGR5 was compared with other GPCR family proteins. From the comparison results, it was observed that the binding site of TGR5 share great similarity with other GPCR family proteins including β2-Adrenergic receptor[17], δ-opioid receptor[23] and human A_{2A} adenosine receptor.[24] The results evident that the computational methods used in the study correctly predicted the agonist binding site of TGR5. The binding mode analysis of two agonists was compared with the experimentally observed results. From the comparative observation, it was found that both compounds produced almost similar results and well correlating with each other. Overall, the interaction pattern and docking scores obtained for both the complexes suggests that INT-777 was well interacted with TGR5 than INT-767. Cross comparison of interacting residues among two agonists revealed that Ser21, Tyr89, Tyr240 and Ser270 were formed interactions with both the compounds. The comparison results concluded that the explicit agonists play an important role in the orientation of four common amino acids such as Ser21, Tyr89, Tyr240 and Ser270 and maintaining the stability of TGR5. The interaction pattern of the two complexes indicates that these common residues involved in the formation of hydrogen bonding interactions possess high affinity towards the ligands. The binding site of TGR5 can accommodate different potential drug molecules.

Binding free energy calculations

The two agonists complexed with TGR5 were rescored using Prime/MM-GBSA. The predicted binding energy (ΔG_{bind}) of the complexes $TGR5_{INT-777}$ and $TGR5_{INT-767}$ was -163.766 kcal/mol and -152.319 kcal/mol, respectively. Binding free energy of the $TGR5_{INT-777}$ system was much higher than that of $TGR5_{INT-767}$ system. The van der Waals interactions were significant contributions to the binding of ligands. Protein-ligand interactions were associated with the favourable polar contributions. The predicted van der Waals (-63.283 kcal/mol) and GB/SA solvation energy (102.362 kcal/mol) of $TGR5_{INT-777}$ was much stronger than the $TGR5_{INT-767}$ complex van der Waals (-57.506 kcal/mol) and GB/SA solvation energy (76.927 kcal/mol). According to the energy components of the binding free energies, van der Waals and non-polar solvation terms are the major favorable contributors to ligand binding. High values of these

energy components are associated with strong binding of ligands. Thus, due to the high values of binding free energy (-163.766 kcal/mol), van der Waals interactions (-63.283 kcal/mol) and solvation energy (102.362 kcal/mol), the binding of INT-777 was slightly stronger than INT-767. Overall, the binding free energies correlated well with the experimental activity.

Molecular dynamics-phase II

Stability analysis of $TGR5_{INT-777}$ and $TGR5_{INT-767}$

Most of the drugs required cell membrane permeation to reach their biological target and the process is crucial for rational drug discovery and development.[25] Drugs may cross through the phospholipid bilayer and through proteinaceous pores or transporters. Only these two main routes are possible for drugs to transport across biological membrane.[26] To understand this mechanism, protein-ligand complex dynamics was performed in lipid bilayer environment. In previous report, it was suggested that the small hydrophobic molecules are easily transported into the lipid bilayer.[27, 28] The agonists used in the study were hydrophobic in nature.[13] Based on the hydrophobic nature of the molecules, they enter easily into plasma membrane and binds with the active site of TGR5 receptors. After docking and MD simulations, the stability of two complexes was verified using backbone RMSD of cα atoms. The calculated RMSD of cα atoms of two complexes are shown in **Figure 7a**, indicates that both the complexes have a similar deviation throughout the simulation of their starting structure to 30 ns. The calculated RMSD reveals that the stability of both the complexes reached maximum equilibrium within 1 ns of MD simulation. The complexes did not show much deviation after the initial deviation before 1 ns. The RMSD curves show the minimum deviation between two complexes indicating their equilibration and stability. The calculated average of these complexes confirmed the above said statement. The average RMSD of $TGR5_{INT-777}$ was 0.21 nm whereas the average RMSD of $TGR5_{INT-767}$ complex was 0.21 nm.

From the generated RMSF plot using backbone cα atoms of $TGR5_{INT-777}$ and $TGR5_{INT-767}$, it was suggested that only very less flexibility observed for the residues which are responsible for the formation of α-helices. None of the fluctuations occurred beyond 0.8 nm but only one fluctuation (Arg204) was exceeded 0.6 nm over the MD simulation. It was observed that Arg204

has 0.36 nm and 0.76 nm fluctuations in $TGR5_{INT-777}$ and $TGR5_{INT-767}$, respectively. The increased fluctuations observed in Arg204 are attributed to the unstructured loops. RMSF analysis of interacting residues provides an importance of these residues on protein stability. The residues involved in the formation of hydrogen bonding interactions between TGR5 and INT-777 showed very small fluctuations, including Ser21 (0.08 nm), Asn76 (0.18 nm), Tyr89 (0.07 nm), Asn93 (0.07 nm), Ser157 (0.08 nm), Tyr240 (0.07 nm) and Ser270 (0.07 nm). Interacting residues of TGR5 with INT-767 also showed very small residual fluctuations such as Ser21 (0.17 nm), Tyr89 (0.08 nm), Trp237 (0.12 nm), Tyr240 (0.10 nm) and Ser270 (0.10 nm). The active site residues such as Ser21, Tyr89, Tyr240 and Ser270 were found to be conserved in both the complexes. From **Figure 7b**, it was observed that very small residual fluctuations of these conserved residues indicating their stable interactions.

Secondary structure profile analysis of complexed and uncomplexed systems

The secondary structural profile of the investigated two complexes and $TGR5_{apo}$ of the protein over the course of simulation were carried out using the DSSP analysis. Structural changes that contribute to stability differences were analyzed through visualization of plots. The generated plots in **Figure 8** represent the α-helix, turns, coils and other secondary structural elements of $TGR5_{apo}$ (a), $TGR5_{INT-777}$ (b) and $TGR5_{INT-767}$ (C). The major secondary structural modification was observed in the loop region comprised of 300 to 330 residues. Initially, turns are formed in the loop region and subsequently it formed bends at the end of the simulation. In particular, the amino acids from Val303 to Ser307 and Val325 to Asn330 formed bends. Further, the secondary structural analysis was carried out for two complexed systems. **Figure 8b** represents the significant structural changes in the region of residues from 150 to 170 of TGR5. These residues were located in the loop region underwent β-sheet conformation due to binding of agonist INT-777. The degradation of loops into β-sheet during the simulation period of 30 ns was observed through visual inspection. Likewise, in **Figure 8c**, the complexed system $TGR5_{INT-767}$ also shows significant structural changes between the residual positions of 150 to 190. A small region comprising residues from 328 to 330 was unstable over the simulation.

Apart from these loop regions, other α-helical conformations were found to be conserved and stable during the simulation period. The comparative results confirmed the stability of TGR5 in both bound and unbound states over the simulation period and the secondary structure calculations were correlated with the fluctuations of residues. Overall, secondary structure

profile results revealed that the conformational changes to turns, bends and β-sheets were occurred in the unstructured loop regions. Large fluctuations and changes in the loop regions were associated with the binding of agonists, thereby affecting the secondary structure elements of TGR5. Hence, the α-helical regions of TGR5 were mostly rigid and no major changes were observed. Undoubtfully, it was demonstrated that all three simulated systems, including $TGR5_{apo}$, $TGR5_{INT-777}$ and $TGR5_{INT-767}$ were more stable and only small amount of lose in secondary structure. Modifications of the protein conformation of three simulated systems were confirmed with the Rg and RMSD.

SASA

SASA was calculated for two complexed systems to measure the interaction between the protein-ligand complexes and solvents. Conformational changes of the protein upon binding of agonists were predicted using SASA. Binding affinity of ligand with a protein is associated with the solvation effects, including the desolvation of protein cavity, ligand and solvent molecules rearrangement.[29] The calculated average SASA value of the complex $TGR5_{INT-777}$ was 779 nm and complex $TGR5_{INT-767}$ was 825 nm. The comparison plot of SASA for two complexes is shown in **Figure 9a**. Relative SASA can predict the conformational changes of protein upon binding of ligands.[30] According to the results of SASA; it was observed that binding of agonists induced small conformational changes of protein.

Radiation of gyration

Structural changes and compactness of two protein-ligand complexes were evaluated using calculated Rg value and represented in **Figure 9b**. The Rg measures the mass of atoms that relative to the center of mass of the complex. $TGR5_{INT-767}$ complex exhibited wider fluctuation from 0 to 5 ns than $TGR5_{INT-777}$ complex and significantly decreased at the end of the simulation. The average Rg value of $TGR5_{INT-767}$ over the simulation period of 30 ns was 2.25 nm whereas for $TGR5_{INT-777}$, it was 2.20 nm. The structural changes and compactness of protein after binding of the two compounds were similar. Also, it was reported that movement of agonists within the binding site was associated with the compactness of protein.[31] The differences between the Rg values of two complexes were less than 0.04 nm. For both the complexes, the Rg values were stabilized around 5 ns.

H-bonding at the TGR5 active site

The existing hydrogen bonding interactions with TGR5 binding site residues with two agonists were evaluated throughout the simulation of 30 ns and plotted in **Figure 10**. The plot shows that the agonists formed stable hydrogen bonding interactions with the active site residues. The residues present in binding site such as Ser21, Asn76, Tyr89, Asn93, Ser157, Trp237, Tyr240 and Ser270 were found to be conserved after MD simulations of 30 ns. The predicted hydrophobic interactions in molecular docking study such as Ala17, Leu18, Leu24, Ala25, Leu68, Pro69, Pro72, Trp75, Pro92, Phe96, Leu97, Leu263 and Leu266 were also found to be conserved and stable throughout the simulations. Both the complexes were stabilized by 4-5 hydrogen bonding interactions. Further, there was no additional hydrogen bonding interactions formed over the simulation of 30 ns. The predicted stable hydrogen bonding and hydrophobic interactions are evident for the binding affinity of the protein and strengthen of the two agonists.

Essential dynamics

PCA was performed using g_covar to extract the large-scale concerted motions that are essential for the activity of the protein. ED was performed to generate a covariance matrix of 990 backbone cα atoms present in $TGR5_{apo}$, $TGR5_{INT\text{-}767}$ and $TGR5_{INT\text{-}777}$. About 3N-6 modes of possible internal fluctuations in a system of N atoms (six degrees of freedom) are required to determine the translation and external rotation of the system. PCA was performed for unbound and two bound complexes such as $TGR5_{apo}$, $TGR5_{INT\text{-}767}$ and $TGR5_{INT\text{-}777}$ with the covariance matrix of 20.36 nm^2, 16.18 nm^2 and 11.19 nm^2, respectively. Covariance matrix was constructed from the trajectories after removing translational and rotational movements. **Figure S3** illustrates the correlation map generated using gromacs inbuilt tool g_covar. The covariance matrix plot describes the correlation between all pairs of backbone cα atoms of protein. The plot represents the correlated (positive) and anticorrelated (negative) motions of backbone atoms as they move around their average position.

The covariance matrix of backbone cα atoms of protein was diagonalized to get the eigenvectors and eigenvalues. The calculated eigenvectors displayed as bars in **Figure 11** shows that first two eigenvectors of $TGR5_{INT\text{-}777}$ and $TGR5_{INT\text{-}767}$ accounts more than 80% of the collective motions of the protein backbone cα atoms. Eigenvector analysis was performed to determine the modes of fluctuation that contribute significantly to the overall motion of the protein. The dynamic behavior of the protein was confined with first two eigenvectors and fluctuation of protein was indicated by eigenvectors. The eigenvector with the highest eigenvalue

is considered the first principal component, whereas the eigenvector with the second highest eigenvalue is considered the second principal component. The direction of motion is represented by the eigenvectors and the amount of motion along with the eigenvectors is represented by eigenvalues.[32] The 2D plot of the first two eigenvectors (PC1 and PC2) of both the complexes and $TGR5_{apo}$ is represented in Figure 11. The amplitude of eigenvector along the multidimensional space is represented by the eigenvalues. Distribution of more dots in the plot represents the conformational changes of protein in unbound as well as bound state.

Biological implications

To the best of our knowledge**,** MD simulations of $TGR5_{INT-777}$ and $TGR5_{INT-767}$ complexes in the lipid environment have not been performed. Therefore, the study was performed to understand the underlying mechanism of the interaction between TGR5 receptor and its agonists in a lipid environment. The aim of the present study was to (1) develop the 3D-structure of human TGR5 using multiple threading method (2) predict the stability of modeled $TGR5_{apo}$ using MD simulation of 100 ns (3) identify active site of TGR5 through docking studies using experimentally proved agonists and (4) perform MD simulations of $TGR5_{INT-777}$ and $TGR5_{INT-767}$ complexes to identify the important residues responsible for hydrogen bonding, hydrophobic and π-π stacking interactions. The results obtained from this study concluded that both $TGR5_{apo}$ and $TGR5_{INT-777}$, $TGR5_{INT-767}$ complexes were stable throughout the MD simulation period. Also, the results indicate that both INT-777 and INT-767 were strongly interacted with human TGR5. Binding energy calculations also suggested that the residues involved in binding possess high affinity towards both agonists.

3. Materials and methods

Homology model development and validation

The amino acid sequence of human TGR5 protein, also known as G-protein coupled bile acid receptor 1 (GPBAR1) was downloaded from UniProtKB (Accession number: Q8TDU6). The employed sequence is composed of 330 amino acids. In order to find suitable templates for TGR5, BlastP search was performed against Protein Data Bank (PDB). From searching results, it was inspected that no suitable template was found to share more than 35% of sequence identity and they were not able to satisfy query coverage of 100%. Previous studies on GPCR have been reported that the improved quality and reliability of homology models can be achieved using multiple templates.[33, 34] Hence, I-Tasser[35], a multiple threading program was used to build the

model of TGR5 using multiple templates. Four variants of PPA methods, including PSI-BLAST profiles, hidden Markov model, Smith-Waterman and Needleman-Wunsch alignment algorithms were used by I-Tasser to search the possible folds through target threading against PDB structure library. Full-length model was produced by reassembling the threading templates (aligned regions) whereas the threading unaligned regions are built by *ab initio* modeling. Replica-exchange Monte-Carlo simulations algorithm were used for the refinement of predicted model. The quality and stereo-chemical property of the model was assessed by validation programs including ProCheck[36], Verify3D[37] and Errat[38]. Seven transmembrane helices, intracellular (ICLs) and extracellular loops (ECLs) were predicted using transmembrane prediction server TMHMM[39]. The quality of consistency between the native fold and sequence and the energy of residue-residue interactions were calculated using ProSA[40] server. The energy is transformed to a Z-score by using the formula:

$$Z_{S,C} = (E_{S,C} - \hat{E}_{S,C}) / \sigma_S$$

Where $E_{S,C}$ and $\hat{E}_{S,C}$ are residue-residue interaction energy and average residue-residue interaction energy of sequence S in conformation space C, respectively. σ_S is the associated standard deviation. Based on the various analyses, the best potential model suitable for further work was selected.

MD simulation-Phase I

System preparation

The constructed model was refined through molecular dynamics (MD) using Gromacs v4.6.3[41] package with GROMOS96 53A6 force field.[42] The POPC (1-palmitoyl-2oleoyl-sn-glycero-3-phosphatidylchlorine) bilayer with 512 molecules were generated by replication of the pre-equilibrated 128 POPC bilayer and corresponding force field parameters were obtained from the ABT site at http://compbio.biosci.uq.edu.au/atb/. Orientation of protein and membrane alignment was done using LAMBADA, available at http://code.google.com/p/lambada-align/. In LAMBADA, the protein was oriented with its hydrophobic belt (HB) parallel to the XY plane and translating the membrane in Z direction. The modeled protein of TGR5 was carefully embedded into a fully hydrated POPC lipid bilayer using InflateGRO2, available at http://code.google.com/p/inflategro2/.[43] In this process, the lipid bilayer was inflated and about 20 shrinking steps were performed to reach the experimental area per lipid value for pure POPC.

Simple point charge (SPC) model for water molecules and Cl^- counter ions were used throughout the simulations. The dimension box with the size of 12.973 X 12.973 X 10.770 nm was used to place up to 96636 water molecules in order to solvate the entire system. About 9 chloride counter ions were added to replace the water molecules and to neutralize the electrical charges of the system. The total number of atoms in the whole system consisting of TGR5, POPC, SOL and CL ions became 1, 24,914. A round of energy minimization using steepest descent with maximal force of 100 kJ/mol^{-1} nm^{-1} was executed to relax lipid molecules and to interact each molecule with another molecule in the system. The position of protein backbone atoms was restrained harmonically with an isotropic force constant of 100000 kJ/mol^{-1} nm^{-2} in order to avoid changes in the position of protein atoms during energy minimization.

MD simulation details

After energy minimization, the whole system was equilibrated with *isothermal-isochoric* (NVT) for 1 ns by applying the reference temperature of 300 K[23, 44, 45] that gradually increased from 100 K by stepwise reassignment of velocities at every 2 ps. Additional 2 ns simulation of *isothermal-isobaric* (NPT) was carried out with the reference pressure of 1 bar and coupling constant of 0.1 ps. Electrostatic interactions were calculated using Particle Mesh-Ewald summation scheme for long-range electrostatics. The LINCS constrain algorithm[46] was used to constrain all bond lengths and SETTLE algorithm[47] was used to constrain the geometry of all covalent bonds containing water molecules. After equilibration, 100 ns of MD simulation were carried out to obtain a relatively stable structure of TGR5 embedded in the membrane.

Preparation of protein

The modeled structure of TGR5 was subjected to Protein preparation wizard workflow implemented in Schrödinger, LLC, New York, NY, USA, 2014. Hydrogen atoms were added and water molecules that present 5Å away from the ligand were removed. The side chains that are not involved in the formation of salt bridges were neutralized. Bond orders, charges and atom types were assigned. Optimized Potentials for Liquid Simulations (OPLS)-2005 force field with an implicit solvation model was used for the energy minimization of modeled protein. Further, positions of hydroxyl, thiol hydrogen atoms, protonation state and tautomers of the His, Asn and Gln residues were selected using protein assignment script provided by Schrödinger. Restrained minimization was performed until the average root mean square deviation (RMSD) of heavy atoms reached 0.3Å.

Identification of potential binding sites

Stable structure of TGR5 obtained from MD simulation of 100 ns was used for molecular docking analysis. Before docking, SiteMap v3.0[48] module of Schrödinger was used to investigate the active site of the modeled TGR5. Identification of one or more possible binding site regions on the protein surface those are suitable for binding ligands were generated using SiteMap. Potential binding sites were identified using Site score, an important property generated by SiteMap. Among the five generated potential binding sites, a site with good SiteMap score was considered as the best binding site for docking studies.

Induced fit docking

The Induced Fit Docking (IFD)[49] protocol from Schrödinger suite was applied for molecular docking studies. Molecular docking was performed with two well-known agonists of TGR5 such as INT-777 and INT-767.[50] Two dimensional structures of two reported agonists were sketched using a chemical drawing interface ChemSketch v12.01 from ACD/Labs. Conversion of three dimensional structure and minimization was performed using LigPrep v2.9 module of Maestro (Schrödinger, LLC, New York, NY, 2014). The OPLS_2005 force field was used for the optimization of ligands. Initial glide docking was performed using a softened potential with van der Waals radii scaling of 0.7Å. Maximum of 20 poses per ligand were retained and used to sample the protein plasticity using Prime program in the Schrödinger suite. Residues with at least one atom located within 5.0Å of each corresponding 20 ligand poses were subjected to a conformational search and minimization. After minimization, the retained 20 ligand poses was redocked by Glide XP (extra precision) using default parameters. The Glide XP scoring function is calculated based on the ChemScore. Protein-ligand hydrogen bond quality based on geometric criteria can be evaluated using ChemScore. The Glide XP scoring function is presented in the following equation[51]:

$$\text{XP GlideScore} = E_{\text{coul}} + E_{\text{vdW}} + E_{\text{bind}} + E_{\text{penalty}} \quad (1)$$

$$E_{\text{bind}} = E_{\text{hyd\_enclosure}} + E_{\text{hb\_nn\_motif}} + E_{\text{hb\_cc\_motif}} + E_{\text{PI}} + E_{\text{hb\_pair}} + E_{\text{phobic\_pair}} \quad (2)$$

$$E_{\text{penalty}} = E_{\text{desolv}} + E_{\text{ligand\_strain}} \quad (3)$$

Where, $E_{\text{hyd\_enclosure}}$ is an improved model of hydrophobic interactions. $E_{\text{hb\_nn\_motif}}$ and $E_{\text{hb\_cc\_motif}}$ are special neutral-neutral hydrogen bond motifs and charged-charged hydrogen bond

motifs, respectively. E_{hb_pair} and E_{phobic_pair} are ChemScore-like hydrogen bond and lipophilic pair terms. E_{desolv} is a water scoring model that can be used to evaluate the basic physics of solvation within the confines of the protein-ligand complex active-site region. E_{ligand_strain} is penalizing strain energy in rigid-receptor docking considered to be a single most difficult component of an empirical scoring function. The IFD score used for final ranking of compounds is based on the following equation[52]: GlideScore + 0.05 x PrimeEnergy. The ligand interaction energy is calculated using GlideScore, whereas the total energy of the system was calculated using Prime energy. The predicted protein structures with large energy gaps can be eliminated using Prime energy.

Binding free energy calculations

Prime/MM-GBSA[53] module of Schrödinger suite was employed to calculate the binding free energy for a top ranked poses of two ligands. The OPLS_2005 force field and generalized-born/surface area (GBSA) continuum solvation model was used for calculating the energies of the complexes. The binding energy (ΔG_{bind}) was calculated using the following equation:

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{solv} + \Delta G_{SA}$$

$$\Delta E = E_{complex} - E_{protein} - E_{ligand}$$

Where, E_{MM} is the difference in energy between protein-ligand complex and the sum of the energies of apo protein and ligand. $E_{complex}$, $E_{protein}$, and E_{ligand} are the minimized energies of the protein–ligand complex, protein, and ligand, respectively. G_{solv} is the difference between GBSA solvation energy of protein-ligand complex and the sum of the corresponding energies for the protein and ligand. ΔG_{SA} is the difference between the surface area energy of protein-ligand complex and the sum of the corresponding energies for the protein and ligand.

MD simulation-Phase II

Total of two different protein-ligand complex systems, including $TGR5_{INT-777}$ and $TGR5_{INT-767}$ obtained from molecular docking analysis, were built for explicit MD simulation to confirm the stability and binding mode. The topology files and force field parameters for two ligands (INT-777 and INT-767) were generated using PRODRG web server.[54] After topology determination for two complexes, LAMBADA was used for the orientation and protein-membrane alignment. Each complex (protein and ligand) was inserted into the membrane carefully using Gromacs inbuilt tool g_membed.[45] After insertion of complex into the membrane, the same protocol mentioned under the section of molecular dynamics-phase I was followed for

the remaining steps. Each complex was fully equilibrated with NVT and NPT were further used for the MD simulation of 30 ns.

Essential dynamics

All the trajectories obtained after MD simulations of 100 ns ($TGR5_{apo}$) and 30 ns ($TGR5_{INT-777}$ and $TGR5_{INT-767}$) were further analyzed by Essential dynamics (ED). In ED, the eigenvectors, eigenvalues and their projection along with the first two principal components were calculated. Principal Component Analysis (PCA) was performed with g_covar and g_anaeig. The generation of non-mass weighted covariance matrix is the first step in ED. For an N atom of system, 3N × 3N covariance matrix is denoted as *A*. N is the number of atoms in the system. The covariance matrix of A, denoted as *C* is defined using the following equation:

$$C = A^T A \qquad (1)$$

Where, T is the transpose of the covariance matrix that obtained by exchanging the rows and columns of the matrix. The eigenvectors of the covariance matrix represent the principle components. It appears into a simple eigenvalue problem:

$$Cx = \lambda x \qquad (2)$$

Where, λ is the eigenvalue associated with the eigenvector x. About 3N eigenvectors and associated eigenvalues are present in an N atom of system. This equation is simplified using the following:

$$(C\text{-}\lambda I)x = 0 \qquad (3)$$

$$D = U^{-1}CU \qquad (4)$$

Where, I is the identity matrix and D is the diagonal matrix of the covariance matrix. The equation (3) can be obtained by diagonalizing the covariance matrix (D). In equation (4), the matrix U contains the eigenvectors and D is the matrix of the corresponding eigenvalues.

Analysis

Atomic interactions and protein trajectories were analyzed using Gromacs inbulit tools such as g_rms, g_rmsf, g_hbond, g_gyrate, g_energy, g_sas and do_dssp. All the trajectories were analyzed graphically using OriginPro software.

4. Conclusion

In the present study, TGR5 model has been developed to understand the stable conformation of protein due to the absence of crystal structure. Validation results revealed that the constructed model was much more reliable and suitable to calculate the structural

conformation and conformational changes upon binding of ligands. Binding mode of two well-known agonists was carried out to understand the binding properties and mechanism of action of interactions. From the combined results of docking and free energy calculations, it was found that the active site residues of TGR5 strongly interacted with two potent agonists. Further, conformational properties of three models including $TGR5_{apo}$, $TGR5_{INT-767}$ and $TGR5_{INT-777}$ were analyzed by employing MD simulation. The prediction of binding site of TGR5 could help in the discovery and designing of different novel potent agonists. The present study may provide valuable benchmark in structural modeling and binding site prediction of TGR5 receptor.

Acknowledgement

Authors are grateful to thank the Department of Bioinformatics, Alagappa University, Karaikudi, India for the laboratory facilities provided for this work.

References

1. J. C. Li, X. F. Shen and X. L. Meng, A traditional Chinese medicine JiuHuangLian (Rhizoma coptidis steamed with rice wine) reduces oxidative stress injury in type 2 diabetic rats, *Food. Chem. Toxicol.*, 2013, 59, 222-229.

2. T. Kim, W. Jee, K. H. Jeong, J. H. Song, S. Park, P. Choi, S. Kim, S. Lee and J. Ham, Total synthesis and dual PPARα/γ agonist effects of Amorphastilbol and its synthetic derivatives, *Bioorg. Med. Chem. Lett.*, 2012, 22, 4122-4126.

3. S. Sancheti, S. Sancheti and S. Seo, Antidiabetic and antiacetylcholinesterase effects of ethyl acetate fraction of *Chanomeles sinensis* (Thouin) Koehne fruits in streptozotocin-induced diabetic rats, *Exp. Toxicol. Pathol.*, 2013, 65, 55-60.

4. IDF Diabetes Atlas (2009). Internatioanl Diabetes Federation. 4th edition.

5. J. Zhu, M. Ning, C. Guo, L. Zhang, G. Pan, Y. Leng and J. Shen, Design, synthesis and biological evaluation of a novel class of potent TGR5 agonists based on a 4-phenyl pyridine scaffold, *Eur. J. Med. Chem.*, 2013,69,55-68.

6. R. J. Hodge, J. Lin, L. S. V. Johnson, E. P. Gould, G. P. Bowers, D. J. Nunez and on behalf of the SB-756050 project team, Safety, pharmacokinetics and pharmacodynamic effects of

a selective TGR5 agonist, SB-756050, in type 2 diabetes, *Clinical Pharmacology in Drug Development.*, 2013, 2, 213-222.

7. M. R. Herbert , D. L. Siegel, L. Staszewski, C. Cayanan, U. Banerjee, S. Dhamija, J. Anderson, A. Fan, L. Wang, P. Rix, A. K. Shiau, T. S. Rao, S. A. Noble, R. A. Heyman, E. Bischoff, M. Guha, A. Kabakibi and A. B. Pinkerton, Synthesis and SAR of 2-aryl-3-aminomethylquinolines as agonists of the bile acid receptor TGR5, *Bioorg. Med. Chem. Lett.*, 2010, 20, 5718-5721.

8. H. Duan, M. Ning, X. Chen, Q. Zou, L. Zhang, Y. Feng, L. Zhang, Y. Leng and J. Shen, Design, synthesis, and antidiabetic activity of 4- phenoxynicotinamide and 4-phenoxypyrimidine-5-carboxamide derivatives as potent and orally efficacious TGR5 agonists, *J. Med. Chem.*, 2012, 55, 10475-10489.

9. H. Duboc, Y. Tache and A. F. Hofmann, The bile acid TGR5 membrane receptor: from basic research to clinical application, *Dig. Liver. Dis.*, 2014, 46, 302-312.

10. M. Baptissart, A. Vega, S. Maqdasy, F. Caira, S. Baron, J. M. Lobaccaro and D. H. Volle, Bile acids: from digestion to cancers, *Biochimie.*, 2013, 95, 504-517.

11. G. Rizzo, D. Passeri D, F. De Franco, G. Ciaccioli, L. Donadio, G. Rizzo, S. Orlandi, B. Sadeghpour, X. X. Wang, T. Jiang, M. Levi, M. Pruzanski and L. Adorini, Functional characterization of the semisynthetic bile acid derivative INT-767, a dual farnesoid X receptor and TGR5 agonist, *Mol. Pharmacol.*, 2010, 78, 617-630.

12. A. Roda, R. Pellicciari, A. Gioiello, F. Neri, C. Camborata, D. Passeri, F. D. Franco, S. Spinozzi, C. Colliva, L. Adorini, M. Montagnani and R. Aldini, Semisynthetic bile acid FXR and TGR5 agonists: Physicochemical properties, pharmacokinetics, and metabolism in the rat, *J. Pharmacol. Exp. Ther.*, 2014, 350, 56-68.

13. R. Pellicciari, A. Gioiello, P. Sabbatini, F. Venturoni, R. Nuti, C. Colliva, G. Rizzo, L. Adorini, M. Pruzanski, A. Roda, A. Macchiarulo, Avicholic Acid: A Lead Compound from Birds on the Route to Potent TGR5 Modulators, *ACS. Med. Chem. Lett.*, 2012, 3, 273-277.

14. R. H. McMahan, X. X. Wang, L. L. Cheng, T. Krisko, M. Smith, K. El Kasmi, M. Pruzanski, L. Adorini, L. Golden-Mason, M. Levi, H. R. Rosen, Bile acid receptor activation modulates hepatic monocyte activity and improves nonalcoholic fatty liver disease, *J. Biol. Chem.*, 2013, 288, 11761-11770.

15. B. K. Yap, M. J. Buckle, S. W. Doughty, Homology modeling of the human 5-HT1A, 5-HT 2A, D1, and D2 receptors: model refinement with molecular dynamics simulations and docking evaluation, *J. Mol. Model.*, 2012, 18, 3639-3636.

16. C. Andreini, L. Banci, I. Bertini, C. Luchinat, A. Rosato, Bioinformatic comparison of structures and homology-models of matrixmetalloproteinases, *J. Proteome. Res.*, 2004, 3, 21-31.

17. B. K. Kobilka, G protein coupled receptor structure and activation, *Biochim. Biophys. Acta.*, 2007, 1768, 794-807.

18. K. Speranskiy, M. Cascio, M. Kurnikova, Homology modeling and molecular dynamics simulations of the glycine receptor ligand binding domain, *Proteins.*, 2007, 67, 950-60.

19. A. Cordomi, G. Caltabiano and L. Pardo, Membrane protein simulations using AMBER force field and Berger lipid parameters, *J. Chem. Theory. Comput.*, 2012, 8, 948-958.

20. J. U. Bowie, Solving the membrane protein folding problem, *Nature.*, 2005, 438, 581-589.

21. J. Wang and T. Hou, Develop and test a solvent accessible surface area-based model in conformational entropy calculations, *J. Chem. Inf. Model.*, 2012, 52, 1199-1212.

22. J. S. Dixon, J. M. Blaney, Docking: predicting the structure and binding affinity of ligand-receptor complexes. In: Y. C. Martin, P. Willet (Eds.), Designing bioactive molecules: Three dimensional techniques and applications. American Chemical Society, Washington DC, 1998, pp. 175-198.

23. F. Collu, M. Ceccarelli, P. Ruggerone, Exploring the binding properties of agonists interacting with a δ-opioid receptor, *PLoS. One*., 2012, 7, e52633

24. D. Sabbadin, A. Ciancetta, S. Moro, Bridging molecular docking to membrane molecular dynamics to investigate GPCR-ligand recognition: The human A2A adenosine receptor as a key study, *J. Chem. Inf. Model.*, 2014, 54, 169-183.

25. D. Bemporad, C. Luttmann, J. W. Essex, Computer simulation of small molecule permeation across a lipid bilayer: Dependence on bilayer properties and solute volume, size, and cross-sectional area, *Biophys. J.*, 2004, 87, 1-13.

26. D. B. Kell, What would be the observable consequences if phospholipid bilayer diffusion of drugs into cells is negligible?, *Trends Pharmacol Sci.,* 2014, S0165-6147, 00170-00179.

27. G. M. Cooper, The Cell: A Molecular Approach.2nd edition. Sunderland (MA): Sinauer Associates, 2000.

28. H. Lodish, A. Berk, S. L. Zipursky, Molecular Cell Biology, 4th edition, New York: W. H. Freeman, 2000.

29. S. Nunez, J. Venhorst and C. G. Kruse, Assessment of a novel scoring method based on solvent accessible surface area descriptors, *J. Chem. Inf .Model.*, 2010, 50, 480-486.

30. J. Marsh, S. A. Teichmann, Relative solvent accessible surface area predicts protein conformational changes upon binding, *Structure.*, 2011, 19, 859-867.

31. W. I. Tou, S. S. Chang, C. C. Lee and C. Y. Chen, Drug design for neuropathic pain regulation from traditional Chinese medicine, *Sci. Rep.*, 2013, 3:844. doi: 10.1038/srep00844.

32. S. Kundu, D. Roy, Temperature-induced unfolding pathway of a type III antifreeze protein: Insight from molecular dynamics simulation, *J. Mol. Graph. Model.*, 2008, 27, 88-94.

33. T. Yarnitzky, A. Levit and M. Y. Niv, Homology modeling of G-protein-coupled receptors with X ray structures on the rise, *Curr. Opin. Drug. Discov. Devel.*, 2010, 13, 317-325.

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

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

36. R.A. Laskowski, M.W. MacArthur, D. S. Moss and J. M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Cryst.*, 1993, 26, 283-291.

37. C. Colovos and T. O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, *Protein. Sci.*, 1993, 2, 1511-1519.

38. R. Luthy, J. U. Bowie and D. Eisenberg, Assessment of protein models with three-dimensional profiles, *Nature.*, 1992, 356, 83-85.

39. A. Krogh, B. Larsson, G. von Heijne and E. L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, *J. Mol. Biol.*, 2001, 305, 567-580.

40. M. J. Sippl, Recognition of errors in three-dimensional structures of proteins. *Proteins. Struct. Funct. Bioinf.*, 1993, 17, 355-362.

41. B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, GROMACS 4: Algorithms for Highly Efficient, Load- Balanced, and Scalable Molecular Simulation, *J. Chem. Theory Comput.*, 2008, 4, 435-447.

42. C. Oostenbrink, T. Soares, N. A. Vegt and W. Gunsteren, Validation of the 53A6 GROMOS force field, *Eur. Biophys. J.*, 2005, 34, 273-284.

43. T. H. Schmidt and C. Kandt, LAMBADA and InflateGRO2: efficient membrane alignment and insertion of membrane proteins for molecular dynamics simulations, *J. Chem. Inf. Model.*, 2012, 52, 2657-2669.

44. G. Bussi, T. Zykova-Timan, M. Parrinello, Isothermal-isobaric molecular dynamics using stochastic velocity rescaling, *J. Chem. Phys.,* 2009, 130, 074101.

45. M. G. Wolf, M. Hoefling, C. Aponte-Santamaria, H. Grubmuller and G. Groenhof, g_membed: Efficient insertion of a membrane protein into an equilibrated lipid bilayer with minimal perturbation, *J. Comput. Chem.*, 2010, 31, 2169-2174.

46. B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, LINCS: A linear constraint solver for molecular simulations, *J. Comput. Chem*., 1997, 18, 1463-1472.

47. S. Miyamoto and P. A. Kollman, Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models, *J. Comput. Chem*., 1992, 13, 952-962.

48. T. Halgren, Identifying and characterizing binding sites and assessing druggability, *J. Chem. Inf. Model.,* 2009, 49, 377-389.

49. R. Farid, T. Day, R. A. Friesner and R. A. Pearlstein, New insights about HERG blockade obtained from protein modeling, potential energy mapping, and docking studies, *Bioorg. & Med. Chem.*, 2006, 14, 3160-3173.

50. A. Baghdasaryan, T. Claudel, J. Gumhold, D. Silbert, L. Adorini, A. Roda, S. Vecchiotti, F. J. Gonzalez, K. Schoonjans, M. Strazzabosco, P. Fickert and M. Trauner, Dual farnesoid X receptor/TGR5 agonist INT-767 reduces liver injury in the Mdr2-/- (Abcb4-/-) mouse cholangiopathy model by promoting biliary HCO^-_3 output, *Hepatology.*, 2011, 54,1303-1312.

51. R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, D. T. Mainz, Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, *J. Med. Chem*., 2006, 49, 6177-6196.

52. W. Sherman, T. Day, M. P. Jacobson, R. A. Friesner, R. Farid, Novel procedure for modeling ligand/receptor induced fit effects, *J. Med. Chem*., 2006, 49, 534-553.

53. *Prime, version 3.5*, Schrödinger, LLC, New York, NY, 2014.

54. A. W. Schuttelkopf and D. M. F. van Aalten, PRODRG: a tool for high-throughput crystallography of protein-ligand complexes, *Acta Crystallogr. Sect. D: Biol. Crystallogr.*, 2004, 60, 1355-1363.

Figure captions

Figure 1: Cartoon representation of modeled structure of TGR5 with seven transmembrane helices.

Figure 2: Structural modification of $TGR5_{apo}$ embedded in POPC bilayer during the simulation time of 0 ns (a) and 100 ns (b). Changes in the loop regions are highlighted in magenta dashed circles.

Figure 3: RMSD of backbone atoms of $TGR5_{apo}$ over the MD simulation period of 100 ns.

Figure 4: The RMSF analysis of TGR5 residues during the MD simulation period of 100 ns.

Figure 5: Time evolution of calculated solvent accessible area (a) and Radius of gyration (b) of TGR5 during the whole simulation of 100 ns.

Figure 6: Binding modes and docking interactions of INT-777 (a) and INT-767 (b) with TGR5. Hydrogen bond interactions formed between agonists and active site residues of TGR5 are shown in purple dashed lines.

Figure 7: RMSD from starting structure evaluated for the backbone cα atoms (a) and Residual fluctuations of each residue (b) during the MD simulation of 30 ns for $TGR5_{INT\text{-}777}$ (Orange) and $TGR5_{INT\text{-}767}$ (Purple) complexes.

Figure 8: Secondary structural changes of $TGR5_{apo}$ (a), $TGR5_{INT\text{-}777}$ (b) and $TGR5_{INT\text{-}767}$ (c) observed using DSSP during the MD simulations.

Figure 9: The calculated solvent accessible surface area (a) and Radius of gyration (b) for $TGR5_{INT\text{-}777}$ and $TGR5_{INT\text{-}767}$ depicted for entire simulation of 30 ns.

Figure 10: Time evolution of the number of intermolecular hydrogen bonds formed between active site residues of TGR5 with INT-777 (dark cyan) and INT-767 (purple).

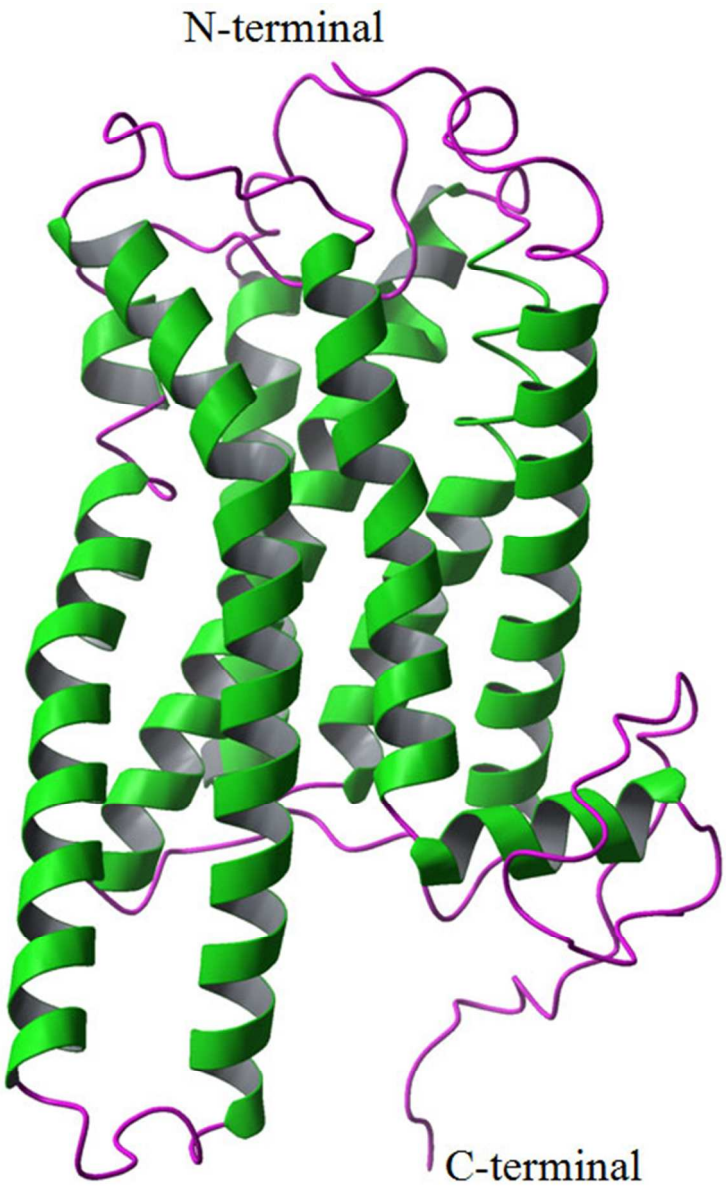
Figure 11: Eigenvector index versus eigenvalues for the first ten eigenvectors of two protein-ligand complexes (a) and projection of the motion of the $TGR5_{apo}$ (b), $TGR5_{INT\text{-}777}$ (c) and $TGR5_{INT\text{-}767}$ (d) in phase space along with first two (PC1 and PC2) principal eigenvectors.

Supplementary figures

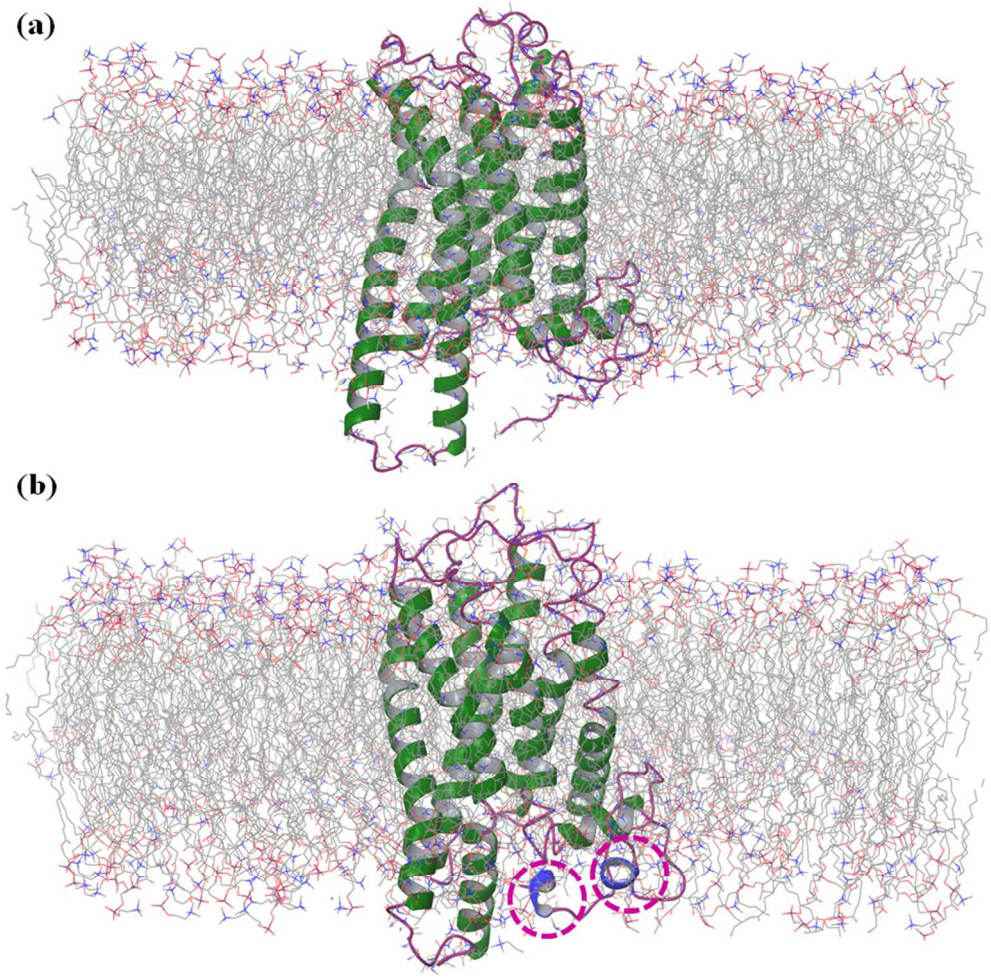
S1: Residue profile of the generated model TGR5 in Ramachandran plot.

S2: Energy plot (a) and Quality index (b) of the TGR5 model generated by ProSA server. Z-score represents the index value.

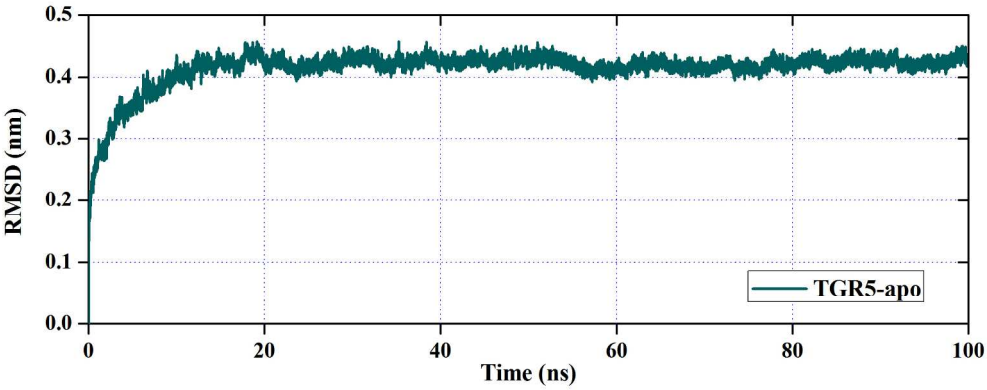
S3: Correlation maps of $TGR5_{INT\text{-}777}$ (a) and $TGR5_{INT\text{-}767}$ (b) and $TGR5_{apo}$ (c) generated using g_covar. Positive (red) and negative (blue) motions between atom pairs are depicted.



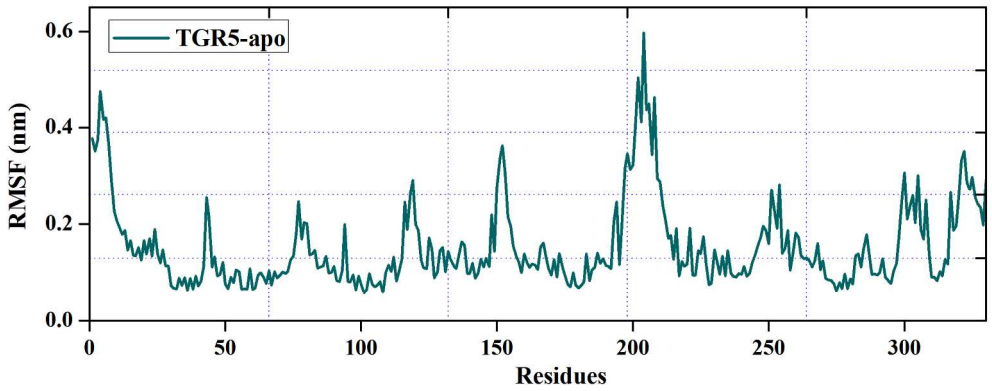
40x65mm (300 x 300 DPI)



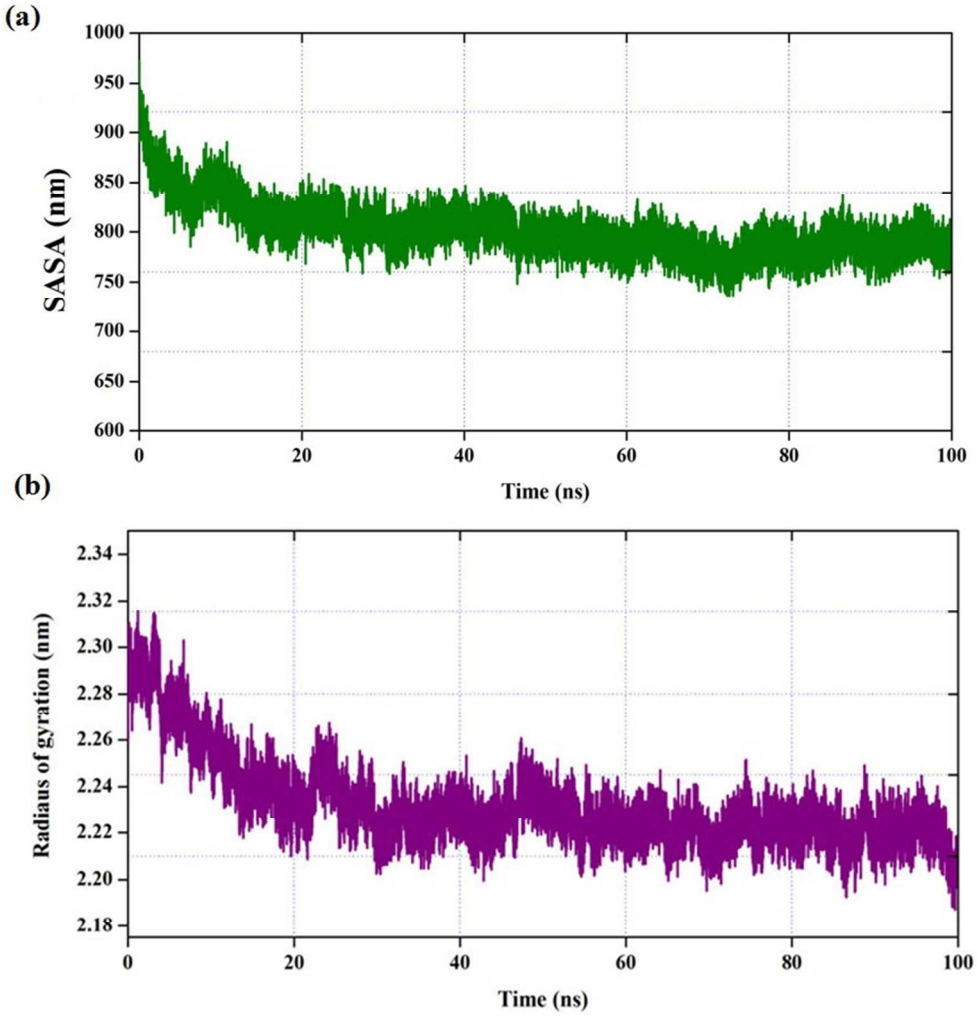
82x80mm (300 x 300 DPI)



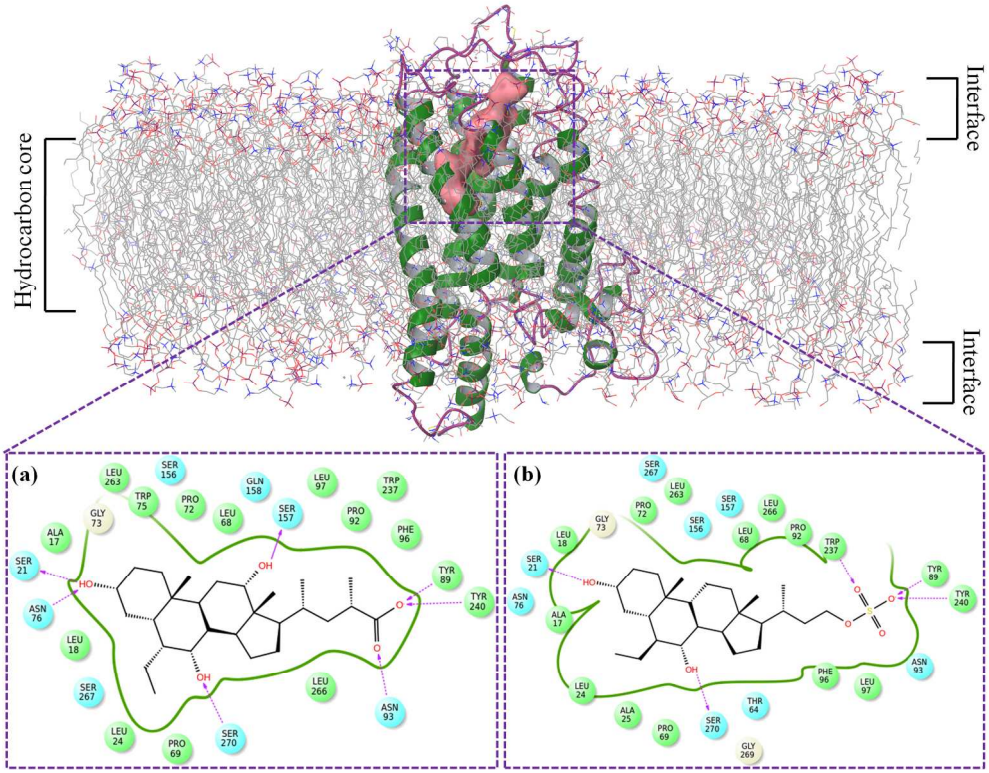
217x86mm (300 x 300 DPI)



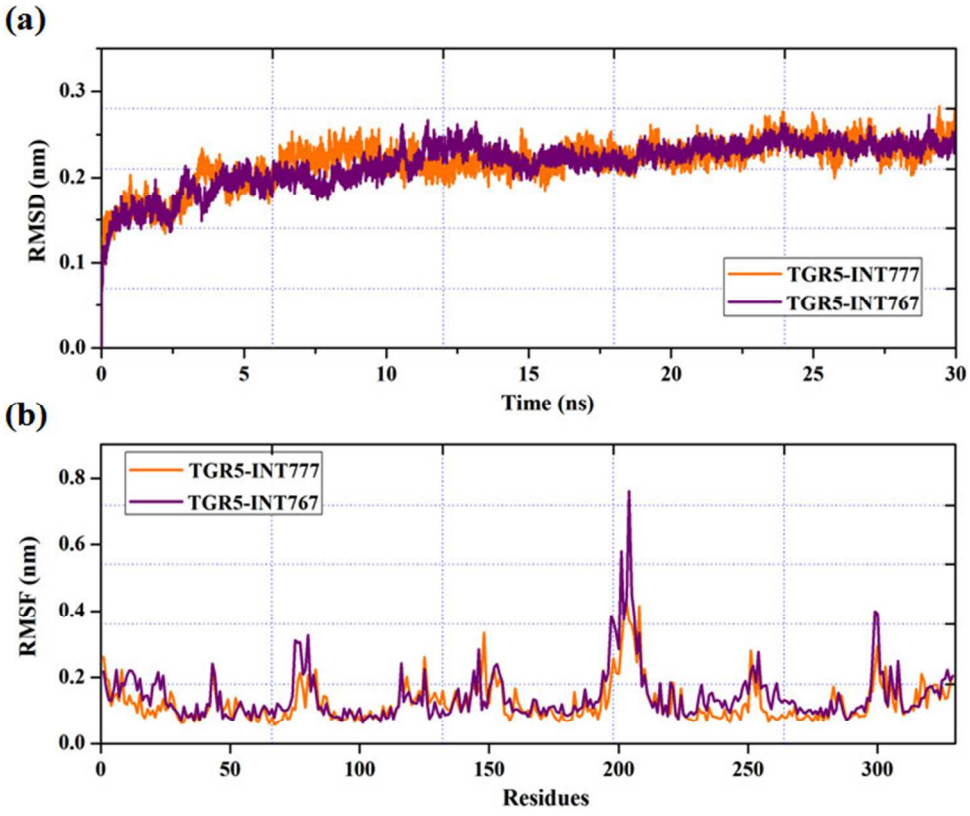
213x86mm (300 x 300 DPI)



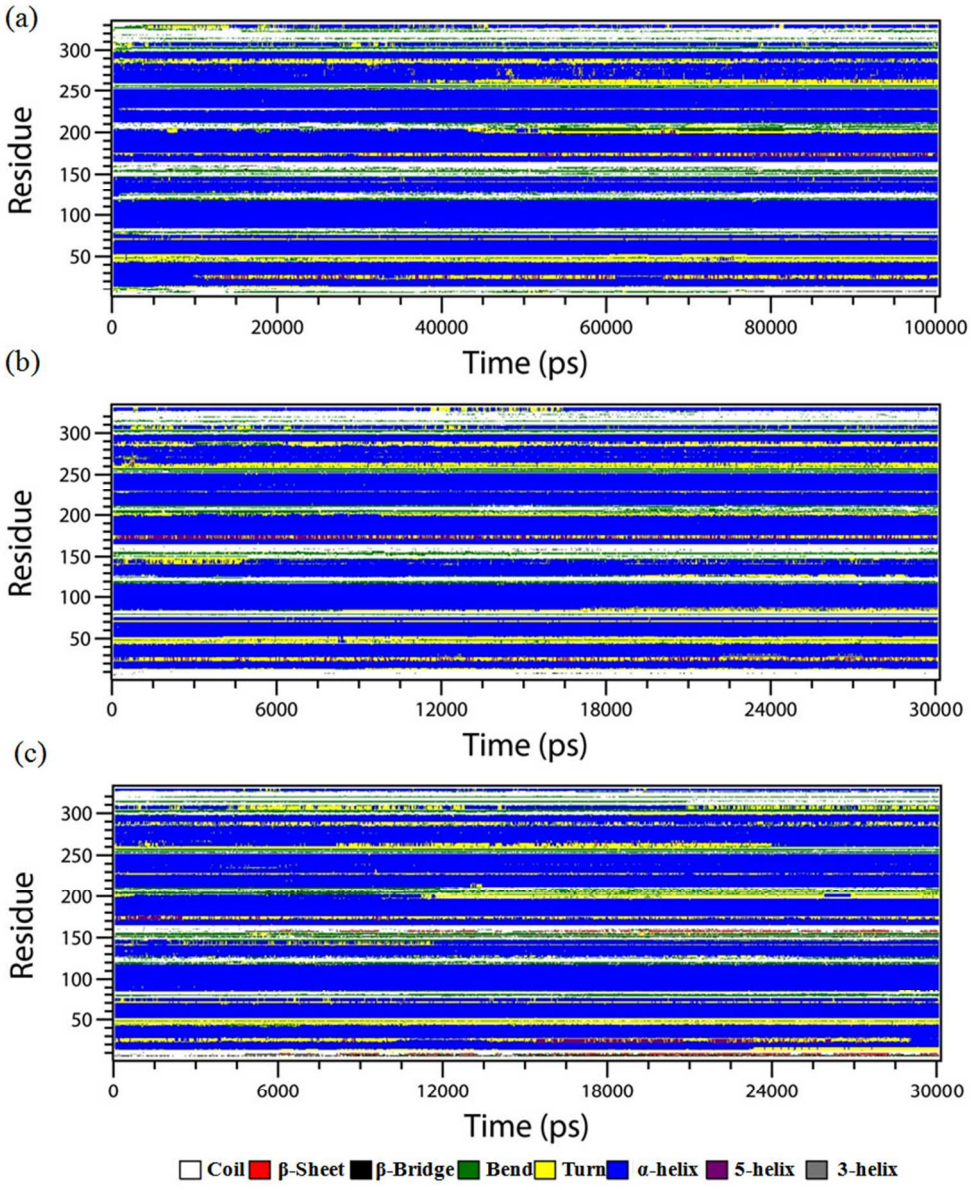
75x78mm (300 x 300 DPI)



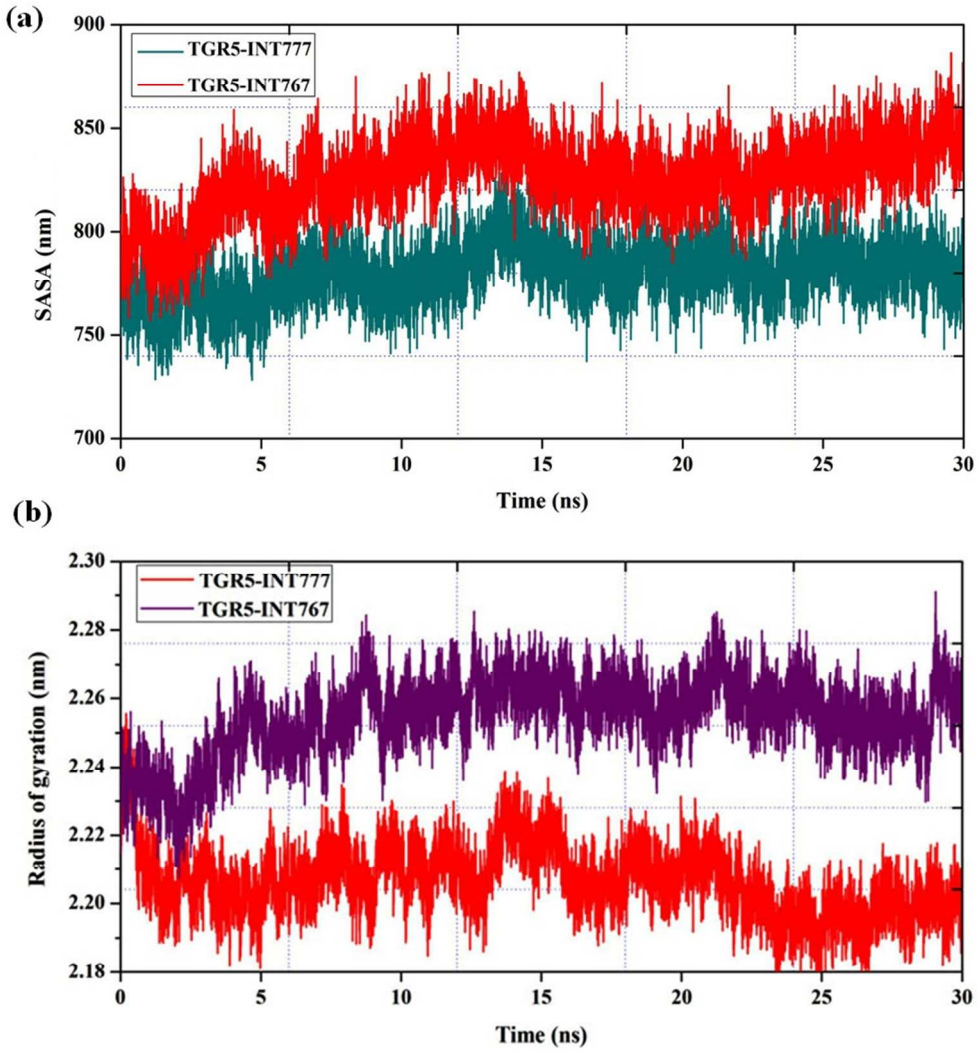
177x136mm (300 x 300 DPI)



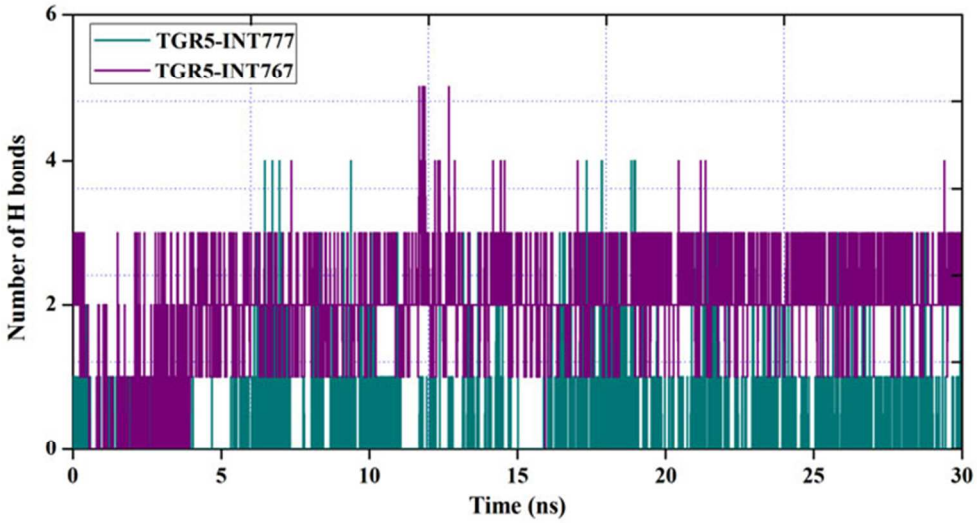
65x54mm (300 x 300 DPI)



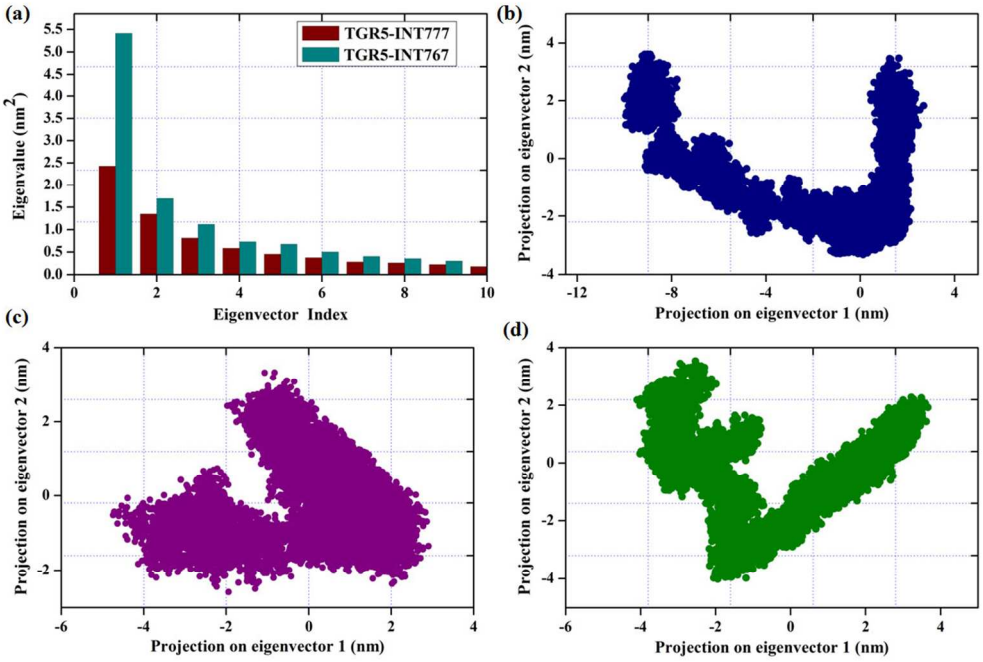
75x91mm (300 x 300 DPI)



84x91mm (300 x 300 DPI)



62x34mm (300 x 300 DPI)



107x72mm (300 x 300 DPI)