J Mol Model (2011) 17:1565–1577
DOI 10.1007/s00894-010-0860-z

ORIGINAL PAPER

Multiple templates-based homology modeling enhances structure quality of AT1 receptor: validation by molecular dynamics and antagonist docking

Pandian Sokkar · Shylajanaciyar Mohandass · Murugesan Ramachandran

Received: 5 June 2010 / Accepted: 24 September 2010 / Published online: 6 October 2010

Abstract We present a comparative account on 3D-structures of human type-1 receptor (AT1) for angiotensin II (AngII), modeled using three different methodologies. AngII activates a wide spectrum of signaling responses via the AT1 receptor that mediates physiological control of blood pressure and diverse pathological actions in cardiovascular, renal, and other cell types. Availability of 3D-model of AT1 receptor would significantly enhance the development of new drugs for cardiovascular diseases. However, templates of AT1 receptor with low sequence similarity increase the complexity in straightforward homology modeling, and hence there is a need to evaluate different modeling methodologies in order to use the models for sensitive applications such as rational drug design. Three models were generated for AT1 receptor by, (1) homology modeling with bovine rhodopsin as template, (2) homology modeling with multiple templates and (3) threading using I-TASSER web server. Molecular dynamics (MD) simulation (15 ns) of models in explicit membrane-water system, Ramachandran plot analysis and molecular docking with antagonists led to the conclusion that multiple template-based homology modeling outweighs other methodologies for AT1 modeling.

Keywords AT1 receptor · Explicit membrane · Homology modeling · I-TASSER · Molecular docking · Molecular dynamics · Multiple templates

P. Sokkar · M. Ramachandran
School of Chemistry, Madurai Kamaraj University,
Madurai 625 021 Tamil Nadu, India

S. Mohandass · M. Ramachandran (✉)
School of Biological Sciences, Madurai Kamaraj University,
Madurai 625 021 Tamil Nadu, India
e-mail: rammurugesan@yahoo.com

Introduction

The renin-angiotensin system (RAS) is an enzymatic cascade initiated by the action of renin. Renin acts upon angiotensinogen to generate inactive decapeptide, angiotensin I (AngI). AngI in turn is hydrolyzed by angiotensin-converting enzyme (ACE) to produce potent pressor octapeptide, AngII [1]. AngII binds to two distinct receptors, AT1 and AT2 receptors, belonging to the G-protein-coupled receptor (GPCR) superfamily that has seven transmembrane spanning helices [2]. Binding of AngII to AT1 mediates several events that lead to hypertension, congestive heart failure and chronic renal failure; hence AT1 serves as a potential therapeutic target [1]. In general, the GPCRs are partially active in their native state. Binding of agonists to the native state of a GPCR causes activation of the receptor, whereas binding of antagonists or inverse agonists causes transition of native state to inactive state [3, 4]. Hence native state of a GPCR holds importance in antagonist designing. The superfamily of GPCRs represents the majority of current drug targets. However, structural information of GPCRs is limited due to experimental difficulties in determining their structures.

Despite the development in protein structure determination techniques (NMR and X-ray crystallography), the gap between available sequences and structures is ever increasing. Computer aided protein modeling techniques have been developed to bridge this gap [5]. Homology modeling is a widely used computational approach to produce comparatively high-resolution models. It makes use of the fact that evolutionary related proteins share a similar structure [5]. Models of a protein with unknown structure (target) can be built based on an alignment of a protein of known structure (template).

Homology modeling has been extensively adopted for modeling GPCRs [6, 7]. The structure of AT1 receptor has

not yet been determined by experimental techniques. Molecular cloning and site-directed mutagenesis experiments on AT1 receptor have given insight into the topological aspects of AT1 receptor [8, 9]. It belongs to the rhodopsin/β-adrenergic receptor subfamily of GPCRs with seven hydrophobic regions (each corresponding to a transmembrane helix) consisting of 21 amino acids [9]. Several attempts have been made to model the human AT1 receptor structure using mainly rhodopsin and β-adrenergic receptor 3D structures as templates [10–13]. These models tend to be less accurate since the sequence similarity between AT1 and its templates is less than 30%. Molecular dynamics (MD) simulations with explicit solvent treatment have been most frequently employed to assess the stability and to refine low quality models. Baleanu-Gogonea and Karnick modeled whole rat AT1 receptor and validated with 1 ns MD simulation [14]. However, models resulting from low sequence homology inevitably contain errors in secondary structure definition and packing of secondary structure elements. During MD refinement of such models, the errors in the packing normally mount up the interatomic forces. This in turn leads to an initial distortion in the structure for several nanoseconds. After the initial distortion, repacking of secondary structure elements takes place, suggesting the need to perform longer MD simulation [15]. According to the reports by Mobarec et al., use of multiple templates in modeling GPCR proteins slightly increases the model quality, in terms of higher predictive power in flexible ligand-rigid protein docking experiments [16]. Further, the improvement in model quality by the use of multiple templates has been reported in CASP experiments [17]. Hence incorporation of multiple templates and increasing the simulation time would be a good strategy to model AT1 receptor.

In this study we have investigated models of AT1 receptor calculated by three comparative modeling techniques namely, (i) simple homology modeling with bovine rhodopsin as template, (ii) multi-template homology modeling and (iii) threading with I-TASSER server [18]. We present here the results of 15 ns MD simulations of AT1 receptor embedded in a hydrated membrane environment for all the three models. Similarities and differences between the models observed in simulations and docking of seven antihypertensive drugs are discussed.

Methodology

Template selection

Human AT1 receptor sequence (P30556), obtained from NCBI protein database [19], was used for all modeling calculations reported in this paper. In order to find suitable templates for the comparative modeling of AT1 receptor, a sequence-similarity search was done with query AT1 sequence against protein data bank [20] (PDB: http://www.pdb.org/) using BLASTp program [21] available on the NCBI website (http://www.ncbi.nlm.nih.gov/), using default threshold E-value of 10 and inclusion threshold value of 0.005. Templates were selected based on the sequence similarity, structural resolution and overall fold.

Homology modeling of AT1 receptor

Bovine rhodopsin-based modeling

The human AT1 sequence was aligned with bovine rhodopsin crystal structure (1GZM) by the use of the ALIGN2D command of the MODELLER 9v6 program [22]. This command implements a global dynamic programming method for comparison of two sequences, but also relies on a variable gap penalty function, favoring gaps in structurally reasonable positions (such as, solvent exposed and outside secondary structure segments). Such a sequence-structure alignment procedure is highly accurate for low target-template sequence identity cases where gaps are very common. Based on the alignment information, the 3D model was built by satisfaction of spatial restraints using MODELLER program. This method employs extraction of spatial restraints from two sources (homology-derived and CHARMM22 force field [23] -derived), followed by optimization with conjugate gradients and molecular dynamics to minimize the violations of spatial restraints. Secondary structure restraints were applied to the TM helices during model building. The loop regions in the model were optimized with the loop optimization protocol of MODELLER.

Multiple templates-based modeling

Bovine rhodopsin (1GZM), squid rhodopsin (2Z73), bovine rhodopsin (stabilizing mutant, 3C9M), human β2-adrenergic receptor (2RH1) and turkey β1-adrenergic receptor (2VT4) were selected as templates for the modeling of AT1. The use of several templates generally increases the model accuracy as it combines information from multiple template structures. The utility of multiple templates for comparative modeling relies on the accuracy of their multiple structure alignment. SALIGN [24] (as implemented in MODELLER) was employed to construct multiple structure alignments of templates. SALIGN creates pairwise alignment by dynamic programming optimization using a scoring function that is dependent of the sequence and structure features. These features include amino acid residue type, residue position, residue accessible surface area, residue secondary structure state and the conformation

of a short segment centered on the residue. Multiple alignments are then constructed by assembling the individual pairwise alignments. A detailed description of SALIGN procedure is described elsewhere [24]. The target sequence was then aligned with this multiple structure-based alignment. The model building from the target-template alignment was done similarly.

Threading of AT1 receptor

AT1 structure was modeled using I-TASSER server, which is a protein structure modeling approach based on the secondary-structure enhanced profile-profile threading alignment (PPA) and the iterative implementation of the Threading ASSEmbly Refinement (TASSER) program. In this approach, the target sequence is first threaded through a PDB structure library to search for the possible folds by four simple variants of PPA methods employing the hidden Markov model, PSI-BLAST profiles, Needleman-Wunsch and Smith-Waterman alignment algorithms. The threading aligned regions are used to reassemble full-length models while the threading unaligned regions are built by ab initio modeling. The model is refined by replica-exchange Monte Carlo simulations [18].

Disulfide linkages

AT1 receptor has been reported to contain two disulfide linkages (CYS18-CYS275 and CYS101-CYS180) [25]. These two linkages were introduced in the models by the use of the AMBER-8 package [26]. Energy minimization was carried out using AMBER03 force field [26] to minimize any geometrical violation.

Model comparison and assessment

TM-score program was used for comparative evaluation of the modeled 3D-structures [27]. TM-score is a structure comparison algorithm, and it exploits a variation of Levitt–Gerstein (LG) weight factor [28] that weights the residue pairs at smaller distances relatively stronger than those at larger distances. Therefore, the TM-score is more sensitive to the global topology than to the local structural variations. The quality of the modeled structure was assessed by Ramachandran plot occupancy of residues using PROCHECK server [29].

Preparation of simulation systems

Visual molecular dynamics (VMD) program was used to prepare the systems [30]. The AT1 structure was embedded in the center of a 75×75 Å palmitoyl oleoyl phosphatidylcholine (POPC) lipid bilayer. Rectangular TIP3P water [31] box was used to solvate protein atoms with the minimum thickness of 12 Å. Lipid and water molecules within 2.5 Å from the protein atoms were removed. No counter ions were added to neutralize the system. The final model contained ~140 lipid molecules. First energy minimization was carried out with protein atoms fixed, using 1000 steps steepest descent, followed by 1000 steps conjugate gradient. Second energy minimization was carried out similarly, without any restraints.

Molecular dynamics simulations

CHARMM27 force field [23] with NAMD [32] was used for all energy minimization and MD calculations unless otherwise specified. The SHAKE algorithm was used to constrain bond stretching [33]. The systems were equilibrated for 0.5 ns with positional restraints on AT1 atoms at 293 K to gently relax the lipids, followed by 250 ps run with protein backbone atoms fixed. After equilibrating the system, unrestrained MD was performed for 15 ns with the time step of 2 fs. The impulse-based Verlet-I/r-RESPA method was used to perform multiple time-stepping: 4 fs for short-range non-bonded forces, and 2 fs for bonded forces [34]. Use of non-periodic boundary condition enabled us to perform longer time scale MD. Langevin dynamics was employed to maintain the temperature at 315 K. Langevin damping coefficient was set to 1/ps. The SWITCH cutoff method was used for non-bonded calculations with the first cutoff at 1 nm and last cutoff at 1.2 nm. A neighbor list, used for calculating the non-bonded interactions, was kept to 1.4 nm and updated every eight steps. Trajectories of the simulations were stored at every 1 ps interval. Initial 250 ps were discarded from the productive run. All simulations were carried out by the use of distributed computing setup by 5-10 dual-core CPU nodes connected through Local-Area Network (LAN) in our lab.

Trajectory analysis

VMD, CARMA [35] and EUCB tools were used to analyze molecular dynamics trajectory files. All non-protein atoms were removed, and protein atoms in all the frames were superimposed on the first frame of the trajectory to remove global (rotational and translational) movements. Root mean square deviation (RMSD) of backbone atoms was calculated with reference to the starting structure. Secondary structure of the protein snapshots during the simulation was analyzed with the STRIDE program [36] as implemented in VMD-

TIMELINE plug-in. Backbone hydrogen bonds (HB) were calculated with the donor-acceptor angle cutoff at 45° and distance cutoff at 0.36 nm.

Molecular docking

Small molecule preparation

Arguslab program (www.arguslab.com) was used for small molecule structure generation and optimization. Initially, structures of antagonists were optimized with universal force field (UFF) [37] to get reasonable geometry. The geometry was further refined by PM3 semi-empirical [38] method.

Macromolecule preparation

Docking studies were done using two approaches.

Approach 1 Initial structures were taken as such for docking. In the case of structures optimized by MD, a region was picked at near-end with less than 0.02 nm RMSD for at least 2 ns. Fifty intermittent conformations in this region were extracted. Average structure for the 50-conformation ensemble was calculated using Molmol viewer (http://www.mol.biol.ethz.ch/groups/wuthrich_group/). A conformation, which is structurally close to the average structure, was picked and energy minimized for 500 steps of conjugate gradient to quench the velocities of the atoms. Energy minimization of solvated protein is likely to introduce distortions in the protein structure. Hence, we adopted simple gas-phase energy minimization for our models. This energy minimized structure was used for docking studies.

Approach 2 The above-mentioned structures were docked with Ang II and the complex was subsequently energy minimized for 1000 steps of conjugate gradient to optimize the sidechain geometry for ligand binding. Then the Ang II coordinates were removed from the complex and again energy minimized for 100 steps of conjugate gradient to remove the unnatural strains. During the energy minimization steps, the backbone atoms of AT1 receptor were frozen.

Autodock settings

All the docking calculations were performed by using AutoDock 4.0 [39] in conjunction with Autodock Tools [40]. The protein models were first modified by adding polar hydrogens and merging non-polar hydrogens to the parent carbon atoms. The macromolecule was kept rigid, while all the torsional bonds of ligands were set free to rotate. The docking area was defined by a grid box set to cover the extracellular half of the protein. For each ligand, 50 separate docking calculations were performed using the Lamarckian genetic algorithm local search method with

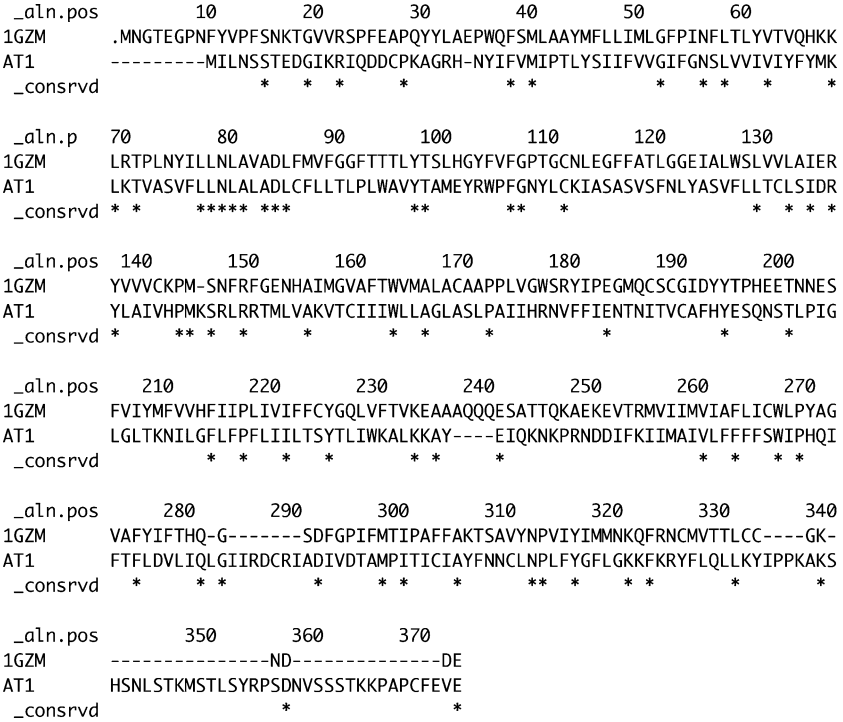

Fig. 1 Sequence-structure alignment between crystal structure of bovine rhodopsin (PDB id: 1GZM) and human AT1 receptor sequence. '*' indicates a match and '-' indicates a gap

default parameters; maximum energy evaluations, 2.5×10^7; population size, 150; mutation rate, 0.02 and crossover rate, 0.8. The docking results from each of the 50 calculations were clustered on the basis of RMSD between the Cartesian coordinates of the ligand atoms and were ranked according to the free energy of binding. The structure with lowest free energy of binding in a highly populated cluster was chosen as the optimal docking pose.

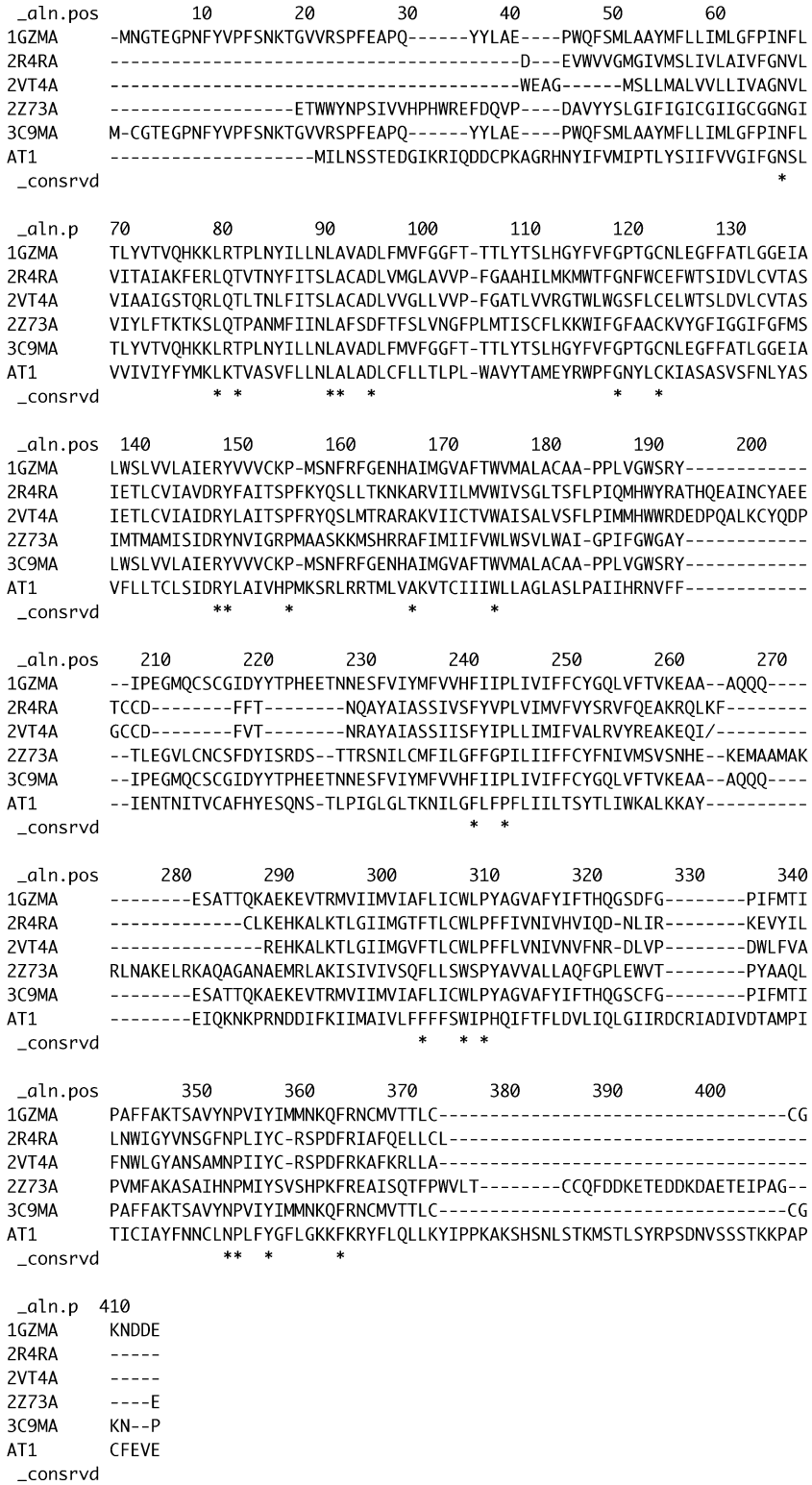

Fig. 2 Multiple structures-sequence alignment between template crystal structures (see the text) and AT1 sequence. Templates were aligned first by 3D-structure alignment and then AT1 sequence was added to the alignment

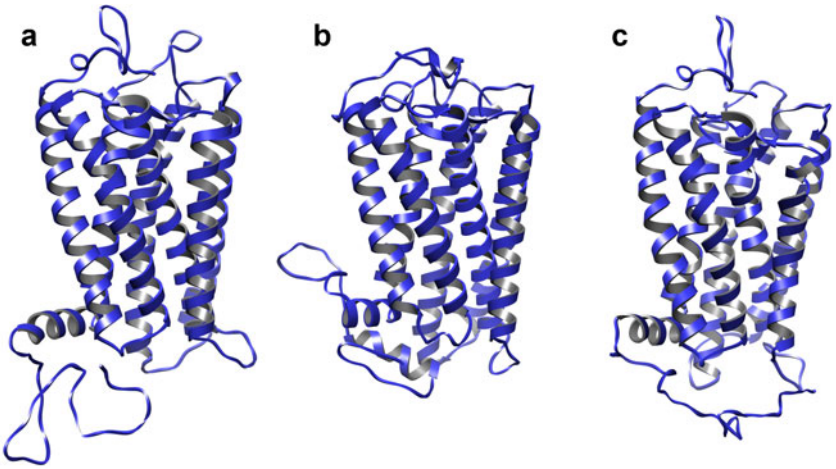

Fig. 3 Structures of initial AT1 models. (**a**) Model 1: based on bovine rhodopsin (1GZM) as template, (**b**) model 2: based on multiple templates and (**c**) model 3: threading model by I-TASSER. Pictures were generated using PyMol

Results and discussion

Structure modeling of AT1 receptor

Sequence-similarity search was done using NCBI-BLASTp – a conventional alignment tool for protein sequences. Sequence similarity is the main criterion to select a template for homology modeling. The sequence similarity between AT1 and known GPCR protein structures is very low (<28%). Therefore, other protein hits which do not belong to GPCR superfamily were screened in the BLAST search. Hence besides the sequence similarity criterion, we used 7-TM architecture as another criterion. The high resolution structure was selected as template when more than one structure was available. AT1 amino acid sequence alignment shows a significant percentage of identity with human β2-adrenergic receptor (2RH1: 28% identity, 50% positives, and 7% gaps for the alignment length 220), squid rhodopsin (2z73: 26%, 42%, and 12% for 284), turkey β1 adrenergic receptor (2VT4: 25%, 45%, and 9% for 307), bovine rhodopsin (1GZM: 20%, 42%, and 4% for 311) and mutant bovine rhodopsin (3C9M: 20%, 43%, and 2% for 261). Even though the target-template sequence similarity falls in the twilight region (<30%), it is still possible to get probable model of AT1 based on these templates since all these proteins belong to rhodopsin superfamily having seven transmembrane helices topology. Sequence-structure alignment has been reported to be superior to sequence-sequence alignment for the purpose of homology modeling and hence the former was used for model building. Alignment of AT1 sequence with 1GZM is shown in Fig. 1. A 29-residue insertion (330-359) in the C-terminal loop was modeled ab initio (model 1).

Since the target-template sequence similarity is less than 30%, it was decided to use multiple templates to improve the model accuracy. For this purpose, the best five hits (2RH1, 2Z73, 2VT4, 1GZM and 3C9M) were selected from the BLAST search. The multiple structure alignment of templates was constructed before aligning it with the target sequence. As shown in Fig. 2, the alignment has relatively fewer amino acid insertions (alignment positions 327-334 and 377-384) when compared to bovine rhodopsin-AT1 alignment. Model building was done by taking locally best regions from template structures as per the target-template structure alignment (model 2). Loop optimization was not performed separately because modeling from multi-templates often produces better results than simple optimization protocols.

AT1 sequence was submitted to I-TASSER 3D structure prediction server, which produced five similar models for AT1. All the models were found to have 7-TM topology and the model with best c-score was chosen. For the ease of discussion, rhodopsin-based model, multiple-templates based model and I-TASSER model will be referred to as model 1, model 2 and model 3, respectively. These three initial models are shown in Fig. 3.

Model quality assessment

All the models were submitted to various model quality assessment programs (MQAP). Results from various MQAPs were inconsistent since these programs are

Table 1 Comparison of 3D structures of AT1 models

AT1 models	RMSD (nm)	TM-score
Model 1, 2	0.788	0.9595
Model 1, 3	0.553	0.9781
Model 2, 3	0.755	0.9618

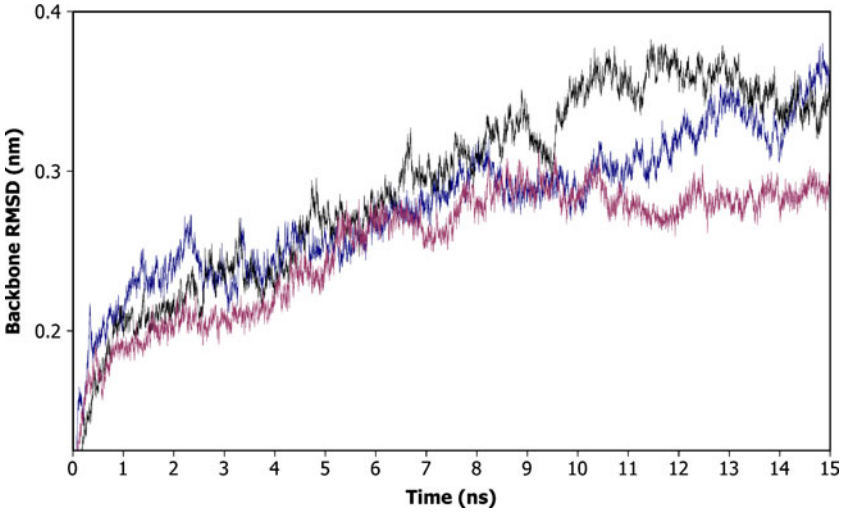

Fig. 4 Root mean square deviation (RMSD) of backbone atoms in coordinates as a function of the simulation time for model 1 (blue), model 2 (magenta) and model 3 (black)

optimized with mostly globular proteins, and therefore, inappropriate to use for membrane proteins. Hence only local errors were taken into account to assess model quality by the use of PROCHECK web-server. Initial models showed low or no violations in their geometrical properties. Ramachandran plot occupancy of residues (non-glycine and non-proline) in the most favored regions was found to be 90.2%, 85.7% and 68.0% for model 1, 2 and 3 respectively. Occupancy of residues in disallowed regions was found to be 0.3%, 2.1% and 7.3%. Model 3 was found to have a high percentage of outliers in Ramachandran plot. However, these residues correspond to loop regions suggesting that loops in the I-TASSER model were not optimized effectively.

AT1 has been reported to contain two-disulfide linkages between CYS18-274 and CYS101-180 positions. However, the models (1-3) did not contain any disulfide bonds. Hence disulfide linkages were introduced and the models were subjected to energy minimization to eliminate unnatural atom contacts.

RMSD and TM-score were used to quantitatively compare the 3D-structures of the models (Table 1). High RMSD values between the models suggest that they significantly differ from each other. Whereas the high TM-scores (on a scale of 0-1) suggest that the topology and overall fold of the models are highly similar and the high RMSD values are mainly due to disoriented segments and loops. For instance, the RMSD between model 1 and 2 was found to be 0.788 nm (Table 1), leading one to think that these two are unrelated proteins. But a TM-score of 0.9595 between them explains that the RMSD is due to difference in orientation of segments and loop conformations.

MD simulation

In cases of low sequence homology, errors in the secondary structure definition and packing of secondary structure elements are common in the generated models. Hence it is

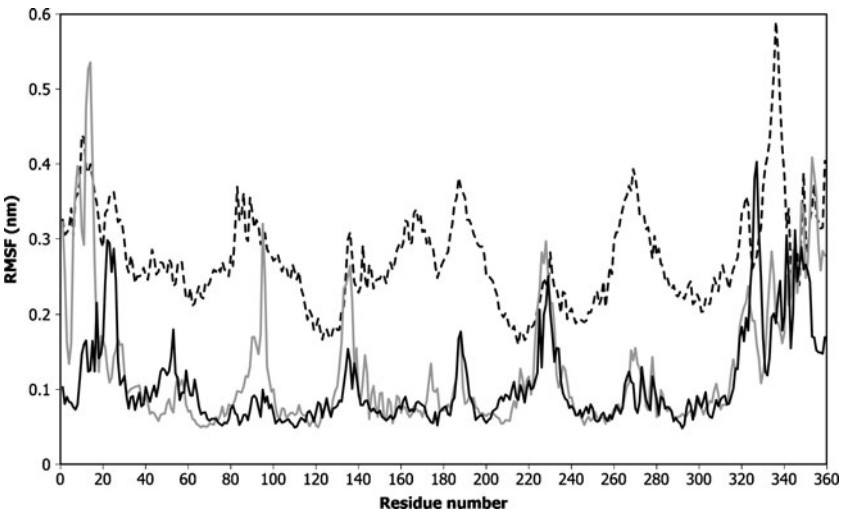

Fig. 5 Root mean square fluctuation (RMSF) of Cα atoms in coordinates for each residue averaged over the duration of the MD simulation. RMSF of model 1 (black dotted line), model 2 (black solid line) and model 3 (gray line) are shown in the picture

essential to critically assess the structure before using them for sensitive applications such as rational drug design. These models were subjected to 15 ns molecular dynamics simulation in explicit membrane-water system to assess the stability and to refine the structures. Periodic boundary condition and particle mesh Ewald (PME) are generally recommended to use in MD simulations. Nevertheless, absence of these conditions in our calculations did not affect the stability of the simulation as it can be seen from our results (RMSD plots and radius of gyration plots). To quantify the extent to which our models represent both a realistic and a stable state of the protein, in the timescales studied, the following parameters were assessed: the root mean square deviations (RMSD) of its backbone, secondary structure, total number of hydrogen bonds (nHbonds), mass weighted radius of gyration (Rg) and root mean square fluctuation of Cα-atoms with respect to the initial values.

Analysis of RMSD of the backbone from the initial structure can help to assess the stability of models. In MD simulations, conformations that are close to global energy minimum or trapped in a minimum would result in a flat RMSD curve for a significant period of time, since thermal fluctuations would be relatively small. RMSD profiles of models as a function of simulation time are given in Fig. 4. Initial structural change during 0-1 ns may be due to equilibration of the models with the surroundings. In the case of model 1 (blue curve), the RMSD was stabilized considerably around ~5 ns. Steady increase in RMSD was observed during ~6-15 ns simulation, suggesting that the stability of the model is suboptimal. Model 2 attains a plateau and relaxes during ~8-15 ns as shown in Fig. 4 (magenta curve), which can be explained by the fact this model is relatively well structured and stabilized in a potential well. For model 3, the RMSD steadily increases to 0.37 nm during ~12 ns and slightly falls back to 0.31 nm which may be due to unpacking and repacking (black curve in Fig. 4). However, distance-averaged RMSD is not residue specific and high fluctuations in loops alone can result in high RMSD. In order to find out which part of the model is highly flexible, time-averaged RMSD or RMSF of Cα-atoms was analyzed during the simulation. As observed in Fig. 5, all models displayed elevated RMSF in loops especially in N- and C-terminals, although it is relatively high for model 1 (black dotted line). Fluctuations in model 2 (black line) and 3 (grey line) are comparable. It is interesting to note that model 2 has least RMSF among the other candidates in almost all regions, indicating overall stability of the structure while model 1 displayed very high fluctuations, indicating suboptimal nature of the model.

Analysis of secondary structural elements during the simulation is given Fig. 6. Clearly defined, long-lived 7-TM α-helices (magenta patches) and highly fluctuating intermittent structurally variable regions (turns = cyan, loops = white, π-helix = red and 3_{10}-helix = blue) are seen in the figure. In all three models, TM-3 helix is the least affected one by thermal fluctuations. In addition, TM-1 and TM-2 helices in model 1 (Fig. 6a) were relatively more stable and

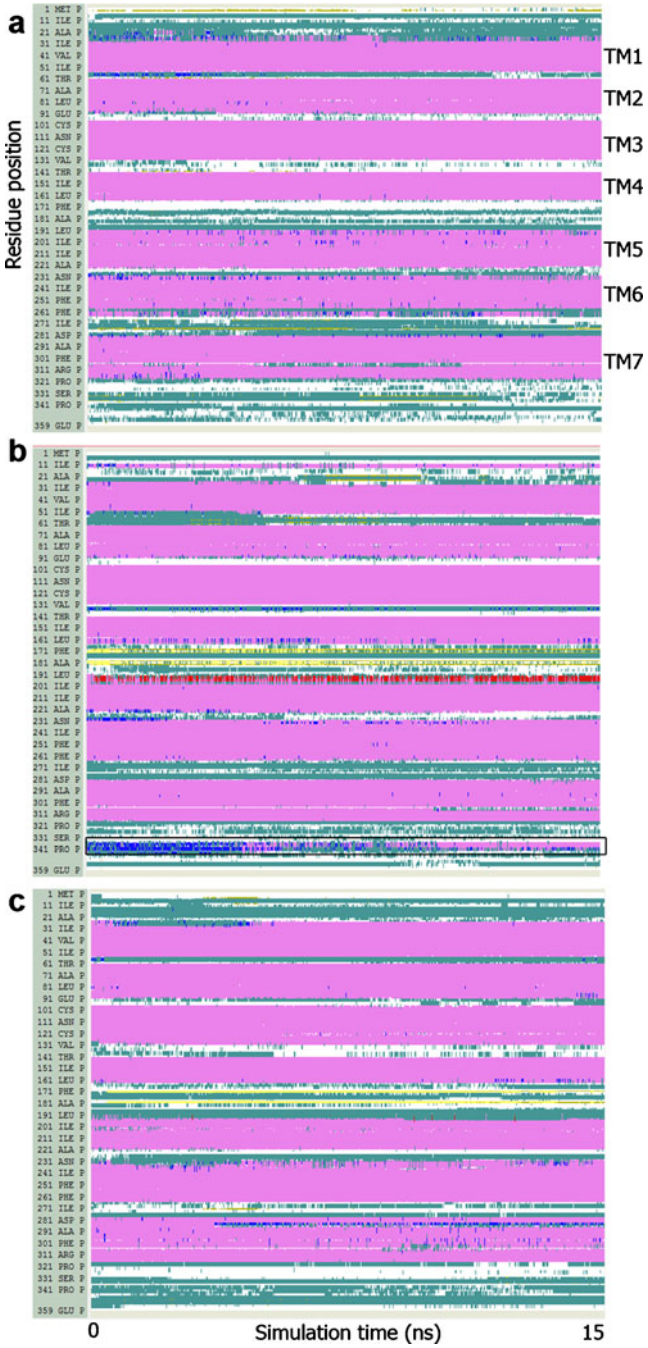

Fig. 6 Stability of secondary structures during the simulation, calculated by STRIDE algorithm. Residue position is in y-axis and simulation time is in x-axis. Each color represents a secondary structure: α-helix – magenta, β-strand – yellow, turns – cyan loops – white, 3_{10}-helix – blue and π-helix – red. Plot shows stable and well-defined TM-helical regions, and highly fluctuating variable regions. (**a**) model 1, (**b**) model 2 and (**c**) model 3. Transition of 3_{10}-helix to a more stable α-helix in model 2 is marked in a box

less affected. While the other four helices (TM-4 to TM-7) and loops were highly variable. In model 2, TM-3 helix was not subjected to any change throughout the simulation period and most of the secondary structural elements were intact during the simulation period as seen in Fig. 6b. Interestingly, a 3_{10}-helix in the C-terminal of initial model assumes a well-defined α-helix during the simulation (highlighted in Fig. 6b). In the case of model 3 (Fig. 6c), the secondary structures were well preserved except in TM-7 helix in which fluctuation was notably high. First few residues of TM-1 helix fluctuate initially to refine and reform the helix that was not disturbed afterwards. Fluctuations in both ends of the helices were, invariably, found in all models.

Radius of gyration (Rg) describes the overall spread of the molecule from its center, which might give additional information such as compactness of the system. Change of Rg of models during the simulation is given in Fig. 7. For models 1 and 3 no specific observation was found. In the case of model 2, after initial rise in Rg, the structure relaxes around 2.33 nm. Moreover, model 2 was found to have least Rg than the others, suggesting it is relatively more compact.

As alternative measures of evaluation, whether simulation has improved the modeled structures, the number of backbone-backbone hydrogen bonds (HB) was also considered. Backbone hydrogen bonding is indicative of regular secondary structure content and it further reflects the compactness of the structure [41]. Decrease in the number of backbone hydrogen bonds indicates the structure is unfolding. At least 5% reduction in HB was observed (Fig. 8) for model 1 and 3 with respect to the initial numbers. However, model 2 showed no significant reduction in HB and fluctuated constantly around the 90 mark. Also it exhibited a greater number of backbone HB.

Representative structures from MD simulation trajectories (denoted as, model 1_MD, model 2_MD and model 3_MD) were selected based on the RMSD profile and assessed with PROCHECK. Model 1_MD, model 2_MD and model 3_MD showed 82.6%, 81.7% and 80.8% occupancies in most favored regions, and 0.9%, 0.9% and 1.8% occupancies in disallowed regions of Ramachandran plot, respectively. When compared with initial models (90.2%, 85.7% and 68.0% in mostly favored, and 0.3%, 2.1% and 7.3% in disallowed regions for model 1, 2 and 3 respectively), quality of model 1 was considerably reduced whereas model 3 showed refinement during the simulations. In the case of model 2 no significant change in quality was observed. But, number of outliers in Ramachandran plot was slightly less in the MD structure. Because the Ramachandran plot occupancy values for the MD structures are very similar, we analyzed their structural similarity to check if the simulation has moved the structures towards a common center in the potential energy surface. RMSD and TM-score differences between the MD structures are comparable to the differences between the initial structures used for optimization (data not shown). This fact suggests that the MD structures are as distinct as the initial structures.

Molecular docking

Finally, the quality of models was assessed by molecular docking experiment by the use of Autodock 4.0. Seven AT1 antagonists, candesartan, eprosartan, irbesartan, losartan, olmesartan, telmisartan and valsartan were used for docking with initial structure and a representative structure from MD. The IC_{50} values of these antagonists were known under similar conditions [42], and hence were chosen for

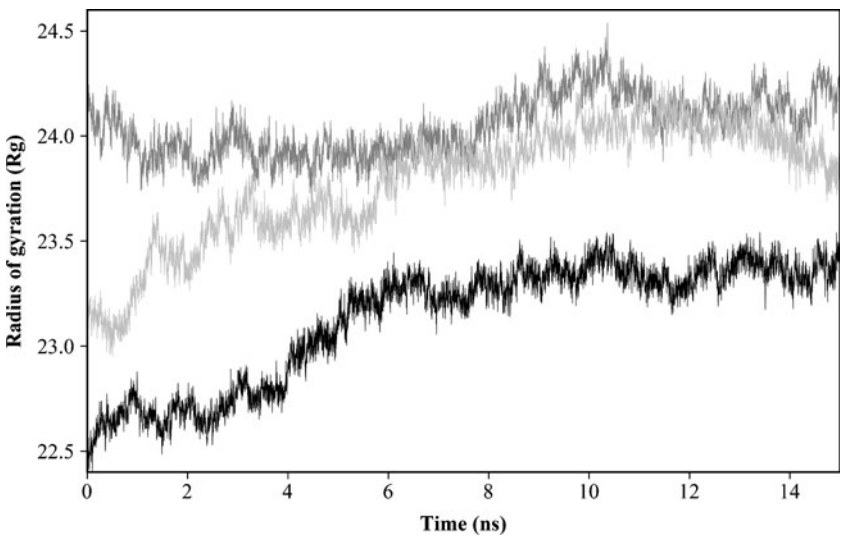

Fig. 7 Mass-weighted radius of gyration as a function of simulation time for model 1 (light grey), model 2 (black) and model 3 (dark grey)

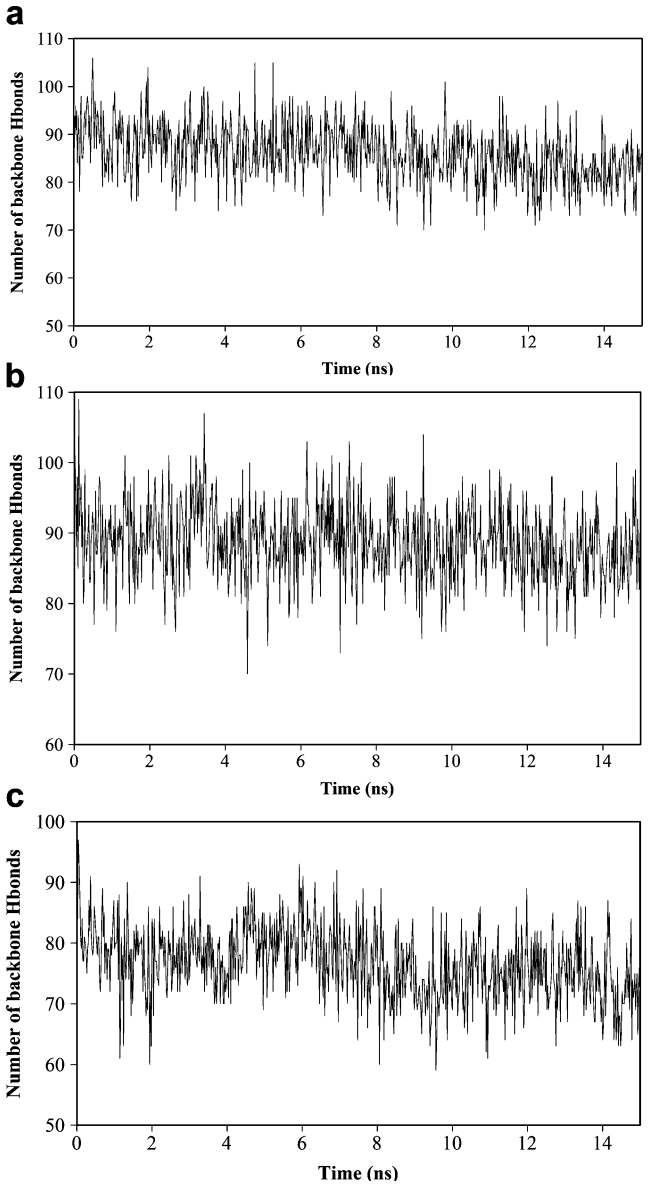

Fig. 8 Number of backbone hydrogen bonds as a function of simulation time for models 1 (**a**), 2 (**b**) and 3 (**c**)

docking studies. These compounds showed low binding affinity towards initial and MD structures of all the models. However, the binding site was found to be same as reported earlier for these antagonists (data not shown). This may be due to small errors in the packing of helices and/or sidechain groups. Simple energy minimization of models prior to docking did not improve docking score. Hence, the sidechains of ligand-binding site residues were optimized by energy minimization in the presence AngII. The docking results of sidechains-optimized models are given in Table 2. The sidechains-optimized models showed high affinity binding, although the binding scores of initial models, MD structures 1 and 3 did not correlate with experimental IC_{50} values. This may be due to incorrect binding-pocket configuration. It is observed that estimated free energy of binding for model 2_MD correlated well with the experimentally reported IC_{50} values of the antagonists. Moreover, the binding pose reveals the vital interactions, important for specific ligand binding, to be similar to that reported earlier [2]. For instance, the binding-site of candesartan was found to be an inter-helical space, lined by TM 3-6 helices (Fig. 9a). Biphenyl moiety is juxtaposed with a hydrophobic cavity created by L112, Y113, V108 and F182, and acidic tetrazole ring nitrogen atoms form hydrogen bonds with K199 (Fig. 9b).

Ligand-binding pocket of model 2_MD with bound candesartan is given in the Fig. 9c. The pocket is characterized by narrow and compact groove enclosed by TM3-TM6 helices to accommodate biphenyl and tetrazole moieties of the antagonists. There is a broad hydrophobic cavity created by TM6 and TM7 helices that can accommodate bulky hydrophobic substituents attached to the 4'-position of the biphenyl ring. In the case of model 3_MD, the entire pocket was found to be narrow due to closely packed TM helices and the hydrophobic patches were less as compared to that of model 2_MD. Whilst, model 1_MD was found to have well-structured compact binding pocket for biphenyl-tetrazole moieties, it is sterically hindered for bulkier 4'-substituents, which explains the reduced docking score for bulkier ligands such as telmisartan. Sterically

Table 2 Binding affinities of different antagonists with AT1 models, calculated by Autodock 4 and their correlation with IC_{50} values

AT1 antagonists	Free energy of binding (kcal/mol)						IC_{50} [a] (nM)
	1	1_MD	2	2_MD	3	3_MD	
Candesartan	−10.07	−10.37	−10.51	−11.21	−9.50	−11.27	0.11
Olmesartan	−11.62	−10.81	−10.51	−11.20	−9.75	−10.98	0.13
Eprosartan	−9.08	−11.60	−10.26	−10.67	−8.11	−10.67	0.29
Telmisartan	−8.85	−7.60	−11.03	−10.48	−12.78	−11.21	0.33
Valsartan	−10.65	−8.95	−9.50	−10.12	−9.72	−11.11	0.55
Losartan	−10.54	−12.36	−10.56	−9.93	−9.90	−10.46	2.45
Irbesartan	−10.66	−13.16	−10.39	−9.86	−10.33	−11.62	4.00

[a] IC50 values obtained from literature [39]

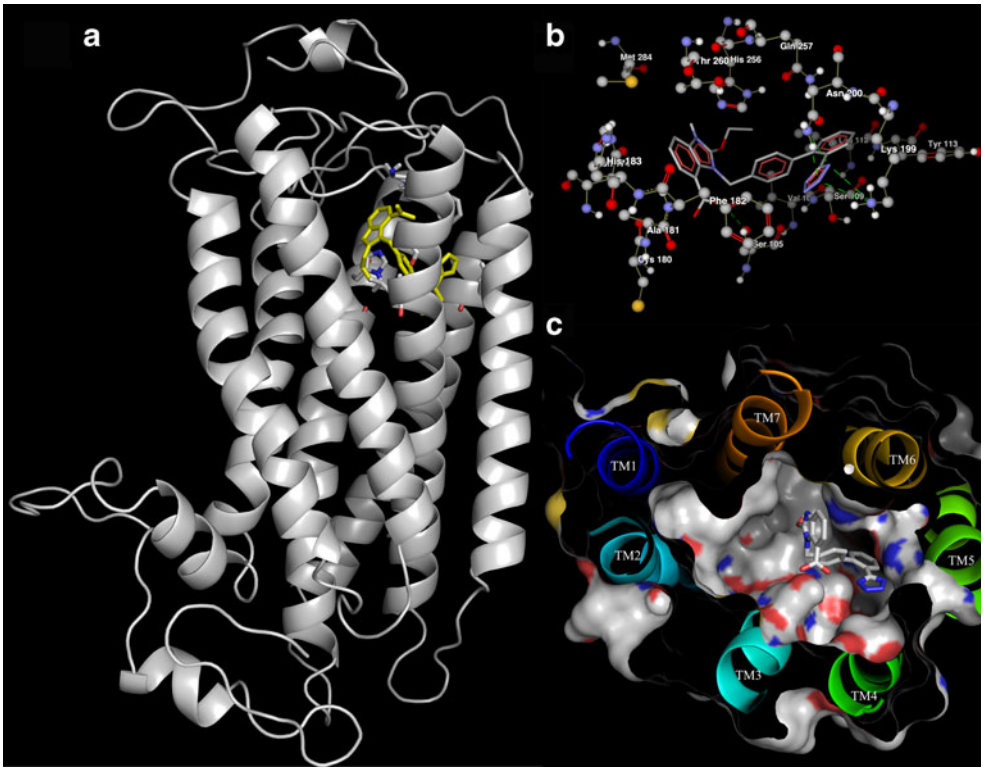

Fig. 9 Representation of ligand-binding pocket and binding mode in model 2_MD. (**a**) Overview of candesartan (rendered as yellow sticks) binding with AT1 receptor (rendered as ribbons). Residues (V108, S109, K199, W253 and H256) important for antagonist binding are rendered as sticks (colored by atom type). (**b**) Candesartan (stick model colored by atom type) and interacting residues (ball and stick colored by atom type) are shown. Hydrogen bonding interactions, tetrazole ring-K199 and ligand carboxylate-S105, are depicted as green dotted lines. (**c**) Ligand binding pocket in surface representation with candesartan (stick model). TM-helices 1-7 are shown in the picture. (Generated using PyMol)

hindered binding pocket due to tight packing of helices and/or sidechains renders the initial models to have low antagonist binding affinity. Upon sidechain optimization, considerable increment in binding affinity was observed, though the docking scores did not correlate with IC_{50} values. This explains the suboptimal packing of helices in initial models and MD structures 1 and 3. Model 2_MD explains trend of ligand binding affinity for seven potent antagonists and hence it can be potentially useful in studying novel antagonists in view of designing better drugs.

The improved quality of model 2 might be due to the incorporation of multiple templates, which provides comprehensive coverage across the target sequence when compared to single-template [16]. In multiple-templates based modeling, each template contributes to the model building and the target structure takes up the conformation of local best region from each template, based on the target-templates sequence alignment.

AT1 structure has been extensively modeled and used by many researchers, but most of their modeling studies are limited to bovine rhodopsin based models and shorter timescale (~1-5 ns) MD simulation. Tuccinardi et al. have reported the homology modeling of AT1 receptor using bovine rhodopsin as template and validated their model by docking with several antagonists [13]. However, the authors have proposed a new binding orientation for the non-peptide antagonists. In the present study, we have considered only the classical binding orientation where tetrazole moiety of the antagonists interacts with K199 through a salt-bridge. Due to the lack of direct experimental evidences, the probable ligand-binding orientation is still unclear. In addition, Tuccinardi et al. have used 1 ns MD of AT1 receptor-losartan complex to refine the model for docking studies. On the other hand, a ligand-free AT1 model is used in the present study for 15 ns MD simulation, mainly to represent the native structure of the receptor. The presence of ligand during the simulation might affect the tight packing of TM helices which may introduce artifacts in the model. Hence, we chose to optimize only the sidechain residues by simple energy minimization of AT1-AngII complex while restraining backbone atoms. It is worth mentioning that the crystal structure of bovine rhodopsin (1GZM) used in our study is truncated (329-348) in the C-terminal region, which causes insertions/deletions in the alignment. However, it has been generally observed that mutation or deletion in the C-terminal tail does not seriously affect the ligand binding affinity [43].

For the first time, we have modeled AT1 receptor by using multiple templates and evaluated the model with relatively longer timescale (15 ns) MD simulation. We also compared three different models obtained by different methodologies (rhodopsin, multiple-templates and I-TASSER based). Longer MD simulations revealed that initial time (up to ~6 ns) of the simulations were less informative, suggesting the need for a longer timescale.

Conclusions

Three different 3D-models for AT1 receptor are compared for stability, quality and ligand-binding by means of relatively longer time scale molecular dynamics calculation in explicit membrane-water system. Our results demonstrate that the accuracy of the routinely used bovine rhodopsin-based AT1 model is limited. The newer approach, AT1 homology modeling with multiple templates, shows enhanced accuracy. Further, this model satisfactorily explains the pattern of antagonist binding. Such a model has potential advantage in rational drug design for AT1 implicated diseases.

Acknowledgments We thank University Grants Commission (UGC), New Delhi, for the award of Senior Research Fellowship to S. Pandian. Thanks are due to UGC-NRCBS, MKU, for providing the Lab facilities.

References

1. MatsusakaT II (1997) Biological functions of angiotensin and its receptors. Annu Rev Physiol 59:395–412
2. De Gasparo M, Catt KJ, Inagami T, Wright JW, Unger T (2000) International union of pharmacology. XXIII. The angiotensin II receptors. Pharmacol Rev 52:415–472
3. Meng EC, Bourne HR (2001) Receptor activation: what does therhodopsin structure tell us? Trends Pharmacol Sci 22:587–593
4. Gether U (2000) Uncovering molecular mechanisms involved inactivation of G protein coupled receptors. Endocr Rev 21:90–113
5. Petrey D, Honig B (2005) Protein structure prediction: inroads to biology. Mol Cell 20:811–819
6. Rayan A (2009) New vistas in GPCR 3D structure prediction. J Mol Model 16:183–191
7. Li M, Fang H, Du L, Xia L, Wang B (2008) Computational studies of the binding site of α1A-adrenoceptor antagonists. J Mol Model 14:957–966
8. Sasaki K, Yamano Y, Bardhan S, Iwai N, Murray JJ, Hasegawa M, Matsuda Y, Inagami T (1991) Cloning andexpression of a complementary DNA encoding a bovine adrenal angiotensin II type-1 receptor. Nature 351:230–233
9. Mukoyama M, Nakajima M, Horiuchi M, Sasamura H, Pratt RE, Dzau VJ (1993) Expression cloning of type 2 angiotensin II receptor reveals a unique class of seven transmembrane receptors. J Biol Chem 268:24539–24542
10. Yamato Y, Ohyama K, Kikyo M, Sano T, Nakagomi Y, Inoue Y (1995) Mutagenesis and molecular modeling of rat angiotensin II receptor. J Biol Chem 270:14024–14030
11. Nikiforovich GV, Marshall GR (2001) 3D model for TM region of the AT-1 receptor in complex with angiotensin-II independently validated by site-directed mutagenesis data. Biochem Biophys Res Comm 286:1204–1211
12. Deraet M, Rihakova L, Boucard A, Perodin J, Sauve S, Mathieu AP, Guillemette G, Leduc R, Lavigne P, Escher E (2002) Angiotensin II is bound to both receptors AT1 and AT2, parallel to the transmembrane domains and in an extended form. Can J Physiol Pharmacol 80:418–425
13. Tuccinardi T, Calderone V, Rapposelli S, Martinelli A (2006) Proposal of a new binding orientation for non-peptide AT1 antagonists: Homology modeling, docking and three-dimensional quantitative structure−activity relationship analysis. J Med Chem 49:4305–4316
14. Baleanu GC, Karnick S (2006) Model of the whole rat AT1 receptor and the ligand-binding site. J Mol Model 12:325–337
15. Fan H, Mark AE (2004) Refinement of homology-based protein structures by molecular dynamics simulation techniques. Protein Sci 13:211–220
16. Mobarec JC, Sanchez R, Filizola M (2009) Modern homology modeling of G-protein coupled receptors: which structural template to use? J Med Chem 52:5207–5216
17. Fernandez-Fuentes N, Rai BK, Madrid-Aliste CJ, Fajardo JE, Fiser A (2007) Comparative protein structure modeling by combining multiple templates and optimizing sequence-to-structure alignments. Bioinformatics 23:2558–2565
18. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9:40
19. Pruitt KD, Tatusova T, Maglott DR (2005) NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 33(Database Issue):D501–D504
20. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. Nucl Acids Res 28:235–242
21. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410
22. Fiser A, Sali A (2003) Modeller: generation and refinement of homology-based protein sequence models. Methods Enzymol 374:461–491
23. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WR III, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102:3586–3616
24. Madhusudhan MS, Webb BM, Marti-Renom MA, Eswar N, Sali A (2009) Alignment of multiple protein structures based on sequence and structure features. Prot Eng Des Sel 22(9):569-574
25. Yamano Y, Ohyama K, Chaki S, Guo DF, Inagami T (1992) Identification of amino acid residues of rat angiotensin II receptor for ligand binding by site directed mutagenesis. Biochem Biophys Res Commun 187:1426–1431
26. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods R (2005) The amber biomolecular simulation programs. J Comput Chem 26:1668–1688
27. Zhang Y, Skolnick J (2004) Scoring function for automated assessment of protein structure template quality. Proteins 57:702–710
28. Levitt M, Gerstein M (1998) A unified statistical framework for sequence comparison and structure comparison. Proc Natl Acad Sci USA 95:5913–5920
29. Laskowski RA, MacArthur MW, Moss DS, Thorton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26:283–291
30. Humphrey W, Dalke A, Schulten K (1996) VMD-visual molecular dynamics. J Mol Graph 14:33–38
31. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935
32. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26:1781–1802
33. vanGunsteren WF, Berendsen HJC (1977) Algorithms for macromolecular dynamics and constraint dynamics. Mol Phys 34:1311–1327

34. Benz RW, Castro-Román F, Tobias DJ, White SH (2005) Experimental validation of molecular dynamics simulations of lipid bilayers: a new approach. Biophys J 88:805–817
35. Glykos NM (2006) CARMA: a molecular dynamics analysis program. J Comput Chem 27:1765–1768
36. Andersen CAF, Palme RAG, Brunak S, Rost B (2002) Continuum secondary structure captures protein flexibility. Structure 10:175–185
37. Rappe AK, Casewit CJ, Colwell KS, Goddard WA III, Skiff WM (1992) UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. J Am Chem Soc 114:10024–10035
38. Stewart JJP (1989) Optimization of parameters for semiempirical methods. J Comput Chem 10:209–220
39. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem 19:1639–1662
40. Sanner MF (1999) Python: a programming language for software integration and development. J Mol Graph Model 17:57–61
41. Sticke DF, Presta LG, Dill KA, Rose GD (1992) Hydrogen bonding in globular proteins. J Mol Biol 226:1143–1159
42. Mire DE, Silfani TN, Pugsley MK (2005) A review of the structural and functional features of olmesartan medoxomil, an angiotensin receptor blocker. J Cardiovasc Pharmacol™ 46:585–593
43. Thomas WG, Thekkumkara TJ, Motel TJ, Baker KM (1995) Stable expression of a truncated AT_{1A} receptor in CHO-K1 cells. J Biol Chem 270:207-213