J Mol Model (2006) 12: 325–337
DOI 10.1007/s00894-005-0049-z

ORIGINAL PAPER

Camelia Baleanu-Gogonea · Sadashiva Karnik

Model of the whole rat AT_1 receptor and the ligand-binding site

Received: 15 April 2005 / Accepted: 22 July 2005 / Published online: 11 January 2006
© Springer-Verlag 2006

Abstract We present a three-dimensional model of the rat type 1 receptor (AT_1) for the hormone angiotensin II (Ang II). Ang II and the AT_1 receptor play a critical role in the cell-signaling process responsible for the actions of renin–angiotensin system in the regulation of blood pressure, water-electrolyte homeostasis and cell growth. Development of improved therapeutics would be significantly enhanced with the availability of a 3D-structure model for the AT_1 receptor and of the binding site for agonists and antagonists. This model was constructed using a combination of computation and homology-modeling techniques starting with the experimentally determined three-dimensional structure of bovine rhodopsin (PDB#1F88) as a template. All 359 residues and two disulfide bonds in the rat AT_1 receptor have been accounted for in this model. Ramachandran-map analysis and a 1 nanosecond molecular dynamics simulation of the solvated receptor with and without the bound ligand, Ang II, lend credence to the validity of the model. Docking calculations were performed with the agonist, Ang II and the antihypertensive antagonist, losartan.

Keywords AT_1 receptor · Ang II · GPCR model · Molecular dynamics simulation · Docking

Introduction

G-protein-coupled receptors (GPCRs) are seven-helix transmembrane (TM) proteins that play an essential role in the transmission of a large variety of external signals. [1] Present in a broad range of organisms, they cause the activation of heterotrimeric guanine nucleotide-binding proteins (G-protein) in response to stimuli as diverse as light, odorants, neurotransmitters, peptides and hormones. Understanding the structure and mechanism of GPCRs is, thus central to many aspects of cellular signaling and control. As a result, many multidisciplinary research projects, including those carried out by pharmaceutical companies to find new therapeutic molecules, are aimed at members of the GPCR superfamily. There is only one experimentally elucidated three dimensional (3 D)-structure for GPCRs [2]. Given the high-level of interest these proteins attract, it is not surprising that predictive methods to derive information about their 3D-structure are essential. Until now theoretical models of TM regions in several GPCRs had been used to rationalize interactions between the proteins and their ligands deduced by site-directed mutagenesis experiments.

Angiotensin II (Ang II) type 1 receptor (AT_1) is a GPCR that regulates all known physiological functions of Ang II, a peptide hormone product of the renin angiotensin system [3]. Ang II binding to the AT_1 receptor plays a critical role in cellular signaling processes responsible for information transfer in cells involved in the regulation of normal blood pressure and water–electrolyte balance. Abnormalities in its function have been linked to many pathological conditions such as cardiac hypertrophy, renal hypertrophy and proliferation of arterial vascular smooth muscle cells. Therefore, design and development of specific improved remedies for hypertension, heart and kidney failures targeting the AT_1 receptor would be significantly enhanced with the availability of the 3D-structure of the receptor and of the binding site for Ang II [4]. The rat AT_1 receptor is composed of a single polypeptide of 359 residues that is predicted to form seven hydrophobic TM α-helices connected by three loop regions on both the extracellular (EC) and intracellular (IC) sides of the embedding membrane [3]. Mammalian AT_1 receptors are highly homologous (up to 95%) and bear 32% overall homology with rhodopsin (Rh) (Fig. 1).

Several attempts have been made to model the human AT_1 receptor structure using bacteriorhodopsin 3D structure [5–7] as a template. Because bacteriorhodopsin is not a

C. Baleanu-Gogonea · S. Karnik (✉)
Department of Molecular Cardiology
at Lerner Research Institute,
The Cleveland Clinic Foundation,
9500 Euclid Avenue,
Cleveland, OH 44195, USA
e-mail: karniks@ccf.org
Tel.: +1-216-4441269
Fax: +1-216-4449263

Fig. 1 The homology between the primary structures of rat AT_1 receptor and bovine Rh. The conserved residues are depicted in *red*. Locations of the two gaps in the Rh crystal structure are highlighted in *purple* and the necessary insertions/deletions are highlighted in *green*. The position of the α-helices are in *yellow*. **GA** 1.1 nm gap; **G** 1.38 nm gap;*73 identical amino acids; 20% identity;*conserved amino acid, or motif

GPCR, direct use of its structure to model the TM-domain structure of the AT_1 receptor is limited. The availability of a 2.8 Å resolution bovine Rh X-ray structure [2] made it possible for Nikiforovich and Marshall to model the rat AT_1 receptor TM-domain [8]. The Rh structure delineates the intra and EC domains as well, which lead to a more complete AT_1 receptor model developed by Deraet et al. [9] and most recently by Oliveira (http://www.gpcr.org/7tm/models/oliveira/index.html). These models, however, lack more than 61 amino acids out of the 359 residues; prompting the modeling and experimental community in constant search of a more complete molecular model of the AT_1 receptor.

Experimental studies have suggested that both integral membrane and extra-membrane regions of the AT_1 receptor play a critical role in the binding of Ang II, isomerization of the receptor to its active state and the activation of the G protein [10–15]. Furthermore, an EC disulfide bond linking the N-terminal tail to the EC-loop 3 exists in the AT_1 receptor in addition to the conserved disulfide bond linking EC-loop 2 to TM3, which is found in 92% of GPCRs [16–18]. Inclusion of these essential structural constraints is therefore important for a realistic model, which could be useful for the interpretation of experimental ligand-binding results and molecular-dynamics (MD) simulation of the receptor activation process. We present such a model deduced from the 359 residues of the rat AT_1 receptor including the two structurally critical disulfide bonds. The model derived from Rh by computational mutagenesis was further validated using Ramachandran-map analysis and MD-simulations. Using the AutoDock3.05 program, we predicted the Ang II binding site and the binding site of an antihypertensive agent, losartan.

Materials and methods

Computational methods

In this study we used sequence homology, force-field energy minimization, MD-simulation, and docking explorations to prepare a model for the AT_1 receptor and identify the binding sites of Ang II agonist.

Transformation of Rh model to AT_1 receptor 3D structure

The molecular model proposed here uses Rh (PDB access No.1F88) as a starting point in modeling [2]. First, we performed sequence alignment of the rat AT_1 receptor primary sequence with that of bovine Rh using CLUSTALW program [19]. From the total of 359 amino acids only 73 were found to be identical. As can be seen from Fig. 1, special attention was paid during the alignment to conserved GPCRs amino acids, and motifs: N46, D74, F94, G97, DRY (residues 125–127), W153, P207, Y215, F249, P255, NPXXY (residues 298–302). The mutation of the remaining 286 amino-acid residues was made sequentially, one-at-a-time following the homology comparison, fol-

lowed by gas-phase energy minimization using the OPLS [20, 21] force field as implemented in the MD-simulation program Gromacs [22, 23]. The deletions and insertions were performed, one amino acid at a time, and then a series of constrained gas-phase energy minimizations was done successively by reducing the constraint in 0.5 Å steps. The crystal structure of Rh has two gaps, one of 11.0 Å (between Glu^{227} and Ile^{228}, in our model), and the second of 13.8 Å (between Ser^{326} and His^{327}, in our model) that were inherited by the new model. The gaps were gradually removed by performing step-wise gas-phase energy minimizations until the missing peptide bond was formed. In each step the gap was reduced by 0.5 Å by performing constrained energy minimization, during which the gap size was kept constant (Fig. 1) [2].

Conservation of AT_1 secondary structure in the TM region

One important issue in eliminating the gaps was the preservation of the secondary structure. During the entire modeling process, the secondary structure of the AT_1 TM domain was conserved by imposing constraints on the α-helix backbone ϕ and ψ torsion angles such that they remained in the allowed range for α-helices. Because Gromacs only allows one to define distance constraints and restraints, we selected a set of distance constraints/restraints that ensured that the ϕ and ψ angles for an α-helix backbone remained in the allowed range of values and at the same time the bond angles associated with the backbone atoms were not severely distorted. The difference between constraints and restraints is the following: distance constraints are strictly enforced by the program to their nominal values using the LINCS algorithm (which is also used to constrain the bond lengths between atoms during a MD simulation), while for distance restraints the program uses a harmonic oscillator penalty function which, depending on the magnitude of the force constant used, can lead to a larger fluctuation in the actual distances (between restrained atoms) obtained at the end of energy minimization. We generated the set of required distance constraints/restraints for all TM α-helices using a custom-made Pymol [4] wizard that runs as a Pymol plug-in. The wizard takes the Pymol selection (in this case the set of atoms of an α-helix sequence) as an input and generates distance constraints for the following bond angles associated with the α-helix backbone: $C_{(-)}$–N–C_α, N–C_α–C and C_α–C–$N_{(+)}$, and similar generates distance restraints for the following sets of terminal atoms that define the ϕ and ψ angles: C atoms of $C_{(-)}$–N–C_α–C torsion angle and N atoms of the N–C_α–C–$N_{(+)}$ torsion angle, respectively. This combination of distance constraints/restraints ensures that the bond angles associated with the α-helix backbone are not distorted. Without using constraints the bond angles tend to be distorted at the expense of keeping the ϕ and ψ torsion angles in the allowed region.

Orientation of the α-helix residues in the TM regions

Inward or outward orientation of amino-acid residues from the α-helices in the TM region was assigned based on their solvent accessibility. The solvent accessible surface area (SASA) of the α-helix residues was calculated using the "G_SAS" program from the Gromacs package [22, 23]. The program written in Python language, which runs alongside with the Pymol molecular viewer, was custom designed to perform the following tasks: (i) the first function (READ_SAS) reads in the solvent accessible surface areas of the residues calculated by the "G_SAS" program and matches the data in the SASA file with those in the protein structure file (number of residues and their type), (ii) the second function "SEL_SARES" takes a Pymol selection (e.g., an α-helix) and uses a cutoff in SASA (in this case the cutoff was 0.5 nm^2) to determine which residues point outward and which ones point in towards the protein core, and (iii) the third function assigns the hydrophobicity of the amino-acid residues in the α-helix sequence and calculates their orientation distribution. The hydrophobicity is obtained from computed log(P) using the small fragment approach, which employs a hydrophobicity scale between 0.5 and 1.0 [24]. According to this scheme, alanine, cysteine, phenylalanine, glycine, isoleucine, leucine, methionine, proline, valine, tryptophan and tyrosine are hydrophobic. "SEL_SARES" function outputs four Pymol selections: two selections are the sets of residues that point outward and in towards the protein core, respectively, and the other two selections are the sets of hydrophobic and hydrophilic residues from the selected α-helix. Thus, the amino-acid residues were classified as pointing in towards the protein core if their SASA was less than 0.5 nm^2, and *vice-versa* for each α-helix (Table 1).

Adjustment of structure and the disulfide bonds

Before the reconstruction of the disulfide bonds Cys^{18}–Cys^{274}, and Cys^{101}–Cys^{180}, the missing amino acids (Asp^{343}–Glu^{359}) were inserted in order to model the entire C-terminal sequence. Fig. 1 shows the homology between the primary structure of Rh and rat AT_1 receptor. The location of the two gaps is highlighted in magenta, the conserved residues are depicted in red, and the necessary insertions in green.

Addition of N-terminus, C-terminus and disulfide bonds

The deletions and insertions were performed one amino acid at a time, followed by energy minimization. Before making any insertion, the gaps (Fig. 1) were removed to ensure preservation of the secondary structure. The residues in the AT_1 receptor for which there were no analogous positions in the Rh model (Fig. 1) were built by

Table 1 Orientational distribution of the amino acid residues in the TM domain of the of AT_1 receptor

α-Helix	Range of residues	Hydrophobic residues[a] [%]			Hydrophilic residues[a] [%]		
		Total [%]	Inward [%]	Outward [%]	Total [%]	Inward [%]	Outward [%]
H1	25–55	83.9	38.7	45.2	16.1	9.7	6.4
H2	62–90	82.8	55.2	27.6	17.2	13.8	3.4
H3	98–132	64.7	50.0	14.7	35.3	29.4	5.9
H4	144–171	82.6	34.8	47.8	17.4	8.7	8.7
H5	193–217	76.0	28.0	48.0	24.0	8.0	16.0
H6	235–266	69.2	42.3	26.9	30.8	26.9	3.9
H7	281–302	78.9	52.6	26.3	21.1	21.1	0.0

[a]The amino acid residues are classified as hydrophobic or hydrophilic based on their hydrophobicity, which is obtained from computational log(P) using the small fragment approach [24]. Hydrophobic amino acid residues have hydrophobicity between 0.5 and 1.0. According to this scheme alanine, cysteine, phenylalanine, glycine, isoleucine, leucine, methionine, proline, valine, tryptophane and tyrosine are hydrophobic

using the SwissPDBViewer [25], in order to model the entire receptor. The disulfide bonds were formed between Cys^{101} in TM3 and Cys^{180} in EC-loop 2 and between Cys^{18} in N-terminal tail and Cys^{274} in EC-loop 3 by using an iterative scheme in which the distance between the sulfur atoms was gradually decreased (0.5 Å one-at-a-time) by constraining the S–S distance and performing energy minimization (using conjugate gradient) with the OPLS force field to allow the rest of the protein atoms to relax and adjust.

Quality of secondary structure model

The quality of the secondary structure was investigated using the program PROCHECK [26]. Ramachandran plots [27] were calculated for each α-helix and loop separately, and for the whole protein. The quality of secondary-structure elements was tested by using the definitions introduced by Morris et al. for the allowed and disallowed regions for the ϕ and ψ angles of the backbone [28].

Preparation of ligand structures

The two ligands were built with ChemDraw and the 2D-structure was converted to a 3 D using Chem3 D. Their geometry was optimized with the built-in MM2 force field [29] of the Chem3 D program. The ligands are the peptide hormone, Ang II, and the non-peptide antagonist, losartan. Gasteiger charges were assigned to all non-peptide ligands by the AutoDock3.05 program [30, 31].

Locating the putative binding site

To locate the binding site of Ang II, and antagonists, we scanned the entire inner space of the AT_1 receptor structure by docking Ang II and the antagonists using the AutoDock3.05 program [31]. The docking procedure employs a genetic algorithm (GA) technique for exploring the conformational space by performing rapid energy evaluation using grid-based molecular affinity potentials. The AutoDock3.05 program uses the AMBER force field to describe the inter- and intra-molecular interactions in proteins [31, 32]. For ligands, the program uses the AMBER force field to estimate the parameters for bonding and nonbonding interactions and Gasteiger atomic charges to describe electrostatic interactions [30]. All amino-acid residues within 6.0 Å from the ligand surface were considered part of the putative binding site for that ligand. Using this procedure we found overlapping putative binding sites for agonists and antagonists.

MD simulation protocol

We performed 1 ns (nanosecond) MD-simulation using the OPLS force field [20] to describe the interactions between protein atoms, and the TIP4P model for water using Gromacs [22, 23]. For both ligand-free and ligand bound AT_1 receptor, we assumed that the relative position of the TM α-helices does not change due to the fact that the membrane constrains the TM α-helices together. We made this assumption because our docking experiments indicated that the α-helices have the proper position for binding the ligand (see sections: *Ang II binding site in the rat AT_1 receptor*, and *Losartan binding site in rat AT_1 receptor*) and consequently their relative positions will not change when the receptor is inserted into the membrane. In order to imitate this membrane constraint we applied a position-restrain potential (the restraining force was 5,000 kJ mol^{-1} nm^{-2}) on the backbone atoms of the α-helices using a harmonic oscillation function. The rest of the atoms (from side chains, backbone of loops and the hydrogen and oxygen atoms connected to TM helical backbone) were free to move.

The model consisting of 5,838 atoms derived from the Rh structure was solvated in a box containing 20,519 TIP4P water molecules. At physiological pH, the AT_1 receptor is positively charged, thus in order to make the simulation system electrically neutral, we added 15 chloride ions (Cl^-) to the simulation box using the "genion" tool that accompanies Gromacs. The simulation system

was set up as an NPT ensemble, i.e., constant number of particles (N), constant pressure (P) and constant temperature (T). The simulation box was coupled to a Nose–Hoover thermostat [33, 34] (300 K and a 0.1 ps (picoseconds) time constant for coupling the system to the thermostat), and a Parrinello–Rahman barostat [35] (1 bar, an isothermal compressibility (β) of 4.5 10^{-5} bar^{-1}, and a 0.1 ps time constant for coupling the system to the barostat). Periodic boundary conditions were used and the electrostatic interactions (long range) were evaluated by using the Ewald summation [36] as implemented in the PME method [37, 38] (the spacing of the Fourier grid was 1.2 Å, the order of the β-spline interpolation function was 4; the value of β—the exponent of the erfc function—is controlled in GROMACS by the parameter ewald_rtol, which represents the strength of the Coulomb interaction at the cutoff (10 Å); we used the default value of 10^{-5}). A 10 Å cutoff was used in the evaluation of van der Waals interactions (short range). The initial velocities were assigned using the Maxwell distribution at 300 K with a random seed.

The simulation protocol has the following steps:

- Adding water and ions: A simulation box large enough to accommodate the protein and minimize the protein image–image interaction, was built (using the "editconf" tool of Gromacs). The distance between the protein surface and the simulation box wall was set to 1 nm. Subsequently, TIP4P water molecules were added to the simulation box using the "genbox" tool of Gromacs. Finally, the chloride ions were added by replacing water molecules using the "genion" tool.
- Energy minimization: 100 steps of energy minimization were performed using the steepest-descent algorithm to relax the water molecules and chloride ions and thus to remove the strong initial forces due to collisions generated by the ion and water insertion process. The energy minimization is only a preparatory step in the simulation.
- Position-restrain MD: 50 ps position-restrain MD simulation were performed, in which all non-hydrogen atoms of the AT_1 receptor were restrained to their initial position using a harmonic oscillator potential (with a force constant of 1,000 kJ mol^{-1} nm^{-2}). This is a pre-equilibration step, which allows water molecules and ions to move closer to the thermodynamic equilibrium state. The use of position-restrain dynamics shortens the time required for the system to reach thermodynamic equilibrium when all atoms are free to move. The integration step was 2 fs (femtoseconds).
- α-Helices backbone position-restrain MD: A 1 ns simulation was performed on the simulation system prepared in Steps 1 to 3. All chemical bonds were constrained to their nominal values using the LINCS algorithm implemented in Gromacs [39]. The integration step was 2 fs and a simulation frame was saved every 50 fs leading to a 20,000 frames trajectory for a 1 ns simulation time. All the distance analyses (presented in the next section) were performed on this trajectory.

Results and discussion

Constructing a complete 3D structure model

As indicated in Fig. 1, only 73 residues (in red and blue) in the rat AT_1 receptor were found to be identical in bovine Rh and 102 residues were similar (i.e., either hydrophobic or hydrophilic). These included amino acids and motifs conserved in the majority of GPCRs: N46, D74, F94, G97, DRY (residues 125–127), W153, P207, Y215, F249, P255, NPXXY (residues 298–302). Thirteen-residue positions in the Rh model (in green) did not align with any residues in the AT_1 receptor. These residues were deleted one-at-a-time followed by energy minimization. The remaining 156 residues that were different in the bovine Rh were mutated one-at-a-time to residues encoding the AT_1 receptor, followed by energy minimization. In the next stage of modeling, the primary sequence of the AT_1 receptor was fully incorporated into the model. This included (i) eliminating the two gaps (Fig. 1) and (ii) insertion of AT_1 residues (Fig. 1) followed by energy minimization. In this way all 359 residues encoded in the rat AT_1 receptor cDNA have been integrated in our model.

Secondary structure quality of the AT_1 receptor model

Ramachandran-plot analysis (not shown) indicates that the seven TM α-helices of the initial AT_1 receptor model have on average 88% of their amino-acid residues in the most favored region and another 11% of the residue in the additional allowed region implying that 99% of the amino acid residues are not in any steric strain. The loops including the N- and C-terminals have on average 52% of the amino-acid residues in the most favored region and 40% of residues in the allowed region. Because loops have higher flexibility, they can have a higher range of the ϕ and ψ angles of the backbone. Including the amino acids in the generously allowed region increases the average value from 92% to 97%. The C-terminal segment has 32% of the amino-acid residues in the most favored region, 49% in the additional allowed region and 12% in the generously allowed region, which gives 93% of the amino acid residues in the allowed region. This value validates our protocol for modeling the C-terminal sequence of the AT_1 receptor.

The model of the AT_1 receptor obtained from the MD simulation is very similar to the initial model. For the TM α-helices, the average percentage of the amino acids in the allowed region increased slightly from 98.7% to 99.5%, while for loops it actually decreases from 97% to 93%. We concluded that the MD simulation validates our initial model, which has over 90% of the amino acid residues in both α-helices and loops with the ϕ and ψ angles in the allowed region of the Ramachandran map.

Orientation of critical residues in the initial model

The seven TM helices in the AT_1 receptor model are of different lengths and are also different in length from the corresponding TM helices of bovine Rh (Fig. 1). The distribution of hydrophobic/hydrophilic residues in each α-helix and their orientation with respect to the core of the protein (see Materials and methods) is shown in Fig. 2a,b. The amino-acid residues pointing toward the protein core with SASA value <0.5 nm^2 and vice versa for each α-helix is shown in Table 1. Except H3 all TM α-helices are predominantly hydrophobic in character (more than 80% of the amino acid residues are hydrophobic). This is expected because the amino-acid residues from the TM region are facing either the membrane or the protein interior, both hydrophobic in character. Thus, the high content of hydrophobic residues ensures the stability of the secondary structure in the TM region. A large percentage of hydrophilic amino-acid residues are pointing inward for, H3 (35.3%), H6 (26.9%) and H7 (21.1%), suggesting that the hydrophilic residues from these α-helices are involved either in binding the hormone or in interhelical bonding. The following residues experimentally shown to be involved in Ang II binding, H3 (Asn^{111}), H6 (His^{256}, Gln^{257}), H7 (Asn^{294}, Asn^{295}, Asn^{298}), are indeed pointing inward. Surprisingly, Lys^{199} with 0.61 nm^2 SASA was found to be oriented inward. This is in agreement with the experimental data (and our docking results shown later) indicating that Lys^{199} binds to Ang II *via* a salt-bridge [10, 11]. This analysis provided a clear understanding of which residues should face the interior of the receptor and which residues should face the membrane lipids.

Elements of secondary structure

The overall helical topology of the free-ligand AT_1 receptor is similar to that found in Rh, but close comparison indicates that position of the kinks and curvature in all TM helices are different (Fig. 2c). In helix 1, Val^{40} induced a kink and in helix 2, Thr^{80} induces the kink. His^{132} located at the bottom of helix 3 induced a kink that resulted in the tilted position of this helix. A similar histidine-induced kink is responsible for the orientation of helix 4 in bovine Rh (Fig. 2d). In helix 4, Ala^{159} rather than Pro^{162}, in helix 5, Leu^{202}/Gly^{203} rather than Pro^{207} and in helix 6, Ser^{252} rather than Pro^{255} induce the kinks. In helix 7 (Pro^{299}) the kink is proline-induced. Helix 7 sharply changes direction near the conserved Pro^{299}, part of the NCXNPXXY motif. The sequence NPXXY, which is a signature motif of the seven TM receptors, plays an important role in the agonist-induced conformational change [1, 40]. This motif may be involved in the interaction with G proteins other than G_q in the AT_1 receptor and is also proposed to interact with Ang

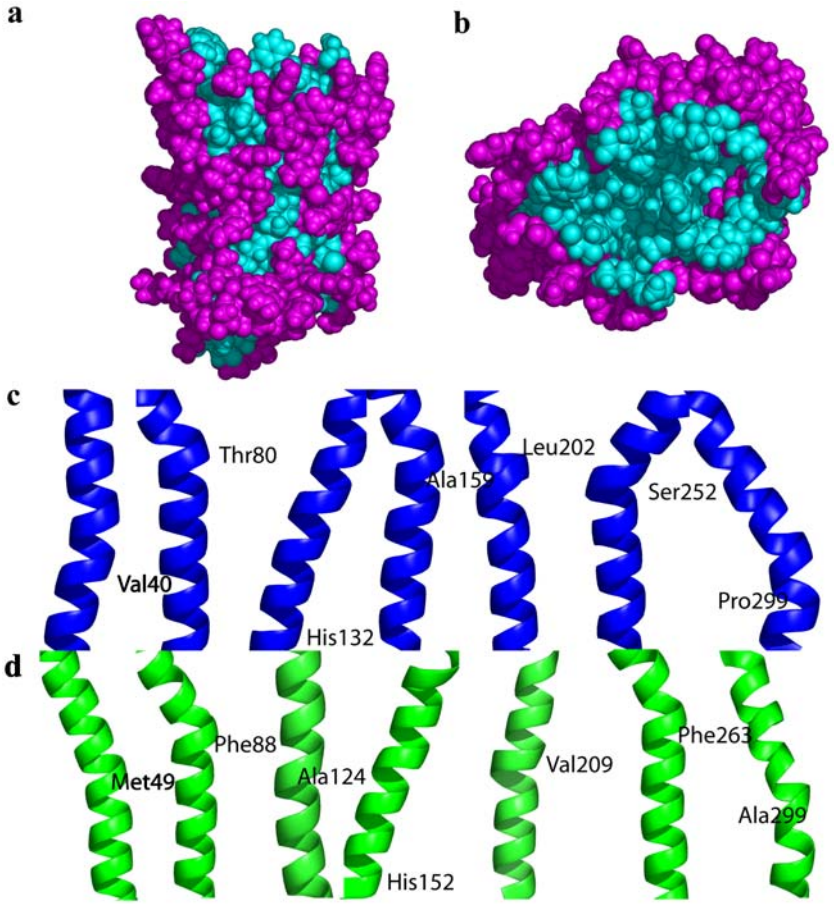

Fig. 2 Rotational orientation of hydrophobic (*magenta*) and hydrophilic (*cyan*) side chains in the rat AT_1 receptor. **a** TM view and **b** EC view. The kinks in TM helices H1–H7 that differ between **c** AT_1 and **d** Rh are shown. The most common type of helical disruption is encoded by local sequence and not necessarily by helix-breaker amino acids such as Pro

II [40]. These results suggest that deviations from α-helicity are encoded locally in the sequence patterns of the amino-acid sequence as previously described [41–43]. Marked differences in TM architecture exist even in structurally related proteins, such as bacteriorhodopsin and Rh. Such variations in TM helices are critical to our understanding of the complex structure–function relationships of the AT_1 receptor for use in future drug discovery.

Elements of tertiary structure

The AT_1 receptor contains four cysteine residues forming two disulfide bonds in the EC domain which confer the subtype-specific thiol-sensitivity to this receptor [16, 17]. Initially, the distance for Cys^{101}–Cys^{180} disulfide bond, which is equivalent to the highly conserved Cys^{110}–Cys^{187} in Rh [17], was 10.7 Å. The distance for Cys^{18}–Cys^{274} disulfide bond that is unique to the AT_1 receptor was 32.4 Å. The new sulfur-sulfur distances 2.02 Å for theCys^{101}–Cys^{180} bond and 2.0 Å for the Cys^{18}–Cys^{274} bond were obtained through an interactive procedure (see Materials and methods). Fig. 3c depicts the AT_1 receptor with the two disulfide bonds. Thiol-reducing agents abolish ligand binding to the AT_1 receptor and disruption of the two-disulfide bonds by mutation impairs protein folding and cell-surface expression [16, 17]. Single-residue mutation of Cys^{18}, Cys^{101}, Cys^{180} and Cys^{274}, individually, by glycine [16] reduces [^{125}I]-Ang II binding by 10-fold, and the double mutation Cys^{101} and Cys^{274}, 100-fold, suggesting that the AT_1 receptor structure is stabilized by the two disulfide bonds.

Figure 3 shows the 3D-structure of our homology model of the rat AT_1 receptor. This is the first computational model incorporating the entire N-terminal domain, the 7TM helices, the IC- and EC-loops, the C-terminal domain and both of the disulfide bonds. Table 2 shows the root mean square (RMS) displacements of the C_α atoms between corresponding α-helices of the AT_1 receptor and Rh. The average RMS is ~1.5 Å, but the calculated RMS for all α-helices at once is 3.2 Å, which implies that the positions of α-helices with respect to each other is different in the AT_1 receptor when compared to Rh. The differences in the relative positions of the α-helices in the two proteins resulting from the modeling process may be real, but the overall 3D-characteristics are similar in these two GPCRs.

The structure of the EC domain seems to be different in the AT_1 receptor compared to that in Rh. The β-strands in the N-tail interacting with EC-loop 1 and EC-loop 3 of Rh has been described as forming a "plug" for the covalently bound 11-*cis*-retinal chromophore pocket in Rh [2]. This structure is absent in the AT_1 receptor (see Fig. 3c,d). It is possible that this is due to the ligands entering and exiting the binding pocket as part of normal function in the AT_1 receptor. However, the end of the TM4 helix (Arg^{167}, Asn^{168}, Val^{169}, Tyr^{170}, and Phe^{171}) and the EC-loop 2

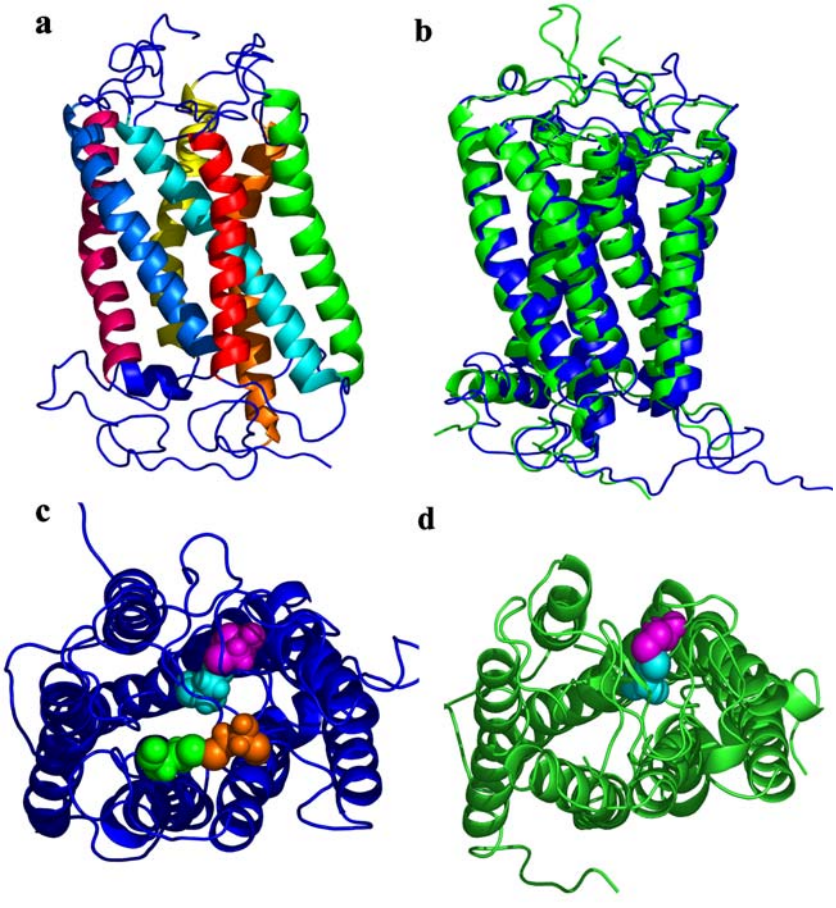

Fig. 3 Comparison of AT_1 and Rh structures. **a** The deduced 3D structure of 359 residues of the rat AT_1 receptor integrated. **b** The rat AT_1 model structure superimposed on bovine Rh crystal structure. The polypeptide back bone of 359 residues of AT_1 (*blue*) compared to 334 residues of bovine Rh (*green*) is shown in. **c** The EC-domain of AT_1 with two disulfide bonds lacks EC-plug. **d** EC-domain of bovine Rh with a single disulfide bond and the EC-plug. The conserved Cys^{101}–Cys^{180} (respectively *magenta* and *cyan*), and non-conserved Cys^{18}–Cys^{274} (respectively *orange* and *green*) of the rat AT_1 receptor, and Rh's Cys^{110}–Cys^{187} (respectively *magenta* and *cyan*) disulfide bonds

Table 2 Root mean square displacement of the TM α-helices of the AT_1 receptor and Rh

Helix	Residue range		Number of residues	RMS[a] [Å]
	AT1b	Rhc		
H1	26–55	35–64	30	0.54
H2	62–90	72–100	29	2.38
H3	99–131	107–139	33	1.19
H4	144–166	151–173	23	1.72
H5	193–217	200–224	25	1.02
H6	235–265	247–277	31	1.72
H7	284–302	287–305	19	2.13
H8	307–319	311–323	13	0.93
All helices			203	3.22

[a]Calculated with SwissPDBViewer
[b]Computational model obtained as described in the Materials and methods section
[c]Crystal structure (PDB #1F88)

residues from Ile^{172} to Thr^{190} form a hairpin that folds into the TM domain. Cys^{180} located in the center of this loop forming the conserved disulfide bond with Cys^{101} at the EC end of TM3. This suggests that the conformation of the EC-loop 2 is conserved in these two receptors.

Interhelical hydrogen-bonds, sequence motifs and polar interactions in the TM regions

In our model, there are 11 hydrogen bonds (six SC type formed between two side chains, and five SB type between a side chain and a backbone nitrogen or oxygen atom) involved in helix–helix interactions that may be essential in preserving the three-dimensional structures, as shown in other membrane proteins [44–46]. The interhelical H-bonds in our AT_1 model were found between H2–H3 (two), H2–H4 (one), H2–H7 (one), H3–H4 (three), H3–H5 (one), H3–H6 (one), H4–H5 (one), and H6–H7 (one). The interhelical hydrogen bond between Asn^{69} and Trp^{153} in AT_1 is also found in the Rh structure (Asn^{78}–Trp^{161}). The residues involved in this bond are highly conserved in GPC Rs. Other conserved residues that are involved in interhelical bonds are Asp^{74} (AT_1) and Asp^{83} (Rh), and Arg^{126} (AT_1) and Arg^{135} (Rh). Some of these are involved in the formation of three polar clamps which are formed by three different residues situated on different TM helices. A high degree of conservation of these H-bonds was observed during the entire MD simulation protocol. The SC bonds between: ND2/Asn^{111} (H3) and OD1/Asp^{74} (H2) (100% conservation during MD simulation), ND/Asn^{298} (H7) and OD2/Asp^{74} (H2) (99.98%), NH/Arg^{126} (H3) and OD1/Asn^{235} (H6) (100%), and the SB type, between the backbone N/Phe^{110} (H3) and OG/Ser^{160} (H4) (100%) were present in all MD simulation time-frames. One polar clamp is between ND2/Asn^{111} and OD1/Asp^{74} (H2), and the backbone O/Ala^{73} (H2) (69.51%). Other highly conserved (during MD simulation) SB type interhelical bonds are: between N/Phe^{110} (H3) and OG/Ser^{160} (H4) (100%), OG/Ser^{109} (H3) and NZ/Lys^{199} (H5) (92.2%), and O/Ala159 (H4) and NZ/Lys199 (H5) (94.45%).

Several motifs were found to induce helix–helix interactions in TM proteins, the GXXXG motif being the most conserved one [47–49]. The TM helices in the AT_1 receptor lack the GXXXG motif, however, other motifs like IXXXI in TM1, TM6, and TM7, AXXXA in TM2 and TM4, SXXXS in TM3, and IG in H5 are present in the TM domain. The sequence motifs, GXXXG, IXXXI, GXXXA and IG are proposed to mediate strong helix–helix associations and the motifs, AXXXA and SXXXS are thought to minimize the steric hindrance of helix backbones, which will enhance the stability of interactions between TM helices [47–49]. Six interhelical hydrogen bonds in our model involve an Asn or Asp residue and two interhelical hydrogen bonds are involving a serine: Ser^{109} from H3 is part of SXXXS motif, and Ser^{160} from H4 is part of AXXXA motif.

The AT_1 receptor has ~30% polar residues, more than normally found in TM helices of membrane proteins (~5%). The polar residues contribute to membrane-protein folding and stability [50, 51]. Involvement of strongly polar residues, Asn and Asp in interhelical interactions in AT_1 suggests a possible role for these in the stabilization of the TM domain.

Ang II binding site in the rat AT_1 receptor

The Ang II binding site predicted from the docking experiments carried out as described in Materials and methods is shown in Fig. 4. Ang II is nestled between TM helices, H2, H3, H4, H5, H6 and H7 as shown in Fig. 4a,b. The 6.0 Å binding site for Ang II in our model demonstrated that the residues determined to be critical in binding the peptide hormone by previous mutagenesis studies carried out in our group are within the binding site predicted in this study (Fig. 4c,d).

TM2—Four residues, Leu^{70}, Asp^{74}, Phe^{77} and Leu^{81} that are also part of the conserved motif of TM2 in GPCRs are predicted to be in the Ang II binding site. The model suggests that Asp^{74} makes a salt-bridge with Arg^{2} of Ang II (2.3 Å). Mutagenesis studies have revealed the importance of Asp^{74} and Phe^{77} in AT_1 receptor structure-function [52, 53].

TM3—Side chains of almost all residues from Ser^{105} through Val^{116} are located in the 6.0 Å binding site and may be involved in extensive interaction with Ang II. Asn^{111} interacts with the Tyr^{4} of Ang II (Fig. 5). The backbone oxygen of Leu^{112} And Ser^{115} make weak H-bonds with the terminal α-NH_3^+ group of Asp^{1}. The Val^{108} and Tyr^{113} interact with Pro^{7} of Ang II. Site-directed mutagenesis results suggest that several other residues, for instance, Ser^{107}, Val^{108}, Leu^{112} and Ser^{115} alter receptor affinity for analogs of Ang II [54–60].

TM4—Six residues (Met^{155}, Ala^{156}, Gly^{157}, Ala^{159}, Ser^{160} and Tyr^{170}) in this helix are predicted to interact with Ang II. Our results indicate that His^{6} of Ang II interacts with this helix. Very little experimental evidence has gathered regarding the role of TM4 in the AT_1 receptor.

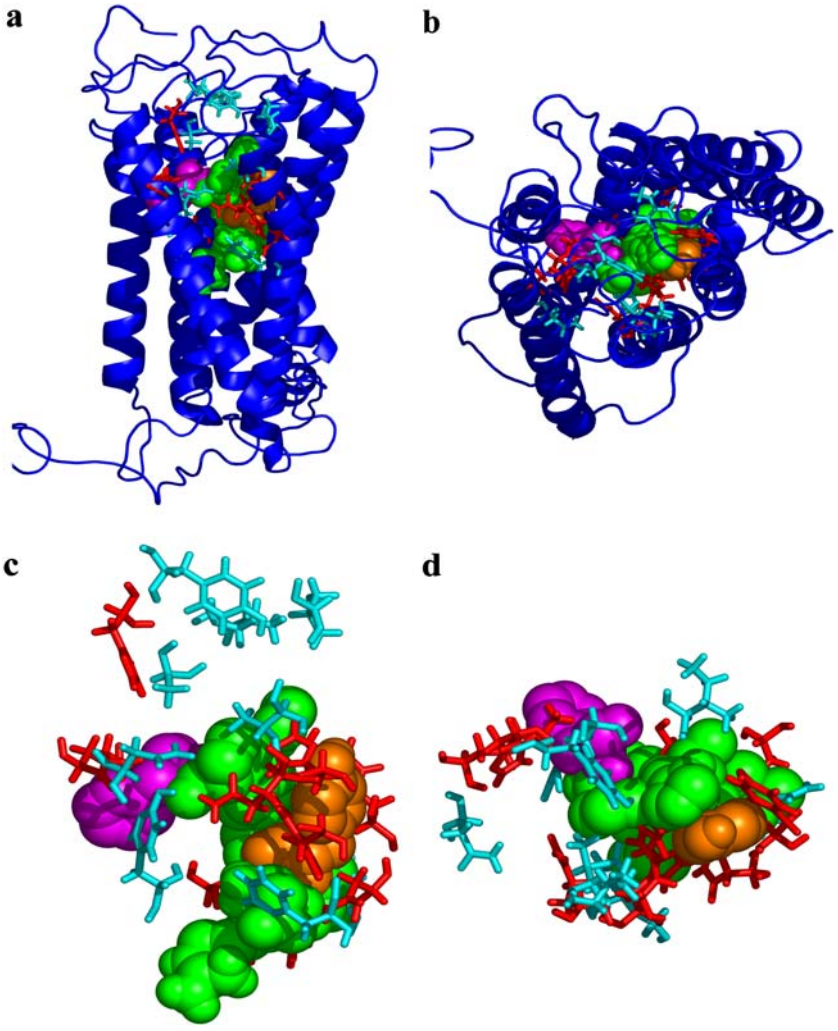

Fig. 4 The predicted hormone-binding site: residues within 6.0 Å of the Ang II binding site in the rat AT_1 receptor. **a** Complete amino acids are shown in the membrane plane view and **b** EC view. The hormone, Ang II in the binding site is shown in space filling model, with Phe^8 (*magenta*), Try^4 (*orange*) and other side chains (*green*). **c** Membrane view and **d** EC view with receptor residues in the 6.0-Å distance. Predicted side chains that were also experimentally determined to interact with specific groups of Ang II are shown in *red*. The predicted residues in *yellow* are targets for future experimental verification

TM5—The role of TM5 in binding Ang II appears to be mediated by a single residue, Lys^{199} (Fig. 5). The interaction of the α-carboxylate group of Phe^8 in Ang II with Lys^{199} is a strong salt-bridge (3.2 Å). Experimental results from Yamano et al. [10] and Noda et al. [11] suggest that Lys^{199} is involved in a pH-sensitive binding of both agonist and antagonist peptide analogs of Ang II. Aromatic acidic docking groups in the AT_1-selective nonpeptide ligands interact with Lys^{199}. It was suggested that the type of interaction of these two residues with different classes of

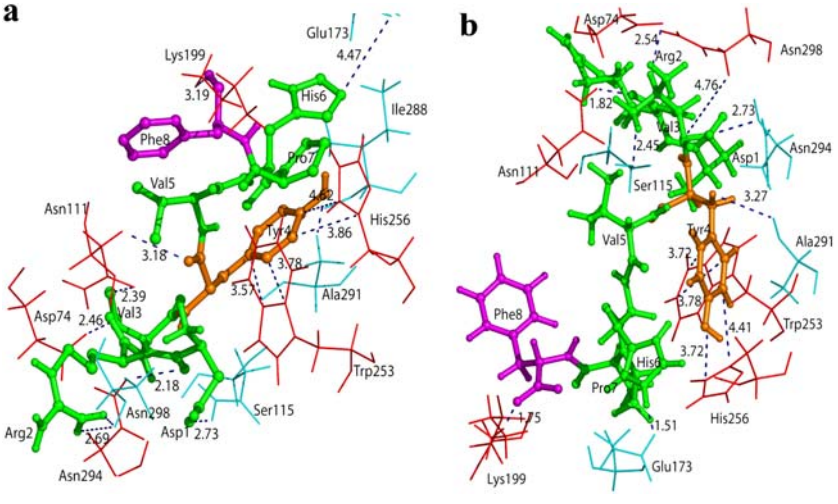

Fig. 5 **a** shows AT_1/Ang II binding complex before the MD simulation, **b** MD simulation of the AT_1/Ang II complex

AT_1 receptor-specific ligands distinguishes agonist from antagonists [11].

TM6—Five residues (Phe^{249}, Trp^{253}, His^{256}, Gln^{257} and Thr^{260}) from TM6 appear to closely interact with Ang II. Among these residues, the interaction of His^{256} with Phe^{8} (Fig. 5) known to be critical for Ang II activation is a stacked attraction of aromatic rings [13, 61, 62].

TM7—Ten different residues (Asp^{281}, Met^{284}, Ile^{288}, Ala^{291}, Tyr^{292}, Phe^{293}, Asn^{294}, Asn^{295}, Cys^{296} and Asn^{298}) in this helix appear to be in the Ang II biding pocket. Our computational results show that Asn^{294} and Asn^{298} make strong H-bonds with the Asp^{1} (Fig. 5) and the backbone oxygen of Asn^{294} and Asn^{298} make H-bonds with the Arg^{2} and Val^{3} of Ang II. Previously, mutational analysis [40, 53, 63] and ligand photocrosslinking [64] studies showed that Phe^{293} and Asn^{294} are important for binding. Residues Tyr^{292} and Asn^{295} were also implicated in Ang II activation [56].

EC-loop 2—The proximal part of this loop, harbors several residues (Ile^{172} Glu^{173}, Asn^{174}, Thr^{178}, Val^{179}, Cys^{180}, Ala^{181} and His^{183}) found in the 6.0 Å binding pocket for Ang II. For example, the model predicts a weak salt-bridge between His^{6} of Ang II and Glu^{173} (4.5 Å) and weak bonding with Phe^{8} of Ang II. Detailed mutational and photo-cross-linking studies have shown that specific residues in EC-loop 2 are important for ligand binding and interact specifically with the first and third side chains in Ang II [12, 14, 60, 65]. These findings support our prediction that EC-loop 2 is an important part of Ang II binding pocket. In addition, unpublished experimental studies in our laboratory indicate that Phe^{182} is quite important to antagonist binding (Takanobu T. and Karnik S., unpublished observations).

Our model does not predict some of the interactions suggested by mutagenesis experiments. For instance, the N-terminal region of Ang II (Asp^{1}–Arg^{2}–Val^{3}–) experimentally found [10–15, 57, 58, 64, 65] to interact respectively with His^{183}, Asp^{281} and Ile^{172} is not recapitulated in our model. All these residues are located in the interhelical loops of the EC domain. The discrepancy suggests that dynamic interaction of Ang II with the receptor is an important issue that should be considered. To account for experimental results, we suggest that AT_1 receptor activation may require an Ang II-induced change in the EC-loop conformation that would greatly improve the interactions of Ang II with the TM domain. Thus, the docking experiments with this model bring in contact the agonism-specifying residues of Ang II, Tyr^{4} and Phe^{8}, with many residues in the TM domain of AT_1 receptor known from site-directed mutagenesis to be involved in signal transduction. We postulate that Ang II binds initially to the EC-loops and exerts activating effects by inducing a motion of EC-loops as proposed earlier [1].

Losartan binding site in rat AT_1 receptor

The predicted binding site of the nonpeptide antagonist, losartan is shown in Fig. 6. Part of the predicted binding site overlaps with that of Ang II. Losartan has several distinct structural units, which are responsible for its strong interaction with the AT_1 receptor. These units are: a biphenyl subunit, a substituted imidazole (with *n*-butyl-, Cl- and -CH_2OH substituents) and a tetrazole ring (that contains four nitrogen atoms). For example, in the case of losartan the phenyl group I (C5–C10) interacts with the

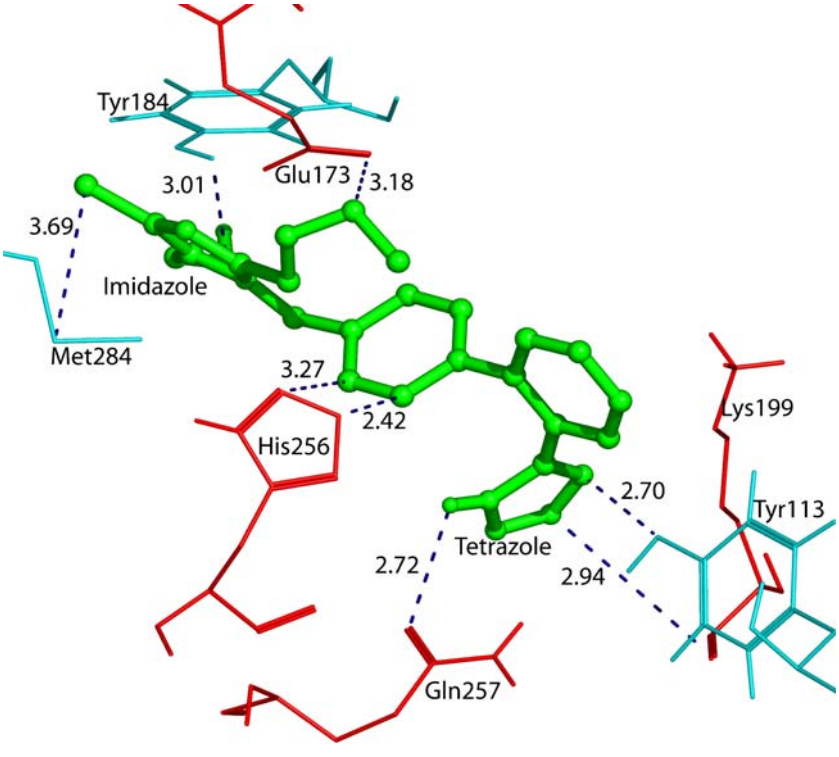

Fig. 6 The predicted binding site of losartan in the AT_1 receptor

His^{256} (TM6) *via* C10 (2.42 Å), Thr^{260} (TM6) *via* C9 (3.17 Å) and interacts weakly with Glu^{173}*via* C7 (3.18 Å) and Ala^{181}*via* C6 (3.71 Å) in EC-loop 2. Phenyl ring II (C11–C16) interacts strongly with Tyr^{113}*via* C15 (2.70 Å; TM3), and Lys^{199}*via* C14 (3.37 Å; TM5) and weakly with His^{256}*via* C11 (3.94 Å; TM6). The tetrazole ring (N4–N5–N6–C17) of losartan interacts with Tyr^{113} in TM3 *via* N6 (2.70 Å), Lys^{199}*via* N6 (2.94 Å), and Asn^{200}*via* N4 (3.42 Å) in TM5, and also with His^{256}*via* N3 (4.05 Å), and Gln^{257}*via* N5 (2.72 Å) of TM6. Finally, the substituted imidazole group of losartan (N1–C1–N2–C2–C3) interacts with the EC-loop 2 (Glu^{173}:2.4 Å/C1; Tyr^{184}:2.1 Å/C3), TM6 *via* His^{256}: 3.59 Å/C22 and Val^{264}: 2.76 Å/C18, and TM7 *via* Met^{284}: 2.66 Å/ and Ile^{288}: 2.68 Å/C20. The chlorine substituent of losartan makes a van der Waals interaction with Met^{284} (3.69 Å; TM7). The interaction between the methionine sulfur and the chlorine atom stems from the high polarizability of these two atoms (they are second row elements), which enhances their mutual interactions. The losartan binding site has been studied extensively by mutagenesis [10–15, 61, 66].

Molecular dynamics simulation of AT_1 receptor with and without the bound Ang II

The energy (*E)* and temperature (*T*) fluctuations during the 1 ns simulation for the structure model AT_1 without Ang II and with the Ang II-bound receptor are small and there is no drifting in the average values of *E* and *T*. The fluctuation in *T* is approximately 2 K around the average *T* (which is the same as the reference *T*, i.e., 300 K). Thus, the quality of the MD trajectory of the structure model assessed here is good.

Figure 7 shows a plot of the root mean square deviation (RMSD) of the extra- and intra-cellular loops during MD-simulation. Fig. 7a displays the RMSD for the free AT_1 receptor and Fig. 7b displays the ligand-bound AT_1 receptor. With the exception of the N- and C-terminals, the structure of the free and bound states of the AT1 receptor did not show significant fluctuation during simulation. The C-terminal is substantially larger and its movement periodicity could not be captured in 1 ns. Fig. 7b shows that the structure of the ligand (maroon) is well converged with a short periodicity. The N-terminal of the ligand-bound receptor also displays more dramatic movement with increased periodicity.

Distance analysis of the stability of interactions shows that the 11 interhelical bonds found in our model of the AT_1 receptor are highly stable during the entire simulation. For example, the hydrogen bonds between Asp^{74} (TM2)–Asn^{111} (TM3), Asp^{74} (TM2)–Asn^{298} (TM7), Phe^{110} (TM3)–Ser^{160} (TM4) and Arg^{126} (TM3)–Asn^{235} (TM6) last during the entire simulation. Ser^{109} (TM3)–Lys^{199} (TM5), and Ala^{159} (TM4)–Lys^{199} (TM5), hydrogen bonds which form one of the three polar clamps are highly stable. The second polar clamp between Asp^{74} (TM2), Asn^{111} (TM3) and Asn^{298} (TM7) is present in all time-frames in the trajectory. Only one interhelical hydrogen bond has a low occurrence during the simulation process: Lys^{102} (TM3)-Ala^{163} (TM4).

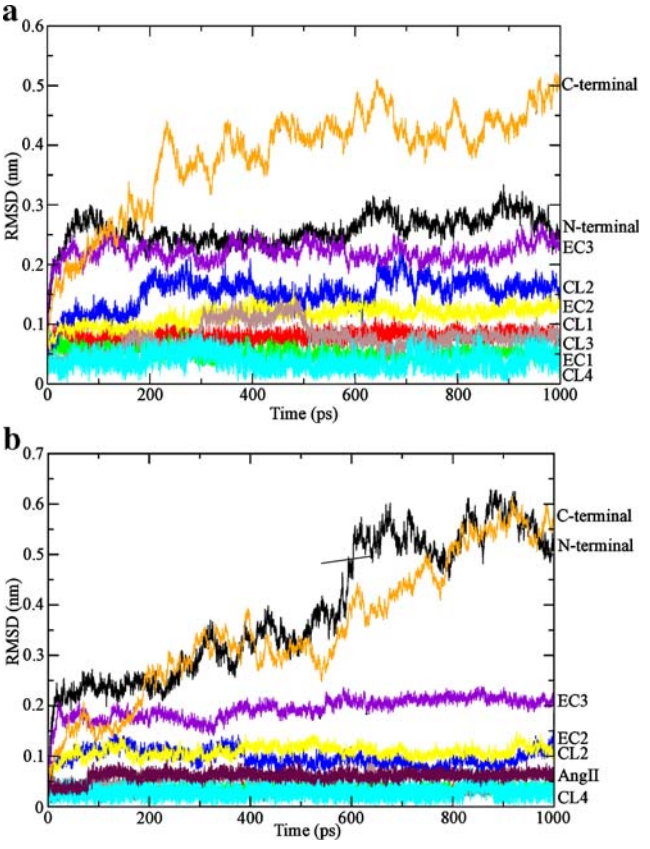

Fig. 7 RMSD of the C_α carbon atoms of the extra- and intra-cellular loops during the MD simulation: **a** free AT_1 receptor and **b** ligand-bound AT_1 receptor. Please note that plots for EC1, CL3 and CL4 are buried under other plots which are clearly visible in Fig. 7b

Variation in the surface area of the binding pocket during the MD simulation was also an important aspect in finding a structure that can better accommodate an octapeptide ligand, such as Ang II. In order to see how the pocket surface is changing during simulation, the fluctuation in distances between five residues, Asn^{111}, Lys^{199}, His^{256}, Gln^{257}, and Asn^{294} (Asn^{111}–Lys^{199}, Lys^{199}–His^{256}, His^{256}–Gln^{257}, Gln^{257}–Asn^{294}, Asn^{294}–Asn^{111}), known to be involved in Ang II binding from experimental data published from different laboratories including ours, were investigated. We selected 10 different time-frame structures and used them in the docking analysis.

The Ang II bound AT_1 model MD simulation was analyzed for distance changes in amino acid residues of the binding pocket in 11 time-frames. The simulation did not affect the accommodation of the ligand in the binding pocket in the initial docking complexes selected (see Fig. 5). For example, the salt bridge that binds Lys^{199} to α-COO^- group of Phe^8 in Ang II is 3.20 Å in the AT_1/Ang II complex before the MD simulation (Fig. 5a), 1.75 Å in the AT_1/Ang II simulated complex (Fig. 5b), and 3.28 Å in the AT_1 (after MD simulation)/Ang II docking complex. The aromatic attraction between His^{256} and Tyr^4 of Ang II is 3.97/4.24 Å in the AT_1/Ang II complex before the MD

simulation was performed (Fig. 5a), 3.62/3.98 Å in the AT_1/Ang II simulated complex (Fig. 5b), and 2.96/2.97 Å in the AT_1 (after MD simulation)/Ang II docking complex. The aromatic attraction between Trp^{253} and Tyr^4 of Ang II is 1.90/2.63 Å in the AT_1/Ang II complex before the MD simulation was performed (Fig. 5a), 4.58/4.72 Å in the AT_1/Ang II simulated complex (Fig. 5b), and 3.78/4.60 Å in the AT_1 (after MD simulation)/Ang II docking complex. The distance between Asn^{111} and Val^3 in Ang II is 3.09 Å in the AT_1/Ang II complex before the MD simulation (Fig. 5a), 4.19 Å in the AT_1/Ang II simulated complex (Fig. 5b), and 4.02 Å in the AT_1 (after MD simulation)/Ang II docking complex.

Conclusion

This is the first GPCR-modeling study in which all of the amino acids encoding the receptor and the disulfide bonds known to be important for receptor conformation have been taken into consideration. The 3D-models of the AT_1 receptor complexes with Ang II, as well as losartan, were obtained only through steric and energy considerations, without using the knowledge of the experimental site-directed mutagenesis results. The site-directed mutagenesis data available were used for independent validation of the predicted binding, which together with the results from the Ramachandran map analysis and MD-simulation suggest that the model presented here is the most plausible. The 3D-model of Ang II complex with the AT_1 receptor shows detailed interaction of the residues of Ang II with many residues of the AT_1 receptor experimentally shown to be involved in agonist binding and signal transduction. Our model suggests additional site-directed mutagenesis studies to test the predicted structure and function of the AT_1 receptor. Our structure can also be used to predict new ligands that would bind specifically to the AT_1 receptor and display better selectivity to block its functions. The current structural model of the inactive state in Rh combined with functional evidences in Rh and other GPCRs clearly suggests that the EC-loops play a critical role in ligand binding as well as activation-induced conformational dynamics (*1*). Furthermore, it is currently believed that the extra-cellular disulfide bonds play a critical role in the transmission of activation-induced conformational changes to the cytoplasmic domain that leads to binding and activation of G-proteins (*1*). Our AT_1 receptor model was further investigated by MD-simulation, in which the influences of interhelical loops on both sides of the membrane are included. Our future goal is to carry out new molecular-dynamics simulations using the current model to decipher structural consequences of mutations and of agonist and antagonist binding to the receptor. Such experiments will be useful to refine the predicted structure further, as well as enhance understanding of the function, and should also be useful for predicting the function of orphan GPCRs. The pdb file with the Cartesian coordinates will be deposited in the Protein Data Bank.

Acknowledgements We thank Dr. Valentin Gogonea, Cleveland State University for useful comments on the manuscript, the members of the Karnik laboratory for valuable discussions and Robin Lewis for manuscript preparation. The R01 (HL57470) grant award to SK from the National Institutes of Health supported this work and SK is an established Investigator of the American Heart Association.

References

1. Karnik SS, Gogonea C, Patil S, Saad Y, Takezako T (2003) Trends Endocrinol Metab 14:431–437
2. Palczewski K, Takashi K, Tetsuya H, Behnke C, Motoshima H, Fox B, Le Trong I, Teller D, Okada T, Stenkamp R, Yamamoto M, Miyano M (2000) Science 289:739–745
3. de Gasparo M, Catt KJ, Inagami T, Wright JW, Unger T (2000) Pharmacol Rev 52:5415–5472
4. DeLano WL (2005) Drug Discov Today 10:213–217
5. Henderson R, Baldwin JM, Ceska TA, Zemlin F, Beckmann E, Downing KH (1990) J Mol Biol 213:899–929
6. Joseph MP, Maigret B, Bonnafous JC, Marie J, Sheraga HA (1995) J Protein Chem 14:381–398
7. Yamato Y, Ohyama K, Kikyo M, Sano T, Nakagomi Y, Inoue Y (1995) J Biol Chem 1270:14024–14030
8. Nikiforovich GV, Marshall GR (2001) Biochem Biophys Res Comm 286:1204–1211
9. Deraet M, Rihakova L, Boucard A, Perodin J, Sauve S, Mathieu AP, Guillemette G, Leduc R, Lavigne P, Escher E (2000) Can J Physiol Pharmacol 80:418–425
10. Yamano Y, Ohyama K, Chaki S, Guo DF, Inagami T (1992) Biochem Biophys Res Commun 187:1426–1431
11. Noda K, Saad Y, Kinoshita A, Boyle TP, Graham RM, Husain A, Karnik SS (1995) J Biol Chem 270:2284–2289
12. Feng YH, Noda K, Saad Y, Liu X, Husain A, Karnik SS (1995) J Biol Chem 270:12846–12850
13. Noda K, Saad Y, Karnik SS (1995) J Biol Chem 270:28511–28514
14. Schambye HT, Hjorth SA, Bergsma DJ, Sathe G, Schwartz TW (1994) Proc Natl Acad Sci USA 91:7046–7050
15. Ji H, Leung M, Zhang Y, Catt KJ, Sandberg K (1994) J Biol Chem 269:16533–16536
16. Ohyama K, Yamano Y, Sano T (1995) Regul Pept 57:141–147
17. Feng YH, Saad Y, Karnik SS (2000) FEBS Lett 484:133–138
18. Karnik SS, Khorana HG (1990) J Biol Chem 256:17520–17524
19. Thompson JD, Higgins DG, Gibson TJ (1994) Nucleic Acids Res 22:4673–4680
20. Jorgensen W, Tirado-Rives J (1988) J Am Chem Soc 110:1657–1666
21. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) J Am Chem Soc 118:11225–11236
22. Berendsen HJC, van der Spoel D, van Drunen R (1995) Comp Phys Comm 91:43–56
23. Lindahl E, Hess B, van der Spoel D (2001) J Mol Model 7:306–317
24. Black SD, Mould DR (1991) Anal Biochem 193:72–82
25. Guex N, Pietsch MC (1997) Electrophoresis 18:2714–2723
26. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) J Appl Cryst 26:283–291
27. Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) J Mol Biol 7:95–99
28. Morris AL, MacArthur MW, Hutchinson EG, Thornton JM (1992) Proteins 12:345–364
29. Allinger NL (1977) J Am Chem Soc 99:8127–8134
30. Gasteiger J, Marsili M (1980) Tetrahedron 36:3219–3228
31. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) J Comput Chem 19:1639–1662
32. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KMJ, Ferguson DM, Spellmeyer DC, Fox T, Caldwell, JW, Kollman PA (1995) J Am Chem Soc 117:5179–5197
33. Nose S (1984) Mol Phys 52:255–268

34. Hoover WG (1985) Phys Rev A31:1695–1697
35. Parrinello M, Rahman A (1981) J Appl Phys 52:7182–7190
36. Ewald PP (1921) Ann Phys 64:253–287
37. Darden T, York D, Pedersen L (1993) J Chem Phys 98:10089–10092
38. Esmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) J Chem Phys 103:8577–8592
39. Hess B, Bekker H, Berendsen HJC, Fraaije J (1997) J Comp Chem 18:1463–1472
40. Hunyady L, Bor M, Baukal AJ, Balla T, Catt KJ (1995) J Biol Chem 270:16602–16609
41. Riek RP, Rigoutsos I, Novotny J, Graham RMJ (2001) Mol Biol 306:349–362
42. Sankararamakrishnan R, Vishveshwara S (1990) Biopolymers 30:287–298
43. Sansom MS (1992) Protein Eng 5:53–60
44. Zhou FX, Cocco MJ, Russ WP, Brugner AT, Engelman DM (2000) Nat Struct Biol 7:154–160
45. Adamian L, Liang J (2002) Proteins 47:209–218
46. Eilers M, Shekar SC, Shieh T, Smith SO, Fleming PJ (2000) Proc Natl Acad Sci USA 97:5796–5801
47. Tourasse NJ, Li W-H (2000) Mol Biol Evol 17:656–664
48. Melnyk RA, Kim S, Curran AR, Engelman DM, Bowie JU, Deber CM (2004) J Mol Biol 279:16591–16597
49. Dawson JP, Weinger JS, Engelman DM (2002) J Mol Biol 316:799–805
50. Zhou FX, Merianos HJ, Russ WP, Brugner AT, Engelman DM (2001) Proc Natl Acad Sci USA 98:2250–2255
51. Gratkowski H, Lear JD, DeGrado WF (2001) Proc Natl Acad Sci USA 98:880–885
52. Bihoreau C, Monnot C, Davies E, Teutsch B, Bernstein KE, Corvol P, Clauser E (1993) Proc Natl Acad Sci USA 190:5133–5137
53. Miura S, Zhang J, Boros J, Karnik SS (2003) J Biol Chem 278:3720–3725
54. Noda K, Feng YH, Liu XP, Saad Y, Husain A, Karnik SS (1996) Biochemistry 35:16435–164424
55. Groblewski T, Maigret B, Larguier R, Lombard C, Bonnafous JC, Marie J (1997) J Biol Chem 272:1822–1826
56. Balmforth AJ, Lee AJ, Warburton P, Donnelly D, Ball SG (1997) J Biol Chem 272:4245–4251
57. Feng YH, Miura S, Husain A, Karnik SS (1998) Biochemistry 37:15791–15798
58. Miura S, Feng YH, Husain A, Karnik SS (1999) J Biol Chem 274:7103–7110
59. Monnot C, Bihoreau C, Conchon S, Curnow KM, Corvol P, Clauser E (1996) J Biol Chem 271:1507–1513
60. Parnot C, Bardin S, Miserey-Lenkei S, Guedin D, Corvol P, Clauser E (2000) Proc Natl Acad Sci USA 97:7615–7620
61. Schambye HT, Hjorth SA, Weinstock J, Schwartz TW (1995) Mol Pharmacol 47:425–431
62. Han HM, Shimuta SI, Kanashiro CA, Oliveira L, Han SW, Paiva AC (1998) Mol Endocrinol 12:810–814
63. Boucard AA, Roy M, Beaulieu ME, Lavigne P, Escher E, Guillemette G, Leduc R (2003) J Biol Chem 278:36628–36636
64. Perodin J, Deraet M, Auger-Messier M, Boucard AA, Rihakova L, Beaulieu ME, Lavigne P, Parent JL, Guillemette G, Leduc R, Escher E (2002) Biochemistry 41:14348–14356
65. Boucard AA, Wilkes BC, Laporte SA, Escher E, Guillemette G, Leduc R (2000) Biochemistry 39:9662–9670
66. Takezako T, Gogonea C, Saad Y, Noda K, Karnik SS (2004) J Biol Chem 279:15248–15257