Homology modeling, molecular dynamics and docking simulations of rat A2A receptor: A three-dimensional model validation under QSAR studies

Article · December 2013

READS
108

6 authors, including:

Marvin A Soriano-Ursúa
Instituto Politécnico Nacional
29 PUBLICATIONS **156** CITATIONS
SEE PROFILE

Omar Deeb
Al-Quds University
38 PUBLICATIONS **375** CITATIONS
SEE PROFILE

Aldo Segura-Cabrera
Institute of Ecology INECOL
24 PUBLICATIONS **118** CITATIONS
SEE PROFILE

José Correa-Basurto
Instituto Politécnico Nacional
140 PUBLICATIONS **636** CITATIONS
SEE PROFILE

Current Trends in Medicinal Chemistry

Original Communication

Homology modeling, molecular dynamics and docking simulations of rat A_{2A} receptor: A three-dimensional model validation under QSAR studies

Luis C. Jimenez-Botello[1], Marvin A. Soriano-Ursúa[2], Omar Deeb[3], Aldo Segura-Cabrera[4], M. Martínez-Archundia[2] and J. Correa-Basurto[2,*]

[1]Escuela de Medicina, Departamento de Ciencias Basicas, Escuela de Ciencias de la Vida ITESM Campus Ciudad de México, Calle del puente 222, Colonia Ejidos de Huipulco, Del. Tlalpan, 14380, México, D. F., [2]Laboratorio de Modelado Molecular y Bioinformática de la Escuela Superior de Medicina, Instituto Politécnico Nacional. Plan de San Luis y Díaz Mirón, Del. Miguel Hidalgo, 11340, México City, México. [3]Faculty of Pharmacy, Al-Quds University, Jerusalem, Palestine. [4]Laboratorio de Bioinformática, Centro de Biotecnología Genómica, Instituto Politécnico Nacional, Boulevard del Maestro esquina Elías Piña, Colonia Narciso Mendoza, 88710, Ciudad Reynosa, Tamaulipas, México

ABSTRACT

Understanding the three-dimensional structure (3-D) of GPCRs (G protein coupled receptors) can aid in the design of applicable compounds for the treatment of several human disorders. To this end, several 3-D models have been obtained in recent years. In this work, we have built the rat adenosine receptor model ($rA_{2A}R$) by employing computational tools. First, the 3-D $rA_{2A}R$ model was built by homology modeling using the human adenosine receptor ($hA_{2A}R$) structure (PDB codes: 3EML) as a template. Then, the $rA_{2A}R$ model was refined by molecular dynamics simulations, in which the initial and refined 3-D structures were used for molecular docking simulations and Quantitative structure-activity relationship (QSAR) studies using a set of known experimentally tested ligands to validate this rA2AR model. The results showed that the hindrance effect caused by ribose attached to agonists play an important role in activating the receptor via formation of several hydrogen bonds. In contrast, the lack of this moiety allows blocking of the receptor. The theoretical affinity estimation shows good correlation with reported experimental data. Therefore, this work represents a good example for getting reliable GPCR models under computational procedures.

KEYWORDS: A_{2A} adenosine receptor, homology modeling, xanthine, drug development, rat brain, Parkinson's disease

1. INTRODUCTION

The A_{2A} adenosine receptor ($A_{2A}R$) is a seven-transmembrane domain receptor (7TM) that belongs to the family A of GPCRs (G protein coupled receptors). This receptor has been implicated in several physiological functions, including some diseases of the central nervous system (e.g., Parkinson's disease) [1]. This adenosine receptor can form heteromeric complexes with D_2-dopamine receptors or with metabotropic glutamate receptors, and it is also known that this 7TM activate proteins involved in intracellular signaling [2]. The $A_{2A}R$ can be blocked or activated by compounds that share very similar

*Corresponding author: jcorreab@ipn.mx

chemical cores (scaffolds). In fact, some xanthine-containing antagonists are used to decrease the motor symptoms in Parkinson's disease (PD) [2, 3]. Additionally, other studies report the use of $A_{2A}R$ agonists to treat hypertension, ischemic cardiomyopathy, inflammation and atherosclerosis [4]. These evidences have provided incentive to study and to develop more selective $A_{2A}R$ ligands [5]. To reach this goal, it is useful to employ the murine models used in drug development for PD treatments and for other diseases of the central nervous system [4-7].

Currently, drug development research explores the ligand and protein targets at an atomic level by constructing three-dimensional (3-D) models of these membrane receptors, whose structures are very difficult to crystallize [8, 9]. This difficulty is related to the location of these proteins inside the lipid bilayer membrane; that is why only few 3-D structures have been solved for these proteins using X-ray methods.

Due to the previously mentioned difficulties, computational tools are widely used to obtain 3-D models of GPCRs, for example by computing *ab initio* calculations, sequence threading and homology modeling. The homology modeling utilizes experimental data obtained from the Protein Data Bank (PDB) named "template". Initially, some reports have used the bovine rhodopsin (the first GPCR characterized) as a template. However, this template yields poor quality structures for the $A_{2A}R$ 3-D models. This is due to distant homology relationship between bovine rhodopsin and other GPCR targets. Hence, bovine receptor would not be an adequate template to construct the $rA_{2A}R$ model [4]. As the structures of $hA_{2A}R$ bound to antagonist [9] and agonist [10] have recently been solved by X-ray methods, it is possible to use both structures as templates for additional adenosine GPCR models to study the molecular recognition properties. Currently, several studies support the use of 3-D structures of $A_{2A}R$ as experimental and theoretical tools for the development and testing of new drugs [11-15].

According to experimental data, there are several compounds that act on $A_{2A}R$ receptors [1, 4-7]. Some ligand-$A_{2A}R$ interactions and affinities have been reported from *in vitro* experiments and from mathematical studies on animal tissues [4, 8, 16-20]. However, some of these ligands have not been studied using theoretical methods and therefore do not explain the ligand recognition on $A_{2A}R$ structure at an atomic level.

The aim of this study was to generate an adenosine receptor model from rat ($rA_{2A}R$), an animal widely used for experimental and preclinical studies. We refined and validated this model constructed by homology modeling. The refinement of this model was through molecular dynamics (MD) simulations, whereas the 3-D model validation was using molecular docking simulations and QSAR studies.

2. Computational procedures

2.1. Ligand retrieval

A set of 48 known $A_{2A}R$ ligands, consisting of 15 antagonists (Table 1) and 33 agonists (Table 2), were docked on all $A_{2A}R$ 3-D models (i.e., the $rA_{2A}R$, the native $hA_{2A}R$ 3-D structures as well as five snapshots retrieved from MD simulations of $rA_{2A}R$). The minimum ligand structure energies were obtained via DFT at the B3LYP/6-31G (d,p) level using Gaussian 98 software [21].

2.2. Homology modeling of the $A_{2A}R$

The amino acid sequence of the $rA_{2A}R$ with accession number NP_445746 was retrieved from NCBI [22]. The Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) found several homologous human proteins with known 3-D structures located in the Protein Data Bank (PDB). To obtain the $rA_{2A}R$ 3-D model, the Swiss Model server [23-25] was used to perform the homology modeling procedure using the $hA_{2A}R$ 3-D structure as a template (PDB code: 3EML) [10].

First, all hydrogen atoms on the $A_{2A}R$ model were added and then minimized in 500 steps with the steepest descendent algorithm, which is employed by GROMOS96 43B1 parameters implemented in the Swiss-PDB Viewer version 3.7. We used the TM-score [26] and TM-align tools [27] (available from the Zhang lab at http://zhang.bioinformatics.ku.edu) to compare the two 3-D models employed in this study. The backbone conformation of both $A_{2A}R$ models was evaluated by Psi/Phi

Table 1. Experimental and calculated affinity values from docking analyses for antagonist ligands tested.

Antagonist ligand	Experimental [a] K_d reported in literature (nM)		Calculated affinity values on the respective models (nM)	
			$hA_{2A}R$	$rA_{2A}R$
KW6002[a]	2	[8]	463	1080
XAC[a]	24	[8]	73.3	135
11ad[b]	14	[20]	1040	2510
11ae[b]	17	[20]	771	2260
11aa[b]	18	[20]	407	1400
11bd[b]	21	[20]	729	819
11ab[b]	22	[20]	422	1010
11ac[b]	25	[20]	1020	1200
11ba[b]	29	[20]	276	432
11af[b]	30	[20]	868	2320
11bc[b]	32	[20]	1320	938
11bb[b]	33	[20]	965	932
DPCPX[a]	340	[8]	1440	1770
Theophylline[a]	1995.26	[16]	16900	17700
Caffeine[a]	2511.89	[16]	30300	36800

[a] From rat striatal membranes. [number] = reference
[b] From pig membranes. [number] = reference

Ramachandran diagrams using the Rampage server [28]. Then, to validate our $rA_{2A}R$ model built by HOMOLOGY MODELING, docking simulations were performed using a set of known ligands (agonists and antagonists).

2.3. Molecular dynamics simulations

As $rA_{2A}R$ is one of the principal targets in this report, we performed MD simulations to refine and to provide evidence of its conformational movements. These simulations were performed by using the GROMACS software package, 4.5.3 version [29]. Structural data for a pre-equilibrated 1-palmitoyl-2-oleoyl-phosphatidylcholine (POPC) bilayer with 128 lipid molecules were downloaded from the Tieleman group (http://moose.bio.ucalgary.ca/index.php?page=Structures_and_Topologies). The protein was then embedded in the pre-equilibrated POPC lipid bilayer using the perl inflategro script from Tieleman's group, performing several rounds of energy minimization. Furthermore, after 23 iterations of scaling down by 0.95, an area of ~74 $Å^2$ per lipid was reached. This value was greater than the experimental value of ~66 $Å^2$ because the inflategro script tends to overestimate the area per lipid [30]. This calculated value, however, was sufficient to continue to the equilibration step.

All systems were equilibrated using simulated annealing under an isothermal-isobaric (NPT) ensemble for 500 ps. The LINCS method [31] was used to restrain all of the heavy atoms in all directions, and this method was also used to restrain the phosphorus atoms of the lipid head groups in the vertical (z) direction, allowing a 2 fs integration step. The use of simulated annealing under NPT ensemble avoided solvent voids that distort the dimensions of the unit cell.

Following simulated annealing, NPT equilibration was performed for 500 ps applying a pressure of 10 MPa in the transverse direction and 0.1 MPa in the vertical direction. Water, lipids and protein were coupled separately to a temperature bath at

Table 2. Experimental and calculated affinity values from docking analyses for agonist ligands tested.

Agonist ligand	Experimental [a] K_d reported in the literature (nM)		Calculated affinities values on the respective models (nM)	
			$hA_{2A}R$	$rA_{2A}R$
19g	1.23	[17]	300	42.3
2g	1.43	[17]	25.6	21.6
23g	1.77	[17]	158	204
3g	2.38	[17]	72.6	412
13g	2.41	[17]	32.1	119
25g	2.42	[17]	51.9	66.3
9g	2.45	[17]	126	452
22g	2.75	[17]	252	303
5g	2.83	[17]	11.8	298
17g	2.87	[17]	461	19800
21g	2.90	[17]	19	2240
11g	2.95	[17]	182	184
12g	2.96	[17]	112	139
8g	3.02	[17]	297	75.6
7g	3.07	[17]	1170	367
15g	3.20	[17]	148	250
29g	3.22	[17]	6.65	21.8
6g	3.25	[17]	270	556
16g	3.27	[17]	229	114
4g	3.29	[17]	104	331
28g	3.37	[17]	5.96	11.6
20g	3.40	[17]	36.9	89.7
14g	3.42	[17]	138	377
10g	3.47	[17]	87.4	484
1g	3.55	[17]	56.1	24.8
26g	3.85	[17]	5.66	6.64
27g	4.69	[17]	7.38	1530
NECA	20	[17]	1180	1210
CGS21680	24	[16]	41.7	43.7
CADO	180	[3]	1120	1310
RPIA	794	[16]	76.4	83.6
CPA	794.33	[3]	230	186
Adenosine	2511.89	[16]	2930	1630

[a] From rat striatal membranes, except for CPA (which is from $hA_{2A}R$ expressed in CHO cells).
[number] = reference

300 K with a relaxation time (τT) of 0.2 ps using a Berendsen thermostat [32]. Each group (protein/cholesterol, lipids, and solvent/ions) was coupled to a separate temperature bath. The parameters developed by Berger *et al*. [33] were applied to the POPC lipids, and the gromos96 53a6 parameter set was used to describe the rest of the systems (protein, solvent and ions). Lennard-Jones interactions were cut off at 1.4 nm, and short-range, non-bonded interactions were calculated with a twin-range cut off scheme (0.9/1.4 nm) updating the neighbor list every five simulation step. Electrostatic interactions were calculated using the particle mesh Ewald algorithm using a fourth-order spline interpolation and a Fourier grid spacing of 0.12 nm. This treatment of electrostatics has been shown to provide an accurate representation of lipid properties [34], and it is also commonly used in simulations of proteins. Following 1000 ps of equilibration, production of this MD simulation was conducted using an NPT ensemble. A pressure of 0.1 MPa was applied in all directions, and all other parameters were the same as those used in the NPT equilibration. All position restraints were removed prior to the production phase. Simulations were conducted using the Argentum cluster at the Centro Nacional de Supercomputo. Coordinates were saved every 2 ps for analysis. All analyses were performed using tools included in the GROMACS package. Subsequently, the dynamic behavior and structural changes of the protein were analyzed by the calculation of the root mean square fluctuations (RMSF) and the root mean square deviations (RMSD) values. The refined model was evaluated using the ANOLEA [35] and ProSA [36] programs.

2.4. Docking simulations

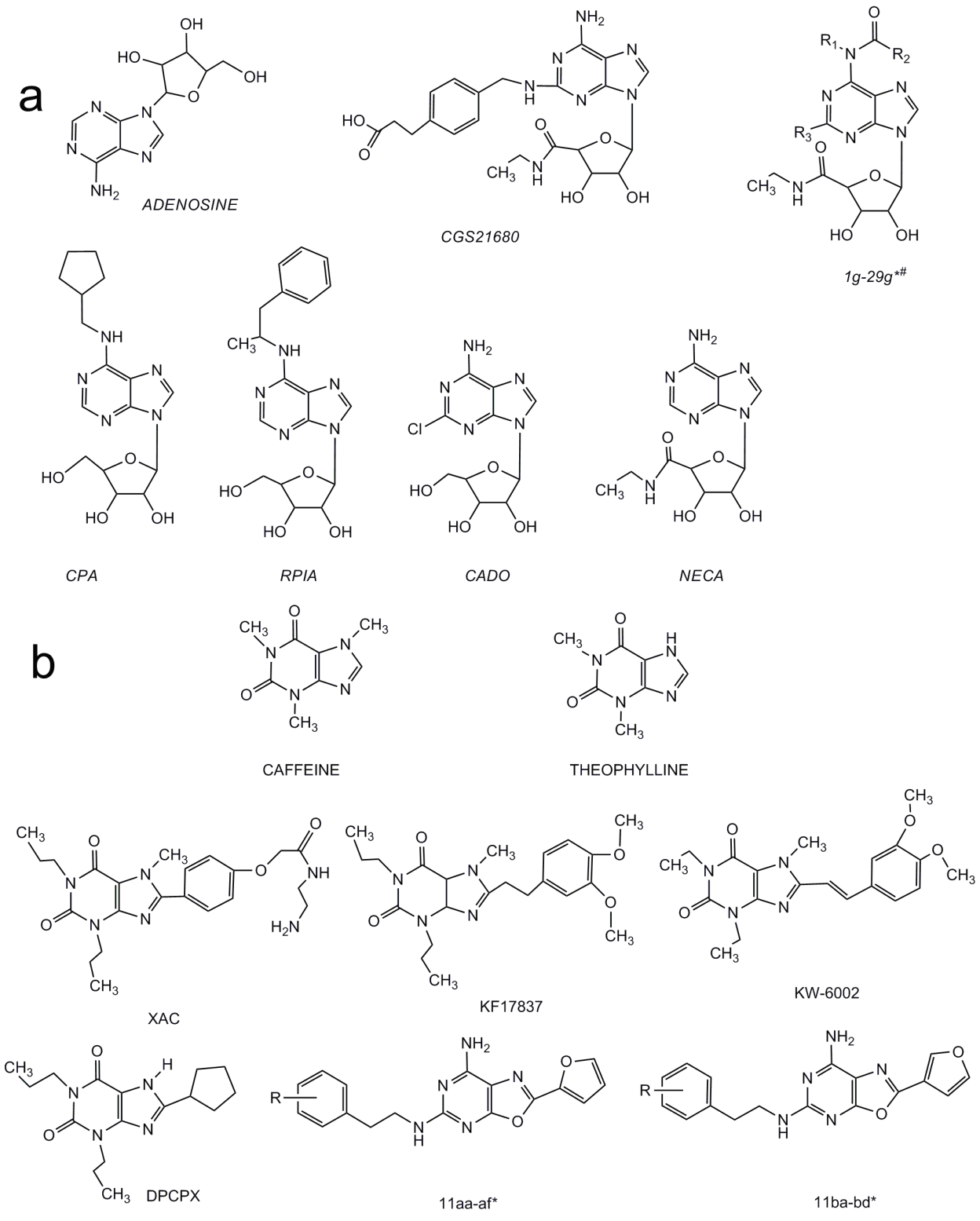
The $rA_{2A}R$ and $hA_{2A}R$ 3-D native structures and snapshots retrieved from the $rA_{2A}R$ MD simulations (taken at 0, 5, 10, 15 and 20 ns) were used to perform docking simulations using Autodock 4.0.1. For this procedure, we used the Lamarckian genetic algorithm, which is a combination of the genetic algorithm method for global and local search to perform energy minimization [37]. Before starting the docking evaluations, the partial atomic charges (Gasteiger-Marsili formalism), all possible routable bonds of the ligands, Kollman charges for $A_{2A}R$ atoms and the hydrogen atoms located in heteroatoms were added at ~ pH 7.4 with the AutoDock Tools 1.5.2 [38].

A cubic grid box (80 x 80 x 80 Å) with grid points separated by 0.375 Å was constructed and centered at a midpoint between the coordinates corresponding to the amino acids in positions Leu85 and Leu249 (3.33 and 6.51, respectively) for $hA_{2A}R$, according to the Ballesteros-Weinstein designation [39]. These amino acids are considered to be the conserved key positions for ligand-interactions in other mammalian species [10, 15, 40]. All docking simulations were performed with an initial population of 100 randomly placed individuals and a maximum number of energy evaluations (1.0×10^7).

The resulting docked orientations within a RMSD of 0.5 Å were clustered together. The lowest energy cluster returned by AutoDock for each compound was used for further analysis. All other parameters were maintained considering the default settings. All of the docking result visualizations were performed by using the Visual Molecular Dynamics (VMD) program version 1.6 [41].

2.5. Molecular descriptors

The 48 known $A_{2A}R$ ligands (Tables 1, 2 and Scheme 1) employed in the docking protocol were tested. The ligand structures were structurally optimized by DFT at the B3LYP/6-31$G_{,d,p}$ level using Gaussian 98 software [21]. Then, the Dragon software was used to calculate 1481 descriptors grouped into 16 groups. In addition, the calculated electronic descriptors obtained from HyperChem Version 7.0 (Hypercube, USA, http://www.hyper.com) were considered as another group. Therefore, 17 groups of descriptors were used in this study. In each group, the calculated descriptors were examined for the presence of constant or near constant values for all molecules. To decrease redundancy in the descriptors data matrix, the correlation among descriptors was examined, and the detected collinear descriptors (i.e., $R^2 > 0.95$) were removed from the data matrix. Accordingly, 589 descriptors were calculated.



Scheme 1. The set of A_{2A} ligands tested (a = agonist and b = antagonists). * Only the pharmacophore of these ligands is depicted. The full structure of this ligand is in references 17 or 20 for 1g-29g and 11aa-af and 11ba-bd. Ligands 18g and 24g are not included.

Table 3. Descriptors selected by MLR for each group.

Group[a]	Descriptors
G1	C-042, C-016, H-050, H-051, C-040, O-059
G2	TI2, VRA1, PJI2, X1A, Ram, IC1, piID
G3	MWC05, SRW09, MWC10, SRW05
G4	BEHm7, BELe8, BELm2, BEHe1, BELv1
G5	JGI10, GGI10, GGI4, JGI4
G6	MATS3p, MATS8v, GATS4e, MATS4v, MATS5v, MATS3e, MATS2p, ATS1e
G7	RNCG, RPCG
G8	AROM, RCI
G9	SHP2
G10	FDI, SPAN
G11	RDF065p, RDF010m, RDF035u, RDF045m, RDF055v
G12	Mor20u, Mor25u, Mor05u, Mor12u, Mor13m, Mor32m, Mor18u
G13	G1v, Dv, E2s, E3u
G14	HATS4u, R5v+, R4m, H3m
G15	nHAcc, nOHp, n=CHR, nCONR2, nNHRPh, nPhX, nCs, nCONHRPh
G16	C-042, C-016, H-050, H-051, C-040, O-059
G17	Ref, HD, ELPH

[a]G1: constitutional descriptors, G2: topological descriptors, G3: molecular walk counts, G4: BCUT descriptors, G5: Galvez topological charge indices, G6: 2D autocorrelations, G7: charge descriptors, G8: aromaticity index, G9: Randic molecular profiles, G10: geometrical descriptors, G11: RDF descriptors, G12: 3D-MoRSE descriptors, G13: WHIM descriptors, G14: GETAWAY descriptors, G15: functional groups, G16: atom-centered descriptors, G17: electronic descriptors.

2.6. QSAR analysis

For QSAR studies, SPSS Software (version 16.0, SPSS, Inc.) was used for the simple multiple linear regression (MLR) analysis. QSAR analysis of the docking results was performed using MLR analysis with stepwise selection and elimination of variables. This was performed to model the relationships between experimental binding affinities (i.e., log K_d values) with a set of 589 descriptors. This evaluation was performed in a similar manner to the procedure that we had used in a previous study [42].

MLR analysis was performed separately for each set of the 17 groups of descriptors (G1-G17, see Table 3) while accounting for the experimental log K_d values. Finally, the best descriptors, chosen both by stepwise selection and by elimination of variables from each group, were gathered into one set. Another MLR was performed on this new set of optimal descriptors obtained from the set of 48 ligands. Furthermore, MLR analysis was carried out for the antagonist set (15 ligands) and for the agonist set (33 ligands) with the best descriptors selected from the original 48 A_{2A}R ligands.

3. RESULTS AND DISCUSSION

3.1. Homology modeling

The BLAST results for $rA_{2A}R$ indicated the highest sequence identity (82%) with $hA_{2A}R$ (NP_000666.2) [27, 34-36], suggesting that $hA_{2A}R$ was the most suitable template for HOMOLOGY MODELING. Moreover, the high percent identity between $rA_{2A}R$ and $hA_{2A}R$ reflects a close homology relationship that could help to explain the same experimental affinity values obtained either from these animal models or from human models, particularly with $A_{2A}R$ agonists derived from adenosine [43].

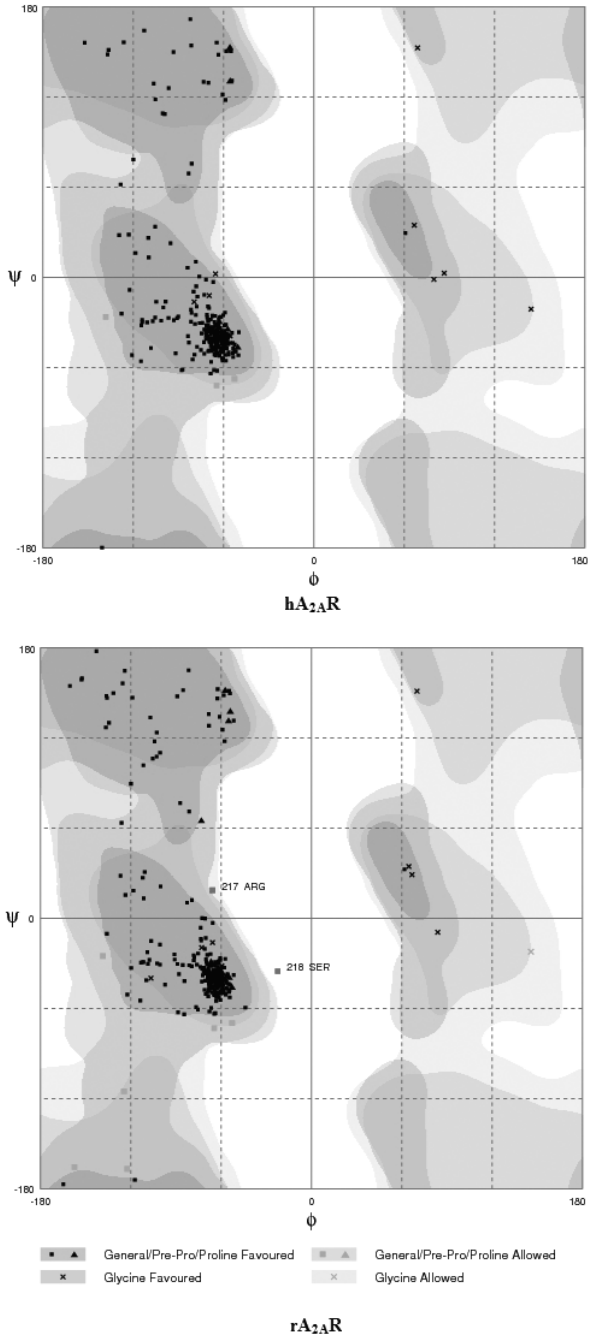


Figure 1. Ramachandran plots of $hA_{2A}R$ (PDB code: 3EML, top panel) and $rA_{2A}R$ (lower panel). The residues in disallowed regions are labeled.

To evaluate the reliability of $A_{2A}R$ models built for docking purposes, we used a Ramachandran plot (Figure 1). These methods identify the Psi/Phi angle distribution in the 3-D model within the allowed or disallowed regions. For $hA_{2A}R$, no residues were located in disallowed conformations, whereas for $rA_{2A}R$, only two residues (Arg217 and Ser218) were found. Additionally, 3-D overlapping of the native human and rat $A_{2A}R$ models shows great similarity, possibly due to the homology modeling procedure (Figure 2). With the TM-score and TM-align programs, a value of TM-score = 0.4510 was obtained, and an RMSD of 0.64 Å between $hA_{2A}R$ and $rA_{2A}R$ was found [26, 27] yielding a better 3-D model than those previously obtained for the $rA_{2A}R$ model (e.g. Kim *et al.*) [44] who reported an RMSD = 1.27 using rhodopsin as the template. Therefore, our $rA_{2A}R$ model can be considered reliable and useful for generating meaningful predictions [4]. Furthermore, we decided to perform MD simulations as performed elsewhere [45]. The long MD simulations permit visualization of the active or inactive structures that can explain the agonist or antagonist conformational movements [46]. In this case, we performed MD simulations to identify the principal structural changes that occur during 20 ns of MD simulations, which is enough

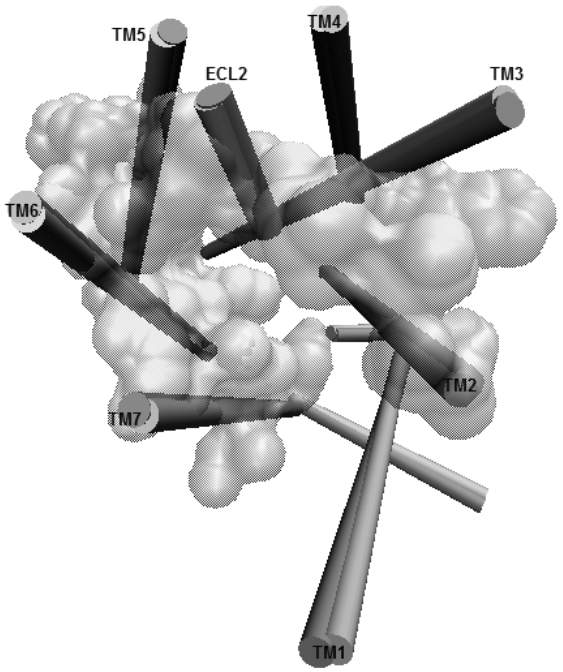


Figure 2. The superimposed $hA_{2A}R$ (dark) and $rA_{2A}R$ (light) 3D models (only the alpha helixes are depicted for clarity). The transparent surface shows the amino acids that are in contact with the co-crystallized ligand (ZM241385) in $hA_{2A}R$ and in same position on $rA_{2A}R$.

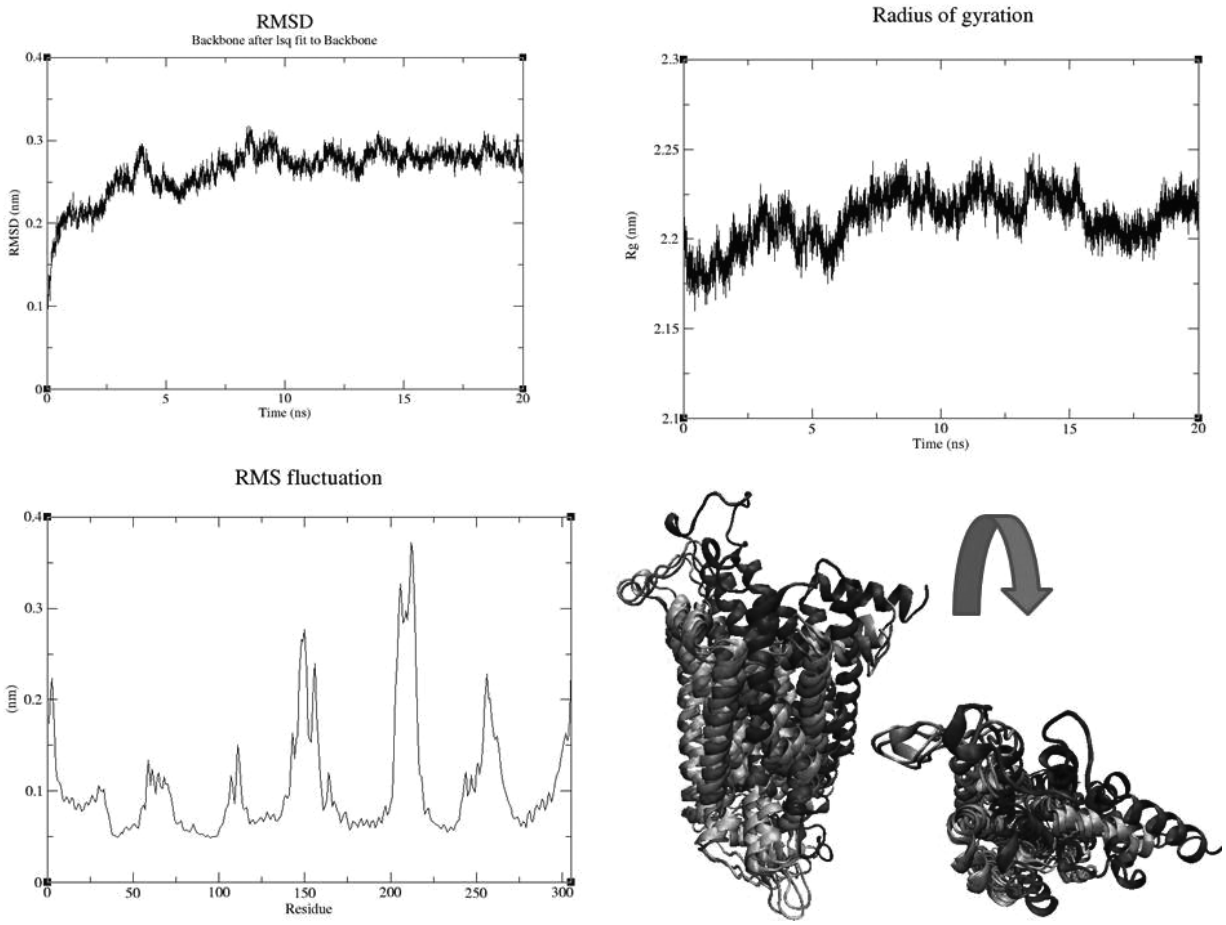


Figure 3. MD simulations of $rA_{2A}R$ to test some structure snapshots. Panels indicate the root-mean square deviation (RMSD), the radius of gyration (Rg), the root-mean square fluctuation (RMSF) and the 3-D structure alignment for all snapshots retrieved from the MD simulations.

time to reach convergence according to the RMSD and the radius of gyration, (Figure 3). During this time, it was clear that there is a greater movement in the loops in comparison with the zones that belong to the TM domain, which the 3-D model indicates is more stable than the loops. Despite the notable structural movements, it is not easy to show the active or inactive conformations because they occur at microsecond time scales [47].

3.2. Comparative analysis of experimental vs. predicted ligand affinities for the $A_{2A}R$ models

Once the $rA_{2A}R$ model was obtained by homology modeling, a docking study was performed, considering ligands whose experimental affinity data are available (Scheme 1 and Tables 1 and 2). After docking simulations, the log K_d values were obtained from free energy (ΔG) values resulting from the ligand-protein interactions (Eq. 1).

$$\Delta G = \Delta G_{vfW}\sum_{i,j}\left(\frac{A_{i,j}}{r_{ij}^{12}}-\frac{B_{ij}}{r_{ij}^{6}}\right)+\Delta G_{hbond}\sum_{ij}E(t)\left(\frac{C_{ij}}{r_{ij}^{12}}-\frac{D_{ij}}{r_{ij}^{10}}\right)+\Delta G_{elec}\sum_{ij}\frac{q_i q_j}{\epsilon(r_{ij})r_{ij}}+\Delta G_{tor}N_{tor}$$
$$+\Delta G_{sol}\sum_{ij}\left(S_i V_j+S_j V_i\right)e^{\left(-r_{ij}^2/2\sigma^2\right)} \quad \text{(Eq. 1)}$$

where the five ΔG terms on the right-hand side are coefficients empirically determined using linear regression analysis from a set of protein-ligand complexes with known binding constants. The summations are performed over all pairs of ligand atoms, *i*, and protein atoms, *j*. The *in vacuo* contributions include three interaction energy terms: a Lennard-Jones 12-6 dispersion-repulsion term, a directional 12-10 hydrogen bonding term and a screened Coulombic electrostatic potential [37].

The *Kd* values from $hA_{2A}R$ and $rA_{2A}R$ were used to perform a correlation study (Figure 4 A-C) with experimental data and we could identify a good relationship for the antagonist ligands (R^2 = 0.6712 and 0.7968, respectively) but not for the agonist ligands (Figure 4D-F). On the other hand, as it is not reliable to make any assertions based upon the initial 3-D structure, we performed a MD simulation (Figure 3) to explore the conformational space and to try to explore the local conformational changes. Then, we selected these snapshots as targets for docking of agonists or antagonists, respectively. The results (Figure 5) suggest that $rA_{2A}R$ maintains its inactive conformation (based on the relationship between the experimental and theoretical data). This is because the $rA_{2A}R$ maintain the best relationship with the antagonist ligands for both the native $rA_{2A}R$ and the snapshots retrieved from the $rA_{2A}R$ MD simulations [48, 49]. Figure 5 depicts the binding pose of the best relationship with the agonist (20 ns, R = 0.28) or antagonist (0 ns R = 0.88). The figure also depicts adenosine as the best example of agonist, which shows interactions with His273 (via formation of hydrogen bonds), and this arrangement is in agreement with recent

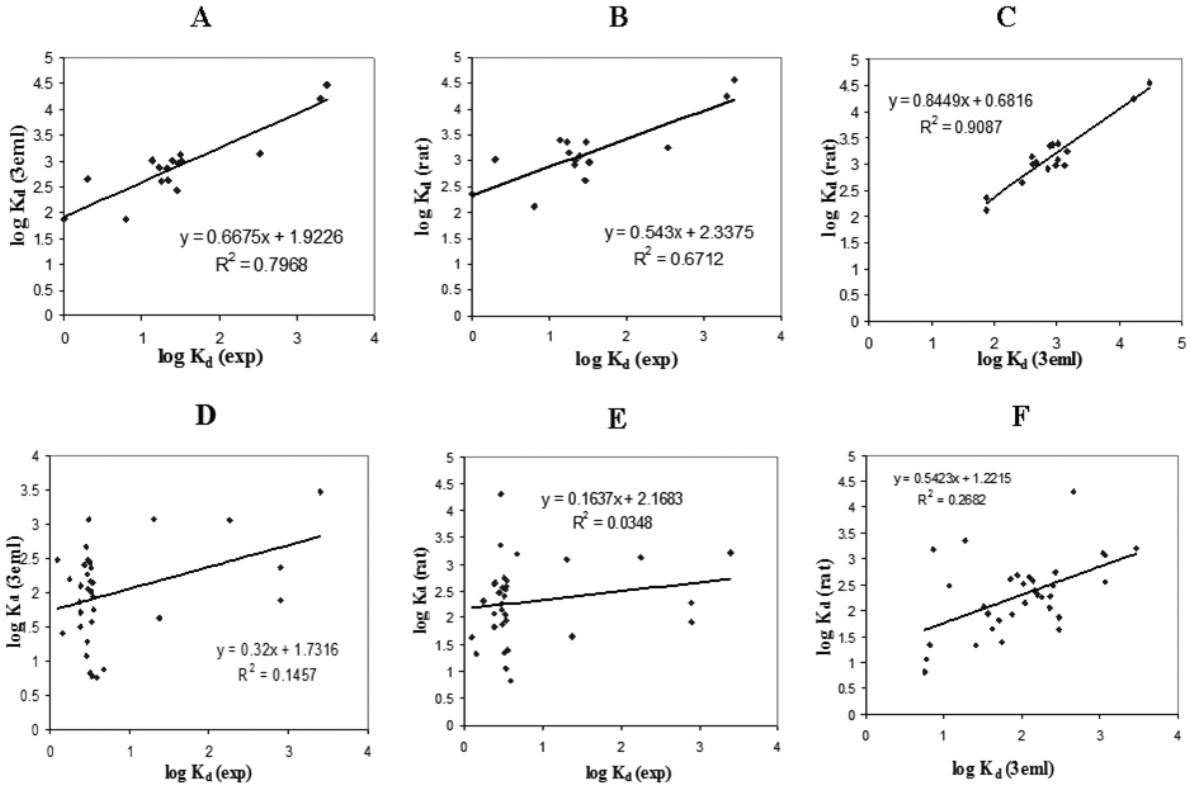


Figure 4. Antagonist ligands A) Correlation between log K_d obtained from $hA_{2A}R$ docking versus experimental log K_d values, **B)** between log K_d obtained from $rA_{2A}R$ docking versus experimental log K_d values, and **C)** between log K_d obtained from $rA_{2A}R$ docking versus log K_d observed from the $hA_{2A}R$ docking values. **Agonist ligands**, **D)** correlation between log K_d obtained from $hA_{2A}R$ docking versus experimental log K_d values, **E)** between log K_d obtained from $rA_{2A}R$ docking versus experimental log K_d values and **F)** between log K_d obtained from $rA_{2A}R$ docking versus log K_d observed from the $hA_{2A}R$ docking values.

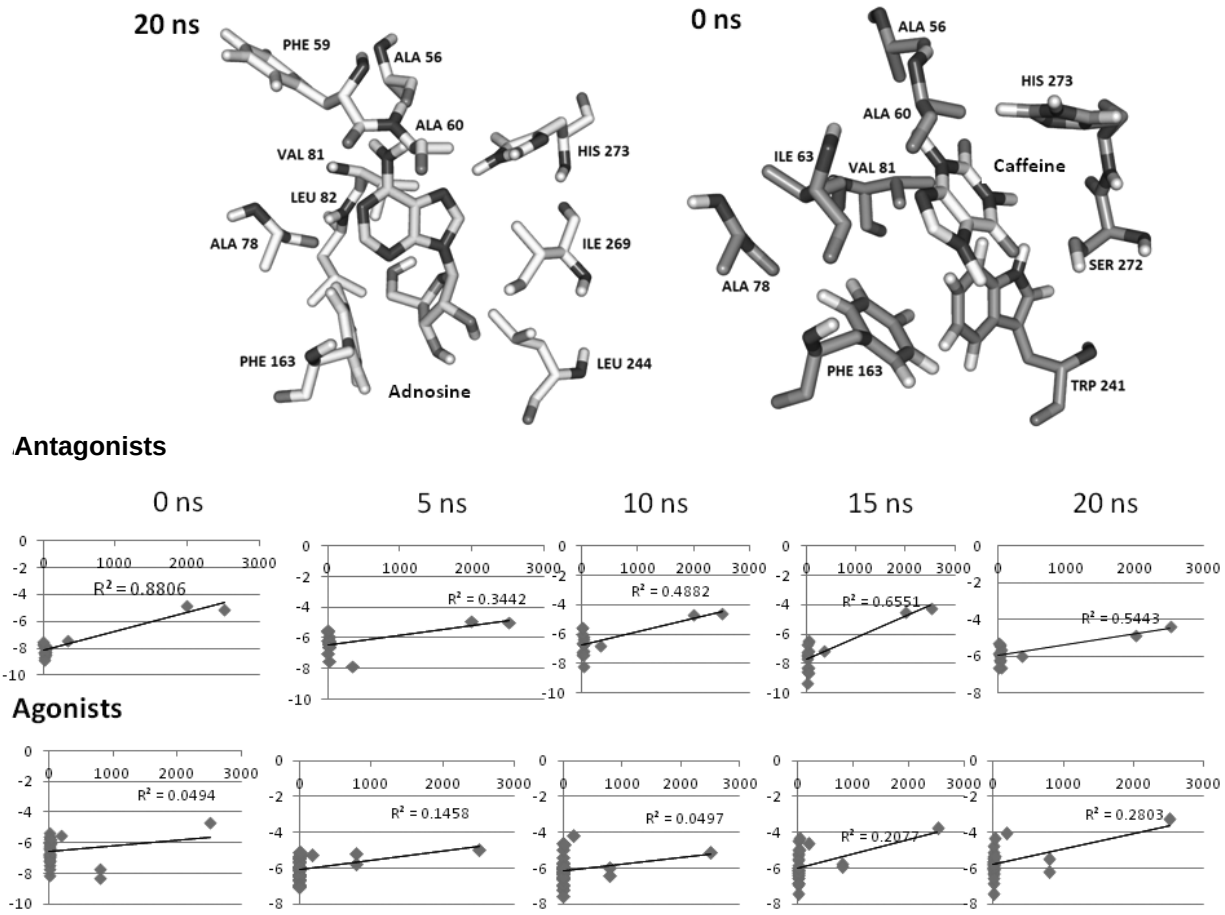


Figure 5. (**below**) relationship between Kd values (x-axes) and free energy values (ΔG, Kcal/mol), (**above**) adenosine **(left)** and caffeine (**right**) docked on the structures that exhibit the best relationship at 20 ns and 0 ns, respectively.

X-ray structure data obtained for $hA_{2A}R$ [10]. For the antagonist models, we depicted caffeine, which participates in a π-π interaction with Phe163, and these interactions are reported in the crystal structure of the human adenosine receptor $A_{2A}R$ in complex with a high affinity subtype-selective antagonist bound ZM132485 [9].

3.3 Specific interactions of selected ligands on $A_{2A}R$ and their biological relationship

To explore the specific interactions of ligands on $A_{2A}R$, a close view of each ligand in its highest affinity conformation obtained by docking was analyzed. The $hA_{2A}R$ protein was solved in complex with ZM132485 [10]. This complex shows that $hA_{2A}R$ recognizes ZM132485 in a binding site constituted by aromatic (Phe168, Trp246 and Tyr271), basic (Asn253, Asn181, His250 and His264,) acidic (Glu169) and hydrophobic (Leu85, Met177, Leu249, Ile274, Leu267, Ala265 and Met270) residues. These residues form a pocket capable of recognizing ligands (agonists or antagonists) which share five-six cycle core reported elsewhere. Furthermore, these interactions explain why most of the ligands tested here recognize this site (Figure 6). In addition, one of the most important findings is that the ribose moiety is oriented towards His278 and His273 for both the human and the rat adenosine receptors (Figures 6A and C, respectively). These results are also in agreement with the experimental data published elsewhere [10].

For the caffeine docking studies, results showed that the ligand establishes interactions with the side chains of Phe168 and Phe163 by π-π

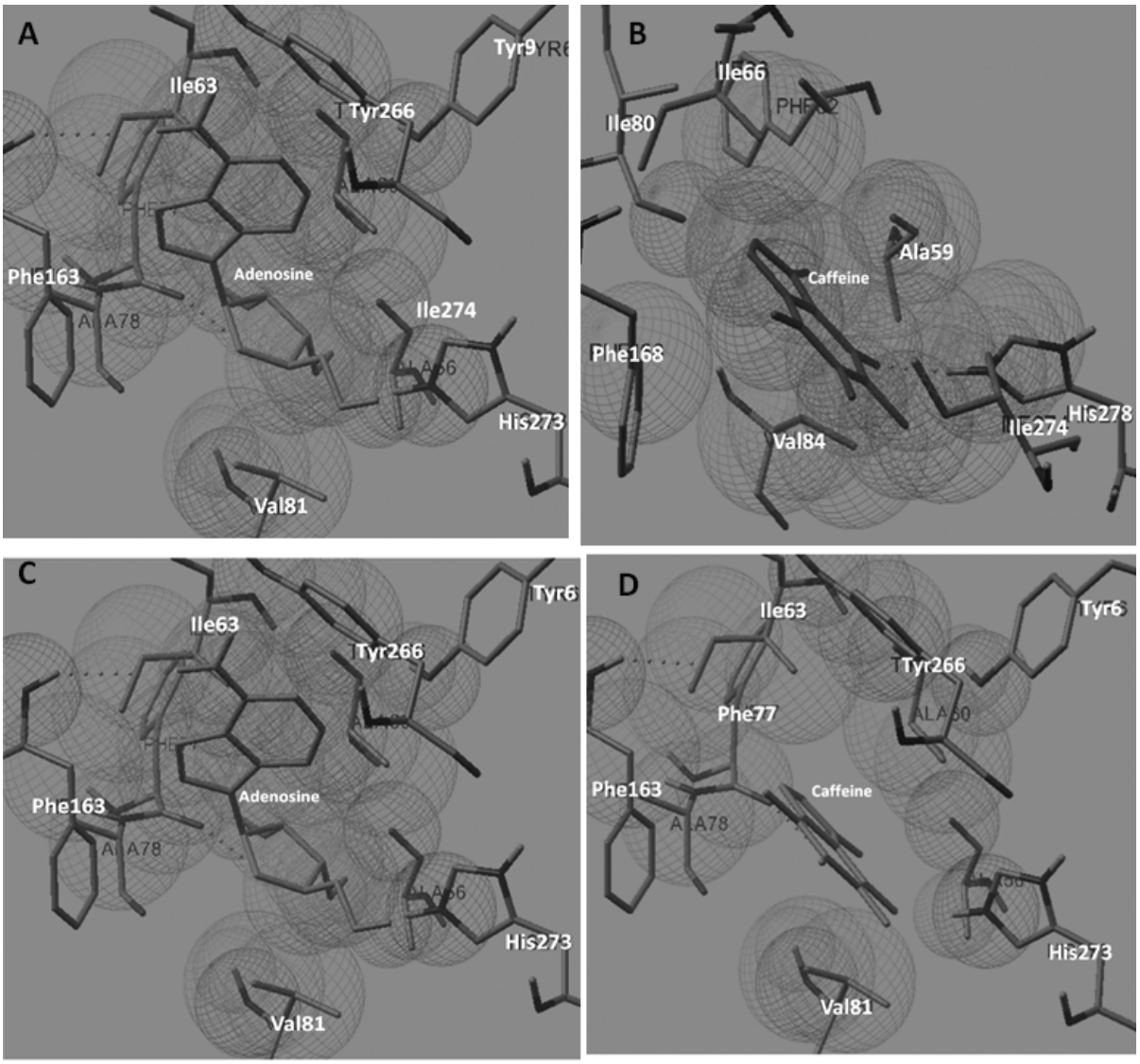


Figure 6. Adenosine on $hA_{2A}R$ **(A)** and $rA_{2A}R$ **(C)** and caffeine on $hA_{2A}R$ **(B)** and $rA_{2A}R$ **(D).**

interactions on $hA_{2A}R$ (Figure 6B) and $rA_{2A}R$ (Figure 6D), respectively. Similar interactions were found for adenosine at the bi-cycle moiety. The adenosine purine moiety has a similar spatial disposition on $hA_{2A}R$ to that of the purine moiety of ZM132485. Indeed, the highest affinities for both complexes between $r/hA_{2A}R$ and caffeine or adenosine exhibit the same binding site because they are overlapped. In general, all of the $A_{2A}R$ ligands tested have the same bi-cycle moiety. The main differences between them are in the side chain of the ligands, and they differ by a ribose located in the agonist ligands. These data explain the differences between the $A_{2A}R$ binding sites and the affinity binding values (Table 2 and Figure 6). For example, theophylline has fewer hindrance effects than the others, allowing it to reach the same site more deeply (Figure 6 and 7) interacting with residues in the TMs 2,3,6 and 7 (Supplementary Figure 1). The major obstacle for accessing the binding site is the micro-switch formed by the conserved residue at position 6.48 (Trp246 on $hA_{2A}R$ and Trp241 on $rA_{2A}R$). Conversely, the large conformational changes of the Trp residues mentioned during 7TM activation permit ligand accessibility [50]. All of the ligands

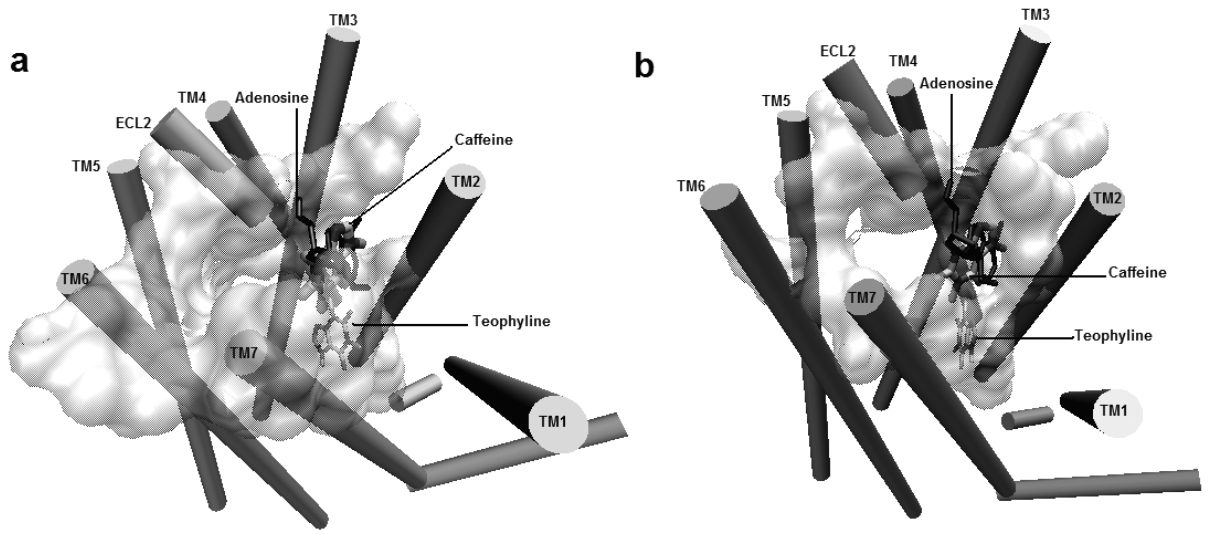


Figure 7. The $hA_{2A}R$ (a) and $rA_{2A}R$ (b) complexes with adenosine (dark) and two antagonists: caffeine and theophylline (light). The amino acids considered key in 7TM receptors by Fatakia *et al*. [40] are depicted as a transparent surface in each case.

tested showed a similar binding site except for theophylline, which reaches a deeper binding site due to its small volume. This behavior could be associated with the different responses found in clinical assays studying the effect of methyl-xanthines commonly included in diets [17, 51, 53]. For instance, Tan *et al*. [52] found that reduction of PD risk varies between green and black tea consumption. Yu *et al*. [53] found greater nervous system stimulation by caffeine than by theophylline. These differences could also be related to the variations in caffeine content found in each tea [51, 54]. Additionally, this difference could involve other mechanisms of action of xanthine compounds, including the inhibitory activity on phosphodiesterase [55]. One explanation for these different effects could be stimulation at different sites on $A_{2A}R$ that play an important role in the activation of the receptor. The differences could be related to the selectivity given by extracellular loop amino acids in $A_{2A}R$ and other 7TMs [56].

3.4. MLR-QSAR results and their implications in ligand binding

Table 3 summarizes the descriptors chosen from the MLR of experimental log K_d values with the 17 groups of descriptors (G1-G17). The final regression models obtained from MLR analysis and executed on the best sets of descriptors are shown below (See Table 4 for a clarification of the descriptors' names).

Final MLR regression model for both agonist and antagonist ligands

log K_d = 8.638868 (±2.073478) - 0.37185 (±0.107628) Mor20u

- 3.45225 (±0.559607) MATS3p + 0.630345 (±0.11321) Mor25u

- 2.09376 (±0.351586) MATS8v + 0.921764 (±0.142635) C-042

- 5.2713 (±1.759311) ATS4e - 3.11057 (±0.76633) AROM

- 0.63123 (±0.162349) GGI4 + 0.647461 (±0.286081) BEHm7 (1)

N = 49, R = 0.980, R^2 = 0.961, R^2_A = 0.952, S = 0.198, R^2_{CV} = 0.959, RMSE = 0.177.

Here, N is the number of compounds, R is the regression coefficient, R^2 is the coefficient of determinations, R^2_A is the adjusted coefficient of determination, R^2_{CV} is the cross validation coefficient of determination, and RMSE is the root-mean square error.

MLR regression model for antagonist ligands

log K_d = 2.471 (±0.187) - 9.481 (±1.182) GGI10 -1.828 (±0.469) MATS3p (2)

N = 16, R = 0.973, R^2 = 0.946, R^2_A = 0.938, S = 0.225, R^2_{CV} = 0.943, RMSE = 0.204.

MLR regression model for agonist ligands

log K_d = 3.551 (±0.490) + 0.812 (±0.204) C-042 - 0.816 (±0.570) MATS8v

- 0.686 (±0.150) C-040 - 0.644 (±0.185) C-041 - 27.475 (±7.193) JGI4 (3)

N = 33, R = 0.977, R^2 = 0.954, R^2_A= 0.946, S = 0.193, R^2_{CV} = 0.952, RMSE = 0.175.

Equation (1) shows that the most significant descriptor of the experimental log K_d for both agonist and antagonist ligands is the 2D autocorrelation descriptor ATS4e, where log K_d is inversely proportional to ATS4e. Other significant descriptors are MATS3p (2D autocorrelation descriptor) and the aromaticity index (AROM), which increases as log K_d decreases.

Equation (2) shows that the most significant descriptor for the antagonist ligands is the Galvez topological charge index of order 10 (GGI10), while equation (3) shows that the most significant descriptor for the agonist ligands is the Galvez mean topological charge index of order 4. The regression models described in equations (1)-(3) have comparable regression and cross validation parameters. However, the RMSE for the regression model of antagonist ligands (0.204) is higher than those for other two models described in equations (1) and (3), where the RMSE = 0.177 and 0.175, respectively. Additionally, for equation (2), the value of R^2_{CV} = 0.943 is lower than those for equations (1) and (3), where R^2_{CV} = 0.959 and 0.952, respectively. The cross validation of parameters indicates that the regression models obtained for the whole data set (consisting of both agonist and antagonist ligands) is comparable to the regression model obtained for the agonist ligands. However, the regression models described in equations (1) and (3) are slightly better than the model described in equation (2). The predicted and observed log K_d for the regression models described in equations (1)-(3) are shown in Figure 8.

As shown in the QSAR results, steric and charge factors in ligands are strongly related to the experimental calculated affinity for ligands. This supports the requirement of singular hindrance effects for ligands in $A_{2A}R$ recognition, and also it supports the idea that differences could be expected between caffeine and ligands with smaller hindrance effects.

Table 4. Summary of descriptors.

Descriptor name	Description
Mor20u	3D-MoRSE - signal 20/unweighted
MATS3p	Moran autocorrelation – lag 3/weighted by atomic polarizabilities
Mor25u	3D-MoRSE - signal 25/unweighted
C-042	corresponds to X—CH...X (among atom centered fragments descriptors)
ATS4e	Broto-Moreau autocorrelation of a topological structure - lag 4/weighted by atomic Sanderson electronegativities
GGI4	Galvez topological charge index of order 4
BEHm7	highest eigenvalue n. 7 of Burden matrix/weighted by atomic masses
AROM	Aromaticity
MATS8v	Moran autocorrelation – lag 8/weighted by atomic van der Waals volumes
GGI10	Galvez topological charge index of order 10
C-010	atom centered fragment CHX3
C-041	atom centered fragment X-C(+X)-X
C-042	atom centered fragment X—CH..X
JGI4	Galvez mean topological charge index of order 4

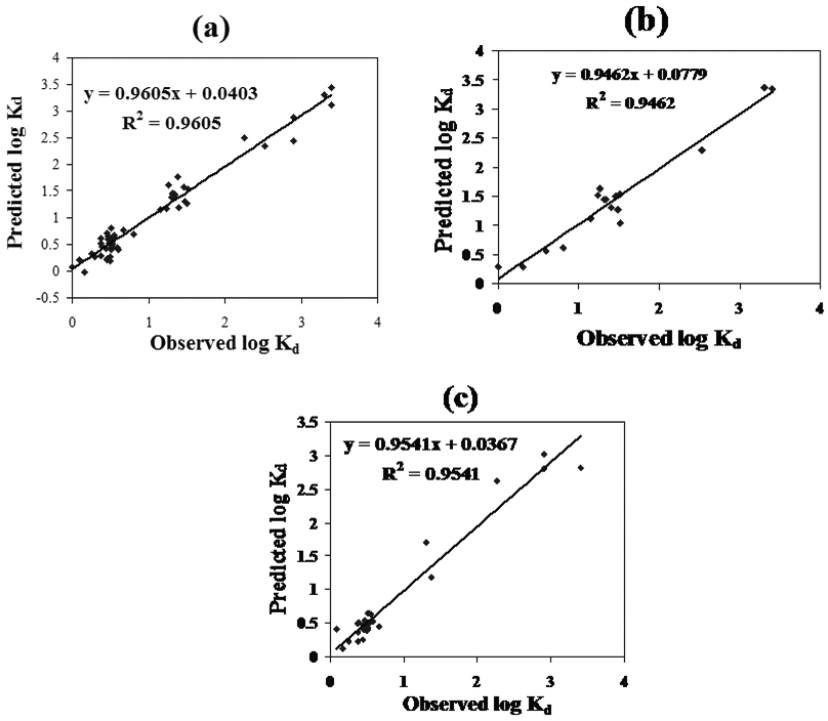
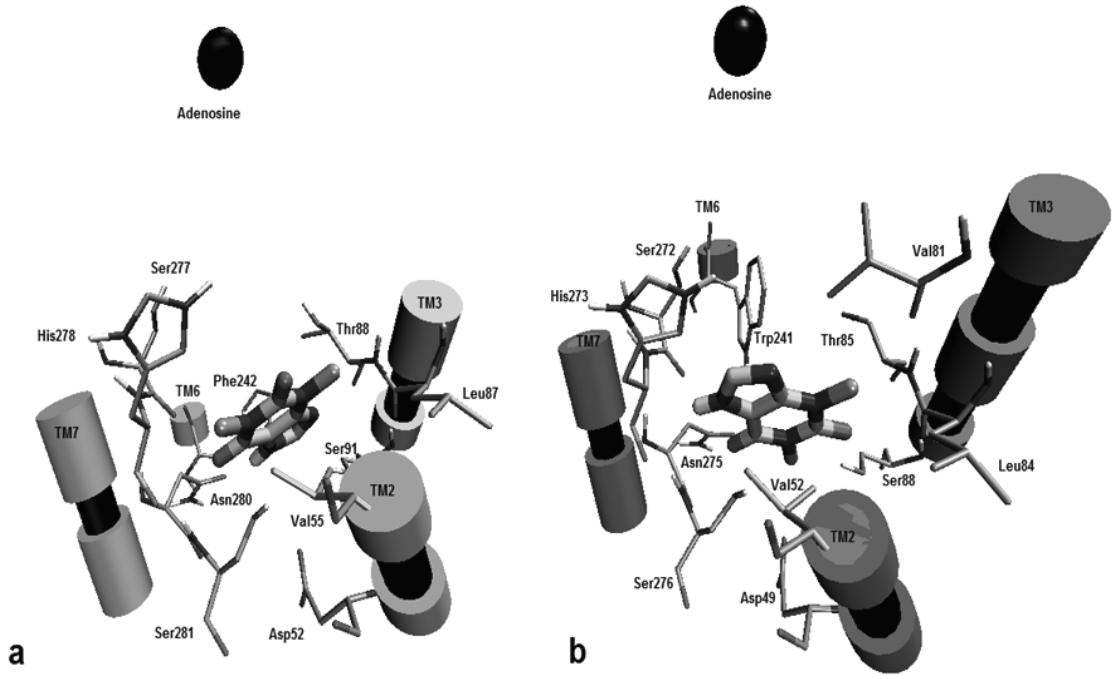


Figure 8. Predicted versus observed log Kd for **(a)** both agonist and antagonist ligands, (**b**) agonist ligands and (**c**) antagonist ligands.



Supplementary Figure 1. Binding site on (**a)** $hA_{2A}R$ and (**b**) $rA_{2A}R$ for theophylline. The core of the adenosine molecule is depicted as a black sphere. Transmembrane domains implicated are depicted as a cartoon marking the segments with amino acids interacting with theophylline. These amino acids are labeled in bond representation.

CONCLUSION

The theoretical results show that the ligands tested have similar binding to hA2AR and $rA_{2A}R$. Additionally, the QSAR results address the importance of the xanthine core contained in $A_{2A}R$ ligands and the hindrance effect on the $A_{2A}R$ affinity. Finally, these findings suggest that the $rA_{2A}R$ in its inactive state can be used for testing $A_{2A}R$ antagonists, but it cannot be used to test agonists despite being summited to MD simulations. The $rA_{2A}R$ model exhibits the best relationship with antagonists even though the adenosine agonist reaches the reported His273 via hydrogen bond interactions.

ACKNOWLEDGEMENTS

We are grateful for the scholarships and financial support from CONACYT, México (132353), ICyTDF (PIRIVI09-9), COFAA and SIP-IPN (20110786), PAPIIT-DGAPA UNAM-215708 and Posgrado en Ciencias Biológicas UNAM. The authors thank the Centro Nacional de Supercomputo, México, for providing access to the "Argentum" cluster.

REFERENCES

1. Ralevic, V. and Burnstock, G. 1998, Rev. Pharmacol., 50, 413.
2. Hernan, M. A., Takkouche, B., Caamaño-Isorna, F. and Gestal-Otero, J. 2002, Ann. Neurol., 52, 276.
3. Tuccinardi, T., Ortore, G., Manera, C., Saccomanni, G. and Martinelli, A. 2006, Eur. J. Med. Chem., 41, 321.
4. Hillion, J., Canals, M., Torvinen, M., Cadado, V., Scott, R., Terasmaa, A., Hansson, A., Watson, S., Olah, M. E., Mallol, J., Canela, E. I., Zoli, M., Agnati, L. F., Ibanez, C. F., Lluis, C., Franco, R., Ferre, S. and Fuxe, K. 2002, J. Biol. Chem., 277, 18091.
5. Mihara, T., Mihara, K., Yarimizu, J., Mitani, Y., Matsuda, R., Yamamoto, H., Akahane, A., Iwashita, A. and Matsuoka, N. J. 2007, Pharm. Exp. Ther., 323, 708.
6. Mingote, S., Font, L., Farrar, A. M., Vontell, R., Worden, L. T., Stopper, C. M., Port, R., Sink, K. S., Bunce, J. G., Chrobak, J. J. and Salamone, J. D. 2008, J. Neuroscience, 28, 9037.
7. Simola, N., Morelli, M. and Pinna, A. 2008, Curr. Pharm. Des., 14, 1475.
8. Yuzlenko, O. and Kiec´-Kononowicz, K. 2009, J. Comput. Chem., 30, 14.
9. Xu, F., Wu, H., Katritch, V., Han, G. W., Jacobson, K. A., Gao, Z. G., Cherezov, V. and Stevens, R. C. 2011, Science, 332, 322.
10. Jaakola, V. P., Griffith, M. T., Hanson, M. A., Cherezov, V., Chien, E. Y., Lane, J. R., Ijzerman, A. P. and Stevens, R. C. 2008, Science, 322, 1211.
11. Mustafi, D. and Palcewski, K. 2009, Mol. Pharmacol., 75, 1.
12. Ivanov, A. A., Barak, D. and Jacobson, K. A. 2009, J. Med. Chem., 52, 3284.
13. Yuzlenko, O. and Kiec-Kononowicz, K. 2009, J. Comput. Chem., 30, 14.
14. Cristalli, G., Lambertucci, C., Marucci, G., Volpini, R. and Dal-Ben, D. 2008, Curr. Pharm. Des., 14, 1525.
15. Volpini, R., Dal Ben, D., Lambertucci, C., Marucci, G., Mishra, R. C., Ramadori, A. T., Klotz, K. N., Trincavelli, M. L., Martini, C. and Cristalli, G. 2009, Chem. Med. Chem., 4, 1010.
16. Kull, B., Arslan, G., Nilsson, C., Owman, C., Lorenzen, A., Schwabe, U. and Fredholm, B. B. 1999, Biochem. Pharmacol., 57, 65.
17. González, M. P., Terán, C., Teijeira, M. and Helguera, A. M. 2006, Eur. J. Med. Chem., 41, 56.
18. Jankovic, J. 2008, Ann. Neurol., 63, 267.
19. Jaakola, V. P., Lane, J. R., Lin, J. Y., Katritch, V., Ijzerman, A. P. and Stevens, R. C. 2010, J. Biol. Chem., 285, 13032.
20. Holschbach, M. H., Bier, D., Stüsgen, S., Wutz, W., Sihver, W., Coenen, H. H. and Olsson, R. A. 2006, Eur. J. Med. Chem., 41, 7-15.
21. Frisch, M. J., Trucks, J. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Robb Cheeseman, J. R., Zakrzewski, V. G., Montgomery, J. A. Jr., Stratmann, R. E., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C.,

Adamo, C., Clifford, S., Ochterski, J., Peterson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Baboul, A. G., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Andres, J. L., Gonzalez, C., Head-Gordon, M., Replogle, E. S. and Pople, J. A. 1998, Gaussian 98, Revision A.9, Gaussian, Inc., Pittsburgh, PA.

22. Chern, Y., King, K., Lai, H. L. and Lai, H. T. 1992, Biochem. Biophys. Res. Commun., 185, 304.
23. Arnold, K., Bordoli, L., Kopp, J. and Schwede, T. 2006, Bioinformatics, 22, 195.
24. Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. 2003, Nucl. Acids Res., 31, 3381.
25. Guex, N. and Peitsch, M. C. 1997, Electrophoresis, 18, 2714.
26. Zhang, Y. and Skolnick, J. 2004, Proteins, 57, 702.
27. Zhang, Y. and Skolnick, J. 2005, Nucl. Acids Res., 33, 2302.
28. Lovell, S. C., Davis, I. W., de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. and Richardson, D. C. 2003, Proteins, 50, 437.
29. van der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E. and Berendsen, H. J. C. 2005, J. Comput. Chem., 26, 1701.
30. Kandt, C., Ash, W. L. and Tieleman, D. P. 2007, Methods, 41, 475.
31. Hess, B., Bekker, H., Berendsen, H. J. C. and Fraaije, J. G. E. M. 1997, J. Comput. Chem., 18, 1463.
32. Berendsen, H. J. C., Postma, J. P. M., Gunsteren, W. F., Dinola, A. and Haak, J. R. 1984, J. Chem. Phys., 81, 3684.
33. Berger, O., Edholm, O. and Jahnig, F. 1997, Biophys. J., 72, 2002.
34. Patra, M., Karttunen, M., Hyvo¨ nen, M. T., Falck, E., Lindqvist, P. and Vattulainen, I. 2003, Biophys. J., 84, 3636.
35. Melo, F., Devos, D., Depiereux, E. and Feytmans, E. 1997, Proc. Int. Conf. Intell. Syst. Mol. Biol., 5, 187.
36. Wiederstein, M. and Sippl, M. J. 2007, Nucleic Acids Res., 35, W407.
37. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. 1998, J. Comp. Chem., 19, 1639.
38. Huey, R., Morris, G. M., Olson, A. J. and Goodsell, D. S. 2007, J. Comp. Chem., 28, 1145.
39. Ballesteros, J. A. and Weinstein, H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. 1995, In Receptor Molecular Biology, Sealfon, S. C. (Ed.) Academic Press, 366.
40. Fatakia, S. N., Costanzi, S. and Chow, C. C. 2009, PLoS One, 4, e4681.
41. Humphrey, W., Dalke, A., Schulten, K. 1996, J. Mol. Graph, 14, 33.
42. Deeb, O., Rosales-Hernández, M. C., Gómez-Castro, C. Garduño-Juárez, R. and Correa-Basurto, J. 2010, Biopolymers, 93, 161.
43. Cristalli, G., Lambertucci, C., Taffi, S., Vittori, S. and Volpini, R. 2003, Curr. Top. Med. Chem., 3, 387.
44. Kim, S. K., Gao, Z. G., Van Rompaey, P., Gross, A. S., Chen, A., Van Calenbergh, S. and Jacobson, K. A. 2003, J. Med. Chem., 46, 4847.
45. Grossfield, A. 2011, Biochim. Biophys. Acta, 1808, 1868.
46. Simpson, L. M., Wall, I. D., Blaney, F. E. and Reynolds, C. A. 2011, Proteins, 79, 1441.
47. Provasi, D. and Filizola, M. 2010, Biophys. J., 98, 2347.
48. Khafizov, K., Lattanzi, G. and Carloni, P. 2009, Proteins, 75, 919-930.
49. Rodríguez, D., Piñeiro, Á. and Gutiérrez-de-Terán, H. 2001, Biochemistry, 50, 4194.

50. Nygaard, R., Frimurer, T. M., Holst, B., Rosenkilde, M. M. and Schwartz, T. W. 2009, Trends Pharmacol. Sci., 30, 249.
51. Tan, L. C., Koh, W. P., Yuan, J. M., Wang, R., Au, W. L., Tan, J. H., Tan, E. K., and Yu, M. C. 2008, Am. J. Epidemiol., 167, 553.
52. Barranco-Quintana, J. L., Allam, M. F., Del Castillo, A. S. and Navajas, R. F. 2009, J. Am. Coll. Nutr., 28, 1.
53. Yu, G., Maskray, V., Jackson, S. H., Swift, C. G. and Tiplady, B. 1991, Br J Clin. Pharmacol., 32, 341.
54. Hannila, S. S. and Filbin, M. T. 2008, Exp. Neurol., 209, 321.
55. Genders, A. J., Bradley, E. A., Rattigan, S. and Richards, S. M. 2011, Am. J. Physiol. Endocrinol. Metab., 301, E342.
56. Soriano-Ursúa, M. A., Trujillo-Ferrara, J. G. and Correa-Basurto, J. 2010, J. Med. Chem., 53, 923.