SUPPORTING INFORMATION

Bridging molecular docking to membrane molecular dynamics to investigate GPCR-ligand recognition: the human A_{2A} adenosine receptor as a key study.

Davide Sabbadin[#], Antonella Ciancetta[#] and Stefano Moro[#]*

[#] *Molecular Modeling Section (MMS), Dipartimento di Scienze del Farmaco, Università di Padova, via Marzolo 5, Padova, Italy.*

Contents of Supporting Information

Table S1. Predicted binding conformations of the selected co-crystallized ligands.

PDB ID [a]:	Cluster number	Gold Score	Pose R.M.S.D.[Å] [b]	ChemPLP	Ligand-protein hydrogen bonds [c]
3EML					
	1-7	**53.41**	Core (< 1 Å) 2.00	**77.17**	E169, N253
	8	51.62	3.06	69.64	E169, N253
	9	45.02	3.58	68.51	E169, D170
	10	50.21	2.47	67.77	-
3UZA					
	1	**53.49**	0.69	**63.24**	N253
	2	47.38	5.60	60.95	-
	3	48.68	4.83	57.04	N253
	4	45.75	4.50	56.55	N253
	5	41.03	6.41	52.52	E169
	6	39.09	4.89	52.18	-
	7	28.99	5.69	51.30	E169
	8	45.46	2.95	51.05	N253
	9	45.48	6.80	50.19	-
	10	47.89	7.72	49.87	-
3UZC					
	1	**56.98**	0.33	**69.87**	N253, H278
	2	53.97	1.71	69.84	N253, H278
	3	41.37	3.43	62.53	-
	4	49.77	4.97	61.14	N253, H278
	5	48.60	4.70	59.98	N253, H278
	6	46.02	4.60	59.30	-
	7	39.83	5.50	57.11	-
	8	45.48	5.53	56.56	E169
	9	35.52	6.58	56.55	E169, N253
	10	-148.66	5.31	-78.44	-
3RFM					
	1	34.91	0.74	**36.99**	N253
	2	**38.63**	4.33	36.95	-
	3	34.54	3.85	36.95	N253
	4	36.79	3.20	31.84	N253
	5	34.30	3.50	31.12	N253
	6	35.53	3.41	21.72	N253
	7	36.25	3.88	20.50	N253
	8	34.80	3.21	14.72	-
	9	34.48	3.14	11.92	-
	10	36.06	4.21	9.29	-
2YDV					
	1	**64.25**	0.29	**67.03**	T88, N253, E169
	2	56.85	4.98	43.56	N253, E169
	5	56.27	2.27	50.51	N253
	7	51.74	3.08	23.62	N253, E169

[a] ZM 241385 (PDB ID: 3EML), T4G (PDB ID: 3UZA), T4E (PDB ID: 3UZC), caffeine (PDB ID: 3RFM) and NECA (PDB ID: 2YDV); [b] R.M.S.D. values with respect to the corresponding crystal structures; [c] Binding site residues that establish hydrogen bonds with the ligand.

Table S2. Biophysical parameters describing the considered membrane-embedded ligand-protein complexes: total number of lipids (upper panel) and average area per lipid headgroup (lower panel).

Total number of lipids

PDB ID	# of lipids composing the bilayer	# of lipids composing the intracellular leaflet
3EML	110	54
3UZA	110	54
3UZC	112	54
3RFM	115	55
2YDV	113	56

Average area per lipid headgroup (APL) [Å^2]

	$eAPL_{1ns}$	$iAPL_{1ns}$	$eAPL_{9ns}$	$iAPl_{9ns}$
3EML	64.93 ± 1.02	68.81± 1.03	66.67 ± 3.65	67.97 ± 3.23
3UZA	66.64 ± 1.17	69.33 ± 1.24	63.06 ± 2.97	69.91 ± 3.72
3UZC	64.06 ± 0.99	70.26 ± 2.07	61.11 ± 2.39	67.29 ± 2.76
3RFM	61.68 ± 1.21	68.12 ± 2.24	62.21 ± 3.26	68.14 ± 2.93
2YDV	63.98 ± 1.24	68.76 ± 1.83	64.51± 2.75	69.78 ± 3.04

Table S3. Slope coefficients of the linear functions f(x)=m· x fitted on the DSF data obtained from MD trajectories.

PDB ID [a]:	Cluster number	Gold Score	M.D. average R.M.S.D. [Å] [b]	DSF_{ele}	$wDSF_{ele}$	DSF_{hyd}	$wDSF_{hyd}$	MD R.M.S.D. at 60 ns [Å]
3EML								
	1-7 XRAY	**53.41**	2.52 ± 0.76	-16.01	-8.55	118.62	62.49	Core<2Å
	8	51.62	5.33 ± 1.87	-14.26	-3.95	104.19	31.42	6.97
	9	45.02	5.55 ± 1.35	-12.75	-4.15	13.62	41.69	Core<2Å
	10	50.21	2.07 ± 0.74	-16.59	-6.84	115.78	47.96	Core<2Å
3UZA								
	1 XRAY	**53.49**	0.99 ± 0.43	-16.37	-7.40	126.18	58.66	2.25
	2	47.38	2.69 ± 1.10	-2.73	-0.71	96.19	35.65	4.94
	3	48.68	1.95 ± 0.97	-11.92	-7.14	96.15	54.09	4.96
	4	45.75	2.70 ± 1.05	-5.60	-2.04	98.74	33.62	5.51
	5	41.03	4.18 ± 1.27	-7.94	-2.79	71.66	25.61	6.70
	6	39.09	4.18 ± 1.97	-6.05	-1.87	102.26	37.39	5.09
	7	28.99	3.72 ± 1.80	-1.20	-0.34	83.29	22.60	10.96
	8	45.46	4.50 ± 2.27	-10.31	-3.08	86.84	28.54	7.69
	9	45.48	6.38 ± 1.89	-7.15	-1.05	49.60	7.82	7.37
	10	47.89	3.98 ± 3.41	-8.67	-3.02	82.56	31.19	20.86
3UZC								
	1 XRAY	**56.98**	2.63 ± 0.96	-21.05	-5.90	171.34	51.26	2.17
	2	53.97	1.78 ± 0.53	-13.14	-11.30	158.73	127.83	2.24
	3	41.37	3.43 ± 1.30	-3.70	-1.67	130.47	50.07	4.40
	4	49.77	2.75 ± 0.93	-7.08	-3.10	130.85	55.85	5.86
	5	48.60	3.00 ± 0.90	-9.57	-7.68	133.46	101.18	4.70
	6	46.02	3.35 ± 1.54	-9.98	-3.88	142.36	53.609	5.49
	7	39.83	3.14 ± 1.50	-10.27	-4.49	139.4	55.06	7.64
	8	45.48	2.86 ± 1.41	-8.29	-3.53	149.54	70.13	6.31
	9	35.52	4.96 ± 1.89	-9.42	-4.23	131.72	49.95	5.73
	10	-148.66	3.89 ± 1.37	-11.39	-3.46	160.71	60.93	5.79
3RFM								
	1	34.91	4.68 ± 2.58	-3.64	-1.34	27.10	9.47	2.42

	2	**38.63**	3.99 ± 1.15	-2.97	-1.79	34.60	21.83	5.17
	3	34.54	3.31 ± 1.56	-3.22	-0.99	28.10	11.33	4.62
	4	36.79	3.41 ± 1.13	-4.96	-0.86	29.59	5.14	>100
	5	34.30	38.27 ± 46.46 [c]	-4.27	-2.17	26.55	15.73	3.96
	6	35.53	1.88 ± 1.12	-3.51	-1.55	16.54	7.84	5.75
	7	36.25	4.50 ± 1.22	-4.51	-1.15	25.57	6.28	2.89
	8	34.80	3.36 ± 1.07	-4.92	-1.71	26.76	8.43	3.88
	9	34.48	4.82 ± 1.56	-0.74	-0.17	18.36	6.40	5.07
	10	36.06	4.04 ± 1.05	-1.39	-0.43	12.20	4.77	3.62
2YDV								
	1 XRAY	**64.25**	1.15 ± 0.33	-25.72	-22.49	53.04	45.96	1.26
	2	56.85	4.13±1.17	-14.65	-8.33	48.62	27.73	2.54
	5	56.27	1.57±0.48	-21.68	-5.39	42.97	10.62	2.46
	7	51.74	2.70±0.67	-13.38	-7.28	45.36	21.41	3.00

[a] ZM 241385 (PDB ID: 3EML), T4G (PDB ID: 3UZA), T4E (PDB ID: 3UZC), caffeine (PDB ID: 3RFM) and NECA (PDB ID: 2YDV); [b] Average ligand (R.M.S.D.) fluctuation values during the unrestrained MD simulation (60 ns); [c] During the MD simulation the ligand is displaced from the binding pocket.

Figure S1. Pose 1 (a) and 2 (b) of compound T4E (PDB ID: 3UZC) prior (gray sticks) and after (green sticks) 60 ns of MD simulations: Side chains of the amino acids in contact with the selected poses are displayed as sticks. Hydrogen atoms are not displayed and the view of TM7 is partially omitted. Hydrogen bond interactions of the starting and final conformation are highlighted as yellow and blue dashed line, respectively. Electrostatic IEFs (c, values expressed in kcal/mol) for selected residues after 60 ns of MD simulation.

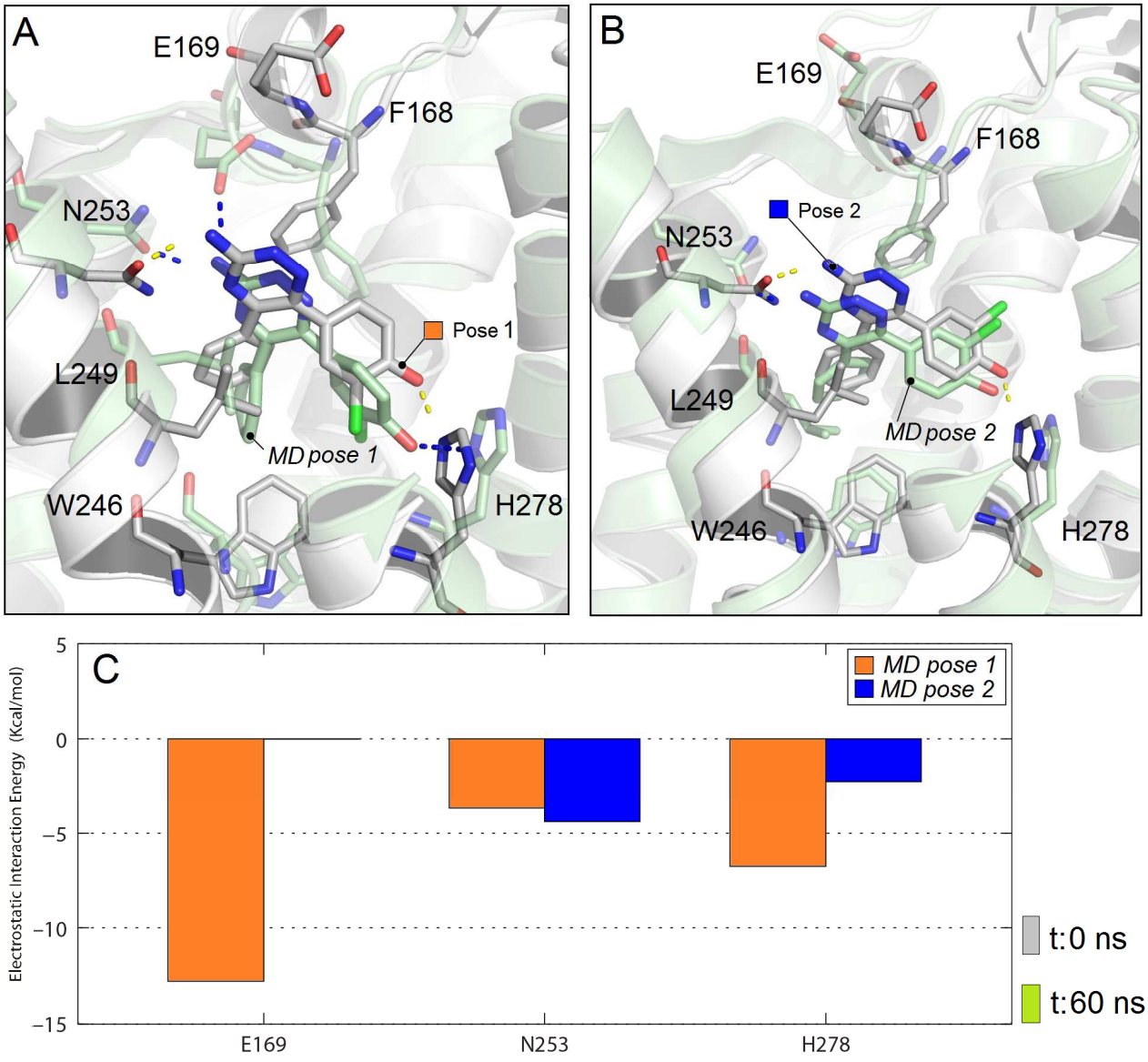
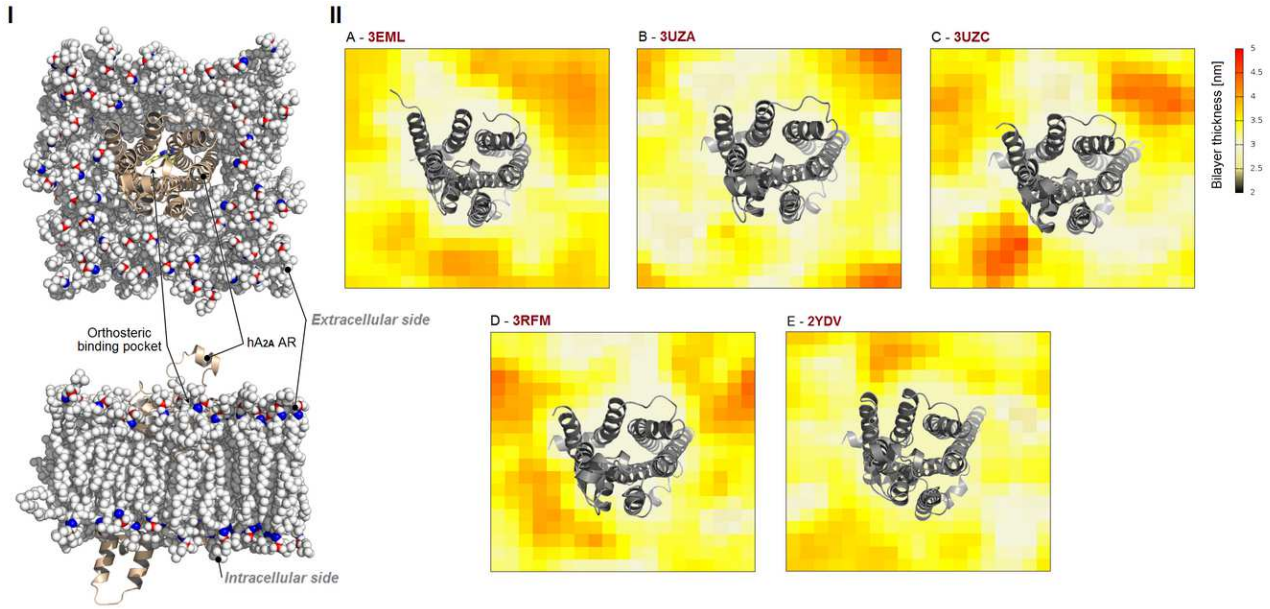


Figure S2. Representation of the human A_{2A} AR embedded in a POPC lipid bilayer (white spheres). Phosphorous atom (red) and the carbon atom (blue) bound to the phosphate group are highlighted in colors. Bilayer thickness at the end of the equilibration phase of MD simulation for each hA_{2A} AR-ligand complex: (a) ZM241385 (PDB ID: 3EML), (b) T4G (PDB ID: 3UZA), (c) T4E (PDB ID: 3UZC), (d) caffeine (PDB ID: 3RFM), (e) NECA (PDB ID: 2YDV).



In silico inspection of the central and intracellular water molecules clusters. The presence of a stable cluster of water molecules in the middle of the 7TM bundle (CC, left central panel in Figure S2) has been monitored by computing the permanence time of water molecules within a range of 7 Å from Ser91 (3.39), thus covering approximately 15 Å of distance from two functionally important and conserved residues such as Trp246 (6.48) and Tyr288 (7.53). This specific cluster is characterized by a persistent triad of water molecules which creates an hydrogen bond network among Asn24 (1.50), Asp52 (2.50), Ser91 (3.39), Trp246 (6.48), Asn280 (7.45), and Gln284 (7.49), thus interfacing TM2, TM3, TM6 and TM7. In the complex with ZM 241385 (Figure S2a, right central panel) the triad is composed of TIPP1512 and TIPP5091 (residing for the 98% of the MD simulation) and TIPP8386, which takes part in the cluster for about the 20% of the simulation time. TIPP5091 instead occupies the position that has been assigned to a sodium ion in the high-resolution hA_{2A} AR crystal structure that has been recently solved (PDB ID: 4EIY). The analysis of the CC in the complexes with T4G and T4E (Figure S2b and c, right central panel) highlights that the bridging role is played by water molecules that reside from 15% to 25% of the total simulation time and share similar geometrical positions to the ones found in the complex with ZM 241385. Moreover, the same CC has been also evidenced in the structure containing caffeine (Figure S2d, right central panel), where the water molecules reside from 18% to 8% of the simulation time. In the agonist (NECA) bound human A_{2A} AR the strong compling between the cited aminoacid located in the middle of the 7TM bundle is mediated by TIP5366, TIP3138, TIP1641, TIP5564 and TIP3784 which reside from 30% to 87% of the simulation time.

An important cluster of resident water molecules (IC, left central panel in Figure S2) is located in the proximity of the DRY motif (D, Asp101 (3.49); R, Arg102 (3.50); Y, Tyr103 (3.51)). In the complex with ZM 241385 (Figure S2a, lower panel), a water molecule (TIP1894, permanency 20% of the MD simulation) is hydrogen bonded to the side chain of Glu228 (6.30) and two other persistent water molecule molecules (TIP14 and TIP4707, resident for 89% and 45% of the entire MD trajectory, respectively) are located near Asp101 (3.49). In the analysis of the complex with T4G (Figure S2b, lower panel), a water molecule (TIP 3157) interacts with Asp101 (3.49) and Arg102 (3.50) for about the 20% of the simulation time, whereas two other molecules (TIP6955 and TIP7839) are located close to Asp101 (3.49) for the 15% and 19% of the simulation time, respectively. Another water molecule (TIP1894, residency 18% of the MD simulation) is hydrogen bonded to the side chain of Glu228 (6.30). A triad of water molecules (TIP6848, TIP7011, and TIP6555) has been evidenced in the proximity of Asp101 (3.49), Arg102 (3.50) and Glu228 (6.30) in the complex with T4E (Figure S2c, lower panel), whereas in the structure with caffeine (Figure S2d, lower panel) a similar cluster is composed by four water molecules (TIP1158, TIP634, TIP7965 and TIP2403) resident from the 25% to the 34% of the MD trajectory. In the agonist (NECA) bound human A_{2A} AR the most resident water molecules (TIP6917 and TIP4664) are resident in the proximity of Arg102 and Asp101 from the 18% to the 25% of the MD trajectory, respectively. No resident water molecules is found in the proximity of Glu228 which is not found to interact with Arg102 for the entire length of the MD trajectory.

Figure S3. Permanence analysis of clustered water molecules at the hA_{2A} AR in complex with the selected antagonists: (a) ZM241385 (PDB ID: 3EML), (b) T4G (PDB ID: 3UZA), (c) T4E (PDB ID: 3UZC), (d) caffeine (PDB ID: 3RFM), (e) NECA (PDB ID: 2YDV). Left central panel: ribbon representation of the hA_{2A} AR embedded in a POPC lipid bilayer (VDW spheres). Side chains of important residues are displayed as gray sticks and hydrogen atoms are not displayed. The extracellular (EC) cluster is located inside the orthosteric binding pocket, the central (CC) cluster in the middle of the receptor and the intracellular (IC) cluster in the proximity of the DRY motif (D101, Asp101; R102, Arg102;Y103, Tyr103 - not shown). Time residence of water molecules is reported as percentage of the total unrestrained MD simulation time (60 ns): Stable water molecules (residence >5%) are depicted with enhanced radius sticks, molecules with a permanence time < 2% are displayed with enhanced transparency sticks, and fast-exchanging water molecules involved in crucial ligand-protein interactions are marked with an asterisk. Hydrogen bond interactions are highlighted as a red dashed line.

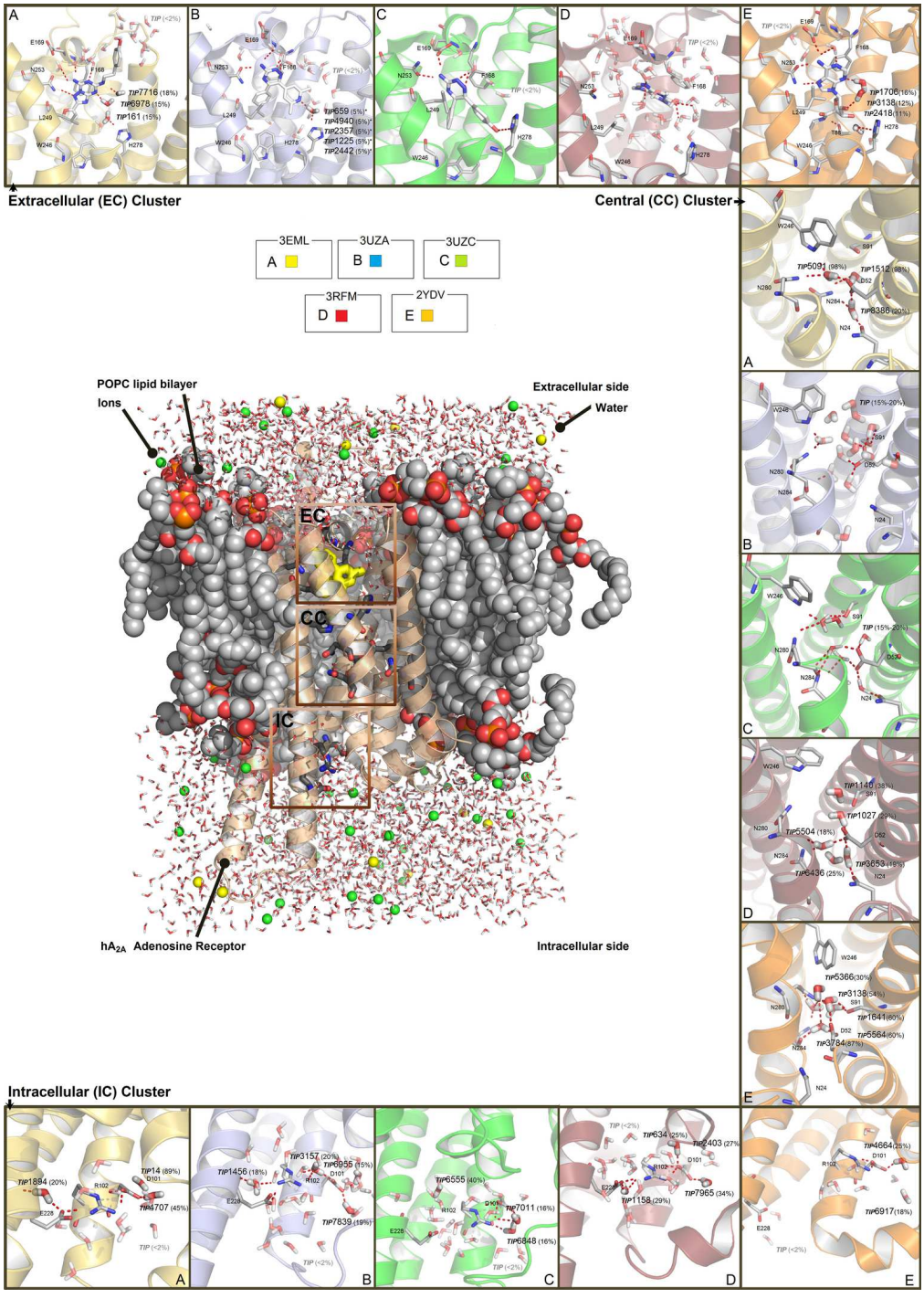
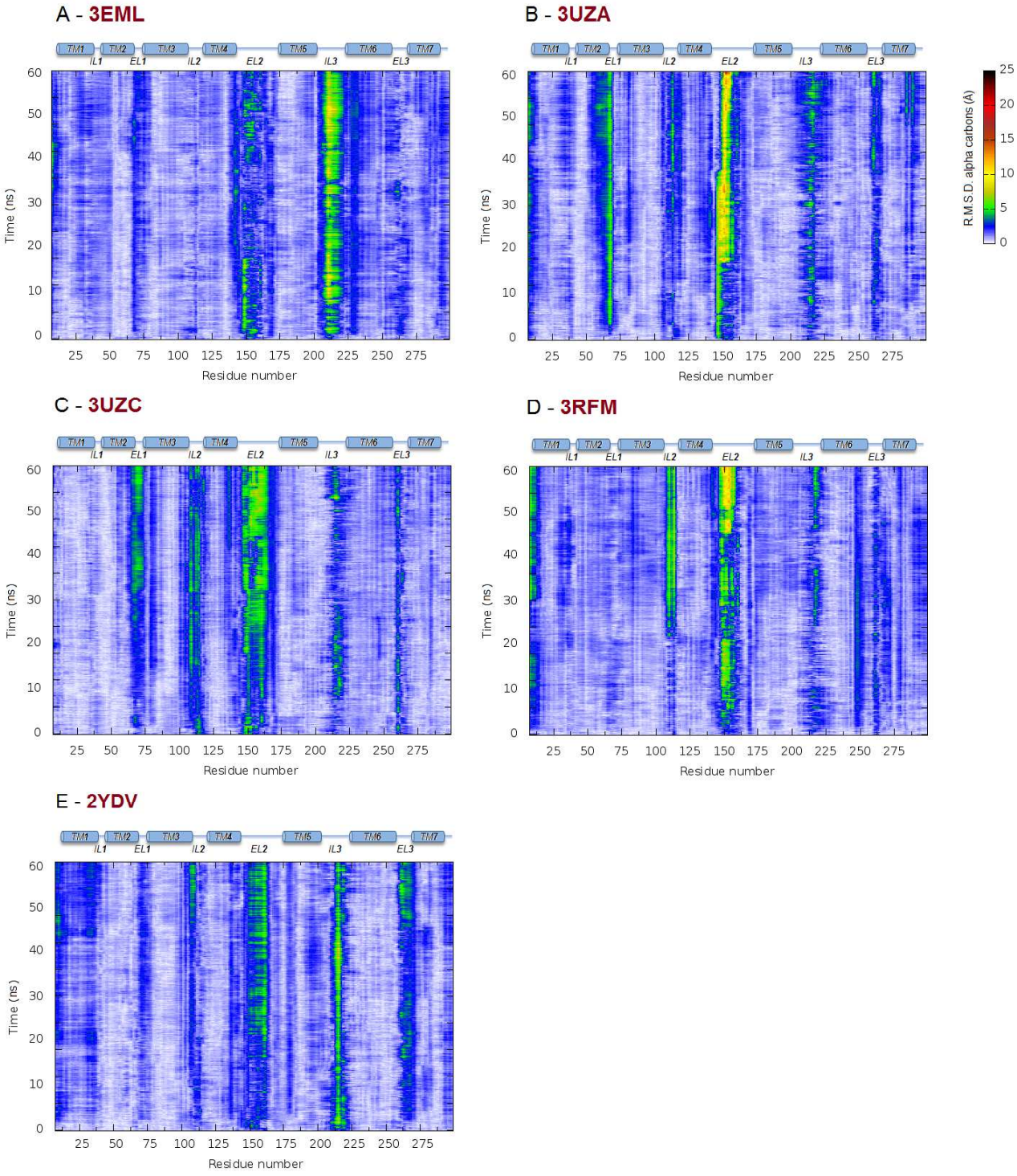


Figure S4. Per residue Cα (R.M.S.D.) fluctuation analysis of hA_{2A} AR models during the unrestrained MD.



Video S1 caption. Evolution of pose 1 of ZM241385 (white sticks) in comparison with the crystallographic information (yellow sticks), inside the hA_{2A} AR orthosteric binding pocket, during the MD simulation (60ns). The MD trajectory has been compressed into a 12 seconds movie. Ribbon representation of the receptor protein and a CPK spheres representation of the POPC lipid bilayer is shown. Hydrogen atoms are not displayed and the view of TM7 is partially omitted. Side chains of the most important amino acids in contact with selected compound are displayed as gray sticks. Hydrogen bonding interactions are highlighted as yellow dashed lines and ligand-protein hydrogen bond acceptor/donor distance is reported above the dashed line. The most persistent water molecules that solvate the ligand are displayed as red spheres.

Video S2 caption. Evolution of pose 1 of T4G (white sticks) in comparison with the crystallographic information (yellow sticks), inside the hA_{2A} AR orthosteric binding pocket, during the MD simulation (60ns). The MD trajectory has been compressed into a 12 seconds movie. Ribbon representation of the receptor protein and a CPK spheres representation of the POPC lipid bilayer is shown. Hydrogen atoms are not displayed and the view of TM7 is partially omitted. Side chains of the most important amino acids in contact with selected compound are displayed as gray sticks. Hydrogen bonding interactions are highlighted as yellow dashed lines and ligand-protein hydrogen bond acceptor/donor distance is reported above the dashed line. The most persistent water molecules that solvate the ligand are displayed as red spheres.

Video S3 caption. Evolution of pose 1 of T4E (white sticks) in comparison with the crystallographic information (yellow sticks), inside the hA_{2A} AR orthosteric binding pocket, during the MD simulation (60ns). The MD trajectory has been compressed into a 12 seconds movie. Ribbon representation of the receptor protein and a CPK spheres representation of the POPC lipid bilayer is shown. Hydrogen atoms are not displayed and the view of TM7 is partially omitted. Side chains of the most important amino acids in contact with selected compound are displayed as gray sticks. Hydrogen bonding interactions are highlighted as yellow dashed lines and ligand-protein hydrogen bond acceptor/donor distance is reported above the dashed line. The most persistent water molecules that solvate the ligand are displayed as red spheres.

Video S4 caption. Evolution of pose 1 of caffeine (white sticks) in comparison with the crystallographic information (yellow sticks), inside the hA_{2A} AR orthosteric binding pocket, during the MD simulation (60ns). The MD trajectory has been compressed into a 12 seconds movie. Ribbon representation of the receptor protein and a CPK spheres representation of the POPC lipid bilayer is shown. Hydrogen atoms are not displayed and the view of TM7 is partially omitted. Side chains of the most important amino acids in contact with selected compound are displayed as gray sticks. Hydrogen bonding interactions are highlighted as yellow dashed lines and ligand-protein hydrogen bond acceptor/donor distance is reported above the dashed line. The most persistent water molecules that solvate the ligand are displayed as red spheres.

Video S5 caption. Evolution of pose 1 of NECA (white sticks) in comparison with the crystallographic information (yellow sticks), inside the hA_{2A} AR orthosteric binding pocket, during the MD simulation (60ns). The MD trajectory has been compressed into a 12 seconds movie. Ribbon representation of the receptor protein and a CPK spheres representation of the POPC lipid bilayer is shown. Hydrogen atoms are not displayed and the view of TM7 is partially omitted. Side chains of the most important amino acids in contact with selected compound are displayed as gray sticks. Hydrogen bonding interactions are highlighted as yellow dashed lines and ligand-protein hydrogen bond acceptor/donor distance is reported above the dashed line. The most persistent water molecules that solvate the ligand are displayed as red spheres.

Graphics have been created with Pymol (http://pymol.org/) and Povray (http://www.povray.org/).