Deciphering the Agonist Binding Mechanism to the Adenosine A1 Receptor

Giuseppe Deganutti1,†,‡, Kerry Barkan2,§,†, Barbara Preti3, Michele Leuenberger3, Mark Wall4, Bruno G. Frenguelli4, Martin Lochner3, Graham Ladds2,†,* and Christopher A Reynolds1,†,*

1 Centre for sport, Exercise and Life Sciences, Faculty of Health and Life Sciences, Coventry University, Alison Gingell Building, CV1 5FB, UK

2 Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge, CB2 1PD, UK

3 Institute of Biochemistry and Molecular Medicine, University of Bern, 3012 Bern, Switzerland

4 School of Life Sciences, University of Warwick, Gibbet Hill Rd, Coventry CV4 7AL, UK

§ New address: Sosei Heptares, Granta Park Steinmetz Building, Cambridge, CB21 6DG, UK

† These authors contributed equally

‡ These authors contributed equally

* Corresponding authors: ad5288@coventry.ac.uk; grl30@cam.ac.uk; ad5291@coventry.ac.uk

KEYWORDS
G protein coupled receptors; GPCRs; adenosine A1 receptor; A1R; supervised molecular dynamics; SuMD, mutagenesis experiments, binding
Despite being amongst the most characterized G protein-coupled receptors (GPCRs), adenosine receptors (ARs) have always been a difficult target in drug design. To date, no agonist other than the natural effector and the diagnostic regadenoson has been approved for human use. Recently, the structure of the adenosine A1 receptor (A1R) was determined in the active, Gi protein complexed state; this has important repercussions for structure-based drug design. Here, we employed supervised molecular dynamics simulations and mutagenesis experiments to extend the structural knowledge of the binding of selective agonists to A1R. Our results identify new residues involved in the association and dissociation pathway, they suggest the binding mode of N6-cyclopentyladenosine (CPA) related ligands, and they highlight the dramatic effect that chemical modifications can have on the overall binding mechanism, paving the way for the rational development of a structure-kinetics relationship of A1R agonists.

1 Introduction

X-ray and cryo-electron microscopy (cryo-EM) structures are determined in equilibrium conditions. This is not representative of the physiological environment, in which the concentration of endogenous effectors and exogenous drugs in the proximity of their target continuously change due to pharmacokinetics. It follows that the affinity of a ligand, which is measured in steady state conditions, is not descriptive of its in vivo binding. The mechanism through which a molecule binds to its biological target, on the other hand, does not change with the local concentration. This consideration has produced increasing attention towards the binding kinetics as key aspect to consider in structure-based drug design (SBDD). Understanding the binding and unbinding
pathways of congeneric ligands would unlock the development of favorable structure-kinetics relationships (SKR) and foster rational guidance for medicinal chemistry modifications.

G protein-coupled receptors (GPCRs), which are target for roughly 34% of the FDA approved drugs\(^1\), represent an invaluable reservoir for therapeutic strategies, however, the majority of the GPCRome is currently not drugged\(^2\). One of the prototypical GPCR family still orphan of therapeutics is the adenosine receptors (ARs). Adenosine acts ubiquitously on four different ARs (e.g. A\(_1\)R, A\(_2A\)R, A\(_2B\)R and A\(_3\)R) contributing to the broad range of purinergic signalling\(^3,4\). The main effect mediated by adenosine comprises of the inhibition or stimulation of adenylyl cyclase, the activation of phospholipase C, intracellular Ca\(_{2+}\) regulation, and the mitogen-activated protein kinases (MAPK) pathways\(^5\). Agonist-activated A\(_1\)R couples to inhibitory G proteins (G\(_{i/o}\)) to trigger numerous physiological effects. For example, it produces negative chronotropic and inotropic effects in the heart\(^6\), reduces the neuronal firing rate by blocking neurotransmitter release in the central nervous system (CNS), and inhibits lipolysis and renin release\(^7\).

Therapeutic approaches based on stimulating the A\(_1\)R could pave the way for hitherto unavailable tools to treat the central nervous system (CNS) and cardiovascular diseases\(^5,8\). From this standpoint, the SBDD of novel agonists can exploit the recent cryo-EM structure of A\(_1\)R in complex with both G\(_{i2}\) and adenosine\(^9\). Like all the other GPCRs, A\(_1\)R presents a transmembrane domain (TMD) formed by seven \(\alpha\)-helixes spanning the cytosolic membrane and shaping the orthosteric and the intracellular G\(_i\) protein binding sites (Figure S1a). Three intracellular loops (ICL1-3) and three extracellular loops (ECL1-3) interconnect the TM helices. ECL2, the longest A\(_1\)R loop, orients almost perpendicularly to the plane of the membrane both in the active\(^9\) and inactive\(^10,11\) A\(_1\)R states, in contrast to ECL2 of A\(_2A\)R, which is almost parallel to the membrane\(^10\). ECL2 is important for A\(_1\)R ligands\(^12\); it has been implicated in the intermediate states that
anticipate the orthosteric complex, therefore acting as a selectivity filter. Moreover, positive allosteric modulators can to bind to it\textsuperscript{13,14}. The key orthosteric interactions between adenosine and A\textsubscript{1}R (Figure S1b) are hydrogen bonds with residues N254\textsuperscript{6,55}, S277\textsuperscript{7,42}, H278\textsuperscript{7,43}, and a π-π interaction that involves F171\textsuperscript{ECL2}. AR agonists bearing small N\textsuperscript{6}-cycloalkyl groups, like the N\textsuperscript{6}-cyclopentyladenosine (CPA, Figure 1), display A\textsubscript{1}R selectivity\textsuperscript{15}. It is proposed that the ligand selectivity for A\textsubscript{1}R is driven by a small hydrophobic pocket underneath ECL3, due to the presence of T270\textsuperscript{7,35} in place of M270\textsuperscript{7,35} (A\textsubscript{2A})\textsuperscript{16,9,10}.

Here, we extensively studied the A\textsubscript{1}R recognition of adenosine, 5’-N-carboxamidoadenosine (NECA), CPA, and the recently characterized agonists HOCPA and BnOCPA\textsuperscript{17} (Figure 1). Binding and unbinding pathways were simulated by means of supervised molecular dynamics (SuMD)\textsuperscript{18,19} and the outcomes tested in mutagenesis experiments to identify A\textsubscript{1}R residues involved along the route towards and from the orthosteric site. We propose the binding conformation of N\textsuperscript{6}-cyclopentyl agonists, and how chemical modifications could impact the binding mechanisms of these selective A\textsubscript{1}R agonists. Our results can be framed within the dynamic nature of ligand-receptor complex formations and highlight the dramatic effect on drug binding kinetics triggered by chemical substitutions.
**Figure 1. A₁R agonists considered.** Adenosine is the endogenous effector; NECA is a nonselective exogenous ARs agonist; CPA represents a prototypical A₁R selective agonist; HOCPA and BnOCPA are A₁R selective analogues of CPA.

## 2 Methods

### 2.1 Experimental Methods

#### 2.1.1 Compounds

Adenosine, NECA ((2S,3S,4R,5R)-5-(6-aminopurin-9-yl)-N-ethyl-3,4-dihydroxyoxolane-2-carboxamide), CPA, were purchased from Sigma-Aldrich and dissolved in dimethyl-sulphoxide (DMSO). HOCPA and BnOCPA was synthesised as described in Knight *et al* (compounds 6 and 7 respectfully). CA200645, a high affinity AR xanthine amine congener (XAC) derivative containing a polyamide linker connected to the BY630 fluorophore, was purchased from HelloBio (Bristol, UK) and dissolved in DMSO. The concentration of DMSO was maintained to 1.1% for NanoBRET ligand-binding experiments using CA200645.

#### 2.1.2 Generation of mutant A₁R constructs

The NanoLLuc(Nluc)-tagged human A₁R pcDNA3.1+ construct used to generate stable HEK 293 cell lines was kindly gifted to us by Stephen Hill and Stephen Briddon (University of Nottingham). Mutations within the A₁R were made using the QuikChange Lightening Site-Directed Mutagenesis Kit (Agilent Technologies) in accordance with the manufacturer’s instructions. All oligonucleotides used for mutagenesis were designed using the online Agilent Genomics ‘QuikChange Primer Design’ tool and purchased from Merck. All constructs were confirmed by in-house Sanger sequencing.
2.1.3 Cell culture and transfection’s

HEK 293 cells in a single well of 6-well plate (confluency ≥ 80%) were transfected with 2 μg of DNA using polyethyleneimine (PEI, 1 mg/ml, MW = 25,000 g/mol) (Polysciences Inc) at a DNA:PEI ratio of 1:6 (w/v). Briefly, DNA and PEI were added to separate sterile tubes containing 150 mM sodium chloride (NaCl) (total volume 50 μl), allowed to incubate at room temperature for 5 minutes, mixing together and incubating for a further 10 minutes prior to adding the combined mix dropwise to the cells. 48 hours post-transfection, stable Nluc-A1R expressing HEK 293 cells were selected using 600 μg/mL Geneticin (Thermo Fisher Scientific) whereby the media was changed every two days. HEK 293 cell lines were routinely cultured in DMEM/F-12 GlutaMAX (Thermo Fisher Scientific) supplemented with 10% FBS (F9665, Sigma-Aldrich).

2.1.4 Analysis of A1R cell-surface expression using flow cytometry

HEK 293 cells and WT or mutant A1R expressing HEK 293 cells were harvested using a non-enzymatic cell dissociation solution (Sigma-Aldrich) and washed with PBS prior to counting. 0.5 x 10⁶ cells were washed three times in flow buffer (PBS supplemented with 1% BSA and 0.03% sodium azide) before re-suspending in 50 μl flow buffer containing anti-Nluc polyclonal primary antibody raised in rabbit (Kindly gifted by Promega) at 1:100 dilution and incubated at room temperature for 1 hour. All samples were washed three times with flow buffer and re-suspended in 50 μl flow buffer containing Allophycocyanin (APC)-conjugated anti-Rabbit IgG (Thermo Fisher Scientific, 31984) at 1:150 dilution and incubated for 1 hour at room temperature in the dark. The cells received a final three washes and were re-suspended in 300 μl flow buffer.

Analysis was conducted using a BD AccuriTM C6 Plus Flow Cytometer which is equipped with a blue (488 nm) and red (640 nm) laser, two light scatter detectors (FSC and SCC) and four
fluorescence detectors (FL1 Em. λ 530/30 nm, FL2 Em. λ 585/40 nm, FL3 Em. λ 570 and FL4 Em. λ 675/25 nm). FL4 optical filters were chosen APC (Ex. λ 633 nm and Em. λ 660 nm). Unstained cells and HEK 293 cells without A₁R expression were used as controls for autofluorescence and unspecific antibody binding, respectively. All data is collected by the flow cytometer and analysis can be conducted at any time using the BD AccuriTM C6 software. The images for figures were made in FlowJo® (V7.6.5), an analysis platform for single-cell flow cytometry analysis.

2.1.5 BRET assays for binding

Saturation binding at WT and mutant A₁R for the determination of CA200645 K_{D} was conducted using NanoBRET. The Nluc acts as the BRET donor (luciferase oxidizing its substrate, furimazine) and CA200645 acted as the fluorescent acceptor. Here, HEK 293 cells stably expressing WT or mutant Nluc-A₁R 24 hours post plating (10,000 cells/well of a 96-well plate) where pre-treated with 0.1 µM furimazine for 5 minutes prior to stimulated with CA200645 at a range of concentrations (0 to 300 nM). Filtered light emission at 450 nm and > 610 nm (640-685 nm band pass filter) was measured over a 30-minute period using a Mithras LB 940 following stimulation with a range of CA200645 concentrations and the raw BRET ratio calculated (610 nm/450 nm). Non-specific binding was determined by saturating concentrations of DPCPX (1 µM) and subtracted from the BRET ratio. CA200645 binding reached equilibrium after 5 minutes (Figure S2) and the BRET ratio was taken at 10 minutes and then fit to the ‘One site – Specific binding’ model built into Prism.

As described previously^{20}, NanoBRET competition binding assays were conducted to determine the affinity (pKi) of various A₁R compounds. Briefly, following a 5 minute pre-incubation with 0.1 µM furimazine, cells were co-stimulated with CA200645 (used at 25 nM, as previously reported^{20,21}) and increasing concentration of unlabeled ligand and emission at 450 nm and > 610
nm immediately measured. The BRET ratio at 10 minutes post stimulation was fitted with the ‘one-site – Ki model’ derived from the Cheng and Prusoff correction\textsuperscript{22}, built into Prism to determine affinity (pKi) values for all unlabeled agonists at the A\textsubscript{1}Rs. The determined K\textsubscript{d} of CA200645 at the mutant A\textsubscript{1}R was taken into account during analysis. Nonspecific binding was determined using a high concentration of unlabelled antagonist, at 1 µM DPCPX.

### 2.1.6 Data analysis

All experiments were conducted in duplicate (technical replicates) to ensure the reliability of single values. Statistical analysis, performed using Prism 8.0, was undertaken for experiments where the group size was at least n = 3 and these independent values used to calculate statistical significance (*, p< 0.05; **, p<0.01; ***, p<0.001; ****, p<0.0001) using a one-way ANOVA with a Dunnett’s post-test for multiple comparisons.

### 2.2 Computational Methods

#### 2.2.1 Biological targets and ligands force field parameters

All ten systems (Table 1) were prepared for molecular dynamics (MD) simulations using the CHARMM36\textsuperscript{23,24}/CGenFF 3.0.1\textsuperscript{25–27} force field combination. Initial ligand force filed topology and parameter files were obtained from the ParamChem webserver\textsuperscript{25}. Adenosine and NECA are already well-parameterized in the CGenFF force filed. Optimized parameters for HOCPA and BnOCPA from our previous work\textsuperscript{28} were used and transferred to CPA.

#### 2.2.2 Protein preparation

The active A\textsubscript{1}R structure (Table 1) was retrieved from the Protein Data Bank\textsuperscript{29} (PDB code 6D9H\textsuperscript{9}). For SuMD binding simulations, the agonists were placed at least 30 Å from the binding site in five different systems. For SuMD unbinding, the experimental coordinates (PDB 6D9H)
were used to simulate the adenosine, while representative frames (ligand orthosteric conformations close to the experimental bound adenosine on A₁R) from the SuMD binding were used to start the SuMD unbinding of NECA, CPA, HOCPA and BnOCPA. The A₁R intracellular loop 3 (ICL3) was modelled using Modeller 9.19\textsuperscript{30}. For all the 10 systems (Table 1), hydrogen atoms were added by means of the pdb2pqr\textsuperscript{31} and propka\textsuperscript{32} software (considering a simulated pH of 7.0); the protonation of titratable side chains was checked by visual inspection. The resulting receptor was inserted in a square 90 Å x 90 Å 1-palmitoyl-2-oleyl-sn-glycerol-3-phosphocholine (POPC) bilayer (previously built by using the VMD Membrane Builder plugin 1.1, Membrane Plugin, Version 1.1 at: http://www.ks.uiuc.edu/Research/vmd/plugins/membrane/), through an insertion method\textsuperscript{33}. The receptor orientation was obtained by superposing the coordinates on the corresponding structure retrieved from the OPM database\textsuperscript{34}. Lipids overlapping the receptor transmembrane helical bundle were removed and TIP3P water molecules\textsuperscript{35} were added to the simulation box by means of the VMD Solvate plugin 1.5 (Solvate Plugin, Version 1.5. at <http://www.ks.uiuc.edu/Research/vmd/plugins/solvate/>). Finally, overall charge neutrality was reached by adding Na\textsuperscript{+}/Cl\textsuperscript{−} counter ions up to the final concentration of 0.150 M, using the VMD Autoionize plugin 1.3 (Autoionize Plugin, Version 1.3. at <http://www.ks.uiuc.edu/Research/vmd/plugins/autoionize/>).

2.2.3 Systems equilibration and general MD settings

The MD engine ACEMD\textsuperscript{36} was employed for both the equilibration and productive simulations. The equilibration of the membrane systems was achieved in isothermal-isobaric conditions (NPT) using the Berendsen barostat\textsuperscript{37} (target pressure 1 atm) and the Langevin thermostat\textsuperscript{38} (target temperature 300 K) with low damping of 1 ps\textsuperscript{-1}. A four-stage procedure was performed (integration time step of 2 fs): first, clashes between protein and lipid atoms were reduced through 2000
conjugate-gradient minimization steps, then a 2 ns long MD simulation was run with a positional constraint of 1 kcal mol$^{-1}$ Å$^{-2}$ on protein and lipid phosphorus atoms. During the second stage, 20 ns of MD simulation were performed constraining only the protein atoms, while in the last equilibration stage, positional constraints were applied only to the protein backbone alpha carbons, for a further 20 ns. Globular protein equilibration was achieved in two steps: after 500 cycles of conjugate-gradient minimization, the system was simulated for 5 ns, employing an integration time step of 2 fs, in the isothermal-isobaric conditions (NPT). Productive trajectories (Table S4) were computed with an integration time step of 4 fs in the canonical ensemble (NVT). The target temperature was set at 300 K, using a thermostat damping of 0.1 ps$^{-1}$; the M-SHAKE algorithm$^{39,40}$ was employed to constrain the bond lengths involving hydrogen atoms. The cut-off distance for electrostatic interactions was set at 9 Å, with a switching function applied beyond 7.5 Å. Long-range Coulomb interactions were handled using the particle mesh Ewald summation method (PME)$^{41}$ by setting the mesh spacing to 1.0 Å.

### 2.2.4 The supervised MD (SuMD) protocol.

The supervised molecular dynamics (SuMD) is an adaptive sampling method$^{42}$ for speeding up the simulation of binding$^{18,43}$ and unbinding processes$^{19,44}$. In the simplest SuMD implementation, sampling is gained without the introduction of any energetic bias, by applying a tabu–like algorithm to monitor the distance between the centers of mass (or the geometrical centers) of the ligand and the predicted binding site or the receptor. However, the supervision of a second metric of the system can be considered$^{44}$. A series of short unbiased MD simulations are performed, and after each simulation, the distances (collected at regular time intervals) are fitted to a linear function. If the resulting slope is negative (for binding) or positive (for unbinding) the next
simulation step starts from the last set of coordinates and velocities produced, otherwise, the simulation is restarted by randomly assigning the atomic velocities.

**2.2.5 Settings for SuMD binding to the A₁R.**

To simulate the agonists’ binding to the A₁R (Table 1, Video S1-S5) the distance between the centroid of the ligand and the centroid of the orthosteric residues N254₆,₅₅, F17₁ₑ₇₂, T27₇₇,₄₂, and H2₇₈₇,₄₃ was supervised during 500 ns long time windows until it reached a value less than 4 Å. Then?

**2.2.6 Settings for SuMD unbinding from the A₁R.**

For the SuMD unbinding (Table 1, Video S1-S4), differently from the SuMD binding algorithm, the length (Δt) of the short simulations increases along the dissociation pathway, according to the formula:

\[ Δt = Δt₀Ntᵢ \]  \hspace{1cm} (1)

\( Δt₀ \) is the duration of the very first MD time window and \( Ntᵢ \) is a factor that is picked from three user-defined values (\( Nt₁ \), \( Nt₂ \), and \( Nt₃ \)), according to the last ligand-protein distance detected. At the end of each MD run, the ligand-protein distance (\( r_L \)) is compared to three distance threshold values (\( D₁ \), \( D₂ \) and \( D₃ \), also defined by the user), allowing a decision on the value of \( Ntᵢ \) factor according to the following conditions:

\[ r_L \leq D₁ \rightarrow Ntᵢ = 1 \]  \hspace{1cm} (2)

\[ D₁ < r_L \leq D₂ \rightarrow Ntᵢ = Nt₁ \]  \hspace{1cm} (3)

\[ D₂ < r_L \leq D₃ \rightarrow Ntᵢ = Nt₂ \]  \hspace{1cm} (4)

\[ D₃ < r_L \rightarrow Ntᵢ = Nt₃ \]  \hspace{1cm} (5)

Values of 3 Å, 5 Å and 8 Å were used for \( D₁ \), \( D₂ \) and \( D₃ \) respectively, while \( Nt₁ \), \( Nt₂ \), and \( Nt₃ \) were set to 2, 4, and 8 (SuMD time windows of 100 ps, 200 ps, 400 ps and 800 ps).
2.2.7 SuMD path sampling protocol

Further MD sampling (SuMD path sampling, Table 1) was performed using the outputs from each SuMD replica (Video S1-S4) for both binding and unbinding. Each trajectory was aligned on the protein alpha carbon atoms and the frames were clustered according to the ligand root mean square deviation (RMSD) to the starting positions (bin of 1 Å). A frame from each group was randomly extracted and used as a starting point for 20 ns long classic MD simulations. The number of classic MD simulations run varies according to the number of bins, and in turn with the linearity of the SUMD pathways. The BnOCPA binding\textsuperscript{28} and the adenosine SuMD binding and unbinding\textsuperscript{19} trajectories used to seed classic MD simulations were the same as our previous work.

2.2.8 Analysis of the MD trajectories

Only the MD trajectories from the SuMD path sampling were analyzed, as they should favor a relaxation of the systems towards local energy minima poorly sampled though SuMD simulations. Interatomic contacts and root mean square deviations (RMSD) were computed using VMD\textsuperscript{45}. Contacts were considered productive if the distance between two atoms was less than 3.5 Å. Ligand-protein hydrogen bonds were detected using the GetContacts scripts tool (https://getcontacts.github.io), setting a hydrogen bond donor-acceptor distance of 3.3 Å and an angle value of 150° as geometrical cut-offs. Contacts and hydrogen bond persistency are quantified as the percentage of frames (considering all the frames obtained by merging the different replicas) in which protein residues formed contacts or hydrogen bonds with the ligand. The computation takes into account direct and water-mediated interactions.

Distances between atoms were computed using PLUMED 2.3\textsuperscript{46}. The molecular mechanics generalized Born surface area (MM/GBSA) energy was computed with the MMPBSA.py\textsuperscript{47} script (AmberTools17 suite at http://ambermd.org/), after transforming the CHARMM psf topology files
to an Amber prmtop format using ParmEd (documentation at <http://parmed.github.io/ParmEd/html/index.html>). We preferred the MM/GBSA approach over the molecular Poisson-Boltzmann surface area (MM/PBSA) because binding and unbinding paths are not compatible with a grid method\textsuperscript{48}.

### 2.3 Numbering system

Throughout the manuscript, the Ballesteros-Weinstein residues numbering system for the GPCRs\textsuperscript{49} is adopted as superscript.

### Table 1. The ten systems simulated employing SuMD and SuMD path sampling.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>Transition</th>
<th># SuMD replicas</th>
<th>Total SuMD path sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenosine</td>
<td>Binding</td>
<td>9</td>
<td>5.18 $\mu$s</td>
</tr>
<tr>
<td></td>
<td>Unbinding</td>
<td>5</td>
<td>3.09 $\mu$s</td>
</tr>
<tr>
<td>NECA</td>
<td>Binding</td>
<td>6</td>
<td>1.86 $\mu$s</td>
</tr>
<tr>
<td></td>
<td>Unbinding</td>
<td>6</td>
<td>7.10 $\mu$s</td>
</tr>
<tr>
<td>CPA</td>
<td>Binding</td>
<td>9</td>
<td>4.34 $\mu$s</td>
</tr>
<tr>
<td></td>
<td>Unbinding</td>
<td>10</td>
<td>0.97 $\mu$s</td>
</tr>
<tr>
<td>HOCPA</td>
<td>Binding</td>
<td>8</td>
<td>4.06 $\mu$s</td>
</tr>
<tr>
<td></td>
<td>Unbinding</td>
<td>6</td>
<td>2.58 $\mu$s</td>
</tr>
<tr>
<td>BnOCPA</td>
<td>Binding</td>
<td>7</td>
<td>5.82 $\mu$s</td>
</tr>
<tr>
<td></td>
<td>Unbinding</td>
<td>6</td>
<td>4.92 $\mu$s</td>
</tr>
</tbody>
</table>
3 Results

3.1 N⁶-cyclopentyl agonists bind to A₁R with similar fashion

ARs ligands bearing an alkyl ring in position N⁶ (on the adenine scaffold) or C8 (xanthine scaffold) display increased affinity for A₁R over A₂AR\textsuperscript{10}. The reason for this is attributed to T270\textsuperscript{7.35} (A₁R) in place of M270\textsuperscript{7.35} (A₂AR), which shapes a hydrophobic sub-pocket underneath ECL3 that partially accommodates the lipophilic substituent. While the inactive structures of A₁R and A₂AR in complex with the xanthine antagonist PSB36\textsuperscript{10} have univocally shown this structural aspect of the selectivity, no confirmation from crystallography or cryo-EM studies is yet available for ARs agonists.

Besides reproducing the adenosine cryo-EM binding conformation into A₁R (Figure S3, Video S1)\textsuperscript{9}, our simulations sampled the orthosteric binding mode of NECA observed on A₂AR (Figure S3, Video S2)\textsuperscript{50}. In light of this reliability, we propose the likely binding mode of CPA, HOCPA, and BnOCPA (Figure 2, Videos S3-S4). As expected, the adenine ring forms a bidentate hydrogen bond with N254\textsuperscript{6.55} and a π-π stacking with F171\textsuperscript{ECL2}, while the N⁶-cyclopentyl ring inserts in the hydrophobic pocket under ECL3, interacting with T270\textsuperscript{7.35} and L253\textsuperscript{6.54} (Figure 2). Mutagenesis experiments confirmed the importance of L253\textsuperscript{6.54} (Table 2) for the affinity of the agonists. BnOCPA is proposed to bind to A₁R with same features of the smaller ligands CPA and HOCPA (Figure 2b, Video S4). However, during the simulations the oxybenzyl group showed\textsuperscript{28} high flexibility and explored three different orientations (Figure 2b). In two of these conformations, BnOCPA interacted with A₁R residue L258\textsuperscript{6.59} (mode B, Figure 2b) and Y271\textsuperscript{7.36} (mode C, Figure 2b). Binding assays with mutants L258\textsuperscript{6.59}A, F258\textsuperscript{6.59}A, F258\textsuperscript{6.59}T and Y271\textsuperscript{7.36}A (Table 2) confirmed that these two residues are likely involved in the orthosteric complex with BnOCPA.
However, a role during agonist association and dissociation events cannot be ruled out, as also the affinity for CPA and HOCPA (which are not involved in contacts with L258^{6.59} and Y271^{7.36} in the bound state, Figure 2a) was affected (Table 2).

Figure 2. Binding modes of the agonists according to SuMD path sampling simulations. a) CPA (magenta) and HOCPA (pink) engage A_{1}R with the same orientation as adenosine (tan stick); the N^{6}-
cyclopentyl group interacts with L253 and T257. B) BnOCPA orients the oxybenzyl group in three different orientations. Hydrogen bonds with N254 are shown as dashed lines, while hydrophobic contacts are depicted as cyan transparent surfaces.

3.2 Simulations suggest the binding and unbinding paths of A₁R agonists

SuMD and SuMD path sampling delivered detailed insights on the possible transitory metastable states that the agonists experienced along the route to, and from, the orthosteric site (Figure 3, Table S1-S4, Figure S5, Figure S6; Video S1-S4). While the endogenous agonist adenosine reached the orthosteric site with very limited intermediate interactions with ECL2 (Figure 3b, Table S1, Figure S5, Video S1), the other ligands formed contacts with this extracellular vestibule, with CPA and BnOCPA most involved in metastable states (Figure 3d,h, Table S1, Figure S5, Video S3, Video S5). As a general view, the increase in lipophilicity at the N₆ position (CPA, HOCPA and BnOCPA, Figure 3d,f,h, Table S1, Figure S5) and 5’ position (NECA, Figure 3, Figure S4a, Table S1, Figure S5) favored intermediate interactions with ECL2. The importance of ECL2 for the binding of NECA, CPA and the antagonist DPCPX to A₁R has been recently demonstrated. Our simulations suggest I175 and E172 as involved in interactions during both the binding and unbinding (Figure 3d, Figure S4, Figure S6), while E170, I167, N159, and W156 engaged NECA during the binding (Figure S4a, Figure S6). L149, on the other hand, was not involved in direct interactions with the ligands, suggesting an important role in stabilizing the overall secondary structure of ECL2 due to its position at the base of the loop helix. Besides the aforementioned residues, further A₁R side chains were involved along the simulated binding paths for all the agonists considered (Table S1, Table S2 Figure S5).
SuMD unbinding routes (Figure 3c,e,g,i, Video S1-S4) had only limited overlap with the binding routes (Figure 3b,d,f,h, Video S1-S4). The agonists, indeed, established frequent interactions with the top of TM1, TM2, and TM7 (Figure 3c,e,g,i, Figure S4b, Table S2, Table S4, Figure S6). However, in analogy with the binding simulations, CPA, HOCPA, and BnOCPA (all bearing an N\textsuperscript{6} hydrophobic moiety) showed a major involvement of ECL2 (Table S3, Table S4, Figure S6; Video S3-S4). Many of the A\textsubscript{1}R residues involved in both agonist association (Table S1, Table S2, Figure S5) and dissociation (Table S3, Table S4, Figure S6) are part of the orthosteric site. It is therefore not surprising that mutagenesis experiments already pointed out their importance\textsuperscript{12,52–56}.

Interestingly, our unbiased nonequilibrium simulations pointed out several residues, located outside the orthosteric site, involved along the binding and unbinding routes. These residues comprise F8\textsuperscript{1.31}, Q9\textsuperscript{1.32}, Y12\textsuperscript{1.35}, I69\textsuperscript{2.64}, N70\textsuperscript{2.64}, N148\textsuperscript{ECL2}, I175\textsuperscript{ECL2}, L253\textsuperscript{6.54}, T257\textsuperscript{6.58}, L258\textsuperscript{6.59}, H264\textsuperscript{ECL3}, K265\textsuperscript{ECL2}, S267\textsuperscript{7.32}, Y271\textsuperscript{7.36} (Figure 3a). Mutagenesis experiments were designed and performed to confirm computational predictions (Figure 3, Table 2, Table S5).
Figure 3. Simulated binding and unbinding paths of the agonists and the relative position of the A₁R mutants tested. a) Two side views of the A₁R (white transparent ribbon); residues considered for the mutations (Table 2) are shown as grey sticks. The cryo-EM bound adenosine (tan stick) is reported as reference. b) to i) Left-hand panels, position of the agonist centroid during simulations, colored according to the interaction energy with A₁R (white ribbon); residues mutated (Table 2) are shown as sticks and colored according to effect on the affinity (red: decreased affinity; green: increased affinity; black: unaltered affinity). Right-hand panels, A₁R-agonist contacts plotted onto the protein surface and colored according to the contacts occupancy. b) Adenosine binding simulations; c) adenosine unbinding simulations; d) CPA binding simulations; e) CPA unbinding simulations; f) HOCPA binding simulations; g) HOCPA unbinding simulations; h) BnOCPA binding simulations; i) BnOCPA unbinding simulations.

3.3 Mutagenesis experiments reveal novel A₁R residues involved in the binding of agonists

3.3.1 Surface expression of the A₁R alanine mutants and CA200645 Kd

The cell-surface expression of WT and mutant A₁R, as determined by mean fluorescence intensity of APC (%WT), was comparable to WT for Y12¹.35A, N70².64A, T257⁶.⁵⁸A, L258⁶.⁵⁹A, K265ECL²A, S267⁷.³²A, L258⁶.⁵⁹T and L258⁶.⁵⁹F A₁R. When compared to WT, the cell-surface expression was determined to be significantly reduced for F⁸¹.³¹A, Q⁹¹.³²A, N148ECL²A, I175ECL²A, H264ECL²A and L258⁶.⁵⁹G whereas it was significantly increased for I69².⁶⁵A, L25³.⁵⁴A and Y27¹.⁷⁶A. This enhanced or reduced mutant A₁R cell-surface expression did not correlate with changes in CA200645 equilibrium dissociation constant (Kd). For example, the Kd for CA200645 determined in L258⁶.⁵⁹T was significantly increased when compared to WT but showed comparable cell-surface expression. I175ECL²A and L258⁶.⁵⁹G also showed an increased Kd but had a reduced cell-surface expression when compared to WT.
Importantly, the determined changes in CPA, BnOCPA, HOCPA or adenosine affinity (pKi) were determined to be independent of changes in cell-surface expression. For example, despite the significantly elevated cell-surface expression of Y2717.36A, the determined compound affinity of all four tested compounds was significantly reduced when compared to that determined at WT. This is likely due to the high receptor expression in our system and/or reserve. The only exception was L2586.59G whereby compound pKi could not be determined due to the likely combined effect of low cell-surface expression and reduced CA200645 affinity.

3.3.2 General effects of alanine substitutions on agonists affinity

A number of mutations enhanced affinity, while other either decreased affinity or had no overall effect. The mutated residues can be divided into three groups according to their positions relative to the (un)binding paths sampled during the simulations.

3.3.3 TM1, TM2 and TM7 residues

The A1R residues located at the top of TM1, TM2 and TM7 (Figure 3a) were generally involved during both the simulated binding (Figure 3b,d,f,h, Figure S5, Figure S6) and unbinding (Figure 3c,e,g,i, Figure S5, Figure S6). Adenosine, CPA, and HOCPA displayed decreased affinity to A1R mutants Q91.32A and Y121.35A (Table 2) and enhanced the binding to F81.31A compared to the WT, in agreement with the transitory interactions formed with TM1 during simulations (Figure 3b-g, Table S1-S4, Figure S5, Figure S6). The affinity for the agonists diminished on Q91.32A and Y121.35A, but BnOCPA was not significantly affected by F81.31A. This is apparently in disagreement with simulations, which proposed BnOCPA as the agonist more prone to form metastable states in the proximity of F1.318 during the dissociation from the receptor (Figure 3i). However, the fact that BnOCPA was the ligand most prone to interact with F8 also during the
association (Figure 3h) may suggest a degree of compensation between binding and unbinding. That is, the variation of the binding rate can be compensated by an opposite change in the unbinding rate (and vice versa) resulting in unchanged affinity. This could also be the case of N70²⁶⁵A, which did not alter the affinity of the agonists, except for BnOCPA (Table 2). Binding and unbinding simulations suggested that N70²⁶⁵ forms numerous interactions with the ligands (Table S1-S4, Figure S5, Figure S6). The adjacent residue I69²⁶⁴, instead, when mutated to alanine (I69²⁶⁴A) significantly decreased the affinity of all the agonists (Table 2). Besides being involved in frequent interactions with the agonists (Table S1-S4), I69²⁶⁴ is packed in hydrophobic contacts with TM3, possibly stabilizing the neighboring part of ECL2.

Moving to TM7, the agonists NECA, CPA, HOCPA, and BnOCPA, but not adenosine, displayed diminished binding to the S267⁷.32A (Table 2, Table S5), while all of them lost affinity to Y271⁷.36A (Table2, Table S5). The bulky Y271⁷.36 side chain, which occupies an important position at the interface between the orthosteric site and the extracellular vestibule, could participate in numerous intermediate interactions along the (un)binding routes.

3.3.4 TM6 residues

The three A₁R residues mutated on TM6 (L253⁶.54, T257⁶.58, and L258⁶.59, Figure 3a) are part of different protein environments. L253⁶.54 shapes part of the hydrophobic pocket underneath ECL3 partially responsible for A₁/A₂A ligands selectivity (along with L269⁷.₃⁴ and T270⁷.₃⁵). All the agonists displayed diminished affinity to L253⁶.54A (Table 2), with the N⁶-substituted agonists CPA, HOCPA, and BnOCPA most affected, in line with the hydrophobic interactions occurring in the bound state (Figure 2). Interestingly, T257⁶.58A increased the affinity of CPA, HOCPA, and BnOCPA. This could be due to an increase in lipophilicity of the protein environment surrounding
the cyclopentyl group either in the bound complex or along the unbinding pathways, (Figure 3e,g,i). NECA showed reduced affinity to T257.58A (Table S5, Figure S4), confirming the importance of hydrophobic N6-substituents for interactions with the top of TM6. L258.59, which is located at the interface with the membrane, is one of the TM5 and TM6 residues shaping a saddle between ECL2 and ECL3, where agonists tended to form metastable interactions along the simulated (un)binding paths (Figure 3b-i, Figure S5, Figure S6). All the agonists (excepted adenosine) displayed a reduced affinity for L258.59A and L258.59T, suggesting that the bulkier ligands may be more prone to interact with this part of A1R (BnOCPA was the only ligand proposed to interact at some extent with L258.59 in mode B, Figure 2b). The affinity decrease (Table 2) displayed by HOCPA and BnOCPA for L258.59F likely excludes a destabilization of the neighbor protein structure, as residue F258.59 WT A2A R does not change the conformation of the top of TM6.

3.3.5 ECL2 and ECL3 residues

As I175ECL2 formed numerous interactions with the agonists during the simulations (Figure 3b-i, Table S1, Table S2, Figure S5, Figure S6), not surprisingly the I175ECL2A mutation reduced the affinity of all the ligands (Table 2, Table S5). I175ECL2 could contribute to keeping the aromatic side chain of F171ECL2 in an appropriate conformation for interacting with the ligands’ adenine ring in the orthosteric complex (Figure S1). On the other hand, N148ECL2A did not affect the affinity of the agonists, despite the frequent interactions during the binding simulations (Table 2, Figure 3b,d,f,h, Table S1, Table S2, Figure S5, Figure S6).

While H264ECL2A showed diminished affinity (Table 2) for CPA and BnOCPA (which have the most lipophilic N6-group, Figure 1), none of the tested ligands were significantly affected by
K265ECL2A. As K265ECL2 is part of a stable salt bridge with E172ECL2, K265ECL2A is expected to affect the affinity of the ligands due to the reduced hindrance of the orthosteric site. The subtype A2AR bears A265ECL2 in place of K265ECL2 and the residue involved in the salt bridge with E169ECL2 (corresponding to E172ECL2 in A1R) is H264ECL2. The A2AR H264ECL2A mutant does not display a modified affinity for the antagonists ZM241385, despite the increase in off rate57. This could indicate a kinetic compensation due to a faster binding to H264ECL2A (A2AR) and K265ECL2A (A1R). A less bulky alanine side chain, indeed, would favor ligand binding and the unbinding to a similar extent.
Table 2. NanoBRET saturation- and competition-binding assays in WT and mutant Nluc-A1R. CA200645 equilibrium dissociation constant (K\(_d\)) and compound affinity (pK\(_a\)) at WT and mutant Nluc-A1R, as determined by NanoBRET saturation and competition ligand-binding assays, respectively.

<table>
<thead>
<tr>
<th>Cell-surface expression(^a)</th>
<th>n</th>
<th>K(_d) (^b)</th>
<th>n</th>
<th>CPA</th>
<th>n</th>
<th>BnOCPA</th>
<th>n</th>
<th>HOCPA</th>
<th>n</th>
<th>Adenosine</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>100 ±0.1</td>
<td>8</td>
<td>68.6 ±7.5</td>
<td>5</td>
<td>6.95 ±0.02</td>
<td>3</td>
<td>6.21 ±0.04</td>
<td>4</td>
<td>6.37 ±0.03</td>
<td>4</td>
<td>5.90 ±0.08</td>
</tr>
<tr>
<td>F8(^{31}) A</td>
<td>53 ±12**</td>
<td>3</td>
<td>90.7 ±21.5</td>
<td>3</td>
<td>7.34 ±0.08*</td>
<td>4</td>
<td>6.37 ±0.02</td>
<td>3</td>
<td>6.95 ±0.07**</td>
<td>3</td>
<td>6.55 ±0.17***</td>
</tr>
<tr>
<td>Q9(^{32}) A</td>
<td>21 ±2****</td>
<td>3</td>
<td>60.6 ±5.4</td>
<td>3</td>
<td>6.41 ±0.09**</td>
<td>3</td>
<td>5.44 ±0.06****</td>
<td>3</td>
<td>5.71 ±0.05***</td>
<td>3</td>
<td>5.07 ±0.11****</td>
</tr>
<tr>
<td>Y12(^{135}) A</td>
<td>120 ±19</td>
<td>3</td>
<td>75.8 ±6.0</td>
<td>3</td>
<td>6.50 ±0.06**</td>
<td>4</td>
<td>5.44 ±0.06****</td>
<td>4</td>
<td>5.78 ±0.10***</td>
<td>4</td>
<td>5.35 ±0.05****</td>
</tr>
<tr>
<td>I69(^{2,64}) A</td>
<td>158 ±7***</td>
<td>3</td>
<td>147.2 ±13.4*</td>
<td>3</td>
<td>4.94 ±0.06****</td>
<td>3</td>
<td>4.72 ±0.04****</td>
<td>3</td>
<td>4.34 ±0.11****</td>
<td>3</td>
<td>4.64 ±0.13****</td>
</tr>
<tr>
<td>N70(^{2,55}) A</td>
<td>86 ±9</td>
<td>3</td>
<td>51.6 ±3.7</td>
<td>3</td>
<td>6.97 ±0.12</td>
<td>3</td>
<td>5.92 ±0.06*</td>
<td>3</td>
<td>6.05 ±0.09</td>
<td>3</td>
<td>5.83 ±0.05</td>
</tr>
<tr>
<td>N148(^{ECL2}) A</td>
<td>13 ±1****</td>
<td>3</td>
<td>70.7 ±13.2</td>
<td>3</td>
<td>7.20 ±0.13</td>
<td>3</td>
<td>6.15 ±0.03</td>
<td>3</td>
<td>6.15 ±0.05</td>
<td>4</td>
<td>5.71 ±0.07</td>
</tr>
<tr>
<td>I178(^{ECL2}) A</td>
<td>23 ±3 ***</td>
<td>3</td>
<td>153.2 ±27.6**</td>
<td>3</td>
<td>6.30 ±0.06****</td>
<td>3</td>
<td>5.60 ±0.05****</td>
<td>3</td>
<td>5.66 ±0.02****</td>
<td>4</td>
<td>5.00 ±0.11****</td>
</tr>
<tr>
<td>T257(^{6,58}) A</td>
<td>114 ±19</td>
<td>4</td>
<td>128.8 ±16.0</td>
<td>3</td>
<td>7.30 ±0.03**</td>
<td>3</td>
<td>6.77 ±0.05****</td>
<td>4</td>
<td>6.94 ±0.05****</td>
<td>3</td>
<td>5.85 ±0.09</td>
</tr>
<tr>
<td>L253(^{5,64}) A</td>
<td>160 ±24****</td>
<td>3</td>
<td>104.1 ±19.3</td>
<td>3</td>
<td>6.17 ±0.06****</td>
<td>4</td>
<td>5.39 ±0.10****</td>
<td>4</td>
<td>5.49 ±0.13****</td>
<td>3</td>
<td>5.40 ±0.06**</td>
</tr>
<tr>
<td>L258(^{59}) A</td>
<td>82 ±2</td>
<td>3</td>
<td>118.3 ±20.6</td>
<td>3</td>
<td>6.34 ±0.07****</td>
<td>6</td>
<td>6.03 ±0.05****</td>
<td>5</td>
<td>5.85 ±0.08***</td>
<td>5</td>
<td>5.66 ±0.08</td>
</tr>
<tr>
<td>H264(^{ECL3}) A</td>
<td>21 ±6****</td>
<td>3</td>
<td>84.1 ±4.6</td>
<td>3</td>
<td>6.33 ±0.09****</td>
<td>3</td>
<td>5.60 ±0.01****</td>
<td>3</td>
<td>6.02 ±0.18</td>
<td>3</td>
<td>6.21 ±0.05</td>
</tr>
<tr>
<td>K265(^{ECL3}) A</td>
<td>68 ±13</td>
<td>3</td>
<td>54.7 ±8.6</td>
<td>3</td>
<td>6.59 ±0.13</td>
<td>3</td>
<td>6.21 ±0.05</td>
<td>4</td>
<td>6.01 ±0.12</td>
<td>4</td>
<td>5.58 ±0.09</td>
</tr>
<tr>
<td>S267(^{7,36}) A</td>
<td>90 ±6</td>
<td>3</td>
<td>89.9 ±11.7</td>
<td>3</td>
<td>6.66 ±0.06*</td>
<td>3</td>
<td>5.85 ±0.01***</td>
<td>3</td>
<td>5.99 ±0.05***</td>
<td>3</td>
<td>5.66 ±0.08</td>
</tr>
<tr>
<td>Y271(^{3,6}) A</td>
<td>171 ±10****</td>
<td>3</td>
<td>94.4 ±11.1</td>
<td>3</td>
<td>5.82 ±0.03****</td>
<td>3</td>
<td>5.14 ±0.07****</td>
<td>4</td>
<td>5.15 ±0.07****</td>
<td>3</td>
<td>5.15 ±0.05****</td>
</tr>
<tr>
<td>L258(^{59,T})</td>
<td>77 ±18</td>
<td>3</td>
<td>200.4 ±24.6****</td>
<td>3</td>
<td>6.39 ±0.07</td>
<td>3</td>
<td>5.69 ±0.04****</td>
<td>4</td>
<td>5.69 ±0.04****</td>
<td>4</td>
<td>5.45 ±0.05</td>
</tr>
<tr>
<td>L258(^{59,F})</td>
<td>131 ±2</td>
<td>4</td>
<td>123.0 ±17.3</td>
<td>3</td>
<td>6.40 ±0.10</td>
<td>3</td>
<td>6.12 ±0.05****</td>
<td>3</td>
<td>5.79 ±0.08***</td>
<td>3</td>
<td>5.47 ±0.04</td>
</tr>
<tr>
<td>L258(^{59,G})</td>
<td>15 ±4****</td>
<td>3</td>
<td>143.3 ±37.5*</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

\(^a\) Mean fluorescence intensity of APC (% Wild-type (WT)). Cell-surface expression of WT or mutant Nluc-A1R in HEK 293 cells was determined by flow cytometry. Cells were incubated with anti-Nluc antibody followed by APC-conjugated anti-rabbit IgG secondary antibody and the fluorescence detector FL4 (Em. \(\lambda\) 675/25 nm) used to detect APC fluorescence.

\(^b\) CA200645 equilibrium dissociation constant (K\(_d\)) as determined by NanoBRET saturation binding assays.
Compound affinity ($pK_i$) (mean ± SEM) determined through NanoBRET competition-binding assays in WT/mutant Nluc-A1R stably expressing HEK 293 cells. The resulting concentration-dependent decrease in BRET ratio at 10 minutes was used to calculate $pK_i$. Compound affinity could not be determined ( - ) for L253G Nluc-A1R given the low cell-surface expression and reduced binding affinity of CA200645 ($K_d$).

Data are expressed as mean ± SEM obtained in $n$ separate experiments. All individual experiments were conducted in duplicate. Statistical significance (*, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$) compared to WT was determined by one-way ANOVA with Dunnett’s post-test.
4. Discussion

The association and dissociation of a ligand from its receptor is a multistep process, characterized by intermediate metastable states that anticipate (e.g. during the association) or follow (e.g. during the dissociation) the bound state. In this complex scenario, further complicated by more than one possible pathway, it is difficult to rationalize the role played by GPCR metastable binding sites. Preliminary interactions with the EC vestibules, for example, act as selectivity filters by favoring the recognition of a chemotype over other molecules58,59, in some cases driving the affinity and the selectivity60. However, if the stability of an intermediate state is similar to the orthosteric one, segregation may occur, that is, the displayed ligand affinity for the receptor diminishes61.

For the A1R, as a general view, the agonist’s binding and unbinding occurred through two major pathways, namely Path A and Path B (Figure 4a). Path A comprises residues located at the top of TM5, TM6 (L2536.54, T2576.58 and L2586.59) and on the distal region of ECL2 (I175ECL2 and N148ECL2), while Path B is well defined by residues on the top of TM1, TM2, TM7 (F81.31, Q91.32, Y121.35, I692.64, N702.65, S2677.32 and Y2717.36). The analysis of the position of the residues in terms of path and distance from the orthosteric site, the occupancy of the contacts formed during the binding and unbinding, and the effect of the mutation to alanine on the pKi, highlights a complicated pattern (Figure 4b). The endogenous agonist adenosine, which is the smallest ligand considered here, was the least affected by mutations; this is consistent with the relatively straight (un)binding pathways sampled during MD simulations (Figure 3a,b). The mutation of the hydrophobic residues L2536.54, I175ECL2 and I692.64, that are close to the orthosteric site, decreased the pKi due to reduced apolar contacts with the ligands (Figure S5, Figure S6). T2576.58A, which augmented the hydrophobicity of the pocket underneath ECL3, increased the binding affinity of
all the agonists bar adenosine, in line with the role of the N\textsuperscript{6} hydrophobic substituent. The other mutation, beside T257\textsuperscript{6.58}A, that increased the pKi (BnOCPA excepted) was F8\textsuperscript{1.31}A (Figure 4b). F8\textsuperscript{1.31} is located far away from the orthosteric site and is involved in the contacts between TM1 and TM7, therefore its mutation to alanine could have local effects on the neighboring side chains or the flexibility of the TM1 N-terminal.

Strikingly, the mutation of N148\textsuperscript{ECL2}, located distant from the orthosteric site along Path A and highly involved during the binding (Figure 4b, Figure 3) did not change the pKi of any of the agonists, in line with previous findings\textsuperscript{12}. We speculate that its effect on the overall kinetics of binding is neglectable because it is not part of the secondary hydrophobic binding site on ECL2 that has been associated to A\textsubscript{1}R allostery\textsuperscript{13,62} and highlighted also in our simulations (Figure 3). Moving to residues located along Path B, N70\textsuperscript{2.65}A negatively affected only BnOCPA pKi. Notably, BnOCPA was highly prone to hydrogen bond with N70\textsuperscript{2.65} during the unbinding but N70\textsuperscript{2.65} was less involved during the binding (Figure S5, Figure S6). This unbalance was not mirrored by the other agonists, which instead interacted with N70\textsuperscript{2.65} during binding and unbinding and therefore experienced a sort of compensation, with the final effect of not affecting the pKi. This binding/unbinding compensation is evident in the case of K265\textsuperscript{ECL3}A. K265\textsuperscript{ECL2} forms a salt bridge with E172\textsuperscript{ECL2} that is likely to contribute to the (un)binding kinetics of A\textsubscript{1}R ligands, in analogy with the A\textsubscript{2}R salt bridge formed between H264\textsuperscript{ECL3} and E169\textsuperscript{ECL2} \textsuperscript{57}. It follows that the disruption of the K265\textsuperscript{ECL2} – E172\textsuperscript{ECL2} salt bridge by mutating K265\textsuperscript{ECL2} to alanine should speed up both the association, as an agonist find a more accessible orthosteric site, and the dissociation, because of less hindrance for the transition to the EC vestibules.
Figure 4. Scrambled plot of the mutated A₁R residues according to the distance from the orthosteric site along Path A or Path B (a), and the effect of the mutation on the pKi. The contacts formed with the agonists are represented as letters of dimension proportional to the occupancy; contacts during the binding are
reported in the first column, while contacts from unbinding are reported in the second column; residues in close proximity of the orthosteric site are underlined (b).

4.1 Towards the definition of the structure-binding/unbinding path relationships of A₁R agonists.

The (un)binding kinetics have a major impact on both pharmacodynamics\textsuperscript{63–65} and pharmacokinetics\textsuperscript{66} of a drug. Even small structural modifications within a congeneric series of ligands modify the kinetics through modified on and off rates\textsuperscript{67}. The reason for this lies in the enormously higher number of different intermediate states that a ligand can experience along (un)binding paths compared to the orthosteric complexes, where it is restrained by intermolecular interactions and steric hindrances. Protein-ligand recognition events have very complex energy landscapes\textsuperscript{68–71} and small changes in either the structure of the ligand or the protein can alter the nature and the position of the transition states along the (un)binding routes\textsuperscript{72}. To investigate this aspect on the A₁R agonists, we compared the A₁R interaction patterns between adenosine, CPA or BnOCPA (Figure 5) to understand how the introduction of the N\textsuperscript{6} cyclopentyl group on the adenosine scaffold and the introduction of the oxybenzyl group on CPA (BnOCPA) could alter the overall (un)binding mechanisms.

4.1.1 The N\textsuperscript{6}-cycloalkyl ring on the adenosine scaffold.

In general, CPA formed more intermediate interactions than adenosine with the extracellular vestibule of A₁R (Figure 5a,b). The presence of the N\textsuperscript{6}-cycloalkyl substituent produced mode contacts with ECL2 and residues located at the top of TM1, TM5, TM6, TM7, and ECL3 (Figure
5a,b). This scenario is in good agreement with mutagenesis experiments (Table 2) showing CPA affinity (but not adenosine’s affinity) was significantly affected by T2576.58A, L2586.59A, H264ECL3A, and S2677.32A, all located in the proximity of the extracellular vestibule.

4.1.2 The 2-oxybenzyl group on CPA N6-cycloalkyl ring.

The presence of this lipophilic moiety (BnOCPA was the bulkiest ligand considered) favored more interactions with the top of TM5 and TM6 during association (Figure 5c) and with the top of TM1 and TM2 during dissociation (Figure 5d). CPA, on the other hand, was more prone to interact with the ECL2 (Figure 5c,d). The changes in the barycenter of the contacts during unbinding due to the 2-oxybenzyl group could explain the unique profile that BnOCPA showed on A1R mutants F81.31A and N702.65A (Table 2). BnOCPA, indeed, was the only agonist significantly affected by N702.65A, and the only one not affected by F81.31A. We speculate that the F81.31A mutation affects both BnOCPA binding and unbinding (Figure 5c,d), leading to compensation and therefore an unmodified pKi compared to A1R wt (Table 2).
Figure 5. Structural modification of A₁R agonists led to different simulated binding and unbinding mechanisms. a) and b) Heatmap showing the contact difference, during SuMD path sampling, between adenosine and CPA, which differ for the N⁶-cyclopieryl group (highlighted as a dashed circle line); A₁R blue surfaces indicates more contacts formed with adenosine, while red surfaces are indicative of more contacts with CPA. a) Binding simulations; b) Unbinding simulations. c) and d) Heatmap showing the contacts difference, during SuMD path sampling, between CPA and BnOCPA, which differ for the oxybenzyl group (highlighted as a dashed circle line); A₁R blue surfaces indicate more contacts formed with CPA, while red surfaces are indicative of more contacts with BnOCPA c) Binding simulations; d) Unbinding simulations.

5. Conclusion

GPCR-ligand complexes form and dissociate through multistep mechanisms. The importance of intermediate metastable states along the (un)binding paths in modulating the overall affinity of ARs agonist and GPCRs ligands in general is gradually emerging¹²,⁶₀,⁶¹. Here we combined unbiased nonequilibrium MD simulations and mutagenesis experiments to study the dynamic binding and unbinding of A₁R agonists. The in silico analysis suggested several A₁R residues involved on regions so far poorly investigated⁷³. Mutagenesis experiments generally confirmed the computational prediction and allow mapping novel receptor vestibules involved in the association or dissociation of the agonists. For binding, the A₁R ECL2 was involved in numerous preliminary contacts, as well as the top of TM6, TM1, TM2 and TM7. The dissociation from A₁R followed similar but distinct routes, with ECL2 generally less engaged (especially by the more hydrophilic adenosine and NECA). The chemical modifications that increase the agonists’ selectivity towards A₁R (e.g. the introduction of N⁶-cycloalkyl groups) also changed the
(un)binding mechanism, favoring the interactions with the extracellular vestibules in general, and the hydrophobic allostERIC pocket located on ECL2 in particular.

MD simulations and mutagenesis experiments are frequently combined to deliver structural insights on endpoint protein-ligand complexes, however, to the best of our knowledge\textsuperscript{74-76} this is the first time that the whole process of formation/dissociation of several GPCR ligands is mapped with a combined \textit{in silico} and \textit{in vitro} approach. Our results pave the way for the rationalization of SKR for A\textsubscript{1}R, agonists, with potential repercussion on the design of long-awaited clinical agents.
ASSOCIATED CONTENT

The following files are available free of charge:

Supporting Information (pdf)

Video S1-S4 (mp4)

Preprint DOI: https://doi.org/10.1101/2020.10.22.350827

AUTHOR INFORMATION

Present Addresses

§ New address: Sosei Heptares, Granta Park Steinmetz Building, Cambridge, CB21 6DG, UK

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

† These authors contributed equally

‡ These authors contributed equally

* Corresponding authors: ad5288@coventry.ac.uk ; grl30@cam.ac.uk ; ad5291@coventry.ac.uk

Funding Sources

Leverhulme Trust (RPG-2017-255, CAR and GL to fund KB and GD).

ACKNOWLEDGMENT

CAR is grateful for a Royal Society Industry Fellowship.
REFERENCES


22. Cheng Y, Prusoff WH. Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem Pharmacol*. 1973;22(23):3099-3108. doi:10.1016/0006-2952(73)90196-2


49. Ballesteros JA, Weinstein H. Integrated methods for the construction of threedimensional models and computational probing of structure-function relations in G protein-


Deciphering the Agonist Binding Mechanism to the Adenosine A1 Receptor

Giuseppe Deganuti¹, Kerry Barkan²-§, Barbara Preti³, Michele Leuenberger³, Mark Wall⁴, Bruno G. Frenguelli⁴, Martin Lochner³, Graham Ladds² and Christopher A Reynolds¹

¹ Centre for sport, Exercise and Life Sciences, Faculty of Health and Life Sciences, Coventry University, Alison Gingell Building, CV1 5FB, UK

² Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge, CB2 1PD, UK

³ Institute of Biochemistry and Molecular Medicine, University of Bern, 3012 Bern, Switzerland

⁴ School of Life Sciences, University of Warwick, Gibbet Hill Rd, Coventry CV4 7AL, UK

§ New address: Sosei Heptares, Granta Park Steinmetz Building, Cambridge, CB21 6DG, UK

The (un)binding path of several agonists of the adenosine A1 receptor (A₁R) was reconstructed by supervised molecular dynamics (SuMD) simulations and tested through mutagenesis experiments. Residues most involved during the recognition were proposed, highlighting the effect that ligand chemical modifications have on the (un)binding mechanism.