JCTC Journal of Chemical Theory and Computation

Biomolecular Systems

Ensemble-Based Steered Molecular Dynamics Predicts Relative Residence Time of A2A Receptor Binders

Andrew Potterton, Fouad S. Husseini, Michelle WY Southey, Michael J. Bodkin, Alexander Heifetz, Peter Vivian Coveney, and Andrea Townsend-Nicholson

J. Chem. Theory Comput., **Just Accepted Manuscript** • DOI: 10.1021/acs.jctc.8b01270 • Publication Date (Web): 20 Mar 2019

ACS Publications

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Ensemble-Based Steered Molecular Dynamics Predicts Relative Residence Time of A_{2A} Receptor Binders

Andrew Potterton[1], Fouad S. Husseini[2], Michelle W.Y. Southey[3], Mike J. Bodkin[3], Alexander Heifetz[1, 3], Peter V. Coveney[2,4] and Andrea Townsend-Nicholson[1,]*

AUTHOR ADDRESS

[1] Institute of Structural & Molecular Biology, Research Department of Structural & Molecular Biology, Division of Biosciences, University College London, London, WC1E 6BT, United Kingdom

[2] Centre for Computational Science, Department of Chemistry, University College London, London, WC1H 0AJ, United Kingdom

[3] Evotec (UK) Ltd., 114 Innovation Drive, Milton Park, Abingdon, Oxfordshire, OX14 4RZ, United Kingdom

[4] Computational Science Laboratory, Institute for Informatics, Faculty of Science, University of Amsterdam, The Netherlands

* a.townsend-nicholson@ucl.ac.uk

KEYWORDS Computational biology, computational chemistry, ensemble-based molecular dynamics, steered molecular dynamics, SMD, G protein-coupled receptor, GPCR, adenosine receptors, A_{2A} receptor, GPCR ligand binding, residence time

ABSTRACT

Drug-target residence time, the length of time for which a small molecule stays bound to its receptor target, has increasingly become a key property for optimization in drug discovery programs. However, its *in silico* prediction has proven difficult. Here we describe a method, using atomistic ensemble-based steered molecular dynamics (SMD), to observe the dissociation of ligands from their target G protein-coupled receptor in a timescale suitable for drug discovery. These dissociation simulations accurately, precisely, and reproducibly identify ligand-residue interactions and quantify the change in ligand energy values for both protein and water. The method has been applied to 17 ligands of the A_{2A} adenosine receptor, all with published experimental kinetic binding data. The residues that interact with the ligand as it dissociates are known experimentally to have an effect on binding affinities and residence times. There is a good correlation ($R^2 = 0.79$) between the computationally calculated change in water-ligand interaction energy and experimentally determined residence time. Our results indicate that ensemble-based SMD is a rapid, novel and accurate empirical method for the determination of drug-target relative residence time.

INTRODUCTION

G protein-coupled receptors (GPCRs), the largest membrane protein family present in humans, are also an important therapeutic target, giving rise to 34% of approved pharmaceutical compounds and worldwide sales of $890 billion from 2011-2016[1]. Despite this success, the rate of new GPCR pharmaceuticals making it to market has remained constant since the 1990s[2] and GPCRs remain a significantly under exploited drug discovery target, as those compounds that have made it to market target only ~10% of all GPCRs[3]. The A_{2A} receptor is a prototypical class

A GPCR. This class, also known as the rhodopsin-like receptors, account for 85% of all GPCRs[4]. Recently, the view that binding affinity is the definitive parameter for the clinical success of a drug candidate has been complemented by evidence from GPCRs and other receptor systems showing a higher correlation between efficacy and the length of time for which the drug stays bound to its receptor target, residence time[5–7]. The 'ideal' residence time may vary from target to target. Longer residence times permit less frequent dosing. For example when targeting M3 muscarinic acetylcholine, a GPCR, Tiotropium can be dosed less frequently than Ipratropium, as the former has a 50-fold longer residence time.[8,9]. Conversely, shortening residence time may also have benefits by, for example, decreasing off-target toxicity[10]. These considerations suggest that residence time may be more important than affinity, therapeutically, and that the optimization of residence time in addition to binding affinity in the early phases of drug discovery is critical to ensure that more drugs make it to market with fewer drugs failing in clinical trials.

Very little is known about the properties that affect residence time. The presence of buried protein-ligand hydrophilic interactions have been noted to increase residence time[11]. With compensation for the breaking of this bond provided, for example, through hydration, this state is less energetically unfavorable than it otherwise would be. Larger ligands are more likely to yield buried hydrophilic interactions, hence the observed correlation between molecular weight and residence time[10]. The kinetic rates of ligand binding, and thus residence time, have been proven to be adjustable by destabilisation of binding transition-states[12]. If understanding ligand-receptor residence time is critical to support the development of therapies with better efficacy *in vivo*[13],

then accurate and reproducible methods to allow for the rationalization of drug-target residence time need to be developed.

There are two categories of methods that are used to experimentally determine residence times of GPCRs: (i) methods that require a labeled ligand and (ii) label-free methods. Label-free methods, such as surface plasmon resonance (SPR), require purification and immobilization of the GPCR, which is not straightforward. However, SPR has been successfully used to calculate residence time using thermostabilized receptors[14]. Methods that require a labeled ligand account for the majority of kinetic studies performed on GPCRs[15]. Indirect kinetic radioligand binding assays require only one labeled ligand, and new bead-based methods remove the need for previously required filtration steps[16,17]. However, these improvements do not overcome the major limitation of these methods, which is the requirement of a suitable labeled ligand. Radiolabeled ligands are not readily available for all receptors and this is especially true of orphan GPCRs. Experimental methods of residence time determination are limited in their throughput[15] making rapid, accurate and reproducible computational methods the most practicable way to assess residence time for existing compounds and the only way to determine residence time for virtual candidate molecules as part of the drug discovery pathway.

Atomistic molecular dynamics (MD) has been used successfully to calculate macromolecular properties such as binding free energies for a number of different biological systems[18–21]. Single molecular dynamics (MD) simulations behave as random Gaussian processes[22], meaning that obtaining an accurate property of the system is impossible from a single run hence ensemble-averaging should be used for all MD simulations performed in both

academia and industry. In recent times, there has been great success using multiple, relatively short simulations (ensemble averaging) with binding affinity calculators, to obtain convergence of results and meaningful errors[20,21,23,24]. To compute any macroscopic property, be it binding affinity or rate parameters, from a microscopic (molecular) description of matter, statistical mechanics decrees that this must be done by use of ensemble averaging. With the use of ensembles, computational calculations are reproducible therefore two independently run ensembles should produce identical results within the reported error[22].

Use of classical molecular dynamics for the simulation of ligand dissociation is computationally very demanding and cannot be realistically used in drug discovery. Even Anton, the HPC (high-performance computing) cluster designed specifically for MD simulations, is only able to calculate these in the millisecond timescale for a 75,000 atom system[25], which is the size of a GPCR-membrane-water model, and Anton is only able to perform one simulation at a time, making ensemble-based analyses in this manner currently infeasible. In order to viably study ligand dissociation and its related parameter, residence time, one must engage methods to accelerate or enhance sampling. A metadynamics method, which describes the system by collective variables, has previously been used to rank the residence times of ligands, categorizing them in discrete classes (short, medium or long residence times)[26]. Another metadynamics method has attempted to predict absolute residence time values for 4 different recepetors[27]. Other methods have tried to predict relative residence times for a ligand series. A scaled molecular dynamics method, which also enhances sampling by modifying the potential energy surface and employing a reweighting scheme solely based on the populations of different states, has been used to predict the relative residence time of 12 ligands (one ligand series comprising 4 ligands

for each of three different protein systems)[28]. Random-accelerated MD (RAMD) accelerates the rate of ligand dissociation from the receptor. In RAMD, a direction of pulling is chosen randomly then, if a defined level of progress is not made, a new random direction is chosen. Using this method, the time taken for dissociation to occur was compared to experimentally-determined residence time for ligands of HSP-90[29] (heat shock protein 90), obtaining good correlation within ligand series but not across more varied datasets. The time taken to perform RAMD calculations, by definition, is not constant; indeed there is a 40-fold range in simulation times of dissociation[30].

Equilibrium MD simulations are typically used to compute absolute binding free energies; however, non-equilibrium MD can be used to calculate absolute free energies, an equilibrium property, through the use of Jarzynski's inequality[30]. In RAMD, accelerated ligand dissociation is used to predict relative residence time, a parameter related to, but not correlated with, binding free energy. Steered molecular dynamics (SMD), another non-equilibrium MD method, also accelerates ligand dissociation. The time taken for ligand dissociation to occur is constant when using constant velocity SMD, making it impossible to correlate computationally-accelerated dissociation against experimentally-determined residence time. Despite this, constant velocity SMD can be used to identify the changing forces from the bound to the partially dissociated state and this, in turn, can be used to predict relative residence time. The computational cost of this method is therefore constant for different ligands of the same receptor, which is a significant advantage over RAMD[31–34]. In the present paper, we present the development of a robust, ensemble-based SMD method for predicting relative drug-target

residence time by a novel means, namely identifying the molecular interactions that take place during dissociation of a ligand from its receptor.

METHODS

Here, we aim to develop a reproducible and accurate MD protocol that can be used to observe the dissociation of ligands from GPCRs. This will be accomplished by using steered molecular dynamics to forcibly accelerate the dissociation of the ligand from its receptor, observing the receptor residues that are involved during exit. To generate an ensemble average and to ensure that calculations are reproducible[30], the number of replicas required to make up the ensemble, needs to be established. Determination of the appropriate ensemble size, a precondition for accumulating results, is described at the end of the methods section.

Amino acid residues will be defined as initial contacts if any atom of the residue is within 3.5 Å of any atom of the ligand during the first five frames (0.2 ns) of the SMD simulation (after the 2 ns equilibration) in greater or equal to 50% of the replicas in an ensemble. Contacts that form as dissociation progresses and the ligand exits the receptor will be defined as intermediate contacts, with the same distance cut-off as initial contacts but taking place at a distance that is greater than 3.5 Å away from the initial binding pose. Additionally, intermediate contacts must remain within the distance requirement for a minimum of 10 consecutive frames (0.4 ns). Residues that form a hydrogen bond or are involved in π-stacking (as defined by visual inspection of the simulation trajectories) with the ligand will be defined as interactions, regardless of the distance between the two. Residues will be described by their amino acid identity (single letter code) and position (amino acid number) within the specific GPCR with the

Ballesteros and Weinstein numbering[22], a scheme for class A GPCRs, whereby X.50 represents the defined centrally conserved residue on helix X, in superscript, as published previously[35]. In addition to identifying contacts and interactions, the SMD simulations will be used to calculate dissociation energies of the ligands with water. The differences in dissociation energies, calculated between the ligand in its initial receptor-bound location and the extracellular vestibule, will be quantified and compared with experimentally-determined residence time.

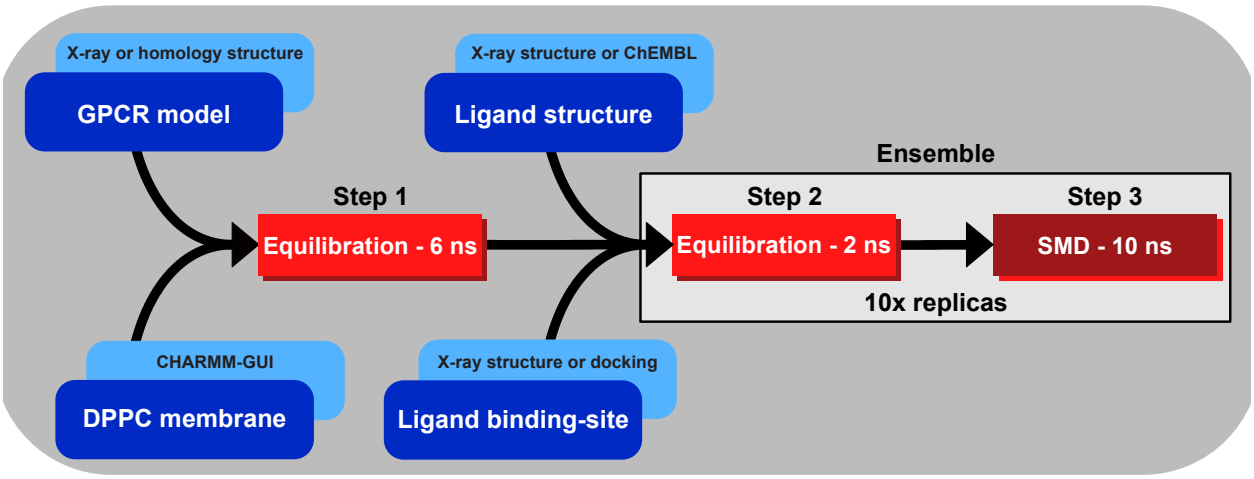


Figure 1. Graphical overview of the SMD protocol, including the two equilibration steps. The replica runs of the ensemble are shown in the light grey box. Timescales and programs/web services used in the protocol are indicated.

There are three stages to the SMD protocol we have developed (see Figure 1). First, the GPCR is placed in a DPPC membrane in the absence of ligand and equilibration is performed for 6 ns, the time it takes for the previously-defined membrane density and membrane thickness to be achieved[24]. Second, each ligand is docked into this structure and equilibrated for a further 2 ns, the time that is needed for the binding pocket of the receptor to accommodate the ligand[36]. Third, SMD is performed for a duration of 10 ns on this equilibrated, ligand-bound, protein-

membrane system. Computational cost is reduced by starting from the protein-membrane equilibration output and repeating steps 2 and 3, in order to produce ten replicas in an ensemble. With this protocol, each replica completed within 10 hours using 256-cores on Grace, an HPC cluster at University College London (UCL) (technical specifications of this cluster can be found at https://wiki.rc.ucl.ac.uk/wiki/RC_Systems#Grace_technical_specs). Although it was deemed to be unnecessary for the development of the protocol, automation would be needed for the robust implementation of this protocol in academia or industry. To increase the scale of compounds analyzed, this protocol could be readily implemented as an extension to the highly scalable BAC software, which offers a high-throughput environment for simultaneous binding free energy determinations for thousands of compounds[19].

Creation of A_{2A} receptor systems

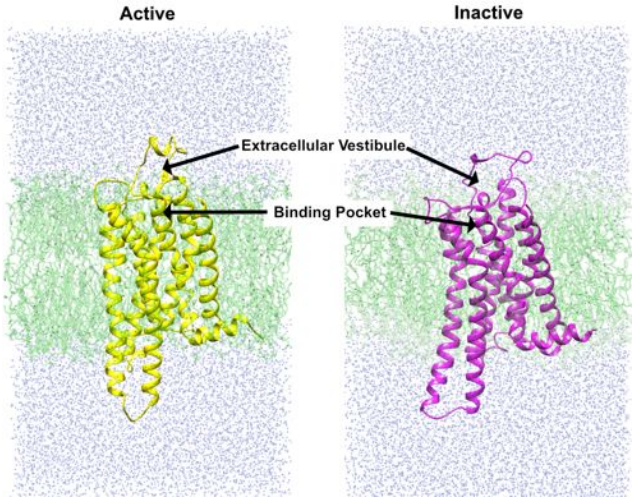


Figure 2. Diagrams of the active and inactive A_{2A} receptor GPCR models. The DPPC membrane is shown in green and water in blue. The location of the extracellular vestibule and binding pocket is indicated.

Models were created for the A_{2A} receptor that were based on the highest resolution X-ray structures available that contained activity-defining thermostabilizing mutations shown to bias the receptor to a particular physiological state[20,21,37]. The active and inactive state models (see Figure 2) were built from PDB accession numbers 4UHR[38] and 5IU4[39], respectively. The cytochrome b562-RIL (bRIL) fusion domain was removed from the 5IU4 model. Using the wild type human A_{2A} sequence accessed from GPCRdb (gpcrdb.org)[40], missing loops from all models were reconstructed with MODELLER 9.12[41] and the thermostabilizing mutations in the model were removed using UCSF's Chimera 1.11.2[42]. The disulfide bonds present in the crystal structures (4 in both the 'active' form and the 'inactive' form) were restored by manually adding CONECT lines to the PDB snapshots using a text editor. These models were then inserted into a 100% 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC) membrane, using CHARMM-GUI[43] with approximately 150 of these membrane molecules needed to make a simulation box of a size sufficient to surround the receptor (75 x 75 Å). Using CHARMM-GUI, approximately16,000 water molecules were added and electroneutrality was achieved by using Na^+ or Cl^- as counterions to balance the net charge of the system to zero. As none of the ligands under study were charged, neutralization was carried out prior to ligand addition. Ligand binding poses from A_{2A} crystal structures were superimposed onto the appropriate equilibrated protein-membrane system, using Chimera's MatchMaker; the coordinates of the binding position were added to the PDB snapshot of the equilibrated model. The structures from which the ligand binding pose was superimposed from are listed in Table 1. If a co-crystallized structure was unavailable for a given ligand, docking was performed using AutoDock 4.2[44] embedded in Chimera. The structures of the ligands used in this study are shown in Figure 3.

Table 1. PDB accession numbers of the structures used for obtaining ligand binding poses.

Ligand	PDB
CGS-21680	4UHR
NECA	2YDV
Theophylline	5MZJ
UK-432,097	3QAK
XAC	3REY
ZM-241,385	5IU4

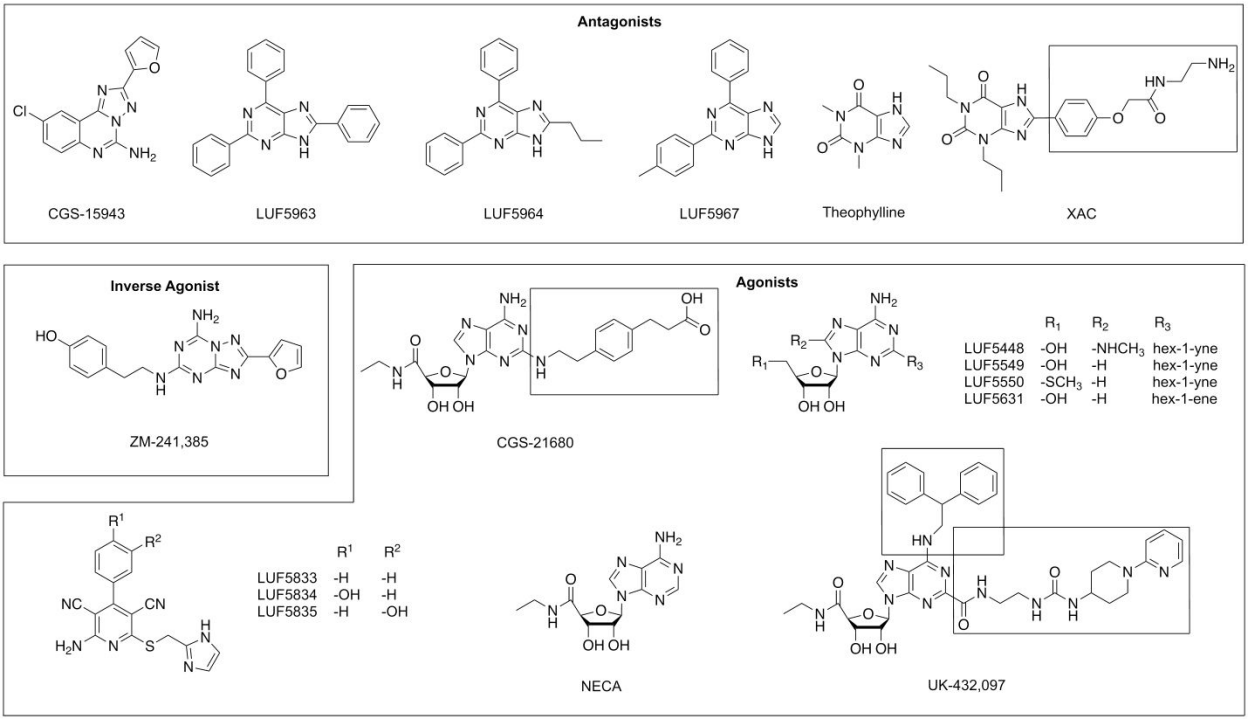


Figure 3. Structures of the 17 A_{2A} receptor ligands used in this study. The 'aromatic ring extensions', referred to in results section, are highlighted by boxes on the 2D structures.

Atomistic MD simulations

Atomistic MD simulations were carried out using NAMD 2.11[45] as the MD engine. To parameterize the system, AMBER 14 forcefields were used, specifically, AMBER GAFF[46], protein14.SB[47], lipid 14[48] and the TIP3P model for the ligand, protein, membrane and water molecules, respectively. Ligand parameterization was carried out using the Antechamber

program[49] within the AM1-BCC charge model[50]. The initial velocity of the atoms were drawn randomly from a Maxwell−Boltzmann distribution at 310 K (the human physiological temperature); replica runs were otherwise identical, apart from these initial velocity seeds. An ensemble of 30 replicas was performed for three ligands, NECA (an agonist), theophylline (an antagonist) and UK-432,097 (the largest ligand in the test set), to determine the optimal number of replicas in an ensemble. For subsequent ligands, 10 replicas were performed on each protein-ligand system. Step 1 of the protocol was conducted in the constant volume, constant temperature (NVT) ensemble. With the exception of this initial equilibration step, all other simulations were conducted with a barostat in the constant pressure, constant temperature (NPT) ensemble. The temperature was controlled at 310 K using Langevin dynamics whilst pressure was kept constant at 1 bar by the Nose-Hoover-Langevin algorithm. The Langevin piston period used was 100 fs with a piston decay of 50 fs. An integration timestep of 2 fs was used for all simulations and the SHAKE constraint algorithm applied to all hydrogen bonds to achieve this. For the calculation of long-range electrostatics interactions, the particle-mesh Ewald (PME) method[51] was used. A cut-off of 10 Å was used to calculate van der Waals interactions. During the 6 ns equilibration of the protein-membrane system (Step 1), decreasing constraints were applied to the backbone of the membrane and the protein. Following this, a PDB formatted snapshot was taken from the last few frames into which the ligands were docked. Equilibration of this protein-ligand-membrane system for 2 ns (step 2) involved an initial constraint on all of the elements in the system except water, followed by a decrease to zero of this constraint. After equilibration, the SMD production run (step 3) was carried out for 10 ns. This time was sufficient for all ligands to be pulled from their binding pockets to the extracellular vestibule. The SMD protocol used a variable force applied to the ligand to keep it moving at a constant velocity of 1

Å/ns (0.000002 Å/timestep); the output of the force used was logged every 1000 steps. The Newtonian spring was attached to the center of the mass of the ligand with a spring constant of 3 kcal/mol/Å^2 (208.4 pN Å). The direction of pulling was held constant for all simulations. Constraints on the protein were found to be necessary to prevent drifting in the direction of pulling. In both active and inactive systems, these restraints were applied in three dimensions to the alpha carbon of five residues at the top of the transmembrane helices (residues $9^{1.35}$, $80^{3.28}$, $177^{5.38}$, $256^{6.58}$ and $270^{7.35}$).

Data analysis

Visualization was performed using VMD 1.9.3[52] and Chimera[53]. Tcl scripts for VMD were generated to produce contact with interaction residue data. Dissociation energies were calculated using NAMD energy 1.4 called by an in-house Tcl script. Automation of the workflow was achieved using bash scripts. Data were analyzed using Python scripts. VMD was used to identify hydrogen bonds, with 3.5 Å and 25° as the hydrogen bond cut-off. In addition, to be classified as a hydrogen bond, the bond must have been seen in >50% of the replicas in the ensemble.

To convert kinetic ligand binding values to equilibrium K_d values, the following equation was used:

$$K_d = \frac{k_{off}}{k_{on}}$$

Along with that, relating the free energy of binding to the equilibrium K_d was done using the following equation:

$$\Delta G = RTlnK_d$$

To determine the dissociation energy (van der Waals and electrostatic interactions) of ligand with water (ΔE_{LW}) for each replica, the following equation was used:

$$\Delta E_{LW} = E_{LW}(t = 10ns) - E_{LW}(t = 0ns)$$

A switching and cut-off distance of 10 Å and 12 Å, respectively, was used to calculate this non-bonded energy. The mean value of ΔE_{LW} for each ensemble was calculated, resulting in an average ΔE_{LW} for each ligand with an associated error.

Potential Mean Force calculations

The force (***F***) applied in the x, y and z directions, and the normalized direction of pulling, referred to as $\mathbf{n_x}$, $\mathbf{n_y}$ and $\mathbf{n_z}$ respectively, was logged every 1000 frames (2 ps). To convert these values to calculate the overall force applied, the following equation was used:

$$\boldsymbol{F}(t) = \boldsymbol{n}_x F_x + \boldsymbol{n}_y F_y + \boldsymbol{n}_z F_z$$

Irreversible work done to the system, as a function of time, was calculated by integrating the force curve and multiplying this by the constant velocity of pulling, ***v***:

$$W(t) = \int_0^t \boldsymbol{v} \cdot \boldsymbol{F}(t) dt$$

Jarzynski's identity[43] was employed to calculate the potential of mean force (PMF) using the following equation (where k_B is Boltzmann's constant (kcal/mol/K), T is temperature (Kelvin) of the bulk system, $\beta = \frac{1}{k_B T}$ and $\sigma_w^2 = \langle W^2 \rangle - \langle W \rangle^2$:

$$PMF(t) = \langle W \rangle - \frac{1}{2}\beta\sigma_w^2$$

This use of Jarzynski's identity minimizes the impact of pulling speed, enabling a better comparison between SMD simulations performed using different parameters[31].

Determination of optimal simulation length and optimal number of runs

The simulation length of the SMD protocol needed to be sufficient for the ligand to be pulled from the binding site to the extracellular vestibule. Initial pulling velocities ranged from 1-300 Å/ns, which led to ligand dissociation on a timescale ranging from 40 ps to 10 ns. The system was very sensitive to the rate of pulling and more rapid rates required the fixing of every atom in the receptor. 1 Å/ns was chosen as a pulling velocity, as it did not require the immobilization of the entire receptor and because this value is on the same order as those previously used in SMD simulations[54]. Pulling ceased after the ligand reached the extracellular vestibule to prevent potentially biased ligand interactions with the membrane environment. Whilst the simulation contains a biological lipid, it does not include known non-lipid components (e.g. cholesterol and other proteins) found in the physiological membrane environments of GPCRs. After the extracellular vestibule has been reached, the pulling direction is no longer limited to the single directional vector that had been imposed by the constraints of the binding pocket and the simulation was stopped to prevent bias.

To determine the number of replicas needed in an ensemble to achieve convergence of results, the SMD protocol was initially performed for an agonist (NECA), for an antagonist (theophylline) and for the largest ligand in the dataset (UK-432,097), for ensembles of 30 replicas each, the number of replicas performed in a previous GPCR MD study[55–57]. NECA and theophylline were selected as representative ligands because 12 of the remaining 15 ligands in the test set have a similar base structure. The convergence of the mean and error as a function of time for these three ligands were explored as values by which to determine the optimum replica number. The mean energy difference between ligand and water, ΔE_{LW}, is the only single continuous output, since residue energy values change over time as the ligand is pulled from the binding site.

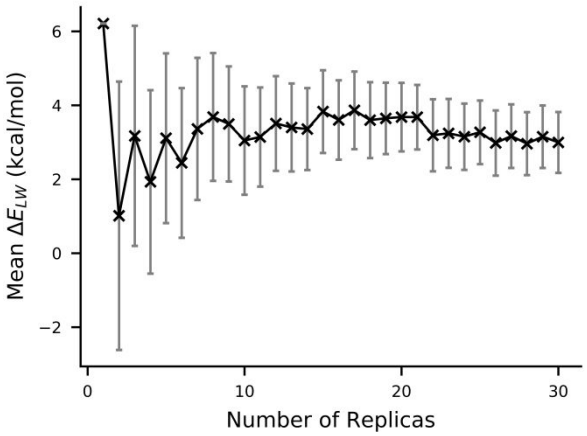


Figure 4. The mean of ΔE_{LW}, the calculated energy difference between the ligand and water, before and after SMD, is shown as a function of number of replicas performed for the antagonist theophylline. Standard deviation, as a bootstrap statistic of the associated number of runs, was used to calculate error.

As shown in Figure 4, increasing the number of replicas in the ensemble results in a statistically significant change in the mean ΔE_{LW} value. This means that performing one-off simulations could result in significantly different ΔE_{LW} values and thus different residence time

prediction values. To examine the effect of ensemble size on error for both agonists and antagonists, the analysis was repeated using the method of bootstrapping[24]. This statistical technique uses resampling with replacement (10,000 times) of the data points (the ΔE_{LW} value of each replica). The mean of each bootstrap is calculated and the standard deviation (σ_{boot}) of these bootstrap averages is calculated. The σ_{boot} value provides an estimate of the error associated with a mean derived from a given sample. Figure 5 shows a sharp decrease in the error of ΔE_{LW} with increasing number of replicas performed, for all three ligands, which steadies after 10 runs. For theophylline (see Figure 5c), the minimum of the decreasing error occurs around 20 runs. However, a decrease of less than 0.6 kcal/mol in the error of ΔE_{LW} requires twice the number of replicas in the ensemble, doubling the computational cost. Ten was determined to be the optimal number of replicas, permitting an effective trade-off between minimization of error, stabilization of the mean and computational cost.

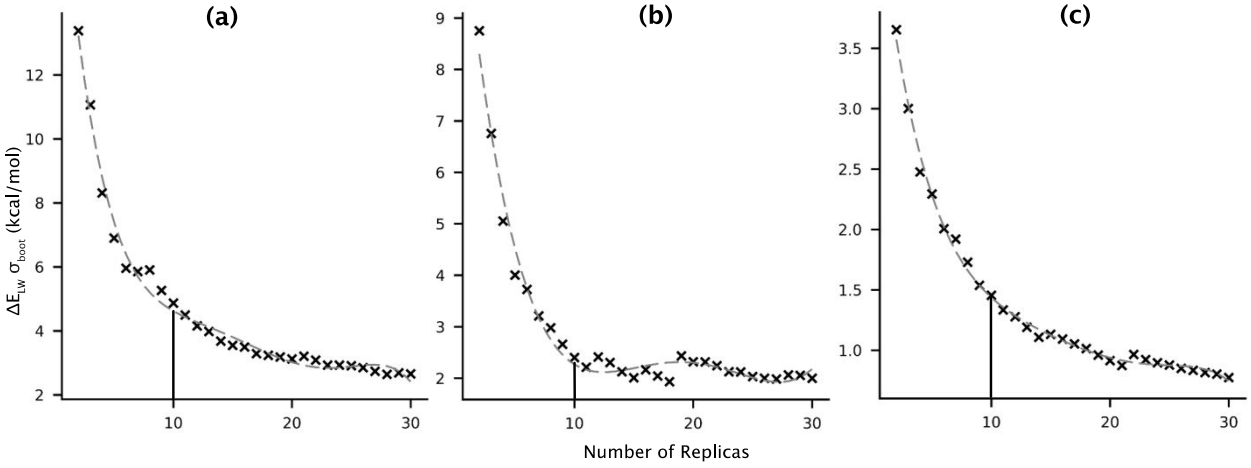


Figure 5. Variation of the bootstrap statistics, σ_{boot}, in ΔE_{LW} is shown as a function of number of replicas performed for the representative ligands UK-432,097 (a) and NECA (b) and Theophylline (c). The optimal number of replicas used in subsequent simulations is shown as a black solid line in all three graphs.

RESULTS

Computational validation

The SMD protocol we have developed assumes that agonists and inverse agonists dissociate from active and inactive states, respectively. This assumption is encoded in the protocol by the use of two distinct starting structures that were based on crystal structures engineered to be trapped either in an active or an inactive state. We tested the validity of the assumption by pulling the inverse agonist ZM-241,385 (see Figure 3) from two starting structures: the inactive structure, for which numerous co-crystallized structures exist, and from the active structure, for which co-crystallized structures have not yet been observed. The contact and potential of mean force (PMF) data for both ensembles were compared with each other and with published experimental SDM data.

The residue contacts made with ZM-214,385 in the active and inactive state ensembles were different. In one specific instance, a receptor residue identified as a contact in the inactive state ensemble was not detected in the active state ensemble. This residue, $K153^{ECL2}$, has been identified by temperature-accelerated MD as a dissociation contact of ZM-241,385 and experimental results show that mutation of this residue leads to a 26% increase in residence time compared with wild type[58]. Three of the contacts identified from the pulling of ZM-241,385 exclusively in the inactive ensemble are confirmed by published data from SDM experiments[59]. None of the contacts identified exclusively in the active ensemble can be confirmed by published experimental data. These findings lead us to conclude that the activity state of the receptor's

starting structure needs to be matched to the properties of the ligand to maximize the detection of biologically-relevant contacts.

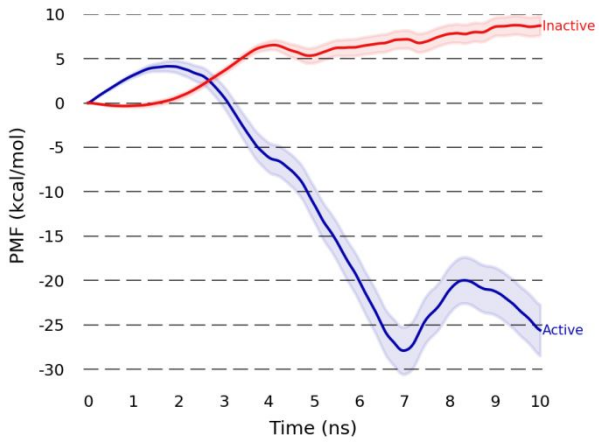


Figure 6. Potential of mean force (PMF) profiles of the inverse agonist, ZM-241,385, being pulled from the active (blue) and inactive (red) states of the A_{2A} receptor. Standard deviation, as a bootstrap statistic, was used to determine error of the 10 replicas, and is shown in the shaded regions.

There is a significant difference in the PMF profile when pulling ZM-241,385 from the active and inactive states of the A_{2A} receptor (see Figure 6). Pulling the ligand from the inactive state causes 9 ± 1 kcal/mol of irreversible work to be done on the system, whereas pulling from the active state results in a negative PMF. A negative PMF indicates that it is less energetically favorable for the ligand to be bound to the receptor than for it to be unbound. It also shows that ZM-241,385 energetically favors binding in the inactive state to binding in the active state. The contact data and the PMF profile results highlight the critical need to use the appropriate receptor state when carrying out this protocol.

Identification of residues involved in ligand dissociation

The dissociation pathways of agonists and antagonists were found to be similar to one another. This is most likely due both a narrow binding pocket and the use of the same pulling direction in all the SMD simulations. There are unique patterns of residues that interact with each of the ligands during dissociation. The residues involved in ligand dissociation can be divided into three categories, as described in the Methods section: (i) initial contacts, (ii) intermediate contacts, and (iii) sustained contacts, Figure 7 shows examples of these types of residues. Each type will be discussed in the context of published experimental data.

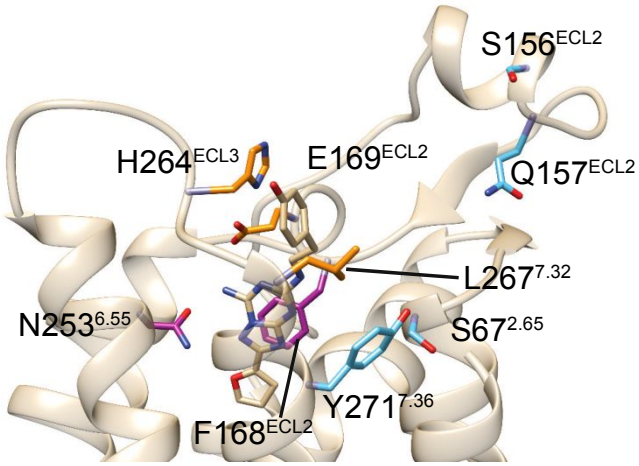


Figure 7. Diagram showing the dominant initial, intermediate and sustained residue contacts of ZM-241,385 colored in purple, light blue and orange, respectively. ZM-241,385 in its initial binding position is shown in beige.

Initial contacts and interactions

Ensembles were performed for 17 A_{2A} receptor ligands with the 10 agonists pulled from the active structure and the 6 antagonists and one inverse agonist pulled from the inactive structure (see Figure 3 for ligand structures). A summary of the interaction energy of the initial residue contacts to the 17 ligands tested is shown in Table 2 (the inverse agonist is included in

the 'antagonist' grouping). The initial residue contacts of the ligands at t=0 in the simulation vary between agonists and antagonists but there are three conserved residue contacts for all 17 ligands: F168^{ECL2}, N253$^{6.55}$ and M270$^{7.35}$. In addition, two of these three residues form specific interactions with many of the ligands tested. F168^{ECL2} forms π-stacking interactions in the orthosteric binding site to all of the ligands, specifically to the methylxanthine or adenine ring of antagonists and agonists, respectively. In particular, it forms π-stacking interactions with a 6-membered ring group for the three non-adenine-based agonists. As a validation of our computational method, the importance of F168^{ECL2} has been confirmed experimentally, as there is no detectable radioligand binding when this phenylalanine residue is mutated to alanine[59]. N253$^{6.55}$ forms a hydrogen bond with all but one of the agonists (LUF5834) and all but one (LUF5967) of the antagonists. Despite similarities in the adenine/methylxanthine moiety of fourteen of the ligands chosen for this study, there appears to be little conservation in the initial residue contacts and the interactions they make with ligands.

Table 2. Initial energies (kcal/mol) of the residue contacts of 17 A_{2A} receptor ligands[1].

Residues	A_{2A} receptor binders																
	CGS-15943	LUF5963	LUF5964	LUF5967	Theophylline	XAC	ZM-241385	CGS-21680	LUF5448	LUF5549	LUF5550	LUF5631	LUF5833	LUF5834	LUF5835	NECA	UK-432,097
A63$^{2.61}$				-0.7		-2.5			0.4	0.3	-0.1		0.3	0.0	0.1		
I66$^{2.64}$				-2.0	-1.0	-2.2				-1.5							
S67$^{2.65}$	-2.1		-1.1	-3.0		**-3.8**		-1.4	-2.1	-1.6	-2.4	-1.7	-1.4	-1.0	-1.0		-2.3
T68$^{2.66}$						-1.3											
A81$^{3.29}$						-0.7											
V84$^{3.32}$	-0.1	0.5	-0.6			-0.8	-0.3	-3.7	-1.7	-1.4	-1.2	-1.1	-2.2	-2.1	-2.2	-3.4	-3.9
L85$^{3.33}$		-0.4	-0.8				-0.8	-4.6		-1.3	-1.4	-1.2	-2.3	-2.4	-2.6	-3.9	-4.2
T88$^{3.36}$								**-5.4**			-0.9		-1.8	-1.6	**-1.9**	**-3.6**	**-5.3**

[1] Residue-ligand pairs that form hydrogen bonds are shown in bold. Error as bootstrap standard deviation (σ_{boot}) is not shown for residue-ligand pairs but is < 1 kcal/mol with the following exceptions: N253$^{6.55}$ – CGS-21680 (1.1 kcal/mol), E169^{ECL2} – LUF5834 (1.8 kcal/mol), E169^{ECL2} – UK-432,097 (2.6 kcal/mol) and N253$^{6.55}$ (1.5 kcal/mol).

Q89$^{3.37}$								-0.9									-1.0
I92$^{3.40}$								0.1								0.1	0.2
E151^{ECL2}																	-0.8
G152^{ECL2}																	-0.1
L167^{ECL2}		-1.9		-3.2		-1.6	-1.2	-2.5	-2.6	-2.5	-2.2	-2.0	-0.9	-0.6			-2.3
F168^{ECL2}	-7.1	-9.9	-6.9	-6.7	-8.3	-8.7	-7.6	-8.1	-8.2	-9.4	-8.0	-8.4	-9.0	-8.2	-8.2	-7.6	-9.0
E169^{ECL2}	**-12**	0.6	-2.0	**-15**	0.1		**-14**	-4.1	-0.7	**-1.0**	**-3.5**	**-3.3**	-3.8	**-15**	-8.7	-1.4	**-23**
M174$^{5.35}$	-0.4	-1.6	-1.7	0.2	-0.5	-1.7	0.0	0.3	-0.6	-0.4	-0.8	-0.4				0.4	-0.3
M177$^{5.38}$	-0.7	-2.4	-1.5		-1.1	-1.7	-1.5	-3.6	-1.6	-2.0	-2.6	-2.3	-3.3	-3.4	-3.8	-3.5	-3.1
N181$^{5.42}$								-1.5					-1.3	-1.6	-1.7	-1.7	-1.6
C185$^{5.46}$								-0.5								-0.3	-0.4
W246$^{6.48}$								-3.1	-1.6	-1.4	-1.1	-1.3	-1.3	-2.1	-1.3	-2.5	-3.2
L249$^{6.51}$	-1.7	-1.6	-1.9		-1.6	-1.7	-1.8	-3.2	-2.8	-2.6	-1.9	**-2.5**	-1.8	-4.1	-1.8	-2.9	-3.8
H250$^{6.52}$		-1.4				-0.7		**-5.8**	-2.0		-1.7		-4.2	-0.6	-4.4	**-5.5**	**-4.6**
I252$^{6.54}$											-1.3						-2.5
N253$^{6.55}$	**-9.8**	**-0.8**	**-1.9**	-4.3	**-3.5**	**0.0**	**-13**	**-6.1**	**-8.0**	**-5.6**	**-5.6**	**-6.2**	**-3.0**	-3.0	**-4.0**	**-5.0**	**-7.8**
T256$^{6.58}$			-0.8														-1.8
S263^{ECL3}			-0.7														
H264^{ECL3}		-3.8	-2.7				-4.2						-0.7	-1.0	**-3.9**		-3.1
A265^{ECL3}							-2.1										
L267$^{7.32}$	-0.4	-0.8	-1.2	-1.3		-2.7	-2.4	-2.3	-0.8	-0.8	-0.9	-0.9	-0.2				-3.0
M270$^{7.35}$	-4.9	-4.3	-4.2	-4.8	-3.0	-4.2	4.2	-4.0	-3.0	-3.5	-3.0	-2.9	-3.4	-3.3	-4.1	-2.3	-7.0
Y271$^{7.36}$		-0.6	-1.0			**-2.1**		-2.5	-1.9	-1.8	-1.7	-1.7					**-7.5**
I274$^{7.39}$	-2.9	-2.3	-2.0	-2.3	-2.6	-1.5		-4.4	-5.9	-6.1	-3.9	-5.5	-3.6	-3.6	-3.6	-4.2	-5.7
S277$^{7.42}$									-2.0	**-3.2**		**-2.9**					
H278$^{7.43}$								**-4.9**	-5.8	**-5.5**	-2.1	**-5.0**				**-5.7**	**-4.9**

There are a number of residues that make contact with agonists but not with antagonists (or the inverse agonist). The following nine residues make contact with a minimum of two agonists: T88$^{3.36}$, Q89$^{3.37}$, I92$^{3.40}$, N181$^{5.42}$, C185$^{5.46}$, V186$^{5.47}$, W246$^{6.48}$, S277^{7,43}, H278$^{7.43}$. A previous analysis of A_{2A} crystal structures has classified four of these residues (T88$^{3.36}$, Q89$^{3.37}$, I92$^{3.40}$, N181$^{5.42}$) as being agonist-only contacts[60]. All nine residues are located below the adenine ring in the binding pocket and are proximal to the ribose moiety, which is not present in antagonists. NECA makes hydrogen bonds with two of these residues: T88$^{3.36}$ and H278$^{7.43}$. Experimental data show that mutagenesis of either of these two residues caused a reduction in binding affinity to agonists including NECA, but not to antagonists[61]. Our protocol confirms that there are five residues involved in the binding site of agonists only and identifies two new initial contacts for future experimental validation.

Intermediate contacts

There is a diverse spread of and very little conservation in the intermediate contacts made for each of the 17 ligands tested (see Table 3). This, to a large extent, is dependent on the initial contacts made by a ligand, as initial contacts and intermediate contacts are non-overlapping groups therefore an intermediate contact cannot be an initial contact. Two residues stand out as intermediate contacts: S6$^{1.32}$ and H264^{ECL3}. H264^{ECL3} is an intermediate contact of 13 ligands; for the remaining 4 ligands, it is an initial contact. S6$^{1.32}$, is universally conserved as an intermediate contact for all but one of the agonists but none of the antagonists. In summary, there is no single residue that forms an intermediate contact with all of the ligands tested, nor is there a single residue that forms an intermediate contact with all of the antagonists or with the inverse agonist.

Table 3. Average energies (kcal/mol) of the intermediate residue contacts[2]

Residues	A_{2A} receptor binders																
	CGS-15943	LUF5963	LUF5964	LUF5967	Theophylline	XAC	ZM-241385	CGS-21680	LUF5448	LUF5549	LUF5550	LUF5631	LUF5833	LUF5834	LUF5835	NECA	UK-432,097
S6$^{1.32}$								-0.7	0.4	-1.3	0.8	0.4	-1.7	-2.4	-1.9	-2.7	
Y9$^{1.35}$		**-1.5**						-1.1	**-1.3**	-1.2	-0.7	**-1.1**	-1.7	-1.2	**-0.8**	**-2.1**	**-2.2**
I10$^{1.36}$										-0.2			-0.1				
E13$^{1.39}$										-6.1							
A63$^{2.61}$	0.1		-0.7				-0.1	-0.5				**-0.5**				-0.4	**-0.8**
I64$^{2.62}$										-0.4			-0.2				
I66$^{2.64}$	-1.2		-0.9				-0.1	-0.6	**-0.5**		**-0.8**	**-0.8**	**-0.9**	-1.2	-1.0	-0.7	0.1
S67$^{2.65}$					**-2.4**		**-0.8**									**-1.2**	
T68$^{2.66}$	-1.0	-0.4	-0.6	**-1.0**			-0.6	-0.7	-0.7	-0.6	-0.6	-0.6		-0.7			
G69$^{2.67}$	-0.6	0.0	-0.3	-0.7		-0.1	-0.3	-0.5	-0.7	-0.6	-0.3	-0.6	0.0		0.1		
A81$^{3.29}$														-0.7	-0.7		
Q89$^{3.37}$																-0.5	

[2] Residue-ligand pairs that form hydrogen bonds are shown in bold. Error as bootstrap standard deviation (σ_{boot}) is not shown for residue-ligand pairs but is < 1.2 kcal/mol with the following exceptions: K153^{ECL2} – CGS-21680 (1.3 kcal/mol), K153^{ECL2} – LUF5550 (1.4 kcal/mol), K153^{ECL2} – LUF5631 (2.4 kcal/mol) and K153^{ECL2} – LUF5834 (1.9 kcal/mol).

K150^{ECL2}											-1.5						
G152^{ECL2}				-0.6		-1.8	-0.4										
K153^{ECL2}	0.6	-4.5	-1.7	**-6.8**		**-8.0**	-1.6	**-4.3**			-0.1	-2.5		-1.0			**-1.6**
H155^{ECL2}						-1.7											
S156^{ECL2}	-0.3	-0.3	-0.7	-0.6		**-1.2**	**-0.8**										
Q157^{ECL2}	**-1.0**	-0.5	-2.0			**-0.9**	**-1.3**	-1.0		-0.4		-0.5					**-1.5**
G158^{ECL2}								-0.4									-0.8
A165^{ECL2}				-0.1			-0.3	0.0	0.1	-0.1		-0.1					
L167^{ECL2}	**-2.9**		-1.8		-1.2										-2.4	-3.0	
E169^{ECL2}						2.8											
D170^{ECL2}				-2.4						1.1	-0.6						
V172^{ECL2}			-0.1									-0.4			0.1		
M174$^{5.35}$													-0.7	-0.8	-0.6		
I252$^{6.54}$			-0.1			-0.5		-0.1	0.0	-0.9		-0.5	-0.1	-0.5	-1.1	-0.1	
F255$^{6.57}$															-0.1		
T256$^{6.58}$		-0.2			-0.1	-0.3	-0.1	-0.8	0.0	-0.3	-0.8	-0.5	-0.5	-0.9	**-1.6**	-0.7	
P260^{ECL3}											0.0	-0.3			-1.1		-0.5
D261^{ECL3}											-1.0						-1.3
S263^{ECL3}	-0.6			-0.6	-0.9	-0.2		-0.6	-0.8	-0.4	-0.9	**-0.9**	-0.7	-0.9	-0.6		-0.8
H264^{ECL3}	**-1.8**			**-2.1**	**-2.6**	**-6.1**		**-3.1**	**-2.6**	**-3.7**	**-3.0**	**-3.1**				**-1.6**	
A265^{ECL3}	**-0.9**	**-0.2**	**-1.6**	-0.8	-1.1	**-1.2**		-0.1	0.4		0.4	-0.2	-0.1	-0.4	-0.5		-0.7
P266$^{7.31}$			-0.3							-0.4	-0.3						
L267$^{7.32}$					-2.0									-2.2	-2.1	-2.1	
Y271$^{7.36}$	**-1.5**			**-1.8**	**-3.6**		**-1.7**						**-2.8**	**-2.5**	**-2.2**	**-2.4**	
I274$^{7.39}$							-1.1										
H278$^{7.43}$														-0.1	-0.3		

Comparison of ZM-241,385 intermediate contacts and SDM radioligand kinetic data

Residue distance was initially used as a cut-off to identify initial and intermediate contacts, with energy values subsequently used to evaluate non-bonded interactions. These data are represented, as a function of time, in Figure 8. There are 13 intermediate contacts formed as ZM-241,385 is pulled from the A_{2A} receptor: A63$^{2.61}$, I66$^{2.64}$, S67$^{2.65}$, T68$^{2.66}$, G69$^{2.67}$, G152^{ECL2}, K153^{ECL2}, S156^{ECL2}, Q157^{ECL2}, A165^{ECL2}, T256$^{6.58}$, Y271$^{7.36}$ and I274$^{7.39}$. Intermediate contacts, by definition, will be not identified in crystal structures where the ligand has been co-crystallized with the receptor. However, experimental data from mutation of these intermediate contacts may offer insight into binding affinity and residence time. Eight of the intermediate contacts identified, I66$^{2.64}$, S67$^{2.65}$, T68$^{2.66}$, K153^{ECL2}, S156^{ECL2}, Q157^{ECL2}, T256$^{6.58}$, Y271$^{7.36}$, have

experimental data[62,63] that can be compared with our computational findings, some of these residues are shown in Figure 7.

The following categories of effect are observed when these residues are experimentally mutated to alanine[59]: (i) increase in residence time with little or no effect on binding affinity (I66$^{2.64}$, S67$^{2.65}$ and K153^{ECL2}); (ii) decrease in binding affinity with little or no effect on residence time (S156^{ECL2} and Q157^{ECL2}); (iii) increase in binding affinity with a decrease in residence time (T256$^{6.58}$); (iv) no detectable binding observed (Y271$^{7.36}$); and, (v) no detectable effect on either binding affinity or residence time (T68$^{2.66}$).

I66$^{2.64}$, S67$^{2.65}$ and K153^{ECL2} all increase residence time when mutated to alanine. These three residues are located at the top of the binding pocket. Their mutation to alanine may make it less energetically favorable for dissociation to occur. In the simulations, S67$^{2.65}$ is shown to form a hydrogen bond that would be lost with experimental mutation to alanine. S156^{ECL2} and Q157^{ECL2} are each located close to the extracellular vestibule (see Figure 2 for the location of the extracellular vestibule in the A_{2A} receptor) and at least 9 Å from the binding pocket. These residues may be involved in ligand docking and it is surprising that they do not have an effect on residence time when mutated. Experimental mutation of T256$^{6.58}$ caused a 151% increase in binding affinity and a 94% decrease in residence time[59]. One crystal structure, PDB accession number 5IU4[59], revealed a water molecule that links T256$^{6.58}$ with other residues important in binding (N253$^{6.55}$ and E169^{ECL2}), which may explain the increase in binding affinity. It was found that experimental mutation of Y271$^{7.36}$ to alanine caused loss of binding for ZM-241,385[40] making it impossible to determine the effect of this residue on residence time. T68$^{2.66}$, which

falls into the final category of effect observed, does not provide a hydrogen bond in any of simulations, unlike S67$^{2.65}$, which does have a significant effect on residence time. Therefore, mutation of T68$^{2.66}$ to alanine may not affect the transient interaction energy made between this residue and the dissociating ligand. Our SMD protocol identifies transient contacts to ZM-241,385 that are experimentally proven to be important in binding, confirming the validity of our method for identifying receptor residues that engage with the ligand as it dissociates from the receptor, although some of these residues clearly have a greater effect on residence time than on binding affinity[59].

Sustained contacts, that is residues that are in contact with the dissociating ligand for the entirety of the 10 ns simulation, also play in a role in contributing to residence time; however, these can also have an effect on binding affinity. There are 5 sustained contacts: E169^{ECL2}, H264^{ECL3}, A265^{ECL3}, L267$^{7.32}$ and M270$^{7.35}$. Experimental data exist for three of these[59]. E169^{ECL2} is an initial binding contact (see Figure 8a) but it continues be a contact of the ligand throughout the entire simulation because of its location on ECL2 towards the top of the binding pocket. Experimental mutation of E169^{ECL2} to glutamine caused a decrease in residence time with an increase in binding affinity. Experimental mutation of H264^{ECL3} to alanine caused a decrease in residence time with an increase in binding affinity. Comparatively, E169^{ECL2} provides much more energy, both initially and throughout the 10 ns, to the ligand than H264^{ECL3} in the SMD simulations (see Figure 8b); this supports experimental findings where the E169Q^{ECL2} mutant had a greater effect on binding affinity and residence time than the H264A^{ECL3} mutant. Conversely L267$^{7.32}$ showed a decrease in binding affinity and an increase in RT when mutated to alanine. Sustained contacts, which can be identified by the SMD protocol

we have developed, play a role in the equilibrium binding affinity and non-equilibrium residence time. They have effects on these that are different to those observed for intermediate contacts.

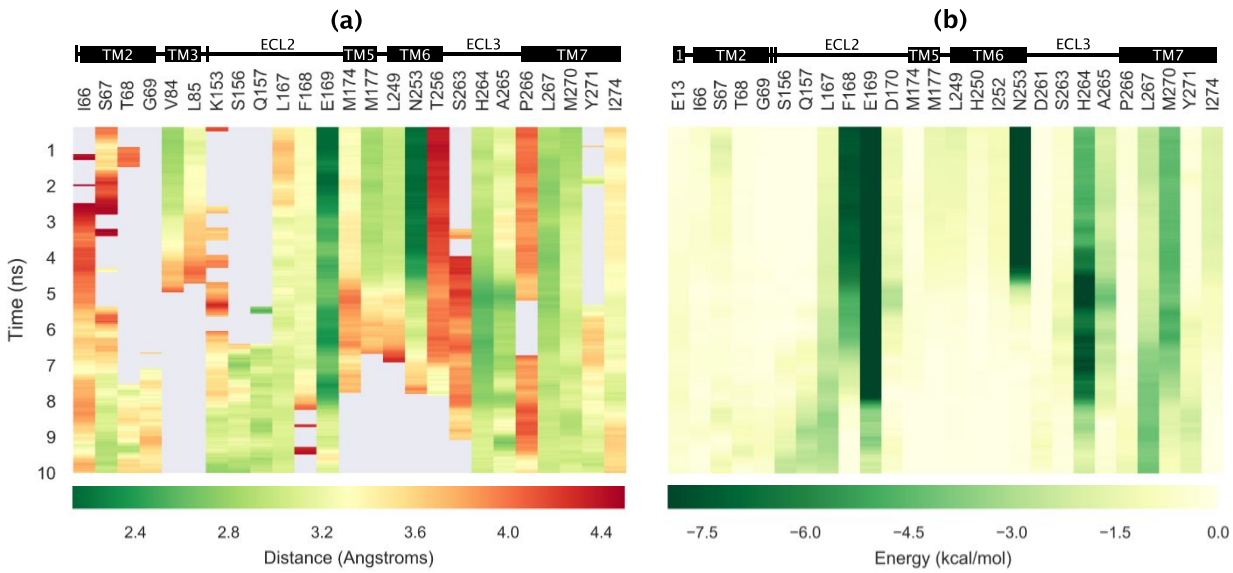


Figure 8. Heatmaps of the dissociation profile of the inverse agonist, ZM-241,385. A schematic representation of the receptor helices and loops involved in forming contacts and interactions is indicated with lettering above the residue numbers (TM = transmembrane domain; ECL = extracellular loop). Contacts are identified by the changing distance between residues of the A_{2A} receptor and the ligand during dissociation (a). Interactions are identified by the sum of van der Waals and electrostatic energy interactions between receptor residues and the ligand (b).

Residues that form multiple interactions

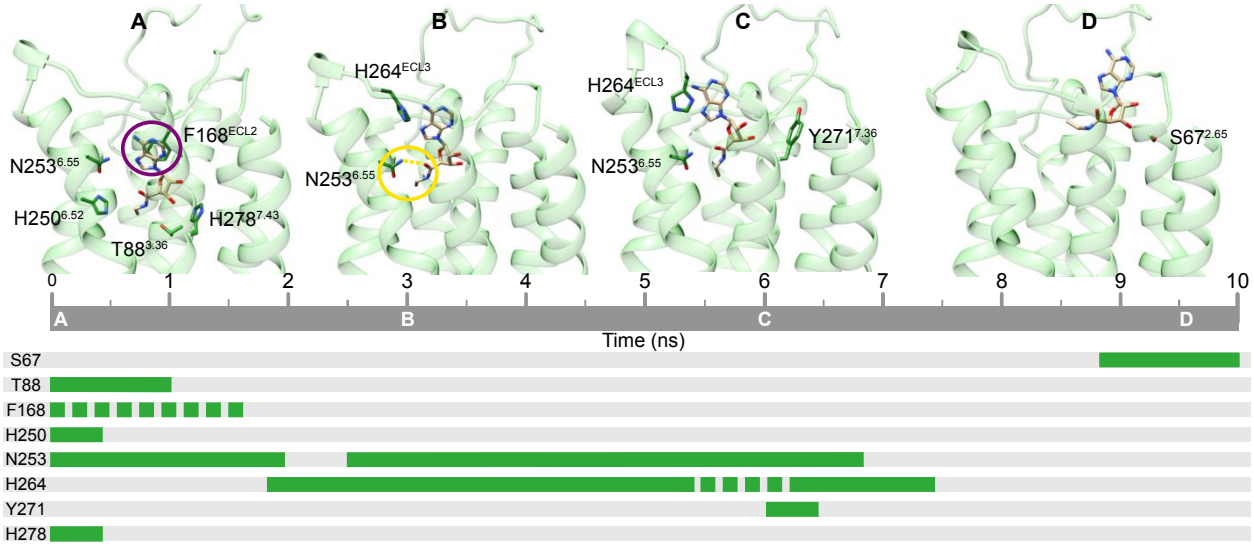


Figure 9. Interactions made by the agonist, NECA, as it dissociates from the A_{2A} receptor. Residues that form hydrogen bonds or π-stacking interactions with the ligand are shown in stick representation and are identified by a one letter amino acid code followed by their location in the receptor. The timeline of the simulation is shown by the dark grey box. Light grey boxes show replica averages for initial and intermediate interactions between ligand and specific receptor residues. The schematic representations shown in A-D are single replica snapshots. Hydrogen bond interactions (example highlighted by yellow ring) are shown as green solid lines, whilst π-stacking interactions (example highlighted by purple ring) are represented by a green dotted line.

Some residues interact with different parts of the ligand at different times during dissociation. This is especially true for A_{2A} receptor agonists, which are generally larger in size than antagonists and contain more moieties able to form interactions. Figure 9 shows the eight residues that the A_{2A} agonist NECA interacts with during dissociation. Two of these are multiple interactions. $N253^{6.55}$ initially makes a hydrogen bond to the adenine ring of NECA, this interaction is also present in the crystal structure (PDB accession code: 2YDV[59]). With the exception of LUF5967 and LUF5834, this particular hydrogen bond is conserved in all other A_{2A} receptor agonists and antagonists tested, making an interaction to the adenine and

methylxanthine rings, of the agonists and antagonists, respectively. With the exception of the LUF583x series, all adenosine receptor agonists tested in this study are closely related to the endogenous ligand adenosine in structure. Only in simulations with these adenosine-related agonists does one see $N253^{6.55}$ re-engaging with the ribose ring or with the ethylcarboxamido extension (in the case of NECA and CGS-21680). $H264^{ECL3}$ is an exception within the identified interacting residues as it makes both π-stacking and hydrogen bond interactions. $H264^{ECL3}$ first makes a hydrogen bond, at t=1.8 ns, with the adenine ring of the agonists, subsequently making π-stacking with the adenine ring of the ligand at t=5.4 ns. These results show that residues can make multiple distinct interactions with different parts of A_{2A} receptor agonists as they dissociate from the receptor.

Ligand-water interaction energy is a predictor of drug residence time

Residence times have been experimentally determined for all of the ligands investigated here (see Table 4). Water has been shown to be important for residence time[64]; therefore the computationally-determined non-bonded interaction energy between ligand and water molecules, E_{LW}, was calculated as a function of time. As shown in Figure 10, there is a good correlation between the change in interaction energy between the ligand and water, ΔE_{LW} (described in methods section), and the experimentally-determined residence times for the 17 A_{2A} receptor ligands tested ($R^2 = 0.79$ where n = 17). There is no correlation between ΔE_{LW} and experimentally-determined free energy of binding ($R^2 = 0.05$), providing further evidence that binding affinity and residence time values do not correlate.

Table 4. Experimental ligand binding data and ΔE_{LW} values of the adenosine A_{2A} receptor ligands used in this study

Ligand	Temp. (Kelvin)	k_{on} (nM^{-1} min^{-1})	k_{off} (min^{-1})	Residence time (mins)	Binding Free Energy (kcal/mol)	Reference for experimental kinetic data	Calculated ΔE_{LW} (kcal/mol)	Calculated PMF (kcal/mol)
CGS15943	283.15	0.30 ± 0.03	0.047 ± 0.005	21 ± 2	-12.69	Guo et al., 2016[65]	21.1 ± 1.7	8.3 ± 0.9
CGS21680	278.15	0.00005 ± 0.00001	0.02 ± 0.00	53.0 ± 0.2	-8.17	Guo et al., 2012[66]	30.7 ± 3.2	8.4 ± 1.0
LUF5448	278.15	0.00028 ± 0.00001	0.06 ± 0.02	16.0 ± 0.3	-8.46	Guo et al., 2012[6]	19.6 ± 2.4	9.0 ± 1.1
LUF5549	278.15	0.0024 ± 0.0005	0.04 ± 0.01	24.0 ± 0.2	-9.89	Guo et al., 2012[6]	18.7 ± 2.2	6.5 ± 0.6
LUF5550	278.15	0.0008 ± 0.0002	0.09 ± 0.02	12.0 ± 0.2	-8.86	Guo et al., 2012[6]	16.6 ± 3.2	4.7 ± 1.1
LUF5631	278.15	0.0008 ± 0.0002	0.05 ± 0.02	21.0 ± 0.4	-9.19	Guo et al., 2012[6]	14.1 ± 2.6	6.1 ± 1.0
LUF5833	278.15	0.0085 ± 0.003	0.16 ± 0.08	6.3 ± 0.5	-9.83	Guo et al., 2012[6]	3.0 ± 2.4	7.2 ± 0.8
LUF5834	278.15	0.011 ± 0.004	0.2 ± 0.1	4.2 ± 0.4	-9.77	Guo et al., 2012[6]	-0.8 ± 3.4	5.9 ± 0.6
LUF5835	278.15	0.016 ± 0.008	0.3 ± 0.1	3.4 ± 0.3	-9.86	Guo et al., 2012[6]	10.4 ± 3.7	12.8 ± 1.1
LUF5963	283.15	0.00023 ± 0.00003	0.044 ± 0.005	23 ± 3	-8.71	Guo et al., 2016[6]	10.4 ± 3.1	10.5 ± 1.2
LUF5964	283.15	0.0026 ± 0.0009	0.18 ± 0.02	5.6 ± 0.6	-9.27	Guo et al., 2016[66]	12.9 ± 2.0	5.5 ± 0.9
LUF5967	283.15	0.0005 ± 0.0001	0.12 ± 0.04	8.3 ± 2.8	-8.45	Guo et al., 2016[66]	15.9 ± 1.2	7.8 ± 0.6
NECA	278.15	0.00050 ± 0.00006	0.03 ± 0.01	35.0 ± 0.2	-9.21	Guo et al., 2012[66]	28.2 ± 2.4	1.7 ± 0.4
Theophylline	283.15	0.00006 ± 0.00001	0.18 ± 0.01	5.6 ± 0.3	-7.16	Guo et al., 2016[6]	3.1 ± 1.5	6.1 ± 0.9
UK-432,097	278.15	0.00050 ± 0.00008	0.004 ± 0.000	250 ± 0.8	-10.31	Guo et al., 2012[66]	61.0 ± 4.8	4.3 ± 0.8
XAC	283.15	0.006 ± 0.001	0.10 ± 0.03	10 ± 3	-10.10	Guo et al., 2016[6]	11.6 ± 3.6	6.7 ± 0.6
ZM-241,385	277.15	0.0210 ± 0.0005	0.014 ± 0.003	71 ± .1	-11.61	Guo et al., 2014[66]	24.7 ± 1.8	8.8 ± 0.6

As dissociation progresses, the ligand becomes increasingly hydrated with water. Ligands that have a low ΔE_{LW} value are comparatively well-hydrated in the binding pocket and hydrophilic interactions are less likely to be shielded from water. As buried hydrophilic interactions have been shown to increase residence time[67] their absence in ligands with low ΔE_{LW} may explain shorter residence times. The ΔE_{LW} value is not just an indication of the number of hydrogen bond capable atoms (hydrogen bond donors plus hydrogen bond acceptors) as evidenced by a poor correlation between these atoms and the ΔE_{LW} values ($R^2 = 0.57$). There is also a weak correlation between the number of hydrogen bond capable atoms and experimental

pRT ($R^2 = 0.48$). ΔE_{LW} is a summary of the of how solvent accessibility of the ligand changes during dissociation, this is dependent on the shape of both the ligand and the protein binding pocket.

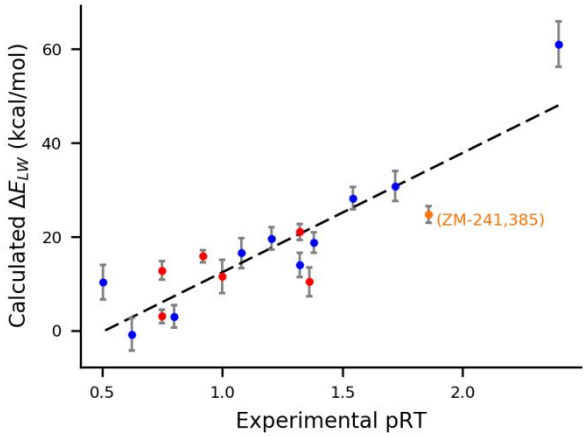


Figure 10. Correlation plot between the $\log_{10}$ of the experimentally determined residence time (pRT) and computationally-determined ΔE_{LW} for the 17 A_{2A} receptor ligands tested. The mean difference in the ligand-water energy (ΔE_{LW}) is calculated from the start to the end of these 10 ns SMD simulations. Error bars represent bootstrap standard deviations (σ_{boot}). Points are colored red or blue for antagonists and agonists, respectively. The inverse agonist ZM-241,385 is labelled and shown in orange.

Four of the ligands in the data set, CGS-21680, NECA, XAC and UK-432,097, are similar in base structure (see Figure 3) but have different residence times values (see Table 4). The rank order of experimentally-determined residence times and computationally-determined ΔE_{LW} is UK-432,097>>CGS-21680>NECA>XAC (see Table 5), which correlates directly with the location and number of aromatic ring extensions on these ligands (see Figure 3). Aromatic ring groups can influence buried hydrophilic interactions particularly when, as is the case for CGS-21680 and UK-432,097, the aromatic ring extension projects towards the extracellular vestibule (see Figure 2) and can prevent water from entering the ligand binding pocket. Aromatic

ring extensions at a distance from the adenine/xanthine ring are flexible and can increase both residence time and ΔE_{LW} by obstructing water solvation in the binding pocket. Support for this can be seen in the ligand structures shown in abstract figure. NECA and CGS-21680 are identical in structure apart from the aromatic ring extension on C2 of the adenine ring of CGS-21680, which has a longer residence time than NECA. Another example of a pair of ligands with a similar base structure but greatly different residence times is that of DPCPX and FSCPX[11], which we would attribute to the large aromatic ring extension on FSCPX. UK-432,097 has aromatic extensions on N^6 and C2 of the adenine ring and has a significantly longer residence time than CGS-21680. Modifying ligands to add a flexible aromatic ring extension to groups that project away from the binding pocket may be an effective design strategy for structural-based drug design.

Table 5. Experimentally-determined RT and computationally-determined ΔE_{LW}.

Ligand	**Calculated ΔE_{LW} (kcal/mol)**	**Experientially-determined RT (Minutes)**
XAC	11.6 ± 3.6	10 ± 3
NECA	28.2 ± 2.4	35 ± 0.2
CGS-21680	30.7 ± 3.2	53 ± 0.2
UK-432,097	61.0 ± 4.8	250 ± 0.8

XAC, with 59 atoms, is one of the largest ligands in the dataset. It has a low ΔE_{LW} (11.6 ± 3.6 kcal/mol) and a short residence time of 10 minutes, indicating that the correlation between ΔE_{LW} and residence time is not due to ligand size nor to the presence of an aromatic ring group. This is evidenced by the poor correlation between number of atoms of the ligands and their associated experimentally-determined pRT for the 17 ligands (R^2 = 0.47 with UK-432,097 included and 0.15 with UK-432,097 excluded from the correlation). The aromatic ring group of

XAC is immediately adjacent to the xanthine ring, so the cyclic group cannot block water from re-hydrating the ligand, which has been desolvated during its introduction into the ligand binding pocket[16]. It is the *position* of the aromatic ring group that influences residence time.

For 16 of the 17 ligands, the interaction between ligand and water correlates closely with residence time and appears to be the determining factor of this parameter. However, in the case of ZM-241,385, the calculated ΔE_{LW} is 24.7 ± 1.8 kcal/mol, which is lower than expected given its long experimentally-determined residence time of 71 minutes[68], suggesting that there are other forces contributing to residence time for this ligand. ZM-241,385's long residence time has previously been attributed to the stabilization of a $E169^{ECL2}$-$H264^{ECL3}$ salt bridge in the A_{2A} receptor that slows ligand dissociation[67]. Interactions between $E169^{ECL2}$ and $H264^{ECL3}$ were detected in the SMD runs we conducted for ZM-241,385 (see Figure 8).

The calculated work done (PMF) on the 17 ligands (see Table 4), the traditional method of predicting binding free energy when using SMD simulations, does not correlate with binding affinity nor residence time values. However, two of the ligands (CGS-21680 and LUF5448) have PMF values which are within error of the experimentally-determined binding free energy values. The structure of the active A_{2A} receptor used has CGS-21680 co-crystallized, however, in order to get accurate binding free energy predicting from using work done, one should use the crystal structure associated with each ligand.

CONCLUSIONS

We have described the development and use of an ensemble-based steered atomistic molecular dynamics (SMD) protocol that can be used to accelerate the process of ligand dissociation, allowing the reproducible identification of transient residue contacts and sustained residue interactions, also enabling the calculation of ΔE_{LW} values that correlate with experimentally-determined residence times. Our simulations explain published experimental SDM kinetic binding data for ZM-241,385 in terms of computationally-identified transient and sustained residue contacts. We found that use of the appropriate receptor state ('active' or 'inactive') is essential for the successful prediction of transient residue contacts. The protocol has been applied here to 17 ligands of the adenosine A_{2A} receptor, a well-characterized GPCR for which numerous high-resolution structures and substantial amounts of kinetic binding data exist, but it can be readily applied to other receptor-ligand complexes that have water exposed binding pockets for which good quality structural data is available.

For the A_{2A} receptor ligands tested, we have demonstrated that changes in water-ligand energy (ΔE_{LW}) from the ligand in the binding pocket to the extracellular vestibule is the most important factor in determining residence time. This is supported by the strong correlation between residence time and ΔE_{LW} ($R^2 = 0.79$). Based on trends observed in the experimental data and findings from the SMD simulations, we propose an aromatic ring extension design strategy for the targeted development of A_{2A} receptor ligands with increased residence times. We have developed and validated a novel and rational strategy for the computational prediction and rationalization of relative residence time using a steered MD-based methodology that is different to other methods as it uses interaction energies, rather than accelerated dissociation times. The SMD-determined relative residence times reliably correlate with experimentally-determined

residence time values for 17 structurally diverse ligands of the A_{2A} receptor and our strategy can be readily extended to other ligand-receptor systems. Automation of the methodology would need to be applied for routine use in academia and in industry, enabling the computational work needed for the calculation of relative residence times for a set of ligands to be achieved within 15 hours (the sum of the wallclock times for equilibration and SMD simulations); use of GPUs would accelerate this protocol even further.

This SMD method, as it currently stands, is less computationally expensive (a reduction of approximately 40%) and faster to carry out than RAMD[30], the other noteworthy relative residence time prediction method. Unlike any method currently used to predict residence time, the computational cost of this SMD method will be constant for a given receptor protein, an attribute that is especially beneficial as it allows one to accurately allocate computational resources to a project. According to their protocol, there is a potential increase of 400% in computational cost over the initial number of trajectories performed, this is peformed in a serial analysis fashion meaning that results take longer to be produced.

RAMD was performed on total of 70 HSP-90 ligands, appreciably more than the 17 A_{2A} receptor ligands in this manuscript, but the correlation obtained by the former method (R^2 = 0.45) is much weaker than the one presented here (R^2 = 0.79) and the correlation between molecular weight (MW) and residence time is stronger than their computed relative residence time values (R^2 = 0.61). The opposite is the case, using our method with the 17 A_{2A} receptor ligand dataset: correlation of MW and residence time (R^2 = 0.47) is much weaker than our correlation between ΔE_{LW} values and experimental residence times. Both our method and

RAMD were tested on a single receptor system. A recent MD method has been used to predict absolute k_{off} values for 6 ligands across 4 different systems[27] however, the fold difference between the predicted k_{off} values and the experimental data ranged from 1 to 10^7, with fewer than half being within a 10-fold difference. We intend to conduct a detailed and systematic comparison of these three methods using the A_{2A} adenosine receptor system described here, in order to provide an objective measure by which to quantify the accuracy of the results and the computational expense of each method.

AUTHOR INFORMATION

Notes

The authors declare no competing financial interest.

ABBREVIATIONS

GPCR – G protein-coupled receptor

PMF - Potential of mean force

SMD – Steered molecular dynamics

ACKNOWLEDGMENTS

We would like to thank our colleagues Dr. David W. Wright and Dr. Shunzhou Wan for their comments on the manuscript. This work was supported by the Biotechnology and Biological Sciences Research Council (grant numbers BB/P004245/1 and BB/M009513/1) and the London Interdisciplinary Bioscience PhD Consortium (LIDo). PVC and ATN are grateful for funding from the EU H2020 CompBioMed Centre of Excellence (grant number 675451). PVC is grateful

for funding from the MRC Medical Bioinformatics project (MR/L016311/1) and the EU H2020 VECMA (grant number 800925) and special funding from the UCL Provost. The authors acknowledge the use of the UCL Legion, Grace and Myriad High Performance Computing Facility (Legion@UCL, Grace@UCL and Myriad@UCL), and associated support services, in the completion of this work. We acknowledge the Leibniz Supercomputing Centre (LRZ) for providing access to SuperMUC. This work was also carried out on the Dutch national e-infrastructure with the support of SURF Cooperative, access to these resources was provided by CompBioMed HPC Centre of Excellence.

REFERENCES

1. Hauser, A. S.; Attwood, M. M.; Rask-Andersen, M.; Schiöth, H. B.; Gloriam, D. E. Trends in GPCR Drug Discovery: New Agents, Targets and Indications. *Nat. Rev. Drug Discov.* **2017**, *16* (12), 829–842.
2. Santos, R.; Ursu, O.; Gaulton, A.; Bento, A. P.; Donadi, R. S.; Bologa, C. G.; Karlsson, A.; Al-Lazikani, B.; Hersey, A.; Oprea, T. I.; Overington, J. P. A Comprehensive Map of Molecular Drug Targets. *Nat. Rev. Drug Discov.* **2016**, *16* (1), 19–34.
3. Shonberg, J.; Kling, R. C.; Gmeiner, P.; Löber, S. GPCR Crystal Structures: Medicinal Chemistry in the Pocket. *Bioorganic Med. Chem.* **2015**, *23* (14), 3880–3906.
4. Fredriksson, R.; Lagerström, M. C.; Lundin, L.-G.; Schiöth, H. B. The G-Protein-Coupled Receptors in the Human Genome Form Five Main Families. Phylogenetic Analysis, Paralogon Groups, and Fingerprints. *Mol. Pharmacol.* **2003**, *63* (6), 1256–1272.
5. Copeland, R. A. The Drug–Target Residence Time Model: A 10-Year Retrospective. *Nat. Drug Disv* **2016**, *15*, 87–95.

6. Guo, D.; Mulder-Krieger, T.; IJzerman, A. P.; Heitman, L. H. Functional Efficacy of Adenosine A2A Receptor Agonists Is Positively Correlated to Their Receptor Residence Time. *Br. J. Pharmacol.* **2012**, *166* (6), 1846–1859.

7. Sykes, D. A.; Dowling, M. R.; Charlton, S. J. Exploring the Mechanism of Agonist Efficacy: A Relationship between Efficacy and Agonist Dissociation Rate at the Muscarinic M3 Receptor. *Mol. Pharmacol.* **2009**, *76* (3), 543–551.

8. Vauquelin, G.; Charlton, S. J. Long-Lasting Target Binding and Rebinding as Mechanisms to Prolong in Vivo Drug Action. *Br. J. Pharmacol.* **2010**, *161* (3), 488–508.

9. Dowling, M. R.; Charlton, S. J. Quantifying the Association and Dissociation Rates of Unlabelled Antagonists at the Muscarinic M3 Receptor. *Br. J. Pharmacol.* **2006**, *148*, 927–937.

10. Tresadern, G.; Bartolome, J. M.; MacDonald, G. J.; Langlois, X. Molecular Properties Affecting Fast Dissociation from the D2 Receptor. *Bioorganic Med. Chem.* **2011**, *19* (7), 2231–2241.

11. Schmidtke, P.; Javier Luque, F.; Murray, J. B.; Barril, X. Shielded Hydrogen Bonds as Structural Determinants of Binding Kinetics: Application in Drug Design. *J. Am. Chem. Soc.* **2011**, *133* (46), 18903–18910.

12. Spagnuolo, L. A.; Eltschkner, S.; Yu, W.; Daryaee, F.; Davoodi, S.; Knudson, S. E.; Allen, E. K. H.; Merino, J.; Pschibul, A.; Moree, B.; Thivalapill, N.; Truglio, J. J.; Salafsky, J.; Slayden, R. A.; Kisker, C.; Tonge, P. J. Evaluating the Contribution of Transition-State Destabilization to Changes in the Residence Time of Triazole-Based InhA Inhibitors. *J. Am. Chem. Soc.* **2017**, *139* (9), 3417–3429.

13. Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug–Target Residence Time and Its

Implications for Lead Optimization. *Nat. Drug Disv* **2006**, *5*, 730–739.

14. Congreve, M.; Andrews, S. P.; Doré, A. S.; Hollenstein, K.; Hurrell, E.; Langmead, C. J.; Mason, J. S.; Ng, I. W.; Tehan, B.; Zhukov, A.; Weir, M.; Marshall, F. H. Discovery of 1,2,4-Triazine Derivatives as Adenosine A2A Antagonists Using Structure Based Drug Design. *J. Med. Chem.* **2012**, *55* (5), 1898–1903.
15. Tautermann, C. S. Impact, Determination and Prediction of Drug–Receptor Residence Times for GPCRs. *Curr. Opin. Pharmacol.* **2016**, *30*, 22–26.
16. Xia, L.; de Vries, H.; IJzerman, A. P.; Heitman, L. H. Scintillation Proximity Assay (SPA) as a New Approach to Determine a Ligand's Kinetic Profile. A Case in Point for the Adenosine A1 Receptor. *Purinergic Signal.* **2016**, *12* (1), 115–126.
17. Guo, D.; IJzerman, A. P.; Heitman, L. H. Importance of Drug–Target Residence Time at G Protein-Coupled Receptors – a Case for the Adenosine Receptors. In *Methods and Principles in Medicinal Chemistry*; 2015; pp 257–272.
18. Sadiq, S. K.; Wright, D. W.; Kenway, O. A.; Coveney, P. V. Accurate Ensemble Molecular Dynamics Binding Free Energy of Multidrug-Resistance HIV-1 Protease. *J. Chem. Inf. Model.* **2010**, *50*, 890–905.
19. Sadiq, S. K.; Wright, D.; Watson, S. J.; Zasada, S. J.; Stoica, I.; Coveney, P. V. Automated Molecular Simulation Based Binding Affinity Calculator for Ligand-Bound HIV-1 Proteases. *J. Chem. Inf. Model.* **2008**, *48*, 1909–1919.
20. Wan, S.; Bhati, A. P.; Zasada, S. J.; Wall, I.; Green, D.; Bamborough, P.; Coveney, P. V. Rapid and Reliable Binding Affinity Prediction of Bromodomain Inhibitors: A Computational Study. *J. Chem. Theory Comput.* **2017**, *13* (2), 784–795.
21. Wright, D. W.; Hall, B. A.; Kenway, O. A.; Jha, S.; Coveney, P. V. Computing Clinically

Relevant Binding Free Energies of HIV-1 Protease Inhibitors. *J. Chem. Theory Comput.* **2014**, *10* (3), 1228–1241.

22. Coveney, P. V.; Wan, S. On the Calculation of Equilibrium Thermodynamic Properties from Molecular Dynamics. *Phys. Chem. Chem. Phys.* **2016**, *18* (44), 30236–30240.

23. Sadiq, S. K.; Wright, D. W.; Kenway, O. A.; Coveney, P. V. Accurate Ensemble Molecular Dynamics Binding Free Energy Ranking of Multidrug-Resistant HIV-1 Proteases. *J. Chem. Inf. Model.* **2010**, *50* (5), 890–905.

24. Altwaijry, N. A.; Baron, M.; Wright, D. W.; Coveney, P. V.; Townsend-Nicholson, A. An Ensemble-Based Protocol for the Computational Prediction of Helix-Helix Interactions in G Protein-Coupled Receptors Using Coarse-Grained Molecular Dynamics. *J. Chem. Theory Comput.* **2017**, *13* (5), 2254–2270.

25. Shaw, D. E.; Bowers, K. J.; Chow, E.; Eastwood, M. P.; Ierardi, D. J.; Klepeis, J. L.; Kuskin, J. S.; Larson, R. H.; Lindorff-Larsen, K.; Maragakis, P.; Moraes, M. A.; Dror, R. O.; Piana, S.; Shan, Y.; Towles, B.; Salmon, J. K.; Grossman, J. P.; Mackenzie, K. M.; Bank, J. A.; Young, C.; Deneroff, M. M.; Batson, B. Millisecond-Scale Molecular Dynamics Simulations on Anton. In *Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis*; New York; pp 1–11.

26. Callegari, D.; Lodola, A.; Pala, D.; Rivara, S.; Mor, M.; Rizzi, A.; Capelli, A. M. Metadynamics Simulations Distinguish Short- and Long-Residence-Time Inhibitors of Cyclin-Dependent Kinase 8. *J. Chem. Inf. Model.* **2017**, *57* (2), 159–169.

27. Sun, H.; Li, Y.; Shen, M.; Li, D.; Kang, Y.; Hou, T. Characterizing Drug-Target Residence Time with Metadynamics: How to Achieve Dissociation Rate Efficiently without Losing Accuracy against Time-Consuming Approaches. *J. Chem. Inf. Model.*

2017, *57* (8), 1895–1906.

28. Rosenbaum, D. M.; Zhang, C.; Lyons, J. A.; Holl, R.; Aragao, D.; Arlow, D. H.; Rasmussen, S. G. F.; Choi, H.-J.; DeVree, B. T.; Sunahara, R. K.; Chae, P. S.; Gellman, S. H.; Dror, R. O.; Shaw, D. E.; Weis, W. I.; Caffrey, M.; Gmeiner, P.; Kobilka, B. K. Structure and Function of an Irreversible Agonist-B2 Adrenoceptor Complex. *Nature* **2011**, *469* (7329), 236–240.

29. Mollica, L.; Decherchi, S.; Zia, S. R.; Gaspari, R.; Cavalli, A.; Rocchia, W. Kinetics of Protein-Ligand Unbinding via Smoothed Potential Molecular Dynamics Simulations. *Sci. Rep.* **2015**, *5*.

30. Kokh, D. B.; Amaral, M.; Bomke, J.; Grädler, U.; Musil, D.; Buchstaller, H. P.; Dreyer, M. K.; Frech, M.; Lowinski, M.; Vallee, F.; Bianciotto, M.; Rak, A.; Wade, R. C. Estimation of Drug-Target Residence Times by τ-Random Acceleration Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2018**, *14* (7), 3859–3869.

31. Jarzynski, C. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.* **1997**, *78* (14), 2690–2693.

32. Martin, H. S. C.; Jha, S.; Coveney, P. V. Comparative Analysis of Nucleotide Translocation through Protein Nanopores Using Steered Molecular Dynamics and an Adaptive Biasing Force. *J. Comput. Chem.* **2014**, *35* (9), 692–702.

33. Martin, H.; Jha, S.; Howorka, S.; Coveney, P. V. Determination of Free Energy Profiles for the Translocation of Polynucleotides through Alpha-Hemolysin Nanopores Using Non-Equilibrium Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2009**, *5*, 2135–2148.

34. Buchsbaum, S. F.; Mitchell, N.; Martin, H.; Wiggin, M.; Marziali, A.; Coveney, P. V.;

Siwy, Z.; Howorka, S. Disentangling Steric and Electrostatic Factors in Nanoscale Transport through Confined Space. *Nano Lett.* **2013**, *13* (8), 3890–3896.

35. Ballesteros, J.; Weinstein, H. Integrated Methods for the Construction of Three-Dimensional Models and Computational Probing of Structure-Function Relations in G Protein-Coupled Receptors. *Methods Neurosci.* **1995**, *25* (C), 366–428.

36. Wang, Y.; Gkeka, P.; Fuchs, J. E.; Liedl, K. R.; Cournia, Z. DPPC-Cholesterol Phase Diagram Using Coarse-Grained Molecular Dynamics Simulations. *Biochim. Biophys. Acta - Biomembr.* **2016**, *1858* (11), 2846–2857.

37. Dakka, J.; Farkas-Pall, K.; Balasubramanian, V.; Turilli, M.; Wan, S.; Wright, D. W.; Zasada, S.; Coveney, P. V.; Jha, S. Enabling Trade-Offs between Accuracy and Computational Cost: Adaptive Algorithms to Reduce Time to Clinical Insight. *Proc. - 18th IEEE/ACM Int. Symp. Clust. Cloud Grid Comput. CCGRID 2018* **2018**, 572–577.

38. Robertson, N.; Jazayeri, A.; Errey, J.; Baig, A.; Hurrell, E.; Zhukov, A.; Langmead, C. J.; Weir, M.; Marshall, F. H. The Properties of Thermostabilised G Protein-Coupled Receptors (StaRs) and Their Use in Drug Discovery. *Neuropharmacology* **2011**, *60* (1), 36–44.

39. Lebon, G.; Edwards, P. C.; Leslie, A. G. W.; Tate, C. G. Molecular Determinants of CGS21680 Binding to the Human Adenosine A2A Receptor. *Mol. Pharmacol.* **2015**, *87* (87), 907–915.

40. Segala, E.; Guo, D.; Cheng, R. K. Y.; Bortolato, A.; Deflorian, F.; Doré, A. S.; Errey, J. C.; Heitman, L. H.; IJzerman, A. P.; Marshall, F. H.; Cooke, R. M. Controlling the Dissociation of Ligands from the Adenosine A2A Receptor through Modulation of Salt Bridge Strength. *J. Med. Chem.* **2016**, *59* (13), 6470–6479.

41. Pándy-Szekeres, G.; Munk, C.; Tsonkov, T. M.; Mordalski, S.; Harpsøe, K.; Hauser, A. S.; Bojarski, A. J.; Gloriam, D. E. GPCRdb in 2018: Adding GPCR Structure Models and Ligands. *Nucleic Acids Res.* **2017**, *46*, 1–7.

42. Fiser, A.; Sali, A. MODELLER: Generation and Refinement of Homology-Based Protein Structure Models. *Methods Enzymol.* **2003**, *374* (20), 461–491.

43. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera - A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612.

44. Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A Web-Based Graphical User Interface for CHARMM. *J. Comput. Chem.* **2007**, *29* (11), 1859–1865.

45. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30* (16), 2785–2791.

46. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–1802.

47. Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174.

48. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713.

49. Dickson, C. J.; Madej, B. D.; Skjevik, Å. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C. Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.* **2014**, *10*,

865–879.

50. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graph. Model.* **2006**, *25*, 247–260.

51. Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *J. Comput. Chem.* **2002**, *23* (16), 1623–1641.

52. Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (98), 10089–10092.

53. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38.

54. González, A.; Perez-Acle, T.; Pardo, L.; Deupi, X. Molecular Basis of Ligand Dissociation in β-Adrenergic Receptors. *PLoS One* **2011**, *6* (9), e23815.

55. Patel, J. S.; Berteotti, A.; Ronsisvalle, S.; Rocchia, W.; Cavalli, A. Steered Molecular Dynamics Simulations for Studying Protein-Ligand Interaction in Cyclin-Dependent Kinase 5. *J. Chem. Inf. Model.* **2014**, *54* (2), 470–480.

56. Lu, H.; Isralewitz, B.; Krammer, A.; Vogel, V.; Schulten, K. Unfolding of Titin Immunoglobulin Domains by Steered Molecular Dynamics Simulation. *Biophys. J.* **1998**, *75* (2), 662–671.

57. Isralewitz, B.; Baudry, J.; Gullingsrud, J.; Kosztin, D.; Schulten, K. Steered Molecular Dynamics Investigations of Protein Function. *J Mol Graph.Model.* **2001**, *19* (1), 13–25.

58. Davison, A. C.; Hinkley, D. V.; Young, G. A. Recent Developments in Bootstrap Methodology. *Stat. Sci.* **2003**, *18* (2), 141–157.

59. Guo, D.; Pan, A. C.; Dror, R. O.; Mocking, T.; Liu, R.; Heitman, L. H.; Shaw, D. E.; IJzerman, A. P. Molecular Basis of Ligand Dissociation from the Adenosine A2A Receptor. *Mol. Pharmacol.* **2016**, *89* (5), 485–491.

60. Jaakola, V.-P.; Lane, J. R.; Lin, J. Y.; Katritch, V.; IJzerman, A. P.; Stevens, R. C. Ligand Binding and Subtype Selectivity of the Human A2A Adenosine Receptor Identification and Characterization of Essential Amino Acid Residues. *J. Biol. Chem.* **2010**, *285* (17), 13032–13044.

61. Carpenter, B.; Lebon, G. Human Adenosine A2A Receptor: Molecular Mechanism of Ligand Binding and Activation. *Front. Pharmacol.* **2017**, *8*, 1–15.

62. Kim, J.; Wess, J.; Van Rhee, A. M.; Schoneberg, T.; Jacobson, K. a. Site-Directed Mutagenesis Identifies Residues Involved in Ligand Recognition in the Human A(2a) Adenosine Receptor. *J. Biol. Chem.* **1995**, *270* (23), 13987–13997.

63. Jiang, Q.; Van Rhee, A. M.; Kim, J.; Yehle, S.; Wess, J.; Jacobson, K. A. Hydrophilic Side Chains in the Third and Seventh Transmembrane Helical Domains of Human A2A Adenosine Receptors Are Required for Ligand Recognition. *Mol. Pharmacol.* **1996**, *50* (3), 512–521.

64. Lebon, G.; Warne, T.; Edwards, P. C.; Bennett, K.; Langmead, C. J.; Leslie, A. G. W.; Tate, C. G. Agonist-Bound Adenosine A2A Receptor Structures Reveal Common Features of GPCR Activation. *Nature* **2011**, *474* (7352), 521–525.

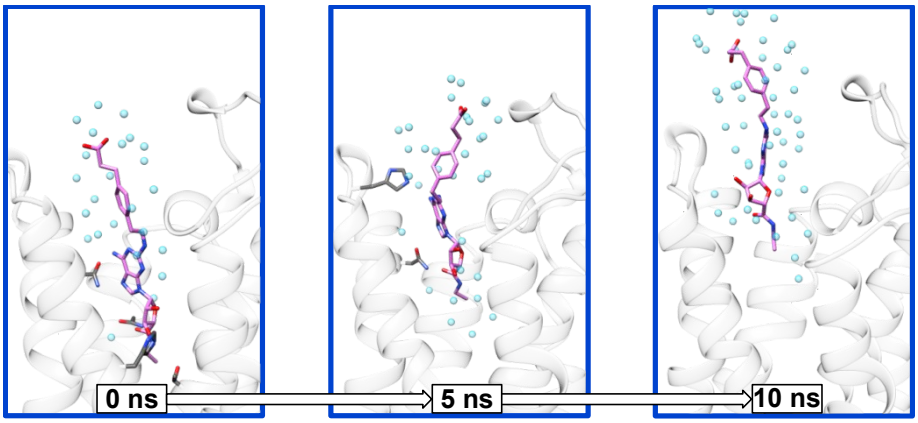
65. Tiwary, P.; Mondal, J.; Morrone, J. A.; Berne, B. J. Role of Water and Steric Constraints in the Kinetics of Cavity–Ligand Unbinding. *Proc. Natl. Acad. Sci.* **2015**, *112* (39), 12015–12019.

66. Guo, D.; Dijksteel, G. S.; Van Duijl, T.; Heezen, M.; Heitman, L. H.; IJzerman, A. P.

Equilibrium and Kinetic Selectivity Profiling on the Human Adenosine Receptors. *Biochem. Pharmacol.* **2016**, *105*, 34–41.

67. Guo, D.; Xia, L.; Van Veldhoven, J. P. D.; Hazeu, M.; Mocking, T.; Brussee, J.; IJzerman, A. P.; Heitman, L. H. Binding Kinetics of ZM241385 Derivatives at the Human Adenosine A2A Receptor. *ChemMedChem* **2014**, *9* (4), 752–761.

68. Dror, R. O.; Pan, A. C.; Arlow, D. H.; Borhani, D. W.; Maragakis, P.; Shan, Y.; Xu, H.; Shaw, D. E. Pathway and Mechanism of Drug Binding to G-Protein-Coupled Receptors. *Proc. Natl. Acad. Sci.* **2011**, *108* (32), 13118–13123.


0 ns
5 ns
10 ns