



OPEN

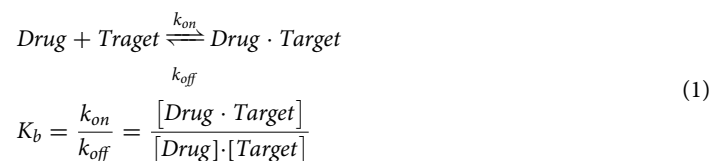
## Bell-Evans model and steered molecular dynamics in uncovering the dissociation kinetics of ligands targeting G-protein-coupled receptors

Muhammad Jan Akhunzada<sup>1</sup>, Hyun Jung Yoon<sup>1,2</sup>, Indrajit Deb<sup>1</sup>, Abdennour Braka<sup>1</sup> & Sangwook Wu<sup>1,2</sup>✉

Recently, academic and industrial scientific communities involved in kinetics-based drug development have become immensely interested in predicting the drug target residence time. Screening drug candidates in terms of their computationally predicted residence times, which is a measure of drug efficacy *in vivo*, and simultaneously assessing computational binding affinities are becoming inevitable. Non-equilibrium molecular simulation approaches are proven to be useful in this purpose. Here, we have implemented an optimized approach of combining the data derived from steered molecular dynamics simulations and the Bell-Evans model to predict the absolute residence times of the antagonist ZMA241385 and agonist NECA that target the A2A adenosine receptor of the G-protein-coupled receptor (GPCR) protein family. We have predicted the absolute ligand residence times on the timescale of seconds. However, our predictions were many folds shorter than those determined experimentally. Additionally, we calculated the thermodynamics of ligand binding in terms of ligand binding energies and the per-residue contribution of the receptor. Subsequently, binding pocket hotspot residues that would be important for further computational mutagenesis studies were identified. In the experiment, similar sets of residues were found to be in significant contact with both ligands under study. Our results build a strong foundation for further improvement of our approach by rationalizing the kinetics of ligand unbinding with the thermodynamics of ligand binding.

Conventional lead optimization methods in drug development through high-throughput screening rely on biochemical and biophysical assays of drug-target binding affinity under *in vitro* thermodynamic equilibrium conditions<sup>1,2</sup>. Notwithstanding their importance in determining the potency of a drug candidate, the traditional equilibrium binding affinity parameters, such as the dissociation constant ( $K_d$ ), inhibition constant ( $K_i$ ), half-maximal inhibitory concentration ( $IC_{50}$ ), and effector concentration ( $EC_{50}$ ), are arguable in terms of their effectiveness *in vivo*<sup>3</sup>. The equilibrium dissociation constant ( $K_d$ ) of a drug-target complex is a major determinant of the concentration or dose of a drug required for sufficient drug-target interaction. A lower  $K_d$  corresponds to an increased amount of the drug, which in turn suggests a better medicinal effect of the drug candidate but may trigger off-target interactions. In a drug-target binding equilibrium (Eq. 1), the ratio of the association rate constant ( $k_{on}$ ) to the dissociation rate constant ( $k_{off}$ ) describes the binding constant ( $K_b$ ), the reciprocal of which defines the dissociation constant ( $K_d$ ), as shown in Eq. (2).

<sup>1</sup>R&D Center, PharmCADD Co. Ltd., 12F, 331, Jungang-daero, Dong-gu, Busan 48792, Republic of Korea. <sup>2</sup>Department of Physics, Pukyong National University, Busan 48513, Republic of Korea. ✉email: s.wu@pharmacadd.com



In Eq. (1),  $k_{on}$ ,  $k_{off}$  and  $K_b$  are the association rate constant, dissociation rate constant and binding constant, respectively.

$$K_d = \frac{1}{K_b} = \frac{k_{off}}{k_{on}} \quad (2)$$

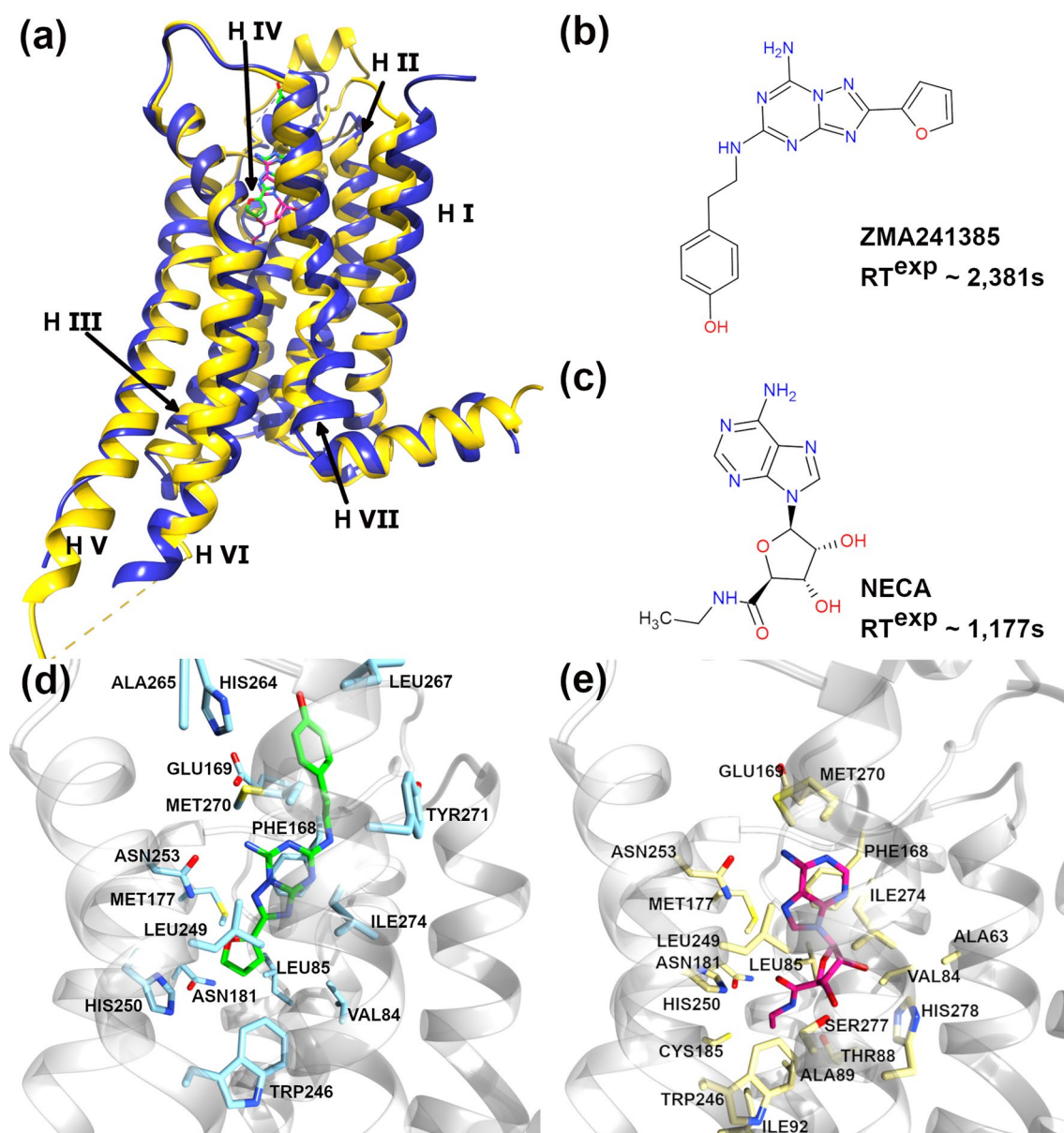
Equation (2) indicates that  $K_d$  would be directly correlated with  $k_{off}$ , considering that chemically similar drug candidates will possess similar  $k_{on}$ <sup>4–6</sup>. In principle, slower unbinding or lower  $k_{off}$  would correspond to better drug efficacy, ensuring long-lasting pharmacological activity. Intriguingly, during the last decade, the pharmaceutical community has become interested in measuring the mean lifetime or the mean residence time (RT) of a drug–target complex. The RT can be defined as the time over which the ligand stays bound in the binding pocket of a receptor and can be expressed as the reciprocal of the dissociation rate constant ( $k_{off}$ ). Both the RT and  $k_{off}$  have already been proven to be more reliable metrics for determining the pharmacological activity of drug candidates in terms of the therapeutic index, dosing, toxicity, and selectivity of drug molecules under open, non-equilibrium *in vivo* conditions compared to the drug–target binding affinity<sup>3</sup>.  $K_d$  determines the pharmacological activity of a drug candidate, whereas RT determines the lifetime of the activity. Increased longevity of the pharmacological activity is of great interest in order to keep the pathological activity of the receptor suspended for a longer period of time. Therefore, screening drug candidates in terms of their RT would be a promising approach in kinetics-driven drug design campaigns, suggesting that the binding affinities and RTs should be simultaneously assessed for the selection of efficient drug candidates. Estimation of the RT has now become a routine *in vitro* measurement using surface plasmon resonance spectroscopy, atomic force spectroscopy, stopped-flow circular dichroism spectroscopy, kinetic capillary electrophoresis, and radiolabeling combined with filtration or dialysis<sup>7</sup>. The RT helps compare the pharmacodynamic–pharmacokinetic durability of ligand hits during the lead optimization stage of drug development.

An in-depth understanding of the drug–target interaction equilibrium requires a complete characterization of the thermodynamics and kinetics of the interactions. Extensive efforts have already been devoted to developing efficient computational methods for capturing the thermodynamics of drug–target interactions in terms of binding energy calculations using molecular docking and molecular dynamics simulation approaches. Molecular docking approaches, despite their simplicity and speed, have shortcomings considering the effect of the explicit solvent environment and configurational entropy of the systems<sup>8</sup>. The widely explored molecular dynamics (MD) simulation approaches, such as the Poisson–Boltzmann or generalized Born surface area (MMPBSA/MMGBSA) approaches, and alchemical methods, such as thermodynamic integration (TI) and free energy perturbation (FEP), demand sufficiently longer simulations to be carried out and complicated postprocessing to estimate the ligand binding energies<sup>8–11</sup>. Umbrella sampling, on the other hand, suffers from the challenges associated with a priori knowledge of reaction coordinates<sup>8</sup>. Regarding the prediction of drug unbinding kinetic constants in accelerating drug candidate selection, computational approaches such as the application of Markov state models to unbiased MD simulations, smoothed potential MD simulations, and enhanced sampling algorithms such as weighted ensemble methods, milestoning, transition path sampling, steered MD, scaled and selectively scaled MD, random acceleration MD, adiabatic-bias MD simulations, and conformational flooding have become very promising<sup>12</sup>.

G-protein-coupled receptors (GPCRs), some of the important and leading therapeutic targets<sup>13–19</sup>, are the largest family of integral membrane proteins and are generally characterized by the presence of seven transmembrane alpha helices separated by alternating extracellular and intracellular loops. Despite multibillion drug sales in the pharmaceutical industry, GPCRs are some of the most underexploited drug targets, as only 10% of GPCRs make it to the drug industry<sup>15,16</sup>. One of the main reasons behind this is the lack of understanding of the atomistic details of the GPCR–ligand interactions in terms of their binding kinetics and thermodynamics, which are the definitive parameters in the clinical success of any drug–target system. Studies have already reported that small-molecule drugs targeting GPCRs and related protein receptors show a high correlation between their efficacy and the RT<sup>20–23</sup>. Adenosine receptors (ARs) are generally classified into four subtypes, A1, A2A, A2B, and A3, and are activated by extracellular adenosine. These receptors are potential therapeutic targets for various conditions, such as sleep disorders, cancer, and dementia, due to their involvement in physiological processes, such as sleep regulation, angiogenesis, and modulation of the immune system<sup>24</sup>. The A2A AR is one of the best structurally characterized GPCRs, with more than 30 crystal structures available<sup>25</sup>.

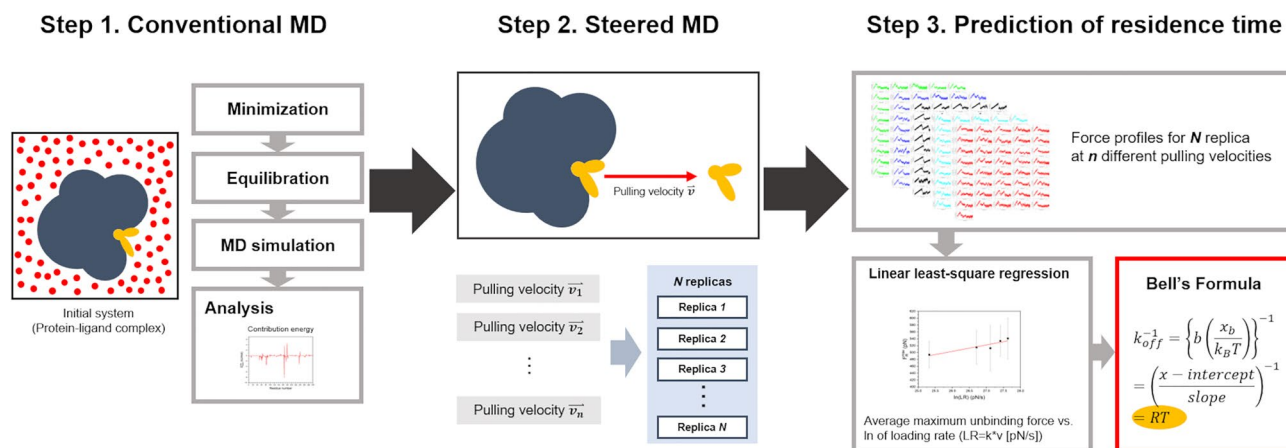
In this study, we develop an efficient computational approach for the prediction of the absolute RTs of drug candidates that target the A2A AR of the GPCR protein family (Fig. 1). Here, we implemented and tested our strategy of using steered molecular dynamics (SMD) simulations and the Bell–Evans model<sup>26–30</sup>, which relates the unbinding force ( $F_R$ ) to the parameters associated with the kinetics and energetics of ligand–receptor systems, as shown in Eq. (3).

$$F_R = \frac{k_B T}{x_b} \ln \frac{F' x_b}{k_b T k_{off}} \quad (3)$$



**Figure 1.** (a) The A2A AR-ligand complexes embedded into the 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipid bilayer; the A2A ARs corresponding to the PDB ids 3EML and 2YDV are shown in blue and yellow, respectively. The receptor helix (H) numbers are indicated in roman numbers. Experimentally determined RTs ( $RT^{\text{exp}}$ ) of ligands (b) ZMA241385 (4-(2-[7-amino-2-(2-furyl)-[1,2,4]triazolo[2,3-a][1,3,5]triazin-5-ylamino]ethyl)-phenol) and (c) NECA (N-Ethyl-5'-Carboxyamido Adenosine) are indicated. The binding pocket residues of the A2A AR within 5 Å of the ligands (d) ZMA241385 (green) and (e) NECA (magenta) are shown in blue and yellow sticks, respectively.

In Eq. (3),  $k_B$  is Boltzmann's constant,  $T$  is the absolute temperature,  $x_b$  is the reaction coordinate corresponding to the separation between the bound and the transition state, projected along the direction of the applied force,  $F$  is the loading rate (LR), defined as the product of the stiffness of the force transducer ( $k$ ) and velocity ( $v$ ), and  $k_{\text{off}}$  is the dissociation rate constant at equilibrium. We predicted the dissociation rate constant ( $k_{\text{off}}$ ) values and the corresponding RT values of the ZMA241385 and NECA ligands that bind to the A2A AR of the GPCR protein family. In addition to predicting the kinetic parameters, the thermodynamics of GPCR-ligand interactions were extensively explored in our study. The MMPBSA<sup>31</sup> approach-based binding energies of the ligands to the A2A AR and the interaction energies of these ligands with the binding pocket residues were estimated. Based on these energetics, structural hotspots in the A2A AR were identified. The results from this work build a strong foundation for further improvement of our approach of predicting the ligand dissociation rate constant ( $k_{\text{off}}$ ) by



**Figure 2.** Workflow used in this work for the estimation of energetics and residence time (RT). In step 1, conventional molecular dynamics (MD) simulation was performed on the initial solvated system; subsequently energy calculations were performed on the simulated trajectory. In step 2, the final coordinate from the conventional MD simulation was subjected to steered molecular dynamics (SMD) simulations with different pulling velocities ( $v$ ) and at a fixed spring force constant ( $k$ ). At each of the pulling velocities SMD simulations were carried out in multiple replicas by reassigning the velocities on the starting structure. In step 3, post-processing of the SMD simulation data was carried out by obtaining the unbinding force profiles along the SMD simulation trajectories and employing the Bell-Evans model.

rationalizing the kinetics of ligand unbinding with the thermodynamics of ligand binding. The steps involved in our approach for predicting the absolute RT and energetics are depicted in Fig. 2.

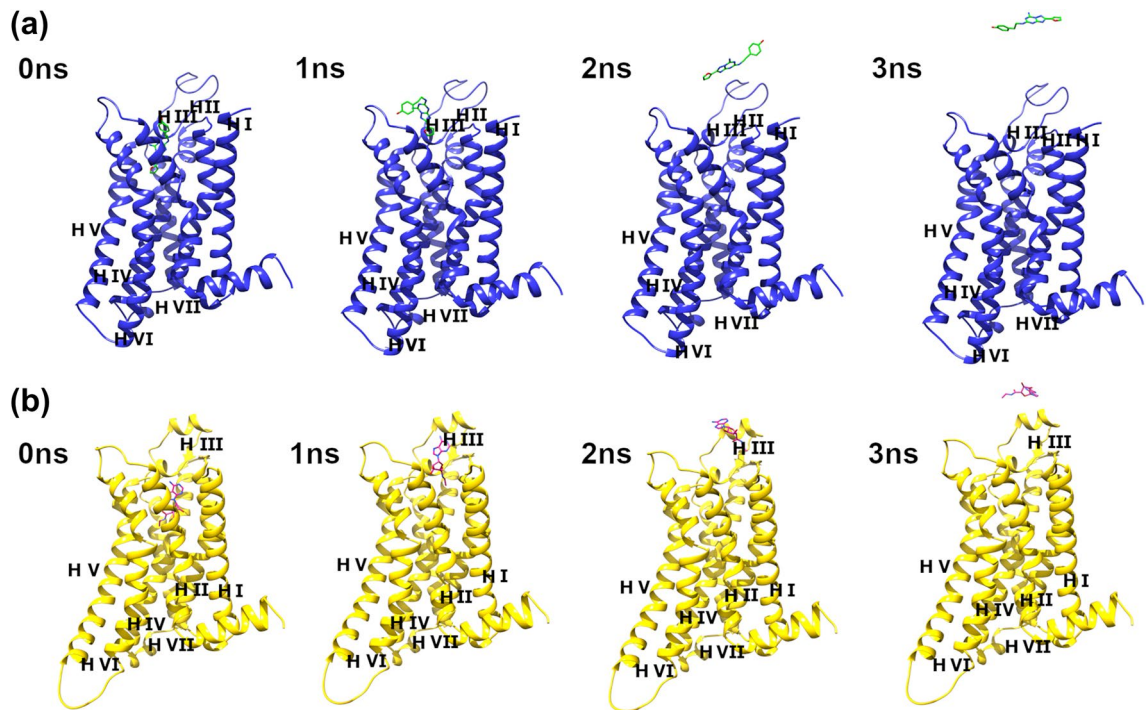
## Results

**Prediction of the absolute residence time (RT) from unbinding steered molecular dynamics (SMD) simulations.** The ZMA241385 and NECA ligands have a common triazolo scaffold, and their reported experimental dissociation rate constants are 0.0252/min and 0.0510/min, respectively, at 294.15 K, which correspond to experimental residence times ( $RT^{\text{exp}}$ ) of  $39.68 \pm 7.56$  min and  $19.61 \pm 3.69$  min, respectively<sup>32</sup>. Before performing the SMD simulations, long conventional all-atomistic MD simulations of 500 ns were performed for both complexes to check whether the ligands unbind within this timescale. Unbinding events were not observed within this 500 ns timescale, which makes these systems ideal for performing SMD simulations to accelerate ligand unbinding. The SMD simulations were performed in a series of slow to fast unbinding simulations with pulling velocities ranging from 0.0001 to 0.0004 nm/ps, 0.0006 nm/ps, 0.0008 nm/ps, and 0.0010 nm/ps and a fixed spring force constant of 600 kJ/mol/nm<sup>2</sup>. Figure 3 shows the unbinding of ZMA241385 and NECA ligands with a pulling velocity of 0.0010 nm/ps along the Z-axis of the lipid membrane that is directed toward the extracellular loops of the protein receptor.

The starting structure for the SMD simulations was the final snapshot of a 100 ns production run in the NPT ensemble. The structural stability of the complexes during this 100 ns simulation was examined through the root mean square deviation (RMSD), residue-wise root mean square fluctuation (RMSF) and radius of gyration ( $R_G$ ) (Supplementary Information, Table S1 and Figs. S1–S3). The average RMSD values for the protein receptors were found to be  $0.21 \pm 0.02$  nm and  $0.21 \pm 0.04$  nm for the ZMA241385 and NECA ligand–receptor complexes, respectively, while the average  $R_G$  values for the protein receptors were estimated as  $2.17 \pm 0.01$  nm and  $2.21 \pm 0.01$  nm, respectively.

The unbinding force ( $F_R$ ) profiles during the SMD simulations were monitored for 41 replica simulations at each of the pulling velocities (Supplementary Information, Figs. S5–S14). In Fig. 4a and b, representative  $F_R$  profiles from one of the 41 replica simulations at each of the pulling velocities are shown. Upon careful monitoring, all 41  $F_R$  profiles (Supplementary Information, Figs. S5–S14) were found to show similar observations. For a simpler understanding, the average of the maximum unbinding forces ( $F_R^{\text{max}}$ ) calculated from the 41 replica simulations at each of the pulling velocities is summarized in Table 1. We also extracted the SMD simulation times ( $T^{\text{max}}$ ) that correspond to the  $F_R^{\text{max}}$  values, and the average of  $T^{\text{max}}$  is reported in Table 1. Interestingly, the average  $T^{\text{max}}$  value at each of the pulling velocities was found to be higher for the ZMA241385 ligand than for the NECA ligand. This is quite reasonable considering that the  $RT^{\text{exp}}$  of the ZMA241385 ligand was reported to be higher than that of the NECA ligand ( $RT^{\text{exp}} = 39.68 \pm 7.56$  min and  $RT^{\text{exp}} = 19.61 \pm 3.69$  min for ZMA241385 and NECA, respectively, at 294.15 K)<sup>32</sup>. Additionally, the gradual decrease in the average  $T^{\text{max}}$  values with the gradual increase in the pulling velocity (Table 1) indicates that the choice of our SMD simulation parameters, i.e., the spring force constant ( $k$ ) and the pulling velocity ( $v$ ), were reasonable. We iteratively increased the number of replica simulations while monitoring the  $T^{\text{max}}$  values at different pulling velocities. We ended up with 41 replicas when we found that the  $T^{\text{max}}$  values gradually decreased with increasing pulling velocity. Comparatively lower standard deviations for the average  $T^{\text{max}}$  values at higher LR values indicate smooth ligand–receptor interactions along the unbinding pathway at higher LR values (Table 1).





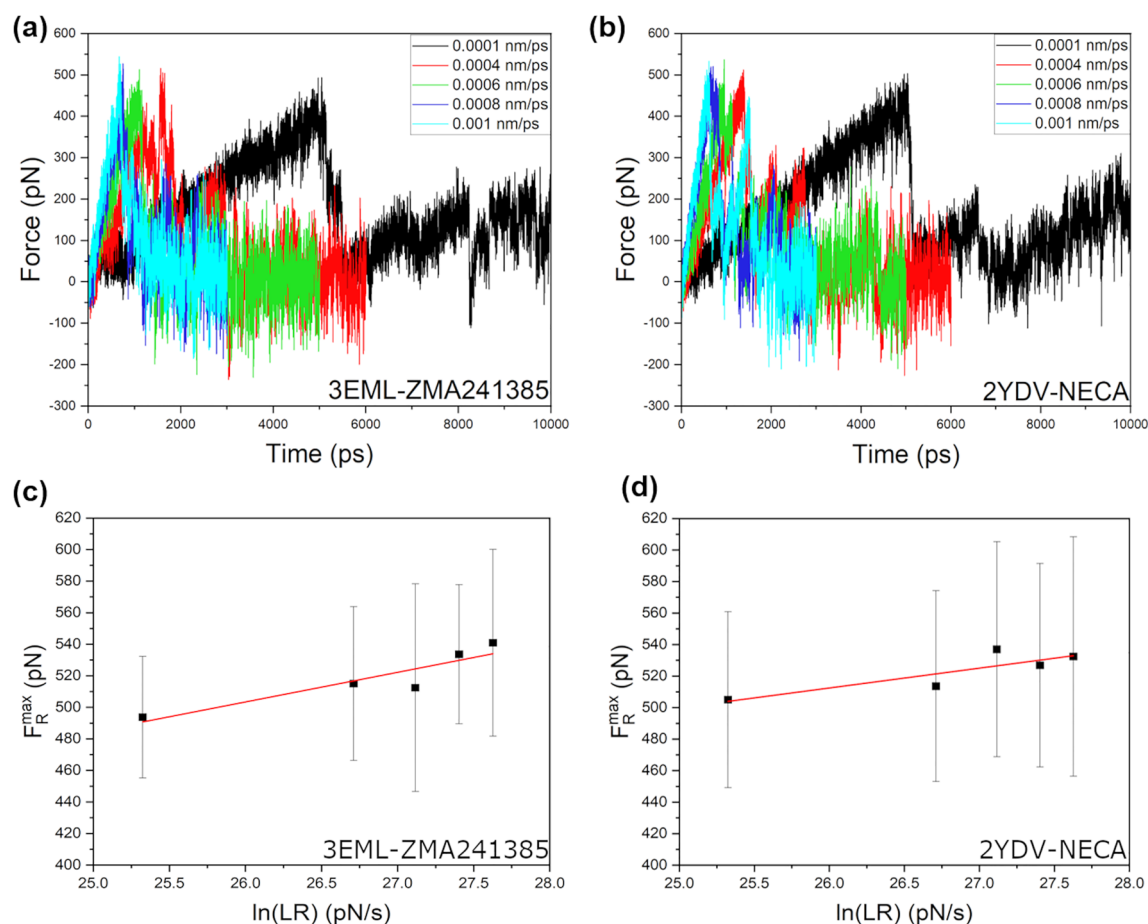
**Figure 3.** Snapshots at different time intervals from a representative SMD trajectory showing the unbinding of ligands with a pulling velocity of 0.0010 nm/ps. **(a)** Global view of the A2A AR (blue) in complex with the ligand ZMA241385 (green). **(b)** Global view of the A2A AR (yellow) in complex with ligand NECA (magenta). The receptor helix (H) numbers are indicated in roman numbers.

Notably, the average  $F_R^{\max}$  values showed mixed observations; no specific trends were observed in relation to the LR values. Additionally, high standard deviations were observed for the average  $F_R^{\max}$ . These observations could be associated with the stochastic nature of the individual ligand–receptor unbinding events inside the binding pocket and strongly correlated with the SMD simulation parameters, such as the pulling velocity and spring force constant. This observation can also be attributed to the reassignment of random velocities at 303.15 K at the start of each of the 41 replica SMD simulations at a particular LR. The initial velocity reassignment ensures random sampling of the initial conformational space for each of the replica simulations while pulling the ligand along the Z-axis. The  $F_R^{\max}$  values have been reported to be sensitive to subtle differences in experimental methods and the initial conformation of the ligand–receptor complex in SMD simulations<sup>29</sup>. Walton et al. found that for SMD simulations of biotin–streptavidin rupture, the unbinding force may vary by  $\sim 20\%$  with the initial configuration of the complex<sup>29</sup>.

We carried out an unpaired *t*-test for the ligand pair at a given pulling velocity for both the average maximum unbinding force ( $F_R^{\max}$ ) and average maximum time ( $T^{\max}$ ) (Table 1). From the results of the *t*-test, based on the conventional criteria, the difference for the  $F_R^{\max}$  at any given pulling velocity was found to be not statistically significant. Interestingly, for  $T^{\max}$ , at pulling velocities of 0.0001 nm/ps, 0.0004 nm/ps, and 0.0008 nm/ps, the differences were found to be significantly different. These results suggest that obtaining statistically significant differences between both the  $F_R^{\max}$  and  $T^{\max}$  values of the two ligands at a given pulling velocity are challenging and are the limitations of ligand unbinding rare-event simulations using SMD approach, when we compare two ligands at each of the pulling velocities. However, we would like to highlight that the main objective of this work is to predict the absolute RT of individual ligands and not to predict the relative RTs. As a secondary outcome of the absolute RT of a given ligand, qualitative ranking of different ligands can be carried out. Therefore, considering the trade-off between the computational cost of performing the simulations in multiple replicas and the reliability of the prediction, we relied only on monitoring of the  $T^{\max}$  values at different pulling velocities for a given ligand in selecting the number of replica simulations.

In this regard, the bootstrapping analysis on the  $F_R^{\max}$  values with increasing number of replicas provides additional confidence (Fig. 5). As shown in Fig. 5a and b, increase in the number of replicas results in significant change in the average  $F_R^{\max}$  values. This observation indicates that average  $F_R^{\max}$  values are sensitive to the choice of number of replicas. However, the standard deviation ( $F_R^{\max} \sigma$ ) at each of bootstrap steps for the average of the  $F_R^{\max}$  is gradually decreasing (Fig. 5c and d). The  $F_R^{\max} \sigma$  values for both the ligands were found to be converging around 40 replicas. Potterton et al. followed similar bootstrapping approach to decide on the choice of optimal number of replicas for the prediction of relative RTs of A2A AR binders. They ended up with selecting 10 replicas after performing 30 replica simulations for an effective trade-off between minimizing the error, stabilization of average property and the computational cost<sup>33</sup>.

Additionally, we generated average unbinding force profiles with the objective of selecting the number of replica simulations for both ligands at different LRs. To generate these profiles, averaging of the unbinding forces



**Figure 4.** Representative unbinding force profiles of one randomly selected replica out of 41 replica SMD simulations at each of the pulling velocities for (a) ZMA241385 and (b) NECA ligand-receptor complexes. Linear least-square regression of the average maximum unbinding forces ( $F_R^{\max}$ ) versus natural logarithms of the LR calculated from the SMD simulations of (c) ZMA241385 and (d) NECA ligand-receptor systems. Standard deviations for the average  $F_R^{\max}$  values are shown as black error bars.

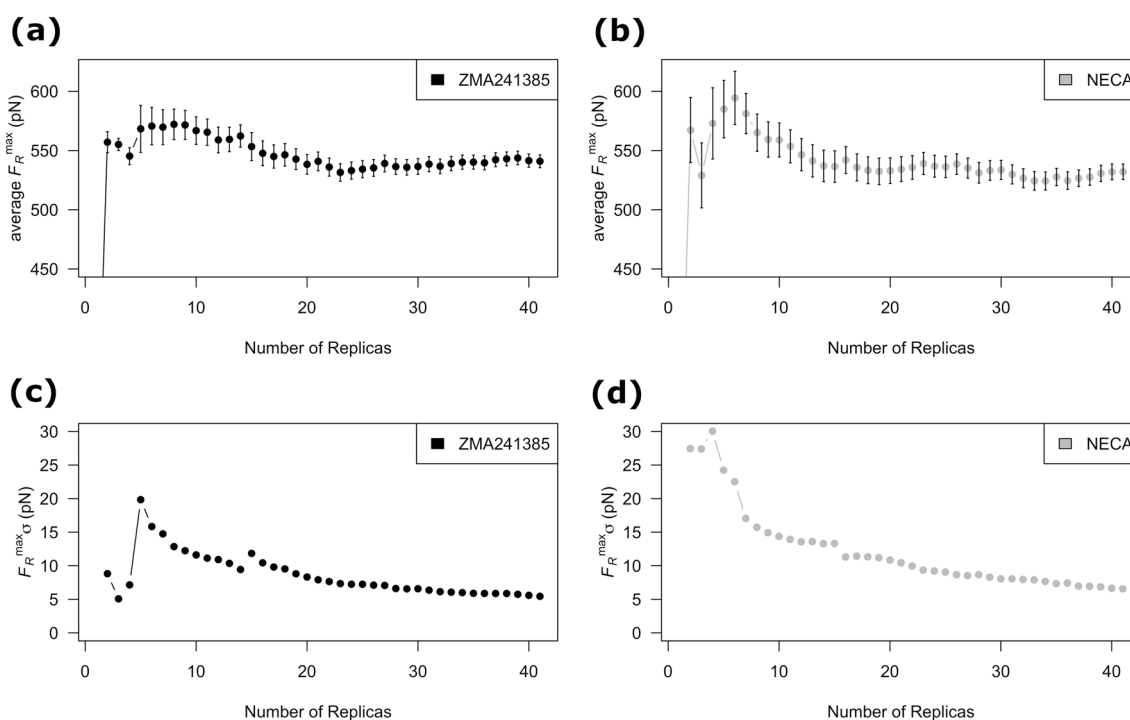
was carried out every 0.1 ps along the SMD trajectories of the 41 replica simulations (Supplementary Information, Fig. S4). For each of the two ligands, along the simulation trajectories, the computational times corresponding to the maxima of the average unbinding force profiles were found to gradually decrease with increasing pulling velocity. The characteristics of the average unbinding force profiles were found to be similar to the individual profiles (Fig. 4a, b and Supplementary Information, Figs. S5–S14 and Table S2). These observations suggest that 41 replicas were sufficient for the generation of a statistically reliable dataset for the employment of the Bell-Evans model.

Linear least-square regression of the average maximum unbinding force  $F_R^{\max}$  (pN) versus the natural logarithm of the LR (pN/s) (Table 1) was carried out (Fig. 4c, d). The goodness of the linear regressions was measured by the R-squared ( $R^2$ ) values, which were found to be 0.84 and 0.74 for the ZMA241385 and NECA ligand-receptor systems, respectively (Table 2). To estimate the dissociation rate constant ( $k_{\text{off}}^{\text{SMD}}$ ) from our SMD simulations, the Bell-Evans model was applied as described in the “Methods” section. The predicted  $k_{\text{off}}^{\text{SMD}}$  values were 0.04/s and 1.17/s for the ligands ZMA241385 and NECA, respectively (Table 2). The ligand absolute RTs estimated from the predicted  $k_{\text{off}}^{\text{SMD}}$  values were found to be 25.00 s and 0.86 s for the ligands ZMA241385 and NECA, respectively. In experimental studies, the ligand RT of ZMA241385 was also determined to be higher than that of NECA ( $\text{RT}^{\text{exp}} = 2380.80 \pm 453.60$  s and  $1176.60 \pm 221.40$  s for ZMA241385 and NECA, respectively, at 294.15 K)<sup>32</sup> (Table 2). However, compared to our prediction, the RTs reported from experimental studies were ~ 95-fold and ~ 1368-fold longer for the ZMA241385 and NECA ligands, respectively.

**Ligand binding affinity from conventional molecular dynamics (MD) simulations.** To better understand the energetics of ligand-receptor interactions inside the binding pocket of the A2A AR, the ligand binding affinities were estimated by calculating the MMPBSA-based ligand binding energy ( $E_{\text{bind}}$ ) and its components: the van der Waals (vdW) interaction energy ( $E_{\text{int}}^{\text{vdW}}$ ), electrostatic interaction energy ( $E_{\text{int}}^{\text{elec}}$ ), polar solvation energy ( $E^{\text{polar}}$ ), and nonpolar solvation energy ( $E^{\text{non-polar}}$ ) (Table 3). The ZMA241385 ligand was found to interact more favorably with the A2A AR, with a significantly lower binding energy of  $-87.42$  kJ/mol compared to the NECA ligand, the binding energy of which was estimated as  $-45.28$  kJ/mol. For both

$v$ (nm/ps)	N	PDB id	Ligand	$\ln(\text{LR})$ (pN/s)	$F_R^{\max}$ (pN)	$t$ -test	$T^{\max}$ (ps)	$t$ -test	$D_{\text{COM-COM}}^{\max}$ (nm)
0.0001	41	3EML	ZMA241385	25.32	$493.72 \pm 38.55$	$P(0.2760) > 0.05$	$5543.11 \pm 751.72$	$P(0.0022) < 0.05$	$0.47 \pm 0.12$
		2YDV	NECA		$505.05 \pm 55.78$		$5050.15 \pm 652.31$		$0.38 \pm 0.17$
0.0004	41	3EML	ZMA241385	26.71	$514.99 \pm 48.76$	$P(0.9091) > 0.05$	$1484.75 \pm 203.58$	$P(0.0001) < 0.05$	$0.46 \pm 0.11$
		2YDV	NECA		$513.60 \pm 60.51$		$1312.75 \pm 174.67$		$0.34 \pm 0.12$
0.0006	41	3EML	ZMA241385	27.12	$512.40 \pm 65.84$	$P(0.1006) > 0.05$	$1029.35 \pm 170.42$	$P(0.8379) > 0.05$	$0.49 \pm 0.12$
		2YDV	NECA		$538.82 \pm 68.26$		$1017.53 \pm 327.11$		$0.48 \pm 0.17$
0.0008	41	3EML	ZMA241385	27.40	$533.64 \pm 44.08$	$P(0.5796) > 0.05$	$793.23 \pm 90.07$	$P(0.0001) < 0.05$	$0.45 \pm 0.10$
		2YDV	NECA		$526.74 \pm 64.54$		$698.75 \pm 101.76$		$0.37 \pm 0.16$
0.0010	41	3EML	ZMA241385	27.63	$540.93 \pm 59.16$	$P(0.5749) > 0.05$	$634.05 \pm 88.46$	$P(0.1501) > 0.05$	$0.37 \pm 0.06$
		2YDV	NECA		$532.02 \pm 75.99$		$601.89 \pm 110.70$		$0.42 \pm 0.14$

**Table 1.** Loading rate (LR), average maximum unbinding force ( $F_R^{\max}$ ), average maximum time ( $T^{\max}$ ) and average maximum distance ( $D_{\text{COM-COM}}^{\max}$ ) where COM of ligand was pulled from the COM of protein residues within 5 Å of ligand at each of the pulling velocities ( $v$ ) from the SMD simulations of ZMA241385 and NECA ligand-receptor complexes. The LR are reported as their natural logarithms. The LR at different  $v$  were calculated by multiplying the spring force constant ( $k$ ), 600 kJ/mol/nm<sup>2</sup>, and the velocities ( $v$ ), 0.0001 nm/ps, 0.0004 nm/ps, 0.0006 nm/ps, 0.0008 nm/ps and, 0.0010 nm/ps. The average of maximum unbinding forces ( $F_R^{\max}$ ) and the corresponding average maximum times ( $T^{\max}$ ) were estimated from the maximum forces and corresponding maximum times at individual replica simulations out of 41 replicas. Similarly average maximum distances ( $D_{\text{COM-COM}}^{\max}$ ) between the COM of the ligand heavy-atoms and the COM of the binding pocket residues corresponds to maximum times at which the maximum forces were recorded for each replica of the system. Results from unpaired  $t$ -test for the ligand-pair at a given  $v$  for both the  $F_R^{\max}$  and  $T^{\max}$  at a confidence interval of 95% and the number of replicas (N) are listed.



**Figure 5.** Variation of the bootstrap statistics, (a) and (b), the average maximum unbinding force ( $F_R^{\max}$ ) and, (c) and (d), standard deviation ( $F_R^{\max} \sigma$ ) are shown as a function of number of replicas for the ligands ZMA241385 (black) and NECA (gray). The representative  $F_R^{\max}$  values used for the bootstrapping analysis were extracted from the SMD simulations at a pulling velocity of 0.0010 nm/ps. The error bars for the, (a) and (b), average  $F_R^{\max}$  values are the  $F_R^{\max} \sigma$  values.

PDB id	Ligand	y-intercept	Slope ( <i>m</i> )	x-intercept ( <i>b</i> )	<i>R</i> <sup>2</sup>	<i>k</i> <sub>off</sub> <sup>SMD</sup> (s <sup>-1</sup> )	RT <sup>SMD</sup> (s)	RT <sup>exp</sup> (s)
3EML	ZMA241385	14.90	18.79	0.79	0.84	0.04	25.00	2380.80 ± 453.60
2YDV	NECA	184.44	12.63	14.61	0.74	1.16	0.86	1176.60 ± 221.40

**Table 2.** Estimated statistics from the linear least-square regression of the average maximum unbinding forces ( $F_R^{\max}$ ) versus natural logarithms of the LR calculated from the SMD simulations of (a) ZMA241385 and (b) NECA ligand-receptor systems, the predicted dissociation rate constant ( $k_{\text{off}}^{\text{SMD}}$ ) and the corresponding residence time (RT<sup>SMD</sup>), and the experimental residence time<sup>32</sup> (RT<sup>exp</sup>) are reported.

PDB id	Ligand	$E_{\text{int}}^{\text{vdW}}$ (kJ/mol)	$E_{\text{int}}^{\text{elec}}$ (kJ/mol)	$E^{\text{polar}}$ (kJ/mol)	$E^{\text{non-polar}}$ (kJ/mol)	$E_{\text{bind}}^{\text{total}}$ (kJ/mol)
3EML	ZMA241385	- 142.27 ± 10.66	- 36.26 ± 9.99	108.47 ± 15.92	- 17.36 ± 0.97	- 87.42 ± 11.87
2YDV	NECA	- 121.89 ± 8.62	- 22.80 ± 8.21	114.90 ± 11.13	- 15.50 ± 0.84	- 45.29 ± 10.26

**Table 3.** Ligand-receptor binding energy ( $E_{\text{bind}}^{\text{total}}$ ) and its components, the vdW interaction energy ( $E_{\text{int}}^{\text{vdW}}$ ), the electrostatic interaction energy ( $E_{\text{int}}^{\text{elec}}$ ), the polar solvation energy ( $E^{\text{polar}}$ ) and, the non-polar solvation energy ( $E^{\text{non-polar}}$ ). The MMPBSA-based calculations were carried out on the last 10 ns of the 100 ns conventional MD simulations performed prior to the SMD simulations. The  $E^{\text{non-polar}}$  was estimated by using the non-polar solvation model of solvent accessible surface area (SASA). The contributions from the net non-bonded interaction energies ( $E_{\text{int}}^{\text{net}}$ ), those were calculated in vacuum, to the binding energies were decomposed into  $E_{\text{int}}^{\text{vdW}}$  and  $E_{\text{int}}^{\text{elec}}$ .

ligands,  $E_{\text{int}}^{\text{vdW}} < E_{\text{int}}^{\text{elec}} < E^{\text{non-polar}} < E^{\text{polar}}$ , indicating that the major contributions to the ligand binding energies are from the net nonbonded interaction energies ( $E_{\text{int}}^{\text{net}}$ ) that come from the summation of the  $E_{\text{int}}^{\text{vdW}}$  and  $E_{\text{int}}^{\text{elec}}$  components. The  $E_{\text{int}}^{\text{net}}$  values for the ZMA241385 and NECA ligands were estimated as - 178.53 kJ/mol and - 144.69 kJ/mol, respectively. Notably, the vdW interaction energy components of the net nonbonded interaction energies were found to contribute more to the stabilization of both ligands in the binding pocket. Other than the nonbonded interaction energies, stabilization of ligand binding through the nonpolar solvation energies was also observed for both ligands.

**Identifying energetic hotspots within the A2A adenosine receptor.** The MMPBSA-based ligand binding energies ( $E_{\text{bind}}^{\text{total}}$ ) were further decomposed at the individual residue level ( $E_{\text{bind}}^{\text{res}}$ ). In Table 4, we list the top 12 residues that contributed < - 1.00 kJ/mol to the total ligand binding energies. In the case of the A2A AR bound to the ligand ZMA241385, except for the residues ASP170 and GLU13, all the residues contributed most favorably to the total ligand binding energy through the molecular mechanics net nonbonded interaction energy ( $E^{\text{MM}}$ ) component, which can be defined as the summation of the vdW and electrostatic interaction energies.  $E^{\text{MM}}$  was also found to be the major component contributing to the ligand binding energy of NECA. For both the ZMA241385 and NECA ligands, the residue ASP170 was found to favorably contribute to the ligand binding energies through the polar solvation energy ( $E^{\text{polar}}$ ) component.

With the objective of identifying structural hotspots within the A2A AR, we list 12 common protein residues in the binding pocket of the A2A AR that are within 5 Å of the ligands ZMA241385 and NECA (Supplementary Information, Table S3). Interestingly, all 12 residues in proximity to the ligands did not favorably contribute to the ligand binding energies (Fig. 6a, c; Supplementary Information, Table S3). Based on the favorable contributions to the ligand binding energy of ZMA241385, the order of protein residues within 5 Å of the ligand is PHE168 > LEU249 > MET270 > ILE274 > MET177 > TRP246 > VAL84 > LEU85, whereas that for the NECA system is PHE168 > LEU85 > VAL84 > TRP246 > MET177. Based on these results, we selected the hotspots as VAL84, LEU85, PHE168, MET177 and TRP246, which are the common residues that contributed favorably to the binding energies of both the ligands ZMA241385 and NECA by < - 1.00 kJ/mol and are present within 5 Å of the ligands inside the binding pocket of the A2A AR (Table 4; Fig. 6b, d). Interestingly, residues LEU249, MET270 and ILE274 were found to favorably contribute to the binding energy of ZMA241385 by - 7.70 kJ/mol, - 6.26 kJ/mol and - 4.19 kJ/mol, respectively, whereas the same residues were found to contribute the least to the binding energy of NECA by - 0.53 kJ/mol, - 0.32 kJ/mol and - 0.61 kJ/mol, respectively (Supplementary Information, Table S3). Notably, GLU169, ASN181, HIS250, and ASN253 unfavorably contributed to the binding energies ( $E_{\text{bind}}^{\text{total}}$ ) of both ligands (Supplementary Information, Table S3).

## Discussion

In this work, we have predicted the dissociation rate constants ( $k_{\text{off}}^{\text{SMD}}$ ) and estimated the corresponding residence times (RT<sup>SMD</sup>) for the two ligands ZMA241385 and NECA that bind to the A2A AR of the GPCR protein family<sup>34,35</sup>. Our estimations of the absolute RTs were on the timescale of seconds, however, they were many folds shorter than those determined experimentally. We have tested the approach of combining the Bell-Evans model and SMD simulations for the prediction. The ZMA241385 ligand has been reported to be an A2A adenosine subtype-selective high-affinity antagonist of naturally occurring caffeine<sup>34</sup>. The NECA ligand, on the other hand, is an agonist of the A2A AR and contains a ribose moiety that is not found in ZMA241385<sup>35</sup> (Fig. 1b, c). Although

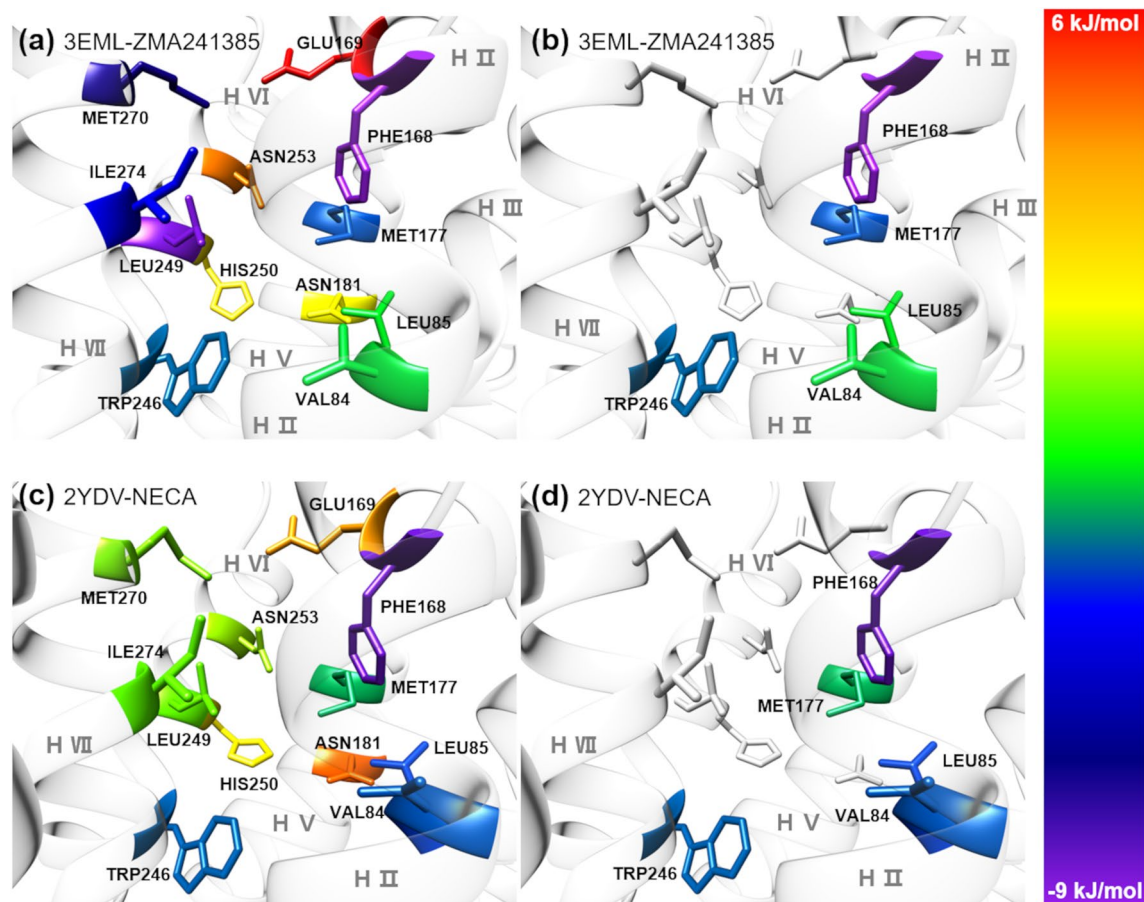


PDB id	Ligand	Residue	$E_{bind}^{res}$ (kJ/mol)	$E^{MM}$ (kJ/mol)	$E^{polar}$ (kJ/mol)	$E^{non-polar}$ (kJ/mol)
3EML	ZMA241385	PHE168	- 8.25 ± 0.15	- 11.72 ± 0.16	4.54 ± 0.12	- 1.07 ± 0.02
		LEU249	- 7.70 ± 0.11	- 8.81 ± 0.11	1.72 ± 0.07	- 0.61 ± 0.02
		MET270	- 6.26 ± 0.18	- 6.91 ± 0.21	1.50 ± 0.08	- 0.85 ± 0.03
		ILE274	- 4.19 ± 0.11	- 4.59 ± 0.12	1.09 ± 0.03	- 0.69 ± 0.02
		MET177	- 2.98 ± 0.08	- 3.74 ± 0.08	0.91 ± 0.05	- 0.15 ± 0.01
		TRP246	- 2.77 ± 0.11	- 3.14 ± 0.11	0.78 ± 0.05	- 0.41 ± 0.01
		ASP170	- 2.72 ± 0.10	- 0.45 ± 0.07	- 2.27 ± 0.08	0.00 ± 0.00
		GLU13	- 2.42 ± 0.11	0.67 ± 0.07	- 3.09 ± 0.11	0.00 ± 0.00
		TYR271	- 1.97 ± 0.19	- 3.52 ± 0.16	2.10 ± 0.13	- 0.55 ± 0.02
		LEU267	- 1.81 ± 0.09	- 1.36 ± 0.10	- 0.16 ± 0.04	- 0.29 ± 0.02
		VAL84	- 1.79 ± 0.08	- 2.67 ± 0.09	1.15 ± 0.04	- 0.27 ± 0.01
		LEU85	- 1.65 ± 0.10	- 2.17 ± 0.10	0.75 ± 0.02	- 0.23 ± 0.01
2YDV	NECA	PHE168	- 8.15 ± 0.19	- 12.26 ± 0.19	5.06 ± 0.14	- 0.95 ± 0.02
		LEU85	- 3.45 ± 0.10	- 6.24 ± 0.14	3.08 ± 0.10	- 0.29 ± 0.01
		VAL84	- 3.22 ± 0.10	- 5.79 ± 0.12	2.98 ± 0.06	- 0.41 ± 0.02
		TRP246	- 3.11 ± 0.11	- 3.72 ± 0.14	1.32 ± 0.08	- 0.71 ± 0.03
		MET177	- 2.32 ± 0.07	- 3.16 ± 0.09	1.21 ± 0.05	- 0.37 ± 0.02
		ALA63	- 1.95 ± 0.12	- 3.71 ± 0.15	2.21 ± 0.10	- 0.45 ± 0.02
		ASP52	- 1.69 ± 0.12	- 2.39 ± 0.13	0.70 ± 0.07	0.00 ± 0.00
		ILE60	- 1.25 ± 0.06	- 0.57 ± 0.05	- 0.51 ± 0.02	- 0.17 ± 0.01
		CYS185	- 1.16 ± 0.11	- 1.73 ± 0.08	0.73 ± 0.06	- 0.16 ± 0.01
		ALA89	- 1.11 ± 0.05	- 1.66 ± 0.05	0.65 ± 0.03	- 0.10 ± 0.01
		ILE66	- 1.10 ± 0.06	- 1.73 ± 0.06	0.72 ± 0.04	- 0.09 ± 0.01
		ASP170	- 1.01 ± 0.04	0.77 ± 0.05	- 1.78 ± 0.06	0.00 ± 0.00

**Table 4.** Top 12 residues of the A2A AR those contributed to the total ligand binding energies ( $E_{bind}^{total}$ ) by  $< -1.00$  kJ/mol. Decomposition of the residue-wise contributions ( $E_{bind}^{res}$ ) into the molecular mechanics net non-bonded interaction energies ( $E^{MM}$ ), polar solvation energies ( $E^{polar}$ ) and, non-polar solvation energies ( $E^{non-polar}$ ) are summarized. Residues within 5 Å of the ligands ZMA241385 and NECA in the binding pocket of the crystal structure of the A2A AR-ligand complex (PDB ids: 3EML and 2YDV, respectively) and that contributed favorably to the binding energies of both the ligands by  $< -1.00$  kJ/mol are shown in bold.

the chemically related regions of the ZMA241385 and NECA ligands have been reported to bind in a similar fashion in the binding pocket of the A2A AR, the antagonist activity of the ligand ZMA241385 was proposed to be associated with the restricted conformational freedom of helix V of the A2A AR<sup>35</sup>. As such, reliable prediction of the kinetic and thermodynamic parameters for these ligands would be of great interest for the development of drug candidates with extended pharmacological activity in vivo. Promisingly, this approach was found to be efficient enough to predict the absolute RTs of these ligands on the timescale of seconds.

From the binding affinity analyses of the ZMA241385 and NECA ligands in terms of their binding energies, a negative correlation was observed between the ligand binding affinities and the dissociation rate constants ( $k_{off}$ ), which agrees with Eq. (2). This observation confirms that stronger binding favors slower unbinding. Our results also suggest that modulating the nonbonded interaction energy between the ligands and the receptor, especially the vdW interaction energy components, would be useful for designing more potent binders. In this regard, identification of the hotspot residues believed to anchor the ligands in the binding pocket points toward a more specific direction. Among the identified hotspot residues VAL84, LEU85, PHE168, MET177 and TRP246, residue TRP246 has been reported to be in significant contact with the ZMA241385 ligand in the X-ray crystal structure<sup>34</sup>. Residue TRP246 has been found to contribute almost equally, by  $-2.77$  kJ/mol and  $-3.11$  kJ/mol, to the ligand binding energies ( $E_{bind}^{total}$ ) of ZMA241385 and NECA, respectively (Table 4). The hotspot residue PHE168 that contributed the most to the  $E_{bind}^{total}$  of both ligands has been reported to be in aromatic  $\pi$ -stacking interactions with both ligands in the X-ray crystal structures<sup>34,35</sup> (Table 4). The other two hotspot residues VAL84 and LEU85 have been found to be in close contact with the ribose moiety of the NECA ligand in the X-ray crystal structure<sup>35</sup>. Comparative studies on the X-ray crystal structures of ZMA241385 and NECA ligands bound to the A2A AR suggested a 2 Å shift of helix III to accommodate the ribose moiety of the NECA ligand<sup>35</sup>. Interestingly, the interaction pattern between the residue VAL84 and ligands was proposed to be specific for agonists, and mutagenesis studies have indicated a major role of this hotspot residue in the activation of the receptor<sup>35-37</sup>. In the X-ray crystal structures, close vdW contact between the hotspot residue MET177 and the NECA ligand has been reported, whereas MET177 is in hydrophobic contact with the ZMA241385 ligand<sup>34,35</sup>. Most interestingly, among the 12 common residues present within 5 Å of both ligands, GLU169, HIS250, and ASN253 unfavorably contribute to the ligand binding energies ( $E_{bind}^{total}$ ) by  $> 0.00$  kJ/mol (Supplementary Information, Table S3). These three residues have been reported to be points of mutations associated with disruption of agonist and/or antagonist binding and can be referred to as anti-hotspot residues<sup>38-40</sup>. Another residue, ILE274, which is also



**Figure 6.** Decomposition of the MMPBSA-based total ligand binding energies ( $E_{bind}^{total}$ ) at the individual residue level ( $E_{bind}^{res}$ ). 12 Common residues within 5 Å of the ligands (a) ZMA241385 and (c) NECA in the binding pocket of the crystal structure of the A2A AR-ligand complex (PDB ids: 3EML and 2YDV, respectively). 5 structural hotspot residues which are common and contribute favorably to the binding energies of both the ligands (b) ZMA241385 and (d) NECA by  $< -1.00$  kJ/mol and are present within 5 Å of the ligands inside the binding pocket of A2A AR. The residues are colored based on their contribution,  $E_{bind}^{res}$ , reported in Supplementary Information, Table S3, to the  $E_{bind}^{total}$  (see colour bar).

reported to be an important mutation point<sup>38–40</sup> and is present within 5 Å of both ligands, favorably contributes to the ligand binding energies of ZMA241385 and NECA by  $-4.19$  kJ/mol and  $-0.61$  kJ/mol (Supplementary Information, Table S3). Along this line, we predicted an additional common residue, ASN181, within 5 Å of ligands ZMA241385 and NECA that unfavorably contributes to the ligand binding energies by  $0.99$  and  $4.58$  kJ/mol, respectively, and can be considered for mutagenesis studies.

Mutation studies of the hotspot residues and chemical modification of the ZMA241385 and NECA ligands and subsequent iterative application of our approach for predicting the kinetic parameters and the computationally less expensive and easy-to-implement MMPBSA approach<sup>31</sup> could be considered a general purpose pipeline for the development of better therapeutics for diseases such as Parkinson's<sup>41,42</sup>, Huntington's<sup>43</sup>, asthma<sup>44</sup>, seizures<sup>45</sup>, pain<sup>46</sup> and many more<sup>47</sup> that are associated with the adenosine class of GPCRs. Notably, from biochemical studies, a 2-phenylhydrazino derivative of ZMA241385, namely, LUF5475, showed comparable and high affinities for the A2A and A2B ARs but was found to be selective for the A2B receptor compared to the A1 and A3 receptors<sup>48</sup>. Atomistic studies on the thermodynamics and kinetics of LUF5475-receptor interactions have yet to be explored. However, further validation of our approach is required for different GPCRs. Mutation studies will be beneficial for rationalizing the binding energies of ligands with respect to their RTs. Specifically, application of our approach for GPCRs with mutated hotspot residues and studying the effect of mutation on the ligand RT will allow additional validation of our strategy.

SMD simulations have already been found to be efficient for predicting the RT of diverse classes of protein receptors and ligand complexes<sup>29,33,49–51</sup>. Wong and coworkers in their two recent consecutive studies reported the applicability of SMD simulations in identifying and designing ligands with long RTs targeting focal adenosine kinase (FAK) as a test system<sup>52,53</sup>. Their most recent work reported a strong correlation between the computational ligand unbinding times for 14 ligands and their experimental dissociation rate constant ( $k_{off}$ ) values<sup>52</sup>. The nonpolar ligand-receptor interaction energy components and the distances of the nonpolar receptor residues from the ligands were found to be strongly correlated to the experimental  $k_{off}$  values. Although their regression model has been proposed to be useful in identifying chemically related ligands with desired dissociation rate

constants, one ligand appeared as an outlier in the regression analysis, and the model was trained for the FAK receptor only. Additionally, Wong and coworkers never attempted to predict the dissociation rate constants ( $k_{off}$ ) and corresponding RTs at the level of individual ligands using the data extracted from the SMD simulations in multiple replicas. The general applicability of our approach to any chemically unrelated ligand targeting any receptor molecule would require additional large-scale validation, which we are currently working on. Interestingly, application of the Bell-Evans model has already been proposed to be very promising in predicting the kinetic parameters for unbinding of streptavidin from biotin using SMD simulations<sup>29</sup>.

Unlike the selectively scaled molecular dynamics (ssMD) approach<sup>54</sup>, a limitation of our approach is associated with the positional restraints applied on the backbone heavy atoms of the receptor during the SMD simulations to prevent the receptor from drifting along the pulling direction. Another significant limitation of our approach is the prediction of absolute RTs on the timescale of seconds, which were found to be ~95-fold and ~1368-fold shorter for the ZMA241385 and NECA ligands, respectively, than those determined experimentally (Table 2). Similar differences have already been reported for the ssMD approach, in which a large discrepancy on the order of  $\sim 10^4$  was observed between the predicted unbinding times at the unscaled potential and experimentally measured RTs<sup>54</sup>. In this regard, an enhanced sampling approach, namely,  $\tau$ -random accelerated molecular dynamics ( $\tau$ -RAMD), has been reported to be very promising in predicting the ligand RT<sup>55–58</sup>. Similar to the SMD simulation approach, acceleration of ligand unbinding can be achieved in the  $\tau$ -RAMD method by applying an additional randomly oriented force in an adaptive manner to the ligand center of mass. The  $\tau$ -RAMD approach has been rigorously tested for membrane-embedded GPCRs, specifically the  $\beta 2$ -adrenergic receptor ( $\beta 2$ AR) and muscarinic acetylcholine receptor M2 (mAChR M2), in their complexes with orthosteric and allosteric compounds<sup>58</sup>. The results from the  $\tau$ -RAMD approach were found to be in robust agreement with those observed in experimental studies<sup>58</sup>. This approach was reported to correctly rank four ligands based on their predicted relative RTs<sup>58</sup>. In the ssMD approach, a short distance cutoff of 10 Å was chosen for the native contacts between the receptor and the ligand to ensure that ligand unbinding was a first-order process<sup>54</sup>. In our work, to employ the Bell-Evans model on SMD data, maximum unbinding forces ( $F_R^{max}$ ) were calculated along the SMD simulation trajectories. In SMD simulation of a ligand–receptor complex in which the ligand is pulled from the binding pocket of the receptor using a user-specified LR, the maximum unbinding force corresponds to the maximum force required to rupture all the native contacts between the ligand and the receptor for initial unbinding but not for complete unbinding<sup>59</sup>. In our work, for many of the unbinding force ( $F_R$ ) profiles, multiple peaks were observed (Supplementary Information, Figs. S5–S14). The initial maximum peaks correspond to the rupture of native contacts, and the following peaks can be assigned to the energy barriers associated with the short-lived interactions between the ligand and the receptor along the unbinding pathway. Interestingly, the average maximum distances ( $D_{COM-COM}^{max}$ ) between the center of mass (COM) of the ligand heavy atoms and the COM of the binding pocket were found to be as short as <5 Å, ranging from 3.7 to 4.9 Å for the 3EML system and 3.4–4.8 Å for the 2YDV system (Table 1). This observation indicates that the absolute RT estimated from the predicted dissociation rate constant ( $k_{off}$ ) corresponds to the initial unbinding of the ligands and provides a plausible explanation for the shorter timescale of the estimated RT compared to the experimentally determined RT. Additionally, the observed differences in the timescale may be associated with the choice of large LR in our study, which is in the range of 25–27 pN/s (as estimated from the natural logarithm of the LR), whereas the experimental LR typically falls below 8 pN/s, as measured for the biotin–streptavidin complex<sup>29</sup>. Notably, Walton et al. proposed a correction to the unbinding force ( $F_R$ ) of  $1/2kx_b$  to eliminate the dependence of  $k_{off}$  and  $x_b$  on the spring force constant ( $k$ ) and reported closer agreement between the kinetic parameters predicted from SMD simulations and those measured from experiments. This correction approach is beyond the scope of our study because we have performed SMD simulations when the spring force constant was fixed at a certain value and only the pulling velocity was varied<sup>29</sup>. Therefore, for constant velocity (fixed spring force constant) SMD simulations, appropriate choice of lower pulling velocities, accounting for lower LR, would be more appropriate for the prediction of the equilibrium dissociation rate constant ( $k_{off}$ ) from non-equilibrium SMD simulations. A lower pulling velocity would require a longer simulation time for the initial unbinding, as can be clearly seen from the unbinding force profiles at different pulling velocities (Supplementary Information, Figs. S5–S14 and Table 1). As such, our choices of pulling velocity and spring force constant were found to be reasonable in establishing a trade-off between the speed and accuracy in predicting the dissociation rate constants ( $k_{off}$ ) for the ZMA241385 and NECA ligands targeting the A2A AR and should be applicable to other systems as well. As a future scope of improving our approach for its wide applicability, we are currently working on extensive validation of our approach on a diverse variety of receptor–ligand systems. The limitations associated with our MMPBSA protocol, and the force field inaccuracies also cannot be ignored<sup>60</sup>. For simplicity and to reduce the computational cost, we disregarded the membrane in our MMPBSA calculations of ligand binding energies considering that the protein receptors, under study, experience similar membrane environments. However, inclusion of membranes implicitly or explicitly in the calculations may have impacted on the environment of the individual protein receptors and subsequently on the binding energies of the ligands to the proteins<sup>61</sup>. In traditional MMPBSA calculations, the polar solvation energy ( $E^{polar}$ ) is calculated considering a continuum of high dielectric constant around the molecular system of interest. Our protein–ligand systems in the presence of membranes may not have replicated such environment appropriately. Regarding the ligand force field parameters, we are planning to perform rigorous parameterization to improve the accuracy of our prediction. Allowing flexibility in the extracellular loops during our SMD simulations would be also interesting to study for capturing the intermediate states and exploring the variability in entrance and/or exit routes.

## Methods

**Computational details.** The molecular mechanics calculations and molecular dynamics (MD) simulations were performed using the GROMACS 2019<sup>62,63</sup> simulation package. The CHARMM36 protein force field<sup>64</sup> was used for the human A2A adenosine receptor, whereas for the 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipid bilayer, the CHARMM36 lipid force field<sup>65</sup> was used. The CHARMM general force field (CGenFF)<sup>66</sup> was used for parameterization of the ZM241385 and NECA ligands. The simulation systems and the initial GROMACS simulation scripts were generated using the CHARMM-GUI webserver<sup>67–69</sup>.

**System preparation.** The primary structures for the simulations of human A2A adenosine receptors in complex with the antagonist ligand ZMA241385 and agonist ligand NECA were obtained from the protein data bank (PDB) via the PDB accession numbers 3EML<sup>34</sup> and 2YDV<sup>35</sup> (Fig. 1b, c). The primary structures were loaded into the CHARMM-GUI webserver<sup>67–69</sup>, where the missing residues were built and the whole system was oriented along the principal Z-axis. The crystallographic water molecules were retained. The fully assembled system from the CHARMM-GUI web server was aligned, utilizing PyMOL molecular graphics software, with the respective coordinates downloaded from the OPM server<sup>70</sup>. The aligned systems for both structures were trimmed by shortening the C-terminal of the receptor for PDB id: 2YDV and by deleting the lysozyme part of the receptor for PDB id: 3EML to obtain a uniform receptor. The modified systems were then further processed through the CHARMM-GUI web server to obtain the final assembled systems with disulfide bonds. The ligand–receptor systems embedded inside the POPC lipid bilayer were built using the CHARMM-GUI membrane builder<sup>68</sup> (Fig. 1a). The systems were solvated with TIP3P<sup>71</sup> water molecules. The solvated systems were neutralized using K<sup>+</sup> ions. Excess KCl salt was added to obtain a final salt concentration of 0.15 M. The initial system parameters of both systems are summarized in Supplementary Information, Table S4.

**Conventional molecular dynamics (MD) simulation.** The initial systems were energy minimized by employing the steepest descent algorithm implemented in GROMACS using an energy threshold of 10 kJ/mol/nm. After minimization, equilibration simulations were performed for 25 ns in the NVT ensemble. Temperature and pressure coupling during the equilibration runs were applied using the Berendsen algorithm<sup>72</sup>. Restraints on heavy atoms of the lipids, protein and ligand backbones and side chains were applied during the equilibration run. After equilibration, 100 ns production runs in an isothermal-isobaric ensemble (NPT) were carried out for both systems, in which all the restraints were lifted. During the production run, temperature coupling was applied using the Nose–Hoover algorithm<sup>73,74</sup>, whereas the Parrinello–Rahman algorithm<sup>75</sup> was used for the pressure coupling. A temperature coupling constant of 1.0 ps and a pressure coupling constant of 5.0 ps were employed. Throughout the production run, the pressure was kept constant at 1 atm. An integration timestep of 2 fs was used for both the equilibration and production runs. Simulations were performed at a temperature of 303.15 K. Nonbonded interactions were truncated at 1.2 nm with the Verlet cutoff scheme<sup>76</sup>. Long-range electrostatic interactions were modeled by the particle mesh Ewald (PME) technique<sup>77</sup>. Hydrogen bonds were constrained using the LINCS<sup>78</sup> algorithm. Simulations were subjected to periodic boundary conditions in a semi-isotropic environment. The system coordinates were saved every 50,000 steps, which corresponded to 100 ps during the production run. Additionally, an independent 500 ns production run simulation was performed for both systems by regenerating the velocities after equilibration.

**Steered molecular dynamics (SMD) simulation.** The velocities after the 100 ns production run were discarded. New random velocities were assigned to the final coordinates of the 100 ns run in 41 replicas, and normal simulations were continued for 50 ps before applying the pulling velocity. The time evolutions of the temperature and kinetic energy for the 50 ps run from one representative simulation out of 41 replicas are presented in the Supplementary Information as Fig. S15. The time evolutions clearly indicate the equilibration of newly assigned velocities within 50 ps. At the end of the 50 ps equilibration simulations, SMD simulations were initiated by pulling the center of mass (COM) of the heavy atoms of the ligands from the COM of the residues of the A2A adenosine receptor within 5 Å of the ligands. The pulling of the ligands was performed along the Z-axis of the lipid membrane, directed toward the extracellular loops of the protein. We applied pulling velocities of 0.0001 nm/ps, 0.0004 nm/ps, 0.0006 nm/ps and 0.0008 nm/ps and 0.0010 nm/ps and a spring force constant of 600 kJ/mol/nm<sup>2</sup> during our SMD simulations. Regarding the choice of the spring force constant, we performed many trials using 200 and 400 kJ/mol/nm<sup>2</sup> for the lowest pulling velocity simulations. We did not observe many unbinding events within 10 ns of the simulations for the spring force constant values of 200 and 400 kJ/mol/nm<sup>2</sup>. Spring force constant values higher than 600 kJ/mol/nm<sup>2</sup> would accelerate the ligand unbinding events but may very abruptly disrupt the receptor–ligand interaction in the binding pocket. Therefore, we selected 600 kJ/mol/nm<sup>2</sup> as the spring force constant. The time lengths of the SMD simulations at each of the velocities were set to ensure complete unbinding of the ligands from the binding pocket of the A2A adenosine receptor. For pulling velocities of 0.0001 nm/ps, 0.0004 nm/ps, 0.0006 nm/ps, 0.0008 nm/ps and 0.0010 nm/ps, SMD simulations were carried out for 10 ns, 6 ns, 5 ns, 3 ns and 3 ns, respectively. A positional restraining force of 1000 kJ/mol/nm<sup>2</sup> was applied on the receptor backbone heavy atoms to avoid drift of the whole receptor in the pulling direction. The coordinates were recorded every 1 ps, and the unbinding forces were written out at each 0.1 ps of the SMD trajectories.

**Choice of optimal number of replicas SMD simulations.** To decide on the optimal number of replicas, the averages of the maximum unbinding forces ( $F_R^{\max}$ ) calculated from the replica simulations at one of the pulling velocities of 0.0010 nm/ps were monitored with increasing number of replicas. This is because  $F_R^{\max}$  is the key parameter extracted from the SMD simulations and is the basis of the calculation of the dissociation rate



constant ( $k_{off}$ ) following Bell-Evans model as described below. To examine the effect of number of replicas on the average  $F_R^{max}$ , we carried out bootstrapping analysis<sup>79</sup>. In our bootstrapping analysis, the  $F_R^{max}$  data points were resampled 10000 times. The average of the  $F_R^{max}$  and standard deviation ( $F_R^{max} \sigma$ ) at each bootstrap was calculated. The  $F_R^{max} \sigma$  provides an estimate of error associated to the average  $F_R^{max}$  calculated at each bootstrap step.

**Employment of the Bell-Evans model.** The dissociation rate constant ( $k_{off}$ ) and the energetic unbinding distance ( $x_b$ ) between the bound state and the energetic maximum for each of the ligands were predicted simply from the slope ( $m$ ) and x-intercept ( $b$ ) of the linear least-square regression of the average maximum unbinding force ( $F_R^{max}$ ) versus the natural logarithm of the loading rate (LR),  $\ln(\text{LR})$ , extracted from the SMD simulations. The LR at different velocities was calculated by multiplying the spring force constant ( $k$ ), 600 kJ/mol/nm<sup>2</sup>, and the velocities ( $v$ ), 0.0001 nm/ps, 0.0004 nm/ps, 0.0006 nm/ps, 0.0008 nm/ps and 0.0010 nm/ps. Following the Bell-Evans model (Eq. 3), the energetic distance ( $x_b$ ) between the bound state and the energetic maximum can be expressed as

$$x_b = \frac{k_B T}{m} \quad (4)$$

where  $k_B$  is Boltzmann's constant,  $T$  is absolute temperature, and  $m$  is the slope, and the dissociation rate constant ( $k_{off}$ ) can be expressed as

$$k_{off} = b \left( \frac{x_b}{k_B T} \right) \quad (5)$$

By combining Eqs. (4) and (5).

$$k_{off} = b \left( \frac{k_B T}{m} \right) \left( \frac{1}{k_B T} \right) = \frac{b}{m} \quad (6)$$

where  $b$  refers to the x-intercept and  $m$  refers to the slope of the linear least-square regression of  $F_R^{max}$  versus  $\ln(\text{LR})$ . The inverse of the  $k_{off}$  value (i.e.,  $1/k_{off}$ ) yields the absolute RT of the ligand.

**Trajectory analysis.** Trajectories were analyzed using the analysis utilities embedded in the GROMACS package. The root mean square deviation (RMSD), root mean square fluctuation (RMSF) and radius of gyration ( $R_G$ ) were computed to confirm the structural stability of the A2A adenosine receptor–ligand complexes. The  $C_\alpha$ -atoms of the protein residues were selected for the RMSD calculations. The RMSD of the ligands was also calculated separately by selecting all ligand heavy atoms. The molecular mechanics Poisson-Boltzmann surface area (MMPBSA) approach was used to calculate the ligand binding energies<sup>31</sup>. In our MMPBSA calculations, we did not consider the membrane environment. However, the importance of taking it into account cannot be completely ignored. In this context, the limitations of our MMPBSA protocol are discussed in the “Discussion” section. The MMPBSA calculations were performed using the `g_mmpbsa` package developed under the Open-Source Drug Discovery Consortium (OSDD)<sup>31</sup>. It is a GROMACS tool for calculating the binding energy using the MMPBSA method without the entropic term. The ligand binding energy ( $E_{bind}^{total}$ ) was estimated as the summation of the vdW interaction energy ( $E_{int}^{vdW}$ ), electrostatic interaction energy ( $E_{int}^{elec}$ ), polar solvation energy ( $E^{polar}$ ), and nonpolar solvation energy ( $E^{non-polar}$ ). The net nonbonded interaction energy ( $E_{int}^{int}$ ), which is the summation of  $E_{int}^{vdW}$  and  $E_{int}^{elec}$ , was calculated in vacuum.  $E^{non-polar}$  was estimated by using the nonpolar solvation model of the solvent accessible surface area (SASA). The last 10 ns of the 100 ns production run simulations were used to perform the binding energy and energy decomposition analysis. The dielectric constant was set to 2 for the calculation of the potential energy in vacuum. All other parameters were set to their default values. Furthermore, per-residue decomposition analyses were performed to obtain the energetic contributions from the amino acid residues to the total ligand binding energies. Additionally, the per-residue energetic contributions ( $E_{bind}^{res}$ ) were decomposed into the molecular mechanics net nonbonded interaction energies ( $E^{MM}$ ), polar solvation energies ( $E^{polar}$ ) and nonpolar solvation energies ( $E^{non-polar}$ ).

## Data availability

Raw data is available upon request to the corresponding author.

Received: 15 April 2022; Accepted: 8 September 2022

Published online: 24 September 2022

## References

- Ciulli, A. Biophysical screening for the discovery of small-molecule ligands. *Methods Mol. Biol. (Clifton, N.J.)* **1008**, 357–388 (2013).
- Hughes, J., Rees, S., Kalindjian, S. & Philpott, K. Principles of early drug discovery. *Br. J. Pharmacol.* **162**, 1239–1249 (2011).
- Copeland, R. A. The drug–target residence time model: A 10-year retrospective. *Nat. Rev. Drug Discov.* **15**, 87–95 (2016).
- Hyre, D. E. *et al.* Cooperative hydrogen bond interactions in the streptavidin–biotin system. *Protein Sci. A Publ. Protein Soc.* **15**, 459–467 (2006).
- Maschera, B. *et al.* Human immunodeficiency virus: Mutations in the viral protease that confer resistance to saquinavir increase the dissociation rate constant of the protease–saquinavir complex. *J. Biol. Chem.* **271**, 33231–33235 (1996).
- Basavapathruni, A. *et al.* Conformational adaptation drives potent, selective and durable inhibition of the human protein methyltransferase DOT1L. *Chem. Biol. Drug Des.* **80**, 971–980 (2012).
- Zheng, X., Bi, C., Li, Z., Podariu, M. & Hage, D. S. Analytical methods for kinetic studies of biological interactions: A review. *J. Pharm. Biomed. Anal.* **113**, 163–180 (2015).

8. Limongelli, V. Ligand binding free energy and kinetics calculation in 2020. *Wiley Interdiscipl. Rev. Comput. Mol. Sci.* **10**, e1455 (2020).
9. Cournia, Z., Allen, B. & Sherman, W. Relative binding free energy calculations in drug discovery: Recent advances and practical considerations. *J. Chem. Inf. Model.* **57**, 2911–2937 (2017).
10. Steinbrecher, T. & Labahn, A. Towards accurate free energy calculations in ligand protein-binding studies. *Curr. Med. Chem.* **17**, 767–785 (2010).
11. King, E., Aitchison, E., Li, H. & Luo, R. Recent developments in free energy calculations for drug discovery. *Front. Mol. Biosci.* **8**, 775 (2021).
12. Bruce, N. J., Ganotra, G. K., Richter, S. & Wade, R. C. Kbox: A toolbox of computational methods for studying the kinetics of molecular binding. *J. Chem. Inf. Model.* **59**, 3630–3634 (2019).
13. Rosenbaum, D. M., Rasmussen, S. G. F. & Kobilka, B. K. The structure and function of G-protein-coupled receptors. *Nature* **459**, 356–363 (2009).
14. Guo, D. *et al.* Molecular basis of ligand dissociation from the adenosine a2a receptor. *Mol. Pharmacol.* **89**, 485–491 (2016).
15. Davenport, A. P., Scully, C. C., de Graaf, C., Brown, A. J. & Maguire, J. J. Advances in therapeutic peptides targeting g protein-coupled receptors. *Nat. Rev. Drug Discov.* **19**, 389–413 (2020).
16. Obot, D. N. *et al.* Advances in the molecular understanding of G protein-coupled receptors and their future therapeutic opportunities. *Future J. Pharm. Sci.* **7**, 194 (2021).
17. Basith, S. *et al.* Exploring G protein-coupled receptors (GPCRs) ligand space via cheminformatics approaches: Impact on rational drug design. *Front. Pharmacol.* **9**, 128 (2018).
18. Wacker, D., Stevens, R. C. & Roth, B. L. How ligands illuminate GPCR molecular pharmacology. *Cell* **170**, 414–427 (2017).
19. Tang, X.-L., Wang, Y., Li, D.-L., Luo, J. & Liu, M.-Y. Orphan G protein-coupled receptors (gpcrs): Biological functions and potential drug targets. *Acta Pharmacol. Sin.* **33**, 363–371 (2012).
20. Hothersall, J. D., Brown, A. J., Dale, I. & Rawlins, P. Can residence time offer a useful strategy to target agonist drugs for sustained GPCR responses?. *Drug Discov. Today* **21**, 90–96 (2016).
21. Guo, D., Hillger, J. M., Ijzerman, A. P. & Heitman, L. H. Drug-target residence time—a case for G protein-coupled receptors. *Med. Res. Rev.* **34**, 856–892 (2014).
22. Cao, A. M. *et al.* Allosteric modulators enhance agonist efficacy by increasing the residence time of a gpcr in the active state. *Nat. Commun.* **12**, 1–13 (2021).
23. Potterton, A., Heifetz, A. & Townsend-Nicholson, A. Predicting residence time of gpcr ligands with machine learning. In *Artificial Intelligence in Drug Design* 191–205 (Springer, 2022).
24. Gessi, S., Merighi, S., Varani, K. & Borea, P. A. Adenosine receptors in health and disease. *Adv. Pharmacol.* **61**, 41–75 (2011).
25. Carpenter, B. & Lebon, G. Human adenosine A2A receptor: Molecular mechanism of ligand binding and activation. *Front. Pharmacol.* **2017**, 8 (2017).
26. Bell, G. I. Models for the specific adhesion of cells to cells: A theoretical framework for adhesion mediated by reversible bonds between cell surface molecules. *Science* **200**, 618–627 (1978).
27. Evans, E. & Ritchie, K. Dynamic strength of molecular adhesion bonds. *Biophys. J.* **72**, 1541–1555 (1997).
28. Evans, E. A. & Calderwood, D. A. Forces and bond dynamics in cell adhesion. *Science* **316**, 1148–1153 (2007).
29. Walton, E. B., Lee, S. & Van Vliet, K. J. Extending Bell's model: How force transducer stiffness alters measured unbinding forces and kinetics of molecular complexes. *Biophys. J.* **94**, 2621–2630 (2008).
30. Xu, L. C. & Siedlecki, C. A. Atomic force microscopy studies of the initial interactions between fibrinogen and surfaces. *Langmuir* **25**, 3675–3681 (2009).
31. Kumari, R., Kumar, R., Consortium, O. S. D. D. & Lynn, A. g\_mmpbsa a gromacs tool for high-throughput mm-pbsa calculations. *J. Chem. Inf. Model.* **54**, 1951–1962 (2014).
32. McNeely, P. M., Naranjo, A. N., Forsten-Williams, K. & Robinson, A. S. A2ar binding kinetics in the ligand depletion regime. *SLAS Discov.: Adv. Life Sci. R&D* **22**, 166–175 (2017).
33. Potterton, A. *et al.* Ensemble-based steered molecular dynamics predicts relative residence time of A<sub>2a</sub> receptor binders. *J. Chem. Theory Comput.* **15**, 3316–3330 (2019).
34. Jaakola, V. P. *et al.* The 2.6 angstrom crystal structure of a human a2a adenosine receptor bound to an antagonist. *Science* **322**, 1211–1217 (2008).
35. Lebon, G. *et al.* Agonist-bound adenosine a 2a receptor structures reveal common features of gpcr activation. *Nature* **474**, 521–525 (2011).
36. Kim, S. K. *et al.* Modeling the adenosine receptors: Comparison of the binding domains of a2a agonists and antagonists. *J. Med. Chem.* **46**, 4847–4859 (2003).
37. Dal Ben, D., Lambertucci, C., Marucci, G., Volpini, R. & Cristalli, G. Adenosine receptor modeling: What does the a2a crystal structure tell us?. *Curr. Top. Med. Chem.* **10**, 993–1018 (2010).
38. Kim, J., Wess, J., van Rhee, A. M., Schöneberg, T. & Jacobson, K. A. Site-directed mutagenesis identifies residues involved in ligand recognition in the human a2a adenosine receptor. *J. Biol. Chem.* **270**, 13987–13997 (1995).
39. Kim, J. *et al.* Glutamate residues in the second extracellular loop of the human a2a adenosine receptor are required for ligand recognition. *Mol. Pharmacol.* **49**, 683 (1996).
40. Jiang, Q., Lee, B. X., Glashofer, M., van Rhee, A. M. & Jacobson, K. A. Mutagenesis reveals structure-activity parallels between human a2a adenosine receptors and biogenic amine g protein-coupled receptors. *J. Med. Chem.* **40**, 2588–2595 (1997).
41. Schwarzschild, M. A., Agnati, L., Fuxe, K., Chen, J. F. & Morelli, M. Targeting adenosine a2a receptors in parkinson's disease. *Trends Neurosci.* **29**, 647–654 (2006).
42. Schapira, A. H. *et al.* Novel pharmacological targets for the treatment of parkinson's disease. *Nat. Rev. Drug Discov.* **5**, 845–854 (2006).
43. Blum, D., Hourez, R., Galas, M.-C., Popoli, P. & Schiffmann, S. N. Adenosine receptors and huntington's disease: Implications for pathogenesis and therapeutics. *Lancet Neurol.* **2**, 366–374 (2003).
44. Brown, R., Spina, D. & Page, C. Adenosine receptors and asthma. *Br. J. Pharmacol.* **153**, S446–S456 (2008).
45. During, M. J. & Spencer, D. D. Adenosine: A potential mediator of seizure arrest and postictal refractoriness. *Ann. Neurol.* **32**, 618–624 (1992).
46. Sawynok, J. & Liu, X. J. Adenosine in the spinal cord and periphery: Release and regulation of pain. *Prog. Neurobiol.* **69**, 313–340 (2003).
47. Benarroch, E. E. Adenosine and its receptors: Multiple modulatory functions and potential therapeutic targets for neurologic disease. *Neurology* **70**, 231–236 (2008).
48. De-Zwart, M. *et al.* Potent antagonists for the human adenosine a2b receptor derivatives of the triazolotriazine adenosine receptor antagonist zm241385 with high affinity. *Drug Dev. Res.* **48**, 95–103 (1999).
49. Braka, A., Garnier, N., Bonnet, P. & Aci-Seche, S. Residence time prediction of type 1 and 2 kinase inhibitors from unbinding simulations. *J. Chem. Inf. Model.* **60**, 342–348 (2019).
50. Salo-Ahen, O. M. H. *et al.* Molecular dynamics simulations in drug discovery and pharmaceutical development. *Processes* **9**, 71 (2020).

51. Akhunzada, M. J. *et al.* The role of Tat peptide self-aggregation in membrane pore stabilization: Insights from a computational study. *Phys. Chem. Chem. Phys.* **19**, 27603–27610 (2017).
52. Spiriti, J. & Wong, C. F. Qualitative prediction of ligand dissociation kinetics from focal adhesion kinase using steered molecular dynamics. *Life* **11**, 74 (2021).
53. Wong, C. F. Steered molecular dynamics simulations for uncovering the molecular mechanisms of drug dissociation and for drug screening: A test on the focal adhesion kinase. *J. Comput. Chem.* **39**, 1307–1318 (2018).
54. Deb, I. & Frank, A. T. Accelerating rare dissociative processes in biomolecules using selectively scaled md simulations. *J. Chem. Theory Comput.* **15**, 5817–5828 (2019).
55. Kokh, D. B. *et al.* Estimation of drug–target residence times by  $\tau$ -random acceleration molecular dynamics simulations. *J. Chem. Theory Comput.* **14**, 3859–3869 (2018).
56. Kokh, D. B., Kaufmann, T., Kister, B. & Wade, R. C. Machine learning analysis of  $\tau$ RAMD trajectories to decipher molecular determinants of drug–target residence times. *Front. Mol. Biosci.* **6**, 36 (2019).
57. Kokh, D. B. *et al.* A workflow for exploring ligand dissociation from a macromolecule: Efficient random acceleration molecular dynamics simulation and interaction fingerprint analysis of ligand trajectories. *J. Chem. Phys.* **153**, 125102 (2020).
58. Kokh, D. B. & Wade, R. C. G protein-coupled receptor–ligand dissociation rates and mechanisms from  $\tau$ ramd simulations. *J. Chem. Theory Comput.* **17**, 6610–6623 (2021).
59. 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.* **54**, 470–480 (2014).
60. Halder, S. *et al.* A multiscale simulation approach to modeling drug–protein binding kinetics. *J. Chem. Theory Comput.* **14**, 6093–6101 (2018).
61. Homeyer, N. & Gohlke, H. Extension of the free energy workflow FEW towards implicit solvent/implicit membrane MM–PBSA calculations. *Biochim. Biophys. Acta Gen. Subj.* **1850**, 972–982 (2015).
62. Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to super-computers. *SoftwareX* **1**, 19–25 (2015).
63. Pronk, S. *et al.* Gromacs 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).
64. Best, R. B. *et al.* Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone  $\phi$ ,  $\psi$  and side-chain  $\chi$ 1 and  $\chi$ 2 dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
65. Klauda, J. B. *et al.* Update of the charmm all-atom additive force field for lipids: Validation on six lipid types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).
66. Vanommeslaeghe, K. *et al.* Charmm general force field: A force field for drug-like molecules compatible with the charmm all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690 (2010).
67. Lee, J. *et al.* Charmm-gui input generator for namd, gromacs, amber, openmm, and charmm/openmm simulations using the charmm36 additive force field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).
68. Jo, S., Kim, T., Iyer, V. G. & Im, W. Charmm-gui: A web-based graphical user interface for charmm. *J. Comput. Chem.* **29**, 1859–1865 (2008).
69. Wu, E. L. *et al.* Charmm-gui membrane builder toward realistic biological membrane simulations. *J. Comput. Chem.* **35**, 1997–2004 (2014).
70. Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: Resources for positioning of proteins in membranes. *Nucleic Acids Res.* **40**, D370–D376 (2012).
71. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
72. Berendsen, H. J., Postma, J. V., Van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).
73. Hoover, W. G., Ladd, A. J. & Moran, B. High-strain-rate plastic flow studied via nonequilibrium molecular dynamics. *Phys. Rev. Lett.* **48**, 1818 (1982).
74. Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268 (1984).
75. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
76. Grubmüller, H., Heller, H., Windemuth, A. & Schulten, K. Generalized verlet algorithm for efficient molecular dynamics simulations with long-range interactions. *Mol. Simul.* **6**, 121–142 (1991).
77. Darden, T., York, D. & Pedersen, L. Particle mesh ewald: An  $n \log(n)$  method for ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
78. Hess, B., Bekker, H., Berendsen, H. J. & Fraaije, J. G. Lincs: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).
79. Davison, A. C., Hinkley, D. V. & Young, G. A. Recent developments in bootstrap methodology. *Stat. Sci.* **18**, 141–157 (2003).

### Author contributions

S.W. conceived the study. M.J.A., H.J.Y., I.D., A.B. and S.W. designed the calculations and interpreted data. M.J.A. and H.J.Y. performed calculations. M.J.A., H.J.Y. and I.D. wrote the manuscript. All authors reviewed and approved the final version of the paper.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-20065-2>.

**Correspondence** and requests for materials should be addressed to S.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022