



HAL
open science

Estimation of Drug-Target Residence Time by Targeted Molecular Dynamics Simulations

Sonia Ziada, Julien Diharce, Eric Raimbaud, Samia Aci-Sèche, Pierre Ducrot,
Pascal Bonnet

► **To cite this version:**

Sonia Ziada, Julien Diharce, Eric Raimbaud, Samia Aci-Sèche, Pierre Ducrot, et al.. Estimation of Drug-Target Residence Time by Targeted Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling*, 2022, 10.1021/acs.jcim.2c00852 . hal-03855904

HAL Id: hal-03855904

<https://univ-orleans.hal.science/hal-03855904v1>

Submitted on 27 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Estimation of drug-target residence time by targeted molecular dynamics simulations

Sonia Ziada¹, Julien Diharce^{1,3}, Eric Rimbaud², Samia Aci-Sèche^{1}, Pierre Ducrot², Pascal
Bonnet^{1*}*

¹Institut de Chimie Organique et Analytique (ICOA), UMR CNRS-Université d'Orléans 7311,
Université d'Orléans BP 6759, 45067 Orléans Cedex 2, France.

²Institut de Recherches Servier, 125 Chemin de Ronde, 78290 Croissy-sur-Seine, France.

³Université Paris Cité and Université des Antilles, Inserm, UMR_S 1134 BGR, 8 rue Maria Helena
Vieira da Silva, 75014 Paris, France

*Authors to whom all correspondence should be addressed:

Samia Aci-Sèche : Tel : +33 238 419 902, E-mail: samia.aci-seche@univ-orleans.fr

Pascal Bonnet: Tel: +33 238 417 254, E-mail: pascal.bonnet@univ-orleans.fr

KEYWORDS: Binding kinetics · Residence time · Molecular dynamics simulation · Drug design ·
Structure-kinetics relationship · Cyclin-dependent kinase 8.

ABSTRACT

Drug-target residence time has emerged as a key selection factor in drug discovery since the binding duration of a drug molecule to its protein target could significantly impact its in vivo efficacy. The challenge in studying the residence time, in early drug discovery stages, lies in how to cost-effectively determine the residence time for systematic assessment of compounds. Today a lack of computational protocols to quickly estimate such a measure still remains, particularly for large and flexible protein target and drugs. Here, we report an efficient computational protocol, based on targeted molecular dynamics, to rank drug candidates from their residence time and to obtain insights into ligand-target dissociation mechanisms. The method was assessed on a dataset of 10 arylpyrazole inhibitors of CDK8, a large, flexible and clinically important target, for which experimental residence time of the inhibitors ranges from minutes to hours. The compounds were correctly ranked according to their estimated residence time scores compared to their experimental values. The analysis of protein-ligand interactions along the dissociation trajectories highlighted the favorable contribution of hydrophobic contacts to residence time and revealed key residues that strongly affect compound residence time.

INTRODUCTION

Drug discovery and development program is a multi-stage complex process. A large proportion of drug candidates fails in the late phase of clinical trials due to a lack of efficacy[1], while those compounds appeared promising in the early stages of the drug discovery programs. The efficacy is the maximum response that a drug can produce in vivo. In order for a drug to have an effect, it needs to bind to its target. Therefore, increasing the target occupancy will increase the efficacy of the drug. One of the parameters that can influence the target occupancy is the drug-target binding kinetic[2]. This last ten years, drug-target binding kinetics is increasingly considered as an important selection criterion in drug discovery, in addition to the traditional focus on drug target-binding affinity[3,4].

Since Swinney et al. suggested that kinetics may provide crucial information[5], the residence time has emerged as an important criterion to evaluate the in vivo efficacy in the early phases of drug discovery 6,7. The positive impact of increasing the residence time on the in vivo efficacy has been demonstrated on multiples targets as G-protein-coupled receptors (GPCRs) 8,9, Histamine H1 Receptor (H1R) 10 or protein belonging to kinase family[11,12,13]. Studies on other therapeutic target proteins show a correlation of the residence time with in vivo efficacy[14,15,16].

Despite the great improvements in experimental methods and the large available panel assay, time-dependent essays needed for the measurement of binding kinetics remain more challenging than the well-established methods used to measure the affinity parameter[17]. Moreover, these experimental methods do not allow the correlation of kinetics data with structural interactions, that is the description of the full (un)binding process at the atomic level including the high-energy transition states and the stability of ground (or metastable) states. Such information would be of great importance to help chemists in the design and synthesis of compounds with optimized kinetics parameters.

In this context, molecular dynamics (MD) simulation is an interesting method to study binding kinetics due to the comprehensive structural view at the atomic level provided by the (un)binding process[18]. The simulation of ligand unbinding with methodologies based on brute-force MD were only achieved on small targets with relatively rigid binding sites and small molecules[19,20,21,22,23,24,25,26]. However, predominant therapeutic targets in drug discovery are larger and more flexible proteins, such as kinases or membrane proteins (GPCR), and often bind large and flexible drugs. Despite progress in computational power, the brute force MD based methods remain computationally too expensive regarding the timescale from milliseconds to hours needed to simulate many clinically relevant ligand unbinding processes. Hence, classical simulations are unsuitable for a routine industrial use where series of compounds must be analyzed during the hit-to-lead and the lead optimization stages. With this rising need in mind, several protocols using biased sampling methods have been developed and applied to compute binding kinetics.

Mollica et al. applied scaled-MD, a simple ranking method that does not require the definition of a reaction coordinate, on several ligands of HSP90, of Glucose-Regulated Protein (Grp78), of adenosine A2A receptor (A2A)[27] and of glucokinase[28], and obtained a correct k_{off} ranking in all cases. In another study, random acceleration molecular dynamics (RAMD) in combination with steered molecular dynamics (SMD) were used to explore ligand (un)binding pathways and to generate potentials of mean force respectively[29]. The difference in transition state barrier calculated from the potentials of mean force is in qualitative agreement with the measured difference in binding kinetics. An RAMD based method called τ RAMD, was applied on 70 diverse drug-like ligands of HSP90 α protein and demonstrated a good correlation ($R^2 = 0.86$) between computed and measured residence time for 78% of the compounds[30]. Very recently, this method was used to characterize the unbinding process for a set of benzene and indole derivatives complexed with T4 lysozyme mutants and provides very good agreement between computed and

experimental residence times[31]. Several approaches derived from metadynamics (MTD) have been developed for the prediction of binding kinetics constants[32,33,34,35,36,37]. Among them, through microsecond time scale simulations, results in good agreement with experiment were obtained for the k_{off} calculation of an urea-based allosteric inhibitor of p38 MAP kinase[38] and the dasatinib of c-Src kinase[37]. Callegari et al. proposed an alternative MTD approach for the ranking of a set of cyclin-dependent kinase 8 (CDK8) inhibitors by their k_{off} in agreement with experiment[33]. In this approach, the relative k_{off} values were estimated from the simulation time required for driving the ligands to the point of dissociation using simulations of dozen nanoseconds. However, in that study, the cyclin C, which is complexed to the kinase and available in the experimental structures[39], was not kept in the simulations, whereas a very recently published study[40], showed that cyclin C is vital for maintaining the structure of CDK8 and providing proper interactions for ligand binding.

In this study, we developed a protocol involving an ensemble of targeted molecular dynamics simulations to allow the ranking of a series of congeneric compounds by their residence time. With a relatively low computational cost, this method is suitable for an industrial use, to analyze chemical series during hit-to-lead or lead optimization stages. The protocol does not require a priori knowledge or hypothesis on the exit pathway and applies sufficiently slow variation to provide a realistic description of the unbinding process. The method was validated on a set of arylpyrazole inhibitors of CDK8[39], a clinically relevant target[41], for which experimental residence time of the inhibitors ranges from minutes to hours. We obtained a good ranking of the compounds according to their computed residence time score in agreement with their experimental residence time. Moreover, we carried out a Structure-Kinetics Relationship (SKR) study to depict the molecular determinants of protein-ligand interaction responsible of a slow dissociation. We hope that it will help optimizing kinetics profile of CDK8-CycC inhibitors.

MATERIEL AND METHODS

Data

We used the data published by Schneider et al.[39], containing a set of 10 arylpyrazole inhibitors of CDK8 complexed with the cyclin C (CDK8-CycC). Experimental residence times of these inhibitors are provided and the crystallographic structures of some of them complexed with CDK8-CycC are also available. We classify the inhibitors into three groups according to their residence time: Short Residence Time (SRT), Medium Residence Time (MRT) and Long Residence Time (LRT) (Figure 1). These compounds display a common pyrazol-5-yl urea scaffold. The SRT inhibitors 1-5 (RT < 1.4 min) comprise an inhibitor without a pendant chain (1), an inhibitor with a methyl derivative (2) and 3 other inhibitors with hydrophilic substituents (hydroxyethyl for 3, morpholinoethyl for 4 and hydroxypropyl for 5). The MRT group contains 3 inhibitors 6, 7, 8 which carry an hydroxybutyl, a morpholinopropyl and a hydroxypentyl chain with residence times of 7, 14 and 57 min, respectively. The third group includes the LRT inhibitors 9 and 10, which have a 1- (2-(4-(3-tert-butyl-1-p-tolyl-1H-pyrazol-5-ylcarbamoyl)- piperazin-1-yl)ethyl) or tert-butoxycarbonylaminopropyl chain and show long residence times of 1626 and 1944 min, respectively.

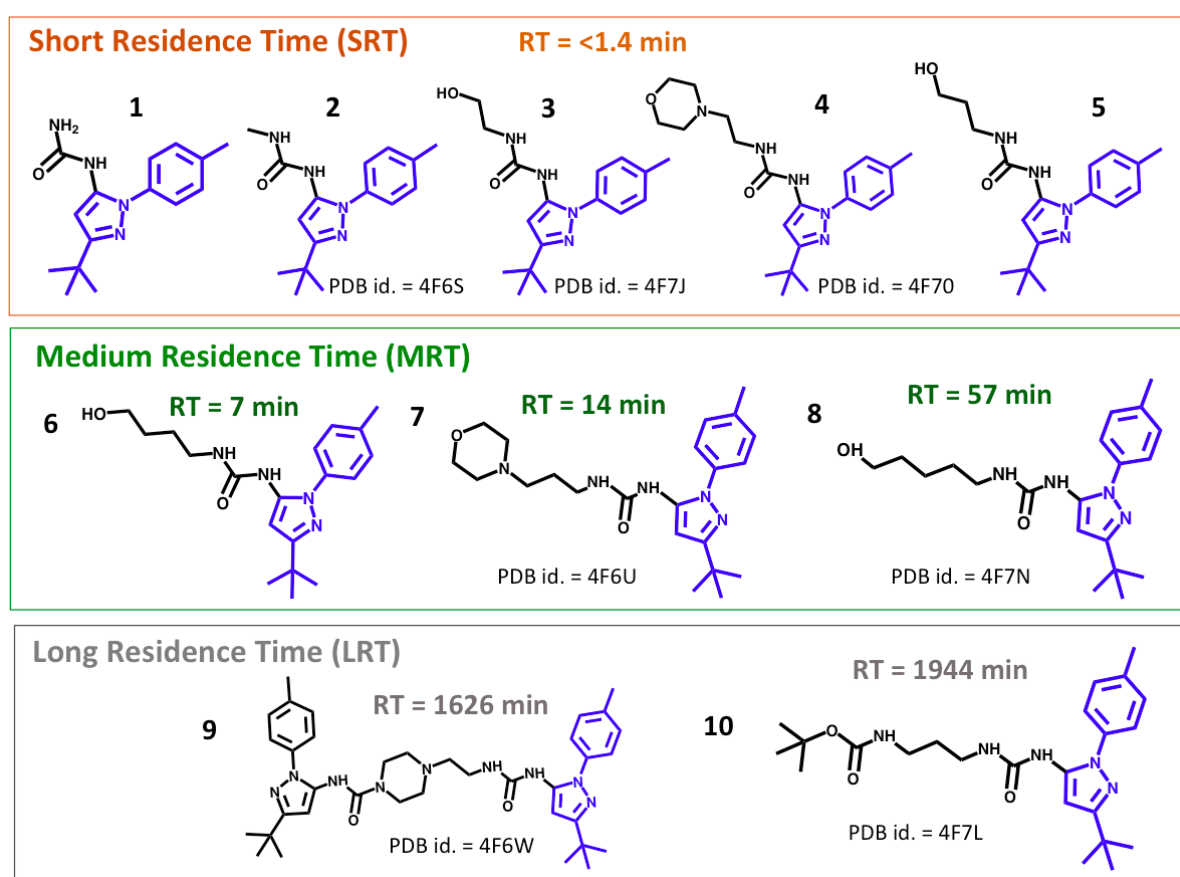


Figure 1. Chemical structures of the ten CDK8 inhibitors, numbered from 1 to 10 (inhibitor id.) and their experimental residence times classified in three groups: SRT MRT and LRT inhibitors. The common 1-(3- tert-butyl-1-p-tolyl-1H-pyrazol-5-yl)urea scaffold is highlighted in blue.

The common 1-(3-tert-butyl-1-p-tolyl-1H-pyrazol-5-yl)urea scaffold of these compounds is anchored in the kinase allosteric pocket (also called hydrophobic pocket) and interacts with the conserved DMG motif through a hydrogen bond (HB) interaction with the backbone nitrogen atom of Asp173 and with the carboxylate group of Glu66 through two HB interactions. The scaffold extends with variable functional groups toward the hinge region or the front pocket (Figure 2).

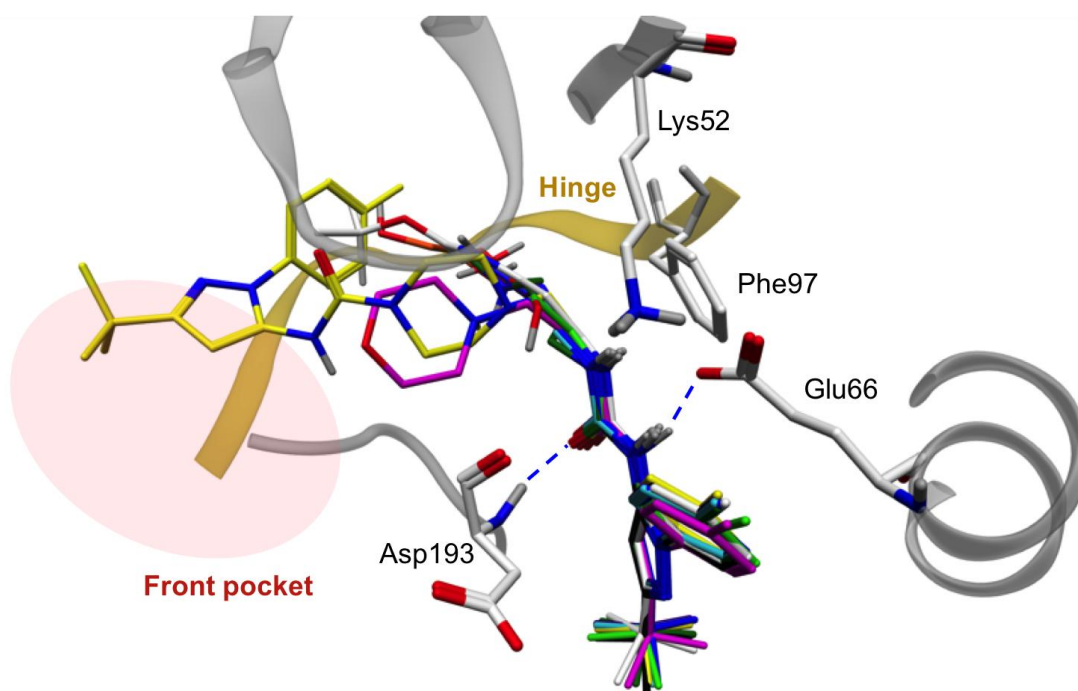


Figure 2. Binding site of the crystal structure of human CDK8 in complex with inhibitors 1-10.

Interactions between the urea of the common scaffold of the inhibitors and the residues Glu66 and Asp173 are shown in blue dashed lines. The protein is represented in white cartoon except the hinge colored in yellow cartoon. A red circle represents the location of the front pocket. Inhibitors 1 to 10 are colored respectively in gray, black, cyan, light green, pink, purple, dark green, blue, yellow and white.

Model Building

Similarly to Mollica *et al.*[28], we make the choice to construct a unique model of CDK8-CycC protein–ligand complex by homology modeling. The other protein-ligand systems were obtained by replacing the inhibitors in that model (chemical replacement), so as to keep the same protocol for inhibitors with no crystallographic structure. Chemical replacement was considered sufficient because the inhibitors of this congeneric series having a crystallographic structure consistently display a conserved orientation within the binding site (Figure 2). The first step thus consists in choosing the crystallographic structure that will be used as template for the homology model

building. All the available experimental structures[39] present 3 missing loops: the activation loop following the key DMG-motif (residues 177 to 193 in the structure of PDB ID 4F6U) which is in DMG-*out* conformation in all available structures and the loops from residues 116 to 120 and residues 240 to 244. After checking the absence of mutations in its sequence, the structure of PDB ID 4F6U was chosen as template structure to construct the model since it presents the best resolution ($R = 2.1 \text{ \AA}$) among the available ones. We have then aligned the UniProt canonical sequence of CDK8 on the PDB database to retrieve homologous structures having the missing regions resolved and the activation loop in the *out* conformation. The sequences alignment was performed with Clustal Omega[42] with a particular care on the alignment of domain kinase conserved motifs. Two crystallographic structures of the human CDK6, (PDB ID 1BI8 and 1G3N) were retained and used as template structures. CDK6 shares 37 % of identity and 63 % of similarity with CDK8 (Figure S1 & S2). Only the missing regions in the target structure were rebuilt to keep the coordinates of the protein solved parts unchanged. MODELLER version 9.16[43] was used to generate a model of CDK8 (residues 1 to 359) complexed to cyclin C (residues 1 to 264) and to compound 8, taking into account the crystallographic molecules of water. The missing C-terminal segments of CDK8 (residues 360 to 464, no significant role or any structuration known) and of cyclin C protein (residues 265 to 283) were not reconstructed. The structure of the complete model was validated using PROCHECK[44] and ProSA-web tools[45] (Figure S3). The CDK8CycC-inhibitor complexes having an available crystallographic structure, namely CDK8CycC complexed with compounds 2, 4, 5, 9-11, were treated as follow. First, the crystallographic structure is aligned to the model. Then, the crystallographic ligand and molecules of water are placed inside the model binding site, and the protein crystallographic structure is deleted. Some residues have been manually rotated to be in agreement with protein-ligand interactions observed in the respective crystallographic structures. The complexes CDK8CycC-inhibitor of compounds 1, 3, 6, 7 were

generated from previous reconstructed models, by manually modifying the chemical structure of the most similar analogues. These manipulations were done with the Molecular Operating Environment (MOE) software version 2016.0802 from the Chemical Computing Group. We last ensure that there is no steric hindrance in our models.

System Preparation

In total, 10 systems (Figure 1) were prepared. The AmberTools 15 suite (Case et al., 2015) was mainly employed for protonation, solvation, neutralization and generation of the systems topology and coordinate files. Regarding the ligands, they were prepared by using the Antechamber tool and parametrized with the GAFF force field. Hydrogen atoms are added with the reduce utility 46,47. All compounds were modeled in their neutral state, except for compound 8 (Figure S4), since the pKa of alkylmorpholines is about 7.4 48. Since the interaction of compound 8 with Ala100 is no longer observed with the protonated morpholine along a previous 1 μ s MD simulation (Figure S4), the morpholine part was finally modeled in its unprotonated state. Partial charges of the ligands were generated with the AM1/BCC method 49. Concerning the protein residues, PROPKA version 3.0 26 was used to check the protonation state of ionizable residue side-chains at pH = 7 and the parameters were assigned using the force field ff14SB 50. The system is then solvated in a rectangular TIP3P water box so that edges of the box are at least 10 Å distant from any solute atom. Finally, Cl⁻ counterions were added to neutralize the positively charged system resulting in a total number of around 110 000 atoms for the whole system.

Simulation protocols

For each of the 10 systems, 11 replicas of simulation were achieved. As the ligand may form slightly different interactions within the protein binding site, each replica was prepared with a new cycle of minimization and equilibration to better reflect this reality. Indeed, slightly different protein-ligand interaction network can be observed among replicas. A four-cycle minimization was

performed with 2000 steps each, minimizing first the solvent, second the residue side-chains, then the solute and finally the whole system. The SHAKE algorithm was applied to constrain bonds involving hydrogen atoms by using a time increment of 2 fs. Temperature regulation at 300K was ensured through Langevin dynamics with a collision frequency of 2 ps⁻¹. The long-range electrostatic interactions were computed by the particle mesh Ewald method beyond a distance of 10Å. The system was slowly heated in NVT ensemble from 0 to 300 K over a period of 50 ps, with a harmonic restraint on the solute (20 kcal.mol⁻¹.Å⁻² force-field constant) to prevent structural distortions. The system was then equilibrated during 10 ns MD simulation in the NPT ensemble at 300K and 1 atm, during which the harmonic restraint is gradually decreased from 20 kcal.mol⁻¹.Å⁻² to 3 kcal.mol⁻¹.Å⁻² in 1.3 ns and then, totally relaxed in 8.7 ns. The pressure relaxation time was set to 1 ps. Brute force MD calculations were performed using the PMEMD.cuda module of the AMBER14 program 51.

Targeted molecular dynamics (TMD)

The TMD is a simulation technique which aims to determine the transition pathway of between two different conformational states: (un)bound, (un)folded, open/close conformation etc.[52]. It consists in constraining the root mean square deviation (RMSD) between the current structure (which is the starting structure at the beginning of the simulation) and a reference structure (RMSD_{current}) to a user-defined value, namely the RMSD_{target}. This value of RMSD_{target} is slowly varied from an initial value to a targeted final value (RMSD_{target_final}), which results in the simulation of the process leading to the final desired state. In AMBER14 program, a harmonic restraining potential (V_{restraint}) is added to the force field, to help the RMSD_{current} reaching the successive values of RMSD_{target} until the final value (RMSD_{target_final}).

$$V_{\text{restraint}} = \frac{1}{2} \times f \times N_{\text{atoms}} \times (\text{RMSD}_{\text{current}} - \text{RMSD}_{\text{target}})^2 \quad \text{Equation 1}$$

Where f is the harmonic force constant, N_{atoms} is the number of restrained atoms, that is, the number of atoms on which the RMSD is calculated. Note that the atomic coordinates are mass weighted in the calculation of RMSD. It exists two approaches of TMD: direct TMD and reverse TMD (TMD^{-1}). In direct TMD, the reference structure corresponds to the final targeted structure, so that the value of $\text{RMSD}_{\text{target}}$ is decreased from the RMSD between the initial and target structure to a value close to 0. In TMD^{-1} , the reference structure corresponds to the initial structure, so that the RMSD value is increased from 0 to a predefined value. Then, the purpose of such simulations is to deviate far from the initial conformation, with no *a priori* information on the searching direction. Therefore, TMD^{-1} is less constraining than direct TMD, since no information on the final desired state is given.

In this study, we apply TMD^{-1} using the equilibrated CDK8CycC-inhibitor complex as reference structure. Therefore, as the $\text{RMSD}_{\text{target}}$ increases, the ligand moves away from the binding site, aided by the addition of the restraining potential ($V_{\text{restraint}}$). The more difficult the sampling at the considered $\text{RMSD}_{\text{target}}$, the higher the energy ($V_{\text{restraint}}$) added. We have verified the stability of the systems, and the overall structure of the protein during our unbinding simulations (see Section S3). The RMSD is calculated on the heavy atoms of the ligand, after aligning the reference and the current structure on the backbone of a set of 22 residues belonging to the binding site. These 22 residues are at a distance of 4 Å from the center of mass of the binding site (Section S2). Since the ligands of our series do not have the same number of heavy atoms, N_{atoms} varies and so the spring constant $f \times N_{\text{atoms}}$ can also vary. To ensure comparability between the results of the diverse ligands, the spring constant $f \times N_{\text{atoms}}$ was kept constant by adapting the value of f . After several tests, the value of $f \times N_{\text{atoms}}$ was fixed to 80 kcal.mol⁻¹. The $\text{RMSD}_{\text{target}}$ is changed by step of 0.01 Å every 0.2 ps from the value of 0.001 to 75.001 Å during a total simulation time of 1500 ps. Hence, at the beginning of the simulation, the ligand is in the binding site and its conformation and position

are sampled at a $\text{RMSD}_{\text{target}}$ value of 0.001 \AA , namely the bound state. After 0.2 ps of sampling at this value, $\text{RMSD}_{\text{target}}$ increases to a value of 0.011 \AA and the ligand is asked to sample at this new value during 0.2 ps etc., until it exits from the protein (the unbound state is defined in the Results and Discussion section). A snapshot was saved every 0.2 ps . TMD runs were performed with the parallelized version of the SANDER module from the AMBER14 program.

Data analysis

The simulations were analyzed using VDM 53, the CPPTRAJ module from AMBER14 program 51 and the Structure Interaction Diagram (SID) module of the Maestro suite (Maestro, Schrödinger, LLC, New York, NY, 2016). The SID module was used to calculate the protein-ligand interactions after converting the trajectories to the maestro format (see Section S6). For each simulation, we analyzed the protein-ligand interactions of a total of 150 conformations (one snapshot every 10 ps). All the data extracted from the simulations were processed and analyzed using the R package version 3.4.1 (R Core Team, 2017). Details about cut-off for residues contacts analysis and definition of interactions considered are given in the Supporting information (See Section S5).

RESULTS AND DISCUSSION

Residence time (RT) and RT_{score}

The residence time is directly related to the energy difference between unbinding transition state energy and the energy of the thermodynamically stable bound state of a drug, as described by the Arrhenius law (section S4). Ligands displaying long residence time are assumed having to cross higher or multiple energy barriers and so need more energy to be expelled from their binding sites. On this basis, the restraining potential added during a TMD^{-1} simulation to cross energy barriers and reach the unbound state could be considered as a relevant estimator of the residence time of a drug. However, in view of the nature of the bias (harmonic restraining potential) which forces the system to sample every micro-state along the pathway, the increase of the restraining potential is not only

associated to the action of “pushing” the ligand to help it crossing high energy barriers, but also to the action of “retaining” the ligand, when the system is on an energetic descent approaching a metastable state. In that connection, we derived a new function $V_{\text{restraint_push}}$ from $V_{\text{restraint}}$ than only encompasses the restraining potential added to cross energy barriers and so, traduces the difficulty encountered by the ligand to escape the binding site. $V_{\text{restraint_push}}$ is defined as:

$$V_{\text{restraint_push}} = V_{\text{restraint}} \times f(\text{RMSD}_{\text{current}}, \text{RMSD}_{\text{target}}) \quad \text{Equation 2}$$

- Where $f(\text{RMSD}_{\text{current}}, \text{RMSD}_{\text{target}}) = 1$, when $\text{RMSD}_{\text{current}} - \text{RMSD}_{\text{target}} < 0$
- And $f(\text{RMSD}_{\text{current}}, \text{RMSD}_{\text{target}}) = 0$, when $\text{RMSD}_{\text{current}} - \text{RMSD}_{\text{target}} > 0$

Accordingly, when the ligand has to overcome an energetic peak, the value of $\text{RMSD}_{\text{current}}$ becomes inferior to the value of the wanted $\text{RMSD}_{\text{target}}$ due to the difficulty to move forward and in this case $V_{\text{restraint_push}} = V_{\text{restraint}}$. On the contrary, when the ligand is about to reach an energetic minimum, the ligand advances faster, so the value of $\text{RMSD}_{\text{current}}$ becomes significantly superior to the value of the wanted $\text{RMSD}_{\text{target}}$. In this case, the increase in $V_{\text{restraint}}$ is not considered, so $V_{\text{restraint_push}} = 0$. Integrating $V_{\text{restraint_push}}$ over time leads to a quantity (noted RTscore) comparable to an action in physics expressed in [energy].[time] which reflects here the overall “action” supplied to pull the ligand out of the binding site.

$$\text{RTscore} = \int_{t_0}^{\text{texit_time}} V_{\text{restraint_push}} = \frac{1}{2} \times f \times \text{Natoms} \int_{t_0}^{\text{texit_time}} (\text{RMSD}_{\text{current}} - \text{RMSD}_{\text{target}})^2 \quad \text{Equation 3}$$

The integral is calculated from the beginning of the simulation ($t = 0$ ns) to the exit time ($t_{\text{exit_time}}$) defined as the time at which the ligand is unbound. Mollica *et al.*[27] used geometrical criteria to select the situation where interactions between the ligand and its target were negligible, defined as a distance between the ligand-site centers of mass of 30Å for the HSP90 and Grp78 systems and of 25 Å for A2A. In this work, we also used a geometric criterion to identify the unbound state by considering the state in which the ligand and its target are separated by 2 layers of water. That

approximately corresponds to the system in which any atoms of the ligand are at least at 6 Å of any atoms of the protein (here the complex CDK8-CycC). We assume that ligands having to cross higher or multiple energy barriers need more energy ($V_{\text{restraint_push}}$) from the TMD⁻¹ protocol to be expelled from the binding site, and so, a higher value of RT_{score} is expected. Therefore, RT_{score} is assumed to be positively correlated to the RT (Section S4). As a ligand could exit through different unbinding paths during TMD⁻¹ simulations, the values of RT_{score} were collected and averaged from the 11 replicated simulations, for each ligand. In Table 1, the average RT_{score} values for compounds 1 to 10 and their standard errors are reported. The Figure 3 shows that the method is able to correctly rank the SRT, MRT and LRT inhibitors according to their residence time on the basis of RT_{score} . The method clearly distinguishes LRT inhibitors from SRT and MRT inhibitors while MRT inhibitors are less well separated from SRT inhibitors in light of the error bars and can be then used for the qualitative prediction of the residence time. Moreover, the time consumed to produce such simulation, and such results, is very reasonable. Indeed, we can treat 5 ligands in one day, only using 11 computers with 8 CPUs inside (1 per replica). The mean time for one simulation is about 5 hours. Clearly, it can be employed for a routine utilization in an industrial context.

Inhibitor id.	RT (minute)	RT_{score} (kcal.mol⁻¹.ps)
1	<1.4	1048 ± 66
2	<1.4	1562 ± 243
3	<1.4	1560 ± 206
4	<1.4	1610 ± 217
5	<1.4	1580 ± 205
6	7	1991 ± 298
7	14	1933 ± 245
8	57	2014 ± 428
9	1626	4494 ± 419
10	1944	3990 ± 289

Table 1. Experimental Residence Times (RT) and values of the estimator of the residence time (RT_{score}) calculated from the simulations for CDK8 inhibitors 1-10. RT_{score} values correspond to the means calculated on the 11 replicas, with their standard errors.

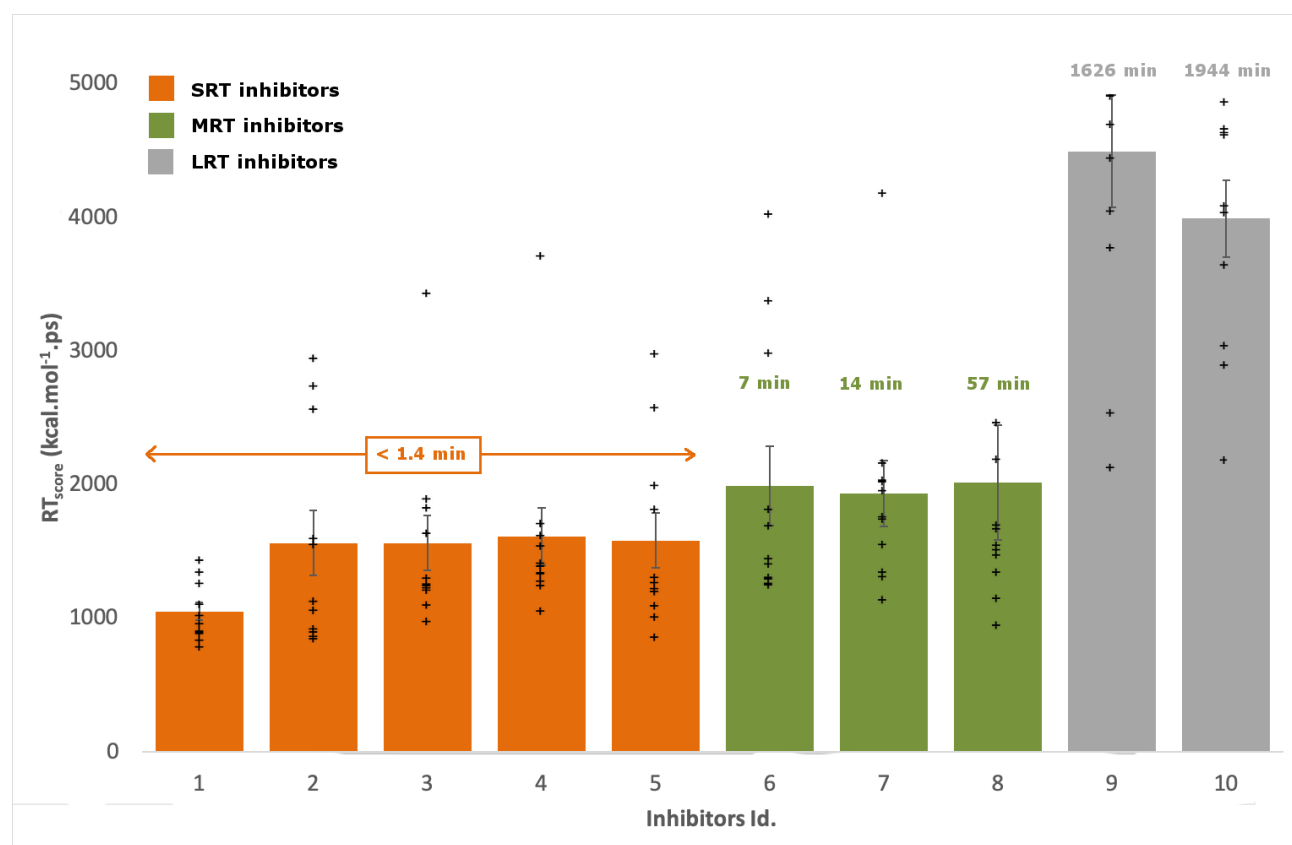


Figure 3. Estimator of the residence time (RT_{score}) calculated from the simulations for CDK8 inhibitors 1-10. The bars are colored according to experimental residence time groups (SRT, MRT and LRT). The error bars are based on the standard error. The RT_{score} values obtained for each separate replica are reported as crosses.

The difficulty in predicting residence time is to set up a method that is applicable to both a simple 1-step kinetics process and a complex process involving several kinetics steps. The association of a drug to its target, and the subsequent dissociation of the drug–target complex, are often controlled by conformational changes, especially involving structural changes in the immediate vicinity of the drug-binding pocket[54]. However, the transition state theory that relates the activation energy to

the kinetics constant (k_{off} , k_{on}) is based on the Arrhenius equation that describes a 1-step kinetics process. In a complex multi-steps kinetics process, given the fact that energy barriers can be multiple, more or less important and depend on the degree of resolution of the energy profile (free energy, restraint energy or other) calculated by numerical methods, determining the k_{off} is complex and challenging. Spirit *et al* combined umbrella sampling free energy simulations and SMD to study the dissociation of 14 ligands from focal adhesion kinase (FAK)[55]. Although they concluded that free energy simulation provided too low barriers to be consistent with the experimental dissociation rates for 3 ligands, they show that a qualitative classification of the ligands is possible by simply using the SMD exit time, related to the SMD force used. In a similar manner than our, this qualitative evaluation is based on the difficulty encountered by the ligand during its exit process. Bortolato and co-workers proposed an approach derived from MTD simulation that only takes into account the first barrier of the bias potential energy. They calculate a kind of RT score defined as the maximum bias potential energy required to move the ligand from the starting energy basin (bound state) to the next[32]. Another approach (Tiwary's approach), also derived from MTD simulation, assumes that the energetic landscape underlying the dissociation process presents few high and sharp barriers resulting thus, in a global Arrhenius behavior, ie a 1-step kinetics process[56,37]. Sun *et al.* tested Bortolato's and Tiwary's approach and were not able to accurately predict the k_{off} of a complex dissociation process involving EGFR kinase protein that presents a deep binding site[35]. Callegari *et al.* applied Tiwary's approach and failed also in predicting the k_{off} . They concluded that the free energy landscape of the CDK8- arylpyrazole inhibitor unbinding process is too complex[33]. The advantage of RT_{score} is that it encompasses all the energy barriers encountered during the dissociation process, which makes the method applicable on complex multistep kinetics processes (Equation 3).

Analysis of the Structure-Kinetics Relationship (SKR)

Pathways

Visual inspection of all the unbinding trajectories revealed three pathways taken by the 10 inhibitors: the “allosteric channel” when the ligand exits through the allosteric pocket (also called hydrophobic pocket), the “ATP channel” when it passes through the front pocket and, a last one when it passes under the hinge, that we called “hinge channel” (Figure 4).

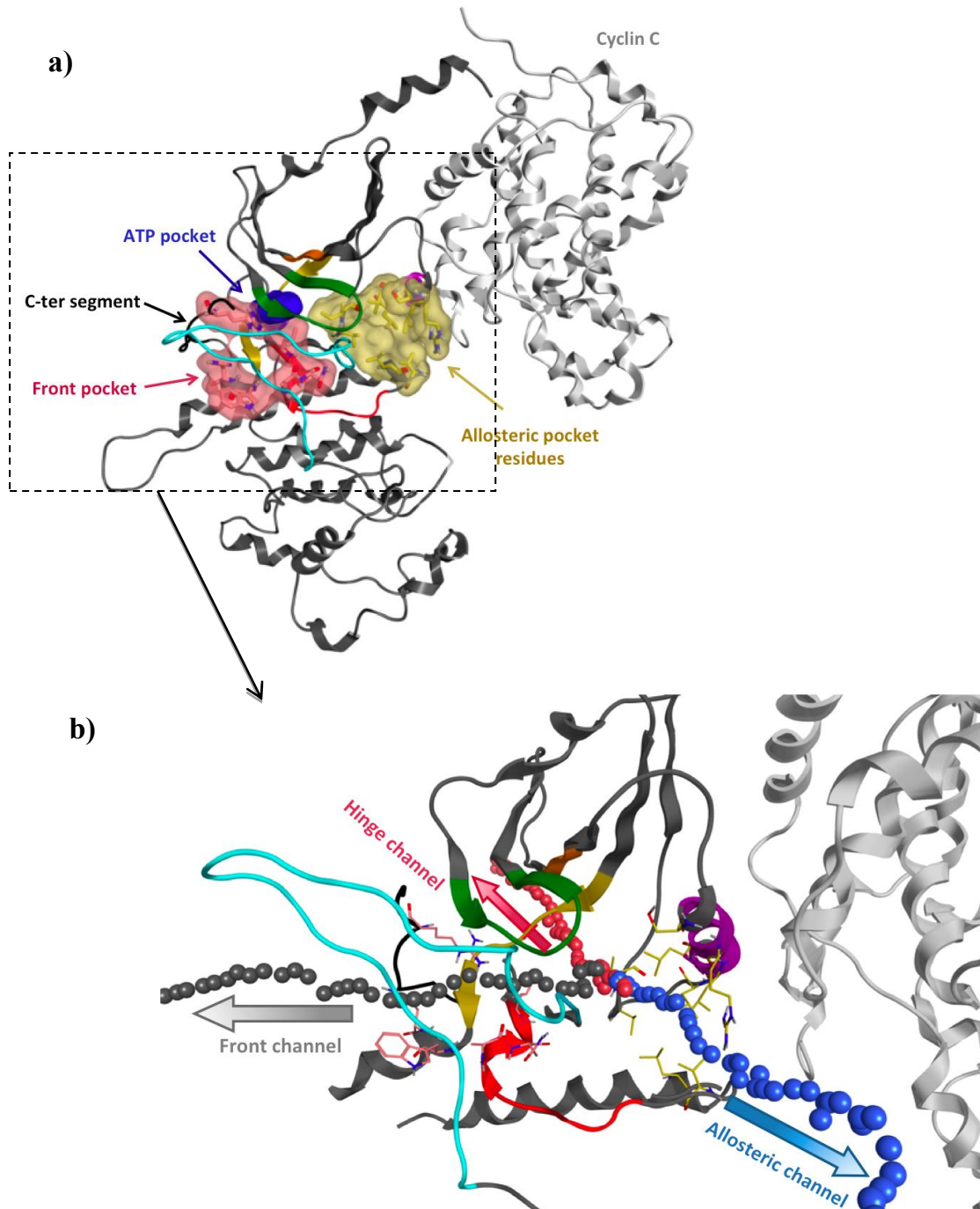


Figure 4. Exit pathways taken by the inhibitors. (a) Representation of CDK8-CycC in cartoons. The cyclin C is in light gray. The kinase domain is in dark gray except the kinase conserved motifs, the C-ter segment (black) and the activation loop (cyan). The conserved motifs of the kinase domain are colored as follows: α -C helix (purple), hinge (yellow), P-loop (dark green), Hyd1 (orange), HRD (red). The activation loop (containing the DMG motif) is coloured in cyan. The allosteric pocket residues, the ATP pocket and the front pocket residues are displayed in surface coloured in yellow, dark blue and pink, respectively. (b) Zoom of picture (a) where an example of each pathway was represented by balls corresponding to the centers of mass of the ligand along the path. The allosteric, ATP and hinge channels were represented in blue, black and red, respectively. The residues of the allosteric pocket and the front pocket are represented in sticks with their carbons colored in yellow and pink, respectively.

LRT inhibitors (9 & 10) take mainly the ATP channel whereas MRT (6-8) and SRT (1-5) inhibitors follow mostly the allosteric channel (Figure 5). The allosteric and the ATP channel have already been observed in several computational studies as possible routes in kinase family [29,57,58]. A type II inhibitor of p38 MAP kinase called BIRB796, having the same scaffold (1-(3-tert-butyl-1-p-tolyl-1H-pyrazol-5-yl)urea) as our series and the same binding mode as MRT inhibitors, has also been shown to exit preferentially through the allosteric channel [58], which is in agreement with our results. Inhibitors exiting through the allosteric channel go either toward the cyclin C partner, interacting with its surface in some cases, or go towards the P-loop or the C-lobe of the kinase. For the ATP channel, the routes also diversify when approaching the solvent, which can be attributed to the movements of the long flexible C-terminal segment and of the activation loop that directly impact the ease of access to the solvent (Figure 4). Regarding the hinge channel, it appears that it has been sampled only 3 times, for compound 4, 8 and 10. The rareness of this exit, and the fact that it is not exclusive to one kind of inhibitors, shed light on its non-importance in the

CDK8-ligand unbinding process. It could also be associated to a simulation artifact, and may not currently exist as a viable biological solution.

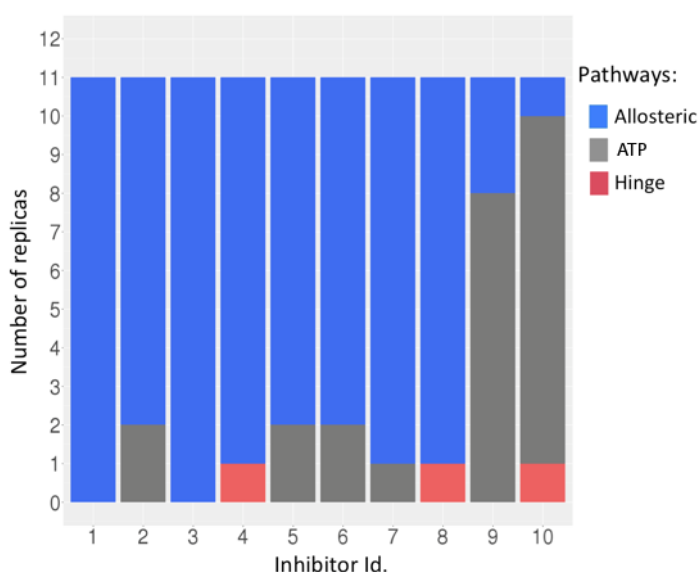


Figure 5. Barplot counting the number of each type of pathway taken by each inhibitor.

We then analyzed protein-ligand interactions along the unbinding process to depict the structure-kinetics relationship. Schneider *et al.* have described the differences in protein-ligand interactions between the inhibitors on the basis of their static crystallographic structures and related them to their residence times. They suggest that hydrogen bonds with the hinge region are indispensable to provide a detectable residence time to compounds leading to MRT inhibitors, whereas large hydrophobic complementarities within the front pocket (Figure 6) significantly optimize the compound residence time leading to LRT inhibitors 39. However, a statistical analysis could be insufficient sometimes to fully describe an unbinding process, which is a dynamical phenomenon by nature.

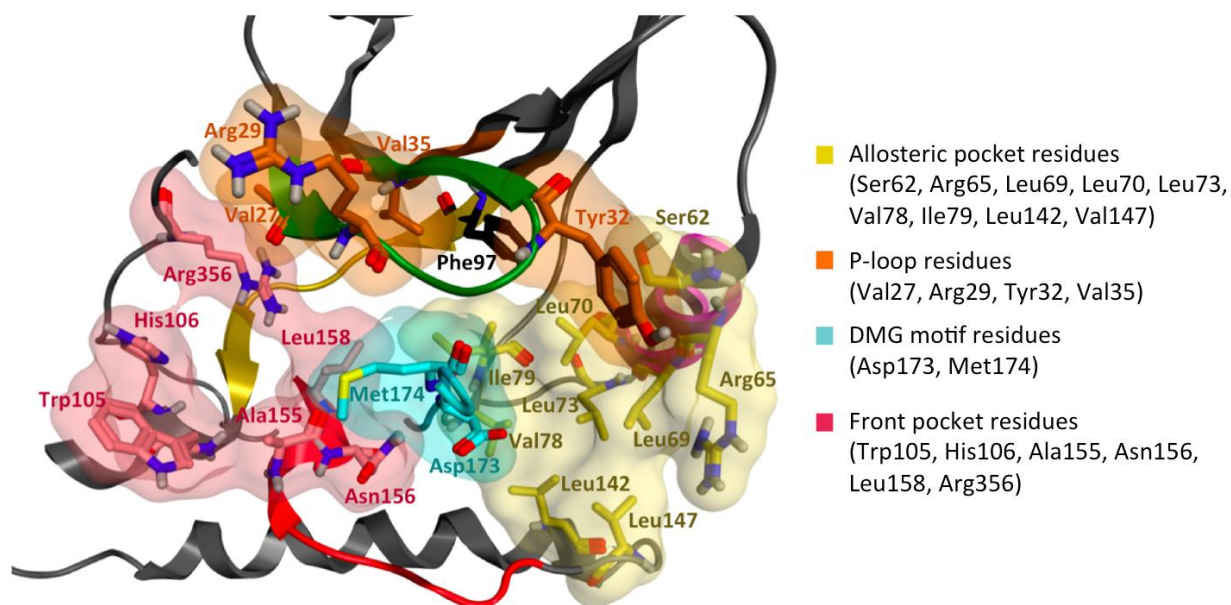


Figure 6. Representation of the regions of the binding site containing some residues that influence the residence time. Nomenclature adapted from (Schneider et al., 2013).

The kinase domain is represented in ribbon coloured in dark gray except the kinase conserved motifs and the C-ter segment (black). The conserved motifs of the kinase domain are coloured as follows: α -C helix (purple), hinge (yellow), P-loop (dark green), Hyd1 (orange), HRD (red), and DMG motif (cyan). The allosteric pocket residues, the P-loop residues, the DMG motif residues and the front pocket residues are displayed in surface respectively in yellow, orange, cyan and pink. They are also represented in stick where the carbon atoms follow the same color code. The residue Phe97, called gatekeeper residue, is represented in stick, with the carbon atom coloured in black.

The MD brings the dynamical view of the phenomenon and, along with several replicas, provides a sampling of the unbinding event leading to a relevant observation. Therefore, in the following section, we analyzed protein-ligand interactions in the bound state and also, along the dissociation path by comparing replicas of a same inhibitor and the inhibitors with each other. The goal is to discuss and hypothesize on possible impact of some protein-ligand interactions on the residence time. Figure 6 will be our reference for the spatial location of residues interacting with the several exiting inhibitors.

SRT inhibitors

For the SRT inhibitors (1 - 5), replicas that establish hydrophobic interactions between the trimethyl group of the scaffold and the residues of the allosteric pocket (Figure 6, yellow surface), in particular Leu70, Leu73, Val78, Ile79, Leu142 and Val147, present higher RT_{score} values. In addition to these interactions, inhibitors 2 to 5 differ from the inhibitor 1 by their ability to establish hydrophobic interactions between the Phe97, known as the "gatekeeper" residue, and the hydrophobic part of their variable fragment (the methyl group for inhibitor 2, and the alkyl chain for the inhibitors 3 to 5). The variable fragment, which is the most flexible part of the inhibitors, is stabilized by the hydrophobic interactions with the gatekeeper Phe97. As a result, the inhibitors are maintained close to the hinge region at the beginning of the simulation, deep in the binding site. In this location, hydrophobic interactions between the allosteric pocket residues on one hand, and the trimethyl group of the scaffold as well as the hydrophobic segment of the variable fragment on the other hand, are promoted, and may be associated to a stable state. Replicas of SRT inhibitors 2 to 5 displaying the highest RT_{score} values always form such hydrophobic interactions involving the gatekeeper Phe97 and the allosteric pocket residues. Consequently, it results in a lower number of hydrophobic contacts for the inhibitor 1, that does not have a variable fragment, compared to other inhibitors (Figure 7, Figure S7). This could be a reason explaining the gap in average RT_{score} values between the inhibitor 1 and the inhibitors 2 to 5 (Figure 3). Callegari *et al.* also observed comparable results, that is a lower predicted residence time for inhibitor 1 compared to inhibitors 2 – 5³³. In the supporting information, we show the positive correlation between the number of hydrophobic contacts (involving the allosteric pocket residues and the Phe97) and the value of RT_{score} (Figure S7). Inhibitors 2, 3 and 5 do not interact with the hinge: inhibitor 2 is too short for such interaction and inhibitors 3 and 5 orient instead their hydroxyl group toward Met174 or Asp173 or Glu66. In addition to the described hydrophobic interactions, we observe for compound

4 (One replica only, with the highest RT_{score}) a HB interaction, involving bridging water molecule, between the oxygen of the morpholine and the Ala100 of the hinge region. However, this interaction is unstable because the inhibitor is not long enough to interact with the hinge region and to maintain, at the same time, the strong HB interactions with Glu66 and Asp173 involving the urea of the scaffold.

MRT inhibitors

From this analysis, we can easily imagine that a longer alkyl chain will stabilize the interaction with the hinge region and lead to an increased residence time. Such inhibitors correspond to the MRT inhibitors where stronger interactions with the hinge region residues are observed without destabilizing the HB interactions involving the scaffold. Indeed, the hydroxyl group of inhibitor 6 establishes stable water bridge interactions with Ala100 and Asp98, (inhibitor 6 with a $RT_{\text{experimental}}$ of 7min) and the hydroxyl group of inhibitor 8 ($RT_{\text{experimental}}$ of 57 min) forms a HB interaction with Asp98 and a water bridge with Ala100. The morpholine of inhibitor 7 interacts with the hinge through a HB interaction with Ala100 and a water bridge with Asp98. These interactions with the hinge region lead to higher average RT_{score} values for the MRT, compared to SRT as show in Figure 3. As suggested by Schneider *et al.*, the HB interactions with the hinge region seem to be indispensable to detect residence time ($RT_{\text{experimental}} > 1.4$ min). Among the replicas of the MRT inhibitors, higher values of RT_{score} are observed when those HB interactions are formed along with hydrophobic contacts involving the allosteric pocket residues and Phe97 (Figure S8 and S9).

Then, we investigated whether there is a relationship between the formed protein-ligand interactions and the taken pathway for SRT and MRT inhibitors, as some of them take the front or the hinge channel instead of the main allosteric channel (Figure 5). We found that all replicas taking the ATP channel display optimized hydrophobic interactions involving the gatekeeper Phe97 and the allosteric pocket residues. Those replicas also obtained higher RT_{score} for (Figure S8 and S9).

However, these interactions are also observed in some replicas exiting through the allosteric channel, meaning that they do not alone determine the taken pathway.

LRT inhibitors

Finally, we analyzed the interactions made by LRT inhibitors. The binding mode of LRT inhibitors differs from SRT and MRT inhibitors since the variable fragment is oriented toward the front pocket (Figure 2) and not the hinge. Schneider *et al.* stated that large hydrophobic complementarities within the front pocket significantly optimize the residence time leading to LRT inhibitors. These interactions are indeed observed, but they are not the ones with the greatest impact on the RT_{score} . Our analysis revealed that the replicas with the highest RT_{score} tend, at the beginning of the simulation, to establish more interactions with the DMG motif (Asp173, Met174) through H-bonds, and with the P-loop residues (Val27, Arg29, Val35) through hydrophobic and HB interactions. Those replicas exit through the ATP channel. On the contrary, the replicas with the lowest RT_{score} tend to present strong H-bonds with Glu66 (involving the urea of the scaffold) and less contacts with the DMG motif and the P-loop residues at the beginning of the simulation. Those replicas exit through the allosteric channel (Figure S10). We did not establish any relationship between the exit through the hinge pathway and the protein-ligand interactions at the beginning of the simulations.

For replicas exiting through the ATP channel, the rise of interactions with the P-loop and the DMG motif is accompanied by an increasing number of hydrophobic contacts with the front pocket residues. Then, the HB interactions between the urea of the scaffold and Glu66 are broken, which leads to increase hydrophobic contacts with the allosteric pocket residues. In this intermediate state, the two extremities of the inhibitor form hydrophobic contacts with the front pocket and the allosteric pocket, respectively. This intermediate state is not observed with SRT and MRT

inhibitors that exit through the ATP channel because the compound is not long enough to interact with the front pocket and the allosteric pocket residues at the same time.

In summary, our SKR analysis suggests that hydrophobic contacts with the allosteric pocket residues (Leu70, Leu73, Val78, Ile79, Leu142 and Val147) and the gatekeeper Phe97, in addition to the HB interactions with the hinge residues (Ala100 and Asp98) may widely contribute to increase the residence time of an inhibitor to medium values ($1.4 \text{ min} < RT_{\text{experimental}} < 57 \text{ min}$). For a LRT inhibitor ($> 57 \text{ min}$), the main positive contributions to residence time are brought by HB and hydrophobic interactions with P-loop residues (Val27, Arg29, Val35) and HB interactions with DMG motif residues (Asp173, Met174) and to a less extent the hydrophobic interactions with the front pocket residues. We observed conserved HB interactions involving the urea of scaffold and residues Glu66 and Asp173 independently of their RT range.

Importance of the hydrophobic contacts

From our SKR results, it appears that the hydrophobic interactions strongly contribute to slow the dissociation process. In line with these results, by analyzing the average total number of each contact type (HB, water bridge, ionic, cation-Pi and Pi-Pi stacking) for each inhibitor during the unbinding process (Figure 7, Figure S11), we found significantly more hydrophobic contacts with LRT inhibitors compared with SRT and MRT inhibitors. The strong impact of the hydrophobic contacts has already been discussed in the literature[59,60]. In that connection, Schmidtke *et al.* demonstrated in their study that the formation of water-shielded hydrogen bonds between a ligand and its receptor protein is a viable strategy to increase the residence time. They showed that the kinetics stability provided by hydrogen bonds depends on their degree of solvent exposure[61]. Gao *et al.*, quantified this effect and revealed that hydrogen bonds can be up to $1.2 \text{ kcal.mol}^{-1}$ stronger in hydrophobic environments[62]. Therefore, we can hypothesize that the high number of hydrophobic contacts involving i) the allosteric pocket residues in one side, and ii) those mediated

by the P-loop and the front pocket residues in the other side, increase the residence time by reducing the solvent exposure of the inhibitor respectively in the allosteric channel side and ATP channel side. In this context, the desolvation energy barrier, required for the ligand to exit the binding site and get solvated, is higher, contributing thus to increase the residence time. Consequently, HB interactions formed with the kinase binding site are stronger and participate to the high residence-time in a synergic way.

HB interactions with bridging water molecules are the most predominant contact in all cases, but it is something quite predictable because of the nature of the phenomenon observed, *i.e.* a dissociation process. Consequently, the high number of bridging water molecules is associated to the exit process, and because the solvent accessibility increases dramatically during the process. We observe that those interactions haven't any direct consequences of the residence time, because of the high volatility of the water molecules.

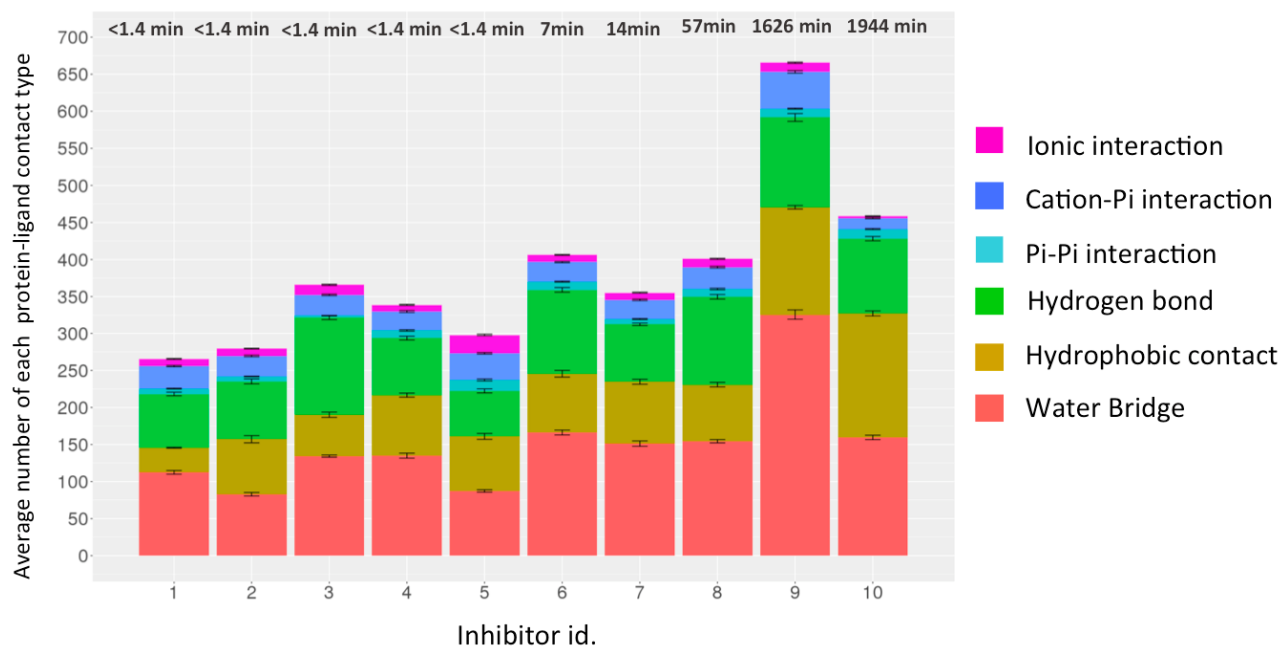


Figure 7. Stacked barplot of the average number of each protein-ligand contact type (pi-pi, HB, etc.) established during the simulation for each CDK8 inhibitors.

For each inhibitor and each replica, the total number of each protein-ligand contact type has been calculated along the simulation. This number is then averaged on the 11 replicas of each inhibitor and we calculated the error bars (based on the standard deviation).

Besides hydrophobic interactions, inhibitors also form ionic and cation- π interactions (Figure 7) mostly with Arg65 and Arg150, in addition to Lys52 and Arg356, through their interaction with the scaffold rings (imidazole and phenyl). Arg65 and Arg150 are two of the three conserved arginines of CDKs family (the third one is Arg178) and are located on both sides of the access gate of the allosteric channel belonging to the α -C helix and to the HRD motif, respectively (Figure 8). Arg150 and Arg65, when interacting with the ligand, tend to decrease the RT_{score} when the exit occurs through the allosteric channel. This effect is clearer with Arg150 than Arg65. Indeed, we systematically observe a low RT_{score} when this interaction is formed (Figure S12). For the interaction with Arg65, it is a bit more complex, because this residue can adopt two conformations: toward the allosteric pocket or toward the Glu99 of the CyclinC (Glu99^{CycC}) (Figure 8). When Arg65 is oriented toward the allosteric pocket the interaction between Arg65 and the inhibitor is relatively stable despite the restraining potential. However, as the inhibitor moves forward, toward the solvent in the allosteric channel, the Arg65 forms HB interaction with Glu99^{CycC}. In this position, Arg65 interacts alternatively with Glu99^{CycC} and with the inhibitor, which drives the ligand toward the cyclin C, so toward the solvent and facilitates the dissociation process. In the literature, Glu99^{CycC} was hypothesized to have an important role in the activation mechanism of CDK8. It has been suggested that Glu99^{CycC} mimics the absent phosphoresidue within CDK8 and interacts with the three conserved arginine residues (Arg65, Arg150, and Arg178) to adjust their orientation and induce an open conformation of the activation loop[63].

Anyway, interactions of the ligand with one of those Arg residues tends to be very favorable. Consequently, those arginine residues seem to possess a crucial role for an easier extraction of the ligand from the binding site, greatly diminishing the RT_{Score} .

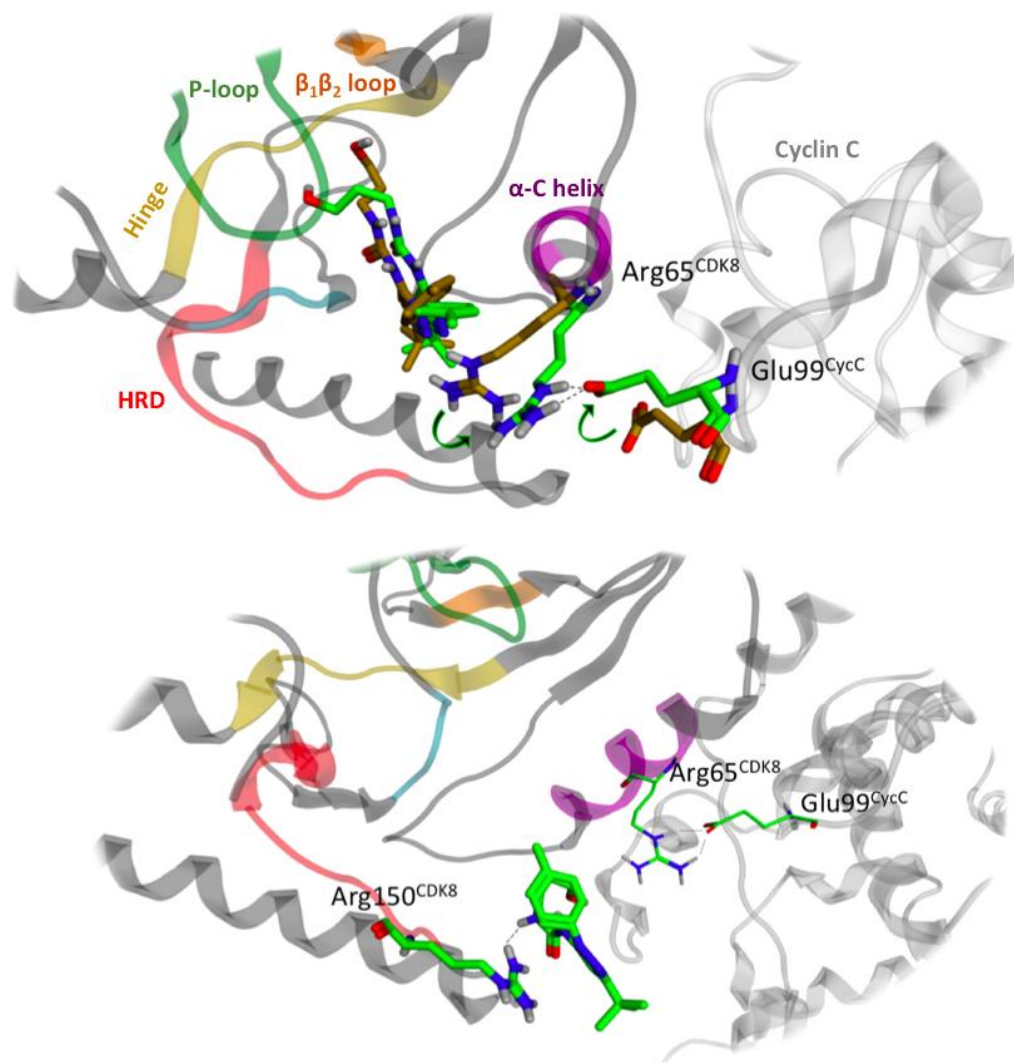


Figure 8. Arg65 and Arg 150, two CDK8 conserved residues that accelerate the dissociation process. (Top) The Arg65^{CDK8} can adopt two conformations: toward the allosteric pocket when it interacts with the ligand through cation-pi or ionic interactions mostly (represented in stick with carbon coloured in brown) or toward the cyclin C when it interacts with the Glu99^{CycC} through a HB interaction (represented in stick with carbon coloured in green stick). (Bottom) The ligand interacts with Arg150^{CDK8} through mostly ionic and cation-pi interaction.

In summary, these results suggest that the interaction of the ligand with Arg65 and Arg150, two residues positioned on both sides of the entrance of the allosteric channel, have a negative contribution to its residence time. In a more general view, this overall protein-ligand interaction study indirectly shows the importance of keeping cyclin C in the system to produce meaningful simulations. Indeed, we showed that the Glu99^{CycC} through its interaction with the Arg65 could have an impact on the residence time. Moreover, besides the role of cyclin C in the biological function of CDK8, a recent study, in good agreement with our results, demonstrates the crucial role of cyclin C in the dynamics and structure of CDK8, particularly in its fundamental role in providing proper interaction for ligand binding through its impact on α -C helix conformation[40]. Indeed, Cholko *et al.* observed that the α C helix of CDK8 adopts an α C-out conformation in the absence of cyclin C, whereby Glu66 moves away from the DMG motif. By losing the H-bond from Glu66, the allosteric binding site collapses, thereby disabling the binding of type-II ligands. Surprisingly, Callegari *et al.* still obtained a correct ranking of these SRT, MRT and LRT inhibitors according to their experimental residence time using a MTD approach while they do not keep the cyclin C in their simulations[33]. Note that the experimental data were measured in the presence of cyclin C 39. Moreover, we observed that some replicas, when they leave the binding site through allosteric channel, are not directly solvated but interact on the surface of the cyclin C. This phenomenon lengthens the exit time of these replicas (Section S6), exit time that is used to estimate the residence time in Callegari *et al.*[33].

CONCLUSION

Drug-target residence time is an important criterion in drug discovery programs to select and optimize lead candidates with improved *in vivo* efficacy, in addition to the traditional focus on drug target-binding affinity[3]. Despite the great improvements in experimental and computational methods combined with the availability of powerful computer resources, the determination of the

residence time is still a challenging task. We presented here a computational method using an ensemble of targeted molecular dynamics simulations to estimate the unbinding kinetics constant of protein-ligand complexes. The method was able to properly rank a set of arylpyrazole inhibitors of cyclin-dependent kinase 8 (CDK8-CycC) according to their experimental residence time ranging from <1.4min to 1944min. One of our major concerns was to develop a method with a relatively low computational cost to be suitable for an industrial use, where dozens of compounds must be prioritized in the hit-to-lead and the lead optimization phases. For a kinase system prepared as mentioned in the Material and Methods section, our method has a throughput of 5 ligands per day (11 replicas per ligand) using 11 computers of 8 cpu. Moreover, another advantage lays in its simplicity and in the fact that it does not require any specific *a priori* knowledge on the exit pathway. The used reaction coordinate (RMSD) has the advantage to induce soft changes since the ligand can increase its RMSD just by changing its conformation without moving ahead its mass center, which provides realistic description of the unbinding process. We subsequently focused on establishing structure–kinetics relationships, to identify the chemical features impacting the residence time. Our results highlight the importance of hydrophobic interactions with the allosteric pocket, the P-loop and the front pocket residues, and the HB interactions with the hinge and the DMG motif residues. This SKR study could be of valuable help in designing CDK8-CycC inhibitors with an optimized kinetics profile and thus an improved *in vivo* profile. All of that associated to a very high-performance methodology, compatible with industrial use.

DATA AND SOFTWARE AVAILABILITY

All PDB structures used to build the initial models of MD simulations were downloaded from the RCSB protein data bank (<https://www.rcsb.org>). Homology modeling were realized by using Modeller software (<https://salilab.org/modeller/>). The licensed Amber14 was used to perform MD simulation (<https://ambermd.org>). The publicly available AmberTools15 was used to analyze the

MD trajectories, in addition with VMD (<https://www.ks.uiuc.edu/Research/vmd/>), Schrödinger suite (<https://www.schrödinger.com>) and homemade programs in python jupyter-notebook. The notebooks are not available because of the industrial collaboration but the methodology is clearly described in the paper so that it can be reimplemented.

ASSOCIATED CONTENT

Supporting information. Details and additional figures on model building and system preparation, TMD setup, impact of the restraining potential on the stability of the system, relationship between RT_{score} and experimental RT, protein-ligand interactions analysis and structure-kinetics relationship (pdf).

AUTHOR INFORMATION

Corresponding Author

samia.aci-seche@univ-orleans.fr

pascal.bonnet@univ-orleans.fr

Funding Sources

This work was supported by the Institut de Recherche Servier.

ACKNOWLEDGMENT

The authors wish to thank the Orléans-Tours CaSciModOT at the Centre de Calcul Scientifique de la Région Centre Val de Loire and the Centre Régional Informatique et d'Applications Numériques de Normandie (CRIANN) for providing computer facilities, ChemAxon for providing academic license free of charge and also the projects CHemBio (FEDER-FSE 2014-2020-EX003677), Techsab (FEDER-FSE 2014-2020-EX011313), the RTR Motivhealth (2019-00131403) and the

Labex programs SYNORG (ANR-11-LABX-0029) and IRON (ANR-11-LABX-0018-01) for their financial support of ICOA, UMR 7311, University of Orléans, CNRS.

REFERENCES

- 1 Arrowsmith, J.; Miller, P. Trial Watch: Phase II and Phase III attrition rates 2011-2012. *Nature Reviews Drug Discovery* **2013**, *12* (8), 569.
- 2 de Witte, W. E.; Danhof, M.; Graaf, P. H.; Lange, E. C. In vivo Target Residence Time and Kinetics Selectivity: The Association Rate Constant as Determinant. *Trends in Pharmacological Sciences* **2016b**, *37*, 831-842.
- 3 Bernetti, M.; Cavalli, A.; Mollica, L. Protein-ligand (un)binding kinetics as a new paradigm for drug discovery at the crossroad between experiments and modelling. *MedChemComm* **2017**, *8*, 534-550.
- 4 Copeland, R. A. The dynamics of drug-target interactions: drug-target residence time and its impact on efficacy and safety. *Expert Opinion on Drug Discovery* **2010**, *5*, 305-310.
- 5 Swinney, D. C. Biochemical mechanisms of drug action: what does it take for success? *Nature Reviews Drug Discovery* **2004**, *3*, 801-808.
- 6 Schuetz, D. A.; de Witte, W. E.; Wong, Y. C.; Knasmueller, B.; Richter, L.; Kokh, D. B.; Sadiq, S. K.; Bosma, R.; Nederpelt, I.; Segala, E. Kinetics for Drug Discovery: an industry-driven effort to target drug residence time. *Drug Discovery Today* **2017**, *22* (6), 896-911.
- 7 Lee, K. S. S.; Yang, J.; Niu, J.; Ng, C. J.; Wagner, K. M.; Dong, H.; Kodani, S. D.; Wan, D.; Morisseau, C.; Hammock, B. D. Drug-Target Residence Time Affects in Vivo Target

Occupancy through Multiple Pathways. *ACS Central Science* **2019**, *5*, 1614-1624.

- 8 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. *British Journal of Pharmacology* **2012**, *166*, 1846-1859.
- 9 Doornbos, M. L.; Cid, J. M.; Haubrich, J.; Nunes, A.; Sande, J. W.; Vermond, S. C.; Mulder-Krieger, T.; Trabanco, A. A.; Ahnaou, A.; Drinkenburg, W. H. Discovery and Kinetics Profiling of 7-Aryl-1,2,4-triazolo[4,3-a]pyridines: Positive Allosteric Modulators of the Metabotropic Glutamate Receptor 2. *Journal of Medicinal Chemistry* **2017**, *60*, 6704-6720.
- 10 Bosma, R.; Witt, G.; Vaas, L. A.; Josimovic, I.; Gribbon, P.; Vischer, H. F.; Gul, S.; Leurs, R. The Target Residence Time of Antihistamines Determines Their Antagonism of the G Protein-Coupled Histamine H1 Receptor. *Frontiers in Pharmacology* **2017**, *8*, 667.
- 11 Wood, E. R.; Truesdale, A. T.; McDonald, O. B.; Yuan, D.; Hassell, A.; Dickerson, S. H.; Ellis, B.; Pennisi, C.; Horne, E.; Lackey, K. A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. *Cancer Research* **2004**, *64*, 6652-6659.
- 12 Puttini, M.; Redaelli, S.; Moretti, L.; Brussolo, S.; Gunby, R. H.; Mologni, L.; Marchesi, E.; Cleris, L.; Donella-Deana, A.; Drueckes, P. Characterization of compound 584, an Abl kinase inhibitor with lasting effects. *Haematologica* **2008**, *93*, 653-661.
- 13 Uitdehaag, J. C.; Man, J.; Willemsen-Seegers, N.; Prinsen, M. B.; Libouban, M. A.; Sterrenburg, J. G.; Wit, J. J.; Vetter, J. R.; Roos, J. A.; Buijsman, R. C. Target Residence Time-Guided Optimization on TTK Kinase Results in Inhibitors with Potent Anti-Proliferative

Activity. *Journal of Molecular Biology* **2017**, *429*, 2211-2230.

- 1 Costa, B.; Da Pozzo, E.; Giacomelli, C.; Barresi, E.; Taliani, S.; Da Settimo, F.; Martini, C.
4. TSPO ligand residence time: a new parameter to predict compound neurosteroidogenic efficacy. *Scientific Reports* **2016**, *6*, 18164.
- 1 Lee, K. S. S.; Liu, J. Y.; Wagner, K. M.; Pakhomova, S.; Dong, H.; Morisseau, C.; Fu, S. H.;
5. Yang, J.; Wang, P.; Ulu, A. Optimized inhibitors of soluble epoxide hydrolase improve in vitro target residence time and in vivo efficacy. *Journal of Medicinal Chemistry* **2014**, *57*, 7016-7030.
- 1 Ramos, I.; Aparici, M.; Letosa, M.; Puig, C.; GavaldÀ , A.; Huerta, J. M.; Espinosa, S.;
6. Vilella, D.; Miralpeix, M. Abediterol (LAS100977), an inhaled long-acting beta-2-adrenoceptor agonist, has a fast association rate and long residence time at receptor. *European Journal of Pharmacology* **2018**, *819*, 89-97.
- 1 de Witte, W. E.; Wong, Y. C.; Nederpelt, I.; Heitman, L. H.; Danhof, M.; Graaf, P. H.;
7. Gilissen, R. A.; Lange, E. C. Mechanistic models enable the rational use of in vitro drug-target binding kinetics for better drug effects in patients. *Expert Opinion on Drug Discovery* **2016a**, *11*, 45-63.
- 1 Nunes-Alves, A.; Kokh, D. B.; Wade, R. C. Recent progress in molecular simulation methods
8. for drug binding kinetics. *Current Opinion in Structural Biology* **2020**, 126–133.
- 1 Huang, D.; Caflisch, A. The free energy landscape of small molecule unbinding. *PLoS*
9. *Computational Biology* **2011**, *7*, e1002002.

- 2 Pan, A. C.; Xu, H.; Palpant, T.; Shaw, D. E. Quantitative Characterization of the Binding and
0. Unbinding of Millimolar Drug Fragments with Molecular Dynamics Simulations. *Journal of
Chemical Theory and Computation* **2017**, *13*, 3372-3377.
- 2 Dickson, A.; Lotz, S. D. Multiple Ligand Unbinding Pathways and Ligand-Induced
1. Destabilization Revealed by WExplore. *Biophysical Journal* **2017**, *112*, 620-629.
- 2 Teo, I.; Mayne, C. G.; Schulten, K.; Lelièvre, T. Adaptive Multilevel Splitting Method for
2. Molecular Dynamics Calculation of Benzamidine-Trypsin Dissociation Time. *Journal of
Chemical Theory and Computation* **2016**, *12*, 2983-2989.
- 2 Buch, I.; Giorgino, T.; De Fabritiis, G. Complete reconstruction of an enzyme-inhibitor
3. binding process by molecular dynamics simulations. *Proceedings of the National Academy of
Sciences* **2011**, *108*, 10184-10189.
- 2 Plattner, N.; Noé, F. Protein conformational plasticity and complex ligand-binding kinetics
4. explored by atomistic simulations and Markov models. *Nature Communications* **2015**, *6*, 7653.
- 2 Votapka, L. W.; Jagger, B. R.; Heyneman, A. L.; Amaro, R. E. SEEKR: Simulation Enabled
5. Estimation of Kinetics Rates, A Computational Tool to Estimate Molecular Kinetics and Its
Application to Trypsin-Benzamidine Binding. *The Journal of Physical Chemistry B* **2017**, *121*
(15), 3597-3606.
- 2 Olsson, M. H.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent
6. Treatment of Internal and Surface Residues in Empirical pKa Predictions. *Journal of Chemical
Theory and Computation* **2011**, *7*, 525-537.

- 2 Mollica, L.; Decherchi, S.; Zia, S. R.; Gaspari, R.; Cavalli, A.; Rocchia, W. Kinetics of
7. protein-ligand unbinding via smoothed potential molecular dynamics simulations. *Scientific Reports* **2015**, *5*, 11539.
- 2 Mollica, L.; Theret, I.; Antoine, M.; Perron-Sierra, F.; Charton, Y.; Fourquez, J. M.;
8. Wierzbicki, M.; Boutin, J. A.; Ferry, G.; Decherchi, S. Molecular Dynamics Simulations and Kinetics Measurements to Estimate and Predict Protein-Ligand Residence Times. *Journal of Medicinal Chemistry* **2016**, *59*, 7167-7176.
- 2 Niu, Y.; Li, S.; Pan, D.; Liu, H.; Yao, X. Computational study on the unbinding pathways of
9. B-RAF inhibitors and its implication for the difference of residence time: insight from random acceleration and steered molecular dynamics simulations. *Physical Chemistry Chemical Physics* **2016**, *18*, 5622-5629.
- 3 Kokh, D. B.; Amaral, M.; Bomke, J.; Grädler, U.; Musil, D.; Buchstaller, H. P.; Dreyer, M.
0. K.; Frech, M.; Lowinski, M.; Vallee, F. Estimation of Drug-Target Residence Times by tau-Random Acceleration Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation* **2018**, *14*, 3859-3869.
- 3 Nunes-Alves, A.; Kokh, D. B.; Wade, R. C. Ligand unbinding mechanisms and kinetics for
1. T4 lysozyme mutants from τ RAMD simulations. *Current Research in Structural Biology* **2021**, 106–111.
- 3 Bortolato, A.; Deflorian, F.; Weiss, D. R.; Mason, J. S. Decoding the Role of Water Dynamics
2. in Ligand-Protein Unbinding: CRF1R as a Test Case. *Journal of Chemical Information and Modeling* **2015**, *55*, 1857-1866.

- 3 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. *Journal of Chemical Information and Modeling* **2017**, *57*, 159-169.
- 3 Pietrucci, F.; Marinelli, F.; Carloni, P.; Laio, A. Substrate Binding Mechanism of HIV-1 Protease from Explicit-Solvent Atomistic Simulations. *Journal of the American Chemical Society* **2009**, *131*, 11811-11818.
- 3 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. *Journal of Chemical Information and Modeling* **2017**, *57*, 1895-1906.
- 3 Tiwary, P.; Limongelli, V.; Salvalaglio, M.; Parrinello, M. Kinetics of protein-ligand unbinding: Predicting pathways, rates, and rate-limiting steps. *Proceedings of the National Academy of Sciences* **2015**, *112*, E386-E391.
- 3 Tiwary, P.; Mondal, J.; Berne, B. J. How and when does an anticancer drug leave its binding site? *Science Advances* **2017**, *3*, e1700014.
- 3 Casasnovas, R.; Limongelli, V.; Tiwary, P.; Carloni, P.; Parrinello, M. Unbinding Kinetics of a p38 MAP Kinase Type II Inhibitor from Metadynamics Simulations. *Journal of the American Chemical Society* **2017**, *139*, 4780-4788.
- 3 Schneider, E. V.; Böttcher, J.; Huber, R.; Maskos, K.; Neumann, L. Structure-kinetics relationship study of CDK8/CycC specific compounds. *Proceedings of the National Academy of Sciences* **2013**, *110*, 8081-8086.

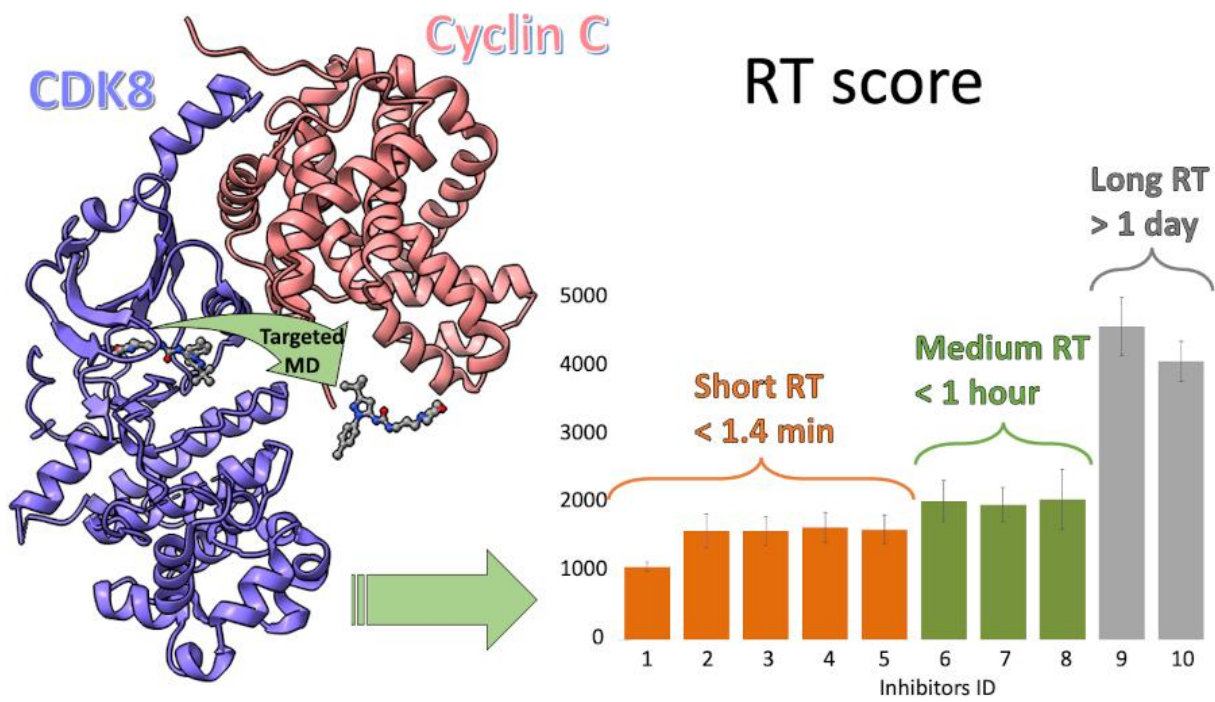
- 4 Cholko, T.; Chen, W.; Tang, Z.; Chang, C. E. A molecular dynamics investigation of
0. CDK8/CycC and ligand binding: conformational flexibility and implication in drug discovery.
Journal of Computer-Aided Molecular Design **2018**, *32*, 671-685.
- 4 Philip, S.; Kumarasiri, M.; Teo, T.; Yu, M.; Wang, S. Cyclin-Dependent Kinase 8: A New
1. Hope in Targeted Cancer Therapy? *Journal of Medicinal Chemistry* **2018**, *61*, 5073-5092.
- 4 Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam,
2. H.; Remmert, M.; Söding, J.; Thompson, J.; Higgins, D. Fast, scalable generation of high-
quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*
2011, *7*, 539.
- 4 Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints.
3. *Journal of Molecular Biology* **1993**, *234*, 779-815.
- 4 Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: a program
4. to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*
1993, *26*, 283-291.
- 4 Wiederstein, M.; Sippl, M. J. ProSA-web: interactive web service for the recognition of errors
5. in three-dimensional structures of proteins. *Nucleic Acids Research* **2007**, *35*, 407-410.
- 4 Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing
6. of a general amber force field. *Journal of Computational Chemistry* **2004**, *25*, 1157-1174.
- 4 Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic atom type and bond type
7. perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling*

2006, 25, 247-260.

- 4 Hall, H. K. Potentiometric Determination of the Base Strength of Amines in Non-protolytic Solvents. *The Journal of Physical Chemistry* **1956**, 60, 63-70.
- 4 Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II Parameterization and validation.. *Journal of Computational Chemistry* **2002**, 23, 1623-1641.
- 5 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. *Journal of Chemical Theory and Computation* **2015**, 11, 3696-3713.
- 5 Case, D.; Babin, V.; Berryman, J.; Betz, R.; Cai, Q.; Cerutti, D.; Cheatham, T.; Darden, T.; Duke, R.; Gohlke, H. *Amber 15*; University of California: San Francisco, **2015**.
- 5 Schlitter, J.; Engels, M.; Krüger, P. Targeted molecular dynamics: a new approach for searching pathways of conformational transitions. *Journal of Molecular Graphics* **1994**, 12, 84-89.
- 5 Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *Journal of Molecular Graphics* **1996**, 14, 33-38.
- 5 Copeland, R. A. Conformational adaptation in drug-target interactions and residence time. *Future Medicinal Chemistry* **2011**, 3, 1491-1501.
- 5 Spiriti, J.; Wong, C. F. Qualitative Prediction of Ligand Dissociation Kinetics from Focal Adhesion Kinase Using Steered Molecular Dynamics. *Life* **2021**, 11 (2), 74.

- 5 Tiwary, P.; Parrinello, M. From Metadynamics to Dynamics. *Physical Review Letters* **2013**,
6. *111*, 230602.
- 5 Patel, J. S.; Berteotti, A.; Ronsisvalle, S.; Rocchia, W.; Cavalli, A. Steered Molecular
7. Dynamics Simulations for Studying Protein-Ligand Interaction in Cyclin-Dependent Kinase 5.
Journal of Chemical Information and Modeling **2014**, *54*, 470-480.
- 5 Sun, H.; Tian, S.; Zhou, S.; Li, Y.; Li, D.; Xu, L.; Shen, M.; Pan, P.; Hou, T. Revealing the
8. favorable dissociation pathway of type II kinase inhibitors via enhanced sampling simulations
and two-end-state calculations. *Scientific Reports* **2015**, *5*, 8457.
- 5 Liu, L.; Michelsen, K.; Kitova, E. N.; Schnier, P. D.; Klassen, J. S. Evidence that Water Can
9. Reduce the Kinetics Stability of Protein-Hydrophobic Ligand Interactions. *Journal of the
American Chemical Society* **2010**, *132*, 17658-17660.
- 6 Schuetz, D. A.; Richter, L.; Amaral, M.; Grandits, M.; Grädler, U.; Musil, D.; Buchstaller, H.
0. P.; Eggenweiler, H. M.; Frech, M.; Ecker, G. F. Ligand Desolvation Steers On-Rate and
Impacts Drug Residence Time of Heat Shock Protein 90 (Hsp90) Inhibitors. *Journal of
Medicinal Chemistry* **2018**, *61* (10), 4397-4411.
- 6 Schmidtke, P.; Luque, F. J.; Murray, J. B.; Barril, X. Shielded Hydrogen Bonds as Structural
1. Determinants of Binding Kinetics: Application in Drug Design. *Journal of the American
Chemical Society* **2011**, *133*, 18903-18910.
- 6 Gao, J.; Bosco, D. A.; Powers, E. T.; Kelly, J. W. Localized Thermodynamic Coupling
2. between Hydrogen Bonding and Microenvironment Polarity Substantially Stabilizes Proteins.
Nature Structural & Molecular Biology **2009**, *16*, 684-690.

- 6 Hoepfner, S.; Baumli, S.; Cramer, P. Structure of the Mediator Subunit Cyclin C and its
3. Implications for CDK8 Function. *Journal of Molecular Biology* **2005**, *350*, 833-842.



For Table of Contents Only