Journal of Medicinal Chemistry

Article

Molecular dynamics simulations and kinetic measurements to estimate and predict protein-ligand residence times

luca mollica, Isabelle Theret, Mathias Antoine, Francoise Perron-Sierra, Yves Charton, Jean-Marie Fourquez, Michel Wierzbicki, Jean A. Boutin, Gilles Ferry, Sergio Decherchi, Giovanni Bottegoni, Pierre Ducrot, and Andrea Cavalli

J. Med. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.jmedchem.6b00632 • Publication Date (Web): 08 Jul 2016

Molecular dynamics simulations and kinetic measurements to estimate and predict protein-ligand residence times

Luca Mollica[1], Isabelle Theret[2], Mathias Antoine[2], Françoise Perron-Sierra[2], Yves Charton[2], Jean-Marie Fourquez[2], Michel Wierzbicki[2], Jean A. Boutin[2], Gilles Ferry[2], Sergio Decherchi[1,3], Giovanni Bottegoni[1,3], Pierre Ducrot[2], Andrea Cavalli*[1,4]*

[1]CompuNet, Istituto Italiano di Tecnologia, via Morego, 30, 16163 Genova, Italy

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

[3]BiKi Technologies S.r.l. via XX Settembre 33/10 16121 Genova, Italy

[4]Dept. of Pharmacy and Biotechnology, University of Bologna, via Belmeloro 6, 40126 Bologna, Italy

Abstract

Ligand-target residence time is emerging as a key drug discovery parameter because it can reliably predict drug efficacy in vivo. Experimental approaches to binding and unbinding kinetics are nowadays available, but we still lack reliable computational tools for predicting kinetics and residence time. Most attempts have been based on brute-force molecular dynamics (MD) simulations, which are CPU-demanding and not yet particularly accurate. We recently reported a new scaled-MD-based protocol, which showed potential for residence time prediction in drug discovery. Here, we further challenged our procedure's predictive ability by applying our methodology to a series of glucokinase inhibitors that could be useful for treating type 2 diabetes mellitus. We combined scaled MD with experimental kinetics measurements and X-ray crystallography, promptly checking the protocol's reliability by directly comparing computational predictions and experimental measures. The good agreement highlights the potential of our scaled-MD-based approach as an innovative method for computationally estimating and predicting drug residence times.

Introduction

Binding and unbinding kinetics are emerging as key factors for predicting drug efficacy in living organisms.[1,2] There are several experimental techniques (e.g. SPR, stopped-flow CD, fluorescence spectroscopy) for studying (un)binding kinetics, but currently no efficient computational approaches for predicting absolute kinetic parameters. The few attempts reported in the literature have mainly been based on brute-force molecular dynamics (MD) simulations,[3] which are highly demanding in terms of time and computational power.[4-6] These methods are also unsuitable for industrial use, where dozens of compounds must be prioritized in the hit-to-lead and the lead optimization phases.[3] Importantly, (un)binding rates cannot be directly computed for pharmacologically relevant systems, since the residence time (t_r) of molecules can be in the order of seconds, minutes, or even hours. Thus, smarter algorithms and effective practical solutions are needed to tackle the problem of kinetic rate estimation.

Researchers recently reported a detailed computational study of protein-ligand dissociation based on extensive metadynamics simulations.[7] This work demonstrated the possibility of studying the mechanisms governing unbinding events, and of disclosing their pathways, rates, and rate-limiting steps. Although this approach has provided remarkable atomistic information, its practical effectiveness is limited by the substantial computational resources required (i.e. many weeks on a highly parallel computational infrastructure) to evaluate each single binding and unbinding kinetic constant pair (i.e. k_{on} and k_{off} values). Moreover, the setup of metadynamics simulations is not trivial and requires a strong expertise by the user. The high error in the k_{off} estimation also prevents an easy application of this approach within the drug discovery pipeline, especially referring to an automated process. A possible alternative is to combine the k_{on} obtained from unbiased simulations with the binding free energy estimated using free energy methods. Although promising, this alternative method is still too inaccurate and computationally demanding for any screening purpose.[5]

We recently reported a novel computational method that addresses the challenge of predicting unbinding kinetics, particularly in the hit-to-lead and lead optimization phases of the drug discovery

process.[8] Rather than trying to predict the absolute off-rate value, $k_{off} = t_r^{-1}$, on individual complexes, we proposed a procedure to identify the correct k_{off}-based ranking relationship among congeneric compounds that bind to a given target, using (possibly) limited computational resources. Our solution is rooted in enhancing the transition probability between different free energy minima during MD simulations by smoothing the potential energy surface (PES) of the system using a scaling factor λ (which ranges from 0 to 1) applied to all the potential energies computed during MD simulations. This procedure flattens the PES associate to biomolecular events and samples a greater amount of conformational space in simulation times that are much shorter than standard MD simulations.

The underlying rationale is that simulating a protein-ligand complex under scaled potential energy conditions facilitates the rupture of the key physical interactions that confer stability to the complex, leading to unbinding in much shorter simulation times. These transitions are considered in a statistical framework that combines multiple replicas, a regressive predictive model, and statistical analysis to establish the confidence of the predictions. The averaged residence times can and have been successfully compared to the experimental ones obtained by inverting the k_{off} values. The details of the method, the analytical treatment of the residence times, and their comparison with experimental data are reported in the Supporting Information S1.

The proposed method is easy to implement and sufficiently accurate in reproducing experimental data.[8] However, several questions remain regarding its general applicability, especially for series of chemically unrelated compounds. To clarify these aspects and to assess the validity of our methodology in a predictive and/or prospective manner, we here focused on glucokinase (GK, GK1, or hexokinase IV), a system of primary interest for both national healthcare systems and the pharmaceutical industry. A member of the hexokinase (HK) family, GK[9] catalyzes the phosphorylation of glucose to glucose 6-phosphate, the rate-limiting step in glycolysis. The pharmacological relevance of GK1 is related to its involvement in type 2 diabetes mellitus (T2DM), a rapidly expanding public health issue affecting over 150 million people worldwide.[10,11] GK plays a critical role in glucose homeostasis, serving as a glucose sensor for glucose-dependent insulin

secretion in pancreatic β-cells and regulating glucose uptake and glycogen synthesis in the liver.[12] GK1 displays several features which are unique among the hexokinase family members: i) as a monomeric enzyme, GK achieves positive substrate cooperativity solely through equilibration between multiple conformations,[13,14] ii) GK has a unique biochemical kinetics that accounts for its role as a glucose sensor, iii) GK has a low affinity for glucose as a substrate with a Ka of ~7 mM, which is within the physiological glucose range, and iv) the enzyme also has no direct feedback inhibition from its product glucose 6-phosphate. Together, these kinetic properties enable GK to be highly responsive to fluctuations in blood glucose levels and to ensure that the glucose metabolic flux is closely tied to glucose concentration. While several classes of diabetic therapies are available for clinical use, there is still a significant need for new therapies with improved efficacy. A promising area of current research involves the use of small molecule allosteric activators of GK (GKAs) to lower blood glucose and normalize insulin secretion. Therapeutically, it seems that activating GK1 in the liver and pancreas would be an effective strategy for lowering blood glucose by upregulating hepatic glucose utilization, downregulating hepatic glucose output, and normalizing glucose-stimulated insulin secretion.[15,16]

In the present work, we investigate the unbinding of a novel series of promising GK modulators at the atomic level, using computational simulations coupled to experimental kinetic measurements and X-ray crystallography. We were able to correctly rank most of the ligands according to their residence time and to depict some structure-kinetics relationships (SKRs). In parallel, we challenge our approach, highlighting its major advantages and limitations. Finally, we provide some perspectives on the methodology in terms of its potential to become the computational method of choice for predicting drug residence times.

Results and Discussion

The main hypothesis under which the original method[8] is able to correctly predict residence time distributions is related to chemical similarity, namely the presence throughout a series of only minor chemical modifications (e.g. small functional group replacement relative to the parent compound). The overall protocol comprises the following steps: i) an initial model for each protein-ligand complex is built. Ideally, one protein-ligand complex has been experimentally determined by X-ray crystallography, and chemical replacement is used to generate three-dimensional models of congeneric compounds bound at the target binding pocket. ii) Multiple replicas of scaled molecular dynamics (MD) of the partially restrained system are performed and stopped when the ligand is unbound. iii) The average simulated and experimental unbinding times are normalized with respect to one randomly selected ligand (in this case the ligand for which the experimental structure is available) in order to convert the experimental residence times ratio to a scaled ratio via an exponential relation. iv) A bootstrap analysis on the simulated unbinding times per target is carried out to assess the statistical significance of the observations and to establish whether an increased number of replicas is needed.

In the present study, two different issues were addressed: i) determination of computational residence times for all the available ligands for GK1, their ranking and comparison with the experimental data, and ii) identification of the physicochemical determinants of residence time variations within the series for the ligands and protein (i.e. depiction of the SKRs).

Measurement of GKA binding kinetics. Following our initial work,[17] we exploited the intrinsic fluorescence properties of **1** (RO4389620)[17] (Figure 1) to use it as a fluorescent reporter in a competitive displacement kinetic experiment to rapidly assess the residence time ($1/k_{off}$) of several GKAs, in the presence of glucose to lock GK in its closed conformation. The observed dissociation transients were monophasic, suggesting a single-step dissociation. The GKAs' association kinetics was also measured, using the quenching of the GK intrinsic fluorescence as readout. As previously

observed[17] with **2** (TAFMT) and **3** (LY121260) (Figure 1), association transients were mainly monophasic. The equilibrium K_D was calculated from the on- and off-rate constants. Data are summarized in Table 1.

Residence time prediction. Initially, we focused on molecules for which crystal structures in complex with GK1 were available, namely compounds **1** and **2** (Figure 1). The presence of a well-defined crystal structure is crucial for determining the physicochemical properties of binding; hence, it is also a fundamental prerequisite for applying our method. Moreover, we decided to include in this first round of calculations ligand **3** (Figure 1) that shares a scaffold similar to **1** and size with the aforementioned molecules. We therefore investigated these three systems first, to challenge the method with similar ligands and to understand the level of scaling needed to investigate whole series of ligands.

Ligands **1** and **2** were co-crystallized[18] with GK1 (**1**, resolution 1.7 Å, PDB ID 4NO7; **2**, resolution 2.3 Å, PDB ID 3F9M) without strongly altering the protein structure, which maintains the same folding features of previously resolved GK1 structures (Figure 2a). These ligands bind the same region of the protein, namely the pocket comprised between the C terminus, the 206 – 218 helix, and the 65 – 69 loop. Despite its length, the 65 – 69 loop is well-resolved in the experimental X-ray structure and its geometrical features are clearly visible (Figure 2b, 2c). In the apo- form of GK1 (e.g. 3IDH, resolution equal to 2.14 Å), the heavy atoms of this loop systematically show a higher B factor relative to the holo- forms. This clearly indicates that the presence of the ligand reduces the conformational flexibility of this region. We assumed that this feature is retained also for the complex of GK1 with ligand **3**; hence we placed it in the binding site by superimposing its structure with the scaffold of ligands **1** and **2** (Figure 2d).

We simulated the three systems using two different scaling factors, $\lambda = 0.4$ and $\lambda = 0.5$. A more pronounced potential scaling ($\lambda = 0.4$) pushed the system in a largely diffusive regime, whereas a gentler potential scaling ($\lambda = 0.5$) required longer simulations to overcome barriers and sample

different minima. In this sense, the two different levels of scaling can highlight different details of the ligand-protein interaction during release. By varying the level of potential scaling, we observed the passage from a regime where minima were still present and the energy landscape was slightly more realistic ($\lambda = 0.5$) to a dominantly diffusive behavior ($\lambda = 0.4$) where the differences in minima are flattened. As shown in Table 1 and in Figure 3, the residence times in both cases have the same trend, i.e. t_r (**1**) > t_r (**3**) > t_r (**2**), in very good agreement with the experimental data. A lower value of λ (i.e. 0.4) has the obvious advantage of reducing the computational cost, translating to a roughly 10x speedup compared to a λ value of 0.5. However, when considering residence times in close proximity (as here), a larger λ value increases separation and returns a higher confidence in the ranking.

We then sought to address whether new and structurally different ligands could be adapted starting from the coordinates of a different holo structure, and how this strategy could lead to correct predictions or inaccuracies. We selected two ligands, **4** (GKA50) and **5** (S-44520) (Figure 1), which share a common structural core but differ in the size and nature of their substituents. We built the complexes for compounds **4** and **5** by chemical replacement (see the Supporting Information S2a/b), since experimental crystal structures for these molecules in complex with GK1 were not available. Their starting conformations were obtained by adapting the structure of GK1 in complex with **1**, i.e. by matching the orientation of the new ligand with the experimentally available pose according to their common scaffold. In principle, starting conformations could also be generated by molecular docking. Here, we adopted the aforementioned procedure because a conserved orientation within the binding site is consistently displayed by quite large series of GK1 inhibitors, for which the three-dimensional structure is available. We were thus confident that our molecules should have been oriented accordingly. We performed a series of unbinding simulations of **5** with two λ factors and compared the computational predictions ($t_r \sim 14$ ns for $\lambda = 0.4$, $t_r \sim 45$ ns for $\lambda = 0.5$) with the experimental results obtained via fluorescence stopped-flow ligand displacement kinetics (see the Materials and Methods section for details). The correlation was strictly dependent on the λ factor

(Figure 3). In detail, simulations with λ = 0.5 showed a residence time that, as confirmed experimentally, was between those of **1** and **3**, whereas λ = 0.4 led to a rather high value for ligand **5** with respect to the other compounds.

This case shows that the effect of a larger molecular size (the radius of gyration of ligand **5** is 5.3 Å, whereas the one for ligands **1**, **2** and **3** is ~ 4.3 Å) becomes dominant if the scaling factor is too low (i.e., the potential is too heavily scaled) and the related energy landscape goes flat. This suggests that, if the potential energy scaling is too pronounced, the proposed method could provide residence times that are longer than those expected according to the experimental results. When treating molecules that are partially diverse in terms of structure from those present in the crystal structure, an explorative approach could rely on a double run with two scaling factors and check consensus and possibly discrepancy between them. An even higher and unexpected estimate of residence time was found for **4** with λ = 0.4, for which the residence time value was almost equal to that of **5**. However, whereas computational predictions for **5** were in line with the experimentally determined residence time when λ = 0.5, **4** could not be accurately ranked with respect to the other members of the series even when the higher scaling factor was applied. This behavior could partially be related to the size of the molecule: since it is even larger than **5**, **4** (its radius of gyration is 5.8 Å) displays a less diffusive behavior regardless of the λ scaling and was hence assigned a longer residence time even when the system potential was more gently scaled (λ = 0.5). As recently reported,[8] a considerable change in size within a series of congeneric ligands of the same protein can greatly alter the solvation/desolvation enthalpy of the unbinding process, greatly affecting the ability of the method to predict the relative residence time for a much larger molecule. Together with a notable deviation from the linearity of distributed data in similar cases (see the recent example of Grp78[8]), this results in a lower predictability of the molecule's residence time. Moreover, despite its similarity with ligand **5**, ligand **4** displays more flexibility and degrees of freedom, suggesting that these also could be

important physiochemical features that are able to influence the kinetic properties of a ligand during its release from the binding site.

Finally, we focused on two racemic mixtures of enantiomers **6** (S-49164) and **7** (S-49513) (**6a-6b** and **7a-7b** in Figure 1; see also the Supporting Information S2c/d), which required simulations on each enantiomer to computationally predict residence times. In particular, we carried out 40 runs for each pair of enantiomers, estimating the residence time averaging all of them. These compounds displayed an overall structure that is partially shared by **3** (Figure 2d). As in the previous case, we adapted the new structures on an existing crystal structure (GK1 in complex with **3**), superimposing the amide moiety and the heterocyclic ring, and keeping their orientation in the binding pocket. The relative prioritization of the ligands is retained in the computational model, averaging the residence times obtained by simulating both the enantiomers (Figure 3, Table 1). Again, agreement with the experimental data was fairly good also for these compounds.

We challenged our novel scaled-MD-based protocol in two different scenarios: i) ranking unbinding kinetics for a series of strictly congeneric compounds, and ii) properly ranking and predicting kinetics for structurally unrelated molecules. We further showed that the method can reliably rank ligands of the same series bearing minor chemical modifications, in agreement with our recent report.[8] When the chemical similarity is lower, because of differences in size and in the chemical groups contacting the target, the method provides predictions that are less accurate but still reliable (see below). With the sole exception of **4**, we can observe rather linear correlations for both the λ values, with the noise increasing, as expected, for $\lambda = 0.4$ relative to $\lambda = 0.5$ (correlation coefficients for residence times normalized as recently described[8] (details are reported in the section S1 of the Supporting Information) are 0.72 and 0.91 for $\lambda = 0.4$ and $\lambda = 0.5$, respectively; see Figure 3b). At the lower λ value, the more diffusive nature of simulations could be responsible for larger errors, pointing to the need for more accurate and CPU-intensive simulations to achieve more reliable predictions of unbinding kinetics. This is perfectly in line with expectations of the approach, which could be

adopted using lower λ values for returning fast results in screening campaigns (especially when the ligands display a high degree of chemical similarity), whereas higher λ values could be adopted for accurate but more computationally-demanding predictions.

Structure-kinetics relationships (SKRs). Despite the loss of detail that is associated with scaled MD simulations, it is still possible to identify the chemical features responsible for longer residence times, thus establishing structure-kinetics relationships (SKRs). This could be of valuable help in designing novel GK1 inhibitors with an improved kinetics profile.

Ligands **1**, **2**, and **3** share some similarities in the chemical structure (i.e. an amide moiety connected at the carbonyl side to a 5- or 6-atom aromatic ring). In particular, the amide group provides stability to the ligand-protein complex donating a hydrogen bond to the backbone of R63 (Fig. 4a). However, due to several differences in flexibility and steric hindrance, these compounds overall establish quite heterogeneous patterns of interaction (Figure 1c-e). In this respect, both **1** and **3**, together with an increased number of rotatable bonds, display a common polar sulphonyl group, which is missing in **2**. This group is accommodated in a region of the protein where it can contact multiple polar groups (Figure 4a), creating a network of interactions that stabilizes the protein–ligand complex and that is eventually conducive to longer computed residence times of **1** and **3** with respect to **2**. This is further confirmed by compounds **4** and **5**, which, bearing electron donor oxygen atoms in the same region of the sulphonylic groups of ligands **1** and **3**, were associated with longer computed residence times across the series. The larger size of ligand **4** and **5** might also contribute to extending computed residence time even further, increasing the enthalpic contribution to the overall unbinding activation energy.

In addition, we noticed that the shape of the ligands plays a role in influencing the residence times. Interestingly, the ligands that display a more pronounced linear shape (**2**, **6**, and **7**) exhibit a significantly shorter residence time than those with a T-shaped geometry (**1**, **3**, **4**, and **5**), with an average t_r ~ 27 ns for the former molecules and ~ 90 ns for the latter ones. The GK activators binding

site has a rather linear shape, and hence it is easily accessible to ligands that possess a linear geometry, whereas molecules bearing other shapes than linear require an induced fit binding mechanism. In particular, Tyr125 side chain should move apart to accommodate T-shaped ligands, a geometrical rearrangement that can likely help accounting for the longer residence times experienced by T-shaped molecules relative to the linear ones. Therefore, the present computational approach can help discern among rather different scaffolds, and hence prioritize, for subsequent chemical synthesis, those most promising from a residence time and lead efficacy standpoints.

Finally, we turned our attention to aliphatic rings of **1** and **3** (cyclopentanone in **1**, cyclohexane in **3**), which were docked between T65 and Y215, and that pointed toward the solvent. These rings are located on the other side of the loop T64 – V69 (see Figure 1) with respect to the sulphonyl functional group in ligands **1** and **3**. For ligand **2**, the same region of the protein is occupied by an N-methyl-imidazole ring (which has much greater affinity for water molecules than the aforementioned cycles) and by moieties rich in methyl groups in ligands **4** and **5**. In this respect, a more hydrophobic element of the ligand in this region seems to provide more stabilization to the complex and to be another key determinant for the unbinding kinetics.

To understand in more detail which structural elements of the protein are relevant for the unbinding process, we analyzed all the trajectories, considering their statistical distribution in terms of contacts between ligands and single residues. Considering both scaling potentials adopted in the present study (i.e. $\lambda = 0.4$ or $\lambda = 0.5$), we identified a set of residues that were transiently involved in all the unbinding processes observed (Figure 5 and Figure S3). We considered a residue relevant for unbinding when a contact with the ligand occurred in more than 66% of the replicas, for both sets of simulations. We could thus monitor at a glance the unbinding events providing a residue level description of the dissociation paths. As expected, simulations with $\lambda = 0.4$ had a more pronounced diffusive character and thus provided more transient and aspecific protein-ligand contacts. Conversely, simulations with $\lambda = 0.5$ contributed more significantly to the identification of specific

contacts along the unbinding paths. The contact maps reported in Figure 6 clearly exemplify this dual behavior (all the other maps are available in Figures S4 and S5 of the Supporting Information). Our analysis allowed the identification of two groups of amino acids: one encompassing residues that were trivially clustered around the binding site (residues 56-66, 89-95, 154-156, 196-199, 203, 206-218, 231-234, 444-454), and the other formed by a group of amino acids (residues 241-248, 393-394) located roughly 20 to 25 Å away from the binding site.

Diabetes is heavily characterized by genetic mutations that correspond to a huge variety of phenotypes. Interestingly, the most relevant pathological mutations identified in diabetes and glucose homeostasis[19] (Figure 5b) correspond to residues that, according to our analysis, appear to be relevant in controlling the unbinding process. GKAs can only be accommodated in their binding site once the glucose is bound to the protein via an allosteric modulation. At the same time, the protein's enzymatic efficiency is significantly reduced for many GK1 mutants[19] that are related to diabetes and similar metabolic diseases. Hence, good GKAs could be designed by accounting for the extreme variability of the binding site and the portion of the surrounding region that is relevant for unbinding. For example, mutants V91L, W99R, and Y214C (Figure 5b) severely unbalance the biochemistry of glucose in patients affected by diabetes. Hence, their effect must be related to the change in local dynamics and to steric hindrance in regulating the action of any potential GKA as well as their dissociation.

Conclusions

Residence time is emerging as a key parameter for analyzing and optimizing lead candidates with improved in vivo efficacy, within modern drug discovery programs that take it into account during the hit-to-lead and lead optimization phases. Here, we have shown that residence time predictions are possible using scaled-MD-based approaches, and that the computational protocol can be fine tuned by comparing computational data and experimental kinetics measurements. Different scaling factors can be used and, depending on the ligand set, the best tradeoff between speed and accuracy can be

sought. To properly identify scaling factors, we compared computational predictions with experimental data, with $\lambda = 0.5$ emerging as more suitable for good correlations (relative to $\lambda = 0.4$). This was mainly because the series of ligands included compounds that were not strictly analogous. However, when investigating series of structurally analogous molecules, a lower λ value can be used, allowing faster simulations.[8] Building and expanding on the previously reported SMD-based method,[8] the present work suggests, a practical approach to set the ideal scaling factor to adopt on the basis of the needs of the users and the chemical nature of the ligands. We subsequently focused on SKRs, an accurate description of the structural determinants responsible for a good residence time can be of paramount importance when designing new compounds with improved in vivo profiles. SKRs can be depicted from a ligand-based standpoint, from a structure-based analysis, or from both, provided that structural information for the ligand and target is available. In this context, by means of a statistical analysis of protein-ligand contacts during the unbinding events, we have been able to identify those residues that affect the residence time and overall kinetic and release profile of a compound. In addition, we observed that the ligand shape could be responsible for an induced fit effect, which could explain the different residence times of differently shaped compounds. In conclusion, the present computational methodology is very promising and practical. However, further case studies are needed to highlight the potential and limitations of scaled MD simulations for drug binding kinetics studies, with the goal of making this approach the method of choice for estimating and predicting residence times.

Experimental section

Chemical synthesis

The synthesis of ligands **1-4** was already reported.[17-18] The ones of ligands **6**, **7** and **5** is reported below, with reference (numbers in italics) to the schemes of the synthetic routes in the Supporting Information section S6. All the ligands have been purified using HPLC and characterized using NMR spectroscopy and MS, assessing for all the molecules purity more than 95%.

6-[[(1R,1aR,6aR)-1a-methyl-6,6a-dihydro-1H-cyclopropa[a]indene-1-carbonyl]amino]pyridine-3-carboxylic acid (6). Cyclopropanation reaction of 3-methylindene ***8*** with ethyl diazoacetate was performed in the presence of copper sulfate to afford a mixture of diastereoisomers ***9*** and ***10*** (70/30). After separation by chromatography on silica gel, the ester function of diastereoisomer ***9*** was hydrolyzed to the corresponding acid ***11***. Compound ***11*** was then transformed in acid chloride with oxalyl chloride, and condensed with 2-pyridin acid to afford the desired amide ***12*** as a racemic mixture (compound **6** in the main text), with 98.7% HPLC purity.

^{1}H NMR (400 MHz, 300K, DMSO-d_6): δ 13.10 (m, O**H**), 10.86 (s, N**H**), 8.78 (s, 1H), 8.22 (m, 2H), 7.37 (d, 1H), 7.19 (m, 3H), 3.24 (dd, 1H), 2.98 (d, 1H), 2.38 (dd, 1H), 1.66 (d, 1H), 1.62 (s, 3H); HRMS (ESI/FIA) for $C_{18}H_{16}N_2O_3$: $[M+H]^+$(calculated) = 309.123918, $[M+H]^+$(found) = 309.1231.

6-[[(1R,1aR,6aR)-6a-methyl-1a,6-dihydro-1H-cyclopropa[a]indene-1-carbonyl]amino]pyridine-3-carboxylic acid (7). Cyclopropanation reaction of 2-methylindene ***13*** with ethyl diazoacetate was performed in the presence of rhodium acetate to afford a mixture of diastereoisomers ***14*** and ***15*** (63/37). After separation by chromatography on silica gel, the ester function of diastereoisomer **7** was hydrolyzed to obtain acid ***16***. The acid function was then transformed in acid chloride with oxalyl chloride, and condensed with 2-aminopyridine methyl ester to afford, after saponification, the desired amide ***17*** (compound **7a/b** in the main text), as a racemic mixture, in 99% purity (HPLC).

^{1}H NMR (400 MHz, 300K, DMSO-d_6): δ 13.11 (m, O**H**), 10.93 (s, N**H**), 8.79 (s, 1H), 8.23 (m, 2H), 7.33 (m, 1H), 7.20 (m, 1H), 7.11 (m, 2H), 3.13 (dd, 2H), 2.84 (d, 1H), 1.70 (d, 1H), 1.48 (s, 3H); HRMS (ESI/FIA) for $C_{18}H_{16}N_2O_3$: $[M+H]^+$(calculated) = 309.123918, $[M+H]^+$(found) = 309.1255.

(E)-3-[6-[(3,5-diisopropoxybenzoyl)amino]-3-pyridyl]prop-2-enoic acid (5). 3,5-dihydroxy methyl benzoate ***18*** is dialkylated by isopropyl iodide to obtain the ester ***19***, hydrolyzed to acid ***20***. The acid function is transformed in acid chloride with oxalyl chloride, then condensed with 2-amino-pyridine-5-acrylate ***22***, obtained from 2-amino-5-iodo-pyridine ***21*** and methyl acrylate under Heck

conditions. Resulting condensation ester ***23*** is finally hydrolyzed to afford the desired acid ***24*** (compound **5** in the main text**)**, in 99% purity (HPLC).

^{1}H NMR (400 MHz, 300K, DMSO-d_6): δ 12.5 (s, O**H**), 8.7 (s, 1H), 7.6 (d, J=15Hz, 1H), 6.6 (d, 1H), 4.7 (m, 2H); HRMS (ESI/FIA) for $C_{21}H_{24}N_2O_5$: $[M+H]^+$(calculated) = 385.176348, $[M+H]^+$(found) = 385.1771.

Protein expression and purification

Human glucokinase isoform 1 (pancreatic) was expressed and purified as previously described.[17]

Steady-state kinetics

GK activity was measured by monitoring the rate of G6P formation using the G6PDH/NADP (glucose-6-phosphate dehydrogenase/nicotinamide adenine dinucleotide phosphate)-coupled enzyme assay. Initial rate measurements were carried out at 25 °C in buffer A (HEPES/Na 50 mM, NaCl 100 mM, $MgCl_2$ 5 mM, TCEP 2 mM, pH 7.1) on a PHERAstar microplate reader (BMG Labtech) by following the increase of absorbance at 340 nm due to the reduction of NADP. Assays were performed in a 384-well microplate with a final volume of 50 µl per well. To determine glucose or Mg-ATP-related kinetic parameters, glucose concentration was varied (0-50 mM) while maintaining a saturating (4 mM) Mg-ATP concentration. A lag phase observed in the early time course of the reaction was omitted for the determination of the initial rate. This lag phase is likely due to the coupled-enzyme assay, as already described.[17] Initial rate data obtained with varying glucose concentrations were fitted to the Hill equation. Nonlinear least-squares regression analysis was performed with SigmaPlot 9.01.

Dissociation rate measurements by competitive displacement kinetics

Because of the relatively fast dissociation kinetics of the GKA studied, the stopped-flow rapid mixing technique was used, with the same setup described for the association kinetics measurements. For all GKA except ligand **2**, displacement experiments were performed by mixing an equal volume (75 µL) of the pre-equilibrated GK/GKA complex with an excess of competing ligand **2** in buffer A, complemented with 5% DMSO and 100 mM Glc. GK, GKA, and ligand **2** concentrations after mixing were 1 µM, 5 µM, and 50 µM respectively. Fluorescence traces were analyzed as previously described.[17]

Protein crystallization and X-ray diffraction data collection

The crystallization samples were solutions of GK (at 6 or 10 mg ml/1) in 20mM HEPES, 50 mM KCl, 2 mM tris(2-carboxyethyl)phosphine (TCEP), 100 mM glucose, 5% (w/v) glycerol pH 7.5. When present, activators (**1** and **2**) were diluted to a concentration of 80 mM and added to the protein solution prior to concentration and AMP-PNP was added to a final concentration of 10 mM with the addition of 1 mM $MgCl_2$.

Initial crystals were obtained under conditions similar to those already reported[14] and showed the same unit cell parameters. However, in our hands, these crystals did not diffract well and grew very slowly. After extensive screening in a matrix of various PEGs buffered using HEPES pH 7.5, a different crystal form was obtained with space group P212121. This crystal was further used for preliminary seeding experiments.

All crystals used in this work were grown under the same conditions, at room temperature using hanging-drop vapor diffusion in Linbro plates. 2 ml protein solution was mixed with 1 ml reservoir solution (3232ro plates. 2 p 100 mM HEPES pH 7.5). After 3 d of equilibration, the drops were seeded with crushed crystals diluted in reservoir solution.

Crystals appeared within 12 h in the presence of TAFMT and notably more slowly in its absence. They grew to their final dimensions within 2 d (with TAFMT) or longer (without TAFMT). Crystals

were obtained over a range (32–38%) of PEG concentrations and cooled under liquid nitrogen without the need for further cryoprotection.

Data sets were collected on the SLS PX6d beamline (Villigen, Switzerland) and were reduced using XDS.[20] To avoid any bias on the densities of the ligands, the following procedure was used for molecular replacement and structure refinement. Firstly, molecular replacement was performed using the program MOLREP.[21] As a model, we used the structure 1V4S,[14] from which we removed all ligands or bound waters. This was followed by refinement omitting all ligands using REFMAC.[22] After several refinement cycles, water molecules were built iteratively by ARP/wARP.[23] At this stage, the densities of the ligands were clearly visible and were then refined. Further refinement using TLS was also performed.[24] The structures have been deposited with the PDB ID 4NO7[18] (**1**) and 3F9M[18] (**2**), the latter being a refinement of the structure 3D18.

MD and scaled MD simulations

Each compound was geometrically optimized via a quantum mechanical approach: electron density calculations were performed in NWChem[25] using the basis set 6-31G* at the Hartree-Fock (HF) level of theory. Partial charges were derived using the RESP methodology[26] as implemented in Antechamber, leading via a GAFF parameterization to a complete topological description of each ligand to be used for classical simulations. The coordinates of the HF/6-31G* optimized ligands are reported in the Supporting Information, section S7.

Protein-ligand experimental complexes were used as a starting point of the MD simulations. When the experimental protein-ligand complex structure was not available the ligand was placed in the binding site according to its best superimposition with the most corresponding moiety of the available experimental structures (PDB ID: 4NO7 and 3F9M).

The protein-ligand complexes were then used as a starting point for molecular dynamics simulations performed in a GROMACS 4.6.1[27] version customized to perform scaled molecular dynamics (scaled MD)[28] implemented by BiKi Life Sciences (BiKi Technologies, Genoa, Italy) as recently described.[8]

All the complexes were placed in the geometrical center of parallelepiped-shaped boxes of volume equal to ~ 650 nm^3. The simulation boxes were then solvated using tLeap, with ~70000 TIP3P water molecules.[29] Some water molecules were replaced with sodium ions in order to preserve the electro-neutrality of the system according to need, i.e. the charge of the protein plus the charge of the ligand, which varied according to the case considered. The system was minimized with the steepest descent method, followed by equilibration of the restrained protein (isotropic 1000 kJ mol^{-1} nm^{-1} force applied to each heavy atom of the protein backbone) in NPT (up to 400 ps, pressure = 1 atm) and NVT (up to 400 ps) ensembles at 300 K via a standard MD procedure. Electrostatics were treated with the cutoff method for short-range interactions and with the Particle Mesh Ewald method for the long-range ones (rlist = 1 nm, cutoff distance = 0.9 nm, vdW distance = 0.9 nm, PME order = 4).[30] The constant temperature conditions were provided by using the V-rescale thermostat,[31] a modification from Berendsen's coupling algorithm.

A series of partially unrestrained (see below) scaled MD production runs were performed for each complex until the occurrence of the unbinding event, defined as the situation where interactions between the ligand and the binding site are no longer present, corresponding to a distance between the ligand and the binding site centers of mass of 30 Å. This threshold has been chosen in order to achieve complete solvation of the ligand (no remaining hydrogen bonds nor contacts between the small molecule and the protein) mimicking the experimental conditions adopted for the experimental measurement of the k_{off} values. Twenty simulations were performed for each protein-ligand system and for each scaling factor λ. We used two different scaling factors, $\lambda = 0.4$ and $\lambda = 0.5$, thus having 40 simulations for each complex and 280 simulations overall.

The restraints were adopted as previously described,[8] i.e. weakly (50 kJ mol^{-1} nm^{-1}) restraining all the backbone heavy atoms with the exception of those in the binding site, defined as those belonging to the residues within 6 Å of the surface of the ligand computed on the starting crystal structure (without hydrogen atoms). The residues left unrestrained were: 57 58 59 60 61 62 63 64 65 92 93 94 155 206 208 210 211 214 216 217 218 231 248 447 448 449 451 452 453 454. The unbound state was determined using the same criterion of Mollica et al.[8]

All the simulations were set up using the BiKi software package and performed on a set of in-house machines, equipped with two esacore Intel Xeon processors and 2 NVIDIA GTX 780 GPUs, for a total of 1300 CPU days.

Abbreviations used

ADP, adenosine diphosphate; AMBER, assisted model building with energy refinement; ATP, adenosine triphosphate; CD, circular dichroism; CPU, central processing unit; GPU, graphical processing unit; DMSO, dimethyl sulphoxide; GAFF, generalized AMBER force field; GK, glucokinase; GKA, glucokinase activator; GROMACS, Groningen machine for chemical simulations; HEPES 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid; HF, Hartree-Fock; HK, hexokinase; HPLC, high performance liquid chromatography; MD, molecular dynamics; MS, mass spectrometry; NADP, nicotinamide adenine dinucleotide phosphate; NMR, nuclear magnetic resonance; NVT, canonical ensemble (amount of substance (N), volume (V), and temperature (T) are conserved); NPT, isothermal-isobaric ensemble (amount of substance (N), temperature (T), and pressure (P) are conserved); PES, potential energy surface; PME, particle mesh Ewald; RESP, restrained electrostatic potential; SKR, structure kinetics relationship; SMD, scaled molecular dynamics; SPR, surface plasmon resonance; T2DM, type 2 diabetes mellitus; TCEP, tris(2-carboxyethyl)phosphine; vdW, van der Waals.

Supporting Information

Mathematical derivation of the k_{off} and residence time scaling, starting poses for ligands **4**, **5**, **6a**, and **7a**, surface representation of the unbinding paths in the GK1 structure, ligand protein contact maps during the unbinding process for the single ligands, ligand protein contact maps during the unbinding process for the enantiomeric mixtures, scheme of reactions for the synthesis of compounds **7**, **6** and **5**, ligand's optimized structure coordinates with RESP charges in .mol2 format.

Corresponding Authors Information

Phone: +39 010 71781530

Fax: +39 010 7170187

E-mail: andrea.cavalli@iit.it

Phone: +33 (0)1 55 72 28 15

Fax: +33 (0)1 55 72 52 15

E-mail: pierre.ducrot@servier.com

References

1. Copeland, R. A.; Pompliano, D. L.; Meek, T. D. Drug-target residence time and its implications for lead optimization. *Nat. Rev. Drug Discov.* **2006,** *5*, 730-739.

2. Swinney, D. C. Biochemical mechanisms of drug action: what does it take for success? *Nat. Rev. Drug Discov.* **2004,** *3*, 801-808.

3. De Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of molecular dynamics and related methods in drug discovery. *J. Med. Chem.* **2016**, *59*, 4035-4061.

4. Decherchi, S.; Berteotti, A.; Bottegoni, G.; Rocchia, W.; Cavalli, A. The ligand binding mechanism to purine nucleoside phosphorylase elucidated via molecular dynamics and machine learning. *Nat. Commun.* **2015,** *6*, 6155.

5. Buch, I.; Giorgino, T.; De Fabritiis, G. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2011,** *108*, 10184-10189.

6. Shan, Y.; Kim, E. T.; Eastwood, M. P.; Dror, R. O.; Seeliger, M. A.; Shaw, D. E. How does a drug molecule find its target binding site? *J. Am. Chem. Soc.* **2011,** *133*, 9181-9183.

7. Tiwary, P.; Limongelli, V.; Salvalaglio, M.; Parrinello, M. Kinetics of protein-ligand unbinding: Predicting pathways, rates, and rate-limiting steps. *Proc. Natl. Acad. Sci. U. S. A.* **2015,** *112*, E386-391.

8. Mollica, L.; Decherchi, S.; Zia, S. R.; Gaspari, R.; Cavalli, A.; Rocchia, W. Kinetics of protein-ligand unbinding via smoothed potential molecular dynamics simulations. *Sci. Rep.* **2015,** *5*, 11539.

9. Matschinsky, F. M. Glucokinase as glucose sensor and metabolic signal generator in pancreatic beta-cells and hepatocytes. *Diabetes* **1990,** *39*, 647-652.

10. Wild, S.; Roglic, G.; Green, A.; Sicree, R.; King, H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care* **2004,** *27*, 1047-1053.

11. Zimmet, P.; Alberti, K. G.; Shaw, J. Global and societal implications of the diabetes epidemic. *Nature* **2001,** *414*, 782-787.

12. Agius, L.; Peak, M.; Newgard, C. B.; Gomez-Foix, A. M.; Guinovart, J. J. Evidence for a role of glucose-induced translocation of glucokinase in the control of hepatic glycogen synthesis. *J. Biol. Chem.* **1996,** *271*, 30479-30486.

13. Heredia, V. V.; Thomson, J.; Nettleton, D.; Sun, S. Glucose-induced conformational changes in glucokinase mediate allosteric regulation: transient kinetic analysis. *Biochemistry* **2006,** *45*, 7553-7562.

14. Kamata, K.; Mitsuya, M.; Nishimura, T.; Eiki, J.; Nagata, Y. Structural basis for allosteric regulation of the monomeric allosteric enzyme human glucokinase. *Structure* **2004,** *12*, 429-438.

15. Grimsby, J.; Berthel, S. J.; Sarabu, R. Glucokinase activators for the potential treatment of type 2 diabetes. *Curr. Top. Med. Chem.* **2008,** *8*, 1524-1532.

16. Sarabu, R.; Berthel, S. J.; Kester, R. F.; Tilley, J. W. Novel glucokinase activators: a patent review (2008 - 2010). *Expert Opin. Ther. Pat.* **2011,** *21*, 13-33.

17. Antoine, M.; Boutin, J. A.; Ferry, G. Binding kinetics of glucose and allosteric activators to human glucokinase reveal multiple conformational states. *Biochemistry* **2009,** *48*, 5466-5482.

18. Petit, P.; Antoine, M.; Ferry, G.; Boutin, J. A.; Lagarde, A.; Gluais, L.; Vincentelli, R.; Vuillard, L. The active conformation of human glucokinase is not altered by allosteric activators. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 929-935.

19. Matschinsky, F. M. Assessing the potential of glucokinase activators in diabetes therapy. *Nat. Rev. Drug Discov.* **2009,** *8*, 399-416.

20. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* **2010,** *66*, 133-144.

21. Vagin, A.; Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **2010,** *66*, 22-25.

22. Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **2011,** *67*, 355-367.

23. Cohen, S. X.; Ben Jelloul, M.; Long, F.; Vagin, A.; Knipscheer, P.; Lebbink, J.; Sixma, T. K.; Lamzin, V. S.; Murshudov, G. N.; Perrakis, A. ARP/wARP and molecular replacement: the next generation. *Acta Crystallogr. D Biol. Crystallogr.* **2008,** *64*, 49-60.

24. Painter, J.; Merritt, E. A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D Biol. Crystallogr.* **2006,** *62*, 439-450.

25. Valiev M., B. E. J., Govind, N., Kowalski, K., Straatsma, T.P., van Dam, H.J.J., Wang, D., Nieplocha, J., Apra, E., Windus, T.L., de Jong W.A. . NWChem: a comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.* **2010,** *181*, 1477.

26. Bailey, C. I., Cieplak, P., Cornell, W. D. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993,** *97*, 10269–10280

27. Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013,** *29*, 845-854.

28. Sinko, W.; Miao, Y.; de Oliveira, C. A.; McCammon, J. A. Population based reweighting of scaled molecular dynamics. *J. Phys. Chem. B* **2013,** *117*, 12759-12768.

29. 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.* **1983,** *79*, 926–935

30. Darden, T.; Perera, L.; Li, L.; Pedersen, L. New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure* **1999,** *7*, R55-60.

31. Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007,** *126*, 014101.

32. Efron, B. Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **1979,** *7*, 1-26.

Table 1

Experimental K_d, k_{on}, k_{off}, and residence time ($t_{r,exp}$), and in silico residence time ($t_{r,comp}$) for each compound examined. K_ds are expressed in μM, k_{on} in $s^{-1}M^{-1}$, k_{off} in s^{-1}, $t_{r,exp} = 1 / k_{off}$ in s, and $t_{r,comp}$ in ns. The activity of the enzyme is reported (in percentage) with reference to the yield of glucose conversion in the presence of the ligand (details are provided in Materials and Methods). The $t_{r,comp}$ values are reported for both λ values used: an estimation of the error is reported as standard error (se) and via a bootstrapping procedure (bs).[32] The experimental measurements for ligands from **1** to **5** have been performed on the enantiomers indicated in Fig.1, whereas the data for ligands **6** and **7** have been collected for the racemic mixtures.

Ligand	K_d	k_{on}	k_{off}	$t_{r,exp}$	Act. (%)	λ = 0.4		λ = 0.5	
						$t_{r,comp,se}$	$t_{r,comp,bs}$	$t_{r,comp,se}$	$t_{r,comp,bs}$
1	0.83	0.1×10^6	0.12	8.3	90	5.5 ± 0.8	5.5 ± 0.9	105.4 ± 9.2	105.1 ± 10.1
2	0.36	1.1×10^6	0.43	2.3	170	3.8 ± 0.2	3.7 ± 0.3	29.4 ± 5.1	29.3 ± 5.3
3	0.40	0.9×10^6	0.37	2.7	130	5.0 ± 0.5	5.0 ± 0.4	38.7 ± 7.2	38.9 ± 7.1
4	3.3	0.2×10^6	0.61	1.6	20	12.9 ± 2.9	12.8 ± 2.8	92.5 ± 7.5	92.9 ± 7.3
5	0.72	0.2×10^6	0.16	6.3	60	14.1 ± 0.7	14.3 ± 0.6	99.4 ± 6.6	99.7 ± 6.7
6	2.0	0.7×10^6	1.4	0.7	n.a.	6.5 ± 1.1	6.5 ± 1.0	25.7 ± 3.4	25.9 ± 3.9
7	4.4	1.3×10^6	5.6	0.2	n.a.	6.2 ± 1.2	6.3 ± 1.1	24.6 ± 2.6	24.7 ± 3.0

Figure 1

Chemical structures of the ligands used in complex with GK1. The protonated structures are reported for the acidic ligands, whereas the ionic form was used in the simulation.

Figure 2

X-ray structure of the three complexes of GK1 with **1, 2** , and reconstructed pose of **3** . a) Overall structure of GK1 in complex with 1 (red), 2 (blue), and 3 (yellow) superimposed in the binding site. b-d) Local structure of the binding pocket for 1 (b), 2 (c), and 3 (d).

Figure 3

Normalized experimental vs. computational residence times for $\lambda = 0.4$ (a) and $\lambda = 0.5$ (b) (errors for the computational residence times have been normalized as well, according to error propagation). The ratios between experimental or computational data and a reference value have been plotted: the residence time values for ligand **3** have been used as a reference due to their average placement within the overall data distribution.[8] The experimental residence times (on the abscissa) were scaled, after the aforementioned referencing scaling, using the same factor λ used for scaling the potentials during the SMD simulations. Analytical details of the whole procedure are reported in the section S1 of the Supporting Information.

A linear fit was applied to the whole datasets (regression lines are not shown) as well with the exclusion of the ligand **4**, leading to a correlation factor R equal to 0.41 (0.72 excluding ligand **4**) for $\lambda = 0.4$ and to 0.81 (0.92 excluding ligand **4**) for $\lambda = 0.5$.

Figure 4

Binding site of GK1 and its interactions with ligand **1**. Hydrogen bonding between the amide proton of the ligand and the R63 backbone carbonyl is reported as a dotted line.

Figure 5

Unbinding processes for key GK1 residues (in red), identified as explained in the text, are annotated on the protein 3D structure. a) An overall representation. b) A detailed representation displaying all the contacts in proximity of the binding site. (b) also reports a set of residues (in green using a stick representation; the spheres correspond to the backbone alpha carbon atoms), These residues correspond to pathological mutants in T2DM, which severely affect the enzymatic properties of GK1. Compound **1** is reported in yellow sticks. The residues highlighted in red are 56-66, 89, 92-95, 154-156, 196-199, 203, 206-218, 231-234, 241-248, 393-394, and 444-454.

Figure 6

The mono-dimensional contact map between ligand **5**) and GK1 is reported for two different levels of scaling adopted during scaled MD simulations, $\lambda = 0.4$ (black) and $\lambda = 0.5$ (red). The numbering of residues is reported on the abscissa and the percentage of contacts computed for all simulation replicas is reported on the ordinate. Details of the method adopted for generating these maps are described within the text.

Figure 1

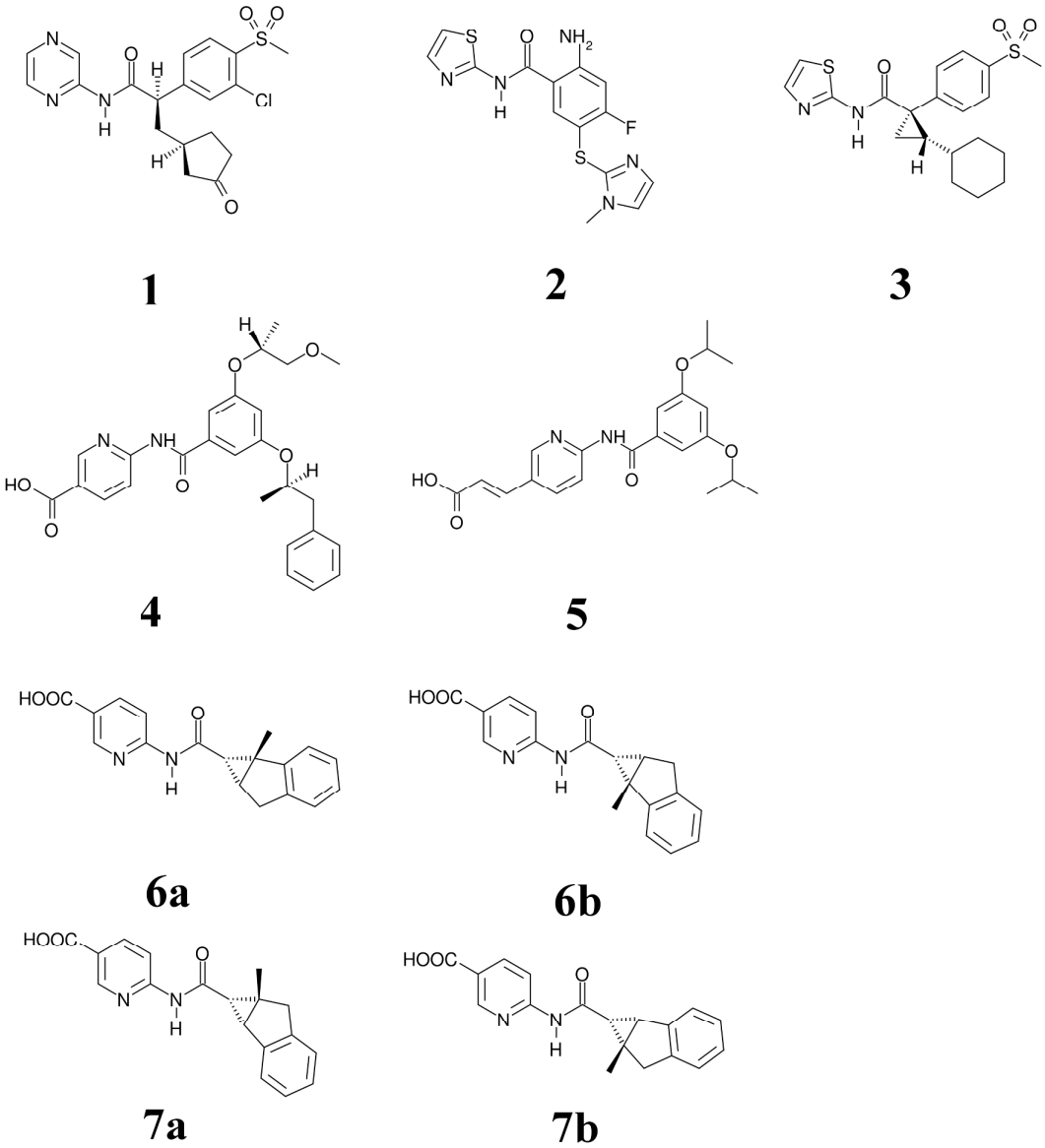

Figure 2

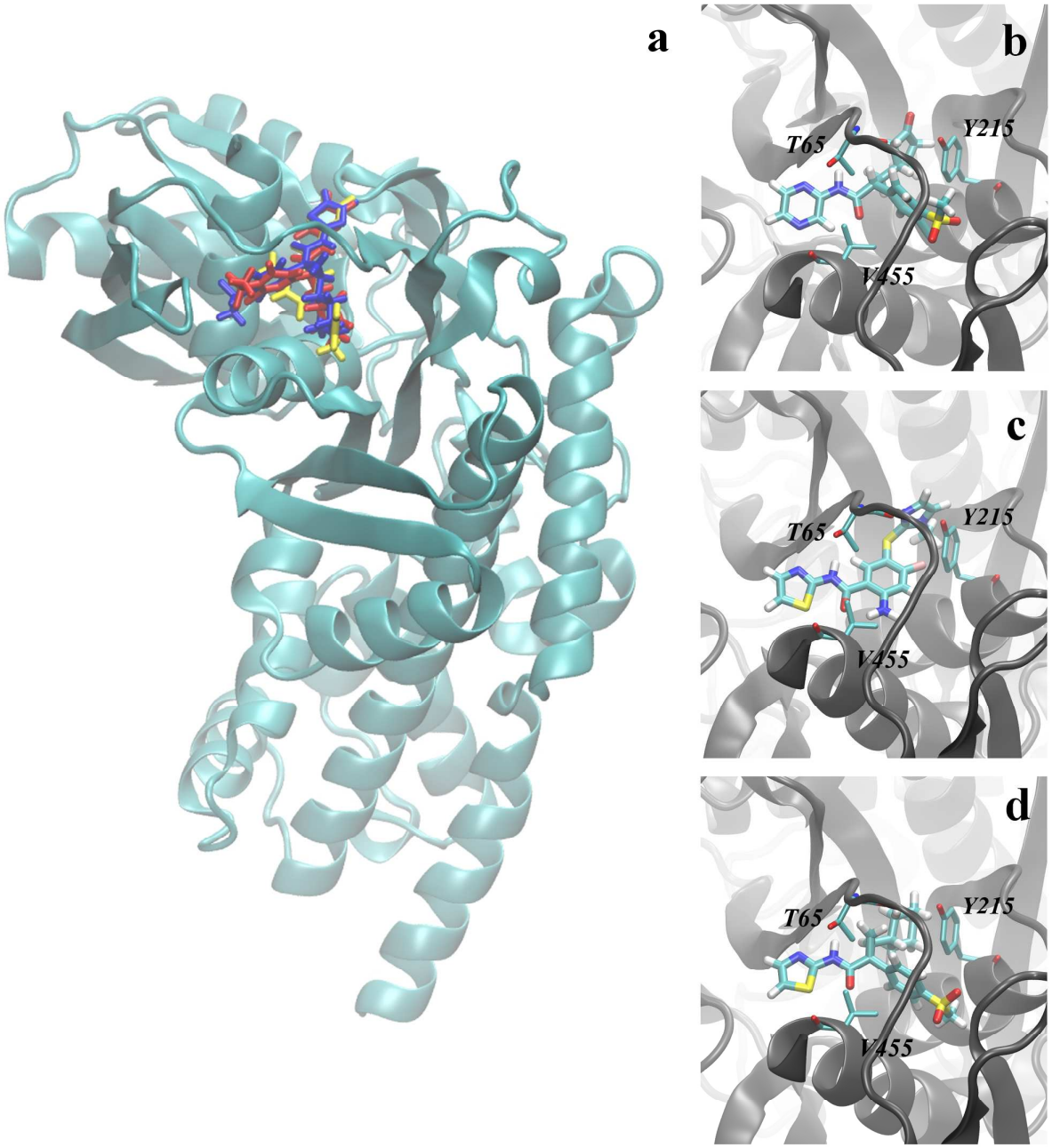

Figure 3

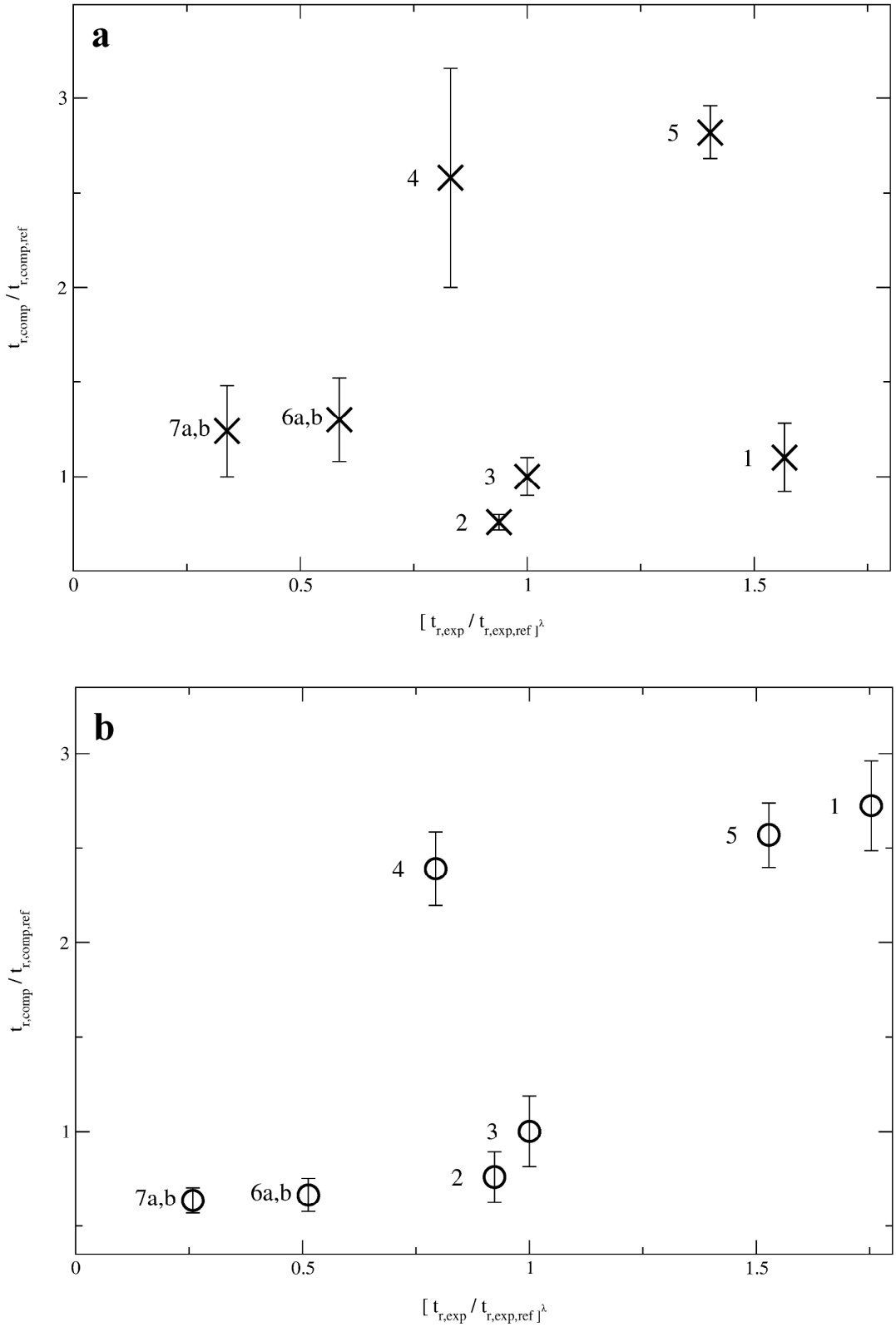
a
b
$t_{r,comp} / t_{r,comp,ref}$
$[t_{r,exp} / t_{r,exp,ref}]^{\lambda}$
0
0.5
1
1.5
2
3
1
2
3
4
5
6a,b
7a,b

Figure 4

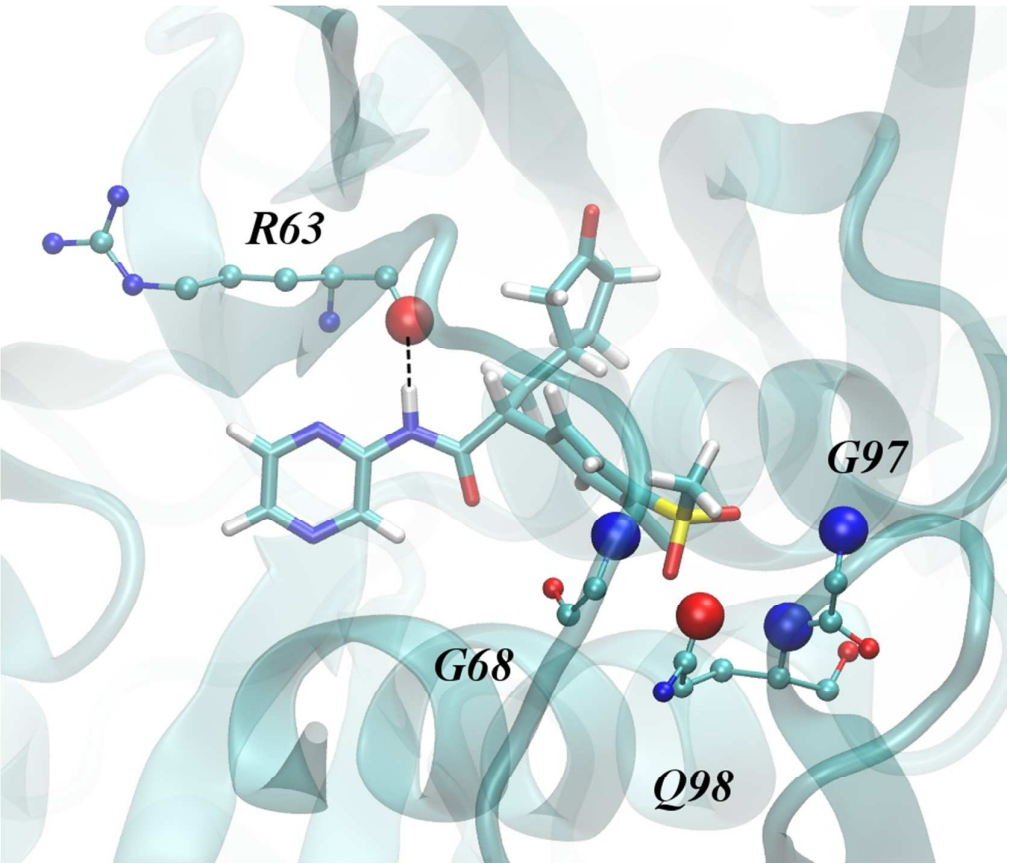

Figure 5

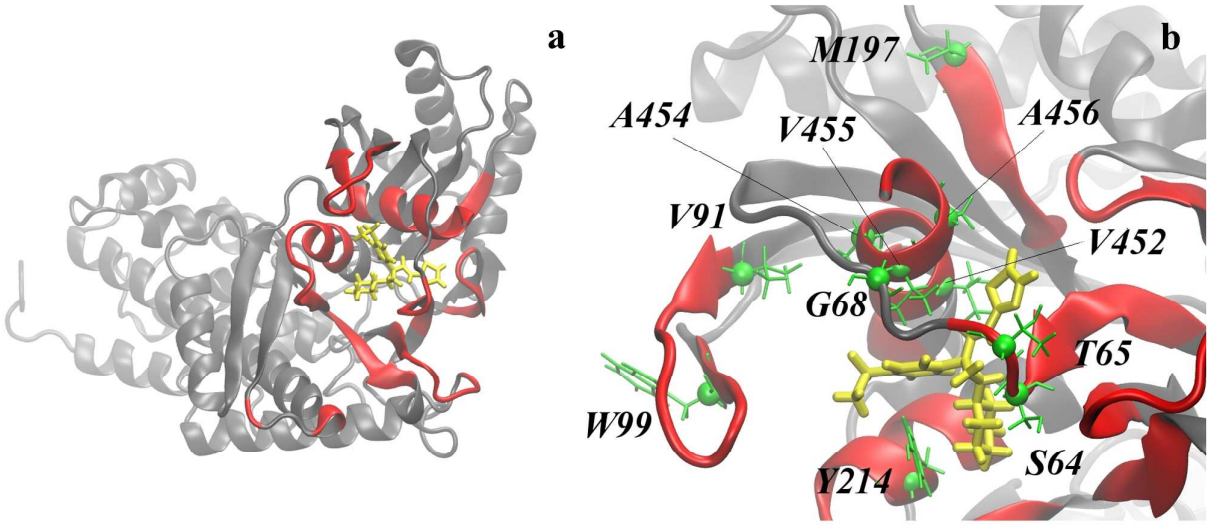

Figure 6

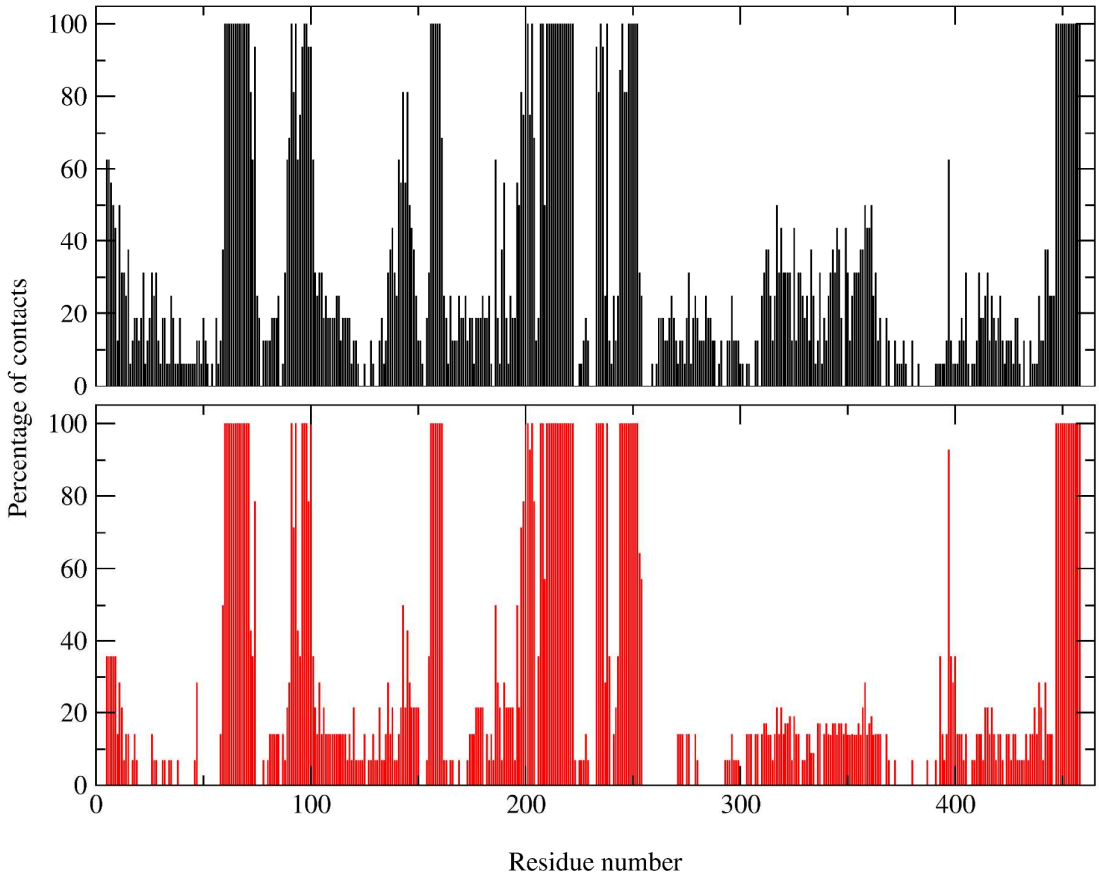
Percentage of contacts
100
80
60
40
20
0
100
80
60
40
20
0
0
100
200
300
400
Residue number

Table of Content Graphic