

Predicting residence time of GPCR ligands with machine learning

Running Header: Predicting residence time with ML

Andrew Potterton^{1,2}, Alexander Heifetz² and Andrea Townsend-Nicholson*¹

*Corresponding Author: Andrea Townsend-Nicholson

Email: a.townsend-nicholson@ucl.ac.uk

¹ Structural and Molecular Biology, Darwin Building, University College London, Gower Street,
London, WC1E 6BT, United Kingdom

² Evotec (U.K.) Ltd., 114 Innovation Drive, Milton Park, Abingdon, Oxfordshire, OX14 4RZ, United
Kingdom

Predicting residence time of GPCR ligands with machine learning

Abstract

Drug-target residence time, the duration of binding at a given protein target, has been shown in some protein families to be more significant for conferring efficacy than binding affinity. To carry out efficient optimisation of residence time in drug discovery, machine learning models that can predict that value need to be developed. One of the main challenges with predicting residence time is the paucity of data. This chapter outlines all of the currently available ligand kinetic data, providing a repository that contains the largest publicly available source of GPCR-ligand kinetic data to date. To help decipher the features of kinetic data that might be beneficial to include in computational models for the prediction of residence time, the experimental evidence for properties that influence residence time are summarised. Finally, two different workflows for predicting residence time with machine learning are outlined. The first is a single-target model trained on ligand features; the second is a multi-target model trained on features generated from molecular dynamics simulations.

Key Words

Residence time; Machine learning; Drug discovery; GPCR; Binding kinetics; Molecular dynamics

1. Introduction

Drug-target residence time, the inverse of the rate of ligand dissociation has been observed, for some targets, to be more influential than equilibrium binding affinity in conferring efficacy [1–5]. Several examples that correlate residence time to *in vivo* efficacy have been published [1–5]. An analysis of 50 drugs that act on 12 different targets has revealed that 70% of drugs with long residence times have a higher efficacy than comparable drugs with short residence times [1]. The evidence for association of residence time and efficacy is most abundant for GPCRs. The efficacy of muscarinic acetylcholine M₃ receptor agonists was found to correlate solely with their associated residence times and not with

their binding affinities [2]. Similarly, with A_{2A} adenosine receptor agonists, residence time was the only binding measure found to correlate with *in vivo* efficacy [4]. The correlation between residence time and efficacy has also been observed with antagonists. Residence times of antihistamine antagonists were found to correlate with their ability to inhibit the H_1 histamine receptor in cells [5]. Thus, in many different GPCR examples, efficacy of agonists and inhibition strength of antagonists has been shown to be better correlated with residence time than with binding affinity. It should be noted that extending residence time not only impacts efficacy, it may also have an effect on drug-dose intervals. Tiotropium has a 50-fold longer residence time compared to Ipratropium, both ligands of the M_3 muscarinic receptor, meaning that Tiotropium can be dosed less frequently [6, 7].

The difference in residence time of the ligand for its target compared with off-target proteins determines the probability of having off-target side effects. Classically, target selectivity is measured as a ratio of the binding affinity values for the off-target compared to the target protein (equilibrium selectivity). As noted by Copeland et al. (2006) [8], the concentration of drug in blood plasma is not constant and, as such, dissociation rates of the drug for the different proteins determine the temporal selectivity of the drug. In other words, upon administration of drug, initial selectivity is governed by the difference in binding affinities between target and off-target proteins but, as the concentration of drug in the plasma decreases over time, the difference in residence times for the target and off-target also determines the overall selectivity of a drug. Whilst equilibrium selectivity is important to optimise, kinetic selectivity can be utilised to achieve overall selectivity in cases where there is an off-target protein with high sequence similarity that will yield low equilibrium selectivity [9].

Concerns have been raised about optimising for drug-target residence time in drug discovery. It has been shown that residence time is only responsible for occupancy if it exceeds clearance time by mathematical models, and for many small-molecule pharmaceuticals on the market this is not the case [10]. There is also the question of protein turnover; if a target has a high turnover there may be no need to extend residence time. For example, a Tyrosine kinase inhibitor with a very long residence time of a week but a high receptor turnover had an *in vivo* occupancy of less than 50% after 24 hours

[11]. Extending residence time should be considered in the context of both protein turnover and ligand clearance time. However, extended target occupancy may not be the only mechanism by which long residence times lead to better efficacy. Sustained antagonist binding with long drug-target residence time may prevent transient agonists from binding as the target would already be occupied [12].

As a result of these different findings, it has been suggested that it is critical to consider drug-target residence time in both the hit-to-lead and lead optimisation phases of drug discovery [8]. To do this effectively, however, we need computational tools that predict and rationalise residence time, a property that is more difficult to determine experimentally than binding affinity. Machine learning (ML) has been at forefront in drug discovery for many years, partly due to the fact it can make regressions in complex data that humans cannot understand, and is a good candidate to consider when searching for a methodology for the computational prediction of drug-target residence time. In order to make an accurate ML regression model, we need to select the correct input features, which are then fed into the ML model to give a predicted value output. Choice of feature is critical to the success of the ML regression model in making accurate predictions that will converge with experimental results.

1.1 Features that confer long drug-target residence time

In order to extend residence time without affecting equilibrium binding affinity, the stability of the transition binding state needs to be achieved. Details of these transition states are currently difficult to assess experimentally but are able to be observed using computational simulations (molecular dynamics). With the recent advancement of X-ray free-electron lasers, the identification of ligand binding transition states is becoming experimentally more feasible using time-resolved crystallography [13]. Until the availability of X-ray free-electron lasers becomes more widespread, however, a combination of computational simulations and experimentally-determined information of the ligand derived in the bound state will need to be used to help uncover the molecular determinants of drug-target residence time.

Suggestions about features that are important for conferring long drug-target residence time has been shown in both computational and experimental studies. A survey of Pfizer's database of 2000 compounds with residence time values reveals a correlation between extended residence time and ligand size [14]. This is supported by an analysis of all available GPCR ligand kinetic data (500 compounds) as shown in Figure 1, which reveals a weak positive correlation between measures of size and residence time. In some cases, molecular weight was found to be correlated solely with residence time and not with binding affinity [15]. In other instances, there is a strong correlation between residence time and ligand molecular weight [16]. In these cases, care should be taken to ensure that the method used to determine residence time outperforms this simple linear correlation is not just predicting ligand size as a proxy for the determination of residence time.

<Insert Figure 1 here>

In addition to ligand size, interactions with water are known to be important in determining residence time. Buried hydrophilic interactions, interactions that are shielded by water, have been proven both computationally and experimentally to extend residence time [17]. These interactions have a higher energy barrier, meaning that they are more stable and less transient. Water has also been shown to be an important factor for conferring long residence time in a GPCR receptor (the A_{2A} receptor) and compounds predicted to have reduced solvation in the bound state were found to have an increased residence time [18]. In our previous publication using steered molecular dynamics, the change in solvation between the bound and unbound ligand was observed to strongly correlate with residence time, partly because ligands with larger changes in solvation energy are more likely to make buried hydrophilic interactions with the protein [19].

The flexibility of the binding site has been noted to also affect residence time [20]. A ligand can affect the stability of the binding site by stabilising interactions within the protein, for example ZM-241,385 stabilizes an intra-protein salt bridge in the A_{2A} receptor, increasing residence time. Disruption of this salt-bridge, by site-directed mutagenesis, reduces the residence time of this ligand 16.8-fold from 84

minutes to 5 minutes [20]. Tiotropium, the previously mentioned Muscarinic subfamily antagonist, has a 10-fold longer residence time for the M_3 than the M_2 receptor subtype [21, 22]. Molecular dynamics simulations have revealed that the second extracellular loop is more flexible in the M_2 receptor [23]. Interactions made by the ligand to the receptor that decrease protein flexibility of ECL2 have been shown to increase residence time.

Matched molecular pair (MMP) analysis was recently used to try to understand structural-kinetic relationships [24]. MMP analysis was generally more successful in explain ligand association rates, with an increase in a ligand's polarity resulting in a slow-down in the association rate. Only a few MMP transformations were identified where the residence time was significantly extended, whereas the binding affinity and the association rate remained largely unperturbed. One of these transformations was the removal of the hydroxyl group from Tiotropium (Des-Hydroxy Tiotropium) which decreases residence time by 56-fold for the M_3 muscarinic receptor. This hydroxyl group in Tiotropium is known to form a hydrogen bond to Asn507 [25]. Due to surrounding aromatic residues, this interaction might be a buried hydrophilic interaction, hence why the presence of the hydroxyl group extends residence time.

1.2 Previous ML methods that have been used to predict RT

There have been a number of published methods that have attempted to predict drug-target residence time using ML but far fewer than the number of ML-based methods that attempt to predict binding affinity values. This is likely due to two reasons: first the severe lack of training data for residence time compared to binding affinity; and, second, the relatively recent identification of the role of residence time in drug discovery [3]. The lack of training data can be seen in the fact that the majority of current residence time prediction methods are trained on small numbers (under 100) of compounds. In addition, these methods have been used for only two protein targets, HIV-1 protease and HSP90, further highlighting the issue of the scarcity of training data.

A summary of different ML methods that have been used to predict drug-target residence time is shown in Table 1. One of the first published methods was a QSKR (quantitative structure kinetic relationship) model that used VolSurf descriptors of mainly water to predict the residence times of 37 HIV-1 protease inhibitors [26]. Another method used COMBINE analysis, which uses the electrostatic and Van der Waals interactions made by specific protein residues to the ligand as features (with different weights) in a PLS (partial-least square) model [27]. Another ML-based method, used protein-ligand interaction fingerprints from random acceleration molecular dynamics trajectories of ligands dissociating from HSP90 [28]. By doing this, not only was a support vector (SV) regression model developed that can predict residence times of prospective HSP90 ligands, it also allowed for ligand interactions with specific protein-residues to be noted as important in extending residence time, helping to guide structure-based drug design for HSP90 ligands. The general accuracy of all these methods is around 1 log unit. These studies were primarily validated on small datasets, making the true predictive nature of the difficult to assess.

The ability to deploy any of these ML models is very much dependent on the specific protein system under study as one can only predict residence time of a compound for a protein target when there is enough kinetic binding data for that protein with which to train a ML model. This constrains the application of these methods to well-studied protein systems and makes them less suited to drug discovery, which often involves developing pharmaceuticals against novel protein targets (first-in-class targets).

2. Materials

This section details instructions of the Python libraries that need to be installed to perform the ML methods detailed here.

2.1.1 To install PyQSAR [29], one needs to create a Python 2.7 environment (NOTE 1).

```
conda create --name py2 python=2.7
```

- 2.1.2 For ligand feature generation, Mordred [30] will be used. This must be installed in the Python 2.7 environment.
- 2.1.3 In a separate Python 3.7 environment, install the following packages: matplotlib, RDKit, pandas, and scikit-learn.

3. Methods

This chapter describes two ML methods that can predict drug-target residence time. The first is a ligand only method, in other words, features from the ligand alone are used to train the model. The second method incorporates features of the ligand and its protein target. Structural data or a high-quality homology model is required to carry out the second method.

3.1 Kinetic binding data

For either method, one requires training data to train the supervised ML models. This needs to be experimentally-determined ligand residence time values for a specified target(s). One can obtain ligand kinetic data from ChEMBL [31] by searching by activity type, however, the data is much sparser than for binding affinity endpoints and is often far from a comprehensive source of all of the kinetic data available in the literature. A recently published database, KOFFI-DB (<http://koffidb.org>), contains kinetic parameters of ligand binding derived from surface plasmon resonance. Currently approximately 1000 k_{off} values are available in this database. By far, the largest collection of kinetic data is the KIND (KINetic Dataset) which contains 3812 entries collated from 21 publications and data produced by the EU-IMI consortium K4DD (<http://k4dd.eu>) for a wide range of targets including ion channels, kinases and GPCRs [24]. Even though this may seem impressive, it is vastly smaller than equivalent databases for binding affinity. For single targets, kinetic ligand binding data is quite sparse, generally fewer than 100 entries. The paucity of kinetic data is one of the main challenges of using ML to predict residence time.

In the examples shown in this chapter, GPCR data scraped from publications and manually curated will be used. These 536 entries are approximately double the number of entries found in the KIND for GPCRs [24] and can be downloaded from <https://potterton48.github.io> (see **Note 2**). Temperature corrected residence time values were calculated to minimise the impact of the difference in temperatures used by different research groups in different publications when performing the kinetic assay. All residence time values were corrected to the mean temperature used across all experiments (294.15 Kelvin) using Arrhenius' equation (see Equation 1). As the change in temperature was small (under 15 Kelvin), the frequency factor (A) was assumed to be constant.

$$k_{off} = Ae^{-\frac{E_A}{RT}} \quad (Eq\ 1)$$

It is important to make sure that the training data are relatively representative of the test compounds. For this reason, the peptide entries for the NK₁ receptor were removed from the example GPCR training set.

3.2 QSKR on ligand-only features for a single target

The first method, the ML model trained on ligand features alone, is a QSKR multi-linear regression model for a single target. An open-source Python library, PyQSAR [29], will be used to carry out the QSKR modelling.

- 3.2.1 Create a table with three columns: the compound name, the compound SMILES string and the associated experimentally-determined, temperature-corrected residence time. These data will be used as the training and testing data of the QSKR model.
- 3.2.2 Investigate the spread of the residence time values, does it follow a normal distribution? If it does not, you may want to transform the data to achieve this. In the GPCR example, a \log_{10} transformation was used for this purpose.

- 3.2.3 Launch the python 2.7 environment with “conda activate py2” on macOS/Linux or “activate py2” on Windows, in order to use PyQSAR. Then launch a Jupyter Notebook session.
- 3.2.4 Import the following libraries: pandas, numpy, Mordred, RDKit, multiprocessing, pyqsar, and scikit-learn.
- 3.2.5 Use the pandas library to load the table with the SMILES strings and associated residence time values into Python 2.7.
- 3.2.6 To generate features of the compounds using Mordred, RDKit molecules first must be created for each of the compounds from their SMILES strings.

```
mols = [Chem.MolFromSmiles(mol) for mol in \
df['SMILES'].values.tolist()]
```

This command loops through each value of the “SMILES” column in the DataFrame and generates a RDKit molecule from that, storing it in a list.

- 3.2.7 Generate approximately 1500 ligand features using Mordred. This may take some time, but approximately 10 compounds will be processed per second on a single CPU. Mordred is freely available, but programs such as Dragon [32] can be used to generate even more ligand features (*see Note 3*).
- 3.2.8 Remove features with no values generated. This can happen when a descriptor does not apply to a particular ligand. If only a limited number of ligands are missing values for a given descriptor, the data could be filled in with, for example, the median value.
- 3.2.9 Scale the data within each column using Scikit-learn.
- 3.2.10 Now split the data into training and test (hold-out) sets. Typically, an 80:20 split is used to maximise the amount of data for training. There are several different methods one can employ to split the data, the simplest being a random split (*see Note 4*).
- 3.2.11 Cluster the features to find highly correlated features reducing the number of overall features that need to be searched using the genetic algorithm.

- 3.2.12 Carry out feature selection using a genetic algorithm, selecting only a single feature from a cluster in order to prevent highly correlated features being selected at the same time. The target information and the experimentally-determined residence time will need to be provided. The “components” argument in the feature selection function determines the final number of features to be included in the QSKR model (see **Note 5**). This step may take several minutes to hours depending on the number of ligands and computing power. This command will return the names of the selected features that have the best predictive power.
- 3.2.13 Save both the test and training data for the features that have been selected as .csv/pickle files. Also save the experimentally-determined, temperature-corrected residence time values as a .csv/pickle files.
- 3.2.14 Open a Jupyter Python 3 Notebook, by closing the Python2.7 environment “conda deactivate py2” and then relaunching Jupyter Notebook. Load the following modules: scikit-learn, pandas, matplotlib.
- 3.2.15 Load the saved feature data and target data (experimentally-determined, temperature-corrected residence time) into Python into two separate pandas’ DataFrames.
- 3.2.16 Train the multilinear regression model on the training data using the features selected in the feature selection stage. Use that trained model to predict values for the test set. The quality metrics that will be applied are the root mean squared error (RMSE) and R^2 (see **Note 6**). For the A_1 receptor kinetic data, the following results were achieved: $RMSE_{Train} = 0.30$, $RMSE_{Test} = 0.48$, $R^2_{Train} = 0.76$, $R^2_{Test} = 0.67$.
- 3.2.17 Plot the results using Matplotlib or any other plotting software/package (see Figure 2 for the results for the A_1 receptor QSKR model). By investigating outliers on the plot, one can determine why the model fails in some instances and what can be done to improve it.

<Figure 2 here>

3.3 Multi-target QSKR model trained on features obtained from molecular dynamics simulations

In cases where there is insufficient data with which to develop a single-target QSKR model, multi-target modelling can be used to increase the amount of data. Multi-target models require some information of the protein either through explicit representation of the protein the ligand is targeting (e.g. by inputting the protein sequence) or through interactions made to the protein by the ligand. These models move towards more generalised models for predicting ligand kinetic rates as opposed to models only applicable to single receptors or to a single ligand series for a single target.

To extend the COMBINE workflow mentioned in the Introduction [27] to a multi-target model, protein family numbering schemes can be used to find equivalent residues to feed into the model as training data. For example, for GPCRs, one could use the GPCRdb's modified Ballesteros and Weinstein numbering scheme [33] to find ligand energy values (VdWs and electrostatics) to specific residue positions (see Figure 3).

<Figure 3 here>

One of the issues of using protein-ligand interactions to train multi-target models is obtaining sufficient data for training the model. The PDBbind database [34] addressed this issue for binding affinities by collating all protein-ligand structures in the PDB with associated binding affinity data. Due to the lack of ligand kinetic data, the overlap between ligand kinetic data and associated PDB structures is very small. Therefore, one has to use predicted structures of the protein-ligand structures to increase the data. To increase the reliability of these predicted poses (dockings), one can use ensembles of short molecular dynamics (MD) simulations following docking. Training models on MD data [35] has been shown to give good performance in predicting values such as log P [36]. A short outline of a method to develop a multi-target QSKR model trained on MD simulation data is described below.

3.3.1 Obtain protein-ligand starting structures. In preference, use X-ray or Cryo-EM structures as starting structures. Failing that, one should use docking to predict ligand binding poses. Bear

in mind, that the quality of the resulting data obtained (output) heavily depends on the quality of the starting structure (input).

- 3.3.2 Perform high-throughput molecular dynamics simulations, using a setup that is as automated as possible. There are some tools that can help with this such as HTMD [37]. The goal is to perform ensembles of simulations to ensure reproducibility of the results. These ensembles are replica simulations where the only difference is the starting velocities assigned to each atom.
- 3.3.3 Obtain properties from the simulations such as hydrogen bonds, RMSD, RMSF, protein-ligand interaction fingerprints. VMD [38], Chimera [39] or MDAnalysis [40] can help extract these features endeavouring to obtain features appropriate for the system and feature to be modeled (*see Note 7*)
- 3.3.4 Apply several different ML methods, using the splitting strategies outlined in this chapter. Evaluate which method (and hyperparameters) are best using a validation set or by k-means validation.

4. Notes

- 4.1 PyQSAR [29] is a Python 2.7 library. As Python 2.7 has already passed its “end of life” and thus is no longer properly supported, only work that needs to be should be carried out in Python 2.7 (because it calls upon PyQSAR). Python 3 should be used as the default for everything else. To support these different Python environments as easily as possible, Anaconda should be used.
- 4.2 All published GPCR-ligand kinetic data was collected into a database. The primary data obtained from each published was the ligand name, SMILES string, the temperature the kinetic assay was performed at, k_{on} , k_{off} , K_D and K_i . Room temperature was assumed to be 294.15 Kelvin. Residence time was calculated as the inverse of k_{off} .

4.3 Other features of the ligand can be used, such as extended-connectivity fingerprints (ECFP) [41].

Multiple descriptors can be tested to see which one gives the best performance for the task in question.

4.4 For any kind of split, it is important to ensure that the target value (residence time in the present case) has a similar distribution in both the training and test sets. With a random split, it is easy to overestimate the power of the model due to compounds in the test set that are very similar to the training set in structure and associated target value. A split that takes ligand structure into account can be performed to avoid this. Time-based splitting is another option that can recreate the scenario of a drug-discovery project.

4.5 Most QSAR models have fewer than 10 features. In the example shown, four have been chosen but the number of features can be varied to determine what is best for a given system. In general, the goal is to use the minimum number of features that give rise to fairly good accuracy in the training data to reduce the chances of overfitting. Ideally, if data permits, a validation set would be used to investigate how many features should be included in the model for optimum performance.

4.6 RMSE has the advantage of being in the same units as the target data, so is easier to understand. R^2 gives an indication whether the model performs better or worse than random predictions. The bigger the difference between training and test RMSE, the more likely that overfitting is occurring.

4.7 For the GPCR QSKR model, the following features were calculated because they had some level of experimental evidence that they contribute to residence time: the interaction energy between water and the ligand, RMSD of the second extracellular loop of the GPCR, RMSD of the binding site, and measures of size of the ligand.

Acknowledgments

This work was supported by the Biotechnology and Biological Sciences Research Council [grant number BB/M009513/1] and the London Interdisciplinary Bioscience PhD Consortium (LIDo). AH and ATN are grateful for funding from the EU H2020 CompBioMed (grant number 675451).

5. References

1. Swinney DC (2004) Biochemical mechanisms of drug action: What does it take for success? *Nat Rev Drug Discov* 3:801–808 . <https://doi.org/10.1038/nrd1500>
2. Sykes DA, Dowling MR, Charlton SJ (2009) Exploring the mechanism of agonist efficacy: a relationship between efficacy and agonist dissociation rate at the muscarinic M3 receptor. *Mol Pharmacol* 76:543–551
3. Copeland RA (2016) The drug–target residence time model: a 10-year retrospective. *Nat Drug Discov* 15:87–95
4. Guo D, Mulder-Krieger T, IJzerman AP, Heitman LH (2012) Functional efficacy of adenosine A2A receptor agonists is positively correlated to their receptor residence time. *Br J Pharmacol* 166:1846–1859
5. Bosma R, Witt G, Vaas LAI, Josimovic I, Gribbon P, Vischer HF, Gul S, Leurs R (2017) The target residence time of antihistamines determines their antagonism of the G protein-coupled histamine H1 receptor. *Front Pharmacol* 8:1–15 . <https://doi.org/10.3389/fphar.2017.00667>
6. Vauquelin G, Charlton SJ (2010) Long-lasting target binding and rebinding as mechanisms to prolong in vivo drug action. *Br J Pharmacol* 161:488–508
7. Dowling MR, Charlton SJ (2006) Quantifying the association and dissociation rates of unlabelled antagonists at the muscarinic M 3 receptor. *Br J Pharmacol* 148:927–937 . <https://doi.org/10.1038/sj.bjp.0706819>
8. Copeland RA, Pompliano DL, Meek TD (2006) Drug–target residence time and its implications for lead optimization. *Nat Drug Discov* 5:730–739
9. Guo D, Dijksteel GS, Van Duijl T, Heezen M, Heitman LH, IJzerman AP (2016) Equilibrium and kinetic selectivity profiling on the human adenosine receptors. *Biochem Pharmacol* 105:34–41 . <https://doi.org/10.1016/j.bcp.2016.02.018>
10. Dahl G, Akerud T (2013) Pharmacokinetics and the drug-target residence time concept. *Drug Discov Today* 18:697–707 . <https://doi.org/10.1016/j.drudis.2013.02.010>
11. Bradshaw JM, McFarland JM, Paavilainen VO, Bisconte A, Tam D, Phan VT, Romanov S, Finkle D, Shu J, Patel V, Ton T, Li X, Loughhead DG, Nunn PA, Karr DE, Gerritsen ME, Funk JO, Owens TD, Verner E, Brameld KA, Hill RJ, Goldstein DM, Taunton J (2015) Prolonged and tunable residence time using reversible covalent kinase inhibitors. *Nat Chem Biol* 11:525–531 . <https://doi.org/10.1038/nchembio.1817>
12. Schuetz DA, de Witte WEA, Wong YC, Knasmueller B, Richter L, Kokh DB, Sadiq SK, Bosma R, Nederpelt I, Heitman LH, Segala E, Amaral M, Guo D, Andres D, Georgi V, Stoddart LA, Hill S, Cooke RM, De Graaf C, Leurs R, Frech M, Wade RC, de Lange ECM, IJzerman AP, Müller-Fahrnow A, Ecker GF (2017) Kinetics for Drug Discovery: an industry-driven effort to target drug residence time. *Drug Discov Today* 22:896–911 . <https://doi.org/10.1016/j.drudis.2017.02.002>
13. Šrajcar V, Schmidt M (2017) Watching proteins function with time-resolved X-ray crystallography. *J Phys D* 50:1–53 . <https://doi.org/10.1016/j.physbeh.2017.03.040>
14. Miller DC, Lunn G, Jones P, Sabnis Y, Davies NL, Driscoll P (2012) Investigation of the

- effect of molecular properties on the binding kinetics of a ligand to its biological target. *RSC Med Chem* 3:449–452 . <https://doi.org/10.1039/c2md00270a>
15. Tresadern G, Bartolome JM, MacDonald GJ, Langlois X (2011) Molecular properties affecting fast dissociation from the D2 receptor. *Bioorganic Med Chem* 19:2231–2241
 16. Kokh DB, Amaral M, Bomke J, Grädler U, Musil D, Buchstaller HP, Dreyer MK, Frech M, Lowinski M, Vallee F, Bianciotto M, Rak A, Wade RC (2018) Estimation of drug-target residence times by τ -random acceleration molecular dynamics simulations. *J Chem Theory Comput* 14:3859–3869 . <https://doi.org/10.1021/acs.jctc.8b00230>
 17. Schmidtke P, Javier Luque F, Murray JB, Barril X (2011) Shielded hydrogen bonds as structural determinants of binding kinetics: Application in drug design. *J Am Chem Soc* 133:18903–18910
 18. Bortolato A, Tehan BG, Bodnarchuk MS, Essex JW, Mason JS (2013) Water network perturbation in ligand binding: Adenosine A2A antagonists as a case study. *J Chem Inf Model* 53:1700–1713 . <https://doi.org/10.1021/ci4001458>
 19. Potterton A, Hussein FS, Southey MWY, Bodkin MJ, Heifetz A, Coveney P V, Townsend-Nicholson A (2019) Ensemble-based steered molecular dynamics predicts relative residence time of A2A receptor binders. *J Chem Theory Comput* 15:3316–3330 . <https://doi.org/10.1021/acs.jctc.8b01270>
 20. Guo D, Pan AC, Dror RO, Mocking T, Liu R, Heitman LH, Shaw DE, IJzerman AP (2016) Molecular basis of ligand dissociation from the adenosine A2A receptor. *Mol Pharmacol* 89:485–491 . <https://doi.org/10.1124/mol.115.102657>
 21. Hegde SS, Pulido-Rios MT, Luttmann MA, Foley JJ, Hunsberger GE, Steinfeld T, Lee TW, Ji Y, Mammen MM, Jasper JR (2018) Pharmacological properties of revefenacin (TD-4208), a novel, nebulized long-acting, and lung selective muscarinic antagonist, at human recombinant muscarinic receptors and in rat, guinea pig, and human isolated airway tissues. *Pharmacol Res Perspect* 6:1–11 . <https://doi.org/10.1002/prp2.400>
 22. Dowling MR, Charlton SJ (2006) Quantifying the association and dissociation rates of unlabelled antagonists at the muscarinic M3 receptor. *Br J Pharmacol* 148:927–937
 23. Jakubík J, Randáková A, Zimčík P, El-Fakahany EE, Doležal V (2017) Binding of N-methylscopolamine to the extracellular domain of muscarinic acetylcholine receptors. *Sci Rep* 7:40381 . <https://doi.org/10.1038/srep40381>
 24. Schuetz DA, Richter L, Martini R, Ecker GF (2020) A structure-kinetic relationship study using matched molecular pair analysis. *RSC Med Chem* 11:1285–1294 . <https://doi.org/10.1039/d0md00178c>
 25. Thorsen TS, Matt R, Weis WI, Kobilka BK (2014) Modified T4 lysozyme fusion proteins facilitate G protein-coupled receptor crystallogenesis. *Structure* 22:1657–1664 . <https://doi.org/10.1016/j.str.2014.08.022>
 26. Qu S, Huang S, Pan X, Yang L, Mei H (2016) Constructing interconsistent, reasonable, and predictive models for both the kinetic and thermodynamic properties of HIV-1 protease inhibitors. *J Chem Inf Model* 56:2061–2068 . <https://doi.org/10.1021/acs.jcim.6b00326>
 27. Ganotra GK, Wade RC (2018) Prediction of drug-target binding kinetics by comparative binding energy analysis. *ACS Med Chem Lett* 9:1134–1139 . <https://doi.org/10.1021/acsmchemlett.8b00397>
 28. Kokh DB, Kaufmann T, Kister B, Wade RC (2019) Machine learning analysis of τ RAMD trajectories to decipher molecular determinants of drug-target residence times. *Front Mol Biosci* 6:1–17 . <https://doi.org/10.3389/fmolb.2019.00036>
 29. Kim S, Cho KH (2019) PyQSAR: A fast QSAR modeling platform using machine learning and jupyter notebook. *Bull Korean Chem Soc* 40:39–44 . <https://doi.org/10.1002/bkcs.11638>

30. Moriwaki H, Tian YS, Kawashita N, Takagi T (2018) Mordred: A molecular descriptor calculator. *J Cheminform* 10:1–14 . <https://doi.org/10.1186/s13321-018-0258-y>
31. Bento AP, Gaulton A, Hersey A, Bellis LJ, Chambers J, Davies M, Krüger FA, Light Y, Mak L, McGlinchey S, Nowotka M, Papadatos G, Santos R, Overington JP (2014) The ChEMBL bioactivity database: An update. *Nucleic Acids Res* 42:1083–1090 . <https://doi.org/10.1093/nar/gkt1031>
32. Mauri A, Consonni V, Pavan M, Todeschini R (2006) Dragon Software: An Easy Approach To Molecular Descriptor Calculations. *Match Commun Math Comput Chem* 56:237–248 . <https://doi.org/10.1016/C2012-0-02727-5>
33. Isberg V, De Graaf C, Bortolato A, Cherezov V, Katritch V, Marshall FH, Mordalski S, Pin JP, Stevens RC, Vriend G, Gloriam DE (2015) Generic GPCR residue numbers - Aligning topology maps while minding the gaps. *Trends Pharmacol Sci* 36:22–31 . <https://doi.org/10.1016/j.tips.2014.11.001>
34. Wang R, Fang X, Lu Y, Wang S (2004) The PDBbind Database: Collection of Binding Affinities for Protein–Ligand Complexes with Known Three-Dimensional Structures. *J Med Chem* 47:2977–2980 . <https://doi.org/10.1021/jm030580l>
35. Riniker S (2017) Molecular Dynamics Fingerprints (MDFP): Machine Learning from MD Data to Predict Free-Energy Differences. *J Chem Inf Model* 57:726–741 . <https://doi.org/10.1021/acs.jcim.6b00778>
36. Wang S, Riniker S (2020) Use of molecular dynamics fingerprints (MDFPs) in SAMPL6 octanol–water log P blind challenge. *J Comput Aided Mol Des* 34:393–403 . <https://doi.org/10.1007/s10822-019-00252-6>
37. Doerr S, Harvey MJ, Noé F, De Fabritiis G (2016) HTMD: High-Throughput Molecular Dynamics for Molecular Discovery. *J Chem Theory Comput* 12:1845–1852 . <https://doi.org/10.1021/acs.jctc.6b00049>
38. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. *J Mol Graph* 14:33–38 . [https://doi.org/10.1016/0263-7855\(96\)00018-5](https://doi.org/10.1016/0263-7855(96)00018-5)
39. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera - A visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612 . <https://doi.org/10.1002/jcc.20084>
40. Michaud-Agrawal N, Denning EJ, Woolf TB, Beckstein O (2011) MDAAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J Comput Chem* 32:2319–2327 . <https://doi.org/https://doi.org/10.1002/jcc.21787>
41. Rogers D, Hahn M (2010) Extended-Connectivity Fingerprints. *J Chem Inf Model* 50:742–754 . <https://doi.org/10.1021/ci100050t>

Figure Captions

1. **Correlation between ligand size and drug-target residence time.** Scatter plots showing the correlation between (top) the drug-target residence times of 500 GPCR ligands and ligand molecular weight (MW) and (bottom) between residence time and the number of rings in the ligand (no_rings). The least squares linear regression lines are shown as a dashed orange lines, the strength of those correlations is shown as R^2 .

2. **Correlation plot of predicted and experimentally-determined (Expt.) residence time (RT) for the QSKR model on A₁ receptor kinetic ligand data.** The blue, open circles are the training data and the orange, opaque triangles are the test data.
3. **Using protein family numbering schemes to find equivalent protein-ligand interactions.** This figure demonstrates how one can use protein family numbering schemes, in this case the GPCRdb's numbering scheme for GPCRs, to find equivalent residues to assign interaction energies to. This example shows two equivalent residues from the first helix the A₁ (PDB accession number: 5UEN) and the A_{2A} receptor (PDB accession number: 3PWH) in blue and peach, respectively. The small heatmap shows example interactions energies, Van der Waals (VdW) and electrostatics (Elect), for these two residue positions.

Table Captions

1. **Summary of the published ML methods that predict drug-target residence time.** ^aThe accuracy on all compounds, excluding 2 outliers, as reported in the original manuscript. ^bDue to the limited number of compounds, the authors performed “leave-one-out” validation in lieu of using a separate test set. ^cNo test set R² accuracy was reported; the mean absolute error of the training set was reported to be 0.48. ^dThis is, strictly speaking, the Q²_{F3} and not the R² of the test set.

Tables

Table 1

Name	ML method	Protein	No. Compounds	Accuracy (R ²)
VolSurf QSKR [26]	PLS regression	HIV-1 protease	37 (28/9)	0.57/0.65
COMBINE [27]	PLS regression	HIV-1 protease	36	0.94 ^a /0.70 ^b
COMBINE [27]	PLS regression	HSP90	70 (57/13)	0.80/0.86 ^b
τRAMD FP [28]	SV regression	HSP90	94 (76/18)	NA ^c /0.56 ^d

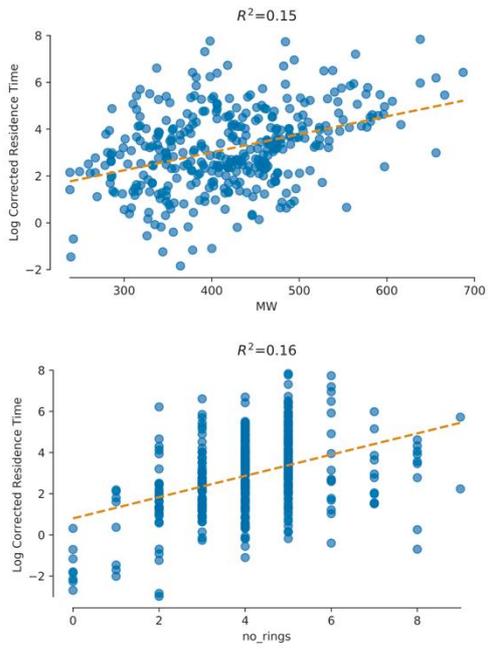


Figure 1

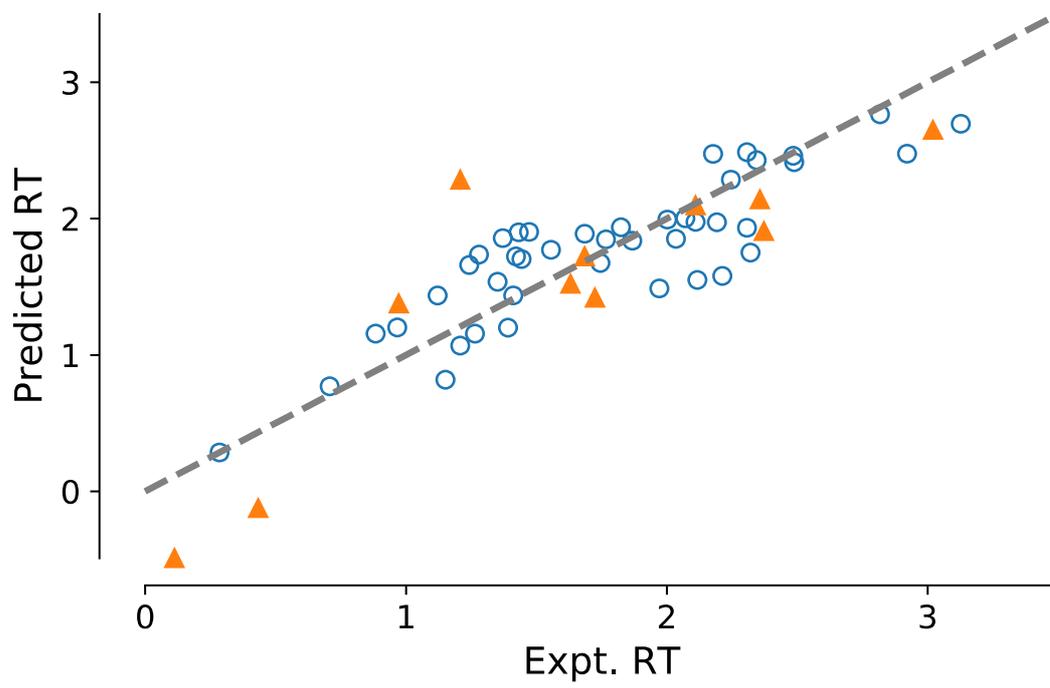


Figure 2

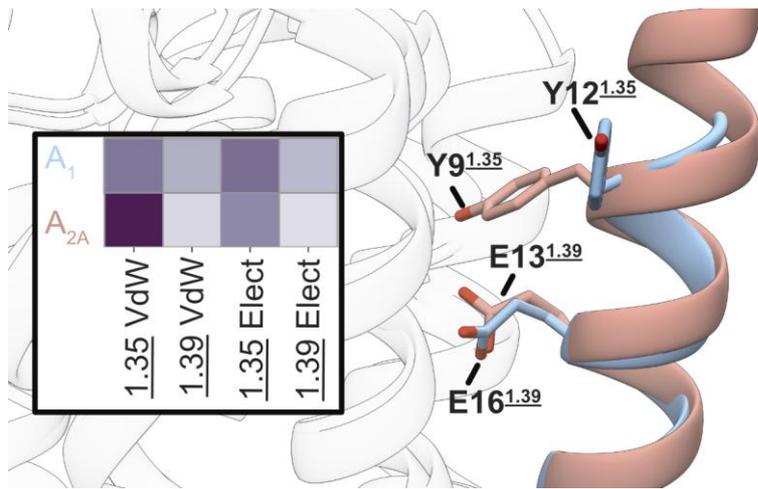


Figure 3