Supporting Information for

Tools

In silico off-target profiling for enhanced drug safety assessment

Jin Liu[a,b,†], Yike Gui[d,b,†], Jingxin Rao[b,c], Jingjing Sun[b,c], Gang Wang[b,c], Qun Ren[d,b], Ning Qu[b,c], Buying Niu[b,c], Zhiyi Chen[e,b,c], Xia Sheng[b,c], Yitian Wang[b,c], Mingyue Zheng[a,b,c,d,e,*], Xutong Li[b,c,*]

[a] *College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, China*

[b] *Drug Discovery and Design Center, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica Chinese Academy of Sciences, Shanghai 201203, China*

[c] *University of Chinese Academy of Sciences, Beijing 100049, China*

[d] *Nanjing University of Chinese Medicine, Nanjing 210023, China*

[e] *School of Pharmaceutical Science and Technology, Hangzhou Institute for Advanced Study, Hangzhou 330106, China*

[†]These authors made equal contributions to this work.

Received 28 November 2023; received in revised form 21 February 2024; accepted 29 February 2024

*Corresponding authors.

E-mail addresses: myzheng@simm.ac.cn (Mingyue Zheng), lixutong@simm.ac.cn (Xutong Li).

Supporting Information of Contents

Text S1. Establishment of off-target panel

In vitro analyzing of candidate drugs against a wide range of targets is an important component of the compound selection process. Compared to screening compounds on known therapeutic targets, promiscuous screening (off-target screening) of compounds is challenging due to the lack of boundaries. Conducting extensive screening on numerous targets can be cost-prohibitive and it may be challenging to predict the relevance of compound-target interactions [1]. To mitigate these challenges, major pharmaceutical companies such as AstraZeneca, GlaxoSmithKline, Novartis, and Pfizer have established internal off-target panels for screening compounds, which reduce the number of molecules tested in subsequent assays. Based on the targets established by these companies, Bowes et al. proposed 44 early drug safety targets that include the toxicity of the central nervous system, immune system, gastrointestinal tract, and heart [2]. AbbVie obtained 70 safety-related targets via a literature search, most of which are included in Eurofins' testing panel [3]. To obtain an optimized target panel, Roche utilized experimental data based on the Bioprint® database [4] and employed a statistical ranking method, resulting in a panel of 50 safety targets [1]. Combining the above targets, we can get an off-target panel containing 90 targets. This panel comprises 50 G protein-coupled receptor protein-related targets, 16 Ion channel-related targets, 9 Enzyme-related receptors, 6 Kinase-related targets, 4 Nuclear receptor-related targets, 4 Transporter-related targets and two other targets.

Text S2. Compound-protein interaction data collection and processing

The data collection and processing process for the ChEMBL database is below, and the PubChem database is collected similarly. According to the gene names of the targets, the steps to collect the corresponding compounds under the corresponding targets from UniProt and ChEMBL databases are as follows：

1. The Retrieve/ID mapping module in UniProt was utilized to convert gene names of safety targets to UniProtKB/Swiss Prot identifiers. The UniProt IDs assigned to this entry were manually annotated with high credibility and low redundancy. The organisms mainly selected were Human, Rat, and Mouse.
2. The corresponding ChEMBL ID in the ChEMBL database (Release 30) was obtained based on the UniProt ID of the safety target.
3. Relevant data was retrieved from the ChEMBL database using PostgreSQL, a powerful relational database management system (RDBMS) that supports crawling data from databases based on filtering conditions. The data filtering condition set for this task is:

（1）target_type = 'SINGLE PROTEIN'

The target types that ChEMBL supports include single protein, protein complex, and protein family. Here, we had chosen single protein.

（2）molecule_type = 'Small molecule'

The selectable molecule types in ChEMBL include oligosaccharides, proteins, small molecules, etc. Here we chose small molecule.

（3）assay_type IN 'B','F'

There are multiple experimental categories supported by ChEMBL, and we had chosen to use the binding ('B') and function ('F') data.

（4）confidence_score = 9

ChEMBL assigns confidence scores to the assay-to-target relationships, reflecting the

reliability of a specific experimental result for a given target type. The scores range from 0 to 9, with a minimum confidence of 4 for known protein targets. We only selected confidence score of 9, ensuring precise experimental data for individual protein targets.

Activity thresholds were determined based on the target protein family, using the guidelines provided by the Illuminating the Druggable Genome consortium (IDG) (https://druggablegenome.net/ProteinFam) [5]. The specific thresholds for each protein family are summarized in Supporting Information Table S3.

Text S3. Principles of enrichment analysis

The *p-value* is the probability of obtaining the observed sample results, or more extreme results, when the null hypothesis is true. If the p-value is very small, it indicates that the probability of such a situation occurring is very low. If it does occur, according to the principle of small probability, we have reason to reject the null hypothesis. The smaller the *p-value* is, the more sufficient our reason to reject the null hypothesis, meaning our results are more significant. In our enrichment analysis, the *p-value* is calculated using the following formula:

$$p-value=\sum_{i=m}^{M}\frac{\binom{M}{i}*\binom{N-M}{n-i}}{\binom{N}{n}}=1-\sum_{i=0}^{m-1}\frac{\binom{M}{i}*\binom{N-M}{n-i}}{\binom{N}{n}} \quad (1)$$

In Eq. (1), the denominator is the process of binding n targets from a total of N targets for drugs. The numerator is M targets denoting the number of targets belonging to the specific ADR term, of which, i is predicted as positive targets for the drug. The p-value refers to the sum of probabilities of observing m targets in a certain ADR term or more extreme results, so i ranges from m to M. This represents all possibilities of seeing more targets fall into this ADR term. As the number of hypothesis tests increases, the probability of making error decisions rapidly increases. To control the false positive rate, we have performed multiple hypothesis testing correction, namely *FDR* (False Discovery Rate, *FDR*) correction (Eq. (2)) which is based on Bonferroni correction. Multiple testing correction adjusts each *p-value* to keep the overall error rate less than or equal to the user-specified *p-cutoff value*.

$$FDR=1-precision=\frac{FP}{(TP+FP)} \quad (2)$$

Text S4. Introduction of MLKNN and the data splitting method for ATC classification

Multi-label learning algorithms can be roughly divided into two categories: transformation algorithms and adaptive algorithms. Here, we used an adaptive algorithm, multi-label k-nearest neighbours (MLKNN), to perform multi-label classification [6, 7]. The MLKNN algorithm follows four steps: 1. Finding the K nearest samples to the input sample; 2. Counting the occurrences of each category in the K samples; 3. Utilizing a Bayesian algorithm with the statistics from step 2 to calculate the label probabilities; 4. Outputting class probabilities. In our project, we have a dataset of n drugs represented by$\{(x_i,y_i)\}_{i=1}^{n}$, where x_i and y_i are m-dimensional feature vectors and n-dimensional side effect vectors for drug i, respectively. MLKNN is used to model the relationship between feature vectors and side effect vectors, i.e., $Y=F(X)\colon 2^m\rightarrow 2^n$.

To conduct multi-label ATC classification, we processed the collected drug-ATC code pairs into a multi-label data format, where a single drug may correspond to multiple ATC code labels.

Initially, the test dataset was divided using the iterative hierarchical partitioning method of data in multi-label learning, implemented through the iterative_train_test_split function in skmultilearn.model_selection, with a scale of 0.1. The remaining data was subsequently split into five folds for cross-validation, yielding training and validation datasets for each fold.

Text S5. Introduction of LightGBM, data processing and hyperparameter search

LightGBM, in comparison to XGBoost, primarily adopts some optimization strategies to obtain superior performance in terms of prediction accuracy, model stability, and computational efficiency [8]:

1) Histogram-based decision tree algorithm, which constructs a histogram of width K by discretizing continuous floating-point eigenvalues to find the optimal segmentation point to reduce the computation.
2) Gradient-based One-Side Sampling (GOSS), which retains data instances with large gradients and randomly sampling instances with small gradients to reduce the data volume while maintaining the data distribution.
3) Exclusive Feature Bundling (EFB) algorithm, which achieves dimensionality reduction by bundling a set of mutually exclusive features into one feature.
4) Leaf-wise leaf growth strategy with depth restriction, which improves convergence speed by paying more attention to data with large errors, compared to the traditional level-wise growth strategy.

To conduct toxicity prediction, after data cleaning and deduplication, we obtained a total of 877 toxic compounds and 1229 non-toxic compounds where we labeled toxic drugs as 1, and non-toxic compounds as 0. We first divided the test set at a scale of 0.2. The remaining data was subsequently split into five folds for cross-validation, yielding training and validation datasets for each fold. Different models were evaluated on the same test set.

Optuna is a hyperparameter optimization framework for automated hyperparameter search that can be applied to machine learning and deep learning models [9]; Optuna uses a sampling and pruning algorithm to optimize the hyperparameters, so it is very fast and efficient; it also dynamically constructs the hyperparameter search space in an intuitive way. We use Optuna here to select the best hyperparameters for the model. The hyperparameter settings of LightGBM and ECFP_LightGBM are shown in Supporting Information Table S11.

Text S6. Drug-ADR data collection and establishment of ADR prediction model

The SIDER database (version 4.1, 2015) served as the primary source for drug-ADR data collection. We retained ADR information for 1178 drugs, focusing on frequently occurring ADRs with frequencies ranging from 0.01 to 1, utilizing Preferred Terms (PT)-level descriptions to prevent semantic redundancy. Consequently, we retained 2551 ADRs for 1178 drugs, where '1' represents drug-ADR interactions and '0' represents no interaction.

ADR prediction was conducted using a multi-label KNN model proposed by Zhang et al. [7], with input features being the predicted results of the off-target profiles or molecular ECFP4. For drug-ADR interaction data, the test dataset underwent division using the iterative hierarchical partitioning method of data in multi-label learning, implemented through the iterative_train_test_split function in skmultilearn.model_selection, with a scale of 0.1. Employing

a fixed test set, five-fold cross-validation was performed, and the best parameter combination obtained through hyper-parameter search was k=5 and s=1, where k represents the number of neighbors in the KNN model, and s is a smoothing parameter. In each round of five-fold cross-validation, we calculated the optimal threshold for each ADR. The optimal threshold, determined based on the Youden Index [10], which serves as a summary measure of the receiver operating characteristic curve. It is the sum of sensitivity and specificity minus one. In binary classification, the Youden Index can be employed to select the optimal classification threshold. We applied the optimal threshold for each category to convert predicted probability values into binary classification labels, thereby obtaining the predicted relevant ADRs for each drug.

Besides multi-label models, for each ADR, an RF classifier was established to determine whether a compound possesses this ADR, with the off-target representation or ECFP4 serving as input features. The performance for these models is presented in Supporting Information Table S18.

Text S7. Drug-target-network analysis of Sertindole

Sertindole, an atypical antipsychotic utilized for schizophrenia treatment, faced market withdrawal in 1998 across multiple countries due to cardiotoxic side effects, including arrhythmias and sudden cardiac death. In Supporting Information Fig. S5, we identified nine known targets of Sertindole that overlap with our off-target panel. Notably, our off-target model successfully predicted eight of these targets, attaining binding probabilities exceeding 0.7. On the one hand, the predicted new targets can help explain the known ADRs of drugs. Our predictions suggest Sertindole's interaction with SLC6A4 and HTR7. Inhibiting SLC6A4 [11] or binding to HTR7 [12] could induce insomnia, thereby furnishing a rationale for insomnia induced by Sertindole usage. Hyperglycemia, an infrequent side effect of Sertindole, finds potential association in our projected outcomes with the newly predicted targets HTR3A [13] and ADRB2 [14]. On the other hand, the drug's known binding targets can explain related ADRs. Sertindole's potential interaction with the new target HRH3 carries an affiliation with sedation [15]. Although sedation is linked to Sertindole's known target ADRA2B [16] and HRH1 [17], the prominence of HRH3 in our forecasts intimates a mechanistic link between the drug and new off-targets. Cardiotoxicity-related side effects of Sertindole can be associated with the known target KCNH2 [18], ADRA2B, as well as predicted new targets such as CHRNA4 and SCN2A, etc.

Text S8. Molecular docking using Schrödinger software

Virtual screening was carried out using Maestro (Schrödinger Release 2023-1: Glide, Schrödinger, LLC, New York, NY, 2023). The crystal structures of the targets were retrieved from the Protein Data Bank (PDB), and the PDB IDs for the utilized targets are presented in Supporting Information Table S18. Protein preparation was performed using Maestro's Protein Preparation Wizard with the OPLS3e force field. The prepared protein served as the basis for generating grid files for molecular docking through the Receptor Grid Generation model. The center of the ligand was used to define the receptor grid box's center, allowing docked molecules to sample a 10 × 10 × 10 Å inner search space, with a 20 Å buffer in all directions. All other parameters remained at default settings.

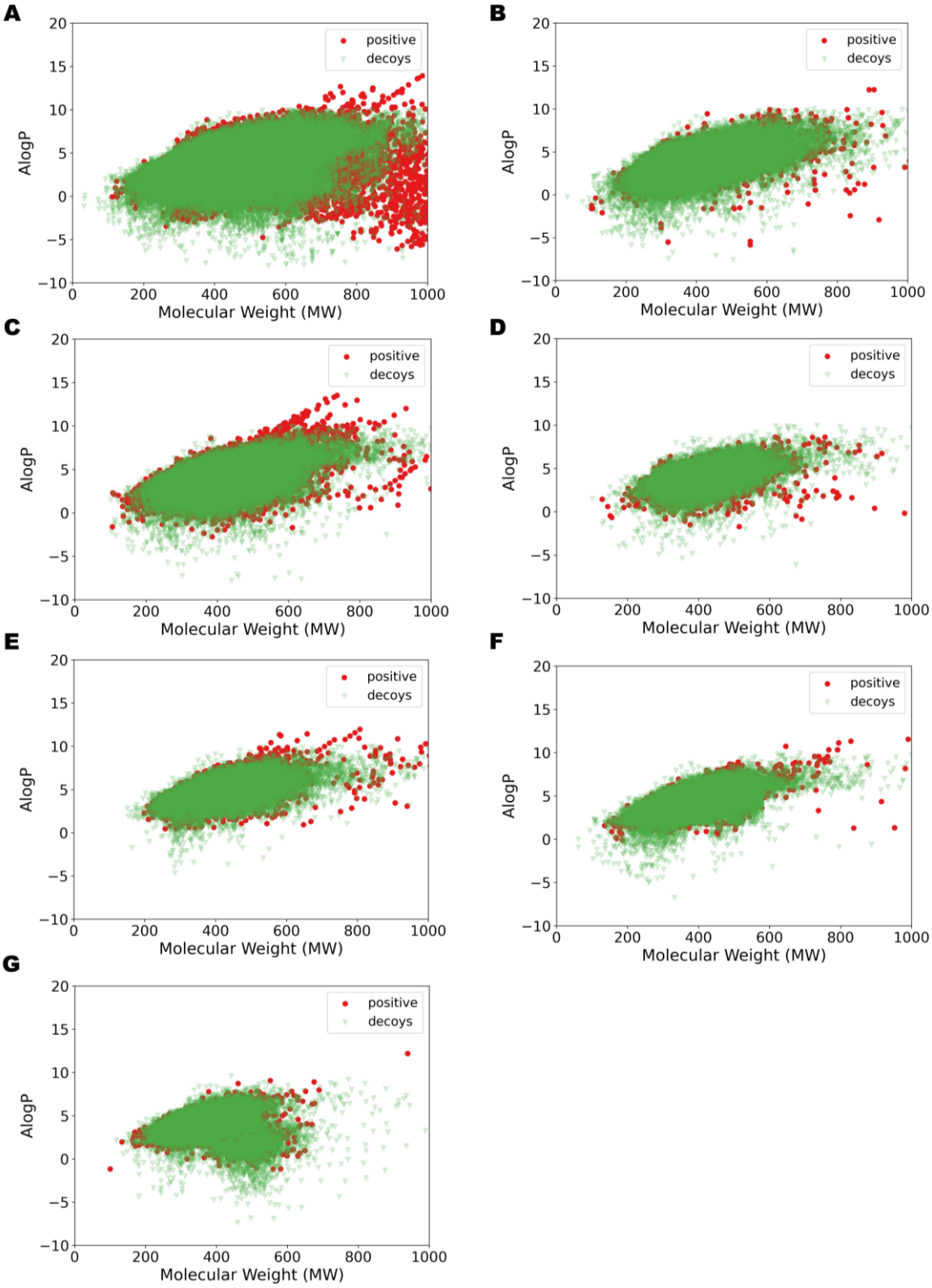

Fig. S1. Physico-chemical spatial distribution of positive molecules and relative decoys under different target classes. **(A)** GPCR. **(B)** Ion channel. **(C)** Enzyme. **(D)** Kinases. **(E)** Nuclear receptor. **(F)** Transporter. **(G)** Other targets.

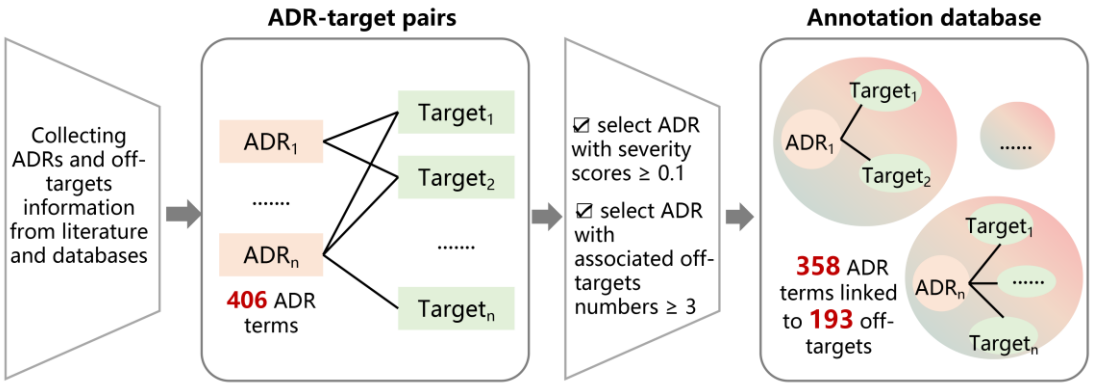

Fig. S2. Establishing ADR-targets mappings.

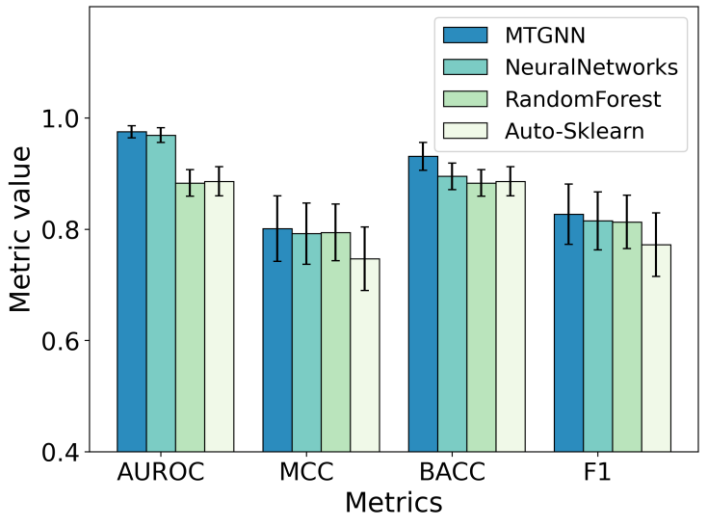

Fig. S3. Performance comparison of four off-target prediction models across all targets on AUROC, MCC, BACC, and F1 metrics.

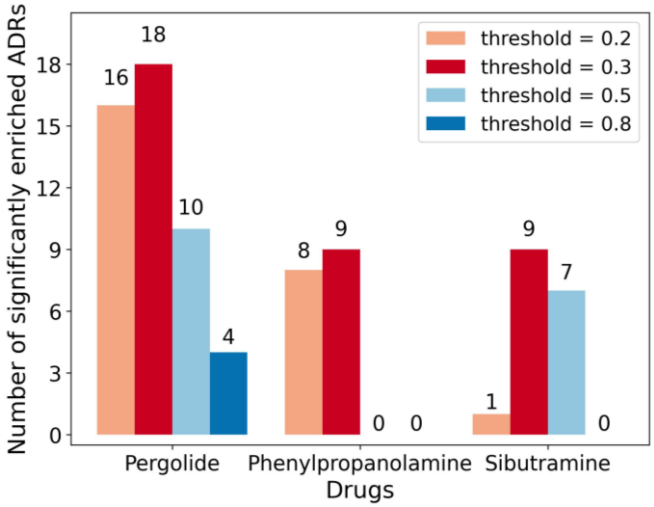

Fig. S4. Number of significantly enriched ADRs for three drugs under different thresholds. The horizontal axis represents drugs, the vertical axis represents the number of significantly enriched ADRs, and different colors indicate different thresholds.

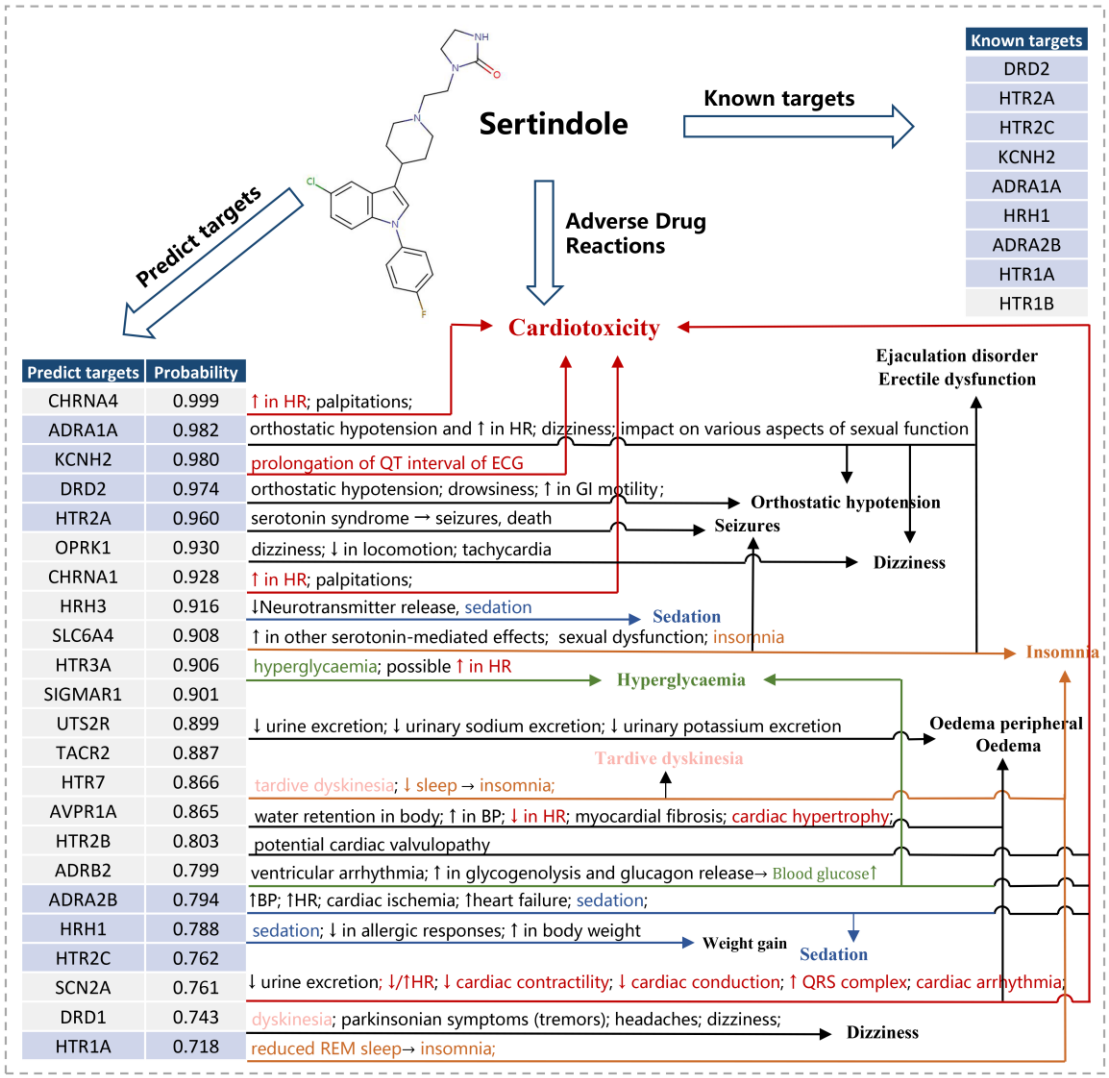

Fig. S5. Drug-target-ADR association diagram for Sertindole. Known targets in the figure are the known targets of the drug (obtained from databases such as ChEMBL, PubChem and DrugBank, which overlap with our off-target panel), Predict targets are the predicted off-targets, and probability is the predicted probability value. The targets with a blue background in the table are predicted targets that overlap with known targets. The side effect descriptions associated with each target are listed after it, and colored ADRs are associated with predicted new targets and correspond to the respective target's side effect description (e.g. Cardiotoxicity is marked in red, corresponding to " ↓ /↑HR", "prolongation of QT interval of ECG", "cardiac hypertrophy" and so on in the target's side effect description; Sedation is marked in blue, corresponding to "sedation" in the target's side effect description; Hyperglycaemia is marked in green, corresponding to "hyperglycaemia" and "Blood glucose ↑" in the target's side effect description). BP: blood pressure; HR: heart rate; GI: glycemic index; ECG: electrocardiography; QT: the time from the start of the Q wave to the end of the T wave, time taken for ventricular depolarisation and repolarisation.

Table S2 List of databases used in the work

Database	Description	URL	Reference
ChEMBL	A manually curated database of biologically active molecules with drug-like properties. It brings together chemical, bioactive and genomic data to help translate genomic information into effective new drugs.	https://www.ebi.ac.uk/chembl/	[19]
PubChem	An open chemical database provides information about molecules such as chemical structures, identifiers, chemical and physical properties, and biological activities.	https://pubchem.ncbi.nlm.nih.gov/	[20]
DrugBank	A web-based database that provides extensive molecular information on drugs, including their mechanisms of action, interactions, and targets.	http://www.drugbank.ca	[21]
SIDER	A database that provides information on marketed drugs and their adverse reactions records.	http://sideeffects.embl.de	[22]
ZINC	A free commercial compound database for virtual screening.	https://zinc15.docking.org/	[23]

Table S3 Activity thresholds corresponding to different types of targets

Target type	Active	Inactive
GPCR	<=100 nM	>100 nM
Ion channel	<=10 μM	>10 μM
Enzyme	<=1 μM	>1 μM
Kinase	<=1 μM	>1 μM
Transporter	<=1 μM	>1 μM
Nuclear receptor	<=100 nM	>100 nM
Others	<=1 μM	>1 μM

Table S4 Hyperparameter search range for MTGNN

Hyperparameter name	Search Range
num_layers	[1, 2, 3, 4, 5, 6, 7, 8, 9]
num_timesteps	[1, 2, 3, 4, 5, 6, 7, 8, 9]
graph_feat_size	[100, 200, 300, 400, 500, 600, 700, 800]
learning_rate	[0.005, 0.0001, 0.00005]
batch_size	[64, 128, 256]
weight_decay	[0.0, 1e-7, 1e-5]

Table S5 Hyperparameter settings of models corresponding to seven target types

	n_layers	n_timesteps	graph_feat_size	learning_rate	batch_size	weight_decay	dropout
GPCR	3	2	200	0.0001	128	0	0.1
Ion channel	3	2	200	0.0001	128	0	0.1
Enzyme	1	4	600	0.0001	64	0	0.1
Kinase	3	2	200	0.0005	64	1e-7	0.1
NR	3	2	200	0.0001	128	0	0.1
Transporter	3	2	200	0.0005	64	0	0.1
others	3	2	200	0.0001	128	0	0.1

Table S6 Parameter settings of Roche's off-target prediction model in the work

Model	**Hyperparameter name**	**Hyperparameter search range / hyperparameter setting**
NeuralNetworks	Hidden units	[256, 512, 1024, 2048]
	Dropout input	[0, 0.1, 0.2]
	Dropout hidden	[0.2, 0.3, 0.4]
	Learning rate	[0.01, 0.001, 0.0001]
	Batch size	[64, 128, 256]
	Hidden layers	2
	Kernel regularizer l2	0.001
RandomForest	n_estimators	100
	min_samples_split	2
	min_samples_leaf	1
Auto-Sklearn	time_left_for_this_task	120
	per_run_time_limit=30	30

Table S7 Model performance of different off-target prediction models under different target types

Target type	Model	AUROC	MCC	BACC	F1
GPCR	MTGNN	**0.9722**	0.7686	**0.9238**	0.7994
	NeuralNetworks	0.9665	**0.7865**	0.8954	**0.8139**
	RandomForest	0.8758	0.7766	0.8758	0.8028
	Auto-Sklearn	0.8987	0.7544	0.8987	0.7818
Ion channel	MTGNN	**0.9703**	**0.8158**	**0.9280**	**0.8420**
	NeuralNetworks	0.9685	0.8150	0.9022	0.8313
	RandomForest	0.8872	0.8133	0.8872	0.8252
	Auto-Sklearn	0.8936	0.7781	0.8936	0.7996
Enzyme	MTGNN	**0.9657**	0.7421	**0.9080**	**0.7675**
	NeuralNetworks	0.9499	0.7198	0.8556	0.7424
	RandomForest	0.8559	**0.7466**	0.8559	0.7625
	Auto-Sklearn	0.8850	0.7124	0.8850	0.7404
Kinase	MTGNN	**0.9679**	**0.7746**	**0.9113**	**0.8069**
	NeuralNetworks	0.9581	0.7555	0.8779	0.7822
	RandomForest	0.8484	0.7287	0.8484	0.7469
	Auto-Sklearn	0.8563	0.7083	0.8563	0.7341
NR	MTGNN	0.9656	0.7476	**0.9179**	0.7778
	NeuralNetworks	**0.9687**	**0.7487**	0.8885	**0.7793**
	RandomForest	0.8853	0.7622	0.8853	0.7954
	Auto-Sklearn	0.8983	0.7508	0.8983	0.7831
Transporter	MTGNN	**0.9946**	**0.8926**	**0.9750**	**0.9074**
	NeuralNetworks	0.9858	0.8575	0.9198	0.8743
	RandomForest	0.9071	0.8625	0.9071	0.8748
	Auto-Sklearn	0.8481	0.6730	0.8481	0.6940
others	MTGNN	**0.9870**	**0.8664**	**0.9547**	**0.8866**
	NeuralNetworks	0.9862	0.8625	0.9253	0.8839
	RandomForest	0.9197	0.8648	0.9197	0.8847
	Auto-Sklearn	0.9248	0.8495	0.9248	0.8721

Table S8. Comparison between multitask off-target prediction models and benchmark models for drug-target interaction prediction

Target type	Model	AUROC	MCC	BACC	F1
GPCR	MTGNN	0.9722	0.7686	**0.9238**	**0.7994**
	TransformerCPI2.0	**0.9730**	**0.8016**	0.8946	0.7555
	TransformerCPI	0.9654	0.7464	0.8874	0.7944
	GraphDTA	0.9474	0.6515	0.8544	0.7178
Ion channel	MTGNN	0.9703	**0.8158**	**0.9280**	**0.8420**
	TransformerCPI2.0	0.9755	0.8102	0.9012	0.7760
	TransformerCPI	**0.9769**	0.7135	0.9006	0.7535
	GraphDTA	0.9475	0.6246	0.8534	0.6805
Enzyme	MTGNN	0.9657	**0.7421**	**0.9080**	**0.7675**
	TransformerCPI2.0	**0.9673**	0.7300	0.8791	0.6812
	TransformerCPI	0.9577	0.6924	0.8747	0.7410
	GraphDTA	0.9484	0.6539	0.8609	0.7084
Kinase	MTGNN	0.9679	0.7746	**0.9113**	0.8069
	TransformerCPI2.0	**0.9787**	**0.8429**	0.9067	**0.8086**
	TransformerCPI	0.9756	0.7382	0.8987	0.7849
	GraphDTA	0.9488	0.6498	0.8480	0.7146
NR	MTGNN	0.9656	**0.7476**	**0.9179**	**0.7778**
	TransformerCPI2.0	**0.9700**	0.7447	0.8807	0.6895
	TransformerCPI	0.9652	0.6989	0.8788	0.7535
	GraphDTA	0.9556	0.7259	0.8667	0.7755
Transporter	MTGNN	**0.9946**	**0.8926**	**0.9750**	**0.9074**
	TransformerCPI2.0	0.9883	0.8473	0.9263	0.8139
	TransformerCPI	0.9934	0.7700	0.9386	0.8001
	GraphDTA	0.9839	0.7036	0.9080	0.7446
others	MTGNN	0.9870	0.8664	0.9547	0.8866
	TransformerCPI2.0	0.9947	0.9205	0.9562	0.9047
	TransformerCPI	**0.9963**	**0.9301**	**0.9653**	**0.9417**
	GraphDTA	0.9889	0.7556	0.9245	0.7912

Table S9 Drug data statistics under the first-level ATC code

Code	Anatomical/Pharmacological Group	Drugs number
A	Alimentary tract and metabolism	453
B	Blood and blood-forming organs	103
C	Cardiovascular system	493
D	Dermatologicals	352
G	Genito urinary system and sex hormones	246
H	Systemic hormonal preparations, excl. sex hormones and insulins	120
J	Antiinfectives for systemic use	476
L	Antineoplastic and immunomodulating agents	336
M	Musculo-skeletal system	185
N	Nervous system	612
P	Antiparasitic products, insecticides and repellents	115
R	Respiratory system	308
S	Sensory organs	293
V	Various	161
The number of Drug-ATC pairs		4253
The number of unique SMILES sequences		3491

Table S10 Comparison of model effects of MLKNN and its ablation experiment for ATC classification

	AUROC	mAP	Rank loss
MLKNN	**0.609**±0. 009	**0.376**±0. 009	**0.377**±0. 006
ECFP_MLKNN	0.560±0. 012	0.310±0. 009	0.415±0. 010

* Values in the table are the mean ± variance of the five-fold cross-training

* The best values of different metrics are bold

Table S11 The hyperparameter search range and optimal settings of LightGBM and ECFP_LightGBM for toxicity prediction

Hyperparameter name	Search Range [range1, range2, step]	Best Hyperparameter	
		LightGBM	ECFP_LightGBM
n_estimators	[100, 1000, 100]	900	900
learning_rate	[0.1, 0.5, 0.05]	0.2	0.45
num_leaves	[20, 1000, 20]	540	940
max_depth	[3, 12, 1]	6	9
min_data_in_leaf	[100, 500, 100]	100	100
lambda_l1	[0, 100, 5]	0	0
lambda_l2	[0, 100, 5]	0	10
min_gain_to_split	[0, 15, 0.1]	0.2	0.0
bagging_fraction	[0.2, 1.0, 0.1]	1.0	1.0
bagging_freq	[1]	1	1
feature_fraction	[0.2, 0.9, 0.1]	0.4	0.5

Table S12 Comparison of LightGBM with other toxicity prediction models

Featurization	Model	AUROC	Accuracy	F1	BACC	MCC
off-target-based	RF	0.896±0.002	0.775±0.002	0.735±0.003	0.792±0.003	0.563±0.006
	SVM	0.880±0.003	0.793±0.005	0.720±0.008	0.779±0.006	0.556±0.010
	XGBoost	0.904±0.003	0.809±0.008	0.753±0.010	0.807±0.008	0.601±0.017
	LightGBM	**0.912**±0.007	**0.824**±0.005	**0.769**±0.006	**0.819**±0.005	**0.630**±0.009
structure-based	ECFP_LightGBM	0.774±0.010	0.705±0.009	0.618±0.009	0.694±0.008	0.381±0.015
	STDNN-SE	0.883±0.006	0.789±0.014	0.749±0.013	0.783±0.012	0.569±0.026
	DTox	0.738±0.039	0.703±0.022	0.560±0.065	0.666±0.038	0.344±0.067

* Values in the table represent the mean ± variance of the five-fold cross-training.

* The optimal values for various metrics are indicated in bold.

Table S13 Drugs and their associated ADRs

Pergolide		**Sertindole**	**Sibutramine**		**Phenylpropanolamine**
Abdominal pain	Tremor	Headache	Angioedema	Hypersensitivity	Anxiety
Dizziness	Anxiety	Insomnia	Urticaria	Suicidal ideation	Arrhythmia
Weight increased	Weight decreased	Orthostatic hypotension	Body temperature increased	Gastrointestinal disorder	Central nervous system stimulation
Asthenia	Insomnia	Sedation	Anxiety	Arrhythmia	Neurotoxicity
Hypercholesterolaemia	Muscle twitching	Ejaculation disorder	Cardiac failure congestive	Orthostatic hypotension	Coordination abnormal
Oedema peripheral	Coordination abnormal	Erectile dysfunction	Salivary hypersecretion	Ventricular tachycardia	Electrocardiogram QT prolonged
Diarrhoea	Dysphagia	Dizziness	Confusional state	Constipation	Constipation
Dyspepsia	Hypersomnia	Pallor	Dysphagia	Depression	Confusional state
Memory impairment	Tardive dyskinesia	Anxiety	Dermatitis contact	Anticholinergic syndrome	Blood pressure increased
Dyskinesia	Rash	Urinary retention	Dizziness	Dysuria	Dizziness
Hypertension	Myoclonus	Increased appetite	Headache	Cardiac arrest	Hallucination
Dysuria	Tachycardia	Hyperglycaemia	Hypotension	Nausea	Hypersensitivity
Constipation	Haematuria	Hyperprolactinaemia	Nervousness	Shock	Nausea
Extrapyramidal disorder	Cardiac failure congestive	Bone pain	Cerebrovascular disorder	Tachycardia	Nervousness
Body temperature increased	Pulmonary hypertension	Neuroleptic malignant syndrome	Urinary retention	Oedema peripheral	Psychotic disorder
Sleep disorder	Nervousness	Oedema	Increased appetite	Sleep disorder	Tachycardia
Headache	Pain	Galactorrhoea	Insomnia	Ureteral spasm	Tremor
Akathisia	Palpitations	Dyskinesia	Dry skin		Vomiting
Orthostatic hypotension	Psychotic disorder	Tardive dyskinesia			Dysphoria
Heart rate increased	Neuroleptic malignant syndrome	Dystonia			Serotonin syndrome
Intracranial pressure increased	Urinary incontinence	Akathisia			Insomnia
		Parkinsonism			Sleep disorder
		Miosis			Sensitisation
		Sexual dysfunction			Cardiac failure
		Tachycardia			Seizure

Table S18 Comparison of ML-based ADR prediction models

	AUROC	mAP	Rank loss
off-target-based MLKNN	**0.905**±0.0008	**0.389**±0.0030	**0.077**±0.0006
ECFP-based MLKNN	0.899±0.0006	0.345±0.0028	0.087±0.0008
off-target-based RF	0.708±0.1460	Not Applicable	Not Applicable
ECFP-based RF	0.659±0.1250	Not Applicable	Not Applicable

Table S19 Molecular docking scores (kcal/mol) for pergolide and sertindole

Drug name	Target name	PDB id	Ref.	Docking score
Pergolide	SLC6A4	5I73	[24, 25]	**-8.752**
Pergolide	HTR7	7XTC	[26]	**-6.458**
Sertindole	CHRNA4	5KXI	[27]	**-6.789**
Sertindole	OPRK1	6B73	[28]	**-9.026**

SI References

1. Bendels S, Bissantz C, Fasching B, Gerebtzoff G, Guba W, Kansy M, Migeon J, Mohr S, Peters JU, Tillier F *et al*: **Safety screening in early drug discovery: An optimized assay panel**. *J Pharmacol Toxicol Methods* 2019, **99**:106609.
2. Bowes J, Brown AJ, Hamon J, Jarolimek W, Sridhar A, Waldron G, Whitebread S: **Reducing safety-related drug attrition: the use of in vitro pharmacological profiling**. *Nat Rev Drug Discov* 2012, **11**(12):909-922.
3. Lynch JJ, 3rd, Van Vleet TR, Mittelstadt SW, Blomme EAG: **Potential functional and pathological side effects related to off-target pharmacological activity**. *J Pharmacol Toxicol Methods* 2017, **87**:108-126.
4. Krejsa CM, Horvath D, Rogalski SL, Penzotti JE, Mao B, Barbosa F, Migeon JC: **Predicting ADME properties and side effects: the BioPrint approach**. *Curr Opin Drug Discov Devel* 2003, **6**(4):470-480.
5. Rodgers G, Austin C, Anderson J, Pawlyk A, Colvis C, Margolis R, Baker J: **Glimmers in illuminating the druggable genome**. *Nat Rev Drug Discov* 2018, **17**(5):301-302.
6. Madjarov G, Kocev D, Gjorgjevikj D, Džeroski S: **An extensive experimental comparison of methods for multi-label learning**. *Pattern Recognition* 2012, **45**(9):3084-3104.
7. Zhang W, Liu F, Luo L, Zhang J: **Predicting drug side effects by multi-label learning and ensemble learning**. *BMC Bioinformatics* 2015, **16**:365.
8. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, Ye Q, Liu T-Y: **Lightgbm: A highly efficient gradient boosting decision tree**. *Advances in neural information processing systems* 2017, **30**.
9. Akiba T, Sano S, Yanase T, Ohta T, Koyama M: **Optuna: A next-generation hyperparameter optimization framework**. In: *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining: 2019*; 2019: 2623-2631.
10. Ilhan Topcu D, Can Cubukcu H: **Optimization of patient-based real-time quality control based on the Youden index**. *Clin Chim Acta* 2022, **534**:50-56.
11. Perlis ML, Corbitt CB, Kloss JD: **Insomnia research: 3Ps and beyond**. *Sleep Med Rev* 2014, **18**(3):191-193.
12. Hedlund PB, Huitron-Resendiz S, Henriksen SJ, Sutcliffe JG: **5-HT7 receptor inhibition and inactivation induce antidepressantlike behavior and sleep pattern**. *Biol Psychiatry* 2005, **58**(10):831-837.
13. Yamada J, Sugimoto Y, Yoshikawa T, Horisaka K: **Hyperglycemia induced by the 5-HT receptor agonist, 5-methoxytryptamine, in rats: involvement of the peripheral 5-HT2A receptor**. *Eur J Pharmacol* 1997, **323**(2-3):235-240.
14. Barth E, Albuszies G, Baumgart K, Matejovic M, Wachter U, Vogt J, Radermacher P, Calzia E: **Glucose metabolism and catecholamines**. *Crit Care Med* 2007, **35**(9 Suppl):S508-518.
15. Ghamari N, Zarei O, Arias-Montano JA, Reiner D, Dastmalchi S, Stark H, Hamzeh-Mivehroud M: **Histamine H(3) receptor antagonists/inverse agonists: Where do they go?** *Pharmacol Ther* 2019, **200**:69-84.
16. Flordellis C, Manolis A, Scheinin M, Paris H: **Clinical and pharmacological significance**

of alpha2-adrenoceptor polymorphisms in cardiovascular diseases. *Int J Cardiol* 2004, **97**(3):367-372.

17. Walsh GM: **Emerging safety issues regarding long-term usage of H(1)receptor antagonists**. *Expert Opin Drug Saf* 2002, **1**(3):225-235.
18. Wang W, MacKinnon R: **Cryo-EM Structure of the Open Human Ether-a-go-go-Related K(+) Channel hERG**. *Cell* 2017, **169**(3):422-430 e410.
19. Mendez D, Gaulton A, Bento AP, Chambers J, De Veij M, Felix E, Magarinos MP, Mosquera JF, Mutowo P, Nowotka M *et al*: **ChEMBL: towards direct deposition of bioassay data**. *Nucleic Acids Res* 2019, **47**(D1):D930-D940.
20. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA, Yu B *et al*: **PubChem in 2021: new data content and improved web interfaces**. *Nucleic Acids Res* 2021, **49**(D1):D1388-D1395.
21. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z *et al*: **DrugBank 5.0: a major update to the DrugBank database for 2018**. *Nucleic Acids Res* 2018, **46**(D1):D1074-D1082.
22. Kuhn M, Letunic I, Jensen LJ, Bork P: **The SIDER database of drugs and side effects**. *Nucleic Acids Res* 2016, **44**(D1):D1075-1079.
23. Sterling T, Irwin JJ: **ZINC 15--Ligand Discovery for Everyone**. *J Chem Inf Model* 2015, **55**(11):2324-2337.
24. Coleman JA, Green EM, Gouaux E: **X-ray structures and mechanism of the human serotonin transporter**. *Nature* 2016, **532**(7599):334-339.
25. Ravna AW, Jaronczyk M, Sylte I: **A homology model of SERT based on the LeuT(Aa) template**. *Bioorg Med Chem Lett* 2006, **16**(21):5594-5597.
26. Thirumaran SL, Lepailleur A, Rochais C: **Structure-activity relationships of serotonin 5-HT(7) receptors ligands: A review**. *Eur J Med Chem* 2019, **183**:111705.
27. Morales-Perez CL, Noviello CM, Hibbs RE: **X-ray structure of the human alpha4beta2 nicotinic receptor**. *Nature* 2016, **538**(7625):411-415.
28. Che T, Majumdar S, Zaidi SA, Ondachi P, McCorvy JD, Wang S, Mosier PD, Uprety R, Vardy E, Krumm BE *et al*: **Structure of the Nanobody-Stabilized Active State of the Kappa Opioid Receptor**. *Cell* 2018, **172**(1-2):55-67 e15.