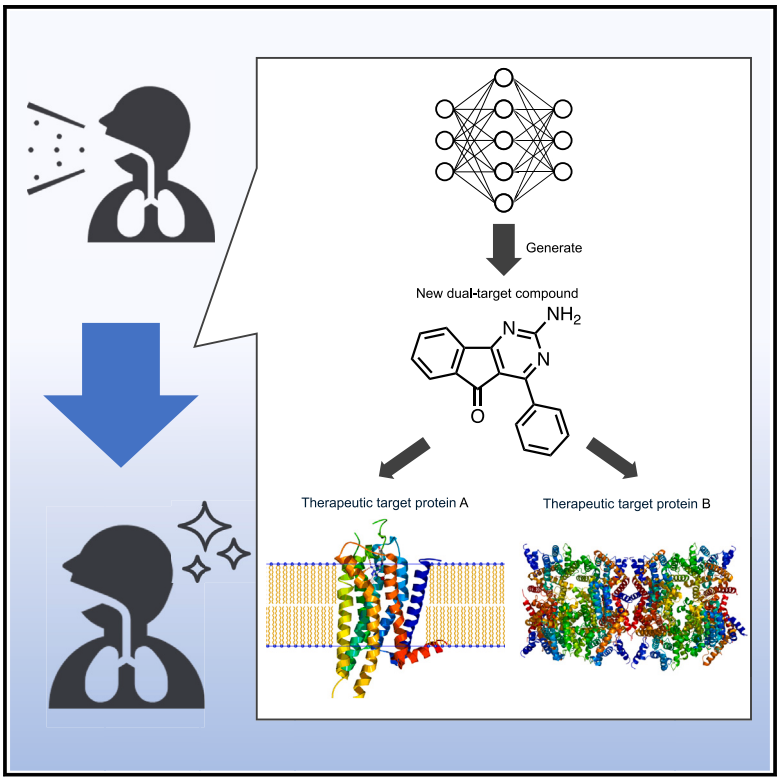
iScience

Article

De novo generation of dual-target compounds using artificial intelligence

Graphical abstract



Authors

Kasumi Yasuda, Francois Berenger, Kazuma Amaike, ..., Koji Tsuda, Kenichiro Itami, Yoshihiro Yamanishi

Correspondence

yamanishi@i.nagoya-u.ac.jp

In brief

Natural sciences; Biological sciences; Bioinformatics; Pharmacoinformatics

Highlights

- We present artificial intelligence-based methods for polypharmacology
- Our method can design compounds that interact with multiple therapeutic targets
- The AI-designed compounds for two therapeutic target proteins were synthesized
- The synthesized compounds successfully interacted with the two target proteins

Yasuda et al., 2025, iScience *28*, 111526
January 17, 2025
https://doi.org/10.1016/j.isci.2024.111526

CellPress

iScience

CellPress
OPEN ACCESS

Article

De novo generation of dual-target compounds using artificial intelligence

Kasumi Yasuda,[1,6] Francois Berenger,[1,2,6] Kazuma Amaike,[3,6] Ayaka Ueda,[3] Tomoya Nakagomi,[3] Genki Hamasaki,[1] Chen Li,[1,5] Noriko Yuyama Otani,[1,5] Kazuma Kaitoh,[1,5] Koji Tsuda,[2] Kenichiro Itami,[3,4] and Yoshihiro Yamanishi[1,5,7,*]

[1]Department of Bioscience and Bioinformatics, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan
[2]Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwa-no-ha, Kashiwa, Chiba 277-8561, Japan
[3]Graduate School of Science, Nagoya University, Chikusa, Nagoya, Aichi 464-8602, Japan
[4]Institute of Transformative Bio-Molecules (WPI-ITbM), Nagoya University, Chikusa, Nagoya, Aichi 464-8602, Japan
[5]Graduate School of Informatics, Nagoya University, Chikusa, Nagoya, Aichi 464-8601, Japan
[6]These authors contributed equally
[7]Lead contact
*Correspondence: yamanishi@i.nagoya-u.ac.jp
https://doi.org/10.1016/j.isci.2024.111526

SUMMARY

Drugs that interact with multiple therapeutic targets are potential high-value products in polypharmacology-based drug discovery, but the rational design remains a formidable challenge. Here, we present artificial intelligence (AI)-based methods to design the chemical structures of compounds that interact with multiple therapeutic target proteins. The molecular structure generation is performed by a fragment-based approach using a genetic algorithm with chemical substructures and a deep learning approach using reinforcement learning with stochastic policy gradients in the framework of generative adversarial networks. Using the proposed methods, we designed the chemical structures of compounds that would interact with two therapeutic targets of bronchial asthma, i.e., adenosine A2a receptor (ADORA2A) and phosphodiesterase 4D (PDE4D). We then synthesized 10 compounds and evaluated their bioactivities via the binding assays of 39 target human proteins, including ADORA2A and PDE4D. Three of the 10 synthesized compounds successfully interacted with ADORA2A and PDE4D with high specificity.

INTRODUCTION

The development of drugs is hampered by enormous cost and time requirements. The typical period between discovery and launch of a drug on the market exceeds 10 years, incurring billions of dollars.[1] Discovering compounds with the desired bioactivities within the vast chemical space (theoretically comprising 10^{60} compounds[2]) is challenging as it requires screening to identify a small fraction of such compounds with the desired phenotypes. Moreover, experimentally verifying the bioactivities of all compounds in such a large chemical space is impossible unless the efficiency of identifying compounds with the desired bioactivities can be improved.

The *de novo* design of drug candidate compounds is expected to yield chemical structures without depending on existing compound libraries.[3] A structure generator algorithm can automatically output a chemical structure based on the initial conditions. Various structure generators using deep generative models, such as the variational autoencoder,[4–6] generative adversarial network (GAN),[7–10] and transformer,[11–13] have been recently reported.[14] The generated structures maximize the lipid solubility of the drug (logP) or a quantitative estimate of drug likeness. Most previously reported methods have focused on a single chemical property or the bioactivity of a single therapeutic protein.

Cocktail therapy, which exploits the synergistic effects of multiple drug combinations, is a potentially effective treatment for multifactorial diseases.[15] The simultaneous regulation of multiple therapeutic targets by several drugs in the cocktail yields synergistic effects. However, harmful side effects can arise from the blind combination of multiple drugs, and simultaneous treatment with multiple drugs can be burdensome for patients. Ideally, one drug that simultaneously acts on multiple therapeutic targets will provide synergistic effects, lessen the side effects, and reduce the cost of cocktail therapies.[15–19] Although the potential of such polypharmacology has been long recognized, the difficulty in rationally designing multitarget compounds has limited its progress.[15,18–21] A computational method for designing compounds that bind to two target proteins was proposed, but the method is applicable only to proteins with known three-dimensional structures and ligand-binding pockets.[19]

In this study, we propose an artificial intelligence (AI)-based method for generating multitarget compounds. To this end, we

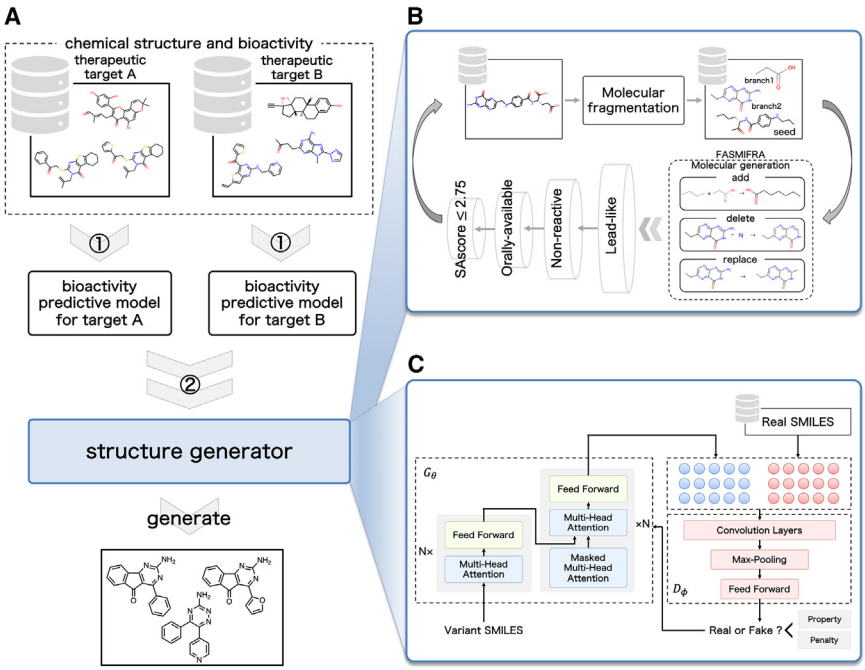


Figure 1. Overview of the proposed method for generating dual-target compounds

(A) Construction of quantitative structure-activity relationship (QSAR) models for predicting the compound bioactivities on different therapeutic targets and generating dual-target compound candidates. Incorporation of bioactivity-value prediction models into (B) the fragment-based structure generator DualFASMIFRA and (C) the deep learning-based structure generator DualTransORGAN.

develop fragment- and deep learning-based structure generators for designing the chemical structures of compounds that can likely interact with two therapeutic targets. The usefulness of the method is demonstrated in a case study involving dual-target compounds for two therapeutic targets of bronchial asthma. This proposed method can open avenues for polypharmacology-based drug discovery.

RESULTS

Overview of the proposed method for designing dual-target compounds

Figure 1 overviews the proposed method for generating dual-target compounds. The proposed method comprises dual-target bioactivity prediction models and dual-target structure generators.

The dual-target bioactivity prediction models (constructed first) calculate the bioactivity scores of a compound based on its structure for both therapeutic targets (Figure 1A). For each therapeutic target, we constructed the quantitative structure-activity relationship (QSAR) model using random forest regressor, which inputs the chemical structure of each compound and outputs its bioactivity (pIC_{50} value). These pIC_{50} values are averaged to obtain the objective function of the dual-target structure generator.

The dual-target structure generators design the chemical structures of dual-target compounds (Figures 1B and 1C). The first generator is a fragment-based structure generator and optimizer, namely DualFASMIFRA (Figure 1B). The structures are generated by a genetic algorithm (GA) that assembles active compound fragments against the target proteins. The GA is a conservative but pragmatic structure generator. The second generator, a deep learning-based structure generator, namely DualTransORGAN (Figure 1C), is a deep generative model based on a GAN with a transformer encoder and decoder. DualTransORGAN generates plausible structures that capture the semantic features of a compound via reinforcement learning using stochastic policy gradients. The proposed algorithms are detailed in the "method details" section.

As a case study, we generated compounds that interact with two therapeutic targets of bronchial asthma: adenosine A2a receptor (ADORA2A) and phosphodiesterase 4D (PDE4D).

Generation of dual-target compounds using the proposed structure generators

The chemical structures of compounds for dual targets, ADORA2A and PDE4D, were generated using DualFASMIFRA. These structures were optimized by combining three techniques: (1) calculation of a scoring function using the QSAR model (2) molecular generation, and (3) optimization. The affinities of the generated compounds to the therapeutic targets were estimated using a scoring function, and the overall optimization process was driven by the GA. Adhering to the best practices in software design, these three techniques were maintained separately to ensure that any component could be changed independently of the others.

The generated chemical structures of the high-scoring dual-target compounds are shown in Figure S1. DualFASMIFRA generated diverse molecular scaffolds with different arrangements of rings and types of atoms, including structures containing fluorene, piperazine, and fused rings with multiple substructures of nitrogen atoms. Additionally, most compounds were expected to be highly planar because they did not have a chiral center and contained only a small proportion of sp carbons. Thus, they can be synthesized via cross-coupling reactions and other techniques.

Figures 2A and 2B) show the scatterplots of the predicted versus observed pIC_{50} values for ADORA2A and PDE4D, respectively. These values were highly correlated for both targets, verifying the optimal performance of the QSAR models. Figure 2C shows the data statistics after one iteration of the GA, including the ratios of elite high-scoring compounds

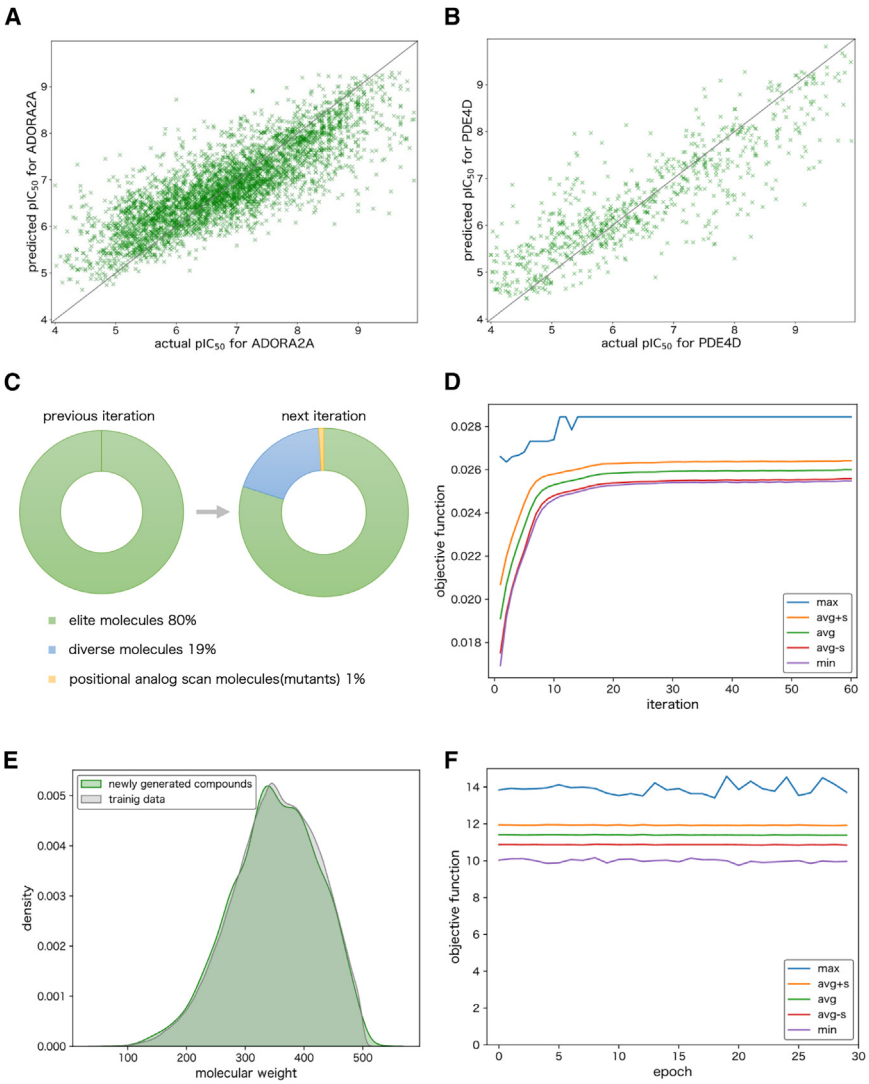


Figure 2. Molecular generation by the proposed structure generators

(A and B) Scatterplots of the predicted versus observed pIC_{50} values using the QSAR models for (A) ADORA2A and (B) PDE4D.

(C) Ratios of elite compounds with high scores in the QSAR models, diverse (randomly selected) compounds, and mutant compounds (positional scan analogs of the merged set of elite and random molecular population) after one iteration of the genetic algorithm in DualFASMIFRA.

(D) Optimization trajectories of the objective functions in DualFASMIFRA. The horizontal axis indicates the generation number of the genetic algorithm, and the vertical axis indicates the predicted molecular weight invariant, pIC_{50} (averaged between the results of ADORA2A and PDE4D).

(E) Distributions of molecular weights of the generated compounds (green) and compounds in the training dataset (gray) applied to DualTransORGAN. The horizontal and vertical axes indicate the molecular weights of the compounds and the densities of the compounds with the corresponding molecular weights, respectively.

(F) Optimization trajectories of the objective functions in DualTransORGAN. The maximum (max), average (avg), and minimum (min) values are shown.

predicted by the QSAR models, randomly selected diverse compounds, and mutant compounds.

In the application of an elitist GA, most (80%) of the generated molecules for the next iteration were high scoring (elite), whereas 19% were selected randomly (the ''diverse'' population) to ensure chemical diversity in the molecular population (Figure 2C). The optional 1% of ''mutant'' molecules allowed for mutations of randomly selected molecules from the previous population using a medicinal chemistry technique called positional analog scanning.[22] Figure 2D shows the optimization trajectories of the objective functions in the structure generator. The predicted average, maximum, and minimum pIC_{50} values for ADORA2A and PDE4D plateaued after approximately 20 iterations. The proportions in the GA (i.e., 80%, 19%, and 1%) were determined via trial and error in several optimization experiments. Other proportions in the GA are possible, but they may exhibit different convergence behavior.

Subsequently, dual-target compounds for ADORA2A and PDE4D were generated using DualTransORGAN. We analyzed 50 compounds with the highest sums of their two bioactivity values and confirmed that the generated structures did not have the same structures as the compounds in the training set. The chemical structures generated by DualTransORGAN were structurally diverse and contained fluorine and sulfur atoms as well as polar functional groups such as hydroxy, carboxy, and cyano groups. The generated structures were rich in steric properties. Moreover, ring structures such as cyclopropane and structures containing chiral carbons were generated. The chemical structures generated by DualFASMIFRA had fewer polar functional groups and larger conjugated systems than those generated by DualTransORGAN. The generated high-scoring chemical structures of the dual-target compounds are shown in Figure S2.

We investigated the distributions of the compounds generated by DualTransORGAN. Figure 2E compares the distributions of molecular weights of these compounds to those in the training dataset. The compounds' molecular weight distributions were almost identical, indicating that DualTransORGAN can reproduce relatively small compounds similar to those in the training dataset in terms of molecular weights. DualTransORGAN generated compounds with molecular weights below 500, consistent with those in the training dataset. Figure 2F shows the optimization trajectories

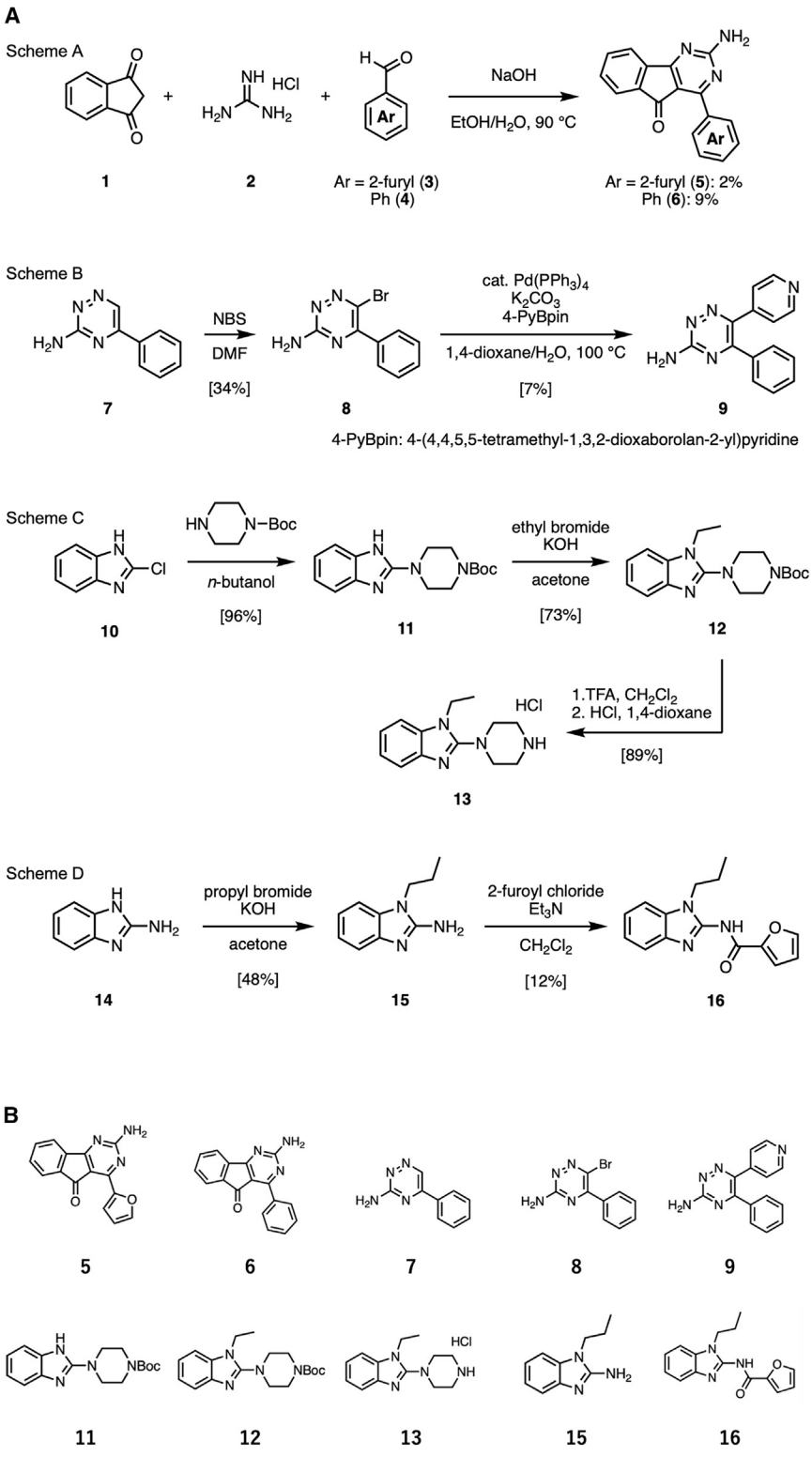


Figure 3. Synthesis of the AI-generated compounds

(A) Computationally generated compounds **5**, **9**, **13**, and **16** are predicted to interact with both ADORA2A and PDE4D. Schemes A, B, C, and D show the synthesis routes of compounds **5**, **9**, **13**, and **16**, respectively. Compound **6** is a derivative of **5**.

(B) The chemical structures of the synthesized 10 compounds and the corresponding compound numbers in (A) are shown.

of the objective functions in the structure generator. The objective function in DualTransORGAN did not change markedly as the epochs increased. The predicted average, maximum, and minimum pIC_{50} values for ADORA2A and PDE4D remained stable across all iterations, implying that the structures of the generated compounds converged along with the convergence of the objective functions. Other parameters in the GAN model are possible, but they may yield different convergence behavior.

Chemical synthesis of selected AI-generated compounds

Among all the AI-generated compounds, we selected the easily synthesizable compounds. Based on the computationally generated chemical structures, we synthesized the 10 compounds shown in Figure 3. Compounds **5**, **9**, **13**, and **16** were anticipated to interact with ADORA2A and PDE4D. Compound **5** was prepared from 1,3-indanedione (**1**), guanidine hydrochloride (**2**), and arylaldehyde in the presence of a base in an ethanol/water mixture.[23] The 4-phenyl compound (**6**) was synthesized along with the 2-furyl compound (**5**) via Scheme A in Figure 3A. The synthesis of **9** is described in Scheme B in Figure 3A. 3-Amino-5-phenyl-1,2,4-triazine (**7**) was prepared as described in a previous study,[24] and **8** was formed via bromination at the C6 position of **7**. **9** was obtained via Suzuki-Miyaura cross-coupling of **8** with 4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl) pyridine. **13** was synthesized by treating commercially available 2-chlorobenzimidazole (**10**) with *N*-boc-protected piperazine. The resulting product **11** was treated with ethyl bromide and subsequently deprotected to obtain **13** (Scheme C in Figure 3A). Compound **16** was obtained by *N*-alkylation of commercially available 2-amino benzimidazole (**14**) with *n*-propyl bromide, followed by condensation with 2-furoyl chloride (Scheme D in Figure 3A). The chemical structures of the 10 synthesized compounds are shown in Figure 3B.

A

Family	Target protein	
	Full name	Gene symbol
GPCR	adenosine A2a receptor	ADORA2A
	adrenoceptor alpha 1A	ADRA1A
	adrenoceptor alpha 2A	ADRA2A
	adrenoceptor beta 1	ADRB1
	adrenoceptor beta 2	ADRB2
	cannabinoid receptor 1	CNR1
	cannabinoid receptor 2	CNR2
	cholecystokinin A receptor	CCKAR
	dopamine receptor D1	DRD1
	dopamine receptor D2	DRD2
	endothelin receptor type A	EDNRA
	histamine receptor H1	HRH1
	histamine receptor H2	HRH2
	cholinergic receptor muscarinic 1	CHRM1
	cholinergic receptor muscarinic 2	CHRM2
	cholinergic receptor muscarinic 3	CHRM3
	opioid receptor delta 1	OPRD1
	opioid receptor kappa 1	OPRK1
	opioid receptor mu 1	OPRM1
	5-hydroxytryptamine receptor 1A	HTR1A
	5-hydroxytryptamine receptor 1B	HTR1B
	5-hydroxytryptamine receptor 2A	HTR2A
	5-hydroxytryptamine receptor 2B	HTR2B
	arginine vasopressin receptor 1	AVPR1A
Transporters	solute carrier family 6 member 3	SLC6A3
	solute carrier family 6 member 2	SLC6A2
	solute carrier family 6 member 4	SLC6A4
Ion Channels	cholinergic receptor nicotinic alpha 4 subunit	CHRNA4
	potassium voltage-gated channel subfamily H member 2	KCNH2
	5-hydroxytryptamine receptor 3A	HTR3A
Kinase	LCK proto-oncogene, Src family tyrosine kinase	LCK
Nuclear Receptors	Androgen Receptor	AR
	nuclear receptor subfamily 3 group C member 1	NR3C1
Other non-Kinase Enzymes	prostaglandin-endoperoxide synthase 1	PTGS1
	prostaglandin-endoperoxide synthase 2	PTGS2
	Acetylcholinesterase	ACHE
	Monoamine oxidase A	MAOA
	Phosphodiesterase 3A	PDE3A
	Phosphodiesterase 4D	PDE4D

B

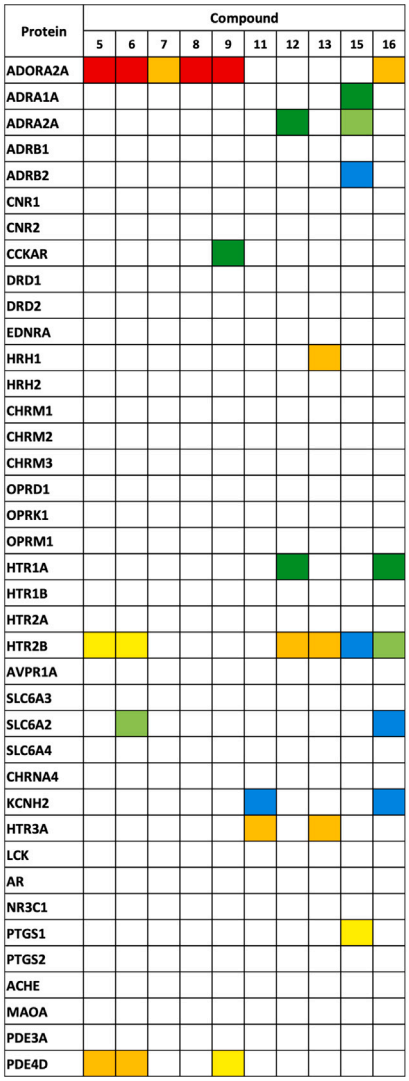


Figure 4. List of target proteins in the binding assays and hit distributions of synthesized compounds for the target human proteins

39 human target proteins including 24 GPCRs, 3 transporters, 3 ion channels, 1 kinase, 2 nuclear receptors, and 6 nonkinase enzymes from the viewpoints of drug safety and pharmacological actions are selected (A). The binding activities of the synthesized compounds **5**, **6**, **7**, **8**, **9**, **11**, **12**, **13**, **15**, and **16** are evaluated on this list including ADORA2A and PDE4D (B). The rows and columns in the heatmap indicate human target proteins and synthesized compounds, respectively. The cells are indicated by the percent inhibition colors of each compound for the 39 proteins. Cells with >100% inhibition scores are highlighted in the darkest red, those with "90% $\leq$ × < 100%" inhibition in orange, "80% $\leq$ × < 90%" inhibition in yellow, "70% $\leq$ × < 80%" inhibition in yellow-green, "60% $\leq$ × < 70%" in green, "50% $\leq$ × < 60%" inhibition in light blue, and "< 50%" inhibition in white.

Biological evaluation of the synthesized compounds

To validate the computationally predicted results, the bioactivities of the 10 synthesized compounds were evaluated by performing binding assays for ADORA2A and PDE4D. To investigate the binding specificity to ADORA2A and PDE4D, we performed additional binding assays for 39 target human proteins. From the drug safety and pharmacological point of view, the target protein set comprised 24 G-protein-coupled receptors (GPCRs), 3 transporters, 3 ion channels, 1 kinase, 2 nuclear receptors, and 6 nonkinase enzymes (Figure 4A).

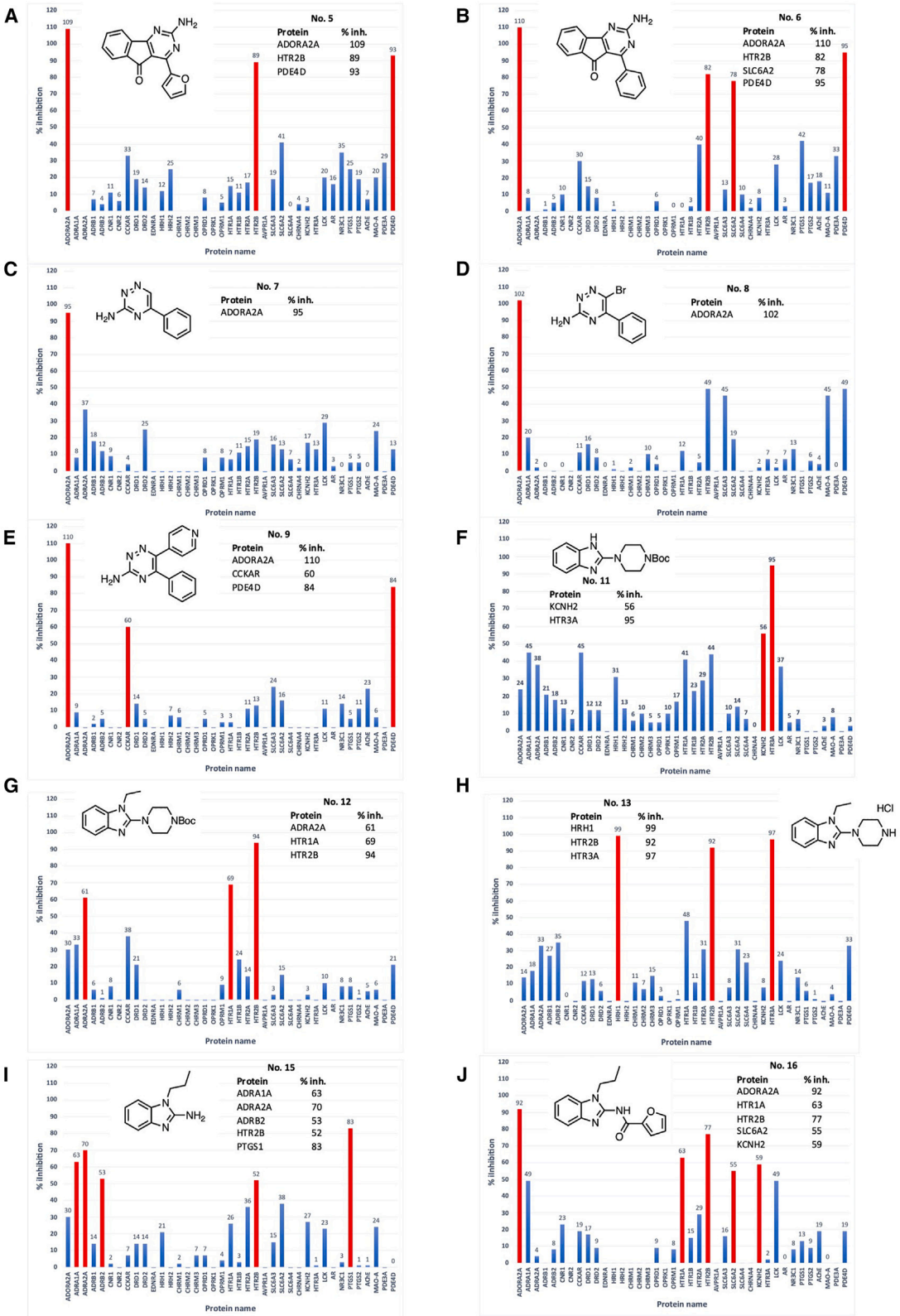
Figure 4 shows the list of 39 target human proteins and the distribution of hit compounds for 39 target human proteins in terms of inhibition scores based on the binding assays. As shown in Figure 4B, the predicted and synthesized compounds using the proposed methods exhibited highly specific binding for a limited number of proteins. For example, compounds **5**, **6**, and **9** showed more than 100% binding inhibition of ADORA2A and more than 80% that of PDE4D. For compounds **7** and **8**, only ADORA2A had >90% inhibition activity. Meanwhile, compounds **11**, **12**, **13**, and **15** showed no binding inhibition of <50% for either ADORA2A or PDE4D. Compound **16** showed >90% binding inhibition against ADORA2A. Moreover, compounds **15** and **16** had >50% inhibitory activity against five proteins. Figure S3 shows a detailed heatmap of the original percentage inhibition scores of the 10 synthesized compounds in the 39 protein binding assays.

The bioactivities and chemical structures of the synthesized compounds

Figure 5 shows the chemical structures of the 10 synthesized compounds and the percentage inhibition values for 39 target proteins in the binding assays. The enlarged results of the binding assays of compounds **5**, **6**, **7**, **8**, **9**, **11**, **12**, **13**, **15**, and **16** are shown in Figures S4–S13, respectively.

Compounds **11**–**13** and **15** interacted with different target proteins, but compounds **5**, **6** (a derivative of **5**), and **9** interacted with both ADORA2A and PDE4D. Compound **5** more selectively bound to both proteins than compound **6**, implying that the aryl group at C4 promotes its protein selectivity. The aryl moiety can be easily replaced by an aldehyde with an aryl group of interest, facilitating the synthesis of more selective compounds. In addition, the synthetic intermediates of compound **9**, i.e., compounds **7** and **8**, bound only to ADORA2A, suggesting that the aryl group at the C6 position of compound **9** facilitated its binding to PDE4D. As the aryl group at C6 was introduced via Suzuki-Miyaura cross-coupling at a later stage of synthesis, the derivatives of compound **9** with various aryl groups could be easily prepared.

The hit compounds with desired activities were synthesized from four AI-generated compounds. The generated structures



(legend on next page)

of compound **9** and 3-amino-5,6-diaryl-1,2,4-triazine derivatives corresponded to known ADORA2A ligands.[25] This implies that the proposed method successfully reproduced known ligands that were absent from the learning dataset. The maximum common substructures[26] of compound 9 and 3-amino-5,6-diaryl-1,2,4-triazine derivatives with the known ADORA2A ligand are shown in Figure S14. The other synthesized hit compounds were not previously reported as ADORA2A and PDE4D ligands. Based on these findings, we concluded that the proposed AI-based methods can accelerate the development of multitarget drugs with desired activities.

DISCUSSION

Summary

In this paper, we developed AI-based methods to generate dual-target compounds against two therapeutic targets, where two bioactivity prediction models were incorporated into structure generators and the combination of fragment- and deep learning-based generators. Using the proposed method, the chemical structures of compounds that will likely interact with any two therapeutic targets can be automatically designed. The proposed method was applied to two therapeutic targets, namely ADORA2A and PDE4D. Then, the AI-generated compounds were synthesized and *in vitro* binding assays were performed. We confirmed that the synthesized compounds successfully bound to ADORA2A and PDE4D. As the proposed method is (in principle) applicable to any combination of therapeutic targets, it is expected to promote further advancements in polypharmacology-based drug discovery.

DualTransORGAN is a generator that learns structural features at the simplified molecular-input line-entry system (SMILES) string level by incorporating a transformer into GAN, enabling the generation of structures with desired chemical properties through reinforcement learning using a stochastic policy gradient. Its main advantage is promoting diversity in the generated structures. In contrast, DualFASMIFRA uses high-speed assembly of substructure fragments based on a GA. Its primary advantage is generating only effective molecules consisting of highly frequent substructures. We believe using both generators together will yield more effective and diverse molecules, which is why we adopted both methods.

By incorporating a bioactivity predictive model for each therapeutic target, the proposed method can generate various compounds that will probably bind to both targets. Herein, random forest was used as the machine learning method for the bioactivity predictive model. However, other machine learning methods such as gradient boosting, logistic regression, kernel regression, and deep neural networks can also be used for the proposed method. An ensemble approach with multiple machine learning models would combine the individual strengths of its constituent models to improve the stability and reliability of the prediction results.

Limitations of the study

The proposed method has some limitations. Here, it was applied to small-molecule substances, but many recent drug modalities include small-, middle-, and macro-molecule substances. Whether the proposed method can usefully predict middle- and macro-molecule substances is unclear. The development of molecular selection criteria for these additional drug modalities is an important research direction. The experimental validation of the generated compounds is another limitation. Herein, the predicted chemical structures were validated only via binding assays. In future work, we would like to work on experimental validations from other viewpoints such as doses and toxicity.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to the lead contact, Yoshihiro Yamanishi (yamanishi@i.nagoya-u.ac.jp).

Materials availability

This study did not generate new unique materials.

Data and code availability

- Compound structure data are available at the ChEMBL database (https://www.ebi.ac.uk/chembl/) (https://doi.org/10.1093/nar/gky1075).
- Bioactivity data for ADORA2A are available at CHEMBL251 in the ChEMBL database (https://www.ebi.ac.uk/chembl/target_report_card/CHEMBL251/).
- Bioactivity data for PDE4D are available at CHEMBL288 in the ChEMBL database (https://www.ebi.ac.uk/chembl/target_report_card/CHEMBL288/).
- The DualFASMIFRA software, TCM@Taiwan, and ChEMBL training samples are on GitHub: (https://github.com/UnixJunkie/FASMIFRA/) (https://doi.org/10.1186/s13321-021-00566-4).
- The code of DualTransORGAN is available at https://yamanishi.cs.i.nagoya-u.ac.jp/dualtarget/, and this code is the original code.
- All data reported in this paper can be made available by the lead contact upon request.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

This research was supported by JSPS KAKENHI (grant numbers 20H05797, 21H04915, and 21K18327).

AUTHOR CONTRIBUTIONS

K.Y.: formal analysis, writing – original draft, and visualization. F.B.: formal analysis, methodology, writing, and investigation. K.A.: methodology, investigation, and writing. A.U.: methodology and investigation. T.N.: methodology and investigation. G.H.: methodology and investigation. C.L.: methodology. N.Y.O.: curation and writing. K.K.: methodology. K.T.: methodology. K.I.: methodology and investigation. Y.Y.: conceptualization, resources, writing – review and editing, visualization, supervision, and funding acquisition.

Figure 5. Experimental validation of synthesized compounds using binding assays of 39 target human proteins

The bioactivities of the 10 synthesized compounds were evaluated on 39 human target proteins, including ADORA2A and PDE4D. (A–J) are the binding assay results of synthesized compounds **5**, **6**, **7**, **8**, **9**, **11**, **12**, **13**, **15**, and **16**, respectively. The bar plots show the percentage inhibition scores of each compound for the 39 proteins. The red bars highlight the proteins with significant responses (>50% inhibition in the binding assay). The associated inhibition scores are also shown.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Synthesis of compounds
  - Synthesis of 2-amino-4-(furan-2-yl)-5*H*-indeno[1,2-*d*ay]pyrimidin-5-one (5)
  - 2-Amino-4-(furan-2-yl)-5*H*-indeno[1,2-*d*ay]pyrimidin-5-one (5)
  - Synthesis of 2-amino-4-phenyl-5*H*-indeno[1,2-*d*ay]pyrimidin-5-one (6)
  - 2-Amino-4-phenyl-5*H*-indeno[1,2-*d*ay]pyrimidin-5-one (6)
  - Synthesis of 5-phenyl-1,2,4-triazin-3-amine (7)
  - 5-Phenyl-1,2,4-triazin-3-amine (7)
  - Synthesis of 6-bromo-5-phenyl-1,2,4-triazin-3-amine (8)
  - 6-Bromo-5-phenyl-1,2,4-triazin-3-amine (8)
  - Synthesis of 5-phenyl-6-(pyridin-4-yl)-1,2,4-triazin-3-amine (9)
  - 5-Phenyl-6-(pyridin-4-yl)-1,2,4-triazin-3-amine (9)
  - Synthesis of *tert*-butyl 4-(1*H*-benzo[*d*]imidazol-2-yl)piperazine-1-carboxylate (11)
  - *tert*-butyl 4-(1*H*-benzo[*d*]imidazol-2-yl)piperazine-1-carboxylate (11)
  - Synthesis of *tert*-butyl 4-(1-ethyl-1*H*-benzo[*d*]imidazol-2-yl)piperazine-1-carboxylate (12)
  - Synthesis of 1-ethyl-2-(piperazin-1-yl)-1*H*-benzo[*d*]imidazole hydrochloride (13)
  - 1-Ethyl-2-(piperazin-1-yl)-1*H*-benzo[*d*]imidazole hydrochloride (13)
  - Synthesis of 1-propyl-1*H*-benzo[*d*]imidazol-2-amine (15)
  - 1-Propyl-1*H*-benzo[*d*]imidazol-2-amine (15)
  - Synthesis of *N*-(1-propyl-1*H*-benzo[*d*]imidazol-2-yl)furan-2-carboxamide (16)
  - *N*-(1-propyl-1*H*-benzo[*d*]imidazol-2-yl)furan-2-carboxamide (16)
  - Binding assays
  - Binding assays for GPCRs, ionic channels, and enzymes, transporters
- METHOD DETAILS
  - Therapeutic targets
  - Algorithm of the dual-target fragment-based structure generator (DualFASMIFRA)
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.111526.

Received: August 9, 2024
Revised: October 6, 2024
Accepted: November 29, 2024
Published: December 3, 2024

REFERENCES

1. DiMasi, J.A., Grabowski, H.G., and Hansen, R.W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. J. Health Econ. *47*, 20–33. https://doi.org/10.1016/j.jhealeco.2016.01.012.
2. Polishchuk, P.G., Madzhidov, T.I., and Varnek, A. (2013). Estimation of the size of drug-like chemical space based on GDB-17 data. J. Comput. Aided Mol. Des. *27*, 675–679. https://doi.org/10.1007/s10822-013-9672-4.
3. Yuan, Y., Pei, J., and Lai, L. (2011). LigBuilder 2: a practical de novo drug design approach. J. Chem. Inf. Model. *51*, 1083–1091. https://doi.org/10.1021/ci100350u.
4. Kusner, M.J., Paige, B., and Hernández-Lobato, J.M. (2017). Grammar variational autoencoder. In Proceedings of the 34th International Conference on Machine Learning, *70*, pp. 1945–1954. https://doi.org/10.48550/arXiv.1703.01925.
5. Dai, H., Tian, Y., Dai, B., Skiena, S., and Song, L. (2018). Syntax-directed variational autoencoder for molecule generation. In Proceedings of the International Conference on Learning Representations. https://doi.org/10.48550/arXiv.1802.08786.
6. Jin, W., Barzilay, R., and Jaakkola, T. (2018). Junction tree variational autoencoder for molecular graph generation. In Proceedings of the 35th International Conference on Machine Learning, *80*, pp. 2323–2332. https://doi.org/10.48550/arXiv.1802.04364.
7. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. Adv. Neural Inf. Process. Syst. *27*, 1–9. https://doi.org/10.48550/arXiv.1406.2661.
8. Cao, N.D., and Kipf, T. (2018). MolGAN: an implicit generative model for small molecular graphs. In ICML18 Workshop on Theoretical Foundations and Applications of Deep Generative Models. https://doi.org/10.48550/arXiv.1805.11973.
9. Kadurin, A., Nikolenko, S., Khrabrov, K., Aliper, A., and Zhavoronkov, A. (2017). druGAN: an advanced generative adversarial autoencoder model for de novo generation of new molecules with desired molecular properties in silico. Mol. Pharm. *14*, 3098–3104. https://doi.org/10.1021/acs.molpharmaceut.7b00346.
10. Guimaraes, G.L., Sanchez-Lengeling, B., Outeiral, C., Farias, P.L.C., and Aspuru-Guzik, A. (2017). Objective-reinforced generative adversarial networks (ORGAN) for sequence generation models. Preprint at arXiv. https://doi.org/10.48550/arXiv.1705.10843.
11. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems. https://doi.org/10.48550/arXiv.1706.03762.
12. Dollar, O., Joshi, N., Beck, D.A.C., and Pfaendtner, J. (2021). Attention-based generative models for de novo molecular design. Chem. Sci. *12*, 8362–8372. https://doi.org/10.1039/d1sc01050f.
13. Liu, X., Ye, K., van Vlijmen, H.W.T., IJzerman, A.P., and van Westen, G.J.P. (2023). DrugEx v3: scaffold-constrained drug design with graph transformer-based reinforcement learning. J. Cheminf. *15*, 24. https://doi.org/10.1186/s13321-023-00694-z.
14. Kaitoh, K., and Yamanishi, Y. (2022). Scaffold-retained structure generator to exhaustively create molecules in an arbitrary chemical space. J. Chem. Inf. Model. *62*, 2212–2225. https://doi.org/10.1021/acs.jcim.1c01130.
15. Anighoro, A., Bajorath, J., and Rastelli, G. (2014). Polypharmacology: challenges and opportunities in drug discovery. J. Med. Chem. *57*, 7874–7887. https://doi.org/10.1021/jm5006463.
16. Hopkins, A.L. (2008). Network pharmacology: the next paradigm in drug discovery. Nat. Chem. Biol. *4*, 682–690. https://doi.org/10.1038/nchembio.118.
17. Peters, J.U. (2013). Polypharmacology – Foe or Friend? J. Med. Chem. *56*, 8955–8971. https://doi.org/10.1021/jm400856t.
18. Morphy, R., Kay, C., and Rankovic, Z. (2004). From magic bullets to designed multiple ligands. Drug Discov. Today *9*, 641–651. https://doi.org/10.1016/S1359-6446(04)03163-0.
19. Proschak, E., Stark, H., and Merk, D. (2019). Polypharmacology by design: a medicinal chemist's perspective on multitargeting compounds. J. Med. Chem. *62*, 420–444. https://doi.org/10.1021/acs.jmedchem.8b00760.

20. Hopkins, A.L., Mason, J.S., and Overington, J.P. (2006). Can we rationally design promiscuous drugs? COSB *16*, 127–136. https://doi.org/10.1016/j.sbi.2006.01.013.

21. Cavalli, A., Bolognesi, M.L., Minarini, A., Rosini, M., Tumiatti, V., Recanatini, M., and Melchiorre, C. (2008). Multi-target-directed ligands to combat neurodegenerative diseases. J. Med. Chem. *51*, 347–372. https://doi.org/10.1021/jm7009364.

22. Pennington, L.D., Aquila, B.M., Choi, Y., Valiulin, R.A., and Muegge, I. (2020). Positional analogue scanning: an effective strategy for multiparameter optimization in drug design. J. Med. Chem. *63*, 8956–8976. https://doi.org/10.1021/acs.jmedchem.9b02092.

23. Undare, S.S., Valekar, N.J., Patravale, A.A., Jamale, D.K., Vibhute, S.S., Walekar, L.S., Kolekar, G.B., Deshmukh, M.B., and Anbhule, P.V. (2016). Synthesis, anti-inflammatory, ulcerogenic and cyclooxygenase activities of Indenopyrimidine derivatives. Bioorg. Med. Chem. Lett. *26*, 814–818. https://doi.org/10.1016/j.bmcl.2015.12.088.

24. Limanto, J., Desmond, R.A., Gauthier, D.R., Devine, P.N., Reamer, R.A., and Volante, R.P. (2003). A Regioselective Approach to 5-substituted-3-amino-1,2,4-triazines. Org. Lett. *5*, 2271–2274. https://doi.org/10.1021/ol034602+.

25. Congreve, M., Andrews, S.P., Doré, A.S., Hollenstein, K., Hurrell, E., Langmead, C.J., Mason, J.S., Ng, I.W., Tehan, B., Zhukov, A., et al. (2012). Discovery of 1,2,4-triazine derivatives as adenosine A2A antagonists using structure based drug design. J. Med. Chem. *55*, 1898–1903. https://doi.org/10.1021/jm201376w.

26. Lisurek, M., Rupp, B., Wichard, J., Neuenschwander, M., von Kries, J.P., Frank, R., Kühne, R., and Kühne, R. (2010). Design of chemical libraries with potentially bioactive molecules applying a maximum common substructure concept. Mol. Divers. *14*, 401–408. https://doi.org/10.1007/s11030-009-9187-z.

27. Siddig, L.A., Bojesomo, R., Khasawneh, M.A., Samadi, A., Paz, A.P., Saadeh, H.A., and Saleh, N. (2022). Bz-8HQ: a novel supramolecular fluorochrome exhibiting multiple stimuli-responsiveness. New J. Chem. *46*, 385–397. https://doi.org/10.1039/D1NJ04998D.

28. Varani, K., Gessi, S., Dalpiaz, A., and Borea, P.A. (1996). Pharmacological and biochemical characterization of purified A2a adenosine receptors in human platelet membranes by [3H]-CGS 21680 binding. Br. J. Pharmacol. *117*, 1693–1701. https://doi.org/10.1111/j.1476-5381.1996.tb15341.x.

29. Houslay, M.D. (2005). The long and short of vascular smooth muscle phosphodiesterase-4 as a putative therapeutic target. Mol. Pharmacol. *68*, 563–567. https://doi.org/10.1124/mol.105.015719.

30. Mackenzie, S.J., and Houslay, M.D. (2000). Action of rolipram on specific PDE4 cAMP phosphodiesterase isoforms and on the phosphorylation of cAMP-response-element-binding protein (CREB) and p38 mitogen-activated protein (MAP) kinase in U937 monocytic cells. Biochem. J. *347*, 571–578. https://doi.org/10.1042/0264-6021:3470571.

31. Ford, A.P., Daniels, D.V., Chang, D.J., Gever, J.R., Jasper, J.R., Lesnick, J.D., and Clarke, D.E. (1997). Pharmacological pleiotropism of the human recombinant α1A-adrenoceptor: implications for α1-adrenoceptor classification. Br. J. Pharmacol. *121*, 1127–1135. https://doi.org/10.1038/sj.bjp.0701207.

32. Sato, S., Hatanaka, T., Yuyama, H., Ukai, M., Noguchi, Y., Ohtake, A., Taguchi, K., Sasamata, M., and Miyata, K. (2012). Tamsulosin potently and selectively antagonizes human recombinant α1A/1D-adrenoceptors: Slow dissociation from the α1A-adrenoceptor may account for selectivity for α1A-adrenoceptor over α1B-adrenoceptor subtype. Biol. Pharm. Bull. *35*, 72–77. https://doi.org/10.1248/bpb.35.72.

33. Gleason, M.M., and Hieble, J.P. (1991). Ability of SK&F 104078 and SK&F 104856 to identify alpha-2 adrenoceptor subtypes in NCB20 cells and guinea pig lung. J. Pharmacol. Exp. Therapeut. *259*, 1124–1132.

34. Lalchandani, S.G., Lei, L., Zheng, W., Suni, M.M., Moore, B.M., Liggett, S.B., Miller, D.D., and Feller, D.R. (2002). Yohimbine dimers exhibiting selectivity for the human α2c-adrenoceptor subtype. J. Pharmacol. Exp. Therapeut. *303*, 979–984. https://doi.org/10.1124/jpet.102.039057.

35. Feve, B., Elhadri, K., Quignard-Boulange, A., and Pairault, J. (1994). Transcriptional down-regulation by insulin of the beta 3-adrenergic receptor expression in 3T3-F442A adipocytes: a mechanism for repressing the cAMP signaling pathway. Proc. Natl. Acad. Sci. USA *91*, 5677–5681. https://doi.org/10.1073/pnas.91.12.5677.

36. McCrea, K.E., and Hill, S.J. (1993). Salmeterol, a long-acting β2-adrenoceptor agonist mediating cyclic AMP accumulation in a neuronal cell line. Br. J. Pharmacol. *110*, 619–626. https://doi.org/10.1111/j.1476-5381.1993.tb13856.x.

37. Jung, M., Calassi, R., Rinaldi-Carmona, M., Chardenot, P., Le Fur, G., Soubrié, P., and Oury-Donat, F. (1997). Characterization of CB1 receptors on rat neuronal cell cultures: binding and functional studies using the selective receptor antagonist SR 141716A. J. Neurochem. *68*, 402–409. https://doi.org/10.1046/j.1471-4159.1997.68010402.x.

38. Melck, D., De Petrocellis, L., Orlando, P., Bisogno, T., Laezza, C., Bifulco, M., and Di Marzo, V. (2000). Suppression of nerve growth factor Trk receptors and prolactin receptors by endocannabinoids leads to inhibition of human breast and prostate cancer cell proliferation. Endocrinology *141*, 118–126. https://doi.org/10.1210/endo.141.1.7239.

39. Munro, S., Thomas, K.L., and Abu-Shaar, M. (1993). Molecular characterization of a peripheral receptor for cannabinoids. Nature *365*, 61–65. https://doi.org/10.1038/365061a0.

40. Bignon, E., Bachy, A., Boigegrain, R., Brodin, R., Cottineau, M., Gully, D., Herbert, J.M., Keane, P., Labie, C., Molimard, J.C., et al. (1999). SR146131: a new potent, orally active, and selective nonpeptide cholecystokinin subtype 1 receptor agonist. I: in vitro studies. J. Pharmacol. Exp. Therapeut. *289*, 742–751.

41. Gao, F., Sexton, P.M., Christopoulos, A., and Miller, L.J. (2008). Benzodiazepine ligands can act as allosteric modulators of the Type 1 cholecystokinin receptor. Bioorg. Med. Chem. Lett. *18*, 4401–4404. https://doi.org/10.1016/j.bmcl.2008.06.053.

42. Dearry, A., Gingrich, J.A., Falardeau, P., Fremeau, R.T., Bates, M.D., and Caron, M.G. (1990). Molecular cloning and expression of the gene for a human D1 dopamine receptor. Nature *347*, 72–76. https://doi.org/10.1038/347072a0.

43. Zhou, Q.Y., Grandy, D.K., Thambi, L., Kushner, J.A., Van Tol, H.H., Cone, R., Pribnow, D., Salon, J., Bunzow, J.R., and Civelli, O. (1990). Cloning and expression of human and rat D2 dopamine receptors. Nature *347*, 76–80. https://doi.org/10.1038/347076a0.

44. Grandy, D.K., Marchionni, M.A., Makam, H., Stofko, R.E., Alfano, M., Frothingham, L., Fischer, J.B., Burke-Howie, K.J., Bunzow, J.R., Server, A.C., et al. (1989). Cloning of the cDNA and gene for a human D2 dopamine receptor. Proc. Natl. Acad. Sci. USA *86*, 9762–9766. https://doi.org/10.1073/pnas.86.24.9762.

45. Hayes, G., Biden, T.J., Selbie, L.A., and Shine, J. (1992). Structural subtypes of the dopamine D2 receptor are functionally distinct: expression of the cloned D2A and D2B subtypes in a heterologous cell line. Mol. Endocrinol. *6*, 920–926. https://doi.org/10.1210/mend.6.6.1323056.

46. Mihara, S., Nakajima, S., Matumura, S., Kohnoike, T., and Fujimoto, M. (1994). Pharmacological characterization of a potent nonpeptide endothelin receptor antagonist, 97-139. J. Pharmacol. Exp. Therapeut. *268*, 1122–1128.

47. De Backer, M.D., Gommeren, W., Moereels, H., Nobels, G., Van Gompel, P., Leysen, J.E., and Luyten, W.H. (1993). Genomic cloning, heterologous expression and pharmacological characterization of a human histamine H1 receptor. Biochem. Biophys. Res. Commun. *197*, 1601–1608. https://doi.org/10.1006/bbrc.1993.2662.

48. Ruat, M., Traiffort, E., Bouthenet, M.L., Schwartz, J.C., Hirschfeld, J., Buschauer, A., and Schunack, W. (1990). Reversible and irreversible labeling and autoradiographic localization of the cerebral histamine H2 receptor using [125I] iodinated probes. Proc. Natl. Acad. Sci. USA *87*, 1658–1662. https://doi.org/10.1073/pnas.87.5.1658.

49. Buckley, N.J., Bonner, T.I., Buckley, C.M., and Brann, M.R. (1989). Antagonist binding properties of five cloned muscarinic receptors expressed in CHO-K1 cells. Mol. Pharmacol. *35*, 469–476.

50. Luthin, G.R., and Wolfe, B.B. (1984). Comparison of [3H] pirenzepine and [3H] quinuclidinylbenzilate binding to muscarinic cholinergic receptors in rat brain. J. Pharmacol. Exp. Therapeut. *228*, 648–655.

51. Clark, M.J., Emmerson, P.J., Mansour, A., Akil, H., Woods, J.H., Portoghese, P.S., Remmers, A.E., and Medzihradsky, F. (1997). Opioid efficacy in a C6 glioma cell line stably expressing the delta opioid receptor. J. Pharmacol. Exp. Therapeut. *283*, 501–510.

52. Martin, N.A., and Prather, P.L. (2001). Interaction of co-expressed μ- and δ-opioid receptors in transfected rat pituitary GH3 cells. Mol. Pharmacol. *59*, 774–783. https://doi.org/10.1124/mol.59.4.774.

53. Maguire, P., Tsai, N., Kamal, J., Cometta-Morini, C., Upton, C., and Loew, G. (1992). Pharmacological profiles of fentanyl analogs at μ, δ and κ opiate receptors. Eur. J. Pharmacol. *213*, 219–225. https://doi.org/10.1016/0014-2999(92)90685-w.

54. Simonin, F., Gavériaux-Ruff, C., Befort, K., Matthes, H., Lannes, B., Micheletti, G., Mattéi, M.G., Charron, G., Bloch, B., and Kieffer, B. (1995). kappa-Opioid receptor in humans: cDNA and genomic cloning, chromosomal assignment, functional expression, pharmacology, and expression pattern in the central nervous system. Proc. Natl. Acad. Sci. USA *92*, 7006–7010. https://doi.org/10.1073/pnas.92.15.7006.

55. Wang, J.B., Johnson, P.S., Persico, A.M., Hawkins, A.L., Griffin, C.A., and Uhl, G.R. (1994). Human mu opiate receptor: cDNA and genomic clones, pharmacologic characterization and chromosomal assignment. FEBS Lett. *338*, 217–222.

56. Martin, G.R., and Humphrey, P.P. (1994). Receptors for 5-hydroxytryptamine: current perspectives on classification and nomenclature. Neuropharmacology *33*, 261–273. https://doi.org/10.1016/0028-3908(94)90058-2.

57. May, J.A., McLaughlin, M.A., Sharif, N.A., Hellberg, M.R., and Dean, T.R. (2003). Evaluation of the ocular hypotensive response of serotonin 5-HT1A and 5-HT2 receptor ligands in conscious ocular hypertensive cynomolgus monkeys. J. Pharmacol. Exp. Therapeut. *306*, 301–309. https://doi.org/10.1124/jpet.103.049528.

58. Maier, D.L., Sobotka-Briner, C., Ding, M., Powell, M.E., Jiang, Q., Hill, G., Heys, J.R., Elmore, C.S., Pierson, M.E., and Mrzljak, L. (2009). [N-methyl-3H3] AZ10419369 binding to the 5-HT1B receptor: in vitro characterization and in vivo receptor occupancy. J. Pharmacol. Exp. Therapeut. *330*, 342–351. https://doi.org/10.1124/jpet.109.150722.

59. Xie, Z., Lee, S.P., O'Dowd, B.F., and George, S.R. (1999). Serotonin 5-HT1B and 5-HT1D receptors form homodimers when expressed alone and heterodimers when co-expressed. FEBS Lett. *456*, 63–67. https://doi.org/10.1016/s0014-5793(99)00918-7.

60. Bonhaus, D.W., Bach, C., DeSouza, A., Salazar, F.H., Matsuoka, B.D., Zuppan, P., Chan, H.W., and Eglen, R.M. (1995). The pharmacology and distribution of human 5-hydroxytryptamine2B (5-HT2B) receptor gene products: comparison with 5-HT2A and 5-HT2C receptors. Br. J. Pharmacol. *115*, 622–628. https://doi.org/10.1111/j.1476-5381.1995.tb14977.x.

61. Saucier, C., and Albert, P.R. (1997). Identification of an endogenous 5-hydroxytryptamine2A receptor in NIH-3T3 cells: agonist-induced down-regulation involves decreases in receptor RNA and number. J. Neurochem. *68*, 1998–2011. https://doi.org/10.1046/j.1471-4159.1997.68051998.x.

62. Thibonnier, M., Auzan, C., Madhun, Z., Wilkins, P., Berti-Mattera, L., and Clauser, E. (1994). Molecular cloning, sequencing, and functional expression of a cDNA encoding the human V1a vasopressin receptor. J. Biol. Chem. *269*, 3304–3310.

63. Giros, B., and Caron, M.G. (1993). Molecular characterization of the dopamine transporter. Trends Pharmacol. Sci. *14*, 43–49. https://doi.org/10.1016/0165-6147(93)90029-j.

64. Gu, H., Wall, S.C., and Rudnick, G. (1994). Stable expression of biogenic amine transporters reveals differences in inhibitor sensitivity, kinetics, and ion dependence. J. Biol. Chem. *269*, 7124–7130.

65. Galli, A., DeFelice, L.J., Duke, B.J., Moore, K.R., and Blakely, R.D. (1995). Sodium-dependent norepinephrine-induced currents in norepinephrine-transporter-transfected HEK-293 cells blocked by cocaine and antidepressants. J. Exp. Biol. *198*, 2197–2212. https://doi.org/10.1242/jeb.198.10.2197.

66. Shearman, L.P., McReynolds, A.M., Zhou, F.C., and Meyer, J.S. (1998). Relationship between [125I] RTI-55-labeled cocaine binding sites and the serotonin transporter in rat placenta. Am. J. Physiol. *275*, C1621–C1629. https://doi.org/10.1152/ajpcell.1998.275.6.C1621.

67. Wolf, W.A., and Kuhn, D.M. (1992). Role of essential sulfhydryl groups in drug interactions at the neuronal 5-HT transporter. Differences between amphetamines and 5-HT uptake inhibitors. J. Biol. Chem. *267*, 20820–20825.

68. Gopalakrishnan, M., Monteggia, L.M., Anderson, D.J., Molinari, E.J., Piattoni-Kaplan, M., Donnelly-Roberts, D., Arneric, S.P., and Sullivan, J.P. (1996). Stable expression, pharmacologic properties and regulation of the human neuronal nicotinic acetylcholine alpha 4 beta 2 receptor. J. Pharmacol. Exp. Therapeut. *276*, 289–297.

69. Huang, X.P., Mangano, T., Hufeisen, S., Setola, V., and Roth, B.L. (2010). Identification of human Ether-à-go-Go related gene modulators by three screening platforms in an academic drug-discovery setting. Assay Drug Dev. Technol. *8*, 727–742. https://doi.org/10.1089/adt.2010.0331.

70. Finlayson, K., Turnbull, L., January, C.T., Sharkey, J., and Kelly, J.S. (2001). [3H] dofetilide binding to HERG transfected membranes: a potential high throughput preclinical screen. Eur. J. Pharmacol. *430*, 147–148. https://doi.org/10.1016/s0014-2999(01)01362-0.

71. Zhou, Z., Gong, Q., Ye, B., Fan, Z., Makielski, J.C., Robertson, G.A., and January, C.T. (1998). Properties of HERG channels stably expressed in HEK 293 cells studied at physiological temperature. Biophys. J. *74*, 230–241. https://doi.org/10.1016/S0006-3495(98)77782-3.

72. Boess, F.G., Steward, L.J., Steele, J.A., Liu, D., Reid, J., Glencorse, T.A., and Martin, I.L. (1997). Analysis of the ligand binding site of the 5-HT3 receptor using site directed mutagenesis: importance of glutamate *106*. Neuropharmacology *36*, 637–647. https://doi.org/10.1016/s0028-3908(97)00044-0.

73. Miller, K., Weisberg, E., Fletcher, P.W., and Teitler, M. (1992). Membrane-bound and solubilized brain 5-HT3 receptors: Improved radioligand binding assays using bovine area postrema or rat cortex and the radioligands [3H]-Gr65630C, [3H]-BRI43694, and [3H]-LY278584. Synapse *11*, 58–66. https://doi.org/10.1002/syn.890110108.

74. Kanda, H., Mimura, T., Hamasaki, K., Yamamoto, K., Yazaki, Y., Hirai, H., and Nojima, Y. (1999). Fyn and Lck tyrosine kinases regulate tyrosine phosphorylation of p105CasL, a member of the p130Cas docking protein family, in T-cell receptor-mediated signalling. Immunology *97*, 56–61. https://doi.org/10.1046/j.1365-2567.1999.00753.x.

75. Traish, A.M., Müller, R.E., and Wotiz, H.H. (1986). Binding of 7 α, 17 α-dimethyl-19-nortestosterone (mibolerone) to androgen and progesterone receptors in human and animal tissues. Endocrinology *118*, 1327–1333. https://doi.org/10.1210/endo-118-4-1327.

76. Honer, C., Nam, K., Fink, C., Marshall, P., Ksander, G., Chatelain, R.E., Cornell, W., Steele, R., Schweitzer, R., and Schumacher, C. (2003). Glucocorticoid receptor antagonism by cyproterone acetate and RU486. Mol. Pharmacol. *63*, 1012–1020. https://doi.org/10.1124/mol.63.5.1012.

77. Mulatero, P., Panarelli, M., Schiavone, D., Rossi, A., Mengozzi, G., Kenyon, C.J., Chiandussi, L., and Veglio, F. (1997). Impaired cortisol binding to glucocorticoid receptors in hypertensive patients. Hypertension *30*, 1274–1278. https://doi.org/10.1161/01.hyp.30.5.1274.

78. Chan, C.C., Boyce, S., Brideau, C., Charleson, S., Cromlish, W., Ethier, D., Evans, J., Ford-Hutchinson, A.W., Forrest, M.J., Gauthier, J.Y.,

et al. (1999). Rofecoxib [Vioxx, MK-0966; 4-(4′-methylsulfonylphenyl)-3-phenyl-2-(5H)-furanone]: a potent and orally active cyclooxygenase-2 inhibitor. Pharmacological and biochemical profiles. J. Pharmacol. Exp. Therapeut. *290*, 551–560.

79. Swinney, D.C., Mak, A.Y., Barnett, J., and Ramesha, C.S. (1997). Differential allosteric regulation of prostaglandin H synthase 1 and 2 by arachidonic acid. J. Biol. Chem. *272*, 12393–12398. https://doi.org/10.1074/jbc.272.19.12393.

80. Riendeau, D., Charleson, S., Cromlish, W., Mancini, J.A., Wong, E., and Guay, J. (1997). Comparison of the cyclooxygenase-1 inhibitory properties of nonsteroidal anti-inflammatory drugs and selective COX-2 inhibitors, using sensitive microsomal and platelet assays. Can. J. Physiol. Pharmacol. *75*, 1088–1095.

81. Warner, T.D., Giuliano, F., Vojnovic, I., Bukasa, A., Mitchell, J.A., and Vane, J.R. (1999). Nonsteroid drug selectivities for cyclo-oxygenase-1 rather than cyclo-oxygenase-2 are associated with human gastrointestinal toxicity: a full in vitro analysis. Proc. Natl. Acad. Sci. USA *96*, 7563–7568. https://doi.org/10.1073/pnas.96.13.7563.

82. Ellman, G.L., Courtney, K.D., ANDRES, V., Jr., and Featherstone, R.M. (1961). A new and rapid colorimetric determination of acetylcholinesterase activity. Biochem. Pharmacol. *7*, 88–95. https://doi.org/10.1016/0006-2952(61)90145-9.

83. Nadarajah, B. (1992). The effect of pralidoxime chloride in the assay of acetylcholinesterase using 5, 5′-dithio-bis (2-nitrobenzoic acid) (Ellman's reagent). J. Anal. Toxicol. *16*, 192–193. https://doi.org/10.1093/jat/16.3.192.

84. Urban, P., Andersen, J.K., Hsu, H.P., and Pompon, D. (1991). Comparative membrane locations and activities of human monoamine oxidases expressed in yeast. FEBS Lett. *286*, 142–146. https://doi.org/10.1016/0014-5793(91)80960-b.

85. Youdim, M.B., and Finberg, J.P. (1991). New directions in monoamine oxidase A and B selective inhibitors and substrates. Biochem. Pharmacol. *41*, 155–162. https://doi.org/10.1016/0006-2952(91)90471-g.

86. Hambleton, R., Krall, J., Tikishvili, E., Honeggar, M., Ahmad, F., Manganiello, V.C., and Movsesian, M.A. (2005). Isoforms of cyclic nucleotide phosphodiesterase PDE3 and their contribution to cAMP hydrolytic activity in subcellular fractions of human myocardium. J. Biol. Chem. *280*, 39168–39174. https://doi.org/10.1074/jbc.M506760200.

87. Hung, S.-H., Zhang, W., Pixley, R.A., Jameson, B.A., Huang, Y.C., Colman, R.F., and Colman, R.W. (2006). New insights from the structure-function analysis of the catalytic region of human platelet phosphodiesterase 3A: a role for the unique 44-amino acid insert. J. Biol. Chem. *281*, 29236–29244. https://doi.org/10.1074/jbc.M606558200.

88. Busse, W.W., and Lemanske, R.F., Jr. (2001). Asthma. N. Engl. J. Med. *344*, 350–362. https://doi.org/10.1056/NEJM200102013440507.

89. Fan Chung, K. (2006). Phosphodiesterase inhibitors in airways disease. Eur. J. Pharmacol. *533*, 110–117. https://doi.org/10.1016/j.ejphar.2005.12.059.

90. Brown, R.A., Spina, D., and Page, C.P. (2008). Adenosine receptors and asthma. Br. J. Pharmacol. *153*, S446–S456. https://doi.org/10.1038/bjp.2008.22.

91. Berenger, F., and Tsuda, K. (2021). Molecular generation by Fast Assembly of (Deep)SMILES fragments. J. Cheminf. *13*, 88. https://doi.org/10.1186/s13321-021-00566-4.

92. Carhart, R.E., Smith, D.H., and Venkataraghavan, R. (1985). Atom pairs as molecular features in structure-activity studies: definition and applications. J. Chem. Inf. Comput. Sci. *25*, 64–73. https://doi.org/10.1021/ci00046a002.

93. Fan, R.E., Chang, K.W., Hsieh, C.J., Wang, X.R., and Lin, C.J. (2008). LIBLINEAR: A library for large linear classification. J. Mach. Learn. Res. *9*, 1871–1874.

94. Abad-Zapatero, C., and Metz, J.T. (2005). Ligand efficiency indices as guideposts for drug discovery. Drug Discov. Today *10*, 464–469. https://doi.org/10.1016/S1359-6446(05)03386-6.

95. Hann, M.M., and Oprea, T.I. (2004). Pursuing the leadlikeness concept in pharmaceutical research. Curr. Opin. Chem. Biol. *8*, 255–263. https://doi.org/10.1016/j.cbpa.2004.04.003.

96. Falcón-Cano, G., Molina, C., and Cabrera-Pérez, M.Á. (2020). ADME prediction with KNIME: development and validation of a publicly available workflow for the prediction of human oral bioavailability. J. Chem. Inf. Model. *60*, 2660–2667. https://doi.org/10.1021/acs.jcim.0c00019.

97. Ertl, P., and Schuffenhauer, A. (2009). Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. J. Cheminf. *1*, 8. https://doi.org/10.1186/1758-2946-1-8.

98. Li, C., Yamanaka, C., Kaitoh, K., and Yamanishi, Y. (2022). Transformer-based objective-reinforced generative adversarial network to generate desired molecules. In IJCAI-ECAI, pp. 3884–3890. https://doi.org/10.24963/ijcai.2022/539.

99. Sutton, R.S., McAllester, D., Singh, S., and Mansour, Y. (2022). Policy gradient methods for reinforcement learning with function approximation. Adv. Neural Inf. Process. Syst. *12*, 1057–1063. https://doi.org/10.48550/arXiv.1706.06643.

100. Arjovsky, M., Chintala, S., and Bottou, L. (2017). Wasserstein Generative Adversarial Networks. In Proceedings of the 34th International Conference on Machine Learning, *70*, pp. 214–223. https://doi.org/10.48550/arXiv.1701.07875.

101. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al. (2019). PyTorch: an imperative style, high-performance deep learning library. Neural Inf. Process Syst. *27*, 8026–8037. https://doi.org/10.48550/arXiv.1912.01703.

102. Kingma, D.P., and Ba, J.L. (2015). Adam: A method for stochastic optimization. In The 3rd International Conference for Learning Representations. https://doi.org/10.48550/arXiv.1412.6980.

103. Mendez, D., Gaulton, A., Bento, A.P., Chambers, J., De Veij, M., Félix, E., Magariños, M.P., Mosquera, J.F., Mutowo, P., Nowotka, M., et al. (2019). ChEMBL: towards direct deposition of bioassay data. Nucleic Acids Res. *47*, 930–940. https://doi.org/10.1093/nar/gky1075.

104. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Müller, A., Nothman, J., Louppe, G., et al. (2011). Scikit-learn: machine learning in Python. J. Mach. Learn. Res. *12*, 2825–2830. https://doi.org/10.48550/arXiv.1201.0490.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
Python (DualFASMIFRA)	Version 3.9.13	https://www.python.org/downloads/
Python (DualTransORGAN)	Version 3.6.10	https://www.python.org/downloads/
RDKit (DualFASMIFRA)	Version 2024.03.5	https://www.rdkit.org
RDkit (DualTransORGAN)	Version 2020.09.1.0	https://www.rdkit.org
PyTorch (DualTransORGAN)	Version 3.6.10	https://pytorch.org
Data		
Compound structures	ChEMBL	https://www.ebi.ac.uk/chembl/ (https://doi.org/10.1093/nar/gky1075)
Bioactivity data for ADORA2A	ChEMBL	https://www.ebi.ac.uk/chembl/target_report_card/CHEMBL251/
Bioactivity data for PDE4D	ChEMBL	https://www.ebi.ac.uk/chembl/target_report_card/CHEMBL288/
Other		
DualFASMIFRA	This study	https://github.com/UnixJunkie/FASMIFRA/ (https://doi.org/10.1186/s13321-021-00566-4)
DualTransORGAN	This study	https://yamanishi.cs.i.nagoya-u.ac.jp/dualtarget/

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Synthesis of compounds

Unless otherwise stated, all reactants or reagents, including dry solvents, were obtained from commercial suppliers and used as received. 2,2-dimorpholino-1-phenylethan-1-one was prepared according to previous report.[23] All reactions were performed using dry solvents under a nitrogen atmosphere in dried glassware using standard vacuum-line techniques. All preparatory and purification procedures were carried out with reagent-grade solvents under air.

Analytical thin-layer chromatography was performed using Merck silica gel 60 F_{254} precoated plates (0.25 mm). The developed chromatograms were analyzed using an untra violet (UV) lamp (254 or 365 nm). Flash column chromatography was performed with KANTO Silica Gel 60N (spherical, neutral, 40–100 μm) or Biotage Isolera equipped with Biotage SNAP Cartridge KP-Sil columns. Preparative thin-layer chromatography (PTLC) was performed using Wakogel B5-F silica coated plates (0.75 mm) prepared in our laboratory. High-resolution mass spectra (HRMS) were obtained from a Thermo Fisher Scientific Exactive (atmospheric pressure chemical ionization, electrospray ionization, ESI). Nuclear magnetic resonance (NMR) spectra were recorded on a JEOL JNM-ECA-500 (H 500 MHz) and JEOL ECS-600 (H 600 MHz, C 150 MHz) spectrometers. Chemical shifts for H NMR are expressed in parts per million (ppm) relative to CD_3OD (3.31 ppm), *d*-DMSO (2.50 ppm), CD_2Cl_2 (5.32 ppm), or $CDCl_3$ (7.26 ppm). Chemical shifts for C NMR are expressed in ppm relative to CD_3OD (49.00 ppm), CD_2Cl_2 (53.84 ppm), or $CDCl_3$ (77.16 ppm). Data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, sext = sextet, m = multiplet), coupling constant (Hz), and integration.

Synthesis of 2-amino-4-(furan-2-yl)-5*H*-indeno[1,2-*d*ay]pyrimidin-5-one (5)

To a 50-mL three-neck round bottom flask containing a magnetic stirring bar were added furan-2-carbaldehyde (**3**) (1.5 mL, 18.1 mmol), 1,3-indanedione (**1**) (2.58 g, 17.7 mmol), NaOH (146 g, 36.5 mmol), ethanol (50 mL) and water (50 mL), and the mixture was stirred for 30 min at room temperature. After guanidine hydrochloride (**2**) (2.56 g, 26.8 mmol) was added, the resulting mixture was heated at 90°C for 32 h. The reaction mixture was poured onto crushed ice. A precipitate was collected by filtration and washed with cold water, dried and purified by flash column chromatography using Isolera (chloroform/MeOH = 10:0 to 10:1) and PTLC (chloroform/EtOAc = 2:1) to afford **5** as a yellow solid (74.6 mg, 2% yield).

2-Amino-4-(furan-2-yl)-5*H*-indeno[1,2-*d*ay]pyrimidin-5-one (5)

H NMR (600 MHz, $CDCl_3$) δ 8.61 (d, *J* = 3.4 Hz, 1H), 7.82 (d, *J* = 7.2 Hz, 1H), 7.76 (d, *J* = 7.2 Hz, 1H), 7.72–7.71 (m, 1H), 7.61 (td, *J* = 7.4, 1.3 Hz, 1H), 7.57 (td, *J* = 7.3, 1.1 Hz, 1H), 6.67 (dd, *J* = 3.6, 1.8 Hz, 1H), 5.82 (s, 2H); C NMR (150 MHz, $CDCl_3$) δ 188.3, 176.9, 164.8, 153.1, 149.6, 146.1, 140.0, 136.8, 134.1, 133.0, 123.7, 121.4, 120.8, 113.0, 110.7.

HRMS (ESI) *m/z* calcd for $C_{15}H_9N_3O_2Na$ $[M+Na]^+$: 286.0587, found 286.0586.

Synthesis of 2-amino-4-phenyl-5*H*-indeno[1,2-*d*ay]pyrimidin-5-one (6)

Compound **6** was synthesized according to procedures reported in the literature.[23] To a Schlenk tube containing a magnetic stirring bar were added benzaldehyde (**4**) (0.11 mL, 1.09 mmol), 1,3-indanedione (**1**) (147 mg, 1.01 mmol), NaOH (81 mg, 1.09 mmol), ethanol (1.5 mL) and water (1.5 mL), and the mixture was stirred for 10 min at room temperature. After guanidine hydrochloride (**2**) (143 mg, 1.50 mmol) was added, the resulting mixture was stirred at 90°C for 11 h. The reaction mixture was poured onto crushed ice. A precipitate was collected by filtration, washed with cold water, and recrystallized from hexane/MeOH to afford **6** as a yellow solid (25.4 mg, 9% yield).

2-Amino-4-phenyl-5*H*-indeno[1,2-*d*ay]pyrimidin-5-one (6)

H NMR (600 MHz, $CDCl_3$) δ 8.08–8.06 (m, 2H), 7.86–7.85 (m, 1H), 7.75–7.74 (m, 1H), 7.62 (td, *J* = 7.5, 1.1 Hz, 1H), 7.57 (td, *J* = 7.5, 1.1 Hz, 1H), 7.54–7.51 (m, 3H), 5.69 (s, 2H); C NMR (150 MHz, $CDCl_3$) δ 188.6, 176.9, 165.9, 164.7, 140.0, 136.7, 135.5, 134.2, 133.0, 131.4, 129.8, 128.2, 123.8, 121.4, 113.2.

HRMS (ESI) *m/z* calcd for $C_{17}H_{11}N_3ONa$ $[M+Na]^+$: 296.0794, found 286.0797.

Synthesis of 5-phenyl-1,2,4-triazin-3-amine (7)

Compound **7** was synthesized according to procedures reported in the literature.[24] To a 100-mL two-neck round bottom flask containing a magnetic stirring bar were added 2,2-dimorpholino-1-phenylethan-1-one (**S1**) (7.30 g, 24.4 mmol), aminoguanidine bicarbonate (**S2**) (3.32 g, 24.4 mmol) and MeOH (35 mL), followed by slow addition of AcOH (4.2 mL, 73.4 mmol). The resulting mixture was stirred at room temperature for 10 min, and then heated 65°C for 21 h. After cooling, the resulting suspension was concentrated to about half its volume, and cooled to 0°C. The resulting solid was filtered and washed with cold MeOH/H_2O (4:1). The collected solid was dried *in vacuo* to afford **7** as an orange solid (2.23 g, 53% yield).

5-Phenyl-1,2,4-triazin-3-amine (7)

H NMR (600 MHz, *d*-DMSO) δ 9.23 (s, 1H), 8.17 (dt, *J* = 6.5, 1.6 Hz, 2H), 7.60–7.55 (m, 3H), 7.25 (s, 2H); C NMR (150 MHz, CD_3OD) δ 164.7, 158.6, 138.2, 135.3, 133.3, 130.2, 128.6.

HRMS (ESI) *m/z* calcd for $C_9H_8N_4Na$ $[M+Na]^+$: 195.0641, found 195.0641.

The H-NMR spectrum was identical to that reported in the literature.[24]

Synthesis of 6-bromo-5-phenyl-1,2,4-triazin-3-amine (8)

Compound **8** was synthesized according to procedures reported in the literature.[25] To a 100-mL two-neck round bottom flask containing a magnetic stirring bar were added 5-phenyl-1,2,4-triazin-3-amine (**7**) (2.10 mg, 12.2 mmol) and DMF (20 mL). The resulting mixture was cooled to −25°C and a solution of *N*-bromosuccinimide (6.52 g, 36.6 mmol) in DMF (15 mL) was added dropwise. The reaction mixture was stirred at room temperature for 18 h. After completion of the reaction, the mixture was poured into sat. $NaHCO_3$ aq. and extracted with diethyl ether. The organic layer was washed with water and brine, and dried over Na_2SO_4. After filtration, the solvent was removed by evaporation. The resulting residue was purified by flash column chromatography using Isolera (hexane/EtOAc = 8:2 to 1:1) to afford **8** as a yellow solid (1.05 g, 34% yield).

6-Bromo-5-phenyl-1,2,4-triazin-3-amine (8)

H NMR (600 MHz, *d*-DMSO) δ 7.74 (dt, *J* = 6.4, 1.7 Hz, 2H), 7.57–7.52 (m, 5H); C NMR (150 MHz, CD_3OD) δ 163.9, 161.0, 136.6, 135.8, 132.0, 130.5, 129.3.

HRMS (ESI) *m/z* calcd for $C_9H_7N_4BrNa$ $[M+Na]^+$: 272.9746, found 272.9747.

The H-NMR spectrum was identical to that reported in the literature.[25]

Synthesis of 5-phenyl-6-(pyridin-4-yl)-1,2,4-triazin-3-amine (9)

To a 50-mL three-neck round bottom flask containing a magnetic stirring bar were added 6-bromo-5-phenyl-1,2,4-triazin-3-amine (**8**) (104.4 mg, 0.42 mmol), 4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)pyridine (**S3**) (96.2 mg, 0.45 mmol), K_2CO_3 (171.2 mg, 1.24 mmol) and 1,4-dioxane (3.0 mL). The resulting mixture was diluted with water (1.6 mL), treated with $Pd(PPh_3)_4$ (25 mg, 0.023 mmol), and stirred at 100°C for 19 h. Then, the mixture was poured into water and extracted with EtOAc. The organic layer was washed with water and brine, and dried over Na_2SO_4. After filtration, the solvent was removed by evaporation. The resulting residue was purified by PTLC ($CHCl_3$/MeOH = 10:1) and recrystallized from $CHCl_3$/hexane to afford **9** as a yellow solid (24.4 mg, 7% yield).

5-Phenyl-6-(pyridin-4-yl)-1,2,4-triazin-3-amine (9)

H NMR (600 MHz, CD_2Cl_2) δ 8.54 (dd, *J* = 4.5, 1.4 Hz, 2H), 7.48–7.45 (m, 3H), 7.37–7.33 (m, 4H), 5.58 (s, 2H); C NMR (150 MHz, $CDCl_3$) δ 161.3, 157.6, 149.9, 148.3, 143.9, 135.2, 130.9, 129.4, 128.7, 123.5.

HRMS (ESI) *m/z* calcd for $C_{14}H_{12}N_5$ $[M + H]^+$: 250.1087, found 250.1086.

Synthesis of *tert*-butyl 4-(1*H*-benzo[*d*]imidazol-2-yl)piperazine-1-carboxylate (11)

Compound **11** was synthesized according to procedures reported in the literature.[27] To a 200-mL three-neck round bottom flask containing a magnetic stirring bar were added 2-chlorobenzimidazole (**10**) (3.00 g, 19.7 mmol), *N*-*tert*-butoxycarbonyl piperazine (**S4**) (3.70 g, 19.9 mmol), and 1-butanol (40 mL). The resulting mixture was refluxed for 3 h. Then, the precipitate was collected and rinsed with diethyl ether and dried over *in vacuo* to afford **11** as a white solid (5.7 g, 96% yield).

tert-butyl 4-(1*H*-benzo[*d*]imidazol-2-yl)piperazine-1-carboxylate (11)

H NMR (600 MHz, CD_3OD) δ 7.42 (dd, *J* = 5.7, 3.3 Hz, 2H), 7.32 (dd, *J* = 5.1, 3.0 Hz, 2H), 3.68 (s, 8H), 1.50 (s, 9H); C NMR (150 MHz, CD_3OD) δ 156.0, 151.8, 131.4, 125.1, 112.5, 82.1, 47.2, 28.6 (missing one peak).

HRMS (ESI) *m/z* calcd for $C_{16}H_{23}N_4O_2$ $[M + H]^+$: 303.1816, found 303.1819.

The H-NMR spectrum was identical to that reported in the literature.[27]

Synthesis of *tert*-butyl 4-(1-ethyl-1*H*-benzo[*d*]imidazol-2-yl)piperazine-1-carboxylate (12)

To a Schlenk tube containing a magnetic stirring bar were added *tert*-butyl 4-(1*H*-benzo[*d*]imidazol-2-yl)piperazine-1-carboxylate (**11**) (1.00 g, 3.31 mmol), ethyl bromide (0.39 mL, 5.01 mmol), KOH (378.0 mg, 6.74 mmol) and acetone (6.6 mL). The resulting mixture was stirred at 40°C for 19 h. Then, the mixture was poured into sat. NH_4Cl aq. and extracted with chloroform. The organic layer was dried over Na_2SO_4. After filtration, the solvent was removed by evaporation. The resulting residue was purified by flash column chromatography by Isolera (hexeane/EtOAc = 1:1 to 2:8) to afford **12** as a white solid (801 mg, 73% yield).

*tert-butyl 4-(1-ethyl-1*H*-benzo[d]imidazol-2-yl)piperazine-1-carboxylate (12)*

H NMR (600 MHz, CD_3OD) δ 7.50–7.47 (m, 1H), 7.41–7.39 (m, 1H), 7.21–7.16 (m, 2H), 4.18 (q, *J* = 7.3 Hz, 2H), 3.65 (s, 4H), 3.23 (t, *J* = 5.1 Hz, 4H), 1.50 (s, 9H), 1.44 (t, *J* = 7.1 Hz, 3H); C NMR (150 MHz, CD_3OD) δ 158.3, 156.4, 141.7, 135.4, 123.1, 123.1, 118.3, 110.8, 81.6, 51.9, 39.9, 28.6, 14.5 (missing one peak).

HRMS (ESI) *m/z* calcd for $C_{18}H_{26}N_4O_2Na$ $[M+Na]^+$: 353.1948, found 353.1943.

Synthesis of 1-ethyl-2-(piperazin-1-yl)-1*H*-benzo[*d*]imidazole hydrochloride (13)

To a screw test tube containing a magnetic stirring bar were added *tert*-butyl 4-(1-ethyl-1*H*-benzo[*d*]imidazol-2-yl)piperazine-1-carboxylate (**12**) (53.0 mg, 0.16 mmol), trifluoroacetic acid (TFA: 0.2 mL) and CH_2Cl_2 (1.8 mL). The resulting mixture was stirred at room temperature for 1 h. After addition of toluene (10 mL), the solvent was removed by evaporation to afford the product as a TFA salt. To a 30-mL flask were added the resulting residue and 4 M HCl/1,4-dioxane (2.0 mL), and the solvent was removed by evaporation. This procedure was repeated four times in total, and the resulting residue was washed with hexane to afford **13** as a white solid (38.1 mg, 89% yield).

1-Ethyl-2-(piperazin-1-yl)-1*H*-benzo[*d*]imidazole hydrochloride (13)

H NMR (500 MHz, CD_3OD) δ 7.74–7.72 (m, 1H), 7.62–7.60 (m, 1H), 7.52–7.50 (m, 2H), 4.35 (q, *J* = 7.3 Hz, 2H), 3.88 (t, *J* = 5.2 Hz, 4H), 3.56 (t, *J* = 5.0 Hz, 4H), 1.59 (t, *J* = 7.3 Hz, 3H); C NMR (150 MHz, CD_3OD) δ 152.4, 132.5, 130.2, 126.7, 126.4, 113.9, 113.0, 47.9, 44.1, 42.3, 14.1.

HRMS (ESI) *m/z* calcd for $C_{13}H_{19}N_4$ $[M + H]^+$: 231.1604, found 231.1601.

Synthesis of 1-propyl-1*H*-benzo[*d*]imidazol-2-amine (15)

To a Schlenk tube containing a magnetic stirring bar were added 2-aminobenzimidazole (**14**) (508 mg, 3.82 mmol), KOH (432.0 mg, 7.70 mmol) and acetone (6.6 mL). After addition of propyl bromide (0.51 mL, 5.63 mmol), the resulting mixture was stirred at room temperature for 3 h. Then, the mixture was poured into sat. NH_4Cl aq. and extracted with chloroform. The organic layer was dried over Na_2SO_4. After filtration, solvent was removed by evaporation. The resulting residue was purified by PTLC (chloroform/MeOH = 3:1 containing 1% Et_3N) to afford **15** as a white solid (323 mg, 48% yield).

1-Propyl-1*H*-benzo[*d*]imidazol-2-amine (15)

H NMR (600 MHz, CD_3OD) δ 7.22 (d, *J* = 7.6 Hz, 1H), 7.15 (d, *J* = 6.9 Hz, 1H), 7.02 (td, *J* = 7.6, 1.4 Hz, 1H), 6.99 (td, *J* = 7.6, 1.4 Hz, 1H), 3.97 (t, *J* = 7.2 Hz, 2H), 1.79 (sext, *J* = 7.4 Hz, 2H), 0.96 (t, *J* = 7.4 Hz, 3H); C NMR (150 MHz, CD_3OD) δ 156.1, 142.5, 135.3, 122.2, 120.5, 115.8, 109.0, 44.6, 23.1, 11.4.

HRMS (ESI) *m/z* calcd for $C_{10}H_{14}N_3$ $[M + H]^+$: 176.1182, found 176.1180.

Synthesis of *N*-(1-propyl-1*H*-benzo[*d*]imidazol-2-yl)furan-2-carboxamide (16)

To a 30-mL two-neck round bottom flask containing a magnetic stirring bar was added furan-2-carboxylic acid (**S4**) (96.0 mg, 0.86 mmol) and $SOCl_2$ (2.0 mL, 28.5 mmol). The resulting mixture was stirred at 80°C for 1 h. Then, the solvent was removed by evaporation. To the resulting residue were added 1-propyl-1*H*-benzo[d]imidazol-2-amine (**15**) (101.0 mg, 0.57 mmol), Et_3N (0.60 mL, 4.0 mmol) and CH_2Cl_2 (2.0 mL). The resulting mixture was stirred at 80°C for 3 h. Then the solvent was removed by evaporation. The resulting residue was purified by flash column chromatography using Isolera (chloroform/MeOH = 95:5 to 85:15) and PTLC (chloroform/MeOH = 50:1) to afford **16** as an orange solid (19.0 mg, 12% yield).

N-(1-propyl-1*H*-benzo[*d*]imidazol-2-yl)furan-2-carboxamide (16)

H NMR (600 MHz, CD_3OD) δ 7.68 (s, 1H), 7.49 (d, *J* = 7.6 Hz, 1H), 7.45 (d, *J* = 7.6 Hz, 1H), 7.31–7.25 (m, 2H), 7.22 (d, *J* = 2.7 Hz, 1H), 6.57–6.57 (m, 1H), 4.26 (t, *J* = 7.2 Hz, 2H), 1.90 (sext, *J* = 7.3 Hz, 2H), 0.99 (t, *J* = 7.6 Hz, 3H); C NMR (150 MHz, CD_3OD) δ 168.5, 153.7, 153.5, 146.3, 131.0, 130.4, 124.3, 124.2, 116.1, 112.9, 112.7, 110.8, 44.7, 22.9, 11.5.

HRMS (ESI) *m/z* calcd for $C_{15}H_{15}N_3O_2Na$ $[M+Na]^+$: 292.1056, found 292.1053.

Binding assays

The interactions between AI-generated compounds and proteins were verified via *in vitro* binding assays. To investigate binding specificity, we performed binding assays for 39 human proteins, including ADORA2A and PDE4D. The 39 human proteins used in binding studies were molecular targets (24 GPCRs, 3 transporters, 3 ion channels, 1 kinase, 2 nuclear receptors and 6 other non-kinase enzymes) expressed in the central nervous, cardiovascular, respiratory, digestive, and excretory systems; namely, ADORA2A, ADRA1A, ADRA2A, ADRB1, ADRB2, CNR1, CNR2, CCKAR, DRD1, DRD2, EDNRA, HRH1, HRH2, CHRM1, CHRM2, CHRM3, OPRD1, OPRK1, OPRM1, HTR1A, HTR1B, HTR2A, HTR2B, AVPR1A, SLC6A3, SLC6A2, SLC6A4, CHRNA4, KCNH2, HTR3A, LCK, AR, NR3C1, PTGS1, PTGS2, ACHE, MAOA, PDE3A and PDE4D.

The binding assay was performed using radioligand binding, spectrofluorimetry, enzymatic activity, or spectrophotometry (see Supplementary Binding Assay section). After measuring the degree to which the added compound sample inhibited the specific binding of the original ligand to the protein, the inhibition rate (%Inh) was calculated. The GPCR binding assay for ADORA2A was performed as follows.[28] ADORA2A expressed in human HEK-293 cells were used in modified Tris-HCl buffer pH 7.4. A15-μg aliquot was incubated with 50-nM [H] CGS-21680 for 90 min at 25°C. Nonspecific binding was estimated in the presence of 50-μM NECA (5′-N-ethylcarboxamidoadenosine). Receptors were filtered and washed before counting to determine the specifically bound [H] CGS-21680. Compounds were screened at 10 μM. The enzymatic activity for PDE4D was as follows.[29,30] Human-recombinant PDE4D expressed in Sf9 insect cells was used. Test compound and/or vehicle was preincubated with 5-ng/mL enzyme in Tris-HCl buffer pH 7.2 for 15 min at 25°C. The reaction was initiated by adding 100-nM fluorescein-labeled cAMP for another 15 min of incubation and terminated by adding the immobilized metal affinity for phosphochemical (IMAP) binding solution. IMAP forms a complex with phosphate groups on nucleotide monophosphate generated from cyclic nucleotides via PDE activity. The amount of complex formed was determined via spectrofluorimetry at 470/525 nm. Compounds were screened at 10 μM. The detailed experimental methods for the remaining 37 proteins are shown in the Supplementary Binding Assay section. All target proteins in the panel target list are given in Table S1.

Binding assays for GPCRs, ionic channels, and enzymes, transporters

Assay no. 1: Human adenosine A2a receptor (ADORA2A)

Human recombinant adenosine A2a receptors expressed in human HEK-293 cells were used in modified Tris-HCl buffer pH 7.4.[28] A15 μg aliquot was incubated with 50 nM [3H] CGS-21680 for 90 min at 25°C. Non-specific binding was estimated in the presence of 50 μM NECA. Receptors were filtered and washed, before counting to determine the specifically bound [3H] CGS- 21680. Compounds were screened at 10 μM.

Assay no. 2: Human adrenoceptor alpha 1A (ADRA1A)

Human recombinant adrenoceptors alpha 1A expressed in human Chem-1 cells was used in modified HEPES buffer pH 7.4.[31,32] A 2 μg aliquot was incubated with 0.6 nM [3H] Prazosin for 60 min at 25°C. Non-specific binding was estimated in the presence of 10 μM phentolamine. Receptors were filtered and washed, before counting to determine the specifically bound [3H] Prazosin. Compounds were screened at 10 μM.

Assay no. 3: Human adrenoceptor alpha 2A (ADRA2A)

Human recombinant adrenoceptor alpha 2A expressed in CHO-K1 cells was used in modified Tris-HCl buffer pH 7.4.[33,34] A 2 μg aliquot was incubated with 1.5 nM [3H] Rauwolscine for 60 min at 25°C. Non-specific binding was estimated in the presence of 10 μM WB-4101. Receptors were filtered and washed, before counting to determine the specifically bound [3H] Rauwolscine. Compounds were screened at 10 μM.

Assay no. 4: Human adrenoceptor beta 1 (ADRB1)

Human recombinant adrenoceptor beta 1 expressed in CHO-K1 cells was used in modified Tris-HCl buffer pH 7.4. A 25 μg aliquot was incubated with 0.03 nM [125I] Cyanopindolol for 120 min at 25°C.[35] Non-specific binding as estimated in the presence of 100 μM S (−)-Propranolol. Receptors were filtered and washed, before counting to determine the specifically bound [125I] Cyanopindolol. Compounds were screened at 10 μM.

Assay no. 5: Human adrenoceptor beta 2 (ADRB2)

Human recombinant adrenoceptor beta 2 expressed in CHO cells was used in modified Tris-HCl buffer pH 7.4.[36] A 50 μg aliquot was incubated with 0.2 nM [3H] CGP-12177 for 60 min at 25°C. Non-specific binding was estimated in the presence of 10 μM ICI-118551. Receptors were filtered and washed, before counting to determine the specifically bound [3H] CGP-12177. Compounds were screened at 10 μM.

Assay no. 6: Human cannabinoid receptor 1 (CNR1)

Human recombinant cannabinoid receptor 1 expressed in rat hematopoietic Chem-1 cells was used in modified HEPES buffer pH 7.4.[37,38] A 5 μg aliquot of membrane was incubated with 2 nM [3H]SR141716A for 60 min at 37°C. Non-specific binding was estimated in the presence of 10 μM CP 55,940. Membranes were filtered and washed, before counting to determine the specifically bound [3H] SR141716A. Compounds were screened at 10 μM.

Assay no. 7: Human cannabinoid receptor 2 (CNR2)

Human recombinant cannabinoid receptor 2 (CNR2) expressed in CHO-K1 cells was used in modified HEPES buffer pH 7.0.[39] A 30 μg aliquot was incubated with 2.4 nM [3H] WIN-55,212-2 for 90 min at 37°C. Non-specific binding was estimated in the presence of 10 μM R (+)-WIN-55,212-2. Membranes were filtered and washed, before counting to determine the specifically bound [3H] WIN-55,212-2. Compounds were screened at 10 μM.

Assay no. 8: Human cholecystokinin A receptor (CCKAR)

1321-N1 (human astrocytoma) stably transfected with a plasmid encoding the human recombinant cholecystokinin A receptors was used to prepare membranes in modified HEPES buffer pH 7.4.[40,41] A 5 μg aliquot of membrane was incubated with 0.11 nM [125I] CCK-8 for 180 min at 25°C. Non-specific binding was estimated in the presence of 1 μM L364,718 (devazepide). Membranes were filtered and washed, before counting to determine the specifically bound [125I] CCK-8. Compounds were screened at 10 μM.

Assay no. 9: Human dopamine receptor D1 (DRD1)

Human recombinant dopamine receptor D1 expressed in CHO cells was used in modified Tris-HCl buffer pH 7.4.[42,43] A 20 μg aliquot was incubated with 1.4 nM [3H] SCH-23390 for 120 min at 37°C. Non-specific binding was estimated in the presence of 10 μM (+)-butaclamol. Receptors were filtered and washed, before counting to determine the specifically bound [3H] SCH-23390. Compounds were screened at 10 μM.

Assay no. 10: Human dopamine receptor D2 (DRD2)

Human recombinant dopamine receptor D2 expressed in CHO cells was used in modified Tris-HCl buffer pH 7.4.[44,45] A 15 μg aliquot was incubated with 0.16 nM [3H] Spiperone for 120 min at 25°C. Non-specific binding was estimated in the presence of 10 μM haloperidol. Receptors were filtered and washed, before counting to determine the specifically bound [3H] Spiperone. Compounds were screened at 10 μM.

Assay no. 11: Human endothelin receptor type A (EDNRA)

Human recombinant endothelin receptor type A expressed in CHO-S1 cells are used in modified Tris-HCl buffer pH 7.4.[46] A 0.5 μg aliquot was incubated with 0.03 nM [125I] Endothelin-1 for 120 min at 37°C. Non-specific binding was estimated in the presence of 0.1 μM endothelin-1. Receptors were filtered and washed, before counting to determine the specifically bound [125I] Endothelin-1. Compounds were screened at 10 μM.

Assay no. 12: Human histamine receptor H1(HRH1)

Human recombinant histamine receptor H1 expressed in CHO-K1 cells was used in modified Tris-HCl buffer pH 7.4.[47] A 10 μg aliquot was incubated with 1.2 nM [3H] Pyrilamine for 180 min at 25°C. Non-specific binding was estimated in the presence of 1 μM pyrilamine. Receptors were filtered and washed, before counting to determine the specifically bound [3H] Pyrilamine. Compounds were screened at 10 μM.

Assay no. 13: Human histamine receptor H2 (HRH2)

Human recombinant histamine receptor H2 expressed in CHO-K1 cells as prepared in K-Na phosphate buffer pH 7.4.[48] A 2 μg aliquot was incubated with 0.1 nM [125I] Aminopotentidine for 120 min at 25°C. Non-specific binding was estimated in the presence of 3 μM Tiotidine. Receptors were filtered and washed, before counting to determine the specifically bound [125I] Aminopotentidine. Compounds were screened at 10 μM.

Assay no. 14: Human cholinergic receptor muscarinic 1 (CHRM1)

Human recombinant cholinergic receptor muscarinic 1 expressed in CHO-K1 cells was used in modified Tris-HCl buffer pH 7.4.[49,50] A 16 μg aliquot was incubated with 0.8 nM [3H] N-Methylscopolamine for 120 min at 25°C. Non-specific binding was estimated in the presence of 1 μM Atropine. Receptors were filtered and washed, before counting to determine the specifically bound [3H] N-Methylscopolamine. Compounds were screened at 10 μM.

Assay no. 15: Human cholinergic receptor muscarinic 2 (CHRM2)

Human recombinant cholinergic receptor muscarinic 2 expressed in CHO-K1 cells was used in modified Tris-HCl buffer pH 7.4.[49,50] A 8 μg aliquot was incubated with 0.8 nM [3H] N-Methylscopolamine for 120 min at 25°C. Non-specific binding was estimated in the presence of 1 μM atropine. Receptors were filtered and washed, before counting to determine the specifically bound [3H] N- Methylscopolamine. Compounds were screened at 10 μM.

Assay No.16: Human cholinergic receptor muscarinic 3 (CHRM3)

Human recombinant cholinergic receptor muscarinic 3 expressed in CHO-K1 cells was used in modified Tris-HCl buffer pH 7.4.[49,50] A 12 μg aliquot was incubated with 0.8 nM [3H] N-Methylscopolamine for 120 min at 25°C. Non-specific binding was estimated in the

presence of 1 μM atropine. Receptors were filtered and washed, before counting to determine the specifically bound [3H] N- Methylscopolamine. Compounds were screened at 10 μM.

Assay no. 17: Human opioid receptor delta 1 (OPRD1)

Human recombinant opioid receptor delta 1 expressed in HEK293 cells was used in modified Tris-HCl buffer pH 7.4.[51,52] A 9 μg aliquot was incubated with 1.3 nM [3H] Naltrindole for 60 min at 25°C. Non-specific binding was estimated in the presence of 1 μM Naltrindole. Receptors were filtered and washed, before counting to determine the specifically bound [3H] Naltrindole. Compounds were screened at 10 μM.

Assay no. 18: Human opioid receptor kappa 1 (OPRK1)

Human recombinant opioid receptor kappa 1 expressed in human HEK-293 cells was used in modified Tris-HCl buffer pH 7.4.[53,54] A 30 μg aliquot was incubated with 0.6 nM [3H] Diprenorphine for 60 min at 25°C. Non-specific binding was estimated in the presence of 10 μM naloxone. Receptors were filtered and washed, before counting to determine the specifically bound [3H] Diprenorphine. Compounds were screened at 10 μM.

Assay no. 19: Human opioid receptor mu 1 (OPRM1)

Human recombinant opioid receptor mu 1 expressed in CHO-K1 cells was used in modified Tris-HCl buffer pH 7.4.[55] A 11 μg aliquot was incubated with 0.6 nM [3H] Diprenorphine for 60 min at 25°C. Non-specific binding was estimated in the presence of 10 μM naloxone. Receptors were filtered and washed, before counting to determine the specifically bound [3H] Diprenorphine. Compounds were screened at 10 μM.

Assay no. 20: Human 5-hydroxytryptamine receptor 1A (HTR1A)

Human recombinant 5-hydroxytryptamine receptor 1A expressed in CHO-K1 cells was used in modified Tris-HCl buffer pH 7.4.[56,57] An 8 μg aliquot was incubated with 1.5 nM [3H]8-OH-DPAT for 60 min at 25°C. Non-specific binding as estimated in the presence of 10 μM metergoline. Receptors were filtered and washed, before counting to determine the specifically bound [3H]8-OH-DPAT. Compounds were screened at 10 μM.

Assay no. 21: Human 5-hydroxytryptamine receptor 1B (HTR1B)

Human recombinant 5-hydroxytryptamine receptor 1B expressed in Chem-1 cells was used in modified Tris-HCl buffer pH 7.4.[58,59] A 2 μg aliquot of membrane was incubated with 1 nM [3H]GR125743 for 90 min at 37°C. Non-specific binding was estimated in the presence of 10 μM 5-HT. Membranes were filtered and washed, before counting to determine the specifically bound [3H] GR125743. Compounds were screened at 10 μM.

Assay no. 22: Human 5-hydroxytryptamine receptor 2A (HTR2A)

Human recombinant 5-hydroxytryptamine receptor 2A expressed in CHO-K1 cells was used in modified Tris-HCl buffer pH 7.4.[60,61] A 30 μg aliquot was incubated with 0.5 nM [3H]Ketanserin for 60 min at 25°C. Non-specific binding was estimated in the presence of 1 μM Mianserin. Receptors were filtered and washed, before counting to determine the specifically bound [3H] Ketanserin. Compounds were screened at 10 μM.

Assay no. 23: Human 5-hydroxytryptamine receptor 2B (HTR2B)

Human recombinant 5-hydroxytryptamine receptor 2B expressed in CHO-K1 cells was used to prepare membranes in modified Tris-HCl buffer pH 7.4.[60] A 30 μg aliquot of membrane protein was incubated with 1.2 nM [3H] LSD for 60 min at 37°C. Non-specific binding was estimated in the presence of 10 μM serotonin. Membranes were filtered and washed, before counting to determine the specifically bound [3H] LSD. Compounds were screened at 10 μM.

Assay no. 24: Human arginine vasopressin receptor 1 (AVPR1A)

Human recombinant arginine vasopressin receptor 1 expressed in HEK-293 cells was used in modified Tris-HCl buffer pH 7.4.[62] A 0.26 μg aliquot was incubated with 0.03 nM [125I] Phenylacetyl-D-Tyr(Me)-Phe-Gln-Asn-Arg- Pro-Arg-Tyr-NH2 for 120 min at 25°C. Non-specific binding was estimated in the presence of 1 μM [Arg8] Vasopressin. Receptors were filtered and washed, before counting to determine the specifically bound [125I] Phenylacetyl-D-Tyr (Me)-Phe-Gln-Asn-Arg-Pro-Arg-Tyr-NH2. Compounds were screened at 10 μM.

Assay no. 25: Human solute carrier family 6 member 3 (SLC6A3)

Human recombinant SLC6A3 expressed in CHO-S cells was used in modified Tris-HCl buffer pH 7.4.[63,64] A 0.4 μg aliquot was incubated with 0.15 nM [125I] RTI-55 for 3 h at 4°C. Non-specific binding was estimated in the presence of 10 μM nomifensine. Transporters were filtered and washed, before counting to determine specifically bound [125I] RTI-55. Compounds were screened at 10 μM.

Assay no. 26: Human solute carrier family 6 member 2 (SLC6A2)

Human recombinant SLC6A2 expressed in dog kidney MDCK cells was used in modified Tris-HCl buffer pH 7.4.[65] A 40 μg aliquot was incubated with 0.2 nM [125I] RTI-55 for 3 h at 4°C. Non-specific binding was estimated in the presence of 10 μM desipramine. Transporters were filtered and washed, before counting to determine specifically bound [125I] RTI-55. Compounds were screened at 10 μM.

Assay no. 27: Human solute carrier family 6 member 4 (SLC6A4)

Human recombinant SLC6A4 expressed in human HEK-293 cell was used in modified Tris-HCl buffer pH 7.4.[66,67] A 9 μg aliquot was incubated with 0.4 nM [3H] Paroxetine for 60 min at 25°C. Non-specific binding was estimated in the presence of 10 μM imipramine. Transporters were filtered and washed, before counting to determine specifically bound [3H] Paroxetine. Compounds were screened at 10 μM.

Assay no. 28: Human cholinergic receptor nicotinic alpha 4 subunit (CHRNA4)

Human recombinant cholinergic receptor nicotinic alpha 4 subunit expressed in SHSY5Y cells was used in modified Tris-HCl buffer pH 7.4.[68] A 30 μg aliquot incubated with 0.6 nM [H] Cytisine for 120 min at 4°C. Non-specific binding was estimated in the presence of 10 μM nicotine. Membranes were filtered and washed, before counting to determine the specifically bound [3H] Cytisine. Compounds were screened at 10 μM.

Assay no. 29: Human potassium voltage-gated channel subfamily H member 2 (KCNH2)

Human recombinant potassium voltage-gated channel subfamily H member 2 expressed in human HEK- 293 cells was used in modified Tris-HCl buffer pH 7.4.[69–71] A 7.5 μg aliquot is incubated with 3 nM [3H]Dofetilide for 60 min at 25°C. Non-specific binding was estimated in the presence of 10 μM Dofetilide. Channel proteins were filtered and washed, before counting to determine the specifically bound [3H] Dofetilide. Compounds were screened at 10 μM.

Assay no. 30: Human 5-hydroxytryptamine receptor 3A (HTR3A)

Human recombinant 5-hydroxytryptamine receptor 3A expressed in human HEK-293 cells was used in modified Tris-HCl buffer pH 7.4.[72,73] A 3.2 μg aliquot was incubated with 0.69 nM [3H]GR-65630 for 60 min at 25°C. Non-specific binding was estimated in the presence of 10 μM MDL-72222. Receptors were filtered and washed, before counting to determine the specifically bound [3H]GR-65630. Compounds were screened at 10 μM.

Assay no. 31: Human LCK proto-oncogene, Src family tyrosine kinase (LCK)

Human recombinant protein kinase LCK expressed in insect cells was used. Test compound and/or vehicle as preincubated with 0.4 μg/mL enzyme in modified HEPES buffer pH 7.4 for 15 min at 37°C.[74] The reaction was initiated by addition of 0.2 mg/mL poly (Glu:Tyr), 10 μM ATP, and 0.25 μCi [γ32P]ATP for another 30 min incubation and terminated by further addition of 3% H3PO4. An aliquot was removed and counted to determine the amount of [32P] Poly (Glu:Tyr) formed. Compounds were screened at 10 μM.

Assay no. 32: Human androgen receptor (AR)

Human androgen receptors obtained from human LNCaP cells were used in modified HEPES buffer pH 7.4.[75] A 70 μg aliquot was incubated with 0.5 nM [3H] Methyltrienolone for 20 h at 4°C. Non-specific binding was estimated in the presence of 1 μM testosterone. Receptors were filtered and washed, before counting to determine the specifically bound [3H] Methyltrienolone. Compounds were screened at 10 μM.

Assay no. 33: Human nuclear receptor subfamily 3 group C member 1 (NR3C1)

Human recombinant glucocorticoid receptor expressed in insect cells was used in modified potassium phosphate buffer pH 7.4.[76,77] A 10 μg aliquot was incubated with 5 nM [3H] Dexamethasone, 2 μg anti-GST antibody, and 0.2 mg SPA beads for 24 h at 4°C. Non-specific binding was estimated in the presence of 10 μM Dexamethasone. Receptors were counted to determine the specifically bound [3H] Dexamethasone. Compounds were screened at 10 μM.

Assay no. 34: Human prostaglandin-endoperoxide synthase 1 (PTGS1)

Human recombinant prostaglandin-endoperoxide synthase 1 expressed in baculovirus infected Sf9 cells were used.[78,79] Test compound and/or vehicle is incubated with 0.44 μg/mL enzyme in modified Tris-HCl buffer pH 8.0 for 15 min at 25°C. The reaction was initiated by addition of 3 μM arachidonic acid and 100 μM Ampliflu Red for another 3 min incubation. Determination of the amount of Resorufin formed through spectrofluorimetry at 535/590 nm. Compounds were screened at 10 μM.

Assay no. 35: Human prostaglandin-endoperoxide synthase 2 (PTGS2)

Human recombinant prostaglandin-endoperoxide synthase 2 expressed in insect Sf21 cells was used.[80,81] Test compound and/or vehicle was preincubated with 34 U/mL enzyme in modified Tris-HCl buffer pH 8.0 for 15 min at 25°C. The reaction was initiated by addition of 3 μM arachidonic acid and 100 μM Ampliflu Red for another 3 min incubation. Determination of the amount of Resorufin formed formed through spectrofluorimetry at 535/590 nm. Compounds were screened at 10 μM.

Assay no. 36: Human acetylcholinesterase (ACHE)

Human recombinant acetylcholinesterase expressed in HEK-293 cells (Sigma, C-1682) was used.[82,83] Test compound and/or vehicle was preincubated with 4.1 ng/mL of enzyme for 15 min at 25°C in phosphate buffer pH 7.4. The reaction was initiated by addition of 0.7 mM acetylthiocholine iodide and 0.5 mM 5,5-dithiobis-2-nitrobenzoic acid for another 20 min incubation. The thiocholine generated reacts continuously with dithiobisnitrobenzoic acid to produce 5-thio-2-nitro-benzoic acid; its spectrophotometric absorbance was read at 405 nm. Compounds were screened at 10 μM.

Assay no. 37: Human monoamine oxidase A (MAOA)

Human recombinant MAOA expressed in insect cells was used.[84,85] Test compound and/or vehicle was preincubated with 4.2 μg/mL enzyme in phosphate buffer pH 7.4 for 15 min at 37°C. The reaction as initiated by addition of 50 μM Kynuramine for another 60 min incubation and terminated by further addition of 1.2 N NaOH. Determination of the amount of 4-hydroxyquinoline formed through spectrofluorimetry at 325/465 nm. Compounds were screened at 10 μM.

Assay no. 38: Human phosphodiesterase 3A (PDE3A)

Human recombinant PDE3A expressed in insect Sf9 cells was used.[86,87] Test compound and/or vehicle was preincubated with 20 ng/mL enzyme in Tris- HCl buffer pH 7.2 for 15 min at 25°C. The reaction was initiated by addition of 100 nM fluorescein labeled cAMP for another 30 min incubation and terminated by addition of IMAP binding solution. IMAP complexes with phosphate groups on nucleotide monophosphate generated from cyclic nucleotides through PDE activity. Determination of the amount of complex formed through spectrofluorimetry at 470/525 nm. Compounds were screened at 10 μM.

Assay no. 39: Human phosphodiesterase 4D (PDE4D)

Human recombinant PDE4D expressed in insect Sf9 cells was used.[29,30] Test compound and/or vehicle was preincubated with 5 ng/mL enzyme in Tris-HCl buffer pH 7.2 for 15 min at 25°C. The reaction was initiated by addition of 100 nM fluorescein labeled cAMP for another 15 min incubation and terminated by addition of IMAP binding solution. IMAP makes a complex with phosphate groups on nucleotide monophosphate generated from cyclic nucleotides through PDE activity. Determination of the amount of complex formed through spectrofluorimetry at 470/525 nm. Compounds were screened at 10 μM.

METHOD DETAILS

Therapeutic targets

We focused on two therapeutic targets, ADORA2A and PDE4D, associated with bronchial asthma. It is an allergic reaction to inhaled allergens that causes airflow obstruction, bronchial hyper-responsiveness, and airway inflammation.[88] Bronchial asthma is currently alleviated using two types of bronchodilators: β_2 stimulants and anticholinergics. ADRB2 increases the cyclic adenosine monophosphate (cAMP) levels and relaxes the bronchial smooth muscle. Anticholinergic drugs increase the intracellular Ca^+ concentration and inhibit bronchial smooth muscle contraction by inhibiting binding to CHRM3. However, long-term use of β_2 stimulants increases the risk of arrhythmia and cardiac arrest, whereas anticholinergic drugs have a slow activity onset and weak action. Therefore, targeting ADORA2A and PDE4D associated with bronchodilation might be more effective than targeting ADRB2 or CHRM3 alone. Compounds that interact with both these targets can potentially delay the progression and treat the symptoms of the disease. As PDE4D inhibits cAMP degradation, it allows intracellular cAMP accumulation and relaxes the bronchial smooth muscle.[89] PDE4D also inhibits bronchial smooth muscle contraction by increasing the intracellular cAMP concentration by targeting ADORA2A.[90]

Algorithm of the dual-target fragment-based structure generator (DualFASMIFRA)

DualFASMIFRA is a dual-target extension of the Fast Assembly of simplified molecular-input line-entry system (SMILES) Fragments (FASMIFRA).[91] The compound structures and their bioactivities for ADORA2A and PDE4D were obtained from the ChEMBL database, where CHEMBL251 corresponded to ADORA2A and CHEMBL288 corresponded to PDE4D. The chemical structures were encoded using a molecular encoding procedure called *unfolded-counted atom pairs*.[92] The algorithm in DualFASMIFRA is a published method.[91]

After molecular standardization, any duplicate compounds were grouped and their median pIC_{50} value was calculated. Activity values were assigned to 4780 compounds for ADORA2A and 981 compounds for PDE4D. Optimization was performed using the L2-regularized support vector regressor from LIBLINEAR,[93] which can be easily trained and provides fast predictions. The *C* parameter was optimized over the exponential range [0.001, 0.002, 0.005, …, 50], whereas epsilon was fixed to zero (its default value). The dependent variable in the model was the molecular weight invariant (MWI) pIC_{50} (similar to the ligand efficiency[94] but divided by the molecular weight instead of the number of heavy atoms). The MWI pIC_{50} is sometimes easier to model than pIC_{50} because the MWI transform does not truncate the modeled variable distribution. To simplify the interpretation of the predicted affinities, some random forest regressors were trained to directly predict the pIC_{50} values (Figures 2A and 2B); 100 trees, maximizing only the mtry parameter). We evaluated the model using R, RMSE, and MAE after 5-fold cross validation. The QSAR model for ADORA2A had an R of 0.66, RMSE of 0.68, and MAE of 0.52. The QSAR model for PDE4D had an R of 0.73, RMSE of 0.75, and MAE of 0.55. Both R scores were >0.65, indicating acceptable quality. The errors were less than 10% in terms of RMSE and MAE. The maximum pIC_{50} value for the test data was 9.92 for both ADORA2A and PDE4D.

In addition, fragments should lie within the applicability domain. In other words, compounds containing the fragments possess known activity against the protein target(s) of interest. Here, the compounds were fragmented and assembled at a high frequency using FASMIFRA in each GA iteration.

The targeted chemical space was constrained using property filters. Compounds passing the lead-like filter[95] should allow space for future medicinal and chemical optimization if required (MolW $\leq$460 Da, $-4.0 \leq$ cLogP $\leq$4.2, rotB $\leq$10, nRing $\leq$4, HBD $\leq$5, HBA $\leq$9). To retain only the compounds predicted as orally available, we trained a classifier on the dataset published in a previous study.[96]

Compounds containing known reactive functional groups were filtered out. To ensure an easy synthesis, we admitted only compounds with an Ertl and Schffenhauer's synthetic accessibility score[97] of 2.75 or lower. Figure 2 (c) shows ratios of elite compounds with high scores in the QSAR models, diverse (randomly selected) compounds, and mutant compounds (positional scan analogs of the merged set of elite and random molecular population) after one iteration of the genetic algorithm in DualFASMIFRA. Figure 2D shows an example of the optimization process, and Figure S1 shows some of the generated compounds.

Algorithm of the dual-target deep learning structure generator (DualTransORGAN)

DualTransORGAN is a GAN-based deep generative model, which is a dual-target extension of TransORGAN.[98] It utilizes a transformer architecture as its generator and employs a stochastic policy gradient reinforcement learning algorithm[99] to design molecules with the bioactivity (pIC_{50}) as the target chemical property.

Specifically, the generator G_θ employs a transformer encoder–decoder architecture for end-to-end learning of SMILES strings. The transformer encoder uses the self-attention mechanism to extract features from a variant SMILES string, whereas the transformer

decoder decodes these features back into the canonical string. Formally, the generator calculates the attention score $\mathcal{A}$ within an SMILES string as follows:

$$\mathcal{A}(Q,K,V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad \text{(Equation 1)}$$

where Q, K, and V represent the queries, keys, and values with corresponding dimensions of d_k for queries and keys, and d_v for values. Then, the generator jointly attends to the features of the different SMILES strings with h projections of multi-head attention (MHA), as calculated by:

$$\text{MHA}(Q,K,V) = \text{Concat}(head_1,\ldots,head_h)W, \quad \text{(Equation 2)}$$

$$head_i = \mathcal{A}\left(QW_i^Q, KW_i^K, VW_i^V\right), \quad \text{(Equation 3)}$$

where W, W_i^Q, W_i^K, and W_i^V denote weight matrices.

The discriminator D_ϕ is implemented as a convolution neural network (CNN) tasked with distinguishing between the generated SMILES-like strings and raw SMILES strings. Formally, let $\boldsymbol{x}_{1:T_x} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{T_x}]$ represent an SMILES string with a maximum length of T_x from the real dataset, and $\boldsymbol{y}_{1:T_y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{T_y}]$ represent a generated SMILES-like string with a maximum length of T_y from the generated dataset, respectively. Then, the source and target matrices can be denoted by:

$$\boldsymbol{X}_{1:T_x} = \boldsymbol{x}_1 \oplus \ldots \oplus \boldsymbol{x}_{T_x}, \boldsymbol{Y}_{1:T_y} = \boldsymbol{y}_1 \oplus \ldots \oplus \boldsymbol{y}_{T_y}. \quad \text{(Equation 4)}$$

Here, $\oplus$ denotes the concatenation operation. Given a window size of t tokens, the feature map of the convolution operation c_{ij} can be calculated as follows:

$$c_{ij} = \text{ReLu}(\omega_j \odot X_{i:i+t-1} + b), \quad \text{(Equation 5)}$$

where ω_j signifies the j-th kernel, $\odot$ denotes element-wise summation of the products, and b corresponds to the bias of the convolutional layer. Then, the max-pooling operator is used to select the maximum value $\widehat{c_j}$ in each pooling region. Finally, the probability of the input string being recognized as true can be determined using a fully connected layer discriminator. To address the training instability and mode collapse challenges often faced in traditional GANs, the Wasserstein distance[100] is employed as the output of the discriminator. This implementation is as follows:

$$W_{D_\phi}(\mathbb{P}_r, \mathbb{P}_\theta) = \sup_{||f||_L \leq 1} \mathbb{E}_{\boldsymbol{x}_{1:T_x}\sim\mathbb{P}_r}[f(\boldsymbol{x}_{1:T_x})] - \mathbb{E}_{\boldsymbol{y}_{1:T_y}\sim\mathbb{P}_\theta}\left[f\left(\boldsymbol{y}_{1:T_y}\right)\right], \quad \text{(Equation 6)}$$

where $\mathbb{P}_r$ represents the real SMILES distribution and $\mathbb{P}_\theta$ is the distribution output of the generator. $\mathbb{E}_{\mathbb{P}_r}$ and $\mathbb{E}_{\mathbb{P}_\theta}$ represent the expected values of distributions $\mathbb{P}_r$ and $\mathbb{P}_\theta$ respectively, for the function f. The upper limit (*sup*) of the difference between each expected value was then obtained.

Due to the discrete nature of SMILES strings, distinguishing samples directly poses a challenge for GANs. To address this non-differentiability issue, a solution is provided through the application of the policy gradient reinforcement learning algorithm. The objective function is shown as follows:

$$J(\boldsymbol{\theta}) = \sum_{\boldsymbol{Y}_{1:T_y}} G_\theta(\boldsymbol{Y}_{1:T_y}|\boldsymbol{X}_{1:T_x}) R\left((\boldsymbol{X}_{1:T_x}, \boldsymbol{Y}_{1:T_y-1}), \boldsymbol{Y}_{T_y}\right), \quad \text{(Equation 7)}$$

where $R(s_0, a)$ represents the expected reward for taking action a starting from the state s_0. The expected reward score is a trade-off among the Wasserstein distance of the discriminator ($D_\phi(\boldsymbol{Y}_{1:T_y})$), the target chemical property of pIC_{50} ($O(\boldsymbol{Y}_{1:T_y})$) and the penalty score for mode collapse ($P(\boldsymbol{Y}_{1:T_y})$). The calculation is as follows:

$$R\left((\boldsymbol{X}_{1:T_x}, \boldsymbol{Y}_{1:T_y-1}), \boldsymbol{Y}_{T_y}\right) = \lambda D_\phi(\boldsymbol{Y}_{1:T_y}) + (1-\lambda)\, O(\boldsymbol{Y}_{1:T_y}) P(\boldsymbol{Y}_{1:T_y}) \quad \text{(Equation 8)}$$

where λ controls the trade-off among the three rewards. $O(\boldsymbol{Y}_{1:T_y})$ denotes the pIC_{50} score of $\boldsymbol{Y}_{1:T_y}$, and $P(\boldsymbol{Y}_{1:T_y})$ represents the penalty for generating duplicate SMILES strings of $\boldsymbol{Y}_{1:T_y}$. This penalty is calculated as the ratio of the number of SMILES strings to the product of the total number of SMILES strings and the number of duplicate strings. All the rewards range from 0 to 1. To improve the stability of the GAN training process, we perform Monte Carlo (MC) tree search N times on the policy G_θ as follows:

$$MC_{G_\theta}((\boldsymbol{X}_{1:T_x}, \boldsymbol{Y}_{1:t}), N) = \left\{\boldsymbol{Y}_{1:T_y}^1, \boldsymbol{Y}_{1:T_y}^2, \ldots, \boldsymbol{Y}_{1:T_y}^N\right\}. \quad \text{(Equation 9)}$$

Then, the mean reward of the MC tree search for N times $\overline{R}$ is

$$\overline{R}((\boldsymbol{X}_{1:T_x}, \boldsymbol{Y}_{1:t-1}), \boldsymbol{Y}_t) = \frac{1}{N}\sum_{n=1}^{N} R\left(\left(\boldsymbol{X}_{1:T_x}, \boldsymbol{Y}_{1:T_y-1}^n\right), \boldsymbol{Y}_{T_y}\right), \boldsymbol{Y}_{1:T_y}^n \in MC_{G_\theta}((\boldsymbol{X}_{1:T_x}, \boldsymbol{Y}_{1:t}), N) \text{ if } t < T_y, \quad \text{(Equation 10)}$$

$$\overline{R}((\boldsymbol{X}_{1:T_x}, \boldsymbol{Y}_{1:t-1}), \boldsymbol{Y}_t) = R\left(\left(\boldsymbol{X}_{1:T_x}, \boldsymbol{Y}^n_{1:T_y-1}\right), \boldsymbol{Y}_{T_y}\right), \text{if } t = T_y. \quad \text{(Equation 11)}$$

Finally, the policy gradient of the generator is computed as

$$\nabla J(\boldsymbol{\theta}) = \frac{1}{T_y}\sum_{t=1}^{T_y}\sum_{\boldsymbol{Y}_t}\overline{R}((\boldsymbol{X}_{1:T_x}, \boldsymbol{Y}_{1:t-1}), \boldsymbol{Y}_t)\nabla_{\boldsymbol{\theta}} \log G_{\boldsymbol{\theta}}(\boldsymbol{Y}_t|\boldsymbol{X}_{1:T_x}, \boldsymbol{Y}_{1:t-1}), \quad \text{(Equation 12)}$$

$$\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} + \nabla J(\boldsymbol{\theta}). \quad \text{(Equation 13)}$$

The DualTransORGAN was trained in Python 3.6.10 environment under Anaconda 4.10.3 (https://www.anaconda.com [accessed 2023-01-26]) using the deep learning library, PyTorch 1.8.1 (https://pytorch.org [accessed 2023-01-26]).[101] The token vector dimension (d_s) and maximum number of tokens (T) of the SMILES input to the generator were set to 128 and 77, respectively. To classify the SMILES tokens, the number of token types and dropout rates were set to 81 and 0.2, respectively. The transformer comprised four encoder and four decoder layers, and the number of heads in multihead attention was set to 4. First, the generator and discriminator were pretrained. The generator was pretrained on 500,000 SMILES strings to learn their semantic and syntactic features. These features underwent pretraining for 250 epochs using an ADAM optimizer[102] with an initial learning rate of 0.001. The discriminator was pretrained for 10 epochs using the previously pretrained generator and 500,000 generated SMILES strings. The initial learning rate was set to 0.001.

After pretraining, the generator was trained using the bioactivity prediction model. First, SMILES strings with structures were generated using the generators that previously relearned the structure characteristics. The activity of each generated structure was predicted using the established bioactivity prediction model, and the existence probability of the structure was determined using the discriminator. The predicted activity and existence probability of the structure were returned as rewards to the generator. The generator then relearned the characteristics and attempted to generate a structure that maximized its reward, i.e., a structure with a higher existence probability and bioactivity than the previous structure. The learning rate was set at 4×10^{-5}, and the maximum number of generated compounds was 5,000.

Then, the discriminator was trained using the generator that previously learned the characteristics based on the bioactivity values. Generators trained via GAN discriminator learning usually created chemical structures based on their bioactivity values. These structures were compared with those of the original data in the discriminator. Furthermore, whether each structure was real or derived was evaluated; structures with a higher probability of being treated were regarded as real structures. The learning rate and loss function were both set to 0.001, and discriminator pretraining was based on the Wasserstein distance. Chemical structures were predicted and generated using the generator.

DualTransORGAN was trained on 500,000 compounds with molecular weights of $\leq$500 g/mol and SMILES length of 10–60. These structural data were obtained from 2,105,464 compounds registered in the ChEMBL database (https://www.ebi.ac.uk/chembl/[accessed 2023-01-26]).[103] The bioactivity data of compound structures with activities against ADORA2A and PDE4D were obtained from ChEMBL. Among these data, 309 and 903 compound structures targeted ADORA2A and PDE4D, respectively. The bioactivity was defined as IC_{50}. The pIC_{50} was then obtained as the negative logarithm of the IC_{50} (in molar concentration).

To obtain the compound structures and their bioactivities on the target proteins, we converted the SMILES representation of the structure data of a compound to a Morgan Fingerprint and represented it as a 2048-dimensional vector. Then, we created a regression model for bioactivity prediction using RandomForestRegressor in scikit-learn (https://scikit-learn.org [accessed 2023-01-26]).[104] The *n_estimators* argument of the RandomForestRegressor was set to 1000. The dataset was split into a training dataset (80%) and a test dataset (20%). The designed regression model was evaluated using the coefficient of determination.

Among the generated compounds, we analyzed 50 compounds with the highest sums of their two bioactivity values, and confirmed that the generated structures did not have the same structures as the compounds in the training set. The results are discussed in the fourth paragraph of the "generation of dual-target compounds using the proposed structure generators" subsection of the results section. The top 50 compounds are listed in Figure S2.

QUANTIFICATION AND STATISTICAL ANALYSIS

These points are not applicable to this study. In this paper, there are no results of statistical tests.