

AgainstOTE: A generative pretrain-tune framework for de novo molecular generation against off-target effects

Yechao Xu,^{1,7} Wen Fang,^{2,7} Shuai Wang,^{3,4,7} Hai-Bin Luo,⁵ Qiang Xu,^{2,*} Fen-Er Chen,^{3,4,*} Jiao Qu,^{2,*} and Yang Sun^{2,6,*}

*Correspondence: yangsun@nju.edu.cn (Y.S.); qujiao@nju.edu.cn (J.Q.); rfchen@fudan.edu.cn (F.C.); qiangxu@nju.edu.cn (Q.X.)

Received: November 30, 2024; Accepted: September 17, 2025; Published Online: September 24, 2025; <https://doi.org/10.59717/j.xinn-med.2025.100158>

© 2025 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- Preclinical AI-based drug discovery plays an important role in the development of innovative drugs.
- Our pre-experiments reveal off-target effects in AI produced molecules, highly risking clinical trial failures.
- A generative pretrain-tune framework is proposed to generate potential candidates against off-target effects.
- AI simulation reveals superior chemical properties and inhibited off-target binding affinity.
- Biological assay is carried out after chemical synthesis of predicted receptor-selective molecules.

AgainstOTE: A generative pretrain-tune framework for de novo molecular generation against off-target effects

Yechao Xu,^{1,7} Wen Fang,^{2,7} Shuai Wang,^{3,4,7} Hai-Bin Luo,⁵ Qiang Xu,^{2,*} Fen-Er Chen,^{3,4,*} Jiao Qu,^{2,*} and Yang Sun^{2,6,*}

¹Department of Computer Science and Technology, Nanjing University, 163 Xianlin Avenue, Nanjing 210023, China

²State Key Laboratory of Pharmaceutical Biotechnology and Nanjing Drum Tower Hospital the Affiliated Hospital of Nanjing University Medical School, School of Life Sciences, Nanjing University, 163 Xianlin Avenue, Nanjing 210023, China

³Engineering Center of Catalysis and Synthesis for Chiral Molecules, Department of Chemistry, Fudan University, Shanghai 200433, China

⁴Shanghai Engineering Center of Industrial Asymmetric Catalysis for Chiral Drugs, Shanghai 200433, China

⁵Key Laboratory of Tropical Biological Resources of Ministry of Education and Hainan Engineering Research Center for Drug Screening and Evaluation, School of Pharmaceutical Sciences, Hainan University, Haikou 570228, China

⁶Jiangsu Key Laboratory of New Drug Research and Clinical Pharmacy, Xuzhou Medical University, Xuzhou 221004, China

⁷These authors contributed equally

*Correspondence: yangsun@nju.edu.cn (Y.S.); qujiao@nju.edu.cn (J.Q.); rfchen@fudan.edu.cn (F.C.); qiangxu@nju.edu.cn (Q.X.)

Received: November 30, 2024; Accepted: September 17, 2025; Published Online: September 24, 2025; <https://doi.org/10.59717/j.xinn-med.2025.100158>

© 2025 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Xu Y., Fang W., Wang S., et al. (2025). AgainstOTE: A generative pretrain-tune framework for de novo molecular generation against off-target effects. *The Innovation Medicine* 3:100158.

De novo molecular generation is committed by the mission to create novel molecules from scratch with desired properties. Learning from massive protein-molecule data, generative models can serve drug discovery, accelerating the development of innovative medical treatment and products curing stubborn diseases. Previous research celebrates phased success by optimizing the joint distribution of molecule-protein pairs. However, our preliminary experiments show that AI-generated molecules exhibit off-target effects, potentially leading to the failure of clinical trials. This indicates precisely binding to the target yet not to others far from settled. Chemical candidate space may be quite narrow, instead of vastly searching only conditioned on target proteins. In this study, we present a generative pretrain-tune framework for de novo molecular generation against off-target effects (AgainstOTE). For pretraining an initialized generative model which is called AgainstOTE-R1 in this context, it exploits the dual-receptor cooperative training mechanism, randomly selecting any other protein as off-target input, providing neural networks the opportunity observing all kinds of structures while focusing on the target. Then it employs structural displacement to create structural randomly shifted samples for simulating conformational changes brought by mobility of proteins. Consistent chemical embeddings are accessed by a regularizer comprising of E(3)-equivariant graph projector and molecular principal component analysis (MolPCA). With structural regularization implemented by a molecular contrastive learning architecture normalizing searching space by evenly distributing molecular features from both target and off-target proteins, AgainstOTE-R1 produces plausible compounds with discriminative structures even when exposed binding sites are similar. For enhancing the physical and biological plausibility, an advanced molecular generation model is tuned from the R1 model. In this stage, more district biochemically meaningful off-target receptors are selected by protein language models as training samples. Additionally, molecular dynamics simulation is exploited in structural displacement for physically modeling temporal trajectories of proteins instead of random sampling. We conduct extensive AI simulation and biochemical experiments to evaluate performance of this framework while designing target-to-sidelobe ratios (TSR) especially to assess off-target effects. Experimental results warn of off-target effects in AI-generated molecules and provide important insights for designing a drug against it.

INTRODUCTION

Innovative drug development is an extremely complex process that covers the discovery of molecular regulators,¹ chemical synthesis,² extensive preclinical testing³ and several phases of clinical trials.⁴ Only if a drug candidate is successful enough to pass all phases and regulatory approval, it can be devoted into the medical system.⁴ De novo molecular generation produces novel potential drug candidates interacting with specific biological binding sites, beginning the loop of creation.⁵ Heavy responsibilities lie in this stage as the generated molecules substantially impact subsequent processes and

huge cost would be in vain if candidates are missed.

An ideal drug can precisely bind with the target protein at the active pocket.⁶ However, the reality is that molecules often uncontrollably bind to other unexpected pockets, which is one of the main reasons for clinical trial failure.⁷ This is called off-target effects, arising from exposed structural features of other pockets similar to those of the target one, where proteins frequently undergo conformational changes brought by the flexibility and mobility of structures due to the cellular response to external stimuli.⁸

For drug discovery, research on machine intelligence largely empowers the procedures. One of the encouraging works is that two out of six compounds designed by Insilico Medicine found to be active in cellular assays against kinase DDR1.^{9,10} Inspired by successful attempts, researchers continuously upgrade the molecular generation models to obtain entities with desired chemical qualities. Masuda et al.¹¹ utilized convolutional neural networks (CNNs) to encode atomic grids into latent space and produce novel molecules that interact reasonably with binding sites through sampling from variable ligand latent space. Francoeur et al.¹² evaluated grid-based models with 3D CNNs, providing critical insights of designing CNNs models. Concurrently, recurrent neural networks (RNNs) are favored by direct steering of de novo molecular generation.¹³ With sequential modeling capability, RNNs provide a pronounced tendency to generate candidates that maintain similar binding modes as referenced molecules.^{14,15} For better representing the continuous latent space and sampling various drug candidates, Lim et al.¹⁶ exploited variational autoencoders (VAEs) to improve de novo molecular generation. As classical generative models, VAEs are widely used in this field with multiple variants,¹⁷⁻¹⁹ as well as generative adversarial networks (GANs)²⁰⁻²³ which optimize the joint distribution of drugs and proteins, resulting molecules with high drug-likeness. As the digital format of molecules are aligned with the graph representations in nature,^{24,25} Peng et al.²⁶ and Zhang et al.²⁷ represented molecules with graph neural networks (GNNs), passing forward the information of atoms, connections and even motifs.

Although existing molecular generation models continuously promote the properties of generated entities by optimizing the distribution of referenced molecules, off-target effects closely relating the biological side effect is not the priority to be considered. Among progress of molecular generation, priorities are witnessed continuously upgrading. In early times, the purpose of generation is to expand the library of chemical candidates.^{14,28} The learning is limited to simulate the structures of ligands already known effective to certain diseases. Afterwards, it is expected to generate molecules satisfied with certain chemical properties.^{11,15,16,18} This expectation is accomplished by inverse molecular design²⁹ which is conditioned on the specific physico-chemical properties including interactions and certain metrics. Recently, molecular generation is mainly focused on the target pockets.^{9,24,25,30,31} Generative models are designed to jointly simulate the correspondence between the structures of ligands and receptors, so novel molecules can be generated given by any unseen pocket. Considering off-target effects, the priority may be altered again where the generation is conditioned not only on target pocket

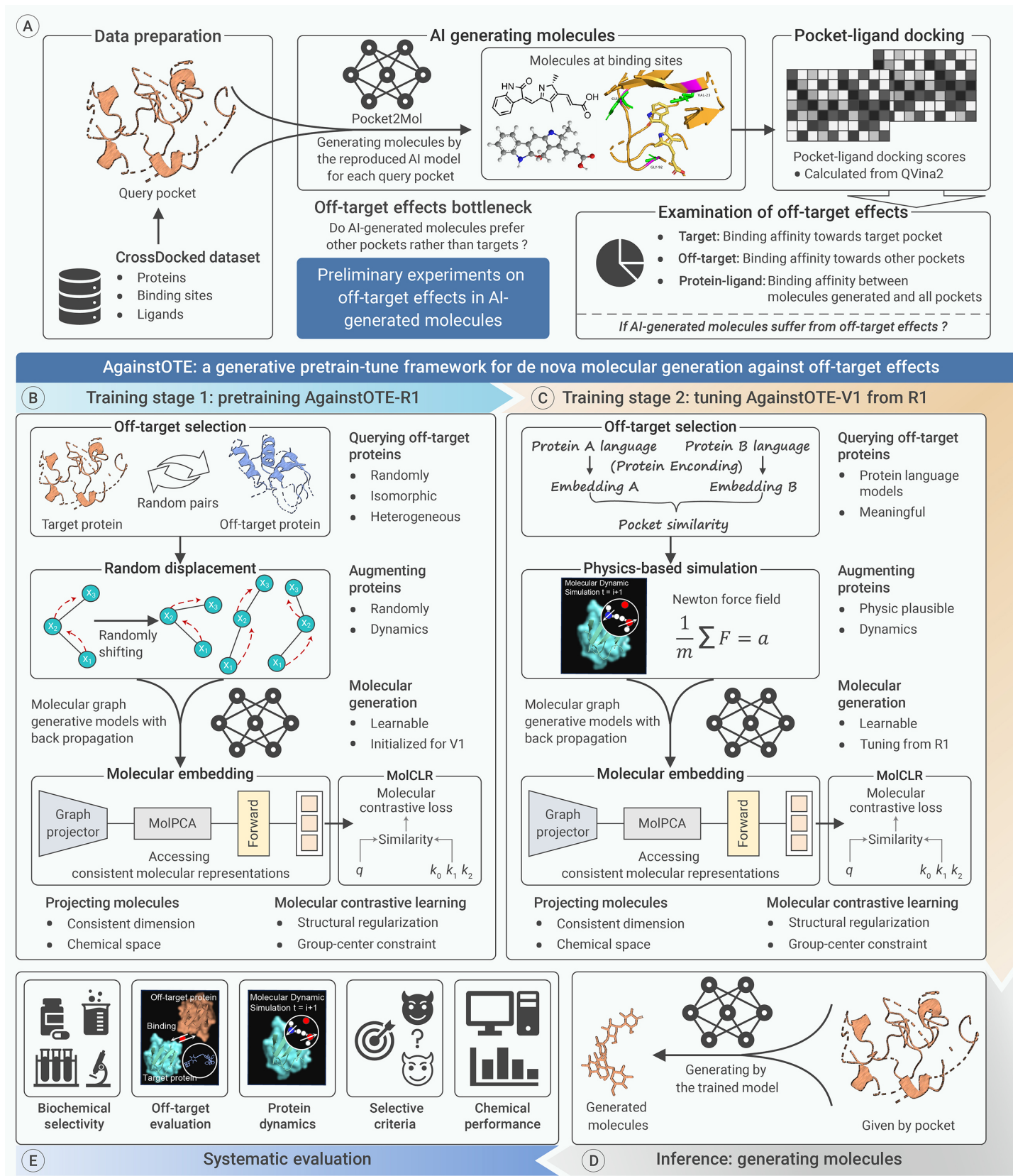


Figure 1. Workflow Our work is consisted of multiple phases. (A) Preliminary experiments on off-target effects in AI-generated molecules are intended to investigate if candidates generated by AI would suffer from off-target effects. A systematic solution starts from (B), where the pretraining of AgainstOTE-R1 is provided. (C) An advanced model AgainstOTE-V1 finetuned from R1 model is also provided. Given by any pocket as (D), the trained model is able to generate molecules with desired properties. In (E), extensive simulation and biochemical evaluation is provided.

ets, but also off-target pockets (i.e., pockets other than the target), narrowing down searching space. But this change must be predicated on the problem that if AI-generated molecules suffer from off-target effects.

In order to examine off-target effects in AI-generated molecules, we conducted preliminary experiments as in Figure 1A. Specifically on open-source CrossDocked dataset,¹² the state-of-the-art (SOTA) molecular gener-

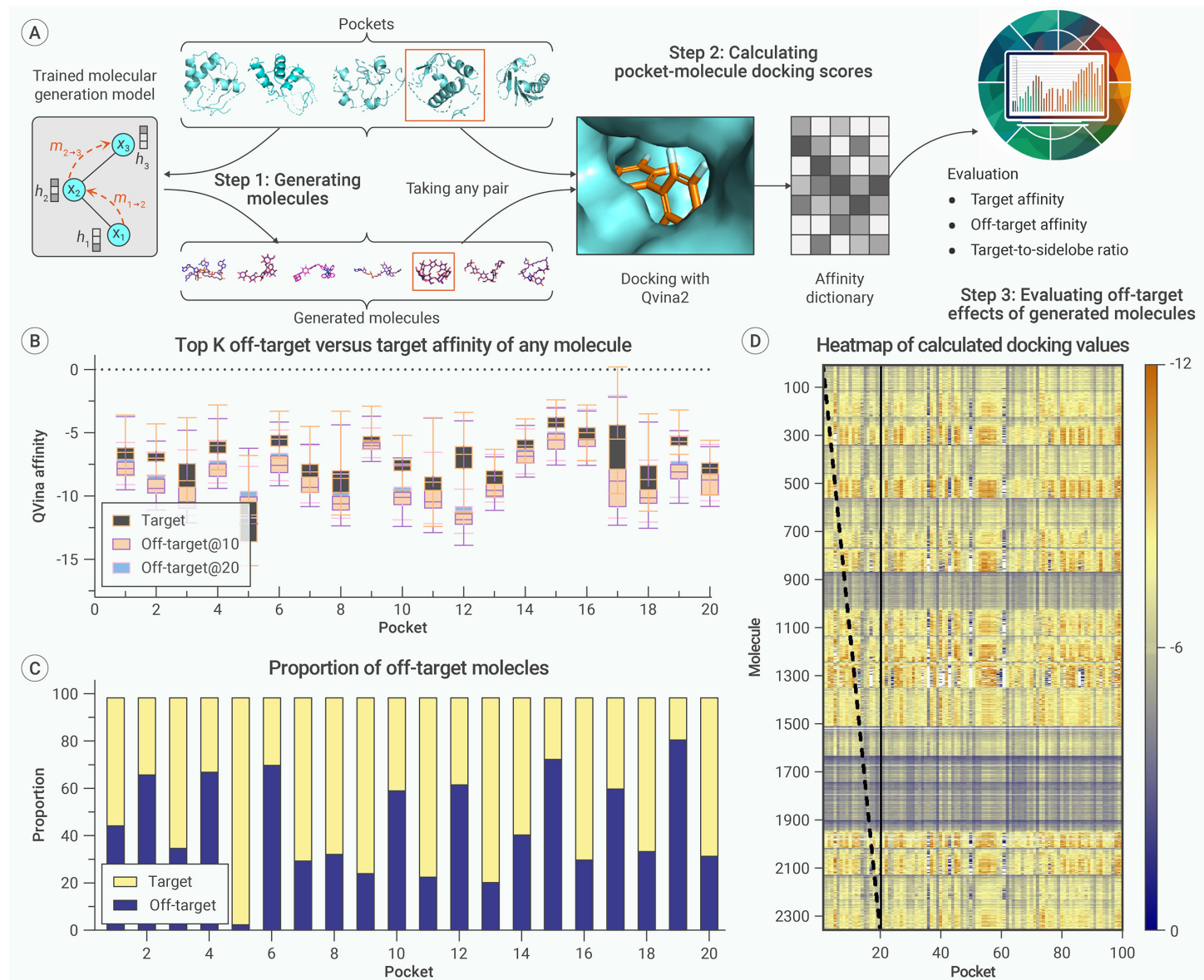


Figure 2. Preliminary experiments and steps of investigating off-target effects (A) Steps of examining off-targets are taken to do preliminary experiments. Three steps include generating molecules, calculating pocket-molecule docking scores and evaluation. (B) Top K off-target versus target affinity of any molecule draws the degree for a generated molecule would bind to off-target proteins. (C) Proportion of off-target molecules indicates the ratio of molecules higher likely binding to other proteins of all generated ones. (D) Heatmap of calculated docking values is from all docking scores calculated, showing the overall docking cases, where the lowest affinity refers to red.

ation model Pocket2Mol⁹ is reproduced and molecules are generated for all proteins in test set by the trained model. Then we calculated docking scores between generated molecules and proteins including both target ones and off-target ones with QVina2.³² Through multiple dimensions of statistical analysis on the dictionary of calculated docking scores, results from Figure 2 show that AI-generated molecules are more highly likely being docked with other proteins than their target ones, implying off-target effects underlying. Detailed results please refer to RESULTS section.

After examining the problem, we designed a generative probabilistic pretrain-tune framework for de novo molecular generation against off-target effects as in Figure 1. For designing AgainstOTE, a critical insight in this study is that molecular features should be evenly distributed in chemical latent space even when exposed structures of proteins are very similar. For example, if the structures of pockets are similar and unluckily generated candidates are not discriminative enough, these candidates are likely to bind with other pockets. For pretraining AgainstOTE-R1 as in Figure 3C, a dual-receptor cooperative training mechanism is exploited where both target and randomly selected off-target proteins are observed in the training pipeline. Structural displacement by randomly sampling from proteins is conducted to simulate conformational changes, creating multiple structural similar

samples. Protein-targeted molecular generation offers the generator to produce corresponding molecules. A regularizer is responsible for projecting generated molecules into chemical embeddings with consistent dimensions. Structural regularization implemented by a molecular contrastive learning architecture as in Figure 3F is specially designed to normalize searching space by enforcing the generator generating compounds with discriminative structures even when exposed binding sites are very close, evenly distributing latent features of generated molecules. In order to provide a physically and biologically plausible model, AgainstOTE-V1 is tuned from pretrained parameters of the R1 model. As in Figure 3A, more district biochemically meaningful off-target proteins are selected by querying protein language models. The protein semantic similarity determines the off-target proteins for tuning. As in Figure 3B, different from random displacement in the R1 model, we use molecular dynamics simulation instead to sample multiple augmented protein conformations following physically temporal trajectories. Both training stages are conducted in an end-to-end manner. Trained models inference novel drug structures given by target pockets as in Figure 1D.

For evaluating AgainstOTE as in Figure 1E, we conducted several stages of experiments on the CrossDocked dataset. It turns out that our systematic solution shows superior chemical properties compared to existing models.

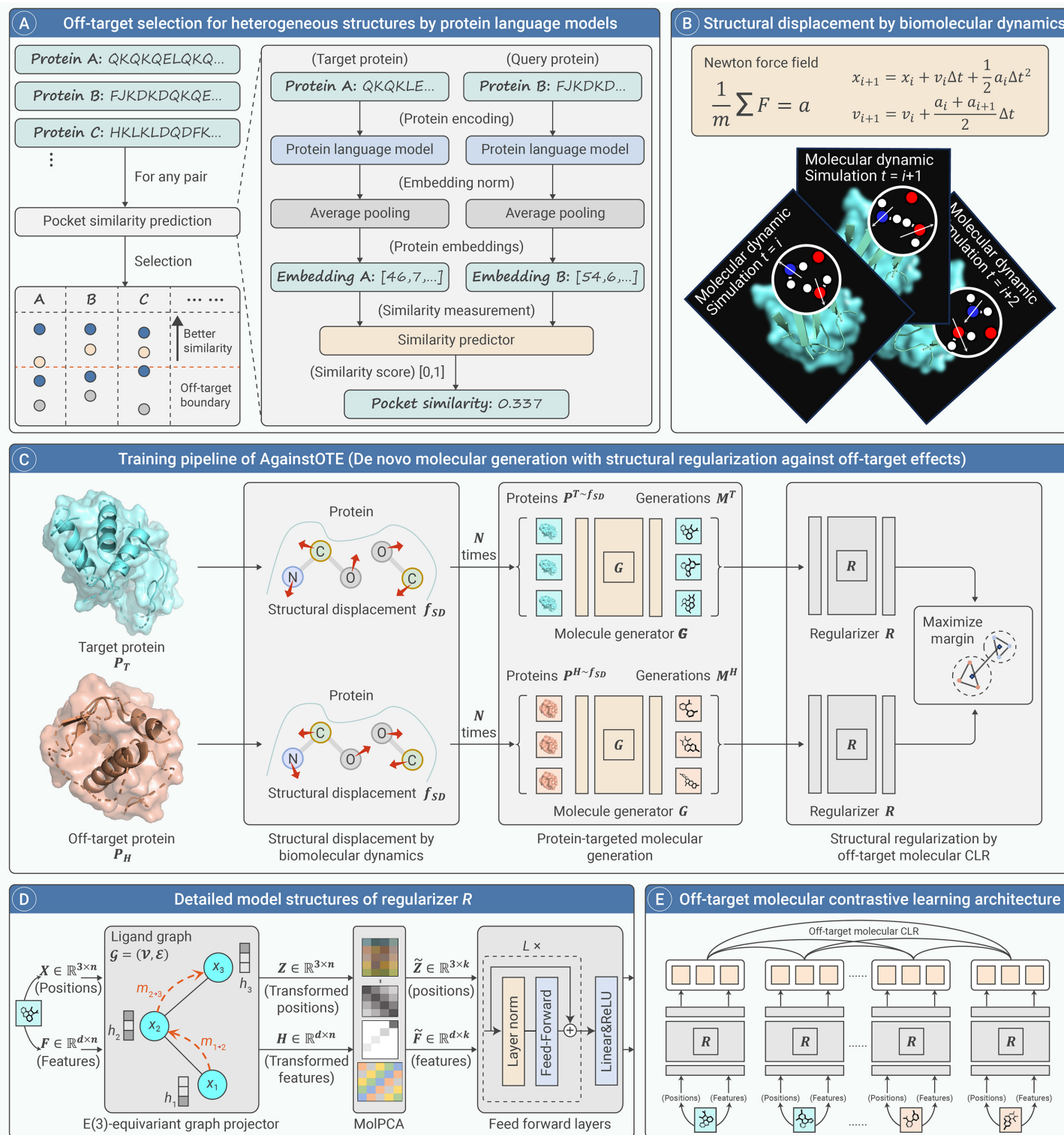


Figure 3. Model details of AgainstOTE (A) Procedure in selection of biochemically meaningful off-target proteins with heterogeneous structures by protein language models. (B) Structural displacement by molecular dynamics simulation for sampling physically varied temporal conformations of proteins. (C) Training pipeline of AgainstOTE illustrates the structures of its common composition. (D) Detailed model structures of the regularizer are specially designed, for regularizing the generator during training to generate molecules satisfying the constraints evenly distributing the molecular features in chemical space. (E) A molecular contrastive learning architecture is used for the regularization-based objective.

Chemical attributes of molecules generated are complied with our intension of being far away from those for off-target proteins. In the optimistic light of off-target effects, binding affinity³² with off-target proteins is evidently increased, lower likely to bind with off-target proteins. Although the affinity with the target is also increased, the degradation of target binding affinity from AgainstOTE-V1 was still in the accessible range. Meanwhile, Again-

stOTE shows outstanding performance in the designed TSR which was a systematic indicator against off-target effects. It needs to be mentioned that lower binding affinity indicates more stable binding states, referring to stronger docking capabilities.³² The distribution of molecular features produced by AgainstOTE is aligned with our intension. Several stages of ablation study reveal the understanding of the pipeline from deeper perspectives.

We selected several targets highly focused in inflammatory or tumor diseases, with developing potential of avoiding side effects in clinical application if precisely targeting is accessible. After generating and visualizing structures of generated molecules, results show that the molecules are plausible and to some degree different from referenced ones yet still well docked at the binding site. This behavior mainly comes from the altered learning priority leading to narrowed searching space considering both target and off-target proteins. At last, a compelling case towards TBK1 versus IKKε kinases is provided, where biochemical selectivity guided by TSR is experimented to directly assess whether off-target affinity is genuinely reduced in practice.

MATERIALS AND METHODS

In this section, we provide detailed descriptions of the generative pretraining framework of AgainstOTE, training and testing. Descriptions of datasets, calculation of the experimental measurement TSR and biochemical experimental methods are presented following.

To fulfill the overall architecture of AgainstOTE as in Figure 3C, we firstly define the general (i) 3D graph representations of ligands and receptors in this study. With this representation, input 3D structures of both target and off-target receptors are processed by (ii) structural displacement, creating more structural similar proteins as augmenting samples. Then we feed the structure-shifted proteins into the molecular generator which produces corresponding molecules highly likely being docked at binding sites. The generator is provided by (iii) protein-targeted molecular generation. Produced molecules are taken by structural regularization. This module is specially designed with an E(3)-equivariant molecular graph projector (encoding the produced molecules into chemical embedding space), MolPCA for consistent representations (extracting the prominent features of non-Euclidean embeddings³³ for same dimensions) and a molecular contrastive learning architecture implemented by group-center constraints (evenly distributing molecular embeddings with proper margins) by operating on molecular embeddings against off-target effects. Both pretraining and tuning stages of AgainstOTE are to optimize molecular generation with regularization, generating plausible molecules binding with the target receptor and away from off-target receptors in molecular embedding space.

3D graph representations of ligands and receptors

Assuming any protein made up of atoms can be denoted by a graph $g = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ represents the atoms as its nodes and $\mathcal{E}$ is the set of edges indicating the connectivity between atoms. $\mathcal{V}$ is consisted of one-hot atom features and 3D positions of nodes, where atom features are denoted by $F \in \mathbb{R}^{d \times n}$ indicating atom types and positions are denoted by $X \in \mathbb{R}^{3 \times n}$, d is the dimension of features and n is the number of atoms. To one atom of the protein, the node feature can be denoted by $f_i \in \mathbb{R}^d$, $f_i \in F$ and $1 \leq i \leq n$, its position is $x_i \in \mathbb{R}^3$. An overall graph representation is $g = (\mathcal{V}, \mathcal{E}) = (F, X, \mathcal{E}) = (\{f_i, x_i | 1 \leq i \leq n\}, \mathcal{E})$. This 3D graph representation is applied on all ligands and receptors in this work. For input of dual receptors in the pretraining stage, one protein is the target protein and another one is randomly selected which can be any other protein. While for the tuning stage, the off-target protein is selected by protein language models.³⁴ Proteins are firstly fed forward the pretrained protein language models and embeddings are accessed as in Figure 3A. Coming after normalization and Cosine similarity, the similarity score is used to determine if the query protein is the off-target one. All queries with low similarity scores below the set threshold are selected.

Structural displacement

Instead of existing drug design models taking only the target protein as input during training, we provide both the target protein P^T and the off-target protein P^H as input. Against off-target effects, generated molecules should be sensitive to different structures of proteins even when the structures are very close. This is why we let both target and off-target proteins be seen during training, and create more structure-similar samples next. We expect to create these close enough structures of input proteins by shifting the positions of atoms with structural displacement. In AgainstOTE-R1, assuming that the random shifted displacement Δx_i of the node i is sampled from the Gaussian distribution $N(\mu, \sigma)$, $\Delta x_i \sim N(\mu, \sigma)$, where μ is the mean value and σ is

the standard deviation of the Gaussian distribution. $\Delta x_i \in \Delta X$, where ΔX denotes displacement with the same number of the corresponding protein. To any input protein graph g , we add ΔX onto X to simulate random displacement of structures. This procedure can be denoted by

$$f_{sd}(X) = X + \Delta X, \quad (1)$$

$$\text{where } \Delta X = \{\Delta x_i | i = 1, \dots, n\}, \text{ and } \Delta x_i \sim N(\mu, \sigma). \quad (2)$$

We conduct respectively N times of structural displacement on both target and off-target proteins, resulting N position shifted proteins $P^{T \sim f_{sd}}$ for the input target protein and N position shifted proteins $P^{H \sim f_{sd}}$ for the input off-target protein. Here is the process of random displacement in AgainstOTE-R1 and the protein dynamics simulation in AgainstOTE-V1 is conducted by the classical physics-based molecular dynamics simulation.^{35,36}

Protein targeted molecular generation

Respectively to each shifted protein $P^{T \sim f_{sd}(i)}$ and $P^{H \sim f_{sd}(i)}$ of $P^{T \sim f_{sd}}$ and $P^{H \sim f_{sd}}$, $1 \leq i, j \leq N$, we expect a molecule generator G to predict the molecular structures. The existing deep generative molecular generator Pocket2Mol⁹ is exploited as G and the major steps of generating each atom are the same as Pocket2Mol. Formally with this generator, the generation procedure based on the position shifted proteins can be denoted by:

$$M_{(k)}^{T(i)} = G(M_{(k-1)}^{T(i)}, P^{T \sim f_{sd}(i)}), M_{(k)}^{H(j)} = G(M_{(k-1)}^{H(j)}, P^{H \sim f_{sd}(j)}), k > 1 \text{ and } 1 \leq i, j \leq N, \quad (3)$$

$$M_{(k)}^{T(i)} = G(P^{T \sim f_{sd}(i)}), M_{(k)}^{H(j)} = G(P^{H \sim f_{sd}(j)}), k = 1 \text{ and } 1 \leq i, j \leq N. \quad (4)$$

The generated M^T and M^H are the candidate molecules at the corresponding binding sites, $M^T = \{M^{T(i)} | 1 \leq i \leq N\}$, $M^H = \{M^{H(j)} | 1 \leq j \leq N\}$. The generated molecules share the same representation as g .

Structural regularization with molecular embeddings

As the molecules generated suffer from off-target effects as examined, we specially design this structural regularization to enforce the molecule generator sensitive to different receptors even when the structures are very close only with position displaced. For accomplishing the structural regularization network as in Figure 3D, generated molecules M^T and M^H are encoded by the E(3)-equivariant graph projector, projecting the molecules from 3D structures into chemical embeddings. MolPCA is conducted on the projected embeddings for consistent molecular vectors for consistent dimensions. These consistent molecular feature vectors are supported for the following constraints.

E(3)-equivariant molecular graph projector. As E(3)-equivariant networks are confirmed to offer good capabilities at modeling 3D graph structures, we design the projector with the E(3)-equivariant graph networks.³⁷⁻³⁹ This projector learns to encode the input molecule with node features and positions, performing graph message passing across complex graph. Assuming for any molecule which is the format of g , X is the positions of atoms and F is the node features, $X \in \mathbb{R}^{3 \times n}$, $F \in \mathbb{R}^{d \times n}$. The projector is comprised of multiple cascaded E(3)-equivariant graph layers, ensuring any independent position displacement of the original structures being exactly reflected in the encoded vectors. To any input molecule g , the calculation in l -th layer is as follows,

$$m_{j \rightarrow i} = \phi^e \left(h_i^{(l)}, h_j^{(l)}, \|x_i^{(l)} - x_j^{(l)}\|^2, f_{j \rightarrow i} \right), \forall (i, j) \in \mathcal{E}, \quad (5)$$

$$m_i = \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} m_{j \rightarrow i}, \forall (i) \in \mathcal{V}, \quad (6)$$

$$x_i^{(l+1)} = \alpha x_i^{(l)} + (1 - \alpha) x_i^{(l)} + \frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \frac{x_j^{(l)} - x_i^{(l)}}{\|x_i^{(l)} - x_j^{(l)}\|} \phi^s(m_{j \rightarrow i}), \forall (i, j) \in (\mathcal{V}, \mathcal{E}), \quad (7)$$

$$h_i^{(l+1)} = (1 - \beta) \cdot h_i^{(l)} + \beta \cdot \phi^h(h_i^{(l)}, m_i, f_i), \forall (i) \in \mathcal{V}, \quad (8)$$

where $\mathcal{N}(i)$ refers to the graph neighbors of node i , α and β respectively represent weights of position-wise (scalar) and representations-wise (vector)

skip connection. Learnable parameters ϕ^* are stacked by shallow neural networks passing scalars and ϕ^b are also stacked by shallow networks passing vectors. The transformed positions after this projector are represented by $Z \in \mathbb{R}^{3 \times n}$ and transformed features are represented by $H \in \mathbb{R}^{d \times n}$. H and Z reflects the sensitive structural difference of any input molecules, with this process going through all generated molecules M^T and M^H .

MolPCA for consistent representations. As the number of atoms of different generated molecules are always different which is a non-Euclidean problem, it is not acceptable for direct calculation between the projected molecular embeddings. For unifying the dimensions of molecule-based embeddings with different numbers of atoms, we present MolPCA. To any transformed molecular embedding H and Z , each node feature is stacked as an itemset in the form of 1-dimensional vector and a transaction is created by stacking all atoms, each position is operated by the same manner. This process can be denoted by $tran(H)$ and $tran(Z)$, resulting the transaction H_{tran} and Z_{tran} . Then singular value decomposition⁴⁰ is conducted on the transactions, $svd(H_{tran})$ and $svd(Z_{tran})$ separately. The mined features are represented by $\tilde{Z} \in \mathbb{R}^{3 \times k}$ and $\tilde{H} \in \mathbb{R}^{d \times k}$, k is the consistent dimension designed by MolPCA.

Finally, we stack multiple feed forward layers to output the final features for calculating the constraint. The feed forward layers are first stacked by L blocks of layer normalization⁴⁰ and feed forward operation.⁴¹ One linear layer and one ReLU layer⁴² are cascaded at last. After the network, the projected embeddings of M^T are denoted by $P^T = R(M^T)$ and the embeddings of M^H are $P^H = R(M^H)$.

Group-center constraints for structural regularization. For regularizing the projected molecular embeddings, the core consideration is that firstly the embeddings from molecules originating from the same input protein should be evenly distributed, required in the molecular group (i.e., a group contains molecules originating from the same input protein such as P^T). Meanwhile, distance between the centers of the group-based embeddings should be still discriminative like molecules in the group. The group constraint requires the structures of generated molecules being different from each other even when the proteins are very similar and the center constraint maintains the larger margin within molecules targeted at different receptors. To fulfill the group constraint, inspired by normalized temperature-scaled cross entropy loss (NT-Xent) function,⁴³⁻⁴⁵ for P^T and P^H , there are $2N$ embeddings in one batch totally,

$$\mathcal{L}_{P \rightarrow T} = \frac{1}{N} \sum_{i,j \in \{1, \dots, N\}} - \log \frac{\exp(\text{sim}(P^{T(i)}, P^{H(j)})/\delta)}{\sum_{j=1}^N (\exp(\text{sim}(P^{T(i)}, P^{H(j)})/\delta) + \exp(\text{sim}(P^{T(i)}, P^{T(j)})/\delta))}, i, j \in \{1, \dots, N\}, \quad (9)$$

$$\mathcal{L}_{T \rightarrow P} = \frac{1}{N} \sum_{i,j \in \{1, \dots, N\}} - \log \frac{\exp(\text{sim}(P^{H(i)}, P^{T(j)})/\delta)}{\sum_{i=1}^N (\exp(\text{sim}(P^{H(i)}, P^{T(j)})/\delta) + \exp(\text{sim}(P^{H(i)}, P^{H(j)})/\delta))}, i, j \in \{1, \dots, N\}, \quad (10)$$

$$\text{where } \text{sim}(A, B) = \frac{A^T B}{\|A\| \|B\|}, \quad (11)$$

where δ is the temperature parameter, $\mathcal{L}_{P \rightarrow T}$ regularizes the distance between all targeted molecules and all off-target ones, $\mathcal{L}_{T \rightarrow P}$ is the same manner.

To fulfill the center constraint,

$$\mathcal{L}_{\text{Center}} = -\log \left\| \frac{1}{N} \sum_{i=1}^N P^{T(i)} - \frac{1}{N} \sum_{i=1}^N P^{H(i)} \right\|. \quad (12)$$

The final group-center constraints are the combination of the group constraint and the center constraint.

Training and testing

For training in the overall framework of both stages, we take the pairs of target and off-target proteins as input. Structural displacement is conducted on input proteins. The training objective is comprised of two parts, molecule generation and structural regularization. For molecule generation, the learning objective is the same as in Pocket2Mol with randomly masking atoms of molecules and for recovering the masked atoms. The loss functions of this

part include the binary cross entropy loss of the frontier prediction $\mathcal{L}_{fro}$, the negative log likelihood of the masked atom positions $\mathcal{L}_{pos}$, the cross-entropy losses for the classification of element type and bond type $\mathcal{L}_{ele}$ and $\mathcal{L}_{bond}$. We denote the objective of this part, $\mathcal{L}_{gen} = \mathcal{L}_{fro} + \mathcal{L}_{pos} + \mathcal{L}_{ele} + \mathcal{L}_{bond}$. For the structural regularization part, the learning objective is composed of the group constraint $\mathcal{L}_{P \rightarrow T}$, $\mathcal{L}_{T \rightarrow P}$ and the center constraint $\mathcal{L}_{\text{Center}}$. We denote the objective of this part is $\mathcal{L}_{reg} = \mathcal{L}_{P \rightarrow T} + \mathcal{L}_{T \rightarrow P} + \mathcal{L}_{\text{Center}}$. These two parts are trained jointly, the molecule generator is optimized while the regularizer demands it against the off-target effect. As a result, in the training state, we aim to update the parameters of our model by minimizing the loss function,

$$\mathcal{L} = \mathcal{L}_{gen} + \mathcal{L}_{reg}. \quad (13)$$

With the model trained, we only need generator G to generate molecules by feeding the target proteins with the procedure in above equations.

CrossDocked dataset

The data source is from CrossDocked¹² dataset, containing originally collected 22.6 million poses of protein-molecule 3D structures, where ligands with over 1000 Da molecular weight are removed and structures are stripped of water and aligned to the binding sites. The ligands with binding pose RMSD greater than 1 Å are filtered out and the subset is created where the ligand is docked to its cognate receptor within 1 Å.²⁶ The filtered subset is consisted of 184,057 pairs of protein-ligand structures with atom positions, atom types, link indexes, link types, binding sites, etc. For utilizing this dataset,^{26,9,11} 100,000 pairs of protein-ligand 3D structures are randomly split as the training data from the clusters and 100 pairs are split from the rest as the testing data for evaluation of the trained model. This processed data supports all experiments conducted in this research.

Calculation of TSR

For measuring the degree of target binding abilities standing out compared to off-target binding abilities, we select 10 proteins sequentially, the trained models of the baseline and ours take these proteins respectively as input and generate around 100 molecules to any one protein. Assuming that the generated molecules are denoted by $\tilde{M}$, the proteins are $\tilde{G}$. We calculate the docking scores between each one generated molecule and each one protein, $S = \{\text{Affinity}(\tilde{m}_i, \tilde{g}_j) \mid \forall \tilde{m}_i \in \tilde{M}, \forall \tilde{g}_j \in \tilde{G}\}$, using QVina2. Among these calculated scores, ones docked with their target proteins are denoted S_{Target} , those with off-target proteins are $S_{\text{Off-Target}}$. With S_{Target} and $S_{\text{Off-Target}}$, we define a technical measurement to intuitively show the off-target degree, called target-to-sidelobe ratio,

$$f_{\text{TSR}}(x, y) = \frac{1}{1 + e^{x-y}}, \quad (14)$$

$$x = \text{TopK}(S_{\text{Target}}), y = \text{TopK}(S_{\text{Off-Target}}), \quad (15)$$

where $f_{\text{TSR}}(\cdot)$ is the calculated results. $\text{TopK}(\cdot)$ is operated at when $K = 5$ here in this study. This measurement is the higher the better and is inspired by peak-to-sidelobe ratios⁴⁶ in the visual tracking field.

Reagents and Materials

The MTT Assay Kit (ST316) was purchased from Beyotime Biotechnology (China); Antibodies against TBK1 (#38066), IKKε (#3416) were purchased from Cell Signaling Technology, Inc (USA); The EGFP-TBK1 plasmids and mCherry-IKKε plasmids were constructed by CORUES BIOTECHNOLOGY (China). β-actin antibody (GB11001-100) was purchased from servicebio (China).

Cell culture

HEK293T cells and MC38 cells were cultured with RPMI 1640 medium separately supplemented with 10% FBS and 1% penicillin/streptomycin.

3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay

The MC38 cells were seeded into 96-well plates at a density of 1×10^4 cells in 100 μl of complete medium and incubated at 37 °C. At the end of each

experiment, removing the 96-well plate from the incubator and discard the culture medium. Add 10 μ L of 5 mg/mL MTT solution to each well. Return the plate to the incubator and incubate for 4 h. Carefully aspirate and discard the supernatant from each well. Add 150 μ L of Dimethyl Sulfoxide (DMSO) to each well to dissolve the formazan crystals. Vortex the plate for 10 minutes to ensure complete dissolution. Measure the absorbance (OD value) at 490 nm using a microplate reader.

Microscale Thermophoresis (MST) assay

HEK293T cells were transfected with EGFP-TBK1 plasmids or mcherry-IKKe plasmids. After 48 h, the cells were lysed and centrifuged. Each ligand diluent was incubated with an equal volume of cell lysate for 10 min. For the measurement, each ligand dilution was mixed with one volume of cell lysis, which led to a final fluorescence of 6000 and final ligand concentrations in the nmol/L to mmol/L range. After 10 min incubation followed by centrifugation at 10,000 g and 4 °C for 10 min, the samples were loaded into Monolith NT.LabelFree zero-background capillaries (NanoTemper Technologies). MO.Affinity Analysis v2.3 software was used to analyze the data.

Cellular thermal shift assay (CETSA)

MC38 cells were treated with Pentoxifylline at different concentrations, and the control group added DMSO of the same volume, and two groups of cells were collected after 2 h. Two groups of cells were washed twice and resuspended with 500 mL of PBS at the last time. Both groups are divided into 10 servings, each of which is 50 mL, and loaded into a PCR tube. Set the PCR meter to 10 temperature gradients (43, 46, 49, 52, 55, and 58 °C) to each sample in the control group corresponds to a temperature for 3 min. The protein was evaluated by Western blot using the indicated antibody.

Western blot

Cells were lysed in lysis buffer supplemented with protease and phosphatase inhibitor. Proteins were quantified by the Bradford assay (HyClone-Pierce). Equal quantities of proteins from different samples were separated on SDS-PAGE gels and transferred onto polyvinylidene difluoride membranes. The cropped membranes were hybridized with different primary antibodies at 4 °C overnight, washed three times with phosphate-buffered saline containing 0.1% Tween20 (PBST), and incubated with the secondary antibodies of the corresponding species for 1 h. Detection was performed using a LumiGLO chemiluminescent substrate system.

Chemistry

The synthetic route of 3,7-dimethyl-1-(5-oxohexyl)-3,7-dihydro-1H-purine-2,6-dione (Pentoxifylline): A mixture of theobromine (0.45 g, 2.5 mmol), 6-chloro-2-hexanone (0.34 g, 2.5 mmol), potassium carbonate (0.69 g, 5 mmol), tetrabutylammonium iodide (10 mmol%) and dimethylformamide (5 mL) was stirred for 8 h in 100 °C. Then the DMF excess was removed on a rotary evaporator. The residue was extracted with petroleum ether (20 mL) to remove remained electrophilic reagent and next the product was extracted with CH_2Cl_2 (20 mL). The undissolved material was filtered off. The filtrate was concentrated under reduced pressure and purified by column chromatography on silica gel.⁴⁷ The synthetic route of 5-chloro-2-(2,4-dichlorophenoxy) phenol (Triclosan): 4-Chloro-2-methoxyphenol (3.17 g, 20 mmol, 1.0 equiv), 1,3-dichloro-4-fluorobenzene (6.60 g, 40 mmol, 2.0 equiv), and Cs_2CO_3 (13.04 g, 40 mmol, 2.0 equiv) were dissolved in dry DMF (50 mL) and stirred at 130 °C for 14 h. The reaction mixture was cooled to room temperature and poured into H_2O (200 mL), and the products were extracted into EtOAc (3×100 mL). The combined organic phases were washed with H_2O (100 mL), dried over Na_2SO_4 , filtered, and concentrated in vacuo to give crude material as a yellow solid, which was purified by recrystallization from $\text{CH}_2\text{Cl}_2/\text{EtOAc}$ to give the pure intermediate product as a white crystalline material. The intermediate was placed into a round-bottom flask, which was evacuated and refilled with Ar, followed by addition of dry CH_2Cl_2 (to prepare a 0.1 M solution). The flask was cooled to -78 °C, and BBr_3 (1.0 M solution in CH_2Cl_2 , 5 equiv) was added dropwise. The reaction mixture was warmed gradually to room temperature within 1 h, and stirring was continued for a further 2.5–3 h. The reaction was terminated by the addition of H_2O , and the products were extracted into CH_2Cl_2 (2×25 mL). The combined organic

phases were dried over Na_2SO_4 , filtered, and concentrated in vacuo to give crude material, which was subsequently purified by preparative HPLC.⁴⁸ Pentoxifylline and Triclosan were characterized by NMR for their structural information, and the results can be found in Figures S1–S2.

RESULTS

Overview

Our work is consisted of several main parts. At first, we conduct preliminary experiments to examine if AI-generated molecules suffer from off-target effects, as in Figure 1A. Supported by the released CrossDocked receptor-ligand pairs of data, the SOTA molecular generation model Pocket2Mol is reproduced and used to generate candidates for each pocket in test set. Then we calculate docking scores between all molecules generated and proteins with QVina2. By statistically analyzing the docking scores, off-target effects in AI-generated molecules are examined. In the second part, we design a generative pretrain-tune framework for molecular generation against off-target effects, as in Figures 1B–C. Trained models take target pockets as input and output molecules with desired properties as in Figure 1D. Extensive experiments are conducted to check the behaviors of this framework as in Figure 1E. Mathematical expressions and more details are available in MATERIALS AND METHODS section.

Preliminary experiments on examining off-target effects in AI-generated molecules

The experimental pipeline is in Figure 2A. In the first step, we reproduced the SOTA 3D molecular generation model Pocket2Mol with the released model trained with the CrossDocked dataset. This generative model takes the protein as input and produces the molecules at the binding site. We took all 100 proteins in test set and generate around 100 molecules for each protein, resulting totally around 10,000 molecules.

As off-target effects illustrate the phenomenon that molecules can be docked with off-target proteins,⁹ in the second step we calculated the docking scores between the generated molecules and proteins including both target and off-target ones with QVina2. For simulating molecules entering the biological environment where a molecule meets many other proteins, we took the molecules generated from the first 20 proteins in order and all proteins in this stage. As around 100 molecules were generated from each protein, docking calculations were densely operated between around 2,000 molecules and 100 proteins, resulting 200,000 times of docking calculation, which costs around 1 month.

In the third step, we conducted multiple dimensions of statistical analysis on these calculated docking scores, including the top K off-target versus target affinity of each molecule as in Figure 2B, proportion of off-target molecules as in Figure 2C and heatmap of all calculated docking values as in Figure 2D. Specifically, top K off-target versus target affinity of each molecule refers to how likely one generated molecule would be docked with other proteins, where the target affinity is represented as docking scores between molecules and target proteins, while the top K off-target scores are mean values of largest K docking scores between molecules and off-target proteins. We compiled top 10 off-target scores, top 20 off-target scores and target scores of any molecule. Proportion of off-target molecules illustrates that the proportion of generated molecules which prefer docking with off-target proteins at their binding sites than the target, where the proportion is represented by the number of these off-target molecules versus the number of molecules prefer their target. The off-target molecule in this study is a molecule that has a lower docking score with any other protein than its target. Heatmap of all calculated docking values shows the docking stability between molecules and proteins completely, any one pixel in this plot is the actual molecule-protein docking value. Through this stage of statistics, it is evident that off-target effects widely exist in AI-generated molecules.

Specifically, through Figure 2B, to molecules generated based on the 20 proteins, the distributions of both top 10 and top 20 docking scores with off-target proteins are showing lower affinity compared to the distributions of docking scores with their targets. This phenomenon happens in 19 out of 20 pockets. It indicates that AI-generated molecules are more likely to bind with other proteins other than the target protein. From the statistics in Figure 2C, to molecules generated from the 20 pockets, the proportions of them are not

satisfactory, almost half of them are higher than 50%, where the proportion means how many off-target molecules would show in the generated molecules. This reveals the high risk of AI-generated molecules that a large number of generated molecules are likely off-target. From statistics in Figure 2D, heatmap of all calculated docking values, as the molecules are targeted at the first 20 pockets, the docking scores between these molecules and the first 20 pockets should have been expected to be the lowest (where the lowest score means the best affinity, denoted by red in the figure), but the heatmap shows the results far from this expectation. The generated molecules are more likely to exhibit lower affinity with other protein at their binding sites. This heatmap comprehensively shows that there is a high risk for AI-generated molecules applying in continuous development even clinical experiments, where off-target effects are widely found in these molecules.

Training stage 1: pretraining AgainstOTE-R1

During the pretraining pipeline of AgainstOTE-R1, for observing both target and off-target proteins in the dual-receptor cooperative training process, we consider (i) 3D graph representations of ligands and receptors of both proteins as input. (ii) Structural displacement is conducted to process the input data, simulating the conformational changes of proteins caused by random motion of atoms, creating multiple structural similar samples. Protein-targeted molecular generation mainly provides the generator to produce the molecules targeted on each of input samples. We specially designed (iii) a regularizer as in Figure 3D to embed the generated molecules into chemical latent space and (iv) the structural regularization to enforce the generator producing compounds with discriminative structures sensitive to different receptors even when they are very similar in a molecular contrastive learning architecture as in Figure 3E.

3D graph representations of ligands and receptors. The representations of molecules are varied from 1D descriptors like simplified molecular input line entry system (SMILES),^{49,50} 2D graphs¹⁷ and 3D structures.^{9,26} SMILES is a standardized lexical representation employed to encode the structure of molecules with textual strings encapsulating the molecular topology such as ring formations, branching patterns and connectivity.⁴⁹ Boyle et al.⁵⁰ introduced DeepSMILES, a probabilistic model dealing with common invalid syntax issues, to generate SMILES strings with desired properties. 2D graphs describe the two-dimensional topological structure of molecules.¹⁷ A molecule is represented as a graph with atoms as nodes and chemical bonds as edges interconnecting. Jin et al.¹⁷ presented a junction tree variational autoencoder that generates the 2D molecular graphs through a two-stage process, with a tree-structured framework encompassing chemical substructures and subsequently utilizing a graph message network to integrate these substructures into a cohesive molecule. In our pipeline, we use 3D graph representations which describe molecules as the three-dimensional topological structures, where the 3D graph consists of nodes and edges. With nodes representing atoms by one-hot encoding and 3D positional features, edges indicating interatomic connectivity, this representation is applied uniformly to all ligands and receptors in this study, allowing our explorations of molecular interactions in 3D space.

Structural displacement. For simulating conformational changes of proteins caused by random movement of atoms due to the cellular response to external stimuli,⁸ multiple augmenting samples of input are created by structural displacement. For randomly shifted displacement,⁵¹ the position offset vectors are sampled from the Gaussian distribution controlled by its mean value and standard deviation. This random offset vector is added to positional features of any input graph. Through this augmentation, we gain multiple augmented versions of both target proteins and off-target ones. This is for continuous operations expected for meanly distributing the distances between molecules even generated from similar proteins undergoing conformational changes. With the structural shifted versions of input, a molecular generator is expected to generate molecules for each one of the structural shifted proteins. The existing deep generative molecular generator Pocket2Mol is exploited as the component of protein-targeted molecular generation. This generator in our pipeline is expected to design molecules given by any input from structural displacement.

Consistent molecular embeddings from the regularizer. This module is specially designed for this pipeline. It is jointly trained with the generator,

enforcing the generator to generate molecules against off-target effects. Structural regularization is accomplished by a regularizer embedding the molecules into projected latent molecular space and group-center constraints evenly distributing the embedded features. The details of the regularizer are accessed in Figure 3D. It comprises of an E(3)-equivariant molecular graph projector, the molecular principal component analysis (MolPCA) and feed forward layers. We place an (i) E(3)-equivariant molecular graph projector at the initial of the regularizer to encode input molecules with node features and positions. As E(3)-equivariant networks³⁷⁻³⁹ are confirmed to offer good capabilities at modeling 3D graph structures, E(3)-equivariant graph networks are expected to perform graph message passing across the complex graph. As the number of atoms from different molecules are always different which is a non-Euclidean problem,³³ it is not acceptable for direct calculation between the projected molecular embeddings. For unifying the dimensions of embedding with different numbers of atoms, we present (ii) MolPCA. In MolPCA, node features and edge features of each input graph vector are transformed into a transaction for continuous mining.⁵² By stacking each node as an itemset in the form of 1D vector and a transaction of node features is created. The transaction of edge features is the same. Principal component decomposition is operated on the created transactions to find the prominent feature vectors and feature values,⁵³ the unified feature vectors are accessed by specifying a fixed dimension. Finally, we stack multiple (iii) feed forward layers consisting of layer normalization,⁴⁰ multilayer perceptron,⁴¹ linear layers and ReLU⁴² to output the final features from the regularizer for continuous calculating the constraints.

Structural normalization in a molecular contrastive learning architecture. A molecular contrastive learning architecture implemented by group-center constraints requires the evenly distribution of molecular features in both group-based embeddings and center-based embeddings, the combination of the group constraint and the center constraint. In this study, a group refers to molecular features originated from the same input protein and the center means the center of features in each group. The group constraint is inspired by normalized temperature-scaled cross entropy loss,⁴³ requiring the structures of generated molecules being discriminative from each other even when the proteins are very similar. The center constraint punishes the distance between the feature centers of different groups, maintaining the difference of generated structures for different receptors. Hyperparameters and settings for experiments are available in Table S1.

Training stage 2: tuning AgainstOTE-V1 from the R1 model

During the tuning stage of AgainstOTE-V1 which is somewhat different from the R1 version, an additional selection of heterogeneous structures for more district determination of off-target proteins is conducted and the random sampling of protein conformations is replaced by a more physical manner with molecular dynamics simulation.

Off-target selection for heterogeneous structures by protein language models. General-purpose protein language models have emerged as powerful tools for extracting biochemically meaningful embeddings from protein amino acid sequences, offering compact representations that encapsulate functional and structural information.^{34,54} By leveraging transformer-based architectures and techniques like masked language modeling,^{55,56} these models can automatically learn complex sequence-function relationships.^{57,58} In this advanced V1 version, we select more district off-target proteins with heterogeneous structures as tuning samples. With implementation of the protein language model,³⁴ we design a pipeline for selection. Proteins are firstly fed forward the pretrained protein language models and embeddings are accessed as in Figure 3A. By a series of normalization and Cosine similarity, the decisive score is calculated for determining if the query protein is the off-target one for the target. By setting a boundary, all queries with low similarity scores are selected as off-target proteins.

Structural displacement by molecular dynamics. Different from random sampling in R1 version, here we use physically temporal simulation of protein structures by molecular dynamics.^{35,36} Molecular dynamics simulation plays an important role in computational biophysics, providing atomic-level insights into biomolecular behaviors over time horizons.⁵⁹ By numerically solving classical Newton's equations of motion for all atoms in a dynamic system, it can reveal motion processes such as protein folding,⁶⁰ ligand binding⁶¹ and

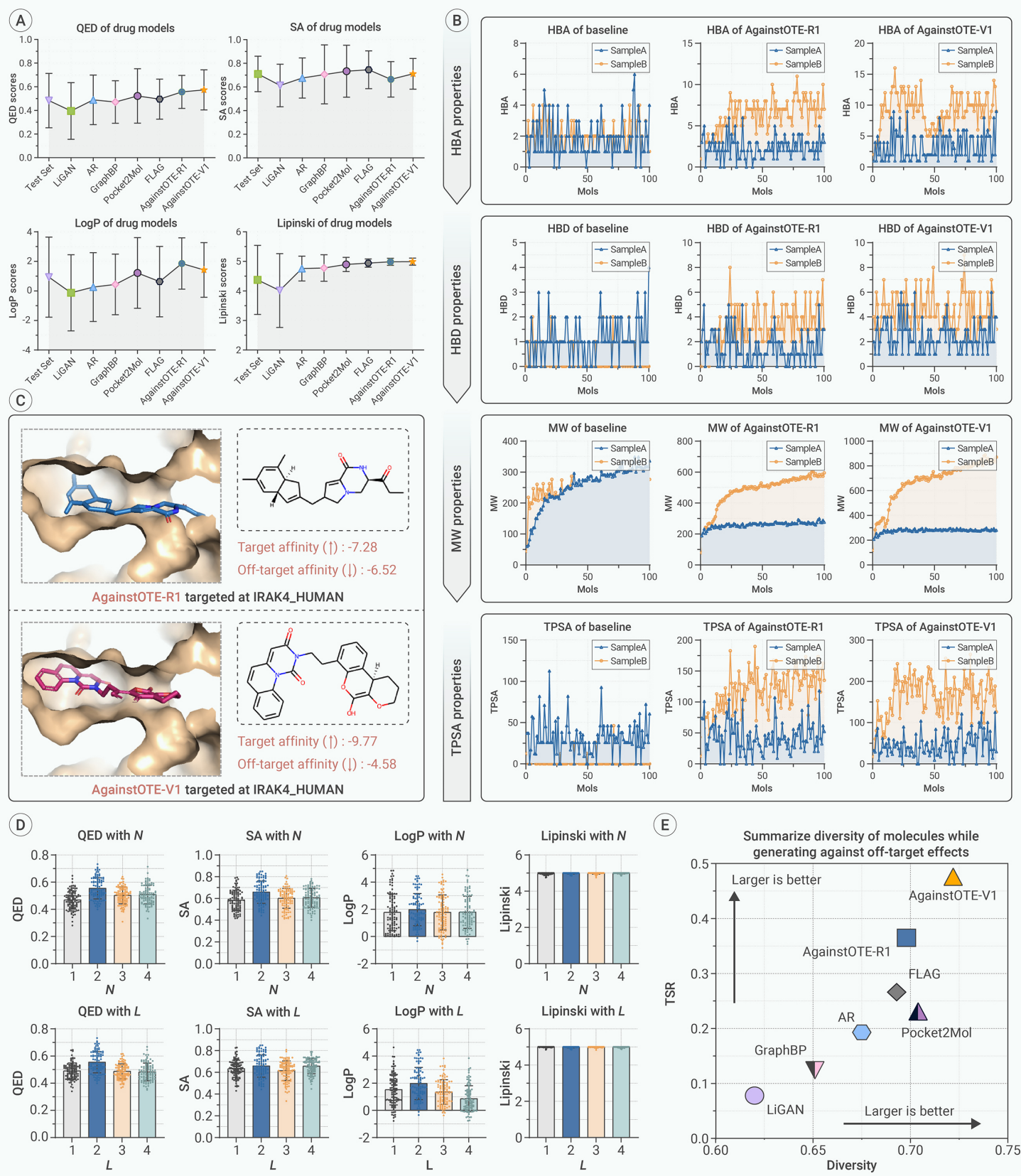


Figure 4. Evaluation in chemical properties (A) Comparison of chemical properties reports chemical properties including QED, SA, LogP and Lipinski of existing models and AgainstOTE. (B) Distributions in HBA, HBD, MW and TPSA of generated molecules from the baseline and AgainstOTE are plotted. (C) An example of molecules generated towards IRAK4 by AgainstOTE is given. (D) Ablation study with N and L investigates the influence of structural displacement and structural regularization in chemical properties. (E) TSR and diversity of drug models are plotted.

conformational changes⁶² that are difficult to observe. By utilizing this simulation as structural displacement, multiple augmented proteins with physical

sampling temporal trajectories are produced, instead of randomly shifted positions of atoms in proteins.

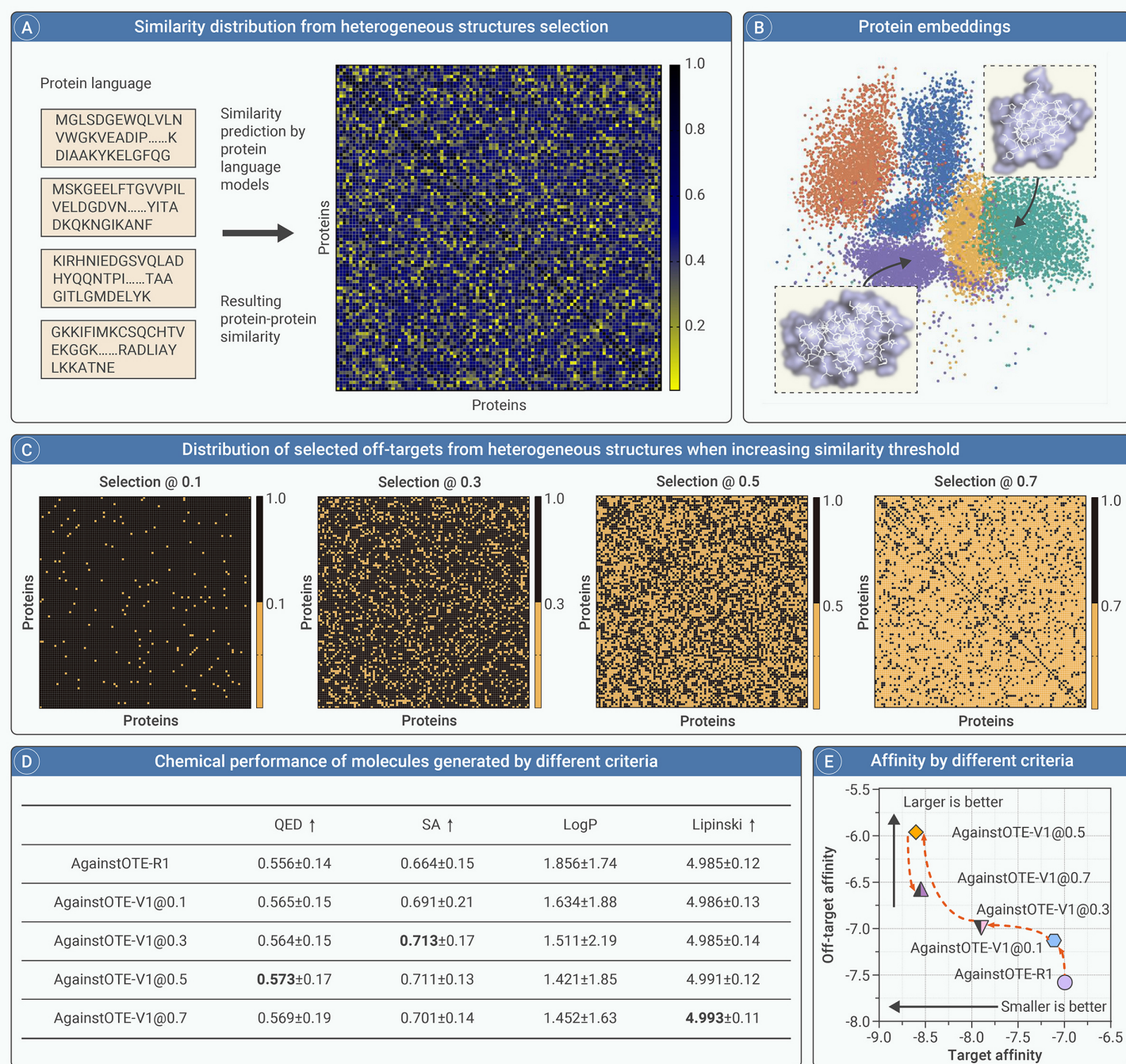


Figure 5. Experiments on selection of off-target proteins by protein language models (A) Similarity matrix from protein pairs is given to show the resulting protein-protein similarity. Each value in the matrix represents the structural similarity between the target protein and query one, where the similarity scores of protein pairs on the diagonal are all one. (B) Protein embeddings of pretrained protein language models are visualized. (C) Distributions of selected off-target receptors at different settings of off-target boundary are given. (D) The chemical performance is calculated for AgainstOTE with different boundary thresholds. (E) The binding ability for changing the boundary thresholds is plotted. Both target and off-target binding affinity are considered.

Overall performance in chemical properties of molecules generated by AgainstOTE

After inferencing novel molecules with trained models given by any query pocket, extensive evaluation was conducted. In Figure 4A, we present the statistical results of several chemical metrics calculated based on molecules generated by existing models including LiGAN,⁶³ AR,²⁶ GraphBP,⁶⁴ Pocket2Mol, FLAG²⁷ and AgainstOTE. The molecules were generated from the 100 proteins in the test set of CrossDocked dataset, and around 100 molecules were produced for each protein. Chemical properties include QED,⁶⁵ SA,⁶⁶ LogP⁶⁷ and Lipinski.⁶⁸ Detailed descriptions of metrics are available in Supplemental Information. Results show that chemical properties of AgainstOTE are optimistic, where highest QED indicating better drug-likeness

than those from existing models, highest Lipinski implying the balance between solubility and permeability better. The partition coefficient related to LogP is in the appropriate range. Another side brought by this pipeline is that synthetic accessibility is struggling from the results in SA, though still acceptable, especially for the finetuned V1 model the synthetic accessibility is highly improved. Numerical numbers are given as in Table S2. Detailed experiments are arranged as below.

Distribution of chemical attributes. In Figure 4B, we display chemical attributes of molecules generated by the baseline Pocket2Mol and ours AgainstOTE. Attributes include number of hydrogen-bond acceptors (HBA), number of hydrogen-bond donors (HBD), molecular weight (MW) and topological polar surface area (TPSA). It can be witnessed that the distribution of

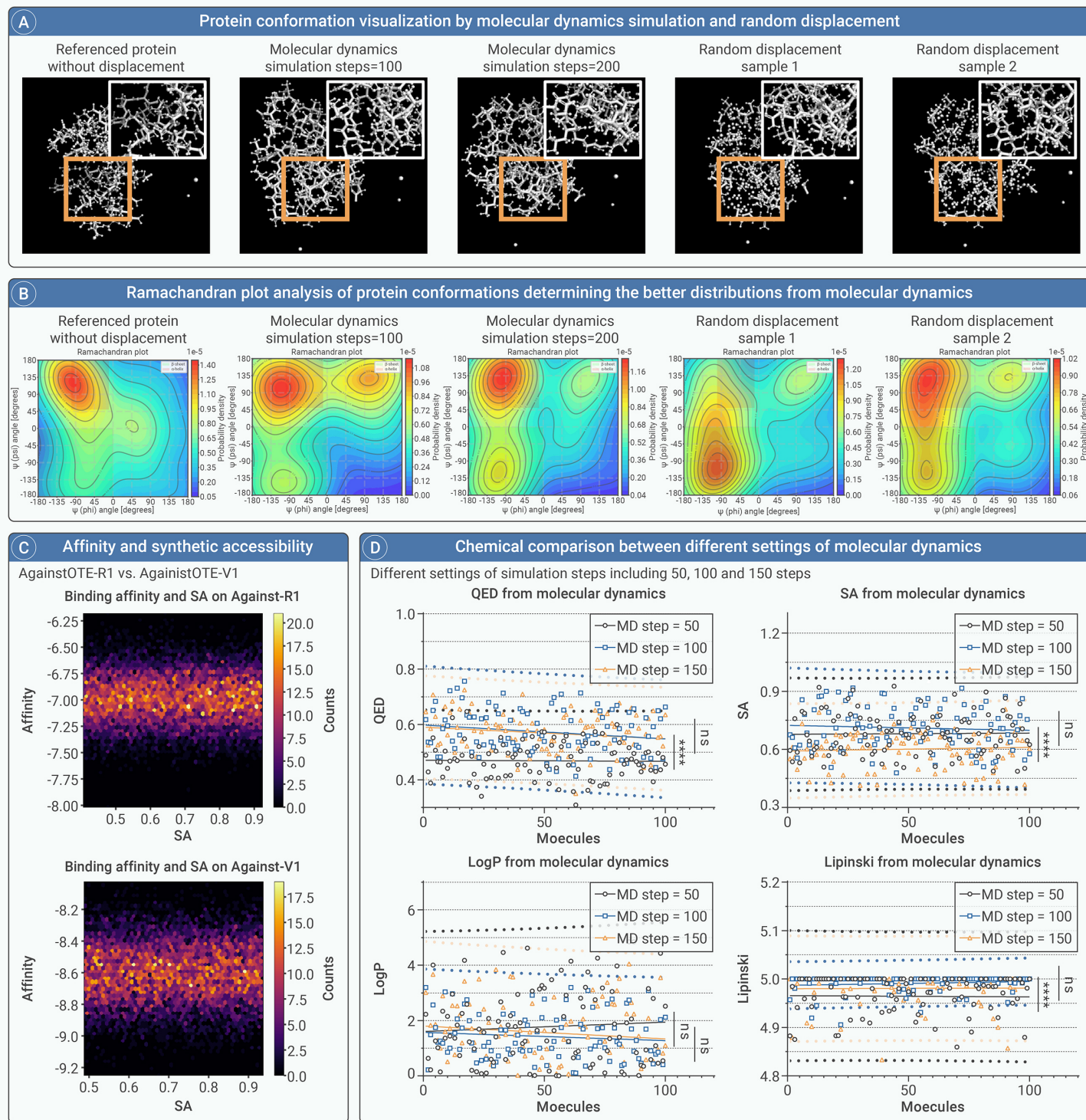


Figure 6. Detailed experiments on structural displacement by molecular dynamics simulation (A) Visualization of protein structures from different time steps after molecular dynamics simulation in AgainstOTE-V1 and structures from random sampling in AgainstOTE-R1. (B) Ramachandran plot analysis on corresponding structures in A is provided. (C) The synthesis accessibility and affinity from R1 version to V1 version are visualized. (D) Chemical performance of different simulation steps is calculated.

two different samples is similar for molecules generated by the baseline in all attributes. However, for both versions of AgainstOTE, distributions of two different samples are far from each other. It illustrates that generated structures for different proteins of AgainstOTE are more discriminative. If structures of the molecules are more similar to the ones belong to other proteins, these molecules are more likely to bind with off-target proteins, causing off-target effects. The distribution from AgainstOTE appears more in line with our expected molecular characteristics against off-target effects. An example targeted at IRAK4 of AgainstOTE is illustrated as in Figure 4C. It can be seen

that in this case the target binding affinity and off-target binding affinity are both improved while the generated structures seem both appropriate.

Influence of N and L from Against OTE-R1 in chemical properties. In Figure 4D, we show the influence of N and L in chemical properties of AgainstOTE-R1 to report the investigation on settings of designed architectures. When $N = 1$, the structural displacement is not working, scores of all metrics are the lowest. This means the structural displacement is necessary in our pipeline, with perturbation noise creating more augmenting samples to simulate the conformational changes of proteins is helpful. On the contrary, if N

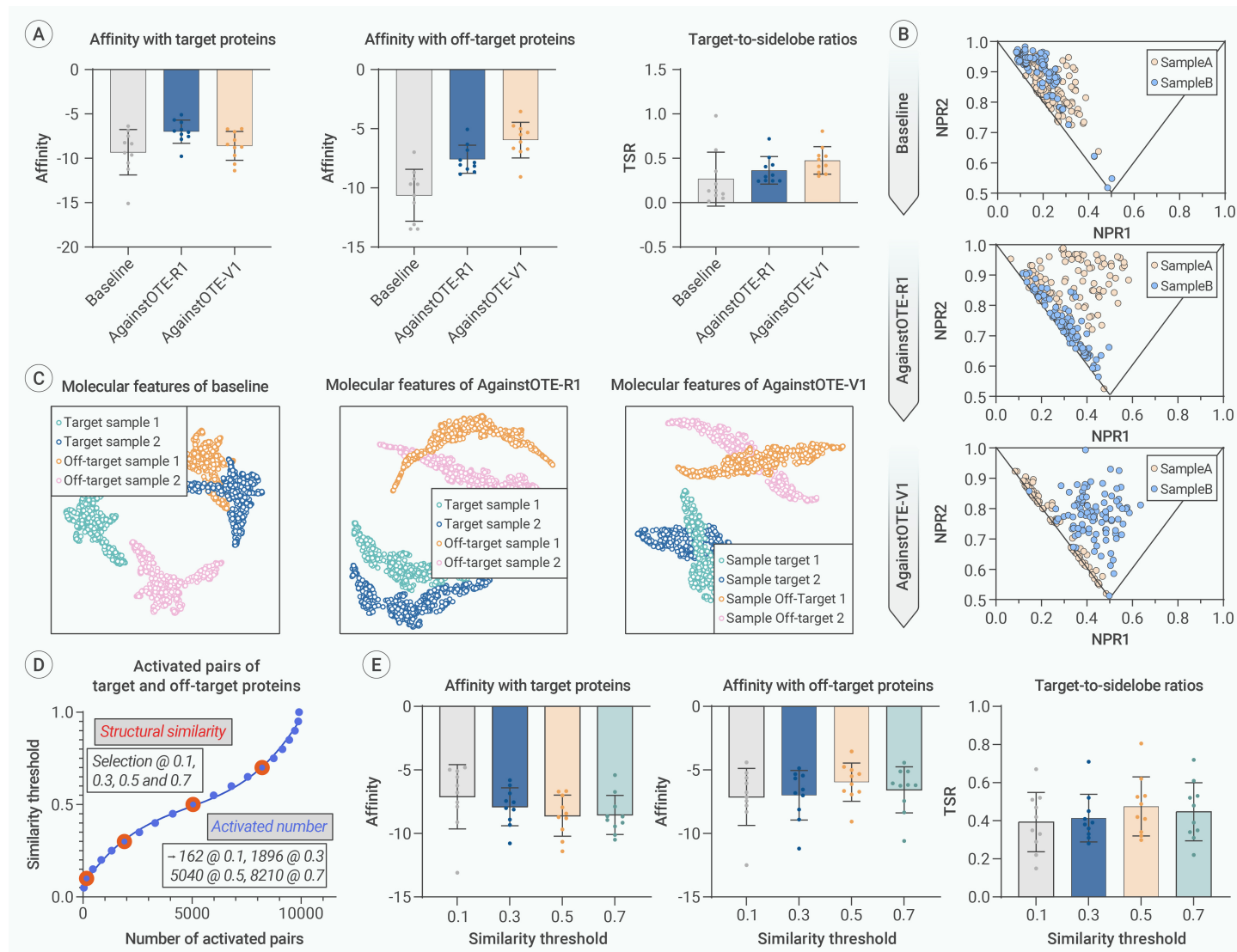


Figure 7. Evaluation of off-target effects (A) Affinity with respectively target proteins, affinity with off-target proteins and target-to-sidelobe ratios of baseline and AgainstOTE are plotted. (B) PMI plots respectively represent features of dimensionality reduction from 3D structures of molecules generated by the baseline and AgainstOTE. From normalized PMI, the distribution of 3D structures can be observed, with the clusters close the 3D structures being similar. (C) Molecular features are visualized for the baseline and ours. There are two target samples and two off-target samples. Distance in the plots implies the similarity in regularized chemical latent space. Considering the structural similarity in off-target selection for test set, (D) numbers of selected pairs of target and off-target proteins are plotted. (E) Evaluation under different structural similarity threshold is presented.

becomes larger, there are more randomly shifted proteins during one time of training. With cost of computation becoming larger, the distribution of these samples is also more difficult. It is observed that $N = 2$ is the best choice. In another side, L indicates stacking how many forward layers in the projector, the nonlinear transforming abilities grow with larger L . Results show that the choice of this parameter should be 2 in R1 version, indicating that the structural regularization being too weak is not a wise choice, yet too strong is also not welcome. The numerical numbers are given as in Table S3.

Diversity and TSR comparison between existing drug models and ours. In Figure 4E, we plot the variations in diversity⁹ of drug models when the ability against off-target effects described by TSR becomes better. Detailed descriptions of diversity are available in Supplemental Information. Results show that AgainstOTE achieves superior performance against off-target effects. With the TSR becomes better, the diversity of generated molecules of our models still show better behavior which may be the reason that modeling structural displacement of proteins is beneficial for generating plausible and diverse results.

Experiments on selection of biochemically meaningful off-target proteins indicate effectiveness

In Figure 5A, proteins were fed forward the pretrained protein language

model³⁴ and Cosine similarity scores were calculated from protein embeddings (Figure 5B). The protein-protein similarity scores on the diagonal are the same ones, so the scores are all one. It can be seen that quite numbers of protein pairs can be selected as candidates for off-target proteins. When settings of off-target boundary thresholds become larger, more and more query proteins are selected as off-target ones for tuning as in Figure 5C. The chemical performance of these different settings in Figure 5D shows the recommended setting of off-target boundary is 0.5 in AgainstOTE-V1. The target and off-target binding affinity also suggest the same settings as in Figure 5E.

Experiments on structural displacement by molecular dynamics simulation

In Figure 6A, 3D structures of the reference protein without displacement, the displaced proteins after different steps of molecular dynamics simulation and the randomly sampled proteins are visualized. In the visual light, these structures appear similar. However, through Ramachandran plot analysis as in Figure 6B, we observe that the proteins following molecular dynamics share similar distributions as the referenced protein, while the distributions from randomly sampled proteins are somewhat different. This indicates that the molecular dynamics simulation in the V1 model is more complied with

physically temporal variations of molecules in nature than random displacement in the R1 version. With respect to this situation, we think the reason about molecular dynamics simulation (V1) appears better than random perturbation (R1) may be in two aspects. One is related to energy minimization processes, where energy minimization in molecular dynamics reduces the system potential energy by optimizing atomic positions, leading to a stable molecular conformation. This avoids the risk resulting the system in high-energy states, e.g., atomic clashes or distorted bond angles which may be caused by random perturbations. Another reason may relate the Newtonian mechanics. Molecular dynamics simulation solves Newton's equations of motion, enabling atoms to temporally evolve on the potential energy surface following physical laws rather than random movement. So the simulation in V1 captures more accurately real molecular behaviors. Meanwhile, the synthesis accessibility and target affinity of AgainstOTE-V1 are also better than the R1 model as in Figure 6C. For investigating the influence of different simulation steps in molecular dynamics, different settings of 50, 100 and 150 steps are experimented as in Figure 6D. Results show that this setting is not sensitive especially when the steps are large enough, where the difference is not evident when steps are 100 and 150.

Evaluation of off-target effects for AgainstOTE

For comprehensively evaluating the ability of models against off-target effects, we designed a systematic measurement called target-to-side ratios (TSR) which highlights the likelihood for a generation model to produce molecules binding to target proteins other than off-target ones, which is inspired from peak-to-sidelobe ratios⁴⁶ in visual tracking field. The specific method of TSR please refer to Methods. In addition to TSR, we also reported affinity with target proteins and affinity with off-target proteins. In this stage, we selected 10 proteins in test set. We use the trained AgainstOTE and the baseline to respectively generated around 100 molecules for each protein. Then docking scores between molecules and all proteins were calculated by QVina2. Affinity to target proteins, affinity to off-target proteins and TSR were plotted in Figure 7A. It can be seen from these plots that TSR of AgainstOTE is much higher than baseline, implying the evident progress against off-target effects. Meanwhile, the affinity with off-target proteins is significantly increased, meaning that the binding possibilities with other proteins are down much. In another side, the affinity with target proteins by AgainstOTE-V1 is decreased at the acceptable range, though this performance from R1 is negative. The detailed numerical data is presented in Table S4.

Analysis of principal moment of inertia. We employed the molecular descriptors of the principal moment of inertia⁶⁹ (PMI) to characterize 3D structures of molecules. As in Figure 7B, by examining the plotted distribution of the normalized PMI, it is observed that generated molecules for different pockets of baseline are clustered in little distance, yet ours are very far from each other. This demonstrates that AgainstOTE series models are capable of producing the structural discriminative molecules for different proteins, avoiding the similar structures of different samples binding to the same protein causing off-target effects.

Distribution of molecular features. We used UMAP⁷⁰ to visualize molecular features embedded by the regularizer. As in Figure 7C, it can be seen that the molecular distances of candidates from target proteins are far from each other yet the distances of molecules from off-target proteins are very close to the baseline. It indicates the baseline is not good at dealing with off-target effects. On the contrary, molecular distances are evenly distributed for AgainstOTE series models, especially the distances between molecules from different proteins are quite far at a proper margin. These results show that molecules generated by ours are more likely to behave against off-target effects.

Distribution of features for proteins. We also used UMAP to visualize features of proteins produced by the trained generator from AgainstOTE. As AgainstOTE is expected to simulate conformations of proteins even when the structures are very similar, as in Figures S3-S4, it shows that the visualized features of proteins are evenly distributed as we expect. Distance between proteins with similar 3D structures is close while maintaining discriminative and the distance between ones with distinctive structures is quite far.

Influence of N and L from AgainstOTE-R1 in evaluation of off-target effects. We also did ablation study of N and L in the angle of being against

off-target effects. The results are available as in Table S5 and Figure S5. For confirming this analysis in candidates from all proteins taken here, we provide the overall performance of molecules generated for each pocket as in Figure S6. From these data, it is observed that TSR and affinity with off-target proteins are improved by a large margin, meanwhile the affinity with target proteins is not in an intense decline in pocket-level results. Additionally, further detailed influence varied with N and L in different pockets is presented as in Figures S7-S16.

Evaluation instructed by structural similarity. Considering the potential bias in this evaluation from off-target selection based on structural similarity, we present results under different similarity threshold. The activated pairs of target and off-target proteins are plotted as in Figure 7D. For pairs resembling proteins in this evaluation setting, binding affinity with target proteins, affinity with off-target ones and TSR are plotted in Figure 7E. These results indicate that structural similarity can influence the evaluation results. On the other hand, results suggest that appropriately introducing proteins with structurally dissimilar structures as off target proteins is beneficial. This may be because off-target effects are widely present, not only in quite similar structural domains, but also in domains with a certain degree of discriminative structures.

Visualization of generated conformations

We selected TBK1_HUMAN with developing potential of avoiding side effects in clinical application and use AgainstOTE series models to generate plausible molecules for these pockets. We visualized molecules at binding sites, 3D structures and chemical formula of generated molecules. In Figure S17, molecules generated are plausible. This indicates that our model has potential to generate different answers to bind with the target, meanwhile less likely to bind with other proteins. More examples are available in Figures S18-S23.

Biological selective evaluation on TBK1/IKKε kinases guided by TSR

To study the quantitative off-target selectivity of AgainstOTE in practical applications, we selected TBK1 (target protein) and its closely related kinase IKKε (off-target protein) as a compelling case, on which we conduct both AI simulation-based evaluation and biological experiments. As in Figure 8A, we provided main procedures especially for biological experiments. We firstly provided quantitative selectivity study as in Figure 8B, including the affinity with both kinases and TSR scores between baseline, AgainstOTE-R1 and AgainstOTE-V1. Results show that our proposed framework helps inhibit off target effects. TSR calculation is conducted on the large-scale bioactive PDBbind⁷¹ database and we report the rankings results sorted by TSR guidance as in Figure 8C, where the higher TSR illustrates better preference towards the target TBK1 and the lower case means the opposite. By chemical synthesis of TSR guided Triclosan with higher TSR preferring TBK1 and Pentoxifylline with lower TSR preferring IKKε, synthesized compounds are used for subsequent biological validation. We then quantitatively measured the binding affinity of the identified compounds toward both TBK1 and IKKε using microscale thermophoresis (MST). Results show that there is higher binding affinity between Triclosan and TBK1 compared to IKKε, while Pentoxifylline exhibits higher binding affinity toward IKKε as in Figure 8D, which is complied with our expectation. Given the greater difference in binding affinity of Pentoxifylline toward TBK1 versus IKKε, we conducted further functional studies on Pentoxifylline. Previous studies have shown that Pentoxifylline inhibits tumor cell proliferation, while IKKε and TBK1 play critical roles in tumor cell proliferation.⁷²⁻⁷⁴ Therefore, we conducted the MTT assay to evaluate the effect of Pentoxifylline on the proliferation of MC38 cells. The results show that Pentoxifylline inhibits the proliferation of MC38 cells at 24 h and 48 h as in Figure 8E. Then we confirmed the engagement between Pentoxifylline and these two kinases using cellular thermal shift assay (CETSA). The data demonstrates that Pentoxifylline significantly enhances the thermal stability of IKKε, maintaining this effect even at higher temperatures, compared to the control group (DMSO). In contrast, Pentoxifylline does not exert a significant enhancing effect on the thermal stability of TBK1 as in Figure 8E.

DISCUSSION

In this section, we discuss several key limitations in this study. (i) From our

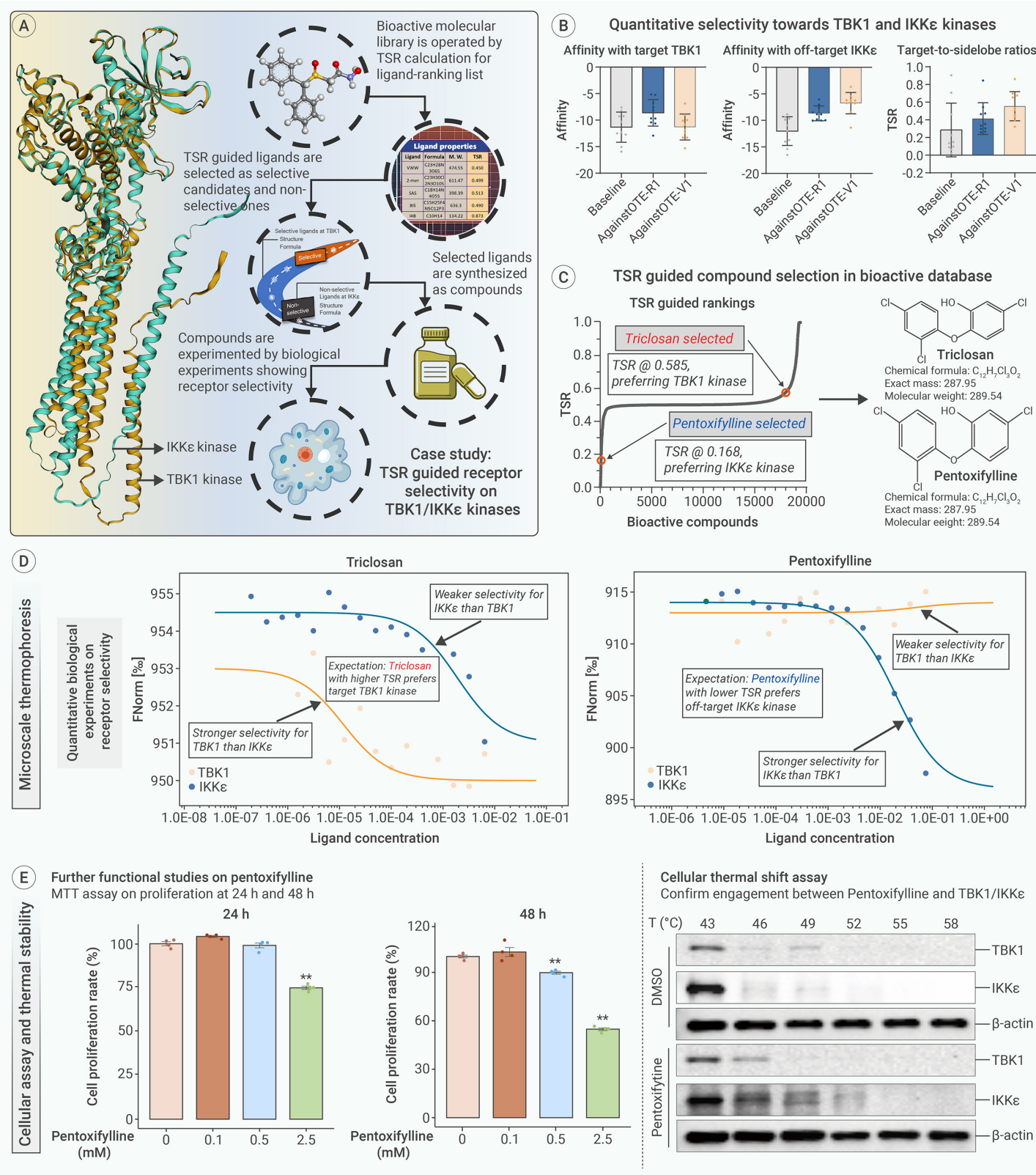


Figure 8. Biological receptor selectivity reveals reduction of risk against off-target effects on target TBK1 and off-target IKKε kinases (A) Main experimental pipeline. This pipeline instructs experimental procedures conducted. (B) Quantitative selectivity towards both kinases is conducted to enhance the evaluation of our framework against off-target effects. We report the TSR scores on PDBbind bioactive compounds towards TBK1 and IKKε kinases as in (C), then we select two compounds respectively prefer TBK1 and IKKε. One is Triclosan which showing higher TSR with the expectation preferring TBK1 and another is Pentoxifylline showing lower TSR with the expectation preferring IKKε. After chemical synthesis, in (D) the interaction of Triclosan and Pentoxifylline with IKKε or TBK1 was measured by microscale thermophoresis (MST). (E) Cell viability of MC38 cell was measured using MTT assay after treating cells with Pentoxifylline (0–2.5 mM) for 24 h or 48 h, *P* values are determined by two-tailed Student's *t*-test. ***p* < 0.01 compared to Pentoxifylline (0 mM). The cellular thermal shift assay (CETSA) was conducted to analyze the thermal stabilization of both kinases at different temperatures in response to Pentoxifylline (2.5 mM) in MC38 cells for 48 h.

experiments, in the optimistic light, most of properties like drug-likeness, solubility, permeability and even the measurement against off-target effects can be positive by AgainstOTE. But in another side, the synthetic availability and the binding affinity with target proteins can also be degraded. This is the most significant trade-off in this study. Although the tuned V1 model improves this problem in some degree, it may potentially weaken desired therapeutic interactions. The main reason leading to this case may be that the hard pockets are often share similar structures as others while behave opposite biologic functions. As a result, the molecules generated with different structures are less likely to interact with the target sites too. Besides, the discriminative molecules with special structures are easy to increase synthesis difficulty.

(ii) For continuous development of a drug product, biological experiments are important for evaluating the molecules. However, unlike biologic verification the chemical synthesis of novel complex molecules is not a mature technology, leading to a critical speed limiting step in the process of developing new drugs. In our future research, the chemical synthesis of novel complex promising molecules will become the most important landmark.

(iii) In practical application, molecules with different water solubility properties may have quite different aggregation behaviors in water. Especially for hydrophobic drugs, they tend to aggregate in water themselves, which may severely affect their interaction with proteins. In the future, we will consider several solutions. (1) Modify the structure of the drug molecules by introducing hydrophilic groups (such as hydroxyl and carboxyl groups) to enhance their interactions with water. (2) Encapsulate the hydrophobic drugs within liposomes or polymeric nanoparticles to form stable nanocarriers. These carriers can protect the drugs from aggregation while improving their bioavailability. (3) Utilize microemulsion or micellar systems to control the release of the drugs, thereby reducing their aggregation in water.⁷⁵ (4) Design environmentally sensitive delivery systems (such as pH-sensitive or temperature-sensitive nanoparticles) that release the drugs under specific conditions.^{76,77} (5) Modify the surface of drug carriers with targeting ligands (such as antibodies or peptides) to increase the concentration of the drugs in target tissues and reduce non-specific aggregation.⁷⁸ (iv) The docking technology implemented by QVina2 in off-target evaluation requires a large amount of computation time, limiting the large-scale full-protein comprehensive experiments. In practical scenes, fast judgement on off-target effects of molecules from a molecular generation system will be in strong demand. This limitation encourages us to develop a more convenient off-target evaluation tool based on other alternative fast docking technologies like training a fast deep learning network to predict the binding energy based on 3D structures of receptor-ligand pairs. Overall, with the rapid advancement of artificial intelligence in the medical and pharmaceutical fields,^{79,80} we believe these limitations will be addressed in the near future.

CONCLUSION

In this study, we focus on off-target effects lied in molecules produced by AI molecular generation models. Preliminary experiments show that molecules generated are obstacle by off-target effects, alarming the risk of uncontrolled binding with proteins other than the target in the biological environment even leading to clinical trial failure. We provide a systematic pretraining framework for de novo molecular generation to alleviate this problem. For pretraining AgainstOTE-R1 in a dual-receptor cooperative training pipeline, both target proteins and randomly selected any other off-target ones are observed by neural networks as input. To simulate the basic conformational changes of proteins, random structural displacement is conducted to create multiple augmenting samples with similar structures. After projecting generated molecules into consistent chemical embedding space by the regularizer, structural regularization term implemented by a molecular contrastive learning architecture is specially designed for regularizing the searching space of the generator by evenly distributing molecular features generated from both target and off-target proteins. By training AgainstOTE-R1 in an end-to-end manner, the trained generator produces plausible compounds meanwhile with lower probabilities binding with other pockets. Meanwhile, an advanced molecular generation model AgainstOTE-V1 is tuned from the R1 model with more district biochemically meaningful off-target proteins selected by protein language models. Additionally, physically molecular

dynamics simulation is utilized in structural displacement instead of random sampling for accessing multiple augmented protein conformations. Extensive experiments are conducted to evaluate various performance of AgainstOTE. In multiple chemical properties including QED, SA, LogP and Lipinski, AgainstOTE behaves positively, showing superior drug-likeness, solubility and permeability. Although synthesis accessibility is decreased in some degree, it is still acceptable especially for the tuned version where this problem is evidently improved. In several chemical attributes including HBA, HBD, MW and TPSA, the distribution of molecules generated by AgainstOTE is more aligned with our intension. Multiple experiments for evaluating the critical designs in AgainstOTE-V1 are conducted. For evaluating off-target effects, we compile affinity to off-target proteins, affinity to target proteins and a designed systematic measurement TSR to evaluate overall off-target effects. Biochemical experiments of receptor selectivity guided by TSR are conducted to directly support the claim regarding improved off-target profiles for practical applications. The statistical results show our efforts against off-target effects are worthwhile. We hope that these results will provide new ideas for target-related pharmacochemical scientists. Experiments show not only the positive side of AgainstOTE, but also the negative side implying off-target effects far from resolved and continuous research becoming necessary.

Through this study, we present some insights and concerns concluded from the experience of the workflow. As off-target effects are widely existed in AI-generated molecules, the chemical searching space may not be as vast as used to be under the joint consideration of both target proteins and off-target proteins. Previously, we try to explore diverse and large molecular libraries for any protein, but considering off-target effects only few numbers of candidates may precisely bind to the target yet not the other proteins. This is the difference brought by this problem. Through AgainstOTE, it can be concluded that observing multiple proteins during one forward-backward calculation, simulating conformational changes and evenly regularizing the projected molecular features may help. But there still lies limitations like difficulties in synthesis and accompanying degraded binding capabilities to the target in some degree. More effective pipelines are expected to further alleviate this problem. Meanwhile, comprehensive systems for evaluating overall performance of AI molecular generation models against off-target effects are needed. In this study we provide a systematic measurement TSR, yet in the future more technical metrics and even the corresponding specialized datasets are expected.

REFERENCES

- Li Y., Zhang L., Wang Y., et al. (2022). Generative deep learning enables the discovery of a potent and selective RIPK1 inhibitor. *Nat. Commun.* **13**:6891. DOI:10.1038/s41467-022-34692-w
- Segler M. H. S., Preuss M. and Waller M. P. (2018). Planning chemical syntheses with deep neural networks and symbolic AI. *Nature* **555**:604–610. DOI:10.1038/nature25978
- Shi H., Doench J. G. and Chi H. (2023). CRISPR screens for functional interrogation of immunity. *Nat. Rev. Immunol.* **23**:363–380. DOI:10.1038/s41577-022-00802-4
- Vamathevan J., Clark D., Czodrowski P., et al. (2019). Applications of machine learning in drug discovery and development. *Nat. Rev. Drug Discov.* **18**:463–477. DOI:10.1038/s41573-019-0024-5
- Moret M., Pachon A. I., Cotos L., et al. (2023). Leveraging molecular structure and bioactivity with chemical language models for de novo drug design. *Nat. Commun.* **14**:114. DOI:10.1038/s41467-022-35692-6
- Nguyen M. N., Sen N., Lin M., et al. (2019). Discovering putative protein targets of small molecules: A Study of the p53 Activator Nutlin. *J. Chem. Inf. Model.* **59**:1529–1546. DOI:10.1021/acs.jcim.8b00762
- Cheng T., Li Q., Zhou Z., et al. (2012). Structure-based virtual screening for drug discovery: A problem-centric review. *AAPS J.* **14**:133–141. DOI:10.1208/s12248-012-9322-0
- Scheiber J., Chen B., Milik M., et al. (2019). Gaining insight into off-target mediated effects of drug candidates with a comprehensive systems chemical biology analysis. *J. Chem. Inf. Model.* **49**:308–317. DOI:10.1021/ci800344p
- Peng X., Luo S., Guan J., et al. (2022). Pocket2Mol: Efficient molecular sampling based on 3D protein pockets. *Proceedings of the 39th International Conference on Machine Learning* **162**:17644–17655. <https://proceedings.mlr.press/v162/peng22b.html>.
- Zhavoronkov A., Ivanenkov Y. A., Aliper A. et al. (2019). Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat. Biotechnol.* **37**:1038–1040. DOI:10.1038/s41587-019-0224-x
- Masuda T., Ragoza M. and Koes D. R. (2020). Generating 3D molecular structures

- conditional on a receptor binding site with deep generative models. *arXiv Preprint arXiv:2010.14442*. DOI:10.48550/arXiv.2010.14442
12. Francoeur P. G., Masuda T., Sunseri J., et al., (2020). Three-dimensional convolutional neural networks and a cross-docked data set for structure-based drug design. *J. Chem. Inf. Model.* **60**:4200–4215. DOI:10.1021/acs.jcim.0c00411
 13. Kotsias P. C., Aris-Pous J., Chen H., et al. (2020). Direct steering of de novo molecular generation with descriptor conditional recurrent neural networks. *Nat. Mach. Intell.* **2**:254–265. DOI:10.1038/s42256-020-0174-5
 14. Segler M. H. S., Kogej T., Tyrchan C., et al. (2018). Generating focused molecule libraries for drug discovery with recurrent neural networks. *ACS Cent. Sci.* **4**:20–131. DOI:10.1021/acscentsci.7b00512
 15. Zhang J. and Chen H. (2022). De novo molecule design using molecular generative models constrained by ligand-protein interactions. *J. Chem. Inf. Model.* **62**:3291–3306. DOI:10.1021/acs.jcim.2c00177
 16. Lim J., Ryu S., Kim J., et al. (2018). Molecular generative model based on conditional variational autoencoder for de novo molecular design. *J. Cheminform.* **10**:31. DOI:10.1186/s13321-018-0286-7
 17. Jin W., Barzilay R. and Jaakkola T. S. (2018). Junction tree variational autoencoder for molecular graph generation. *Proceedings of the 35th International Conference on Machine Learning* **80**:2323–2332. <https://proceedings.mlr.press/v80/jin18a>.
 18. Fu T., Xiao C., Qian C., et al. (2021). Probabilistic and dynamic molecule-disease interaction modeling for drug discovery. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* **pp**:404–414. DOI:10.1145/3447548.3467286
 19. Lam H. Y. I., Pinckert R., Han H., et al. (2023). Application of variational graph encoders as an effective generalist algorithm in computer-aided drug design. *Nat. Mach. Intell.* **5**:754–764. DOI:10.1038/s42256-023-00683-9
 20. Cao N. D. and Kipf T. (2018). MolGAN, an implicit generative model for small molecular graphs. *arXiv Preprint arXiv:1805.11973*. DOI:10.48550/arXiv.1805.11973
 21. Kadurin A., Nikolenko S., Khrabrov K., et al. (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. DOI:10.1021/acs.molpharmaceut.7b00346
 22. Prykhodko O., Johansson S. V., Kotsias P. C., et al. (2019). A de novo molecular generation method using latent vector based generative adversarial network. *J. Cheminfo.* **11**:74. DOI:10.1186/s13321-019-0397-9
 23. Polykovskiy D., Zhebrak A., Vetrov D., et al. (2018). Entangled conditional adversarial autoencoder for de novo drug discovery. *Mol Pharm.* **15**:4398–4405. DOI:10.1021/acs.molpharmaceut.8b00839
 24. Wang Y., Wang J., Cao Z., et al. (2022). Molecular contrastive learning of representations via graph neural networks. *Nat. Mach. Intell.* **4**:279–287. DOI:10.1038/s42256-022-00447-x
 25. Fang Y., Zhang Q., Zhang N., et al. (2023). Knowledge graph-enhanced molecular contrastive learning with functional prompt. *Nat. Mach. Intell.* **5**:542–553. DOI:10.1038/s42256-023-00654-0
 26. Luo S., Guan J., Ma J., et al. (2021). A 3D generative model for structure-based drug design. *Advances in Neural Information Processing Systems* **34**:6229–6239. <https://proceedings.neurips.cc/paper/2021/hash/314450613369e0ee72d0da7f6fee773c-Abstract.html?ref=https://githubhelp.com>.
 27. Zhang Z., Min Y., Zheng S., et al. (2023). Molecule generation for target protein binding with structural motifs. *The Eleventh International Conference on Learning Representations*. <https://openreview.net/forum?id=Rq13idF0F73>.
 28. Tong X., Liu X., Tan X., et al. (2021). Generative models for de novo drug design. *J. Med. Chem.* **64**:14011–14027. DOI:10.1021/acs.jmedchem.1c00927
 29. Sanchez-Lengeling B. and Aspuru-Guzik A. (2018). Inverse molecular design using machine learning: Generative models for matter engineering. *Science* **361**:360–365. DOI:10.1126/science.aat2663
 30. Jiang Y., Zhang G., You J., et al. (2024). PocketFlow is a data-and-knowledge-driven structure-based molecular generative model. *Nat. Mach. Intell.* **6**:326–337. DOI:10.1038/s42256-024-00808-8
 31. Sadybekov A. A., Sadybekov A. V., Liu Y., et al. (2022). Synthon-based ligand discovery in virtual libraries of over 11 billion compounds. *Nature* **601**:452–459. DOI:10.1038/s41586-021-04220-9
 32. Alhossary A., Handoko S. D., Mu Y., et al. (2015). Fast, accurate, and reliable molecular docking with QuickVina 2. *Bioinformatics* **31**:2214–6. DOI:10.1093/bioinformatics/btv082
 33. Jensen K. T., Kao T. C., Tripodi M., et al. (2020). Manifold GPLVMs for discovering non-Euclidean latent structure in neural data. *Advances in Neural Information Processing Systems* **33**. <https://proceedings.neurips.cc/paper/2020/hash/fedc604da8b0f9af74b6cfc0fab2163c-Abstract.html>.
 34. Elnaggar A., Essam H., Salah-Eldin W., et al. (2023). Ankh: Optimized protein language model unlocks general-purpose modelling. *arXiv Preprint arXiv:2301.06568*. DOI:10.48550/arXiv.2301.06568
 35. Gowers R. J., Linke M., Barnoud J., et al. (2019). MDAnalysis: A python package for the rapid analysis of molecular dynamics simulations. *Proceedings of the 15th Python in Science Conference* **pp**:98–105. <https://www.osti.gov/servlets/purl/1565806>.
 36. Eastman P., Swails J., Chodera J. D., et al. (2017). OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.* **13**:e1005659. DOI:10.1371/journal.pcbi.1005659
 37. Satorras V. G., Hoogeboom E. and Welling M. (2021). E(n) equivariant graph neural networks. *Proceedings of the 38th International Conference on Machine Learning* **139**:9323–9332. <https://proceedings.mlr.press/v139/satorras21a.html>.
 38. Wang Y., Shen Y., Chen S., et al. (2023). Learning harmonic molecular representations on riemannian manifold. *The Eleventh International Conference on Learning Representations*. https://openreview.net/forum?id=ySCL-NG_I3.
 39. Stärk H., Ganea O., Pattanaik L., et al. (2022). EquiBind: Geometric deep learning for drug binding structure prediction. *Proceedings of the 39th International Conference on Machine Learning* **162**:20503–20521. <https://proceedings.mlr.press/v162/stark22b>.
 40. Zhang B. and Sennrich R. (2019). Root mean square layer normalization. *Advances in Neural Information Processing Systems* **32**. <https://proceedings.neurips.cc/paper/2019/hash/1e8a19426224ca89e83cef47f1e7f53b-Abstract.html>.
 41. Vaswani A., Shazeer N., Parmar N., et al. (2017). Attention is all you need. *Advances in Neural Information Processing Systems* **30**. <https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>.
 42. Hanin B. and Rolnick D. (2019). Deep ReLU networks have surprisingly few activation patterns. *Advances in Neural Information Processing Systems* **32**. https://proceedings.neurips.cc/paper_files/paper/2019/hash/9766527f2b5d3e95d4a733fcfb77bd7e-Abstract.html.
 43. Chen T., Kornblith S., Norouzi M., et al. (2022). A simple framework for contrastive learning of visual representations. *Proceedings of the 37th International Conference on Machine Learning* **119**:1597–1607. <https://proceedings.mlr.press/v119/chen20j.html>.
 44. Sohn K. (2016). Improved deep metric learning with multi-class n-pair loss objective. *Advances in Neural Information Processing Systems* **29**. <https://proceedings.neurips.cc/paper/2016/hash/6b180037abbeba991d8b1232f8a8ca9-Abstract.html>.
 45. Oord A. V. D., Li Y. and Vinyals O. (2018). Representation learning with contrastive predictive coding. *arXiv Preprint arXiv:1807.03748*. DOI:10.48550/arXiv.1807.03748
 46. Bolme D. S., Beveridge J. R., Draper B. A., et al. (2010). Visual object tracking using adaptive correlation filters. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* **pp**:2544–2550. DOI:10.1109/CVPR.2010.5539960
 47. Skwierawska A. and Pazik A. (2012). Inhibition of impurities formation in the synthesis of N-alkylthiobromines stimulated by microwave irradiation. *Cationic and anionic response of membrane electrodes. J. Incl. Phenom. Macrocycl. Chem.* **74**:145–155. DOI:10.1007/s10847-011-0093-5
 48. Stec J., Fomovska A., Afanador G. A., et al. (2013). Modification of triclosan scaffold in search of improved inhibitors for enoyl-acyl carrier protein (ACP) reductase in *Toxoplasma gondii*. *ChemMedChem* **8**:1138–60. DOI:10.1002/cmdc.201300050
 49. Weininger D. (1988). SMILES, a chemical language and information system. 1. introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **28**:31–36. DOI:10.1021/ci00057a005
 50. O'Boyle N. and Dalke A. (2018). DeepSMILES: An adaptation of SMILES for use in machine-learning of chemical structures. *ChemRxiv*. [chemrxiv:7097960.v1](https://chemrxiv.org/chemrxiv/7097960.v1). DOI:10.26434/chemrxiv.7097960.v1
 51. Liu Y., Jin M., Pan S., et al. (2023). Graph self-supervised learning: A survey. *IEEE Transactions on Knowledge and Data Engineering* **35**:5879–5900. DOI:10.1109/TKDE.2022.3172903
 52. Li Y., Liu L., Shen C., et al. (2017). Mining mid-level visual patterns with deep CNN activations. *International Journal of Computer Vision* **121**:344–364. DOI:10.1007/s11263-016-0945-y
 53. Stretcu O. and Leordeanu M. (2015). Multiple frames matching for object discovery in video. *Proceedings of the British Machine Vision Conference* **pp**:186.1–186.12. DOI:10.5244/C.29.186
 54. Rives A., Goyal S., Meier J., et al. (2020). Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proceedings of the National Academy of Sciences of the United States of America* **118**:e2016239118. DOI:10.1073/pnas.2016239118
 55. Devlin J., Chang M. W. Lee, et al. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* **1**:4171–4186. DOI:10.18653/v1/N19-1423
 56. Liu Y., Ott M., Goyal N., et al. (2019). RoBERTa: A robustly optimized BERT pretraining approach. *arXiv Preprint arXiv:1907.11692*. DOI:10.48550/arXiv.1907.11692
 57. Hsu C., Verkuil R., Liu J., et al. (2022). Learning inverse folding from millions of predicted structures. *Proceedings of the 39th International Conference on Machine Learning* **162**:8946–8970. <https://proceedings.mlr.press/v162/hsu22a.html>.
 58. Lin Z., Akin H., Rao R., et al. (2022). Language models of protein sequences at the scale of evolution enable accurate structure prediction. *bioRxiv Preprint* DOI:10.1101/2022.07.20.500902
 59. Thomson E. C., Rosen L. E., Shepherd J. G., et al. (2021). Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. *Cell* **184**:1171–1187. DOI:10.1016/j.cell.2021.01.037
 60. Hou Y. J., Chiba S., Halfmann P., et al. (2020). SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo. *Science* **370**:1464–1468. DOI:10.

- 1126/science.abe8499
61. Weissman D., Alameh M. G., Silva T., et al. (2021). D614G Spike Mutation Increases SARS-CoV-2 Susceptibility to Neutralization. *Cell Host Microbe*. **29**:23–31. DOI:10.1016/j.chom.2020.11.012
 62. Yurkovetskiy L., Wang X., Pascal K. E., et al. (2020). Structural and functional analysis of the D614G SARS-CoV-2 spike protein variant. *Cell* **183**:739–751. DOI:10.1016/j.cell.2020.09.032
 63. Ragoza M., Masuda T. and Koes D. R. (2022). Generating 3d molecules conditional on receptor binding sites with deep generative models. *Chem. Sci* **13**:2701–2713. DOI:10.1039/d1sc05976a
 64. Liu M., Luo Y., Uchino K., et al. (2022). Generating 3d molecules for target protein binding. *Proceedings of the 39th International Conference on Machine Learning* **162**:13912–13924. <https://proceedings.mlr.press/v162/liu22m.html>.
 65. Bickerton G. R., Paolini G. V., Besnard J., et al. (2012). Quantifying the chemical beauty of drugs. *Nat. Chem.* **4**:90–98. DOI:10.1038/nchem.1243
 66. Ertl P. and Schuffenhauer A. (2009). Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *J. Cheminform.* **1**:8. DOI:10.1186/1758-2946-1-8
 67. Wildman S. A. and Crippen G. M. (1999). Prediction of physicochemical parameters by atomic contributions. *J. Chem. Inf. Comput. Sci.* **39**:868–873. DOI:10.1021/ci990307l
 68. Lipinski C. A., Lombardo F., Dominy B. W., et al. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Del. Rev.* **46**:3–26. DOI:10.1016/s0169-409x(00)00129-0
 69. Jensen R. K. and Nassas G. (1988). Growth of segment principal moments of inertia between four and twenty years. *Med. Sci. Sports. Exerc.* **20**:594–604. DOI:10.1249/00005768-198812000-00013
 70. Becht E., McInnes L., Healy J., et al. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**:38–44. DOI:10.1038/nbt.4314
 71. Wang, R., Fang, X., Lu, Y., et al. (2004). The PDBbind database: Collection of binding affinities for protein-ligand complexes with known three-dimensional structures. *J. Med. Chem.* **47**:2977–2980. DOI:10.1021/jm030580l
 72. Wang Y., Dong L., Li J., et al. (2017). Pentoxifylline induces apoptosis of HepG2 cells by reducing reactive oxygen species production and activating the MAPK signaling. *Life Sci.* **183**:60–68. DOI:10.1016/j.lfs.2017.05.029
 73. Lu J., Yang Y., Guo G., et al. (2017). IKBKE regulates cell proliferation and epithelial-mesenchymal transition of human malignant glioma via the Hippo pathway. *Oncotarget* **8**:49502–49514. DOI:10.18632/oncotarget.17738
 74. Hu L. and Zhang Q. (2024). Mechanism of TBK1 activation in cancer cells. *Cell Insight* **3**:100197. DOI:10.1016/j.cellin.2024.100197
 75. Ait-Touchente Z., Zine N., Jaffrezic-Renault N., et al. (2023). A Exploring the versatility of microemulsions in cutaneous drug delivery: opportunities and challenges. *Nanomaterials* **13**:1688. DOI:10.3390/nano13101688
 76. Zhuo S., Zhang F., Yu J., et al. (2020). pH-sensitive biomaterials for drug delivery. *Molecules* **25**:5649. DOI:10.3390/molecules25235649
 77. Gomte S. S., Agnihotri T. G., Khopade S., et al. (2024). Exploring the potential of pH-sensitive polymers in targeted drug delivery. *J. Biomater. Sci. Polym. Ed.* **35**:228–268. DOI:10.1080/09205063.2023.2279792
 78. Muttenthaler M., King G. F., Adams D. J., et al. (2021). Trends in peptide drug

discovery. *Nat. Rev. Drug Discov.* **20**:309–325. DOI:10.1038/s41573-020-00135-8

79. Huang T., Xu H., Wang H., et al., (2023). Artificial intelligence for medicine: Progress, challenges, and perspectives. *Innov. Med.* **1**:100030. DOI:10.59717/j.xinn-med.2023.100030
80. Dai J., Xu H., Chen T., et al. (2025). Artificial intelligence for medicine 2025: Navigating the endless frontier. *Innov. Med.* **3**:100120. DOI:10.59717/j.xinn-med.2025.100120

FUNDING AND ACKNOWLEDGMENTS

This work was supported by National Key Research and Development Plan (2022YFC3500202), National Natural Science Foundation of China (Nos. 82230116, 82404639), Natural Science Foundation of Jiangsu Province (BK20241249). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

Conceptualization: Yechao Xu, Jiao Qu and Yang Sun. Experimental implementation: Yechao Xu. Methodology: Yechao Xu, Jiao Qu, Yang Sun, Wen Fang and Fen-Er Chen. Investigation and visualization: Yechao Xu, Jiao Qu and Wen Fang. Writing—original draft: Yechao Xu and Jiao Qu. Writing—review & editing: Yechao Xu, Yang Sun, Jiao Qu, Qiang Xu, Wen Fang, Hai-Bin Luo, Fen-Er Chen and Shuai Wang. Data curation: Yechao Xu and Wen Fang. Funding acquisition: Yechao Xu, Jiao Qu, Yang Sun, Hai-Bin Luo, Fen-Er Chen and Shuai Wang. Project administration: Yechao Xu, Jiao Qu, Yang Sun, Qiang Xu and Fen-Er Chen. Resources: Yang Sun, Jiao Qu, Qiang Xu, Fen-Er Chen and Shuai Wang. Software: Yechao Xu, Wen Fang. Supervision: Yechao Xu, Jiao Qu, Yang Sun, Qiang Xu, Hai-Bin Luo, Fen-Er Chen. Validation: Yechao Xu, Jiao Qu, Wen Fang, Yang Sun, Fen-Er Chen and Shuai Wang. All authors contributed to and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

ETHICAL STATEMENT AND PATIENT CONSENT

Not Applicable. This paper does not involve any ethical issues related to humans or animals.

DATA AND CODE AVAILABILITY

The data including training data and testing data used for experiments is released at <https://zenodo.org/records/10642766>. Pre-trained weights of AgainstOTE with the best-ing setting are available from <https://zenodo.org/records/12722246>. The source code and scripts including the production of preliminary experiments, training, testing with AgainstOTE and running experiments can be freely accessed from <https://github.com/miaomiaoToday/AgainstOTE-master>.

SUPPLEMENTAL INFORMATION

It can be found online at <https://doi.org/10.59717/j.xinn-med.2025.100158>