OXFORD

Briefings in Bioinformatics, 2025, **26(4)**, bbae531
https://doi.org/10.1093/bib/bbae531
Problem Solving Protocol

Deep reinforcement learning as an interaction agent to steer fragment-based 3D molecular generation for protein pockets

Xudong Zhang[1], Jing Hou[2], Sanqing Qu[2], Fan Lu[2], Zhixin Tian[3], Alois Knoll[4], Guang Chen[2,*], Shaorong Gao[5,*], Yanping Zhang[5,*]

[1]Shanghai Key Laboratory of Maternal Fetal Medicine, Clinical and Translational Research Center of Shanghai First Maternity and Infant Hospital, School of Computer Science and Technology, Tongji University, Shanghai 200092, China
[2]School of Computer Science and Technology, Tongji University, Shanghai 200092, China
[3]School of Chemical Science and Engineering, Tongji University, Shanghai 200092, China
[4]School of Computation, Information and Technology, Technical University of Munich, 85748, Garching b. München
[5]Shanghai Key Laboratory of Maternal Fetal Medicine, Clinical and Translational Research Center of Shanghai First Maternity and Infant Hospital, Frontier Science Center for Stem Cell Research, School of Life Sciences and Technology, Tongji University, Shanghai 200092, China
*Corresponding authors: Guang Chen. Email: guangchen@tongji.edu.cn; Shaorong Gao. Email: gaoshaorong@tongji.edu.cn; Yanping Zhang. Email: yanpingzhang@tongji.edu.cn

Abstract

Designing high-affinity molecules for protein targets (especially novel protein families) is a crucial yet challenging task in drug discovery. Recently, there has been tremendous progress in structure-based 3D molecular generative models that incorporate structural information of protein pockets. However, the capacity for molecular representation learning and the generalization for capturing interaction patterns need substantial further developments. Here, we propose AMG, a framework that leverages deep reinforcement learning as a pocket–ligand interaction agent (IA) to gradually steer fragment-based 3D molecular generation targeting protein pockets. AMG is trained using a two-stage strategy to capture interaction features and explicitly optimize the IA. The framework also introduces a pair of separate encoders for pockets and ligands, coupled with a dedicated pre-training strategy. This enables AMG to enhance its generalization ability by leveraging a vast repository of undocked pockets and molecules, thus mitigating the constraints posed by the limited quantity and quality of available datasets. Extensive evaluations demonstrate that AMG significantly outperforms five state-of-the-art baselines in affinity performance while maintaining proper drug-likeness properties. Furthermore, visual analysis confirms the superiority of AMG at capturing 3D molecular geometrical features and interaction patterns within pocket–ligand complexes, indicating its considerable promise for various structure-based downstream tasks.

Keywords: structure-based drug design; molecular generation; deep reinforcement learning; interaction modeling

Introduction

Numerous diseases across the globe continue to elude effective treatment. Traditional methodologies employed in drug discovery, such as high-throughput screening, are often beset by drawbacks, including lengthy time frames, substantial financial costs, and notably high rates of failure [1, 2]. In response, structure-based virtual screening (SBVS), a computer-aided drug design method, computationally screens large compound libraries to identify molecules complementary to target proteins and experimentally validates those predicted to bind well [3–5], thereby facilitating the discovery of new ligands for specific targets. However, the molecular libraries available for screening represent but a small fraction of the chemical space [6], with many contained molecules exhibiting undesirable physicochemical properties and low affinity [7]. This constraint often leads to a form of blind screening, where the full potential of the chemical space remains untapped.

In light of the recent strides made in developing deep generative models [8–12] boasting formidable capabilities in molecular generation, there is a promising avenue for establishing dependable molecular libraries through this approach. Most molecular generative models could be classified into ligand-based molecular generative (LBMG) models [12–16] and structure-based molecular generative (SBMG) models [17–23]. Based on the principle that molecules with similar structures tend to exhibit similar properties, LBMG models take a set of molecules with the same targets to fine-tune a model pre-trained on a large-scale dataset, enabling the molecular generation against specific targets [24, 25]. Despite their efficacy for well-characterized proteins, the applicability of LBMG models is still limited for new targets or those with few known ligands [26, 27]. In contrast, SBMG models excel at generating high-affinity molecules for unseen targets by learning the 3D structures of protein–ligand complexes [28–30], offering a promising avenue for de novo drug design.

Current state-of-the-art SBMG models, such as atom-based generation [19–21], have attracted growing interest; however, efficient utilization of chemical rules and empirical knowledge for extensive sampling and generation of realistic sub-structures needs further improvement. DecompDiff [22] and FLAG [23] adopt a fragment-based generation that incorporates chemical priors, significantly mitigates the generation of unrealistic substructures, and enhances the sampling efficiency; however, the sole

Received: July 23, 2024. **Revised:** September 11, 2024. **Accepted:** October 1, 2024

consideration of fragment-level interactions and lack of explicit interaction modeling led to poor generalization in generating high-affinity molecules. Moreover, reliance on a single pocket–ligand complex encoder restricts the potential for pre-training (PT) using the extensive chemical space available, thus limiting the models' ability to leverage vast datasets for enhancement of generalization ability and performance. Typically, these models [19–23] are trained on datasets such as the CrossDocked dataset [31], which features 22.5 million poses of ligands cross-docked into multiple similar binding pockets sourced from the Protein Data Bank (PDB) [32]. Despite its size, the CrossDocked dataset includes only 2,922 pockets and 13,839 ligands. This relatively narrow dataset scope severely limits the models' generalization capabilities, presenting a challenge for the development of SBMG models that can generate high-affinity molecules across a broader spectrum of protein targets.

To address these limitations, we propose AMG, a framework that leverages deep reinforcement learning (DRL) as a pocket–ligand interaction **A**gent to gradually steer fragment-based 3D **M**olecular **G**eneration targeting protein pockets. Specifically, AMG incorporates a two-stage training strategy to capture interaction features explicitly and optimize the molecular generation process. In the first stage (Fig. 1A), the model is trained on the CrossDocked dataset to generate high-affinity molecules for pockets. AMG employs a Physics-Informed Interaction Modeling (PIIM) module to predict and embed atom-level interaction energies. Thereafter, in the second stage (Fig. 1B), a DRL-based interaction agent (IA) is integrated with the trained model, taking interaction energies as actions to steer the generation process at the feature level. The generation process (Fig. 1C) proceeds in a fragment-by-fragment manner, leveraging substructures with prior knowledge to avoid generating unrealistic substructures. Moreover, a novel PT strategy based on contrastive learning and molecular reconstruction is introduced to enhance the generalization ability of AMG (Fig. 1D). Extensive experiments conducted on the CrossDocked dataset and real-world therapeutic targets demonstrate that AMG significantly outperforms representative state-of-the-art baselines in affinity performance and surpasses another fragment-based method in terms of drug-likeness properties. Additionally, visual analysis reveals that AMG performs excellently in capturing the 3D molecular geometric features and interaction patterns of pocket–ligand complexes.

Materials and methods

Dataset usage

To pre-train the ligand encoder and decoder, the Collection of Open Natural Products (COCONUT) dataset [33] was introduced. This dataset was sourced from 53 different data sources and included additional molecules collected from literature sets. It comprises 406 076 unique molecules (without stereochemistry). After filtering out molecules that could not be transformed and sanitized using RDKit, 404 397 molecules with 3D structures were saved in SDF files. The natural products in the COCONUT dataset, having undergone a long process of natural selection, often possess unique bioactivities. Their structural diversity and complexity far surpass many synthetic small molecules, enabling them to interact with a variety of biomolecules, thus making them a significant resource in drug discovery. For PT the pocket encoder, 211 377 proteins with 3D structures were downloaded from the PDB database [32]. Possible protein binding pockets were identified and extracted using Fpocket [34], then filtered based on pocket score, druggability, and volume. Finally, all pockets with 3D structures were saved in PDB format, resulting in the Pocket3D dataset with 277 530 pockets. The CrossDocked dataset [31], containing 22.5 million pocket–ligand complexes, was used for fine-tuning. Following FLAG [23], molecules with a binding pose RMSD greater than 1Å or those that RDKit could not sanitize were excluded. Then, mmseqs2 [35] was used to cluster the data at 30% sequence identity. Finally, 100 000 pocket–ligand pairs were randomly selected for training, and the test set included 100 pockets from the remaining clusters.

Problem formulation

In contrast to previous methods [19–21, 23] that use a single pocket–ligand complex encoder, we design two separate encoders to independently capture the features of pockets and ligands. The 3D geometries of both pockets and ligands are represented by 3D graphs. A protein pocket is represented as $\mathcal{P} = \{(a_i, b_i, s_i)\}_{i=0}^{m}$, where m is the number of atoms in the pocket, a_i is the type of atom represented by a one-hot vector, b_i is the bonding relationship, and $s_i \in \mathcal{R}^3$ represents the 3D Cartesian coordinates. Similarly, a ligand is represented as $\mathcal{L} = \{(u_i, v_i, r_i)\}_{i=0}^{n}$. Atom-level interactions, defined as I, are described in detail in Section 2.3.2. We formulate the generative process for 3D molecular generation as an autoregressive model conditioned on 3D protein pockets and atom-level interactions. The generative model, denoted as ϕ, learns the conditional probability distribution $p(\mathcal{L}|\mathcal{P}, I)$. The 3D molecular graphs are generated in a fragment-by-fragment manner. A fragment is defined as $F_i = (\mathcal{U}i, \mathcal{V}i, \mathcal{R}i)$, where $\mathcal{U}i$ is the set of atoms, $\mathcal{V}i$ is the set of covalent bonds, and $\mathcal{R}i$ is the set of atom coordinates. The generation process is defined as follows:

$$(\mathcal{U}_t, \mathcal{V}_t, \mathcal{R}_t) = \phi(\mathcal{U}_{t-1}, \mathcal{V}_{t-1}, \mathcal{R}_{t-1}, \mathcal{P}, I), \quad t > 1 \tag{1}$$

$$(\mathcal{U}_1, \mathcal{V}_1, \mathcal{R}_1) = \phi(\mathcal{P}, I), \quad t = 1. \tag{2}$$

when $t = 1$, ϕ generates fragment $F_1 = (\mathcal{U}_1, \mathcal{V}_1, \mathcal{R}_1)$ conditioned on pocket $\mathcal{P}$ and interaction I; when $t > 1$, ϕ generates fragment $F_t = (\mathcal{U}_t, \mathcal{V}_t, \mathcal{R}_t)$ conditioned on $\mathcal{P}$, I, and the currently partially generated molecule $F_{t-1} = (\mathcal{U}_{t-1}, \mathcal{V}_{t-1}, \mathcal{R}_{t-1})$.

AMG architecture

Separate pocket and ligand encoders

The global positional relationship of complexes can be effectively captured using a single pocket–ligand complex encoder, which facilitates the placement of generated molecules in appropriate locations within pockets. However, the limited quantity and quality of available data can hinder training, resulting in generated molecules with undesirable properties. Therefore, two separate encoders, one for pockets and one for ligands, are introduced to enable PT with extensive and diverse molecules. The first layer in the pocket and ligand encoders is a linear layer that encodes the atomic attributes of each atom i in pockets and each atom j in ligands into initial embeddings $P_i^{(0)}$ and $L_j^{(0)}$, respectively. The edge embedding e_{ij} is calculated by encoding pairwise distances using Gaussian functions [36]. A 3D graph transformer composed of L transformer layers [37] serves as the backbone. The scaled dot-product attention mechanism in the multi-head self-attention of the lth layer ($1 \le l \le L$) is formulated as follows:

$$head_{P_i}^{m} = \sum_{u \in \mathcal{N}(i)} \text{Softmax}\left(\frac{(W_Q P_i^{(l)})^\top \cdot W_K r_{iu}^{(l)}}{\sqrt{d}}\right) W_V r_{iu}^{(l)} \tag{3}$$

$$head_{L_j}^{m} = \sum_{u \in \mathcal{N}(j)} \text{Softmax}\left(\frac{(W_Q L_j^{(l)})^\top \cdot W_K r_{ju}^{(l)}}{\sqrt{d}}\right) W_V r_{ju}^{(l)} \tag{4}$$

where $head_{P_i}^{m}$ and $head_{L_j}^{m}$ denote the mth attention head, $\mathcal{N}(i)$ and $\mathcal{N}(j)$ are neighbor sets of atoms i and j, $P_i^{(l)}$ and $L_j^{(l)}$ are atom

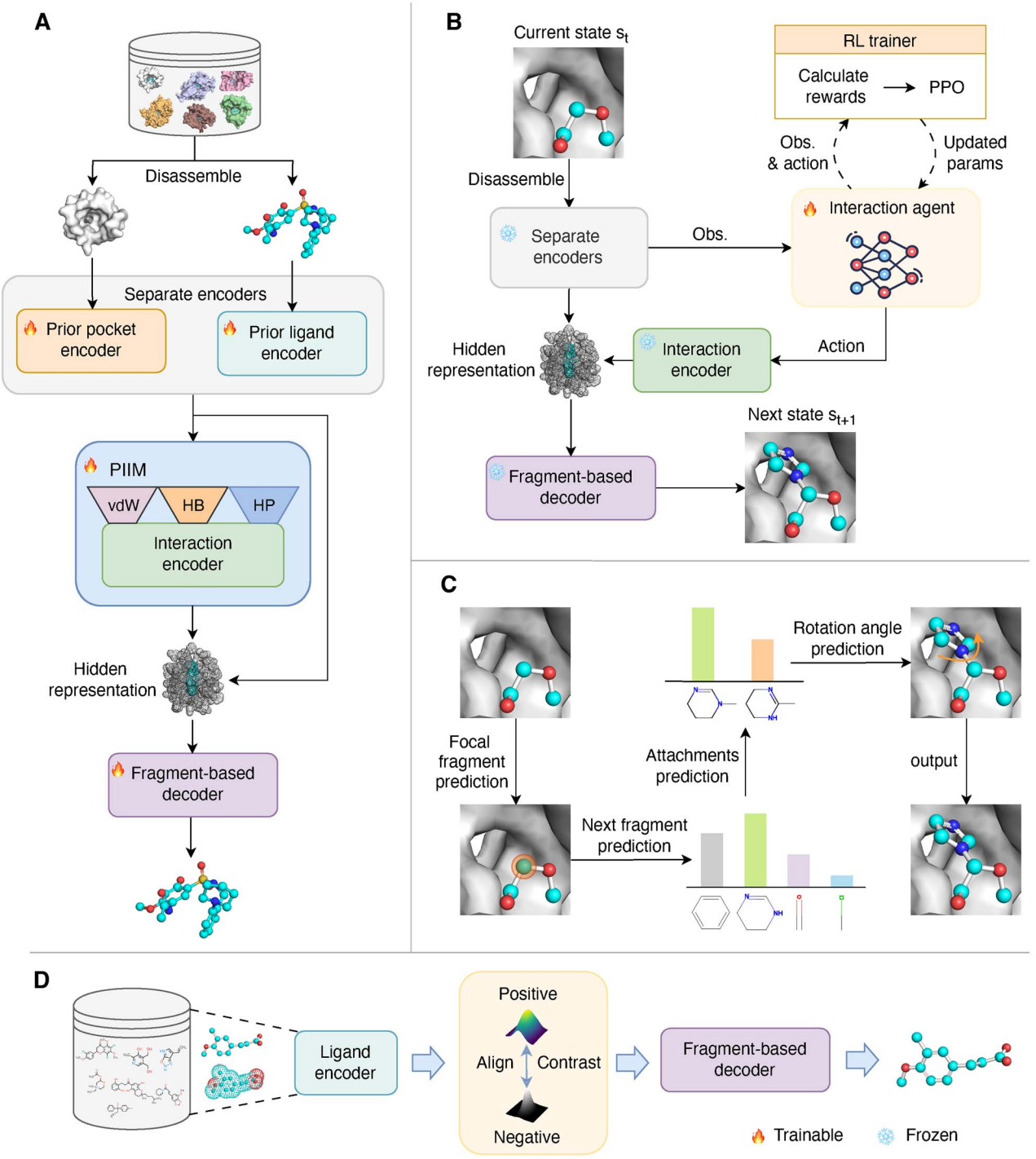

Figure 1. Overview of the AMG framework. (A) The first training stage of AMG. A pocket–ligand complex is disassembled into individual structures to obtain features. These features are fed into the PIIM module to predict atom-atom pair interaction energies. The interaction energies are encoded and fused with pocket and ligand features for molecular reconstruction. (B) The second training stage of AMG. A partially generated molecule s_t is processed by encoders with frozen parameters to generate embedding features. An RL-based IA takes these features to perform actions aimed at maximizing the reward, with optimization achieved through the PPO strategy. The pocket feature (when $t = 1$) or molecular feature (when $t > 1$) is fused with the interaction feature and fed into the decoder to generate the next molecular state s_{t+1}. During inference, the RL trainer is not used, and the IA's parameters are frozen. (C) An illustration of one generation step of the fragment-based decoder. (D) The PT strategy for the ligand encoder and fragment-based decoder. Slight noise perturbations are added to atomic coordinates to create original and augmented views, facilitating ligand encoder PT. The decoder is pre-trained by reconstructing masked molecules.

embeddings. For the relational information, $r_{iu}^{(l)} = \mathbf{Concat}(P_i^{(l)}, e_{iu}^{(l)})$ and $r_{ju}^{(l)} = \mathbf{Concat}(L_j^{(l)}, e_{ju}^{(l)})$. W_Q, W_K, W_V are learnable transformation matrices and d is the dimension size of embeddings. All attention heads are concatenated and further transformed to obtain the final output.

PIIM module

Current SBMG models primarily generate molecules by feeding encoded complex features into a decoder, neglecting the essential interaction information between pockets and ligands. Such neglect typically leads to the generated molecules that only

approximate the spatial structure of binding pockets without establishing stable interactions. Inspired by docking score functions [38, 39] and PIGNet [40], we conceptualize binding affinity as the summation of three interaction types: van der Waals interaction (E_{vdW}), hydrogen bond interactions (E_{HB}), and hydrophobic interaction (E_{HP}). In contrast to PIGNet, which is limited to 1D protein and molecule sequence inputs, our proposed PIIM module predicts atom-level energy components within a 3D structural level, thereby informing the 3D molecule generation process. This module is trained by minimizing an auxiliary mean squared error (MSE) loss, as detailed in Section 2.4.

Specifically, the vdW interaction term is defined as the sum of all potential atom-atom pairwise vdW energy contributions and calculated using the 12-6 Lennard-Jones potential. The vdW energy E_{vdW} is formulated as follows:

$$E_{vdW} = \sum_{i,j} \mathbf{MLP}_{vdW}(C)_{i,j} \left[\left(\frac{d'_{i,j}}{d_{i,j}} \right)^{12} - 2 \left(\frac{d'_{i,j}}{d_{i,j}} \right)^{6} \right] \tag{5}$$

where C denotes a complex feature formed by concatenating encoded pocket and ligand information. The function $\mathbf{MLP}_{vdW}(\cdot)$ takes C as input and predicts the minimum vdW interaction energy, thereby aligning each estimated energy component with its actual counterpart.

The calculations for hydrogen bond interactions E_{HB} and hydrophobic interaction E_{HP} follow a same formula:

$$e_{i,j} = \begin{cases} C & \text{if } d_{i,j} - d'_{i,j} < a_1, \\ C\left(\frac{d_{i,j} - d'_{i,j} - a_2}{a_1 - a_2} \right) & \text{if } a_1 < d_{i,j} - d'_{i,j} < a_2, \\ 0 & \text{if } d_{i,j} - d'_{i,j} > a_2 \end{cases} \tag{6}$$

where $e_{i,j}$ represents the corresponding atom-atom pairwise energy contribution coefficients. The coefficients a_1 and a_2 are predefined, set as 0.7 and 0.0 for hydrogen bond interactions and 0.5 and 1.5 for hydrophobic interaction. E_{HB} and E_{HP} are calculated as the sum of $e_{i,j}$. Subsequently, an auxiliary MLP serves as the interaction encoder, extracting high-dimensional interaction features I:

$$I = \mathbf{MLP}_{inter}(E_{vdW}, E_{HB}, E_{HP}) \tag{7}$$

Finally, a feature fusion function, denoted as $\mathcal{F}$, is utilized to derive the ultimate hidden representation Z of a complex:

$$Z = \mathcal{F}(P, L, I) \tag{8}$$

where the pocket feature P, ligand feature L, and interaction feature I are integrated to form a hidden representation.

DRL-based IA

To mitigate the generation of unrealistic substructures, we adopt a fragment-based generation approach. This approach features a decoder that performs the following steps: focal fragment selection, next fragment prediction, fragment attachments prediction, and rotation angle prediction (Fig. 1C). Detailed descriptions of the decoder are provided in Supplementary Materials. Additionally, this generation approach focuses on fragment-level interactions, which complements the atom-level interaction energies mentioned above, thereby improving the affinity of the generated molecules. Specifically, we propose a pocket–ligand IA based on DRL, treating these interaction energies as an action. This enables AMG to explore optimal affinity and drug-likeness properties in a step-wise manner, ultimately achieving global optimization.

We formulate a Markov Decision Process (MDP) through the application of sequential fragment-based generation and utilize the actor-critic [41, 42] pipeline in RL for molecule design. The actor network takes the current state s_t ($1 \le t \le T$) of the system as input and executes the corresponding action a_t to maximize the reward. The critic network evaluates the actor's performance by considering a_t, s_t, and the subsequent state (s_{t+1}) from taking that action. The action a_t integrates three interaction energy components: E_{vdW}, E_{HB}, and E_{HP}, with integer values ranging from -10 to 0 indicating the intensity of each energy component. The observation state s_t represents an intermediate state at step t:

$$s_t = \mathcal{D}(\mathcal{F}(P, L_{t-1}, \mathbf{MLP}_{inter}(a_{t-1}))) \tag{9}$$

where P denotes the pocket feature, L_{t-1} denotes the feature of partially generated molecule at step $t-1$, a_{t-1} is the action taken by the agent at step $t-1$. $\mathbf{MLP}_{inter}$ takes a_{t-1} as input and outputs a high-dimensional interaction feature. $\mathcal{D}$ signifies the fragment-based decoder that transforms a hidden representation into a molecule.

To achieve multi-objective molecular property optimization, we construct a reward function R, which incorporates the Vina Score, Quantitative Estimate of Drug-likeness (QED), SA, and Lipinski. The reward function is defined as follows:

$$R(s_t) = \alpha \cdot r_v(s_t) + \beta \cdot r_q(s_t) + \gamma \cdot r_s(s_t) + \delta \cdot r_l(s_t) \tag{10}$$

where r_v, r_q, r_s, r_l represent the normalized scores corresponding to each property, weighted by their respective coefficients α, β, γ, and δ. These weights are predetermined as 1, 0.9, 0.9, and 0.9, respectively, to balance their contribution to the overall reward. Since intermediate molecules may not be valid, the reward for all intermediate states s_t, where $t < T$ is set to 0 [43].

Training protocol

PT strategy

The SBMG models, functioning as conditional generative models, are tasked with learning features of both ligands and proteins. Their generation process presents challenges in training due to its intricate design. To address this, we propose a Pre-training strategy based on contrastive learning [44–46] and molecular reconstruction. Inspired by the fact that 3D geometric information or atomic coordinates affect the position and affinity between the generated molecules and the pockets [44], we consider geometrical perturbations that add slight noise to atomic coordinates, enhancing the model's learning ability for atomic coordinates. Specifically, two views, the original geometry and the augmented geometry, are constructed, and an energy-based model with noise contrastive estimation (EBM-NCE) [45] is employed to align the two views above, facilitating PT of the pocket and ligand encoder. For the PT of the decoder, we utilize a molecular reconstruction task. Similar to FLAG [23], we first extract fragments from molecules, then randomly mask some fragments and train the model to recover the masked ones. The losses for the PT of the ligand encoder and fragment-based decoder are as follows:

$$\mathcal{L}_{ligand} = \mathcal{L}_{ebmnce} + \mathcal{L}_{frag} + \mathcal{L}_{focal} + \mathcal{L}_{attach} + \mathcal{L}_{\alpha} \tag{11}$$

where $\mathcal{L}_{ebmnce}$ represents the alignment loss of the two views via EBM-NCE and also serves as the PT loss for the pocket encoder. $\mathcal{L}_{frag}$ denotes the cross-entropy loss for fragment type prediction. $\mathcal{L}_{focal}$ is the binary cross-entropy loss for predicting focal atom-/fragment. $\mathcal{L}_{attach}$ represents the loss for fragment attachment prediction, which aims to maximize the log-likelihood of accurately predicting the molecular graph. $\mathcal{L}_{\alpha}$ is the torsion angle prediction loss, fitting angles with the von Mises distribution [47].

The first training stage of AMG

To capture atom-level interaction information and generate molecules with high affinity, we trained our model on the CrossDocked dataset. The loss consists of the total interaction energy prediction loss, multi-type interaction energy reconstruction loss and molecular reconstruction loss, formulated as follows:

$$\mathcal{L} = \mathcal{L}_{E} + \mathcal{L}_{IR} + \mathcal{L}_{frag} + \mathcal{L}_{focal} + \mathcal{L}_{attach} + \mathcal{L}_{d} + \mathcal{L}_{\alpha} \tag{12}$$

$$E_{total} = \frac{E_{vdW} + E_{HB} + E_{HP}}{T_{rotor}} \tag{13}$$

where L_E denotes the MSE loss for the prediction of total interaction energy, with E_{total} defined in Equation (13) and T_{rotor} representing the rotor penalty. L_{IR} is an auxiliary loss for reconstructing the predicted multi-type interaction energies, which guides the training of the interaction encoder. The molecular reconstruction loss, similar to the task during the PT phase, is composed of L_{frag}, L_{focal}, L_{attach}, L_d, L_{α}. Here L_d is the loss associated with the prediction of distances between atom pairs, contributing to more precise coordinate predictions.

The second training stage of AMG

The proposed IA is built on the actor-critic architecture. As one of the most prominent architectures within reinforcement learning, it excels at searching optimal strategies through low-variance gradient estimates, rendering it highly applicable in various real-world scenarios [48]. However, in molecular generation tasks, our aim extends beyond the mere exploration of new chemical spaces; crucially, we seek to generate molecules that are realistically meaningful. In essence, while we encourage the model to explore new structures, we also expect it to retain, to a degree, the prior knowledge acquired via self-supervised training. To achieve this, we employ the Proximal Policy Optimization (PPO) [49] algorithm for implementation, which is specifically designed to update policies in a manner that enhances learning stability and efficiency. PPO introduces an objective function that constrains the disparity between new and old policies, thereby preventing excessive alterations during policy updates. This approach mitigates training instability, facilitates smoother policy updates, and consequently improves learning outcomes in complex environments.

Consequently, the second stage of training aims to determine the optimal policy within the specified MDP using PPO. Initially, the generator model is parameterized by θ and represented as the policy $G_{\theta}(s_{t+1}|s_t, a_t)$, which controls the transition from the current state-action pair (s_t, a_t) to the next state s_{t+1}. The updates to the parameters are guided by the advantage function, defined as follows:

$$A^{G}(s,a) = Q^{G}(s,a) - V^{G}(s) \tag{14}$$

where $Q^{G}(s,a)$ is the action-value function that evaluates the expected return of taking a certain action a in a given state s. $V^{G}(s)$ denotes the state value function that estimates the expected return of being in state s.

PPO proceeds to update the parameters within a trust region to prevent drastic policy changes during training, as given by the objective function:

$$J(\theta) = E_{G_{\theta_c}}\left[\min\left(\frac{G_{\theta}(a|s)}{G_{\theta_c}(a|s)}A^{G_{\theta_c}}, g(\varepsilon, A^{G_{\theta_c}})\right)\right] \tag{15}$$

$$g(\epsilon, A) = \begin{cases} (1+\epsilon)A & A \geq 0 \\ (1-\epsilon)A & A < 0 \end{cases} \tag{16}$$

where θ_c represents the current policy parameters, and ϵ is a hyperparameter that controls the size of the trust region. The function $g(\epsilon, A)$ ensures that the policy updates are kept within a specified range relative to the old policy, effectively regularizing the parameter updates and maintaining stable learning progress [50].

Evaluation metrics

The evaluation of the generated molecules encompasses various aspects including their binding affinity, drug-likeness properties, structural properties, and the efficiency of the sampling process. **Affinity metrics**: Vina Score measures the binding affinity between generated molecules and protein targets, calculated using Qvina [51, 52], with lower scores indicating stronger predicted binding. High Affinity (HA) represents the percentage of generated molecules that exhibit higher affinity compared to reference molecules in the test set. Mean Percentage Binding Gap (MPBG) [53] measures the average performance of the model in terms of binding affinity. A smaller MPBG indicates a higher average binding affinity. **Drug-likeness properties:** QED calculated by RDKit [54] represents the likelihood of a molecule being a viable drug candidate. Synthetic Accessibility (SA) score measures the complexity of synthesizing a drug, with values normalized between 0 and 1. Higher values suggest easier synthesis. Lipinski (Lip.) counts the number of Lipinski's Rule of Five [55] criteria met by a molecule, which are guidelines for drug-likeness. **Structural properties:** Finger.sim. is the Tanimoto similarity [56], measuring how closely the molecular fingerprints of generated molecules match those in the training set. Frag.sim. is computed by MOSES [10], this denotes the cosine similarity between the distributions of BRICS fragments [57] of the generated molecules and those in the training set. Diversity (Div.) quantifies the structural diversity of generated molecules for a given binding pocket, calculated as 1 − average pairwise Tanimoto similarities. root-mean-square deviation (r.m.s.d.) is used to quantify the structural quality of generated conformations [20]. **Sampling efficiency:** time measures the duration required to generate 100 valid molecules for a specific binding pocket.

Results

To rigorously verify the efficacy of the proposed molecular generative model, we conduct extensive experiments across five critical aspects. These experiments are designed to thoroughly examine the merits and capabilities of AMG in generative viable molecular candidates for therapeutic applications. The evaluation strategies are as follows: (1) Properties evaluation of generated molecules. (2) Quality analysis of molecular structure. (3) Structural similarity and diversity analysis. (4) Ablation study. (5) Interaction patterns analysis on real-world therapeutic targets.

Table 1. The top 1/3/5/10 molecules (sorted by Vina Score) mean binding energies and drug-likeness properties on the test set. The fragment-by-fragment methods are denoted with ⋆. The best and suboptimal results are labeled with bold and underlined

	Vina Score (↓)	HA (↑)	MPBG (↓)	QED (↑)	SA (↑)	Lipinski (↑)
			Top 1			
Testset	−6.986	–	–	0.468	0.707	3.908
Pocket2mol	−8.785	88.0%	−0.95%	0.517	0.778	**4.950**
TargetDiff	−9.278	94.0%	−1.22%	0.426	0.476	4.420
DecompDiff	−8.485	95.0%	−0.99%	0.552	0.656	4.510
ResGen	−8.693	91.0%	−0.96%	**0.567**	**0.796**	4.940
FLAG ⋆	−9.837	96.0%	−1.45%	0.436	0.397	4.280
AMG ⋆	**−10.201**	**98.0%**	**−1.52%**	0.524	0.514	4.820
			Top 3			
Testset	−6.986	–	–	0.468	0.707	3.908
Pocket2mol	−8.525	82.7%	−0.86%	0.525	0.777	**4.947**
TargetDiff	−8.892	92.3%	−1.12%	0.457	0.502	4.453
DecompDiff	−8.116	92.0%	−0.78%	0.544	0.660	4.517
ResGen	−8.429	88.3%	−0.89%	**0.566**	**0.804**	4.913
FLAG ⋆	−9.414	93.0%	−1.34%	0.458	0.406	4.377
AMG ⋆	**−9.765**	**97.3%**	**−1.43%**	0.538	0.514	4.893
			Top 5			
Testset	−6.986	–	–	0.468	0.707	3.908
Pocket2mol	−8.395	80.4%	−0.81%	0.528	0.781	**4.946**
TargetDiff	−8.686	90.8%	−1.07%	0.453	0.511	4.446
DecompDiff	−7.946	90.6%	−0.72%	0.541	0.662	4.492
ResGen	−8.292	85.6%	−0.85%	**0.571**	**0.802**	4.912
FLAG ⋆	−9.159	90.8%	−1.29%	0.468	0.406	4.432
AMG ⋆	**−9.508**	**95.8%**	**−1.37%**	0.553	0.512	4.910
			Top 10			
Testset	−6.986	–	–	0.468	0.707	3.908
Pocket2mol	−8.190	77.1%	−0.76%	0.530	0.780	**4.941**
TargetDiff	−8.370	88.5%	−0.10%	0.466	0.523	4.519
DecompDiff	−7.715	86.4%	−0.65%	0.539	0.666	4.435
ResGen	−8.083	82.4%	−0.80%	**0.579**	**0.800**	4.908
FLAG ⋆	−8.756	87.3%	−1.19%	0.476	0.415	4.493
AMG ⋆	**−9.102**	**92.8%**	**−1.27%**	0.560	0.515	4.932

Properties evaluation of generated molecules

The primary task of SBMG models is to generate molecules with high affinity and proper drug-likeness properties. We first evaluated the performance of AMG and other competitive methods on the test set. The results of the top 1/3/5/10 generated molecules (ranked by Vina Score) are shown in Table 1. Unless specified otherwise, the top molecules throughout this paper are ranked based on the Vina Score. The Vina Scores of molecules generated by AMG surpass the second-ranked FLAG in the top 1/3/5/10 by 3.7%, 4.49%, 3.81%, and 3.95%, respectively. For a fair comparison, the 3D graph transformer was used as the encoder architecture for FLAG, which is its latest official implementation and is consistent with AMG. Furthermore, the molecules generated by AMG show the best HA and MPBG scores, indicating superior mean performance. Despite FLAG achieving high Vina Scores, it fails to generate drug-likeness molecules. We speculate that this may be the result of limited learning capability for intricate fragment features. AMG, in contrast, benefits from the pre-trained ligand encoders and the IA, resulting in a significant improvement in molecular properties, as presented in Fig. 2A. Drug candidate screening is a multifaceted process that involves evaluating not only their affinity but also their properties. Accordingly, we ranked candidates based on their QED and SA scores. The results are presented in Supplementary Tables S1 and S2, respectively. Table S1 indicates that AMG is nearly optimal regarding QED and affinity-related metrics, which suggests a higher likelihood of AMG generating high-affinity, drug-likeness molecules. We speculate that PT with the natural product compounds enhances the ability of AMG to generate bioactive molecules. Table S2 results, filtered by SA, indicate that Pocket2Mol tends to generate molecules easier to synthesize due to its atom-based generation method that favors fewer atoms. In this setting, DecompDiff performed best on the affinity metric. As fragment-by-fragment generation methods, AMG and FLAG occasionally generate molecules composed of a single or very few fragments. Although these molecules exhibit high SA, their simple structures often lead to lower affinity. Notably, when filtered by SA from the top 1 to top 10, we observe that as SA decreases, affinity-related metrics increase. We believe that easier-to-synthesize molecules, likely having simpler structures, may struggle to interact effectively with protein targets. Therefore, we suggest that future research focus on developing models that can balance affinity and synthetic accessibility while exercising caution in generating molecules with overly simplistic structures.

In addition to the quantitative evaluation of the generated molecules, we further visualized the interaction patterns. Two distinct pockets within the test set were focused on the following: `4aaw`, where the reference molecule demonstrated a high Vina Score, and `1fmc`, with a low Vina Score. Our selection criteria for molecules generated by each method were based on identifying those with the lowest Vina Scores, indicative of higher binding affinity. To further assess these molecules, we employed CB-Dock2 [58] for re-docking and subsequent visualization. As depicted in Fig. 2B and 2C, the molecules generated by AMG exhibit lower

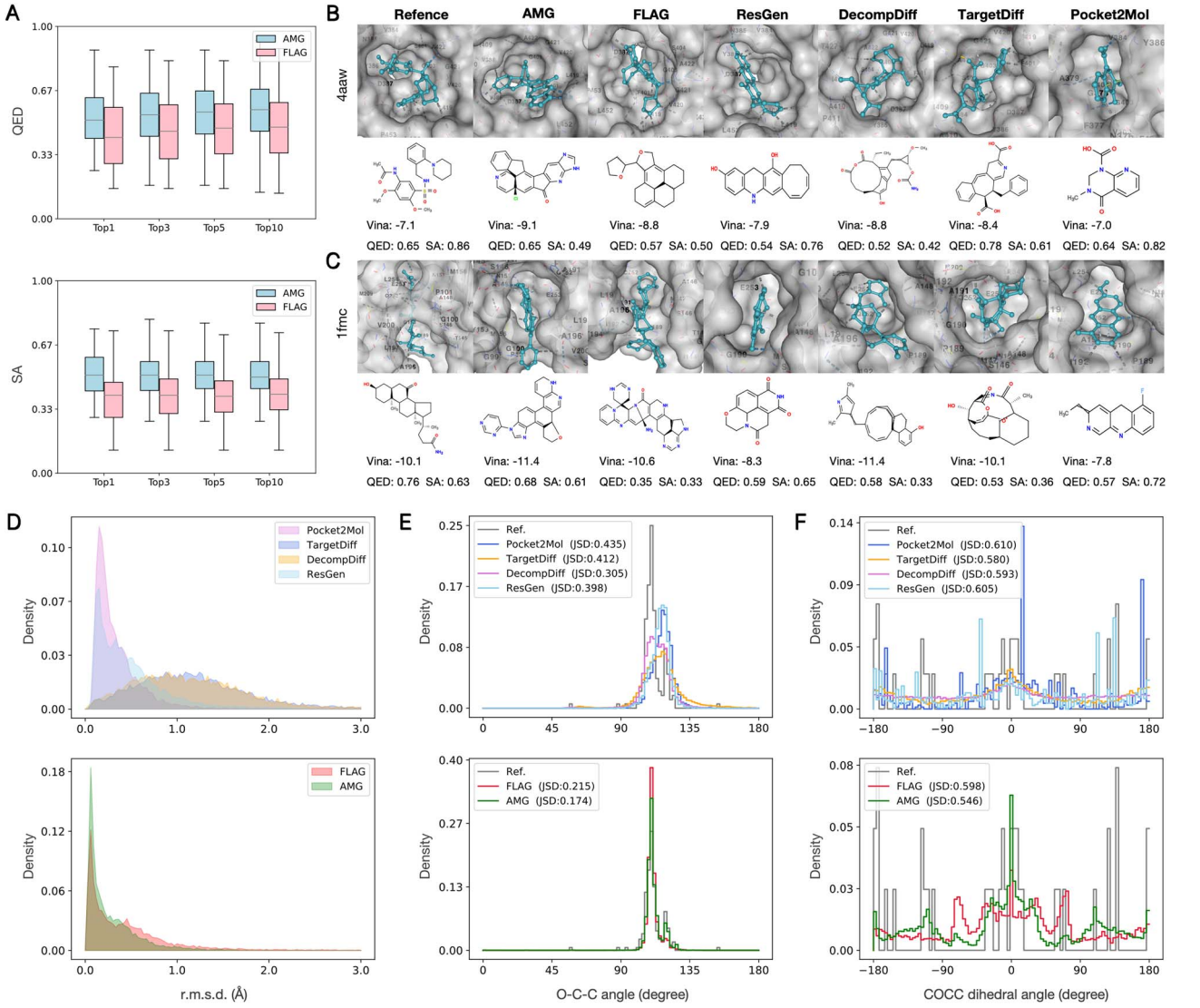

Figure 2. Evaluation of generated molecules. (A) Comparison of QED and SA among the top 1/3/5/10 molecules ranked by Vina Scores generated by AMG and FLAG. (B) Interaction patterns and molecular properties of the top 1 molecules generated for `4aaw` in the test set. (C) Interaction patterns and molecular properties of the top 1 molecules generated for `1fmc` in the test set. (D) Comparison of r.m.s.d. between generated geometry and 20 ETKDG-predicted conformations. (E) Distributions of OCC bond angle between the test set and generated molecules. (F) Distributions of COCC dihedral angle between the test set and generated molecules.

Vina Scores while maintaining proper molecular properties. An intriguing observation from our study is that the molecular structures generated by AMG closely mimic the shape of the reference molecules. This fidelity in replicating the reference molecules' shape underscores AMG's capability to generate molecular candidates that are not only potent in terms of binding affinity but also exhibit structures conducive to effective binding. This outcome may be attributed to the fragment-based decoder that utilizes substructures with prior knowledge to explore molecules, thereby avoiding the generation of unrealistic substructures.

Quality analysis of molecular structure

The generated molecules are expected to possess desired properties and exhibit high-quality structures at both global and substructure levels. For global-level structure evaluation, we analyzed the quality of generated molecular conformations. The ETKDG (classical conformation generation algorithm) [59] and the UFF force field [60] were employed to optimize the molecules generated by each method on the test set. We followed ResGen [20], which generated 20 conformations for each molecule using the aforementioned methods, and then calculated their r.m.s.d. with the original conformations generated directly from the generative models. Figure 2D presents the r.m.s.d. distribution of each method, with smaller values indicating a higher quality of the generated conformation. According to the results, AMG can achieve a significant advancement over previous methods. In addition to high-quality conformations, the generated molecules are expected to be rational at the substructure level. We computed the Jensen–Shannon divergence (JSD) of bond and dihedral angles between the distributions of generated molecules and the test set. A lower JSD value indicates a better similarity between two distributions. As consolidated in Table 2, AMG performs first or second across all angle categories, and FLAG is closely following. We visualized the JSD distributions of OCC (Fig. 2E) and COCC (Fig. 2F) of the generated molecules. An obvious observation is that AMG and FLAG show superior capability in accurately reproducing the distribution of reference molecules, which demonstrates that

Table 2. Jensen–Shannon divergence of bond angles (upper part) and dihedral angles (lower part) between the test set and generated molecules

	Pocket2Mol	TargetDiff	DecompDiff	ResGen	FLAG⋆	AMG⋆
CCC	0.308	0.322	**0.203**	0.299	0.316	0.212
OCO	0.521	0.486	0.483	0.437	0.361	**0.290**
CCO	0.464	0.375	0.260	0.430	**0.146**	0.237
OC=O	0.577	0.552	0.456	0.525	0.408	**0.304**
CCCC	0.450	0.335	0.300	0.348	**0.240**	0.278
cccc	0.575	0.477	0.366	0.536	0.235	**0.158**
COCC	0.610	0.580	0.593	0.605	0.598	**0.546**
CC=CC	0.468	0.489	0.524	0.398	0.387	**0.374**

Table 3. Molecular similarity and diversity analysis on the training set

	Frag.sim. (↓)	Finger.sim. (↓)	Diversity (↑)
Pocket2Mol	0.506	**0.278**	0.812
TargetDiff	0.846	0.304	0.846
DecompDiff	0.916	0.304	0.798
ResGen	0.711	0.286	0.792
FLAG⋆	0.388	0.308	0.857
AMG⋆	**0.340**	0.287	**0.859**

fragment-based models can better capture the distribution of sub-structures. For the more complex dihedral angle COCC, AMG exhibited superior performance compared to FLAG. This may be attributed to the advantages derived from enhanced molecular representation learning capabilities through our PT strategy.

Structural similarity and diversity analysis

The lower similarity between the training set and generated molecules typically indicates that molecular generative models have not merely memorized the features and distribution of the training set. Instead, they possess the capability to generate more innovative and diverse molecules. In our evaluation, we calculated similarity at both the fragment and molecular levels using fragment similarity (Frag.smi.) and molecular fingerprint similarity (Finger.sim.), as detailed in Table 3. As expected, the Frag.smi. of AMG-generated molecules is 0.340, notably superior to other methods. FLAG ranks second, surpassing other methods, indicating that fragment-level generation could effectively avoid generating redundant fragments. Due to its reliance on chemical priors, fragment-based generation tends to yield slightly higher molecular-level similarities than atom-based methods. Furthermore, the Finger.sim. of AMG is 0.287, ranking third and demonstrating superior performance to FLAG. We credit this to AMG's utilization of a more diverse library of fragments constructed from the PT dataset.

In addition to showcasing superior structural similarity, AMG also exhibits higher internal diversity among generated molecules, as illustrated in Table 3. It is equally crucial for a model to generate a diverse range of positive molecules [61]. Therefore, we next analyzed whether the models can navigate a broad chemical space to generate molecules that possess both high affinity and diversity. Specifically, we selected the generated molecules with higher affinity than the reference molecules in the test set and visualized their distribution using TMAP [62], a method that clusters molecules with similar chemical structures together. As shown in Fig. 3A, the structural distribution of molecules generated by AMG covers the test set and spreads far beyond it, illustrating the ability to explore the vast chemical space. In contrast, the structural distributions of molecules generated by TargetDiff, DecompDiff, and ResGen show limited variability across the reference molecules (Fig. 3B). FLAG also excels in exploring the chemical space, albeit prioritizing high diversity at the cost of affinity. Conversely, AMG achieves an optimal balance, securing the highest levels of both diversity and affinity (Fig. 3C). We speculate that this is the result of global-level optimization by the IA, which explores molecules with higher affinity without sacrificing diversity.

Ablation study

Existing SBMG models [17–23] typically capture pocket–ligand complex features by a single complex encoder, enabling learning of the global positional relationship to some extent. However, this manner is limited in capturing the local features of pockets and ligands, thereby affecting the generation process and ultimately compromising the properties of generated molecules. To interpret the contribution of our pre-trained separate pocket and ligand encoders, we employed t-SNE [63] to reduce the dimensionality of two encoded pocket–ligand complexes in the test set. The pre-trained separate encoders excel in extracting local features of pockets and ligands, evident from the distinct clustering of complex features into two classes (Fig. 3D). Conversely, the separate encoders trained exclusively on the CrossDocked dataset without PT (w/o PT), as well as those trained solely on the PT dataset without subsequent fine-tuning (w/o FT), were ineffective at distinguishing these features. Equipped with only a single complex encoder (w/ single encoder), it is difficult to distinguish between pocket and ligand features, leading to poor drug-likeness of the generated molecules.

The impact of different encoder architectures on model performance is also worth exploring. We used SchNet [64] and EGNN [65] as the backbone of encoders and visualized the features they extracted, as shown in Fig. 3E and 3F. Similar to AMG, the features extracted by the pre-trained encoders exhibited distinct clustering. EGNN efficiently captures geometric symmetries, resulting in an orderly clustered distribution in the embedding space. The message passing mechanism of SchNet focuses on local features. It can effectively distinguish different local structural features of proteins and ligands, forming multiple independent clusters. In contrast, AMG, taking 3D graph transformer as backbone, relies on its self-attention mechanism to capture global dependencies, leading to the formation of two distinct clusters. Table S5 shows the quantitative comparison results of the models constructed with different encoders. An observation is that AMG excels in high affinity, while EGNN exhibits a significant advantage in drug-likeness properties. Since our primary contribution does not lie

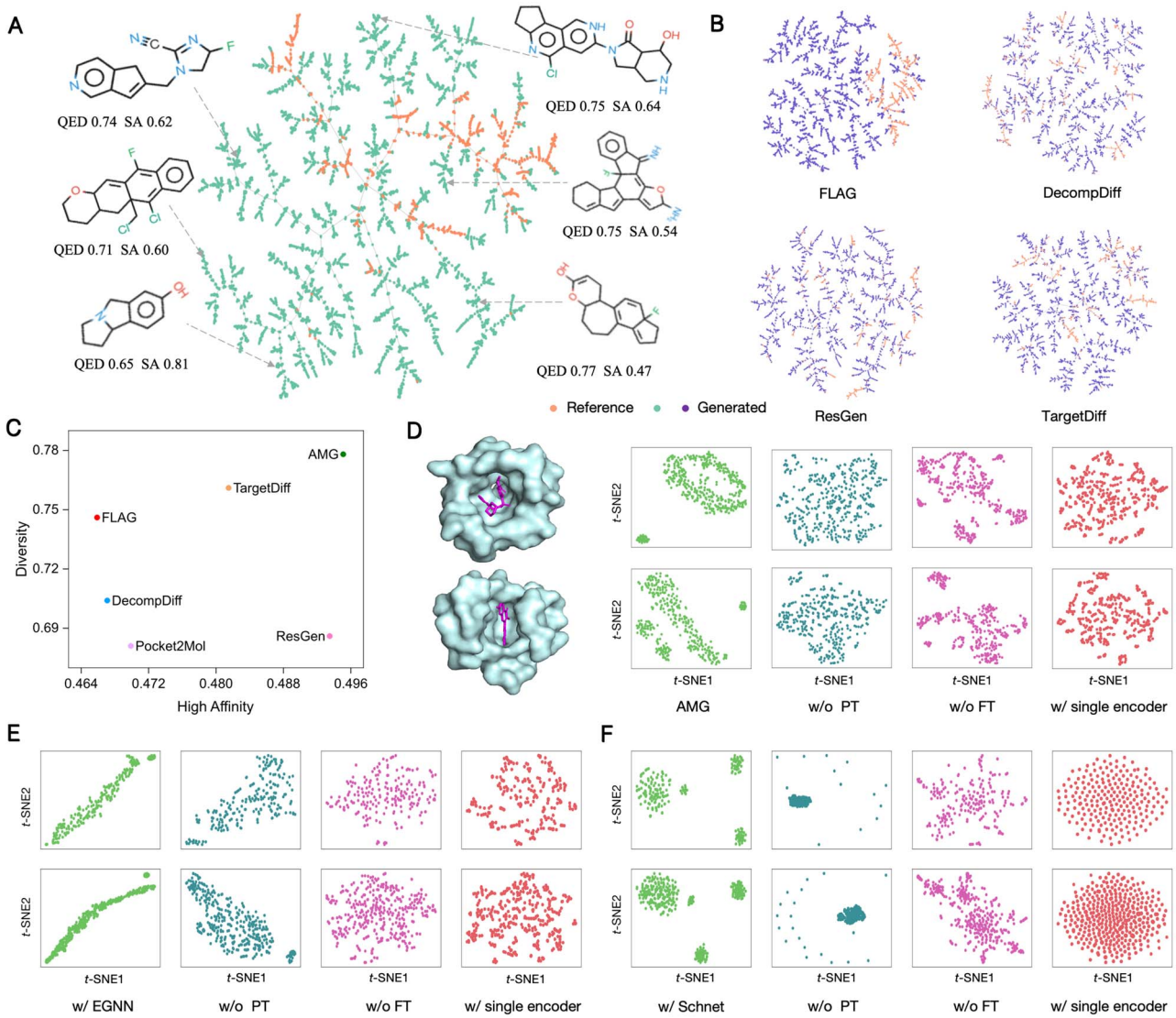

Figure 3. Cluster analysis for generated molecules. (A) TMAP distributions of the test set and AMG-generated molecules with lower Vina Scores than reference molecules. Each point represents a molecule, with closer positions indicating greater structural similarity between the molecules. (B) TMAP distributions of the test set and molecules generated by FLAG, ResGen, DecompDiff, and TargetDiff. (C) Comparison of diversity and high affinity of generated molecules (the closer to the top right, the better the effect). High Affinity is the percentage of generated molecules with Vina Scores lower than the reference molecules in the test set. Diversity is calculated after filtering out low-affinity molecules. (D) Visualization of t-SNE dimensionality reduction for two pocket–ligand features encoded by AMG using 3D graph transformer encoders. (E) Visualization of t-SNE dimensionality reduction for two pocket–ligand features encoded by AMG using EGNN encoders. (F) Visualization of t-SNE dimensionality reduction for two pocket–ligand features encoded by AMG using SchNet encoders.

in the design of the encoder architecture, we selected the 3D Graph Transformer as our backbone, prioritizing high affinity and ensuring fairness through consistency with FLAG's backbone. Future research could explore the integration of more advanced graph-based encoders, all of which can be seamlessly incorporated into our framework. Furthermore, Table S5 reveals that without encoder PT, performance in molecular properties declines significantly, underscoring the critical role of PT strategy across different encoder architectures.

To evaluate the impact of IA and PT strategy on model performance, we performed a quantitative analysis under different settings. Specifically, we compared the model performance without the IA (w/o IA) and without both the PT and IA (w/o PT & IA) to determine the role of these components. The performance of the top 5 generated molecules is presented in Table 4, with additional results available in Supplementary Table S3. As shown in Table 4, the removal of both PT and IA leads to a significant decrease in the Vina Score to −9.196 and the HA to 93.9%, highlighting the critical role of these components in enhancing molecular affinity. When only the IA module is excluded, the Vina Score drops from −9.508 to −9.323, the HA decreases from 95.8% to 94.4%, and the drug-likeness properties also diminish, underscoring the importance of the IA module in improving model performance. In contrast, when both modules are retained (AMG), the model achieves optimal performance across all metrics, confirming the synergistic effect of IA and PT in enhancing overall model performance.

Furthermore, we investigated the impact of various PT strategies on model performance, specifically examining the effects of not PT the ligand encoder (w/o ligand PT), pocket encoder (w/o pocket PT), and decoder (w/o decoder PT), as well as the absence of any PT (w/o PT). The performance of the top 5 generated

Table 4. Ablation studies of the IA module

	Vina Score (↓)	HA (↑)	MPBG (↓)	QED (↑)	SA (↑)	Lipinski (↑)
w/o PT & IA	−9.196	93.9%	−1.28%	0.480	0.437	4.412
w/o IA	−9.323	94.4%	−1.33%	0.544	0.501	4.824
AMG	**−9.508**	**95.8%**	**−1.37%**	**0.553**	**0.512**	**4.910**

Table 5. Ablation studies of the PT strategy

	Vina Score (↓)	HA (↑)	MPBG (↓)	QED (↑)	SA (↑)	Lipinski (↑)
w/o ligand PT	−8.835	90.2%	−1.26%	0.489	0.453	4.650
w/o pocket PT	−9.218	92.6%	−1.31%	**0.584**	0.507	**4.912**
w/o decoder PT	−9.356	91.0%	**−1.38%**	0.380	0.312	4.230
w/o PT	−9.156	92.4%	−1.32%	0.517	0.482	4.798
AMG	**−9.508**	**95.8%**	−1.37%	0.553	**0.512**	4.910

molecules is presented in Table 5, with additional results available in Supplementary Table S4. As shown in Table 5, the Vina Score of w/o ligand PT is −8.835, the lowest among all strategies, while the property metrics QED, SA, and Lipinski show a significant decline compared to AMG. This highlights the crucial role of ligand PT in improving the Vina Score and drug-likeness properties. The Vina Scores for w/o pocket PT and w/o decoder PT are −9.218 and −9.356, respectively, indicating a decline in performance compared to AMG, suggesting that PT of the pocket encoder and decoder also positively contributes to enhancing molecular affinity. Interestingly, the w/o pocket PT strategy yielded the highest QED score at 0.584, slightly surpassing AMG's score of 0.553, indicating that omitting pocket PT may enhance the drug-likeness of molecules. QED is typically used for rough screening, while a decrease of 1 kcal mol^{-1} in energy is believed to increase the activity of the binding molecule by approximately fivefold [20]. Consequently, we prioritize better affinity performance, even if it comes at the expense of a slight reduction in QED performance. Notably, the drug-likeness properties of w/o decoder PT strategy are the poorest, even lower than that of w/o PT strategy. We speculate that different components of the model collaborate during training, and the removal of the decoder PT alone could disrupt the alignment between separate encoders and the decoder. Consequently, the latent representations generated by the encoders may not be effectively utilized by the decoder, resulting in a decline in performance. Conversely, although the w/o PT strategy generally performed poorly, the absence of any PT interference might allow better coordination among the various modules, resulting in some cases where the w/o PT strategy outperforms the w/o decoder PT strategy. In conclusion, we maximized the overall performance of the model through targeted PT strategies, enabling AMG to generate molecules with high affinity while maintaining favorable drug-like properties.

Interaction patterns analysis on real-world therapeutic targets

The efficacy of an SBMG model extends beyond merely generating appropriate molecules for a test set; its true value lies in ensuring robust performance in real-world scenarios. Here, each model was utilized to generate molecules for real-world therapeutic targets and evaluated through quantitative and visual analyses. We selected two real-world therapeutic targets, AMPA (`6nj1`) and AGTR1 (`4yay`), from the PDB [32], and compared the top 5 molecules generated by each method (Fig. 4A). Similar to the observations on the CrossDocked dataset, AMG remains optimal, while other methods notably decline. By appropriately modeling interaction patterns, AMG can authentically capture interaction information. Furthermore, through exploration by the IA for high-affinity molecules, AMG demonstrates outstanding generalization capabilities even when dealing with real-world therapeutic targets. Candidate drugs that adhere to Lipinski's rule of five often exhibit lower attrition rates during clinical trials [66], we hence compared the average Lipinski scores of the top 5 molecules. The Lipinski scores of molecules generated by AMG are notably higher compared to other methods (Fig. 4B). Pocket2Mol, as an atom-based generation method, generates molecules with fewer atoms, resulting in higher Lipinski scores and lower affinity. Other methods suffer from a relative decrease of Lipinski on real-world therapeutic targets due to the absence of pre-trained molecular representation learning and optimization in the generation process. In addition to molecular property metrics, sampling efficiency is also crucial for generative models. This determines whether the model can generate a substantial number of candidate molecules within a given timeframe. Therefore, we measured the time of generating 100 molecules for each target, and the results indicate that AMG exhibits significantly superior sampling efficiency (Fig. 4C). Fragment-based generation can circumvent the inefficient sampling resulting from individually generating atoms, and the generation process optimized through the IA can enhance the success rate of generation, thereby improving sampling efficiency.

The critical step in virtual screening involves visualizing the docked pocket–ligand structures and excluding molecules with undesirable interaction patterns [5]. Therefore, we visualized the interaction patterns between real-world therapeutic targets and generated molecules to assess their rationality. Specifically, the generated molecules with the lowest Vina Score against `6nj1` and `4yay` were selected for visualization using PLIP [67] (Fig. 4D). Molecules generated by AMG show lower Vina Scores and form stable hydrogen bonds (blue solid lines), hydrophobic interactions (gray dashed lines), halogen bonds (green solid lines), and salt bridges (yellow dashed lines). In comparison, molecules generated by other methods have higher Vina Scores and form relatively simplistic interaction patterns. We then offered a more intuitive comparison of the interaction patterns between AMG-generated molecules and reference molecules (Fig. 4E). It can be observed that AMG essentially reproduces the interactions with residues involved in the reference molecular interactions (orange) while

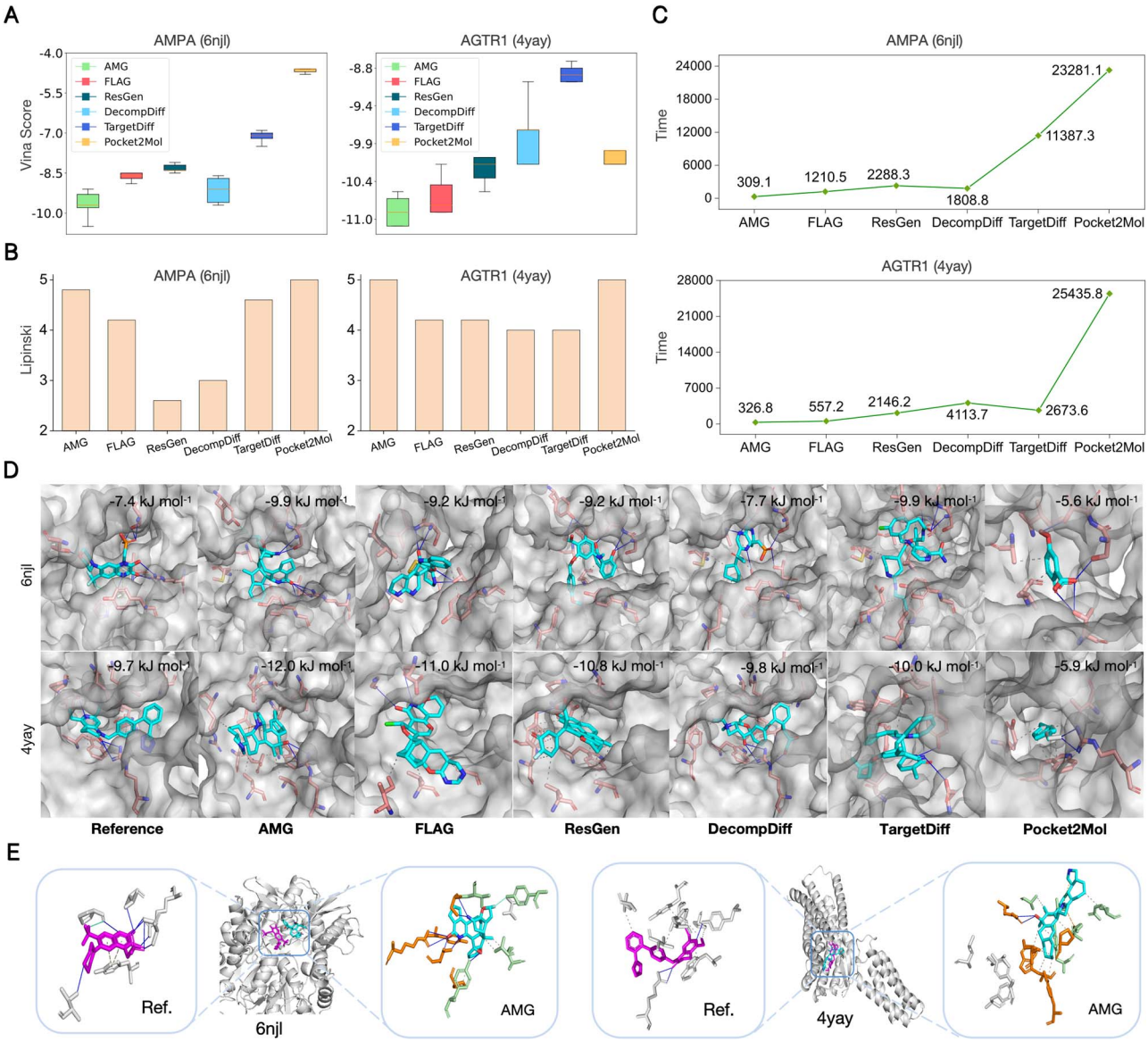

Figure 4. Visualization of generated molecules for real-world therapeutic targets. (A) Comparison of Vina Score for the top 5 molecules generated by each method for real-world therapeutic targets `6njl` and `4yay`. (B) Comparison of Lipinski for the top 5 molecules for `6njl` and `4yay`. (C) Comparison of sampling efficiency (runtime of generating 100 molecules) for `6njl` and `4yay`. (D) Visualization of interaction patterns for generated molecules against real-world therapeutic targets `6njl` and `4yay`. The gray area represents the protein pocket, molecules predominantly in light blue represent the generated molecules, pink denotes residues that can interact with the generated molecules, and the short lines of various colors represent different types of interactions that are consistent with PLIP. (E) Analysis of protein–ligand interaction between reference molecules (purple) and AMG-generated molecules (light blue) for `6njl` and `4yay`. The molecules generated by AMG partially reproduce the residues interacting with reference molecules (orange) and interact with some new residues (green).

also interacting with some new residues (green). Both quantitative and visual analyses demonstrated the excellent generalization of AMG in generating high-affinity molecules. For additional examples of other therapeutic targets, please refer to Supplementary Materials.

Discussion

The development of effective 3D molecular generative models conditioned on protein pockets represents a critical step forward in drug discovery. Here, we have proposed and implemented a DRL framework, termed AMG, for 3D molecular generation targeting protein pockets. Specifically, we designed a two-stage training strategy to achieve this. In the first stage, the model was trained in a supervised manner to reconstruct molecules and employed a PIIM module to predict and embed atom-level interaction energies. In the second stage, a DRL-based IA was integrated with the trained model, taking interaction energies as actions to steer the generation process at the feature level. Furthermore, we incorporated separate encoders for protein pockets and ligands, along with a robust PT strategy. This dual-encoder architecture enables AMG to leverage the extensive prior knowledge available in the pharmaceutical domain.

Given the existence of similar SBDD methods, we briefly discuss the novelty of AMG. Currently, the generation of 3D molecules can be broadly classified into two categories: atom-based generation and fragment-based generation. Methods such as Pocket2Mol [19], TargetDiff [21], and ResGen [20] generate molecules at the atomic level, while DecompDiff [22] and FLAG [23] implement at the fragment level. We adopted the

fragment-based generation approach because pre-constructed fragments inherently possess chemical priors, helping to avoid generating structures that lack practical significance. Solely using supervised learning might result in unreasonable fragment combinations, leading to poor drug properties. For instance, although molecules generated by FLAG exhibit good affinity, they score lower on metrics such as QED, SA, and Lipinski. Therefore, we introduced a PT strategy and RL method to steer the model in optimizing specific properties of the molecules.

Our experiments have demonstrated that AMG outperforms current state-of-the-art SBDD methods. Furthermore, molecules generated by AMG exhibit superior drug-likeness properties compared to those generated by FLAG, addressing common shortcomings in fragment-based approaches. Balancing high affinity with drug-likeness properties is crucial for producing viable drug candidates. The framework's capability to generate realistic molecular structures further underscores its potential practical application in drug discovery. The structural quality of AMG-generated molecules is reflected in their lower r.m.s.d. values and better Jensen-Shannon divergence scores for bond angles and dihedrals, confirming that AMG can generate realistic and stable molecular structures. Additionally, our method's ability to maintain high diversity while preserving high affinity is vital for discovering new therapeutic candidates. The capability of AMG to generate high-affinity molecules with stable interaction patterns on real-world therapeutic targets highlights its generalizability. Visual analysis of interaction patterns confirms that AMG can replicate and enhance key interactions observed in reference molecules, making it a reliable tool in practical applications. The superior Lipinski scores and efficient sampling time of AMG further validate its effectiveness in drug discovery. Our ablation study reveals the significant contribution of the pre-trained separate encoder and IA to the overall performance of AMG. The ability of these components to extract and optimize local features of pockets and ligands enhances the affinity and properties of the generated molecules.

There are some limitations of the current study. Firstly, the PIIM module models the local interactions at the atomic level and then steers fragment-based molecular generation with the extracted atomic-level interaction features. This approach integrates both local atomic-level and global fragment-level interactions to some extent, resulting in molecules with high affinity. However, the current framework lacks explicit modeling of fragment-level interactions, which may lead to the inability of PIIM to adequately simulate certain non-local contributions of aromatic ring interactions to intermolecular interactions, such as $\pi - \pi$ interactions between aromatic ring systems [68]. In future work, researchers could consider more sophisticated and physically meaningful joint modeling of interactions at both the atomic and fragment levels to guide molecular generation/optimization. Secondly, as a fragment-based generation method, AMG continues to face unresolved inherent limitations. For example, when generating molecules for pockets with small cavities, AMG tends to select just a single fragment from the library to complete the generation. This leads to the phenomenon observed in Table S2, where the molecules generated by AMG are highly synthesizable but exhibit lower affinity due to their simplistic structure. Future research could investigate strategies to circumvent this limitation, possibly by integrating atom-based and fragment-based methods for broader applicability. Thirdly, substantial research opportunities remain in optimizing fragment-based molecular generation models through reinforcement learning. Currently, while some studies employing reinforcement learning target SBMGs [69–71], few focus specifically on fragment-based approaches. Although this study presents optimization via an interaction-based agent approach, opportunities for enhancement remain, notably in conducting multi-objective optimization to substantially improve affinity and drug-likeness properties.

In summary, drawing inspiration from generative pre-trained models and physics-informed pocket–ligand interaction modeling, we propose a robust framework that integrates DRL in fragment-based generation to address key challenges in structure-based drug design. Its superior performance in affinity and drug-likeness, along with high-quality structural generation and effective chemical space exploration, positions AMG as a promising tool for generating viable drug candidates. The inventive use of separate encoders, PT strategies, and an IA within the framework ensures its generalization across diverse protein targets, underscoring its potential for various downstream applications in drug discovery. Future research could refine AMG further and broaden its applicability to a wider range of pharmaceutical areas, thereby amplifying its impact in the field.

Key Points

- We introduce a reinforcement learning-based interaction agent to progressively steer fragment-based 3D molecular generation targeting protein pockets.
- A dedicated two-stage training strategy is proposed to explicitly capture interaction features and optimize the interaction agent.
- Two datasets are constructed, including 277 530 3D structures of protein pockets and 404 397 3D structures of natural products, for pre-training our model.
- Evaluations on extensive datasets show the superior performance of our work over other baselines in affinity and drug-likeness, confirming its effectiveness in capturing 3D molecular geometry and interaction patterns.

Acknowledgments

The authors thank the anonymous reviewers for their valuable suggestions.

Supplementary data

Supplementary data is available at *Briefings in Bioinformatics* online.

Conflict of interest: None declared.

Funding

This work is supported by funding from the National Natural Science Foundation of China (No. 32170589, 32370616, 92168205, 22074105), the National Key Research and Development Program of China (No. 2022YFA1103102), the Shanghai Rising Star Program (No.21QC1400900), Xiaomi Young Talents Program and the European Union' s Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreements No. 945539 (Human Brain Project SGA3).

Data availability

The ligand and pocket datasets used for pre-training in this study were sourced from the COCONUT database (https://coconut.naturalproducts.net) and the Protein Data Bank (PDB) (https://

www.rcsb.org), respectively. The dataset used for two-stage training was obtained from CrossDocked2020 (https://bits.csb.pitt.edu/files/crossdock2020).

Author contributions

X.Z. and G.C. conceived the main idea. X.Z. and J.H. developed the code for the core algorithm. J.H., S.Q., and Y.Z. designed and performed the data analysis. F.L. and Z.T. collected and processed the pre-trained datasets. X.Z., G.C., S.Q., F.L., Z.T., and Y.Z. wrote the initial version of the manuscript. A.K. and S.G. designed the experiments and revised the manuscript.

Code availability

The source code and pre-trained models for this study are available at https://github.com/ispc-lab/AMG.

References

1. Bajorath J. Integration of virtual and high-throughput screening. *Nat Rev Drug Discov* 2002; **1**:882–94. https://doi.org/10.1038/nrd941.
2. Schneider G. Virtual screening: an endless staircase? *Nat Rev Drug Discov* 2010; **9**:273–6. https://doi.org/10.1038/nrd3139.
3. Lyne PD. Structure-based virtual screening: an overview. *Drug Discov Today* 2002; **7**:1047–55.
4. Kitchen DB, Decornez H, Furr JR. *et al*. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov* 2004; **3**:935–49. https://doi.org/10.1038/nrd1549.
5. Ghosh S, Nie A, An J. *et al*. Structure-based virtual screening of chemical libraries for drug discovery. *Curr Opin Chem Biol* 2006; **10**:194–202. https://doi.org/10.1016/j.cbpa.2006.04.002.
6. Goel M, Aggarwal R, Sridharan B. *et al*. Efficient and enhanced sampling of drug-like chemical space for virtual screening and molecular design using modern machine learning methods. *Wiley Interdiscip Rev: Comput Mol Sci* 2023; **13**:e1637. https://doi.org/10.1002/wcms.1637.
7. Shoichet BK. Virtual screening of chemical libraries. *Nature* 2004; **432**:862–5. https://doi.org/10.1038/nature03197.
8. Gómez-Bombarelli R, Wei JN, Duvenaud D. *et al*. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Cent Sci* 2018; **4**:268–76. https://doi.org/10.1021/acscentsci.7b00572.
9. Dai H, Tian Y, Dai B. *et al*. Syntax-directed variational autoencoder for molecule generation. In: *International Conference on Learning Representations*, Vancouver Canada: ICLR, 2018.
10. Polykovskiy D, Zhebrak A, Sanchez-Lengeling B. *et al*. Molecular Sets (MOSES): a benchmarking platform for molecular generation models. *Front Pharmacol* 2020; **11**:565644. https://doi.org/10.3389/fphar.2020.565644.
11. Bagal V, Aggarwal R, Vinod PK, Deva Priyakumar U MolGPT: molecular generation using a transformer-decoder model. *J Chem Inf Model* 2021; **62**:2064–76. https://doi.org/10.1021/acs.jcim.1c00600.
12. Hoogeboom E, Satorras VG, Vignac C. *et al*. Equivariant diffusion for molecule generation in 3D. In: *International Conference on Machine Learning*, pp. 8867–87. Baltimore MD USA: PMLR, 2022.
13. Gebauer NWA, Gastegger M, Hessmann SSP. *et al*. Inverse design of 3D molecular structures with conditional generative neural networks. *Nat Commun* 2022; **13**:973. https://doi.org/10.1038/s41467-022-28526-y.
14. Kong X, Huang W, Tan Z. *et al*. Molecule generation by principal subgraph mining and assembling. *Adv Neural Inf Process Syst* 2022; **35**:2550–63.
15. Lemeng W, Gong C, Liu X. *et al*. Diffusion-based molecule generation with informative prior bridges. *Adv Neural Inf Process Syst* 2022; **35**:36533–45.
16. Moret M, Angona IP, Cotos L. *et al*. Leveraging molecular structure and bioactivity with chemical language models for de novo drug design. *Nat Commun* 2023; **14**:114. https://doi.org/10.1038/s41467-022-35692-6.
17. Liu M, Luo Y, Uchino K. *et al*. Generating 3D molecules for target protein binding. In: *International Conference on Machine Learning*, pp. 13912–24. Baltimore MD USA: PMLR, 2022.
18. Ragoza M, Masuda T, Koes DR. Generating 3D molecules conditional on receptor binding sites with deep generative models. *Chem Sci* 2022; **13**:2701–13. https://doi.org/10.1039/D1SC05976A.
19. Peng X, Luo S, Guan J. *et al*. Pocket2mol: Efficient molecular sampling based on 3D protein pockets. In: *International Conference on Machine Learning*, pp. 17644–55. Baltimore MD USA: PMLR, 2022.
20. Zhang O, Zhang J, Jin J. *et al*. ResGen is a pocket-aware 3D molecular generation model based on parallel multiscale modelling. *Nat Mach Intell* 2023; **5**:1020–30. https://doi.org/10.1038/s42256-023-00712-7.
21. Guan J, Qian W W, Peng X. *et al*. 3D equivariant diffusion for target-aware molecule generation and affinity prediction. In: *International Conference on Learning Representations*, Kigali Rwanda: ICLR, 2023.
22. Guan J, Zhou X, Yuwei Yang Y. *et al*. DecompDiff: diffusion models with decomposed priors for structure-based drug design. In: *International Conference on Machine Learning*, pp. 11827–46. Hawaii USA: PMLR, 2023.
23. Zhang Z, Min Y, Zheng S. *et al*. Molecule generation for target protein binding with structural motifs. In: *International Conference on Learning Representations*, ICLR, 2022.
24. Segler MHS, Kogej T, Tyrchan C. *et al*. Generating focused molecule libraries for drug discovery with recurrent neural networks. *ACS Cent Sci* 2018; **4**:120–31. https://doi.org/10.1021/acscentsci.7b00512.
25. Wang X, Gao C, Han P. *et al*. PETrans: de novo drug design with protein-specific encoding based on transfer learning. *Int J Mol Sci* 2023; **24**:1146. https://doi.org/10.3390/ijms24021146.
26. Krishnan SR, Bung N, Vangala SR. *et al*. De novo structure-based drug design using deep learning. *J Chem Inf Model* 2021; **62**:5100–9. https://doi.org/10.1021/acs.jcim.1c01319.
27. Li Y, Pei J, Lai L. Structure-based de novo drug design using 3D deep generative models. *Chem Sci* 2021; **12**:13664–75. https://doi.org/10.1039/D1SC04444C.
28. Anderson AC. The process of structure-based drug design. *Chem Biol* 2003; **10**:787–97. https://doi.org/10.1016/j.chembiol.2003.09.002.
29. Atz K, Grisoni F, Schneider G. Geometric deep learning on molecular representations. *Nat Mach Intell* 2021; **3**:1023–32. https://doi.org/10.1038/s42256-021-00418-8.
30. Isert C, Atz K, Schneider G. Structure-based drug design with geometric deep learning. *Curr Opin Struct Biol* 2023; **79**:102548. https://doi.org/10.1016/j.sbi.2023.102548.
31. Francoeur PG, Masuda T, Sunseri J. *et al*. Three-dimensional convolutional neural networks and a cross-docked data set for structure-based drug design. *J Chem Inf Model* 2020; **60**:4200–15. https://doi.org/10.1021/acs.jcim.0c00411.
32. Burley SK, Bhikadiya C, Bi C. *et al*. RCSB Protein Data Bank (rcsb.org): Delivery of experimentally-determined PDB

structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. *Nucleic Acids Res* 2023; **51**:D488–508. https://doi.org/10.1093/nar/gkac1077.
33. Sorokina M, Merseburger P, Rajan K. *et al*. Coconut online: collection of open natural products database. *J Chem* 2021; **13**: 1–13.
34. Le Guilloux V, Schmidtke P, Tuffery P. Fpocket: an open source platform for ligand pocket detection. *BMC Bioinform* 2009; **10**: 1–11.
35. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat Biotechnol* 2017; **35**:1026–8. https://doi.org/10.1038/nbt.3988.
36. Schlichtkrull M, Kipf TN, Bloem P. *et al*. Modeling relational data with graph convolutional networks. In: *The Semantic Web: 15th International Conference*, pp. 593–607. Heraklion, Crete, Greece: ESWC, 2018.
37. Vaswani A, Shazeer N, Parmar N. *et al*. Attention is all you need. Adv Neural Inf Process Syst 2017;**30**:5998–6008.
38. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 2010; **31**:455–61. https://doi.org/10.1002/jcc.21334.
39. Huang S-Y, Grinter SZ, Zou X. Scoring functions and their evaluation methods for protein–ligand docking: recent advances and future directions. *Phys Chem Chem Phys* 2010; **12**:12899–908. https://doi.org/10.1039/c0cp00151a.
40. Moon S, Zhung W, Yang S. *et al*. PigNet: a physics-informed deep learning model toward generalized drug–target interaction predictions. *Chem Sci* 2022; **13**:3661–73. https://doi.org/10.1039/D1SC06946B.
41. Konda V, Tsitsiklis J. Actor-critic algorithms. Adv Neural Inf Process Syst 1999;**12**:1008–1014.
42. Haarnoja T, Zhou A, Hartikainen K. *et al*. Soft actor-critic algorithms and applications. arXiv preprint arXiv:1812.05905.2018. https://arxiv.org/abs/1801.01290.
43. Goel M, Raghunathan S, Laghuvarapu S. *et al*. Molegular: molecule generation using reinforcement learning with alternating rewards. *J Chem Inf Model* 2021; **61**:5815–26. https://doi.org/10.1021/acs.jcim.1c01341.
44. Liu S, Guo H, Tang J. Molecular geometry pretraining with SE (3)-invariant denoising distance matching. In: *International Conference on Learning Representations*, Kigali Rwanda: ICLR, 2023.
45. Liu S, Wang H, Liu W. *et al*. Pre-training molecular graph representation with 3D geometry. In: International Conference on Learning Representations, ICLR, 2021.
46. Wang Y, Wang J, Cao Z. *et al*. Molecular contrastive learning of representations via graph neural networks. *Nat Mach Intell* 2022; **4**:279–87. https://doi.org/10.1038/s42256-022-00447-x.
47. Senior AW, Evans R, Jumper J. *et al*. Improved protein structure prediction using potentials from deep learning. *Nature* 2020; **577**: 706–10. https://doi.org/10.1038/s41586-019-1923-7.
48. Grondman I, Busoniu L, Lopes GAD. *et al*. A survey of actor-critic reinforcement learning: standard and natural policy gradients. *IEEE Trans Syst Man Cybern C (Appl Rev)* 2012; **42**:1291–307. https://doi.org/10.1109/TSMCC.2012.2218595.
49. Schulman J, Wolski F, Dhariwal P. *et al*. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347. 2017. https://arxiv.org/abs/1707.06347.
50. Chikhaoui K, Ghazzai H, Massoud Y. PPO-based reinforcement learning for UAV navigation in urban environments. In: *2022 IEEE 65th International Midwest Symposium on Circuits and Systems*, pp. 1–4. Fukuoka, Japan: MWSCAS, 2022.
51. Alhossary A, Handoko SD, Yuguang M. *et al*. Fast, accurate, and reliable molecular docking with Quickvina 2. *Bioinformatics* 2015; **31**:2214–6. https://doi.org/10.1093/bioinformatics/btv082.
52. Graff D E, Coley C W. pyscreener: A Python Wrapper for Computational Docking Software[J]. *Journal of Open Source Software*, 2022;**7**:3950.
53. Lin H, Huang Y, Liu M. *et al*. DiffBP: generative diffusion of 3D molecules for target protein binding. arXiv preprint arXiv:2211.11214. 2022. https://arxiv.org/abs/2211.11214.
54. Patrícia Bento A, Hersey A, Félix E. *et al*. An open source chemical structure curation pipeline using RDKit. *J Chem* 2020; **12**:1–16. https://doi.org/10.1186/s13321-020-00456-1.
55. Lipinski CA, Lombardo F, Dominy BW. *et al*. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 1997; **23**:3–25. https://doi.org/10.1016/S0169-409X(96)00423-1.
56. Bajusz D, Rácz A, Héberger K. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *J Chem* 2015; **7**:1–13. https://doi.org/10.1186/s13321-015-0069-3.
57. Degen J, Wegscheid-Gerlach C, Zaliani A. *et al*. On the art of compiling and using'drug-like'chemical fragment spaces. *Chem Med Chem* 2008; **3**:1503–7. https://doi.org/10.1002/cmdc.200800178.
58. Liu Y, Yang X, Gan J. *et al*. CB-Dock2: improved protein–ligand blind docking by integrating cavity detection, docking and homologous template fitting. *Nucleic Acids Res* 2022; **50**:W159–64. https://doi.org/10.1093/nar/gkac394.
59. Riniker S, Landrum GA. Better informed distance geometry: using what we know to improve conformation generation. *J Chem Inf Model* 2015; **55**:2562–74. https://doi.org/10.1021/acs.jcim.5b00654.
60. Rappé AK, Casewit CJ, Colwell KS. *et al*. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J Am Chem Soc* 1992; **114**:10024–35. https://doi.org/10.1021/ja00051a040.
61. Jin W, Barzilay R, Jaakkola T. Multi-objective molecule generation using interpretable substructures. In: *International Conference on Machine Learning*, pp. 4849–59. Vienna Austria: PMLR, 2020.
62. Probst D, Reymond J-L. Visualization of very large high-dimensional data sets as minimum spanning trees. *J Chem* 2020; **12**:1–13. https://doi.org/10.1186/s13321-020-0416-x.
63. Van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res 2008;**9**:2579–2605.
64. Schütt K, Kindermans P-J, Felix HES. *et al*. SchNet: a continuous-filter convolutional neural network for modeling quantum interactions. Adv Neural Inf Process Syst 2017;**30**:991–1001.
65. Satorras VG, Hoogeboom E, Welling M. E (n) equivariant graph neural networks. In: *International conference on machine learning*, pp. 9323–32. PMLR, 2021.
66. Lipinski CA. Lead-and drug-like compounds: the rule-of-five revolution. *Drug Discov Today Technol* 2004; **1**:337–41. https://doi.org/10.1016/j.ddtec.2004.11.007.
67. Salentin S, Sven Schreiber V, Haupt J. *et al*. PLIP: fully automated protein–ligand interaction profiler. *Nucleic Acids Res* 2015; **43**:W443–7. https://doi.org/10.1093/nar/gkv315.

68. Cao D, Chen G, Jiang J. *et al*. Generic protein–ligand interaction scoring by integrating physical prior knowledge and data augmentation modelling. *Nat Mach Intell* 2024;**6**:1–13.
69. Fengqing L, Li M, Min X. *et al*. De novo generation of dual-target ligands using adversarial training and reinforcement learning. *Brief Bioinform* 2021; **22**:bbab333. https://doi.org/10.1093/bib/bbab333.
70. Lee D, Cho Y. Fine-tuning pocket-conditioned 3D molecule generation via reinforcement learning. In: *ICLR Workshop on Generative and Experimental Perspectives for Biomolecular Design*, Vienna Austria: ICLR 2024.
71. Xiuyuan H, Liu G, Zhao Y. *et al*. De novo drug design using reinforcement learning with multiple GPT agents. *Adv Neural Inf Process Syst* 2024;**36**:7405–7418.

Briefings in Bioinformatics, 2025, **26(4)**, bbae531
https://doi.org/10.1093/bib/bbae531
Problem Solving Protocol