nature machine intelligence

ARTICLES

https://doi.org/10.1038/s42256-022-00527-y

Check for updates

Accelerated rational PROTAC design via deep learning and molecular simulations

Shuangjia Zheng[1,2,4], Youhai Tan[1,4], Zhenyu Wang[2], Chengtao Li[2], Zhiqing Zhang[2], Xu Sang[2], Hongming Chen[3] and Yuedong Yang[1]✉

Proteolysis-targeting chimeras (PROTACs) have emerged as effective tools to selectively degrade disease-related proteins by using the ubiquitin-proteasome system. Developing PROTACs involves extensive tests and trials to explore the vast chemical space. To accelerate this process, we propose a novel deep generative model for the rational design of PROTACs in a low-resource setting, which is then guided to sample PROTACs with optimal pharmacokinetics through deep reinforcement learning. Applying this method to the bromodomain-containing protein 4 target protein, we generated 5,000 compounds that were further filtered through machine learning-based classifiers and physics-driven simulations. As a proof of concept, we identified, synthesized and experimentally tested six candidate bromodomain-containing protein 4-degrading PROTACs, of which three were validated by cell-based assays and western blot analysis. One lead candidate was further tested and demonstrated favourable pharmacokinetics in mice. This combination of deep learning and molecular simulations may facilitate rational PROTAC design and optimization.

Since the first proof-of-concept study of proteolysis-targeting chimeras (PROTACs) in 2001 (ref. [1]), PROTACs have emerged as effective tools to selectively degrade disease-related proteins by using the ubiquitin-proteasome system. PROTACs contain three parts: a ligand (warhead) targeting the protein of interest (POI), a ligand that recruits an E3 ubiquitin ligase and a chemical linker that connects the two ligands[1,2]. Because of this hetero-bifunctional structure, PROTACs have the ability to simultaneously bind the POI and E3 ligase, forming a ternary complex and promoting polyubiquitination and degradation of the POI[3]. Therefore, PROTACs only require transient binding to the target protein to induce ubiquitination and degradation[4,5], which is different from conventional occupancy-driven inhibitors that need adequate binding affinity to druggable sites of target proteins. In addition, PROTACs are not limited to occupying druggable active sites and thus have the potential to utilize all surface binding sites of the target protein for modulating 'non-druggable' targets[6,7]. Accordingly, the rational design of PROTACs is much more challenging than traditional small-molecule discovery.

The rational design of PROTACs can be divided into the design of three components. While the discovery of warheads and E3 ligands is not fundamentally different from the regular small-molecule discovery process, the design of linkers is experimentally challenging because the POI and E3 ligase do not interact without efficient PROTACs. Traditional methods have to design new PROTACs through extensive tests and trials[8,9], which are extremely inefficient. Many recent efforts have shifted the goalposts to de novo PROTAC design by generating linkers, because the linker is increasingly known to be critical for the physicochemical properties and degradation activity of PROTACs[10–12]. Unfortunately, linker design remains a formidable challenge due to the structural complexity and dynamics of ternary structures. Another major challenge preventing PROTACs from realizing their therapeutic potential is that the designed molecules do not conform to the accepted drug properties associated with orally administered drugs[13,14]. The engineering of enhanced pharmacokinetics (PK) has proven to be challenging due to the large and flexible nature of PROTACs[15]. Thus, novel approaches are needed to increase the discovery rate of new functional PROTACs.

On the basis of an intelligent exploration of chemical space, de novo molecular design has been greatly advanced by recent breakthroughs in deep generative models[16,17]. Various generative neural networks, such as recurrent neural networks[18–21], transformer neural networks[22,23], autoencoders[24,25] and generative adversarial networks[26,27], have proven to be effective for generating desirable small molecules[28,29], peptides[30] and antibodies[31]. These strategies have also been utilized to generate linkers for PROTACs. For example, Imrie et al.[32] developed a graph-based deep generator (DeLinker), incorporating three-dimensional (3D) structural information directly into the design process. Meanwhile, Yang et al.[33] transformed the linker design into a sentence completion task and introduced a language model (SyntaLinker) to generate novel linkers given simplified molecular-input line-entry system (SMILES) of hit fragments. Both methods demonstrate the generation of various linkers for de novo PROTAC design. However, these methods are limited to training on small molecules, without considering the differences in the design strategies and chemical space between small molecules and PROTACs. Additionally, they did not consider the drug metabolism and PK properties of the generated molecules. This is partly because of the extremely small amount of data publicly available for PROTAC. For example, the largest open-source PROTAC database currently contains only 2,300 samples[34], covering a tiny fraction of the chemical space. Considering such a small sample size, it is challenging to train a model that is capable of generating novel PROTACs simultaneously with desired properties and diversity, as also indicated in other studies in a low-resource setting[22,35]. More importantly, none of these previous generative models achieved experimental validation involving the synthesis of novel PROTACs for in vitro or in vivo testing.

[1]School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, China. [2]Galixir Technologies, Beijing, China. [3]Guangzhou Laboratory, Guangzhou, China. [4]These authors contributed equally: Shuangjia Zheng and Youhai Tan. ✉e-mail: yangyd25@mail.sysu.edu.cn

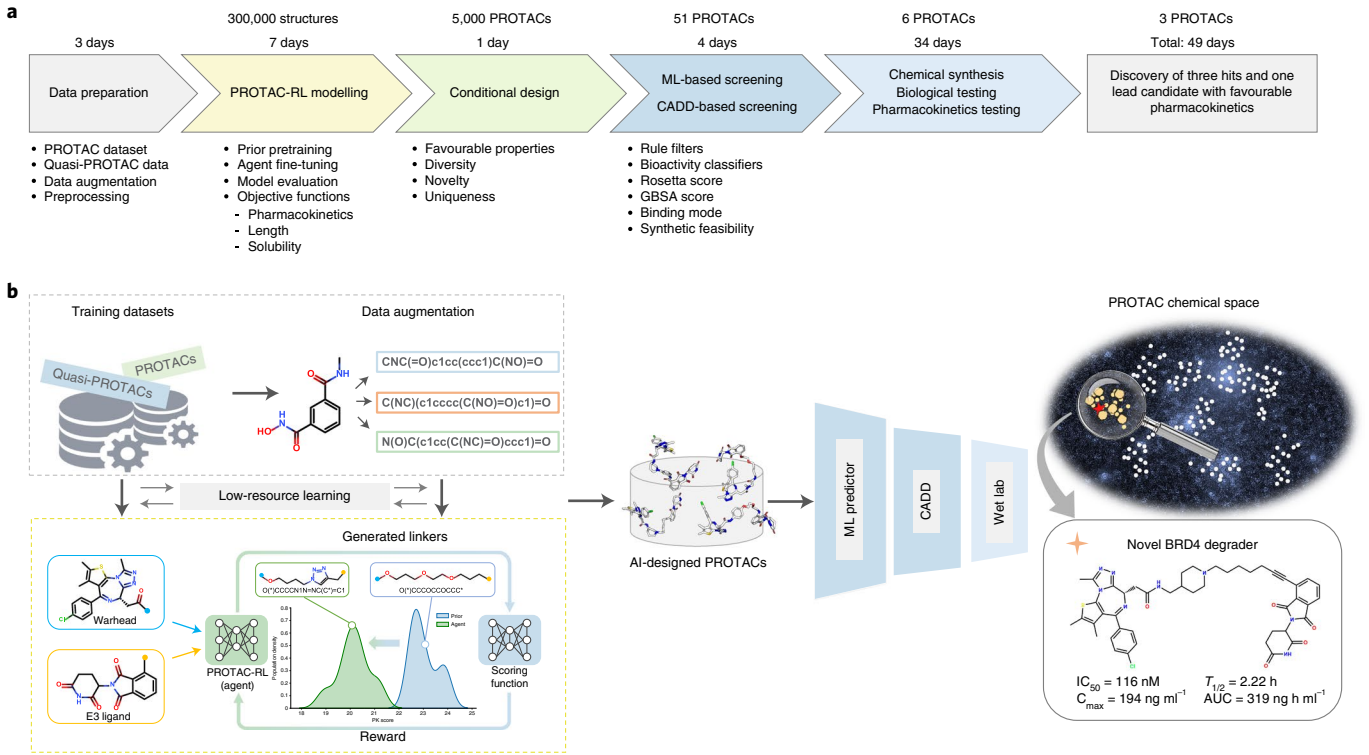


Fig. 1 | Approach overview. a, Overview and time line of the proposed artificial intelligence (AI)-driven approach for accelerated PROTAC design. **b**, The general work flow for the design of lead PROTACs using PROTAC-RL. CADD, computer-aided drug discovery.

In this work, we have developed PROTAC-RL, a novel deep generative model that combines an augmented transformer[36] architecture and memory-assisted reinforcement learning (RL) for rational PROTAC design. The model takes a pair of E3 ligand and warhead as input and outputs designed linkers to generate chemically feasible PROTACs with favourable properties. To overcome the problem of the small amount of training data, we first pretrained a fragment linking model with a transformer neural network using a large collection of quasi-PROTAC small molecules that have a chemical space similar to PROTACs, which is fine-tuned by using actual PROTACs augmented with randomized SMILES[37] of fragments. This trained Proformer model was then fed into a memory-assisted RL with empirical reward functions to generate PROTACs with better PK attributes. As a proof of concept, we selected the bromodomain-containing protein 4 (BRD4) target protein and generated 5,000 PROTACs, which were further clustered and screened through hierarchical machine learning classifiers and physics-driven molecular simulations. According to the synthetic accessibility, we synthesized and experimentally tested six BRD4-degrading PROTACs, of which three showed inhibitory activity against BRD4. One lead candidate simultaneously showed high anti-proliferative potency against the Molt4 cell line and favourable PK in mice (Fig. 1). This rapid discovery (within 49 days) highlights the significant impact that the combination of deep learning and molecular dynamics can have to facilitate efficient PROTAC design and optimization.

Results

Overview of PROTAC-RL. For a given pair of warhead and E3 ligands, rational PROTAC design requires sampling the vast chemical space to generate chemically feasible PROTAC molecules simultaneously with preferred PK properties. To achieve these two aims, our method couples two corresponding modules, namely a 'prior' network (Proformer) that is responsible for sampling the PROTAC chemical space and an 'agent' network (PROTAC-RL) that navigates the prior model towards desirable properties. For the first task, we formulate the problem as a sentence completion process, where the input is the pair of warheads and E3 ligands and the outputs are the designed PROTACs. Both the inputs and outputs are represented in molecular SMILES notations. The basic prior generating model, Proformer, was developed based on transformer neural networks because of their superior performance in molecular modelling (Supplementary Fig. 1). Owing to the small number of PROTAC molecules available (<2,300, namely PROTAC-DB[34]), we trained a basic model from a culled set of 294,675 quasi-PROTAC small molecules (a similar size to PROTAC) and fine-tuned the model on the real PROTAC set. We also improved the fine-tuning performance by augmenting PROTACs with randomized SMILES strings[37]. For the second task, we fed the Proformer into a reinforced learning framework (PROTAC-RL) to navigate the sampling of molecules with desirable properties. The reinforced learning strategy was selected because property functions were usually complicated and non-differentiable.

Proformer enables general PROTAC generation. To evaluate our method, we followed previous fragment-linking works[32,33] to divide the PROTAC dataset into three sets with a ratio of 8:1:1 for training, validating and testing, respectively. For each testing pair of warhead and E3 ligand, the top ten candidates were generated and the percentage of ground-truth linkers generated among the compounds in the test set (that is, the recovery rate) was used to evaluate the performance. We compared Proformer with other state-of-the-art fragment-linking methods, including a graph-based method called Delinker[26], a language-based method called SyntaLinker[27] and their versions retrained on the PROTAC training set. As shown in Fig. 2a, Proformer achieved the highest recovery rate of 43.0 ± 1.8%, significantly higher than those achieved by Delinker (0.6 ± 0.6%) and SyntaLinker (0.6 ± 0%). These low recoveries by the original

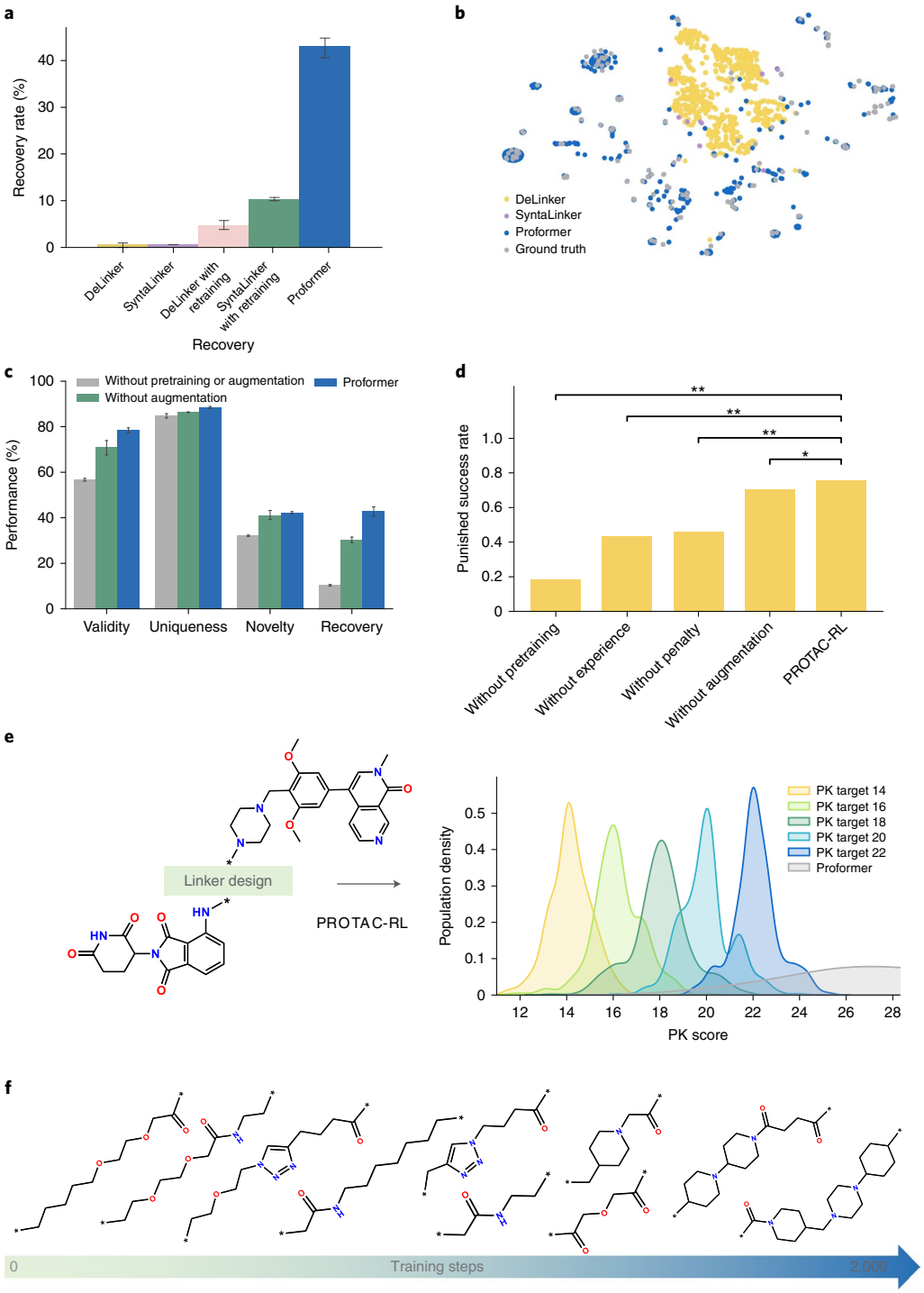


Fig. 2 | Performance of PROTAC-RL. a, Performance comparison of state-of-the-art fragment-linking models. Data presented as mean values. Error bars indicate s.d. ($n = 3$ independent runs). **b**, *t*-Distributed stochastic neighbour embedding (*t*-SNE) of ground-truth PROTACs in the test set and PROTACs generated using different models. **c**, Performance comparison of generative models with different training strategies in terms of validity, uniqueness, novelty and recovery on the fragment-linking task. Data presented as mean ± s.d. ($n = 3$ independent runs). **d**, Ablation study of PROTAC-RL models over average punished success rate on the conditional generation task (total $n = 10$ per group). Two-sided *t*-tests were used (the exact *P* values from top to bottom are 6.54×10^{-5}, 7.19×10^{-5}, 1.31×10^{-3} and 1.90×10^{-2}). $^{*}P < 0.05$. $^{**}P < 0.01$. **e**, The distributions of the PK score for structures sampled from agents towards a target PK score of 14, 16, 18, 20 or 22. The lower the PK score, the better. **f**, The evolution of the generated structures as the PK reward is increased.

Delinker and SyntaLinker models suggest that models trained on small molecules cannot reproduce the chemical space of PROTACs. The projection of the generated molecules (Fig. 2b) indicates that Proformer covered a similar space to the real PROTACs (ground truth), while baseline models are biased. After retraining Delinker and SyntaLinker using the PROTAC training datasets, their recovery rates increased sharply to 4.8% and 10.4%, respectively, but remained much lower than that achieved by our method (43.0%).

This difference could result from the advance represented by our proposed model, since the removal of the SMILES augmentation component caused a decrease of 12.7% in the recovery rate and further removal of the pretraining component caused a decrease of 19.9% (Fig. 2c). In general, the superior performance of our model resulted from a balanced performance in terms of validity, uniqueness and novelty when compared with other methods (Supplementary Table 1).

RL enables conditional PROTAC generation. Another challenge in PROTAC development is the poor PK properties of the designed molecules. To achieve the desired PK scores (Methods), we employed the RL strategy, and introduced the memory experience mechanism[38] and repeat score penalty[39] to alleviate the collapse problem during RL. We used an empirical PK objective function (the lower the better, see Supplementary Methods) and aimed to decrease the average PK score from 25.6 to 14.0, a reasonable PK value for PROTAC molecules. Since our RL is not strongly dependent on specific PROTAC molecules, we randomly selected ten pairs of warheads and E3 ligands from the PROTAC test set. The top 100 candidates per pair were generated after 2,000 steps of RL. The output PROTAC molecules are considered to be successful if they are close to (±2.0) the expected target score. As shown in Fig. 2d, the PROTAC-RL agent model generated molecules with an average PK score of 14.5, with an average punished success rate (percentage of successfully generated unique compounds) of 75.5%. These high success rates result from the synergy of four modules (pretraining, augmentation, memory experience and penalty), while the removal of any module would cause a notable decrease in the success rate. Removal of the pretraining module decreased the performance of the model by five fold (15.0%), probably because the model trained solely on the small training set did not contain enough unique molecules. This reaffirms the importance of pretraining for learning in low-resource scenarios. Similarly, the removal of the penalty or memory experience module decreased the punished success rate to 37.2% and 39.5%, respectively, probably because the modules incremented the global searching ability to avoid becoming trapped in local minima.

As an example, we selected a pair of warhead (BI-7273, targeting BRD7) and E3 ligand (CRBN (Cereblon) recruiter, thalidomide), and aimed to generate PK scores of around 22, 20, 18, 16 and 14. We did not minimize the PK score since a low PK score tends to produce short linkers that are less likely to form appropriate ternary complex structures. The different thresholds were selected to retain a diversity of generated compounds. As shown in Fig. 2e, at these five thresholds, our sampled compounds were mostly distributed around the expected PK scores (±2.0) with percentages of 97.4%, 94.2%, 88.8%, 95.2% and 95.4%, respectively. In contrast, the prior model sampled only <0.5% of compounds with PK scores <18. The trends were similar for the other nine selected pairs (Supplementary Fig. 2). Figure 2f illustrates the evolution of the generated structures along the model RL. The structures become more complex, with increasing numbers of rings and substituents.

To further evaluate the generalizability of our RL-based architecture, we also used octanol–water partition coefficient (log *P*) and linker length as the reward functions (Supplementary Figs. 3 and 4, respectively), again proving that PROTAC-RL was able to guide the prior to generate molecules with desired properties.

In silico PROTAC virtual screening against BRD4. To further indicate our method, we selected BRD4 as a case study because it has emerged as an attractive target for anti-cancer therapy while no clinical PROTAC molecules have been reported[40]. Although Winter et al.[41] developed a BRD4 PROTAC degrader (dBET6) using a bromodomain inhibitor JQ1 (PubChem 46907787) and recruitment of the E3 ubiquitin ligase CRBN (UniProt Q96SW2), this compound faces problems in terms of permeability and PK (with a PK score of 22.24). Here, we aimed to obtain BRD4 substitutes with better PK by launching the RL stage with PK score targets of 20, 18, 16, 14 and 12. We limited the analysis to linkers containing 2–16 atoms since Nowak et al. found this length to be optimal for BRD4 PROTAC activity[10]. During RL, we saved a model every 250 steps between the 1,000th and 2,000th steps and used each of the four models to generate 250 linkers. Finally, we obtained a total of 5,000 PROTACs, containing 4,860 valid SMILEs strings and 2,894 unique ones.

As no standardized protocol exists for screening PROTAC candidates, we designed a comprehensive screening strategy to narrow down the molecules (Fig. 3a). First, to ensure the novelty of the generated PROTACs, we removed molecules with Tanimoto similarity of the Extended-Connectivity Fingerprints (ECFP) fingerprint[42] above 0.6 to any BRD4 PROTAC in the training set (2,645 molecules left), and further filtered the PROTACs containing unfavourable substructures in the Pan-assay interference compounds (PAINS)[43] filters (1,992 molecules left). Then, the remaining PROTACs were clustered according to their Tanimoto similarity using the Butina algorithm[44], obtaining 78 clusters. The average internal Tanimoto similarity of these 1,992 PROTAC linkers is 0.184, indicating that they are not similar to each other. At the same time, we scored all molecules through a PROTAC degradation activity predictor, which was trained on the DC_{50} (the concentration at which 50% of the target protein has been degraded) and degradation percentage data according to the ensemble of four heterogeneous machine learning models, achieving a three-fold area under the receiver operating characteristic curve of 0.945, 0.938 and 0.936, respectively (Fig. 3b; see Supplementary Information for the detailed implementation). Note that the predictor may not generalize well owing to the small number of training samples. Hence, we only used the predictor to filter low-scored molecules, and designed a hierarchical strategy (detailed in the Supplementary Information) to maintain the diversity and high scores (Fig. 3c). The remaining 51 candidates have a better distribution of PK properties (lower PK score) than the originally generated PROTACs (Fig. 3d).

For the selected 51 candidates, we modelled PROTAC-mediated ternary complexes by combining the PRosettaC protocol[45] with molecular dynamics modelling (Methods). We observed that the calculated Rosetta scores and Molecular mechanics with generalised Born and surface area solvation (MM/GBSA) scores for five known BRD4 PROTACs did not always agree (Fig. 3e). Among the selected candidates, 34 showed better Rosetta scores while 3 showed better MM-GBSA scores than dBET6 (Fig. 3f), and the union set of the two resulted in 36 PROTACs. According to their synthetic accessibility, we selected six candidates for experimental validation. The projection of all generated PROTACs and active PROTACs (Fig. 3g) revealed that the six selected compounds are structurally diverse and show low similarity to known active PROTACs.

Wet laboratory characterization. After in silico screening, we identified six lead candidates, and by day 39, these molecules had been successfully synthesized (Fig. 4a–c and Supplementary Fig. 5). We first conducted a primary screening to assess the protein degradation ability on the HEK293T cell line by western blot assay. As shown in Fig. 4d–f, compounds 1–3 showed activity for the reduction of BRD4 protein. Specifically, compound 1 down-regulated the BRD4 level at a concentration of 100 nM, while compounds 2 and 3 induced BRD4 degradation at 300 nM. Compounds 4–6 did not show obvious degradation activity (Supplementary Fig. 5). We next selected compounds 1–3 to examine the in vitro anti-proliferative effect against Molt4 cells. As shown in Fig. 4g–i, these compounds showed different degrees of anti-proliferative activity against the Molt4 cell line, with compounds 1, 2 and 3 showing the concentration of an inhibitor needed to inhibit a biological process or response by 50% (IC_{50}) values of 116 nM, 5.1 μM and 21 μM, respectively. Apart from activity, in vitro inhibition of the human ether-a-go-go-related gene (hERG) channel has been used as an important assay to assess the potential

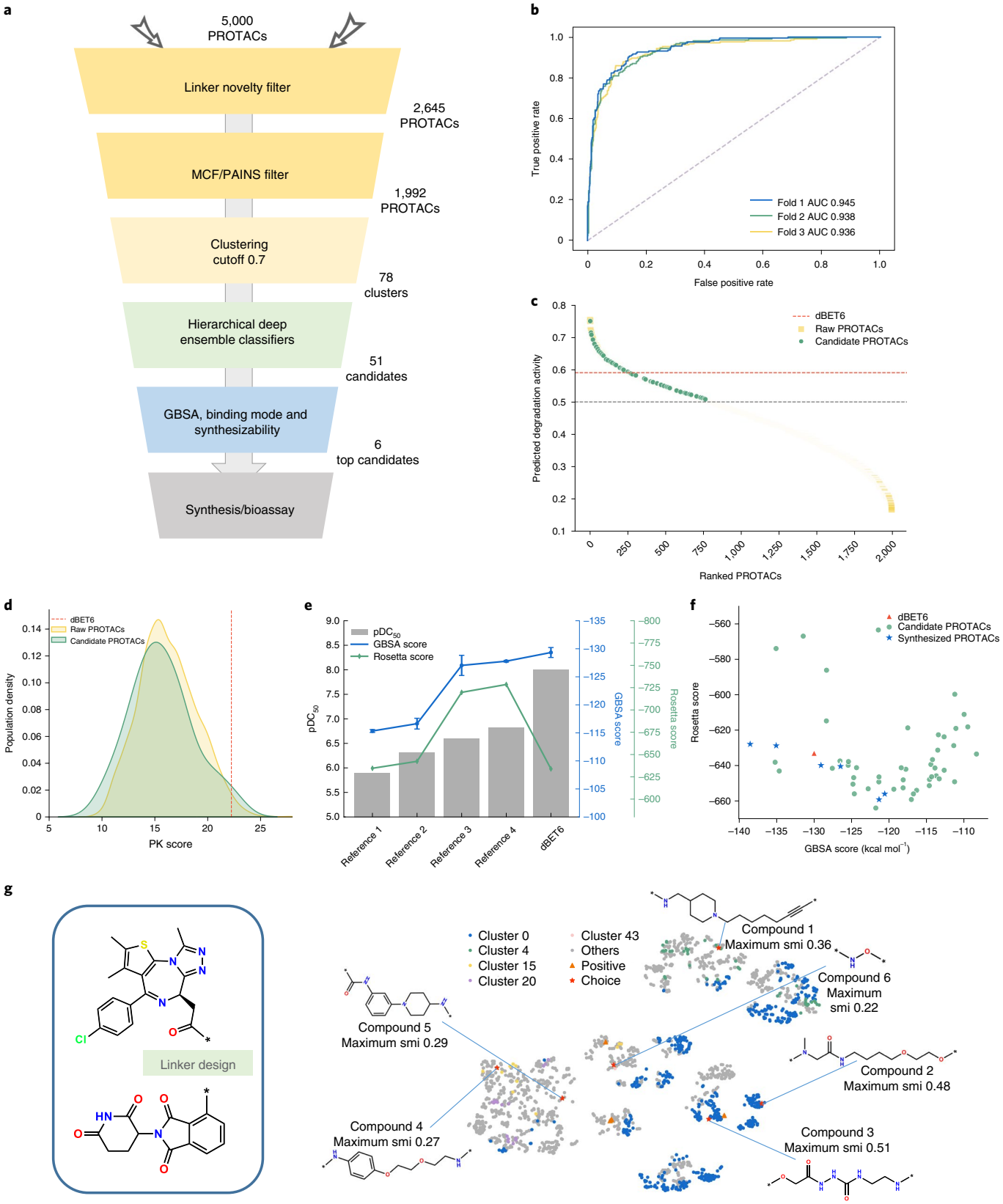
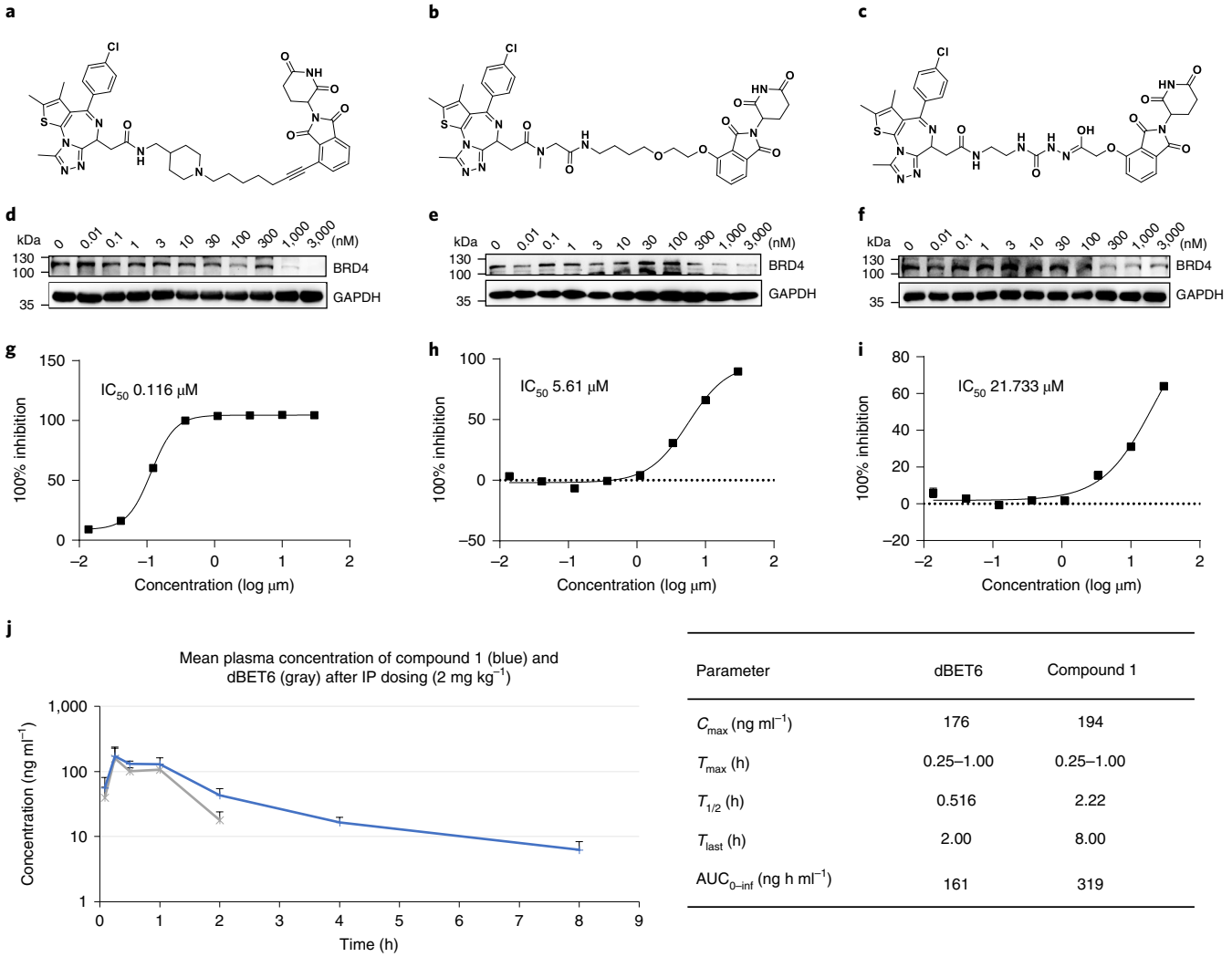


Fig. 3 | In silico post-generation screening. a, In silico post-generation screening work flow. **b**, Receiver operating characteristic curve plot evaluating ensemble activity predictor performance with three-fold cross-validation. **c**, Rank-ordered BRD4-degradation activity prediction scores of raw generated PROTACs and candidate PROTACs. **d**, Distributions of PK scores of generated PROTACs and candidate PROTACs. **e**, Retrospective study of pRosettaC (Rosetta score) and molecular dynamics simulations (GBSA score) with known active and inactive BRD4 PROTACs, showing the average and s.d. from three independent runs. **f**, Scatter plot of Rosetta scores and GBSA scores for candidate PROTACs, synthesized PROTACs and reference compound. **g**, t-SNE of molecules from the whole generated set (grey), the known positive BRD4 PROTACs (orange), the representative clusters (0–43 (colored with blue, green, yellow, light yellow and purple) and the synthesized PROTACs (red). MCF, medicinal chemistry filters; pDC_{50}, log half-maximal degradation concentration; smi, similarity.



Parameter	dBET6	Compound 1
C_{max} (ng ml^{-1})	176	194
T_{max} (h)	0.25–1.00	0.25–1.00
$T_{1/2}$ (h)	0.516	2.22
T_{last} (h)	2.00	8.00
AUC_{0-inf} (ng h ml^{-1})	161	319

Fig. 4 | Bioactivity and PK characterization. a–c, Structures of compound 1 (**a**), 2 (**b**) and 3 (**c**). **d–f**, Immunoblot for BRD4 and GAPDH following 3 h drug incubation with compound 1 (**d**), 2 (**e**) and 3 (**f**). **g–i**, Cell viability of Molt4 cells treated with compound 1 (**g**), 2 (**h**) and 3 (**i**) for 72 h, using CellTiter-Glo (mean ± s.e.m. for $n = 3$ replicates). **j**, PK studies performed in CD1 mice following 2 mg kg^{-1} intraperitoneal injection of compounds formulated in capitol/water (mean ± s.e.m. for $n = 3$ replicates). C_{max}, the maximum observed concentration; T_{max}, time to maximum plasma concentration; $T_{1/2}$, half-life; T_{last}, time of last measurable concentration; AUC, total area from zero to infinity.

cardiotoxicity of drug molecules. We evaluated compound 1 for its inhibition of the hERG channel and found a low hERG inhibition of 27.4 μM (Supplementary Fig. 7), which proves that compound 1 is a strong candidate for further development. Moreover, compound 1 showed favourable physicochemical properties, with water solubility Log *S* of 1.42 and distribution co-efficient (Log *D*) of 3.27.

Because PROTAC-RL is mainly employed to optimize the PK properties of PROTACs, we further tested compound 1 in a rodent model, where compound 1 was delivered to mice intraperitoneally at 2 mg kg^{-1}. The three administrations resulted in a similar half-life of ~2.22 h (Fig. 4j). Intraperitoneal administration conferred a peak plasma concentration (C_{max}) of 194 ng mL^{-1} on initial delivery, which peaked at 0.5 h after delivery. This is much better than the reference compound dBET6, with a C_{max} of only 176 ng mL^{-1} and half-life of 0.52 h. These results distinctly highlight compound 1 as a potent BRD4 degrader with favourable PK. A summary of the properties of compound 1 is provided in Supplementary Table 6.

Structural and mechanism analyses. To understand the mechanistic basis of the degradation, we further compared the ternary complexes between dBET6 and compound 1. As shown in Fig. 5a–c, compound 1 maintained the same major interactions as dBET6, including two crucial hydrogen bonds between the triazole in the BET warhead and BRD4 CYS470 and ASN474, as well as the hydrogen bonds between the glutarimide in the CRBN binder and the CRBN HIS326 and SER327 backbone. This demonstrates that compound 1 could maintain a similar binding mode to the reference compound. Note that compound 4 showed a similar binding mode to the reference and a better predicted degradation activity (according to both the GBSA and pRosettaC values) but without obvious degradation activity. To interpret the potential reasons for this, we performed steered molecular dynamics simulations to evaluate the stability of the PROTAC-mediated ternary complexes for compounds 1 and 4. Interestingly, the complex formulated by compound 1 was shown to be more stable than that by compound 4 (Fig. 5d), which probably explains why compound 4 did not show bioactivity in wet experiments.

Discussion

We report herein a fully automated computational framework that combines conditional generative modelling, machine learning and physics-driven learning for rational de novo design of PROTACs,

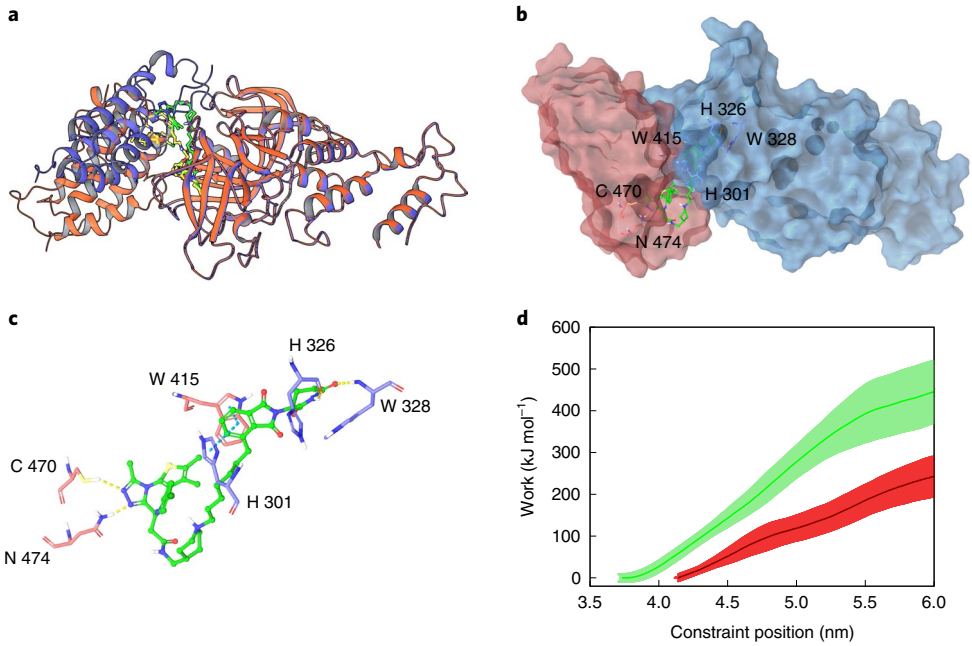


Fig. 5 | Binding mode analysis and atomistic simulation. a, The overlap of the compound 1 complex and the reference dBET6 crystal structure 6BOY. **b**,**c**, Ternary complex (**b**) and interaction mode (**c**) of compound 1. **d**, Pulling work analysis along with the relative position between the centre of PROTAC and that of BRD4 and CRBN, showing the mean results for the complex built by compound 1 (green) and compound 4 (red). Shading shows the error of ten replicas calculated using the s.d.

and experimentally validated these PROTACs for favourable anti-cancer efficacy and PK. The discovered PROTACs show high degradation activity against BRD4, a gene that is overexpressed in different human cancers. Wet laboratory results confirmed the efficiency of the proposed approach for designing new and optimized PROTACs with a very modest number of candidate compounds synthesized and tested. The present design approach in this proof-of-concept study yielded a 50% success rate and a rapid turnaround of 49 days, highlighting the importance of combining artificial intelligence-driven computational strategies with experiments to achieve more-effective drug candidates.

As the proposed model PROTAC-RL is a generic method, it is suitable for a wide range of conditional generation tasks and can handle multiple objective functions simultaneously. Thus, future directions of this research will explore the effect of additional relevant constraints, such as the cell permeation ability or oral bioavailability, on the designed PROTACs using the approach presented herein.

There are also several potential limitations of PROTAC-RL under the current framework. First, as an RL approach, the design of the reward function is crucial to the outputs. However, as there are not enough experimental data for PROTAC property prediction, the choice of robust predictor would be limited. One option is to use some semi-experienced scoring functions, as shown in this study. Another alternative is to apply a docking scoring function to guide the model to generate potential candidates. Secondly, current ternary complex modelling still relies on physics-driven simulations. As our results show, the best results obtained by docking are not necessarily active, and steered molecular dynamics was proven to be helpful. The requirement for extensive modelling simulations may be facilitated by the rapid development of supercomputers. Another alternative worth trying would be to use recent deep learning-based binding prediction methods[46,47]. Third, attachment sites fundamentally affect the degradation and selectivity[12]. PROTAC-RL currently follows the settings of previous methods that predefine the attachment sites according to prior bioactivity data. When crystal structures are available, the solvent-exposed positions identified from those high-resolution co-crystal structures could be determined as attachment sites[48,49]. Otherwise, suitable attachment points may be identified by discovering data for the structure–activity relationship and combining the inspection of explainable artificial intelligence[50] and molecular dynamics simulations. Finally, advances in protein structure prediction and protein–protein interaction prediction may enable more accurate modelling of ternary complex structures[51,52].

Overall, our results suggest that the time is ripe for the application of modern deep learning approaches to PROTAC discovery. Such efforts could increase the rate at which new molecular entities are discovered, decrease the resources required to identify these molecules and decrease the associated costs. We believe that the strategies described in our work will inspire future PROTAC design works.

Methods

Datasets. *PROTAC dataset.* We obtained PROTAC molecules from PROTAC-DB[34], a public, web-accessible database. It collects 1,656 PROTACs from literature and patents, each consisting of a fragment molecule triplet (FMT; warhead, linker and E3 ligand). The dataset included 202 unique warheads, 65 E3 ligands and 806 linkers. The PROTACs were randomly divided into three parts in a ratio of 8:1:1 for training, validating and testing, respectively.

As Arús-Pous et al.[37] indicated that training models with randomized SMILES are generalized to a larger chemical space, we adopted the same approach to augment each FMT (corresponding to the warhead, linker and E3 ligand triplet) in the training set to five FMTs, where one was original, one FMT was a simple exchange of warhead and the E3 ligand with a corresponding change of the linker and three FMTs corresponded to randomized SMILES for both input and output. After this augmentation, the training set was expanded to 6,630 FMTs.

ZINC quasi-PROTAC dataset. Because of the small number of PROTAC molecules available, we culled molecules with similar properties from the ZINC[53] dataset (https://pubs.acs.org/doi/10.1021/acs.jcim.5b00559) for pretraining our model. We selected molecules with molecular weight greater than 500 and further filtered them based on PROTAC properties reported in ref. [54]. Then, each molecule was first cut into three parts using the matched molecular pairs cutting algorithm as proposed by Hussain et al.[55]. Briefly, the algorithm first enumerates all double cuts of nonfunctional groups and acyclic single bonds and transforms each compound into FMTs. The FMTs were then filtered according to PROTAC warhead, E3 ligand

and linker properties, including the molecular weight, the number of carbon atoms, the number of aromatic rings (NAR), the number of donor atoms for H-bonds, the number of acceptor atoms for H-bonds, Kier's flexibility index and the topological polar surface area. Full details of the property filters are provided in the Supplementary Information. Since one molecule might generate multiple fragment triplets, we only kept triplets with the longest linker. Finally, 294,675 unique triplets remained. By using the same augmentation strategy, the training dataset was expanded to 1,178,685 FMTs, namely the ZINC quasi-PROTAC dataset.

PROTAC-RL framework. To generate PROTACs with a given pair of warhead and E3 ligands, we have proposed an RL-based framework. The method consists of two modules: the Proformer module to generate valid, diverse and novel PROTACs and the RL module to navigate PROTACs with desirable properties.

Proformer architecture. As SyntaLinker using SMILES representation was shown to outperform DeLinker using the topological graph, we followed SyntaLinker to input with molecule SMILES. The PROTAC generation was formulated as a sentence completion process, where the input is the SMILES of the warhead and E3 ligands and the output is the SMILES of PROTACs. The fundamental architecture of Proformer is a transformer neural network (Supplementary Fig. 1) which contains multiple encoder–decoder modules[36]. Each encoder layer comprises a multi-head self-attention sub-layer and a position-wise feed-forward network sub-layer. In multi-head attention, several scaled dot-product attention functions work in parallel, allowing the model to concentrate on information from different subspaces at different positions. An attention function computes the dot products of the query ($\mathcal{Q}$) with all keys ($\mathcal{K}$), introducing a scaling factor d_k (equal to the size of the weight matrices) to avoid excessive dot products, and then applies a softmax function to the weights on the values (V). Formally,

$$\text{Attention}\left(\mathcal{Q},\mathcal{K},\mathcal{V}\right)=\text{softmax}\left(\frac{\mathcal{Q}\mathcal{K}^{\intercal}}{\sqrt{d_k}}\right)\mathcal{V} \tag{1}$$

Where $\intercal$ indcates transposition. The feed-forward network sub-layer is activated by the Rectified Linear Unit (ReLU) function[56]. Then, two sub-layers above were linked by layer normalization[57] and a residual connection[58]. Each decoder layer consists of three sub-layers, including a feed-forward network sub-layer and two attention sub-layers. One decoder attention layer is the self-attention sub-layer, and to hinder attending to unseen future tokens, a mask function is utilized in this layer. The other is the encoder–decoder attention layer. It helps the decoder to capture the relationship between the encoder and decoder and to concentrate on crucial parts in the source sequence.

For a given source sequence (the SMILES of the pair of warhead and E3 ligand), its input is processed by one-hot encoding, followed by embedding into a latent representation $\mathcal{H}$. Given $\mathcal{H}$, the decoder output is then normalized with a softmax function, yielding a probability distribution for sampling tokens, iteratively generating an output sequence (SMILES of PROTAC) $Y=(y_1,\ldots,y_m)$ until the ending token '</s>' is occurred. During the training stage, the goal of the model is to minimize the cross-entropy loss $\mathcal{L}$ between the target sequence $M_t=(t_1,\ldots,t_k)$ and the output sequence Y:

$$\mathcal{L}\left(Y,M\right)=-\sum_{i=1}^{k}y_i\log t_i \tag{2}$$

RL. By using Proformer as the prior model, we navigate the generating PROTACs with desired properties through a Markov decision process, where the agent must decide which SMILES character to choose at the current state to maximize the target reward function. Similar to the work of Olivecrona et al.[59], we utilized the probability distributions learned by the prior model as our initial policy. The generation task is finitely episodic. Given a source sentence $x=(x_1,x_2,\ldots,x_n)$, with the Transformer network processing, the decoder finally samples a sentence $\widehat{\mathcal{Y}}=(\mathcal{Y}_1,\mathcal{Y}_2,\ldots,\mathcal{Y}_m)$ of word tokens and ends when the 'EOS' token is sampled. As the episode has been determined, four basic elements in the RL can be recognized: the actions $\mathrm{A}=(a_1,a_2,\ldots,a_m)$ represent the SMILES token sampled from the dictionary, the states $\mathrm{S}=(\mathrm{s}_1,\mathrm{s}_2,\ldots,\mathrm{s}_m)$ are the hidden states which encode the information of the input SMILES and tokens that have been sampled before, the policy π can be represented by a model distribution of sampling learned from data, and the reward R is calculated by a given scoring function after the whole sequence has been sampled. The product of the action probabilities represents the model likelihood of the sequence, where T = steps, formed as

$$\mathrm{P}\left(A\right)=\prod_{t=1}^{T}\pi\left(a_t|x\right). \tag{3}$$

Once the agent generates a complete sequence $\widehat{\mathcal{Y}}$, a terminal reward $R\left(\widehat{\mathcal{Y}}\right)$ is calculated by scoring functions as

$$R\left(\widehat{\mathcal{Y}}\right)=\begin{cases}\max\left(0,1-\frac{1}{\alpha}\text{abs}\left(f\left(\widehat{\mathcal{Y}}\right)-\text{Target}\right)\right),\ \text{if valid;}\\ 0,\ \text{if invalid.}\end{cases} \tag{4}$$

where α is a parameter and $f\left(\widehat{\mathcal{Y}}\right)$ is changeable for each task such as calculating the linker length, log P and the PK score. For rating the desirability of the sequences, max methods are used to make $R\left(\widehat{\mathcal{Y}}\right)\in[0,1]$. Note that the reward $R\left(\widehat{\mathcal{Y}}\right)$ here is a sentence-level reward. So, the goal now is to upgrade the agent policy π in such a way as to increase the expected score for the generated sequences. To keep the learned syntax of SMILES and the distribution of molecular structure in previous datasets, we would like to anchor our new policy to the prior policy. We therefore introduce an augmented likelihood $\log P\left(A\right)_{\mathbb{U}}$ as a sum with prior likelihood and modulated reward of a sequence:

$$\log P\left(A\right)_{\mathbb{U}}=\log P\left(A\right)_{\text{Prior}}+\sigma R\left(\mathcal{Y}\right), \tag{5}$$

where σ is a scalar coefficient. As the goal of the agent is to learn a policy that maximizes the expected return while anchoring to prior policy, we achieved this by minimizing the cost function $J\left(\Theta\right)=-G$.

$$G\left(A\right)=-\left[\log P\left(A\right)_{\mathbb{U}}-\log P\left(A\right)_{\mathbb{A}}\right]^2, \tag{6}$$

which measures the squared difference of the current agent's likelihood over a set of actions A, $\log P\left(A\right)_{\mathbb{A}}$, and the augmented likelihood.

To increase the diversity of the sampled molecules, the score of each generated molecule $\widehat{\mathcal{Y}}_i$ was weighted by assigning a weight of $\frac{1}{\sum\widehat{\mathcal{Y}}_i}$, where $\sum\widehat{\mathcal{Y}}_i$ is the number of repeating molecules $\widehat{\mathcal{Y}}_i$. Experience replay[38,59] was introduced to lead the training process. For each generation batch, a certain number of memories were sampled from the experience buffer with the normalized score as a probability distribution, joining the backward of each step.

Transfer learning for low-resource training. Due to the small number of PROTAC molecules, we trained the Proformer model on the ZINC quasi-PROTAC dataset with a chemical space similar to the PROTAC dataset, then fine-tuned the model on the PROTAC dataset. This strategy, also known as transfer learning[18,60], can allow the model to explore a broader chemical space to learn the SMILES syntax and the relation between input and output, such as the linking site and generated linker, assisting our model to generate more valid and diverse PROTACs. At the same time, the fine-tuning on the real PROTAC molecules guides the learned distribution to the real PROTAC chemical space.

Evaluation metrics. We assessed the generated molecules by using a range of metrics:[32,33]

- **Recovery** means the percentage of ground truth generated among test-set compounds.
- **Validity** refers to the percentage of generated chemically valid molecules with the input of fragments.
- **Uniqueness** is the average percentage of unique linkers in generation for each pair of fragments.
- **Novelty** is the percentage of generated validly unique PROTACs with novel linkers (not present in the training set).

Reward functions. We defined three reward functions, that is, linker length, log P and PK score, from easy to hard.

Linker length is one of the most important properties involved in PROTAC selectivity, the ability to cross the plasma membrane and metabolism[10,48,61,62]. In real-world PROTAC design, drug chemists would first determine the optimal linker length and design linkers according to the goal[8]. To mimic this process and generate linkers of any specific length, the scoring function was defined as

$$S_{\text{length}}\left(X\right)=\begin{cases}\max\left(0,1-\frac{1}{\alpha}\text{abs}\left(\text{LinkerLen}\left(X\right)-\text{Target}\right)\right),\ \text{if valid;}\\ 0,\ \text{if invalid.}\end{cases} \tag{7}$$

α denotes a changeable parameter. 'Target' is a user-predefined targeted linker length.

Log *P* is a basic physicochemical property for PROTACs. A low log P will limit the utility of in vitro data for establishing in vitro–in vivo correlations for the key drivers of achieving acceptable oral PK[63]. We defined the scoring function as

$$S_{\log P}\left(X\right)=\begin{cases}\max\left(0,1-\frac{1}{\alpha}\text{abs}\left(\log P\left(X\right)-\text{Target}\right)\right),\ \text{if valid;}\\ 0,\ \text{if invalid.}\end{cases} \tag{8}$$

PK score. PROTACs often have molecular weight above 700 and exhibit other violations of the Lipinski RO5[64] and/or Veber's rotatable bond/polar surface area rules[65], in a chemical space lying beyond traditional drugs. DeGoey et al.[66] described a composite AbbVie multiparametric score (AB-MPS) metric to calculate oral absorption via *calculated LogP* (CLog P), the number of aromatic rings and the number of rotatable bonds (NRotB). The formula was defined as

$$\text{PK}\left(X\right)=\text{AB-MPS}=\text{abs}\left(\log P\left(X\right)-3\right)+\text{NAR}+\text{NRotB}. \tag{9}$$

The lower the AB-MPS score, the more likely the compound is to be absorbed, and a value of ≤14 is reported to predict a higher probability of oral absorption[13]. However, a large gap in the AB-MPS of PROTACs sharing the same warhead and E3 ligand pair often results in a great difference in linker length and affects the PROTAC activity. To minimize the effect on PROTAC activity, we designed a PK scoring function with the target AB-MPS value as follows:

$$S_{\mathrm{PK}}(X) = \begin{cases} \max\left(0, 1 - \frac{1}{a}\mathrm{abs}\left(PK(X) - \mathrm{Target}\right)\right), & \text{if valid;} \\ 0, & \text{if invalid.} \end{cases} \quad (10)$$

where α is a changeable parameter. Because the AB-MPS formulation includes calculations of log P and the number of rotatable bonds, the PK optimization task can also be treated as a multi-objective optimization task. 'Target' is a user-predefined targeted AB-MPS score.

Molecular simulations. Since the generated PROTACs do not explicitly consider the ternary complexes with the target ligase and protein of interest, we followed the modelling protocol of PRosettaC[45]. In particular, we first downloaded the tertiary complex structure from the Protein Data Bank (ID: 6BOY) as a reference. From the complex structure, we removed the PROTAC linker fragment and divided the complex into two parts: the E3 ligase bound with the E3 ligand and the target protein with the warhead. To build candidate complex structures with different linkers, we first constructed complex models by protein–protein global docking using PatchDock[67], with a distance constraint between the binding pockets of the two ligands. The distance constraint range was determined according to the distribution of distances between two termini of the generated linker conformations (200 in our case). For each linker from this step, 500 complex models were generated. Then, each of the docked complex models was further refined through Rosetta protein–protein local docking and a relaxed protocol to generate ten refined models. For each refined complex model, the missing linker fragment was sampled between the fixed ligands to generate 100 conformations. All the generated conformations were clustered by using the DBSCAN method based on the pairwise r.m.s.d. of each two models, where the cutoff was set to 4 Å. From the largest cluster, the complex models were scored by using the Rosetta packing protocol, and the conformations with the best Rosetta score were selected for further simulations.

As the Rosetta score was indicated to be not strongly correlated with the protein degradation ability in previous studies[45], we further performed molecular dynamics (MD) simulations to capture the movement of mediated ternary complexes. For each generated PROTAC model, we performed three parallel MD simulations with different assigned initial velocities. Each MD simulation run for 50 ns and the last 10 ns (200 conformations) were extracted to calculate the MM-GBSA energy (after removing all water molecules). The 200 energy values were averaged to obtain the final binding affinity. The MD study was conducted by using DESMOND (Schrodinger release 2021-1). For each MD simulation, the mediated ternary complex is first prepared with the Optimized Potentials for Liquid Simulations 4 force field and solvated in the orthorhombic simple point-charge water box. The *NPT* ensemble at 300 K and 1.01325 bar was used in the production run. The time step is 2 fs, and the recording interval is set to 50 ps for trajectories and 1.2 ps for energy. All other parameters are set as default. The GBSA calculations were performed using the Schrodinger Thermal MM-GBSA tool.

To compare the stabilities of different PROTAC–protein complexes, we carried out steered MD simulations using Gromacs 2020.6. Enhanced sampling methods were supported by PLUMED-2.7.2. Only one collective variable centre-of-mass distance (Dcom) between proteins was set. With a force constant (K = 1,000 pN) along Dcom, the bounded proteins were pulled from the natural length of the native structure to a separated state of 6 nm in 4 ns. For each complex, ten rounds of pulling experiments were applied.

Cell viability assay. Cell growth inhibition studies were performed at ChemPartner (Shanghai). The Molt4 cell line was seeded in 96-well plates at a density of 2×10^4 cells per well and treated continuously with different concentrations of compounds 1–3. Cells were incubated with compounds for 72 h at 37 °C under 5% CO_2. Anti-proliferative effects were then assessed by using the CellTiter-Glo luminescent cell viability assay (Promega cat. no. G7570) according to the manufacturer's standards, measuring luminescence using an Envision plate reader. IC_{50} values were calculated using a non-linear regression curve fit in GraphPad Prism 6. For each treatment condition, N = 3 biological replicates were used.

Western blot studies. For western blot analysis, 2×10^6 cells per well were treated with compounds at the indicated concentrations at various times. Cells were collected and lysed in radioimmunoprecipitation assay buffer containing protease inhibitors. An amount of 20 µg of lysate was run in each SDS polyacrylamide gel electrophoresis lane and blotted onto polyvinylidene difluoride membranes. Antibodies for immunoblotting were BRD4 purchased from Bethyl Laboratories (Sangon, Shanghai, China), and glyceraldehyde 3-phosphate dehydrogenase from Proteintech (USA).

PK studies. PK studies were performed at WuXi AppTec (Shanghai). Briefly, two compounds (dBET6 and compound 1) were formulated in solubilizing vehicles (10% DMSO/10% solutol/80% water) and administered intraperitoneally to CD-1 mice in the cassette at 2 mg kg^{-1} each. The animal room environment was controlled (target conditions: temperature 18–26 °C, relative humidity 30–70%, 12 h artificial light/12 h dark). The study was permitted by WuXi AppTec Institutional Animal Care and Use Committee. Serial blood samples were collected at 0.083, 0.25, 0.5, 1, 2, 4, 8 and 24 h post dose and processed for plasma by centrifugation. Plasma drug concentrations were determined by liquid chromatography with tandem mass spectrometry (n = 3 mice per time point). The PK parameters were calculated by fitting the individual plasma concentration–time data to a non-compartmental model using Phoenix WinNonlin 6.3.

hERG channel inhibition assay. Compound 1 was tested for its in vitro effects on the electric current passing through hERG potassium channels stably transfected in a HEK 293 cell line to determine the concentration–response relationship for hERG inhibition by compound 1 using the manual patch-clamp technique. Compound 1 was tested at 0.3, 1, 3, 10 and 30 µM in duplicate. The assay was performed by Pharmaron (Beijing, China).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data used in this paper are publicly available and can be accessed at http://cadd.zju.edu.cn/protacdb/ for the PROTAC-DB dataset, https://zinc15.docking.org/ for the ZINC dataset and https://www.rcsb.org for the protein crystal structure. Source data are provided with this paper.

Code availability

Demo, instructions and codes for PROTAC-RL are available at https://github.com/biomed-AI/PROTAC-RL.

Received: 21 February 2022; Accepted: 28 July 2022;
Published online: 15 September 2022

References

1. Sakamoto, K. M. et al. Protacs: chimeric molecules that target proteins to the Skp1–cullin–F box complex for ubiquitination and degradation. *Proc. Natl Acad. Sci. U. S. A.* **98**, 8554–8559 (2001).
2. Deshaies, R. J. Prime time for PROTACs. *Nat. Chem. Biol.* **11**, 634–635 (2015).
3. Dale, B. et al. Advancing targeted protein degradation for cancer therapy. *Nat. Rev. Cancer.* **21**, 1–17 (2021).
4. Pettersson, M. & Crews, C. M. PROteolysis TArgeting Chimeras (PROTACs)—past, present and future. *Drug Discov. Today Technol.* **31**, 15–27 (2019).
5. Lai, A. C. & Crews, C. M. Induced protein degradation: an emerging drug discovery paradigm. *Nat. Rev. Drug Discov.* **16**, 101–114 (2017).
6. Bai, L. et al. A potent and selective small-molecule degrader of STAT3 achieves complete tumor regression in vivo. *Cancer Cell* **36**, 498–511. e417 (2019).
7. Liu, Z. et al. Design and synthesis of EZH2-based PROTACs to degrade the PRC2 complex for targeting the noncatalytic activity of EZH2. *J. Med. Chem.* **64**, 2829–2848 (2021).
8. Han, X. et al. Discovery of ARD-69 as a highly potent proteolysis targeting chimera (PROTAC) degrader of androgen receptor (AR) for the treatment of prostate cancer. *J. Med. Chem.* **62**, 941–964 (2019).
9. Zoppi, V. et al. Iterative design and optimization of initially inactive proteolysis targeting chimeras (PROTACs) identify VZ185 as a potent, fast, and selective von Hippel–Lindau (VHL) based dual degrader probe of BRD9 and BRD7. *J. Med. Chem.* **62**, 699–726 (2018).
10. Nowak, R. P. et al. Plasticity in binding confers selectivity in ligand-induced protein degradation. *Nat. Chem. Biol.* **14**, 706–714 (2018).
11. Bemis, T. A., La Clair, J. J. & Burkart, M. D. Unraveling the role of linker design in proteolysis targeting chimeras. *J. Med. Chem.* **64**, 8042–8052 (2021).
12. Smith, B. E. et al. Differential PROTAC substrate specificity dictated by orientation of recruited E3 ligase. *Nat. Commun.* **10**, 131 (2019).
13. Edmondson, S. D., Yang, B. & Fallan, C. Proteolysis targeting chimeras (PROTACs) in 'beyond rule-of-five'chemical space: recent progress and future challenges. *Bioorg. Med. Chem. Lett.* **29**, 1555–1564 (2019).
14. Garber, K. The PROTAC gold rush. *Nat. Biotechnol.* **40**, 12–16 (2022).
15. Cecchini, C., Pannilunghi, S., Tardy, S. & Scapozza, L. From conception to development: investigating PROTACs features for improved cell permeability and successful protein degradation. *Front. Chem.* **9**, 672267 (2021).
16. Vamathevan, J. et al. Applications of machine learning in drug discovery and development. *Nat. Rev. Drug Discov.* **18**, 463–477 (2019).

17. Ekins, S. et al. Exploiting machine learning for end-to-end drug discovery and development. *Nat. Mater.* **18**, 435–441 (2019).
18. Segler, M. H. S., Kogej, T., Tyrchan, C. & Waller, M. P. Generating focused molecule libraries for drug discovery with recurrent neural networks. *ACS Cent. Sci.* **4**, 120–131 (2018).
19. Gupta, A. et al. Generative recurrent networks for de novo drug design. *Mol. Inf.* **37**, 1700111 (2018).
20. Kotsias, P.-C. et al. Direct steering of de novo molecular generation with descriptor conditional recurrent neural networks. *Nat. Mach. Intell.* **2**, 254–265 (2020).
21. Zheng, S. et al. QBMG: quasi-biogenic molecule generator with deep recurrent neural network. *J Cheminform* **11**, 5 (2019).
22. Wang, J., Zheng, S., Chen, J. & Yang, Y. Meta learning for low-resource molecular optimization. *J. Chem. Inf. Model.* **61**, 1627–1636 (2021).
23. Zheng, S. et al. Deep scaffold hopping with multimodal transformer neural networks. *J Cheminform* **13**, 1–15 (2021).
24. Gomez-Bombarelli, R. et al. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Cent. Sci.* **4**, 268–276 (2018).
25. Skalic, M., Jimenez, J., Sabbadin, D. & De Fabritiis, G. Shape-based generative modeling for de novo drug design. *J. Chem. Inf. Model.* **59**, 1205–1214 (2019).
26. De Cao, N. & Kipf, T. MolGAN: an implicit generative model for small molecular graphs. Preprint at https://arxiv.org/abs/1805.11973 (2018).
27. Kadurin, A., Nikolenko, S., Khrabrov, K., Aliper, A. & Zhavoronkov, A. 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 (2017).
28. Walters, W. P. & Barzilay, R. Applications of deep learning in molecule generation and molecular property prediction. *Acc. Chem. Res.* **54**, 263–270 (2021).
29. Zhavoronkov, A. et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat. Biotechnol.* **37**, 1038–1040 (2019).
30. Das, P. et al. Accelerated antimicrobial discovery via deep generative models and molecular dynamics simulations. *Nat. Biomed. Eng.* **5**, 613–623 (2021).
31. Mason, D. M. et al. Optimization of therapeutic antibodies by predicting antigen specificity from antibody sequence via deep learning. *Nat. Biomed. Eng.* **5**, 600–612 (2021).
32. Imrie, F., Bradley, A. R., van der Schaar, M. & Deane, C. M. Deep generative models for 3D linker design. *J. Chem. Inf. Model.* **60**, 1983–1995 (2020).
33. Yang, Y. et al. SyntaLinker: automatic fragment linking with deep conditional transformer neural networks. *Chem. Sci.* **11**, 8312–8322 (2020).
34. Weng, G. et al. PROTAC-DB: an online database of PROTACs. *Nucleic Acids Res.* **49**, D1381–D1387 (2021).
35. Altae-Tran, H., Ramsundar, B., Pappu, A. S. & Pande, V. Low data drug discovery with one-shot learning. *ACS Cent Sci* **3**, 283–293 (2017).
36. Vaswani, A. et al. Attention is all you need. In Guyon, I. et al. (eds). *Advances in Neural Information Processing Systems*, 30 (2017). https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf
37. Arús-Pous, J. et al. Randomized SMILES strings improve the quality of molecular generative models. *J. Cheminform.* **11**, 1–13 (2019).
38. Wang, Z. et al. Sample efficient actor-critic with experience replay. Preprint at https://arxiv.org/abs/1611.01224 (2016).
39. Olivecrona, M., Blaschke, T., Engkvist, O. & Chen, H. Molecular de-novo design through deep reinforcement learning. *J Cheminform* **9**, 48 (2017).
40. ClinicalTrials.gov database, https://clinicaltrials.gov/
41. Winter, G. E. et al. BET bromodomain proteins function as master transcription elongation factors independent of CDK9 recruitment. *Mol. Cell* **67**, 5–18 (2017). e19.
42. Rogers, D. & Hahn, M. Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **50**, 742–754 (2010).
43. Baell, J. B. & Holloway, G. A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* **53**, 2719–2740 (2010).
44. Butina, D. Unsupervised data base clustering based on daylight's fingerprint and Tanimoto similarity: a fast and automated way to cluster small and large data sets. *J. Chem. Inf. Comput. Sci.* **39**, 747–750 (1999).
45. Zaidman, D., Prilusky, J. & London, N. PRosettaC: Rosetta based modeling of PROTAC mediated ternary complexes. *J. Chem. Inf. Model.* **60**, 4894–4903 (2020).
46. Paggi, J. M. *et al.* Leveraging nonstructural data to predict structures and affinities of protein-ligand complexes. *Proc. Natl Acad. Sci. U. S. A.* https://doi.org/10.1073/pnas.2112621118 (2021).
47. Zheng, S., Li, Y., Chen, S., Xu, J. & Yang, Y. Predicting drug–protein interaction using quasi-visual question answering system. *Nat. Mach. Intell.* **2**, 134–140 (2020).
48. Paiva, S. L. & Crews, C. M. Targeted protein degradation: elements of PROTAC design. *Curr. Opin. Chem. Biol.* **50**, 111–119 (2019).
49. Cheng, M. et al. Discovery of potent and selective epidermal growth factor receptor (EGFR) bifunctional small-molecule degraders. *J. Med. Chem.* **63**, 1216–1232 (2020).
50. Jimenez-Luna, J., Skalic, M., Weskamp, N. & Schneider, G. Coloring molecules with explainable artificial intelligence for preclinical relevance assessment. *J. Chem. Inf. Model.* **61**, 1083–1094 (2021).
51. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
52. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
53. Sterling, T. & Irwin, J. J. ZINC 15–ligand discovery for everyone. *J. Chem. Inf. Model.* **55**, 2324–2337 (2015).
54. Ermondi, G., Garcia-Jimenez, D. & Caron, G. PROTACs and building blocks: the 2D chemical space in very early drug discovery. *Molecules* **26**, 672 (2021).
55. Hussain, J. & Rea, C. Computationally efficient algorithm to identify matched molecular pairs (MMPs) in large data sets. *J. Chem. Inf. Model.* **50**, 339–348 (2010).
56. Nair, V. & Hinton, G. E. Rectified Linear Units Improve Restricted Boltzmann Machines. In Furnkranz, J. and Joachims, T. (eds) *Proceedings of the 27th International Conference on Machine Learning*, 807-814, (2010). https://icml.cc/Conferences/2010/papers/432.pdf
57. Ba, J. L., Kiros, J. R. & Hinton, G. E. Layer normalization. Preprint at https://arxiv.org/abs/1607.06450 (2016).
58. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Computer Society, 770–778 (2016).
59. Olivecrona, M., Blaschke, T., Engkvist, O. & Chen, H. Molecular de-novo design through deep reinforcement learning. *J. Cheminform.* **9**, 1–14 (2017).
60. Cai, C. et al. Transfer learning for drug discovery. *J. Med. Chem.* **63**, 8683–8694 (2020).
61. Burslem, G. M. et al. The advantages of targeted protein degradation over inhibition: an RTK case study. *Cell Chem. Biol.* **25**, 67–77. e63 (2018).
62. Goracci, L. et al. Understanding the metabolism of proteolysis targeting chimeras (PROTACs): the next step toward pharmaceutical applications. *J. Med. Chem.* **63**, 11615–11638 (2020).
63. Dressman, J. B. & Reppas, C. In vitro–in vivo correlations for lipophilic, poorly water-soluble drugs. *Eur. J. Pharm. Sci.* **11**, S73–S80 (2000).
64. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **23**, 3–25 (1997).
65. Veber, D. F. et al. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* **45**, 2615–2623 (2002).
66. DeGoey, D. A., Chen, H.-J., Cox, P. B. & Wendt, M. D. Beyond the rule of 5: lessons learned from AbbVie's drugs and compound collection: miniperspective. *J. Med. Chem.* **61**, 2636–2651 (2017).
67. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **33**, W363–W367 (2005).

Acknowledgements

This study has been supported by the National Key R&D Program of China (2020YFB0204803, Y.Y.), National Natural Science Foundation of China (61772566, Y.Y.) and Guangdong Key Field R&D Plan (2019B020228001, Y.Y.; 2018B010109006, Y.Y.). We thank R. Hu, W. Lu, L. Shi and J. Zhang for helpful discussions.

Author contributions

S.Z. and Y.Y. contributed the concept and experimental design. S.Z., Y.T. and C.L. contributed the code implementation. Z.W., X.S. and Y.T. contributed the development of the molecular simulations part. S.Z. and Q.Z. contributed to the wet experiment design. Y.Y., S.Z. and Y.T. wrote the manuscript. H.C participated in the discussion and revision of the manuscript. All authors contributed to the interpretation of the results. All authors reviewed and approved the final manuscript.

Competing interests

S.Z., Z.W., C.L., Z.Z. and X.S. work directly or indirectly for Galixir. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42256-022-00527-y.

Correspondence and requests for materials should be addressed to Yuedong Yang.

Peer review information *Nature Machine Intelligence* thanks Guowei Wei and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.