

Structure–function conditioned molecular generation with PocketGPT enables prospective discovery of PARP1 inhibitors

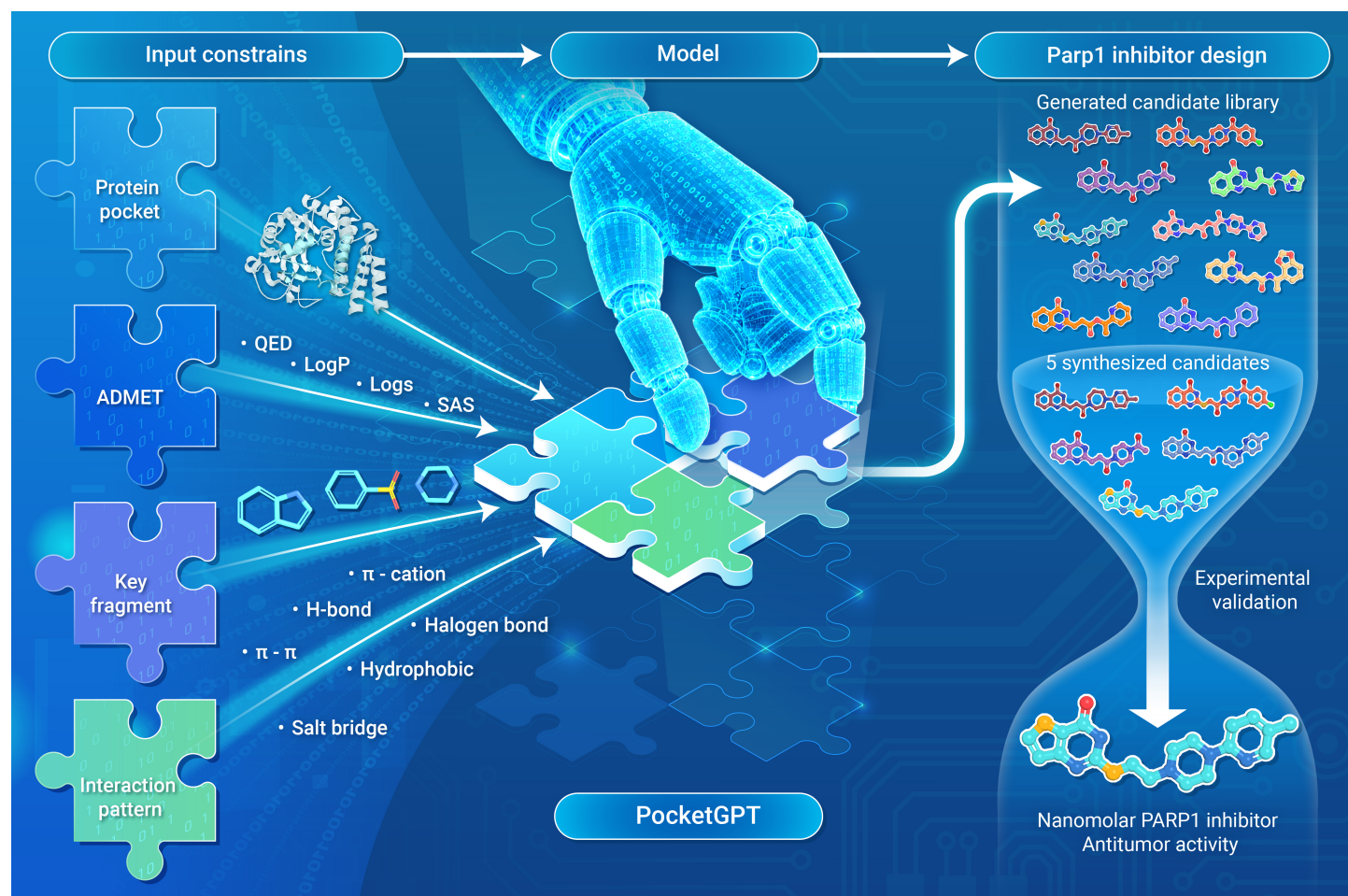
Rong Qian,^{1,2,4} Jilong Duan,^{1,4} Xiaoping Liu,^{1,4} Wenxuan Wang,¹ Tingfei Zhu,^{1,2} Dejun Jiang,¹ Jin Liu,^{2,3} Youchao Deng,^{1,*} Aiping Lyu,^{2,3,*} and Dongsheng Cao^{1,*}

*Correspondence: dengyouchao1206@163.com (Y.D.); aipinglu@hkbu.edu.hk (A.L.); oriental-cds@163.com (D.C.)

Received: March 30, 2026; Accepted: July 7, 2026; Published Online: July 15, 2026; <https://doi.org/10.59717/j.xinn-drugdisc.2026.100028>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- PocketGPT uses protein pocket and functional information to guide controllable molecule design.
- It integrates pocket context, molecular properties, fragments, and interactions to enrich promising candidates.
- A prospective poly(ADP-ribose) polymerase 1 (PARP1) study yielded an antitumor nanomolar inhibitor.

Structure–function conditioned molecular generation with PocketGPT enables prospective discovery of PARP1 inhibitors

Rong Qian,^{1,2,4} Jilong Duan,^{1,4} Xiaoping Liu,^{1,4} Wenxuan Wang,¹ Tingfei Zhu,^{1,2} Dejun Jiang,¹ Jin Liu,^{2,3} Youchao Deng,^{1,*} Aiping Lyu,^{2,3,*} and Dongsheng Cao^{1,*}

¹Xiangya School of Pharmaceutical Sciences, Central South University, Changsha 410013, Hunan, China

²School of Chinese Medicine, Hong Kong Baptist University, Hong Kong SAR 999077, China

³Institute of systems medicine and health sciences, Hong Kong Baptist University, Hong Kong SAR 999077, China

⁴These authors contributed equally to this work

*Correspondence: dengyouchao1206@163.com (Y.D.); aipinglyu@hkbu.edu.hk (A.L.); oriental-cds@163.com (D.C.)

Received: March 30, 2026; Accepted: July 7, 2026; Published Online: July 15, 2026; <https://doi.org/10.59717/j.xinn-drugdisc.2026.100028>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Qian R., Duan J., Liu X., et al. (2026). Structure–function conditioned molecular generation with PocketGPT enables prospective discovery of PARP1 inhibitors. *The Innovation Drug Discovery* 1:100028.

Structure-based molecular generation has become increasingly prominent in drug discovery; however, many generated binders still fail to satisfy the physicochemical and functional constraints required for pharmacological viability. Here, we present PocketGPT, a pocket-conditioned generative framework that demonstrates structure-aware molecular generation can be effectively enabled without explicit three-dimensional ligand construction, by jointly conditioning protein pocket context and medicinal-chemistry constraints during generation. By co-conditioning physicochemical property targets, fragment priors, and protein–ligand interaction fingerprints, PocketGPT enables structure–function co-generative control, allowing directional modulation of drug-like properties while partially retaining binding-relevant interaction patterns across diverse targets. In a prospective design–make–test campaign targeting poly(ADP-ribose) polymerase 1 (PARP1), PocketGPT identified a novel inhibitor with nanomolar binding affinity ($K_d = 4.88$ nM) and favorable tolerability in mice. Together, these results demonstrate that joint structural and pharmacological conditioning, embedded directly in the generative process, constitutes a viable and controllable representation for structure-aware molecular generation, offering a complementary pathway to geometry-first approaches for producing experimentally actionable chemical matter.

INTRODUCTION

The discovery of clinically viable small-molecule therapeutics remains a fundamentally inefficient and costly process, characterized by high attrition rates driven by challenges in efficacy, selectivity, and developability.^{1–3} Despite sustained advances in medicinal chemistry and computational modeling, identifying compounds that simultaneously achieve strong target engagement, favorable pharmacological properties, and translational feasibility continues to represent a central bottleneck in drug discovery. Recent progress in generative artificial intelligence has renewed interest in de novo molecular design as a means to accelerate early-stage drug discovery.^{4–8} Early generative approaches predominantly focused on ligand-based drug design (LBDD), learning statistical regularities from known bioactive compounds to propose novel analogues.^{9–14} While effective within well-characterized chemical spaces, such methods inherently lack explicit structural context and therefore face limitations when extrapolated to novel protein targets or unexplored binding environments.¹⁵ Structure-based drug design (SBDD) partially addresses this limitation by conditioning molecular generation on protein structures or binding pockets, enabling target-specific exploration.^{16–20}

A broad class of recent SBDD methods—including autoregressive graph-based models²¹ and three-dimensional diffusion^{22–25} generators—explicitly model ligand–pocket²⁶ complementarity and have achieved impressive geometric fidelity.^{27–29} However, geometric plausibility alone does not guarantee pharmacological viability.^{30,31} Molecules that are sterically compatible with a binding pocket may remain chemically unstable, synthetically inaccessible, or insufficiently selective, often necessitating extensive downstream filtering, rescoring, or redesign.³² Conversely, approaches that emphasize molecular properties or fragment-level priors may compromise the spatial precision required for high-affinity binding.^{33,34} This recurring tension highlights a

persistent challenge in generative drug design: reconciling structural compatibility with the implicit, multi-dimensional constraints imposed by medicinal chemistry.³⁵ A fundamental question then arises as to whether binding-relevant structural and functional constraints can be enforced during molecular generation without explicitly constructing three-dimensional ligand geometries.

Importantly, this limitation reflects a representation-level mismatch between the objectives optimized during molecular generation and the criteria ultimately used to evaluate translational viability. Many structure-based generative frameworks prioritize geometric complementarity as the primary organizing principle, while pharmacological considerations are frequently enforced only after candidate structures have been constructed. As a result, chemical validity, synthetic feasibility, and interaction specificity^{36,37} are treated as secondary constraints to be retrofitted, rather than emerging coherently during generation, which can limit robustness under realistic multi-constraint design settings.

In this work, we argue for a complementary generative formulation in which medicinal chemistry logic is incorporated directly into the generative process rather than applied as a post hoc correction. Under this formulation, molecular design is framed as a conditional generation problem in which structural context and pharmacological constraints are jointly interpreted during sequence construction.^{38–43} Crucially, explicit three-dimensional ligand geometries need not be generated; instead, binding-relevant structural patterns can arise implicitly from learned chemical regularities conditioned on pocket context. Based on this perspective, we introduce PocketGPT, a structure-aware molecular generation framework that models ligand design as conditional chemical language generation. PocketGPT integrates protein pocket information with multiple drug design constraints—including physicochemical properties, fragment priors, and protein–ligand interaction patterns—directly within a unified generative process. By learning these constraints jointly rather than reconciling them after structure generation, PocketGPT enables modular and composable control over molecular design while remaining compatible with structure-based discovery workflows.

We evaluate PocketGPT across representative structure-conditioned generation tasks, demonstrating its ability to balance geometric compatibility with chemically meaningful constraint satisfaction. Through systematic analyses of controllability, constraint interaction, and performance under increasing constraint complexity, we characterize how semantic conditioning reshapes the distribution of generated molecules. Finally, in a prospective application targeting poly(ADP-ribose) polymerase 1 (PARP1), PocketGPT identified potent and selective inhibitors that were subsequently validated in biochemical and cellular assays, providing experimental evidence that embedding medicinal chemistry logic as a generative prior can yield biologically actionable chemical matter.^{44,45}

RESULTS

An overview of the PocketGPT framework for structure–function–aware molecular generation

A central challenge in structure-based molecular generation is to learn a molecular design distribution that remains responsive to protein pocket context while simultaneously satisfying medicinal-chemistry objectives, with-

out relying on explicit three-dimensional ligand coordinate construction. Rather than treating structural awareness and pharmacological optimization as two sequential steps, PocketGPT formulates molecular generation as a unified conditional modeling problem, in which protein pocket information and downstream design requirements are embedded directly into the generative process as structured control variables (Figure 1).

Formally, PocketGPT learns an autoregressive conditional distribution over molecular sequences, in which ligand generation is conditioned on a heterogeneous set of signals comprising pocket-level structural representations, physicochemical property targets, fragment-level pharmacophoric priors, and protein–ligand interaction fingerprints. Specifically, the model aims to learn $p(x|C_{\text{pocket}}, C_{\text{prop}}, C_{\text{frag}}, C_{\text{lip}})$, where x denotes the molecular token sequence and the conditioning variables jointly specify the desired structural and functional design regime. Importantly, different constrained generation tasks were implemented as different conditioning settings of the same trained model, rather than as independently trained task-specific models. Pocket-only, property-controlled, fragment-conditioned, interaction-conditioned, and fully conditioned generation were obtained by activating the corresponding subset of conditioning variables while keeping the Transformer backbone and semantic-fusion architecture unchanged. The technical difficulty of this setting lies in the fact that these control signals differ substantially in modality, scale, sparsity, and semantic granularity: pocket descriptors encode continuous structural context derived from voxelized three-dimensional pocket geometry, property targets are low-dimensional numerical constraints, fragment annotations introduce discrete chemical priors, and interaction fingerprints provide sparse relational cues associated with binding patterns. A naive fusion of these inputs can easily impair controllability, destabilize generation, or dilute target-specific structural information.

To address this challenge, PocketGPT adopts a modality-aware conditioning architecture in which each control source is first encoded through a dedicated projection module and mapped into a shared latent space with aligned dimensionality. The resulting embeddings are then integrated into a unified semantic conditioning representation, which is prepended to the molecular token sequence as a prefix embedding and used to condition the Transformer backbone during autoregressive decoding.^{46,47} This prefix-based design enables heterogeneous structural and functional constraints to influence the hidden state trajectory from the earliest generation steps, allowing the model to learn a joint conditional distribution over molecules under composable and potentially competing design objectives, thereby reducing reliance on purely post hoc filtering or hard-constrained search.

The model is optimized in two stages to separate generic molecular language acquisition from target-aware conditional alignment. Large-scale pretraining on drug-like molecules establishes a broad chemical prior, capturing syntactic regularities and general patterns of chemically valid composition. Subsequent fine-tuning on protein–ligand complexes aligns this prior with pocket-specific structural semantics and medicinal-design signals, thereby transforming a general molecular language model into a controllable generator conditioned on biologically grounded design context. Because all control variables are introduced as continuous semantic prompts rather than explicit geometric supervision, PocketGPT operates entirely in the SMILES domain while retaining the capacity to integrate diverse sources of structural and pharmacological information within a single generative framework.

Directional semantic responses to isolated structural and functional conditioning

A long-standing question in structure-based molecular generation is whether binding-relevant structural and functional semantics can be activated through conditioning signals alone, without explicit geometric construction. To examine whether PocketGPT satisfies this criterion, we systematically evaluated its generative behavior under single-condition structural and functional prompts, without combining multiple objectives or applying post hoc filtering.

Pocket-derived structural context induces binding-relevant energetic shifts. Structural complementarity between a ligand and its binding pocket is a prerequisite for productive molecular recognition. To assess whether PocketGPT can leverage pocket geometry as a conditioning signal—rather than as

an explicit generative objective—we evaluated its performance on a held-out benchmark of 100 diverse protein–ligand crystal structures whose proteins and binding pockets were excluded from structure-conditioned fine-tuning.

We quantified the global impact of pocket conditioning by comparing molecules generated with pocket context against those produced under unconditioned generation. As shown in Figure 2A, the inclusion of pocket-derived semantic context resulted in a pronounced leftward shift in the docking score distribution. The mean docking score improved from −6.24 under the unconditioned baseline to −8.33 with pocket conditioning, indicating substantial enrichment of energetically favorable candidates. Docking scores⁴⁸ were used exclusively as a coarse proxy for binding-relevant geometric compatibility and were not optimized during training or generation. This analysis was used to assess whether the geometry-derived pocket embedding translated into shape-compatible generated molecules after docking-based pose evaluation.

Across the benchmark set, pocket conditioning consistently improved predicted binding energetics: 81% of targets exhibited lower average docking scores relative to the unconditioned baseline (Figure 2B), and 87 of 100 targets showed statistically significant improvements. Comparable trends were observed when de novo generated molecules were evaluated against biologically validated reference ligands for representative targets (PDB IDs: 1U5C, 2GSQ, 3RAS, and 5WUU), where a subset of generated molecules achieved docking scores on par with co-crystallized ligands (Figure 2C).

For a small subset of targets, including 3AC1 and 6MEP, the effect of pocket conditioning was attenuated. Structural inspection revealed that these pockets are characterized by extreme shallowness or high solvent exposure, limiting the discriminative power of geometry alone. These cases delineate practical boundaries of pocket-conditioned generation and are discussed in Supplementary Note 2.

Physicochemical objectives elicit directional global property responses. Beyond structural context, lead optimization requires precise control over global physicochemical properties to balance potency, safety, and developability. In PocketGPT, such objectives are introduced as semantic conditioning signals, enabling direct modulation during generation rather than post hoc selection.

We first examined single-objective conditioning for four key descriptors: lipophilicity (LogP), drug-likeness (QED), topological polar surface area (TPSA), and synthetic accessibility (SAS). For each property, target values spanning biologically relevant ranges were specified. As shown in Figure 3A, the resulting property distributions formed sharp, well-separated peaks centered near the desired setpoints, with minimal overlap across distinct targets (e.g., LogP = 2.0, 4.0, and 6.0). Comparable behavior was observed for QED, TPSA, and SAS.

These distributions indicate that numerical property prompts induce directional and controllable shifts in chemical space, rather than broad statistical biasing. When multiple physicochemical objectives were specified simultaneously, PocketGPT preserved this directional control across properties, yielding structured and well-separated distributions in the joint property space (Figure S8). Importantly, property modulation was achieved without explicit filtering, demonstrating that global physicochemical objectives are internalized as semantic guidance signals during generation.

Fragment-level conditioning anchors local functional motifs. Medicinal chemistry frequently exploits privileged fragments as functional anchors rather than rigid templates, allowing chemically meaningful variation while preserving biological activity. To evaluate whether PocketGPT reflects this practice, we conducted fragment-conditioned generation across six pharmacologically prevalent motifs.

PocketGPT achieved consistently high fragment retention rates, ranging from 80.1% to 92.7% across fragments (Table 1), indicating reliable anchoring around specified chemical frameworks. Deviations from exact fragment identity were structured rather than random. Qualitative inspection revealed substitutions consistent with classical bioisosterism—for example, frequent replacement of tetrazole rings with 1,2,4-triazoles, preserving hydrogen-bonding capacity and electronic characteristics despite altered topology.

These results suggest that fragment conditioning constrains generation at the level of functional motifs, rather than enforcing rigid subgraph copying. Such behavior closely mirrors medicinal chemistry practice, where functional

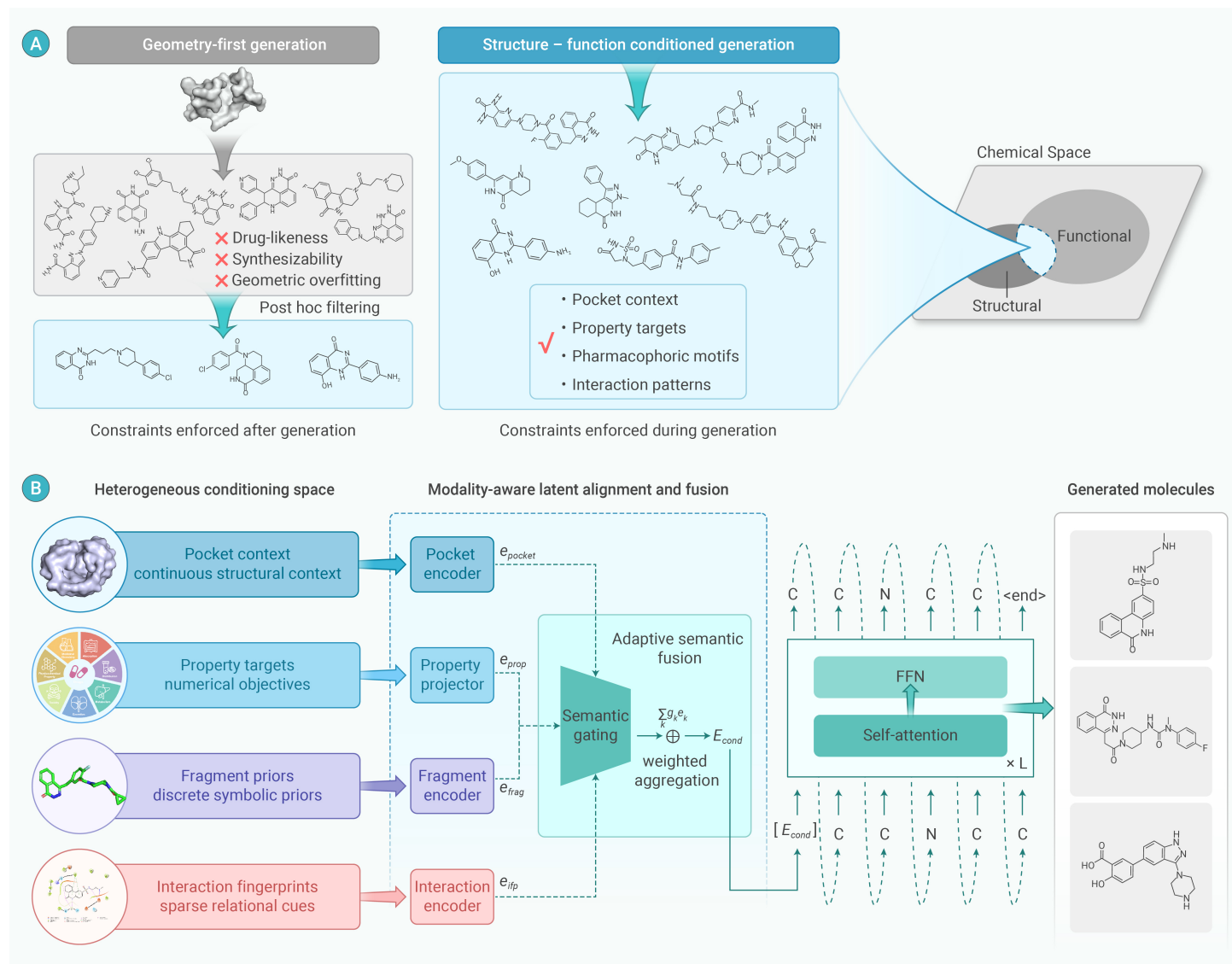


Figure 1. The PocketGPT framework for structure-aware molecular generation with composable semantic priors (A) Paradigm overview. PocketGPT departs from geometry-first generation followed by post hoc filtering, and instead embeds structural context and medicinal-chemistry design objectives directly into the generative process. Multiple sources of prior knowledge—including protein pocket context, physicochemical property targets, fragment priors, and protein–ligand interaction patterns—are integrated to guide generation toward chemically viable and pharmacologically relevant regions of chemical space. (B) Unified neural architecture. Heterogeneous conditioning signals—including 3D pocket embeddings, property targets, fragment priors, and interaction fingerprints—are first mapped into a shared latent space through modality-specific encoders, and then integrated by adaptive semantic fusion to yield a fused semantic prefix embedding E_{cond} . This conditional representation is prepended to the molecular token sequence and used to condition a Transformer decoder for autoregressive SMILES generation, enabling structure-aware molecular design without explicit three-dimensional ligand coordinate construction. Here, “function” refers to pharmacologically relevant design objectives, such as property profiles and binding-relevant interaction patterns, rather than explicit biological activity prediction.

equivalence and bioisosteric replacement are routinely exploited to preserve activity while optimizing downstream properties.

Interaction fingerprints guide functional interaction patterns. Binding specificity arises from structured constellations of non-covalent interactions, including hydrogen bonds, hydrophobic contacts, π – π stacking, salt bridges, π –cation interactions, and halogen bonds. To assess whether PocketGPT captures such functional semantics, we evaluated interaction-conditioned generation using predefined protein–ligand interaction fingerprints.

Analysis of the held-out test set revealed a long-tailed distribution of interaction types (Figure 3B), dominated by hydrogen bonds and hydrophobic contacts. Despite this intrinsic imbalance, PocketGPT preferentially recovered dominant interaction patterns associated with productive binding. For major interaction classes, median F1-scores exceeded 0.7 (Figure 3C), while increased variance was observed for rarer interaction types, consistent with their lower prevalence (Supplementary Information).

These findings indicate that interaction fingerprints guide generation toward functionally relevant interaction patterns, even in the absence of explicit three-dimensional ligand construction.

Robustness under realistic multi-objective design constraints

Rational drug design is inherently a multi-parameter optimization (MPO) problem⁴⁹, requiring the simultaneous satisfaction of structural compatibility, physicochemical viability, and functional requirements. While the previous sections established PocketGPT’s behavior under isolated conditioning signals, a more stringent and practically relevant evaluation concerns its behavior when multiple constraints are imposed concurrently, where trade-offs are unavoidable and provide a critical test of whether semantic guidance remains stable under competing objectives.

To this end, we evaluated PocketGPT under a comprehensive multi-constraint regime across 100 diverse protein targets, simultaneously enforcing pocket structural compatibility, ADMET-related property constraints, fragment retention, and interaction fingerprint alignment. As summarized in Table 2, the concurrent imposition of orthogonal semantic constraints leads to a measurable contraction of the accessible chemical space. Compared to single-objective conditioning, multi-constraint generation (PocketGPT-all) exhibits systematic reductions in validity, novelty, and uniqueness, reflecting the intrinsic cost of exerting explicit control along multiple, partially compet-

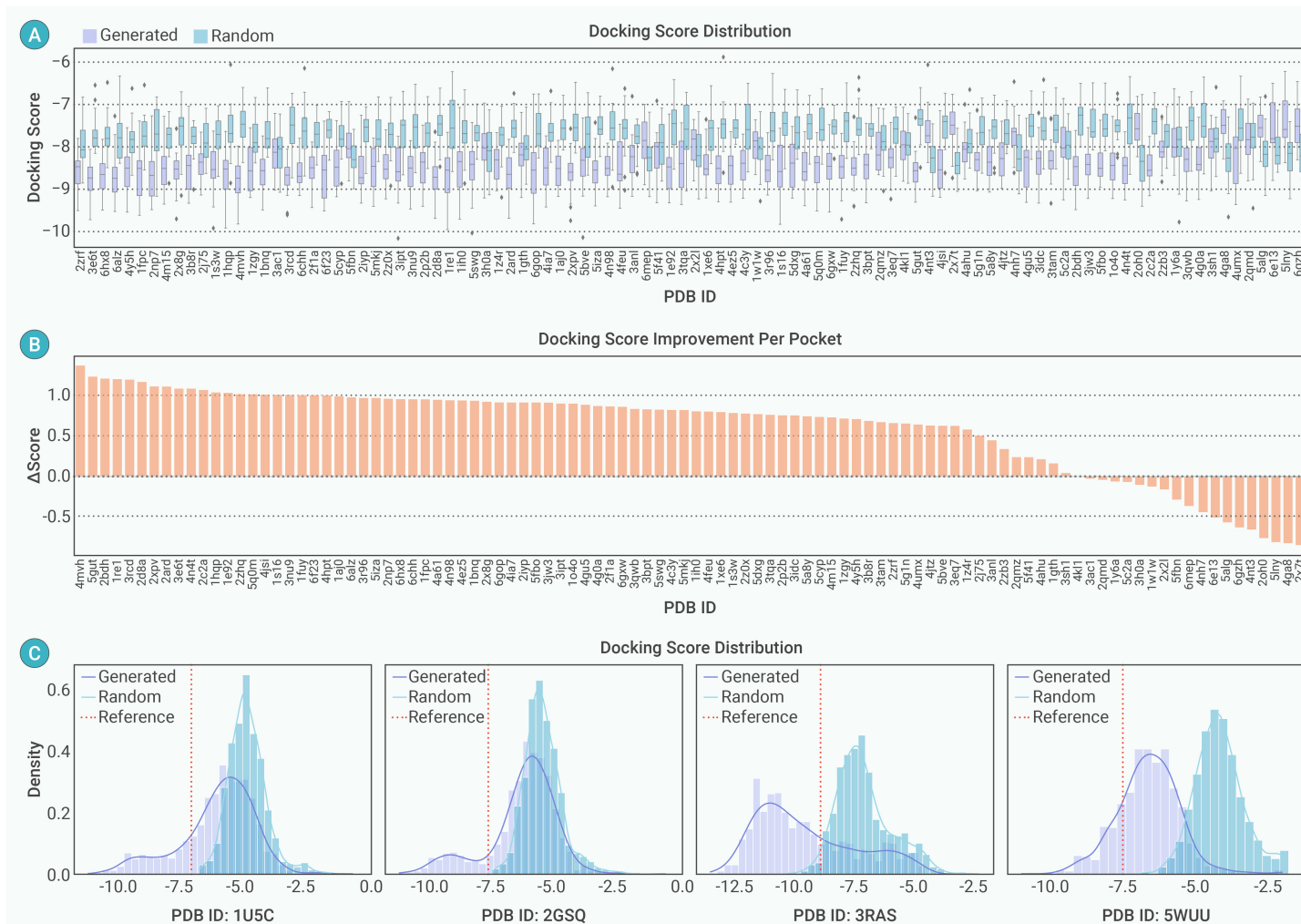


Figure 2. Pocket-derived semantic conditioning shifts generated molecules toward binding-compatible chemical space (A) Docking score distributions for PocketGPT-generated (red) and unconditioned random (blue) molecules across 100 test targets, showing a population-level shift toward lower-energy states under pocket conditioning. (B) Target-wise changes in mean docking score ($\Delta\text{Score} = \text{Baseline} - \text{PocketGPT}$), where positive values indicate more favorable predicted binding energetics induced by pocket-derived context. (C) Representative targets (PDB IDs indicated) illustrating consistent shifts of generated populations toward high-affinity regions relative to random baselines and reference ligands (dashed lines).

ing dimensions.

Consistent with this interpretation, Table 2 shows that while individual semantic constraints can be reliably satisfied in isolation, their concurrent enforcement reduces constraint adherence in a non-uniform manner. QED showed the largest decrease under full conditioning, dropping from 0.93 to 0.61, whereas LogP, TPSA, and SAS retained moderate adherence. This pattern suggests that composite property objectives are particularly sensitive to competition with fragment anchoring and interaction-pattern recovery, because QED is coupled to molecular size, ring systems, polarity, and heteroatom composition. Interaction-fingerprint alignment was also more difficult to preserve under concurrent constraints, consistent with the spatial specificity required for recovering non-covalent interaction patterns after SMILES generation and docking. In practical use, these results suggest that global property targets and interaction fingerprints may be treated as softer preferences during early-stage lead generation, whereas pocket compatibility and fragment-level anchoring may be prioritized more explicitly.

Importantly, this contraction in generative flexibility does not lead to a collapse of binding-relevant behavior. Despite the increased constraint burden, PocketGPT-all maintains favorable Top-10 docking scores (-8.3 on average), outperforming the PocketGPT training-stage baselines and remaining within the range of representative molecular generative models listed in Table 3.^{50,51} Together with the single-condition results, this comparison suggests that PocketGPT-all retains competitive docking-ranked performance while providing additional structure–function controllability through composable semantic conditioning.

This gradual degradation profile suggests a potential advantage of semantic conditioning under multi-objective constraints. Unlike coordinate-level generation, where mutually incompatible geometric and functional objectives may be difficult to satisfy simultaneously, PocketGPT operates in a semantic conditioning space and therefore showed smoother performance degradation in our evaluation, without an abrupt loss of generative viability (Table 3).

To further characterize behavior under concurrent constraints, we analyzed the top-ranked candidates generated under the comprehensive regime (Figure 4). Within the top 20 molecules per target (ranked by docking score), PocketGPT achieved fragment retention and ADMET compliance rates of 0.72 and 0.66, respectively, while interaction fingerprint alignment proved more challenging (0.52), consistent with the stringent spatial requirements associated with specific non-covalent interactions. The distribution of objective satisfaction across generated molecules (Figure 4B) shows that a substantial fraction of candidates simultaneously satisfy multiple constraints: 36.3% meet three objectives and 17.6% meet all four.

Global docking score distributions under the comprehensive constraint regime (Figure 4C) exhibit a consistent leftward shift relative to random generation across all targets, confirming that primary shape complementarity is largely preserved despite the imposition of auxiliary constraints. To further dissect the contribution of individual conditioning sources, we performed an equal-budget ablation analysis across the 100 held-out test targets. For each target and each generation setting, 100 molecules were generated and evaluated using the same docking and post-generation filtering workflow. Detailed docking and post-generation filtering criteria are

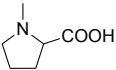
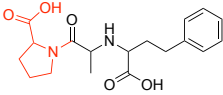
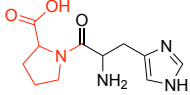
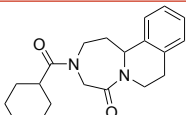
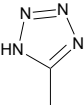
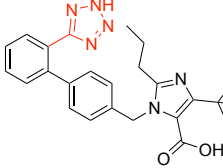
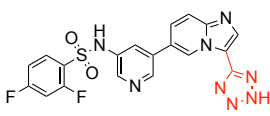
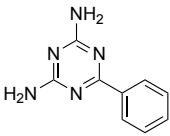
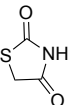
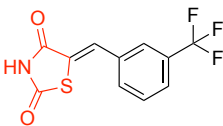
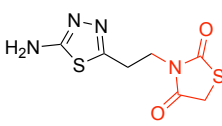
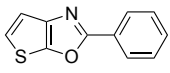
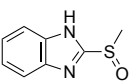
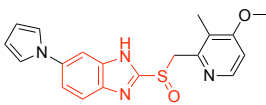
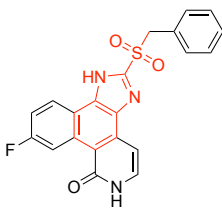
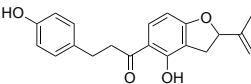
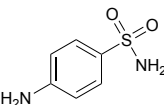
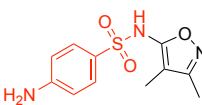
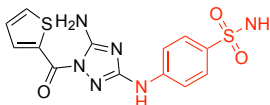
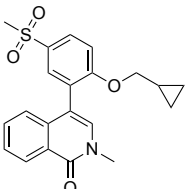
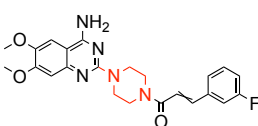
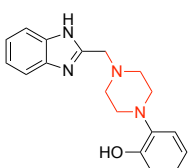
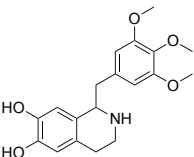


Figure 3. Directional responses to isolated functional and interaction-level conditioning (A) Ridgeline distributions of four physicochemical properties (LogP, QED, TPSA, and SAS) under single-property conditioning at different target values, showing well-separated and target-centered responses without post hoc filtering. (B) Dataset-level statistics of protein–ligand interaction fingerprints, including the frequency of major interaction types (left) and their coverage across test complexes (right), highlighting the long-tailed nature of interaction patterns. (C) Distribution of interaction-type-level recovery F1-scores across individual test complexes. Each point corresponds to a generated molecule, and dashed lines indicate per-complex mean values, illustrating consistent interaction-level recovery under interaction conditioning.

provided in Supplementary Methods S4. Generation-budget-controlled ablation settings and candidate-retention statistics are summarized in Supple-

mentary Note 6 and Table S1. Pocket-only generation produced the strongest docking-ranked enrichment, consistent with the direct role of pocket struc-

Table 1. Summary of benchmark fragment properties and retention performance in fragment-constrained generation

	Privileged Fragment	Hit rate	Hit 1	Hit 2	Miss
Proline		0.912			
Tetrazole		0.906			
Thiazolidinedione		0.894			
Benzimidazole-Sulfonyl		0.801			
Sulfonamide		0.826			
Piperazine		0.927			

tural conditioning in improving binding compatibility. Adding individual functional conditions produced distinct effects: property conditioning improved physicochemical-property compliance, fragment conditioning improved fragment retention, and interaction-fingerprint conditioning improved interaction-level recovery.

Importantly, PocketGPT-all did not necessarily achieve the best docking score as a single metric, reflecting the trade-off introduced by concurrent structural and functional constraints. However, under the same generation budget and identical filtering workflow, PocketGPT-all yielded the largest number of final qualified candidates satisfying docking-score, property, fragment-retention, and interaction-fingerprint criteria (Figure 4D-E). These results indicate that multi-constraint conditioning enriches high-quality candidates before virtual screening, rather than relying solely on post hoc filtering.

Prospective validation and discovery enabled by structure-function conditioned generation

Building on the directional semantic responses characterized under isolated conditioning (Section 2.2) and the robustness observed under coupled multi-objective constraints (Section 2.3), we next asked whether such semantic control is sufficient to support prospective molecular design and experimental validation in realistic discovery settings.

Cross-target generalization of structure-function semantic control. To evaluate whether joint structural and functional conditioning translates into coherent binding behavior across mechanistically distinct target classes, we applied PocketGPT to two therapeutically relevant proteins with distinct binding architectures: the phosphatases PTPN2/PTPN1⁵² (PDB: 7UAD) and the

kinase PI3Kα⁵³ (PDB: 8OW2). In these experiments, pocket geometry, ADMET-related physicochemical constraints, and pharmacophoric anchoring were enforced, while explicit interaction fingerprint conditioning was intentionally withheld. This design provides a stringent assessment of structure-function consistency under partial supervision.

Candidate libraries generated for both targets exhibited pronounced shifts toward favorable binding energetics compared to unconditioned baselines. As shown in Figure 5A, PocketGPT-generated molecules achieved mean docking scores of -8.91 for PTPN2/PTPN1 and -9.02 for PI3Kα, substantially outperforming random controls. Low-dimensional chemical space projections (UMAP) indicate that generated candidates cluster in proximity to reference ligands, suggesting preferential exploration of bioactive regions of chemical space. Analysis of physicochemical properties further confirms adherence to imposed design constraints, with generated molecules exhibiting favorable synthetic accessibility profiles relative to baseline samples (Figure 5B).

Post hoc interaction analysis revealed that, despite the absence of explicit interaction-level supervision during generation, PocketGPT-generated ligands consistently recovered interaction patterns characteristic of known binders (Figure 5D). For PTPN2/PTPN1, generated molecules frequently formed conserved hydrogen bond networks involving ARG222, GLN264, and ASP50, as well as π-π stacking interactions with PHE183. For PI3Kα, recurrent contacts with ASP603 and PHE1016 were observed, accompanied by additional polar interactions with residues such as GLU365 or TYR641. These observations indicate that joint structural and pharmacophoric conditioning can bias generation toward binding-compatible interaction patterns across mechanistically distinct targets.

Table 2. Constraint adherence under explicit semantic conditioning

Property	MolGPT	PocketGPT-single	PocketGPT-all
QED	0.91	0.93	0.61
SAS	0.87	0.89	0.65
LogP	0.93	0.95	0.70
TPSA	0.90	0.89	0.68

Table 3. Generative behavior of PocketGPT variants and representative molecular generation baselines

Method	Validity	Novelty	Unique	Top-10 Docking ↓
PocketGPT-scratch	0.760	0.680	0.713	-6.2
PocketGPT-pretrained	0.998	0.984	0.993	-7.5
MolGPT	0.975	0.783	0.982	-7.9
Pocket2Mol	0.989	0.865	0.992	-8.2
TargetDiff	0.974	0.846	0.975	-8.1
PocketFlow	0.978	0.880	0.985	-8.3
PocketGPT-single	0.987	0.860	0.987	-8.9
PocketGPT-all	0.840	0.771	0.798	-8.3

Note. PocketGPT-scratch denotes the PocketGPT architecture trained directly on the structure-conditioned fine-tuning dataset without large-scale molecular language pretraining. PocketGPT-pretrained denotes the chemically pretrained language model before structure-conditioned fine-tuning and is included as a chemical-prior baseline. PocketGPT-single denotes the pretrained PocketGPT model fine-tuned using pocket structural conditioning only. PocketGPT-all denotes the pretrained PocketGPT model fine-tuned with all conditioning variables, including pocket structure, physicochemical properties, fragment priors, and interaction fingerprints.

Prospective discovery and experimental validation of PARP1 inhibitors.

To evaluate whether the semantic control behaviors characterized above can translate into experimentally actionable molecular designs in a realistic discovery setting, we conducted a prospective design-make-test campaign targeting poly(ADP-ribose) polymerase 1 (PARP1), a clinically validated enzyme in DNA damage response pathways⁵⁴. PARP1 also represents a challenging testbed for molecular generation, as insufficient isoform selectivity has contributed to dose-limiting toxicities in several existing inhibitors. We therefore selected this target to assess the feasibility of molecules generated under structural and functional constraints when evaluated experimentally.

The generative campaign was anchored on the PARP1–Olaparib complex (PDB: 7KK4), which served exclusively as a geometric reference for pocket topology. Molecular generation was conditioned on pocket compatibility and constraints related to physicochemical and synthetic feasibility, as well as pharmacophoric anchoring to the phthalazinone core, a privileged scaffold shared by approved PARP inhibitors (Figure 6A–B).

Under these constraints, PocketGPT generated approximately two million candidate molecules. After deduplication against the training set, 1,944,524 candidate molecules were retained. After removing PAINS-positive⁵⁵ and bRo5 molecules, 920,795 candidates were retained. Docking-score-based⁵⁶ prioritization further reduced the pool to 46,034 molecules, followed by multi-constraint filtering based on predicted binding compatibility, physicochemical properties, scaffold diversity, and synthetic feasibility, yielding 7,873 candidates. From the top-50 docking-ranked molecules in this filtered pool, five representative compounds were selected for chemical synthesis and experimental evaluation. Detailed docking and post-generation filtering criteria are provided in Supplementary Methods S4.

Biochemical evaluation revealed that four of the five synthesized compounds inhibited PARP1 with sub-micromolar potency ($IC_{50} < 2 \mu M$; Figure 6D), indicating substantial empirical enrichment in this prospective campaign. Among these, Compound 2 exhibited the strongest inhibitory activity, exceeding that of the clinical reference Olaparib under matched assay conditions.

PARP2 selectivity was not included among the conditioning or prioritization criteria in the PARP1 campaign. Nevertheless, subsequent profiling against the closely related enzyme PARP2 revealed a pronounced preference

for PARP1. Specifically, Compound 2 showed a greater than 1000-fold reduction in inhibitory potency toward PARP2 relative to PARP1. This selectivity was therefore observed during post-generation experimental profiling, rather than being explicitly optimized during generation. Further studies incorporating counter-target information or dual-pocket selectivity objectives will be needed to systematically analyze and design PARP isoform selectivity.

To further characterize its functional relevance, Compound 2 was subjected to orthogonal biochemical and cellular assays. Isothermal titration calorimetry confirmed a high-affinity interaction with PARP1, yielding a dissociation constant (K_d) of 4.88 nM (Figure 7A–B). To further assess the dynamic stability of the predicted binding mode, we performed a 200 ns explicit-solvent molecular dynamics simulation of the Compound 2–PARP1 complex. The PARP1 backbone remained stable throughout the trajectory, with an average backbone RMSD of 0.154 ± 0.018 nm after protein alignment. The complex also maintained a persistent hydrogen-bonding network, with an average of 2.76 ± 0.54 hydrogen bonds. MM/GBSA analysis over 501 trajectory frames yielded a favorable binding free energy of -46.12 ± 5.54 kcal/mol, mainly driven by van der Waals and electrostatic contributions (Supplementary Note 7 and Figure S9). In BRCA1-deficient MDA-MB-436 breast cancer cells, Compound 2 suppressed cellular proliferation and colony formation more effectively than Olaparib at equivalent concentrations (Figure 7C–D). Consistent with on-target PARP1 inhibition, treatment with Compound 2 resulted in a dose-dependent reduction in poly(ADP-ribose) levels accompanied by accumulation of the DNA damage marker γ -H2AX (Figure 7E), indicating impaired DNA repair signaling.

To assess the preliminary *in vivo* safety profile of Compound 2, mice received daily oral administration of Compound 2 at doses of 100, 200, or 400 mg/kg for 15 consecutive days. Throughout the treatment period, no mortality, abnormal behavior, or overt signs of systemic toxicity were observed. Body weight increased steadily in all treatment groups and remained comparable to that of vehicle-treated controls, indicating that Compound 2 was well tolerated even at the highest tested dose (Figure 7F).

Histopathological examination of H&E-stained sections from major organs, including the heart, liver, spleen, lung, and kidney, revealed no apparent treatment-related pathological alterations (Figure 7G). The overall tissue architecture of all examined organs remained comparable to that of the control

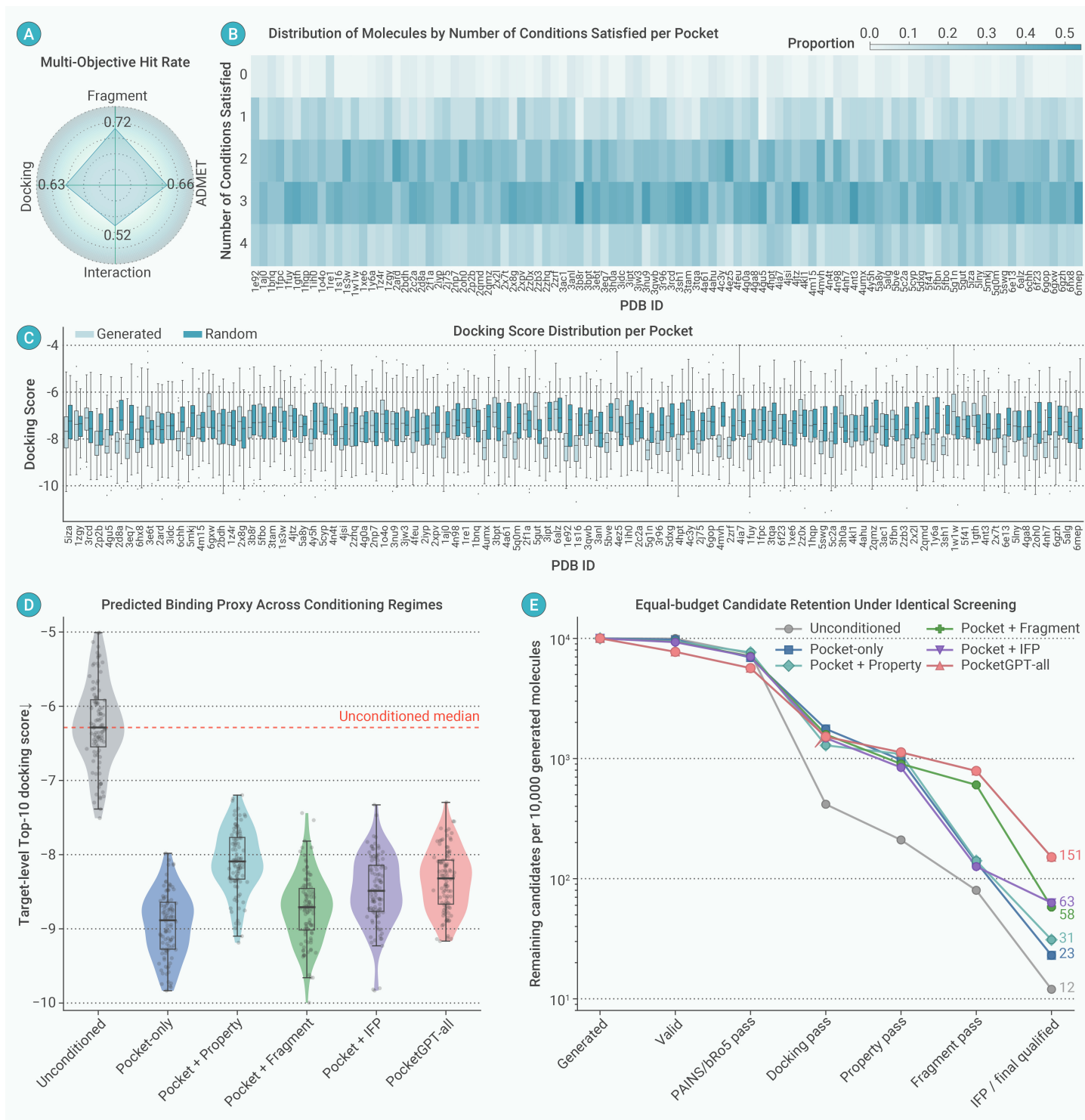


Figure 4. Compositional semantic guidance under concurrent constraints (A) Radar plot illustrating the simultaneous satisfaction of fragment anchoring, binding affinity, physicochemical properties, and interaction recovery for representative generated candidates. (B) Distribution of multi-constraint fulfilment across targets, showing that multiple semantic conditions can be jointly satisfied within a constrained generative library. (C) Docking score distributions under full conditioning, indicating that binding-relevant structural semantics are preserved despite concurrent functional and physicochemical guidance. (D) Equal-budget ablation of predicted binding performance across six conditioning settings, quantified by target-level Top-10 docking scores. Each point represents one held-out target, and lower values indicate more favorable predicted binding. (E) Cumulative retention of generated candidates under equal generation budgets and an identical post-generation filtering workflow. Values are normalized per 10,000 generated molecules. Final qualified candidates were defined using the docking and post-generation filtering criteria described in Supplementary Methods S4.

group. In particular, hepatic lobular organization was preserved without obvious structural disruption. Pulmonary tissues maintained normal alveolar morphology and showed no visible evidence of inflammatory lesions or alveolar interstitial thickening. Renal tissues exhibited intact glomerular and tubular structures. Likewise, no remarkable histopathological abnormalities were observed in the heart or spleen. Collectively, these findings suggest that Compound 2 possesses a favorable preliminary safety and tolerability profile

following repeated oral administration.

Collectively, this prospective design-make-test study demonstrates that structure-function semantic conditioning can give rise to chemically tractable molecules⁵⁷ whose experimentally observed binding potency, isoform preference, and biological effects are coherent across biochemical, cellular, and *in vivo* assays. These results were achieved without iterative structure-based optimization after generation, while the general benchmark

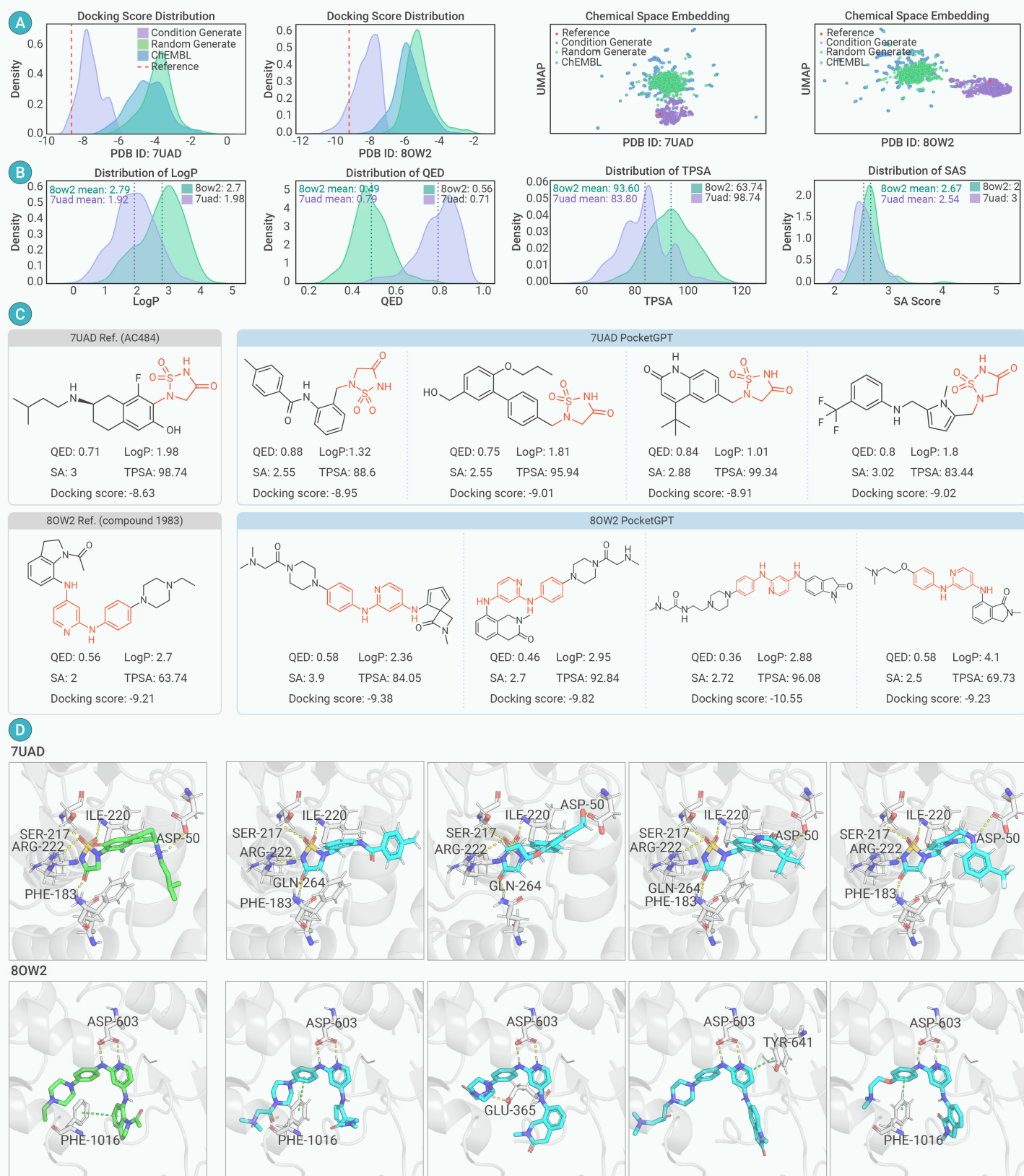


Figure 5. Emergent binding-relevant behavior from semantic conditioning (A) Pocket-conditioned generation induces a systematic shift toward favorable docking energetics and guides molecules toward chemical neighborhoods enriched in known active ligands. (B) Generated libraries satisfy imposed physicochemical constraints and exhibit favorable synthetic accessibility, supporting practical developability. (C) Representative structures preserve essential pharmacophoric scaffolds while adaptively optimizing substituents for pocket complementarity. (D) Without explicit interaction-level supervision, post hoc docking and interaction analysis showed recurrent recovery of reference-like interaction patterns, suggesting that pocket and pharmacophore conditioning can bias generation toward binding-compatible solutions.

and non-PARP cross-target experiments were performed using the base PocketGPT model.

CONCLUSIONS

We have presented PocketGPT, a structure-aware generative framework

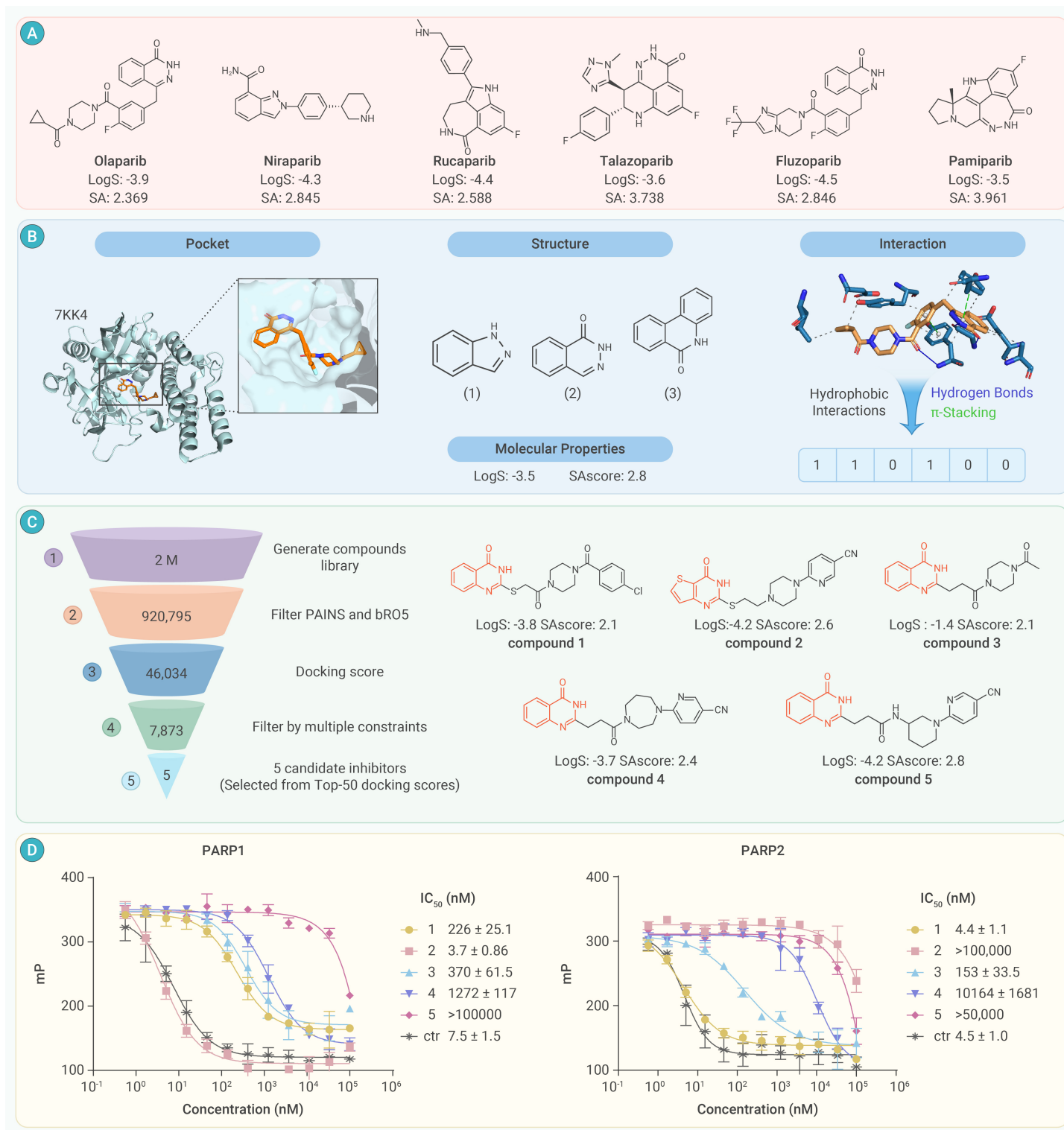


Figure 6. Semantic-prior-driven generative design and experimental validation of PARP1 inhibitors (A) Fragment priors extracted from clinically approved PARP1 inhibitors were used as semantic anchors for generation. (B) PocketGPT was conditioned on the PARP1–Olaparib complex (PDB: 7KK4) together with global physicochemical objectives (LogS and SAS). (C) Candidate prioritization from a large generated library enriched molecules consistent with the imposed semantic constraints. (D) Experimental evaluation of five synthesized candidates identified multiple potent inhibitors, including one exceeding the clinical reference Olaparib under matched assay conditions, supporting the practical utility of the generated candidates.

that reframes structure-based molecular design by decoupling binding-relevant structural and functional semantics from explicit three-dimensional ligand construction. By embedding multidimensional medicinal-chemistry constraints—including physicochemical properties, fragment priors, and protein–ligand interaction patterns—directly into the generative prior, PocketGPT establishes a structural–functional co-generative paradigm that contrasts with geometry-first approaches relying on post hoc filtering or

coordinate-level optimization.

The practical consequences of this paradigm are demonstrated through a prospective computation-to-experiment campaign targeting PARP1. Without explicitly optimizing affinity or isoform selectivity as independent objectives, PocketGPT generated candidates with favorable predicted binding compatibility and developability profiles, ultimately yielding a nanomolar-affinity ($K_d = 4.88$ nM) PARP1 inhibitor with favorable *in vivo* tolerability. Subsequent profil-

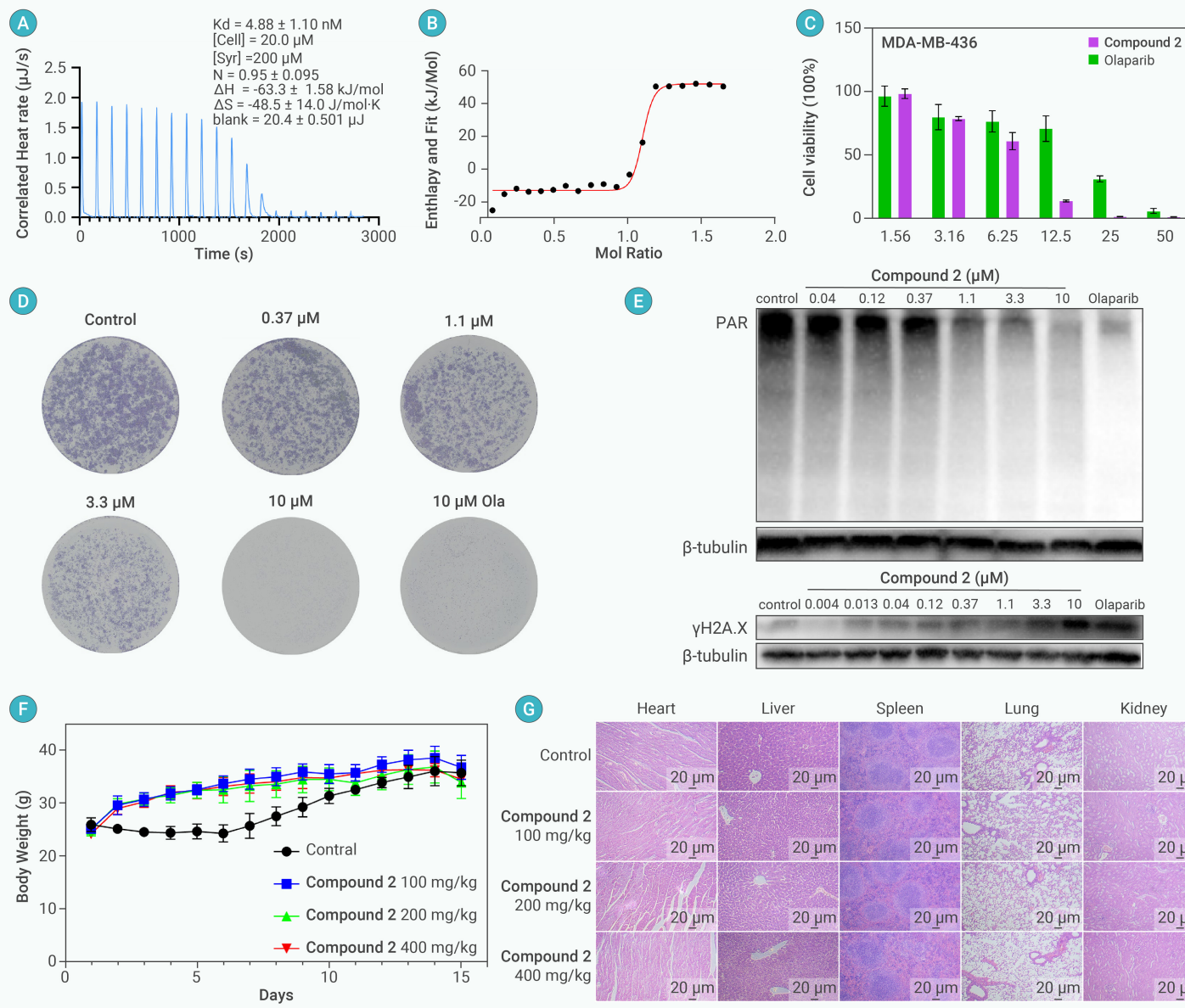


Figure 7. Experimental validation of target engagement and preliminary developability of a representative generated candidate (A-B) Biophysical characterization of Compound 2 by isothermal titration calorimetry confirms direct interaction with the PARP1 catalytic domain, providing evidence of specific target engagement. (C-D) Cellular assays demonstrate that Compound 2 elicits antiproliferative effects in PARP1-dependent cancer models, consistent with on-target activity. (E) Biochemical markers of PARP1 inhibition show dose-dependent suppression of poly(ADP-ribose) synthesis accompanied by induction of DNA damage signaling, supporting functional engagement of the intended pathway. (F) Changes in body weight during the 15-day treatment period. (G) Representative H&E-stained sections of major organs collected at the end of the 15-day treatment period. No obvious treatment-related histopathological abnormalities were observed in any examined tissues. Images were acquired at 200 $\times$ magnification. Scale bar = 20 μ m.

ing revealed strong PARP1-over-PARP2 selectivity for Compound 2, which we interpret as an experimental observation arising from the prospective campaign rather than as a pre-specified design objective. This progression from sequence-level generation to biochemical, cellular, and *in vivo* validation provides experimental support that structure–function conditioning can produce experimentally actionable candidates without explicit ligand geometry supervision.

At the same time, our results delineate clear boundaries of the proposed approach. The reduced novelty and uniqueness observed under fully conditioned generation indicate that strong multi-constraint control can restrict chemical-space exploration. This trade-off may be acceptable in target-focused lead optimization, where binding-relevant motifs, property profiles, and experimental tractability are often prioritized, but it may limit applications aimed at broad scaffold hopping or first-in-class discovery. PocketGPT should therefore be viewed as screening-aware generation rather than screening-free generation: it front-loads learnable and user-specifiable crite-

ria, including pocket compatibility, physicochemical property ranges, fragment priors, and interaction fingerprints, into the generative process, while retaining PAINS filtering and consensus docking as independent quality-control steps. This separation avoids direct optimization of discrete structural-alert rules or pose-dependent docking scores, which may otherwise introduce rule overfitting or scoring-function bias. Future work could mitigate these limitations through diversity-aware decoding, scaffold-level novelty regularization, softer fragment conditioning, constrained decoding, reinforcement learning, active learning, or Pareto-based multi-objective optimization.

More broadly, this work illustrates how rethinking the abstraction underlying structure-based molecular generation—shifting from geometry-centric formulations toward structure–function–aware generative priors—can preserve chemical realism while expanding the space of experimentally actionable designs. We anticipate that continued development along this direction will contribute to more interpretable, controllable, and experimentally grounded AI-assisted drug discovery pipelines.

MATERIALS AND METHODS

Dataset Construction and Preprocessing

Chemical prior pretraining. To establish a broad and chemically diverse prior, we constructed a large-scale ligand-only molecular corpus by integrating the MOSES⁵⁸ benchmark (ZINC Clean Leads subset) and the GuacaMol⁵⁹ benchmark (ChEMBL v24). Standardized SMILES preprocessing and deduplication were applied to ensure chemical validity and consistency, resulting in a final pretraining dataset of approximately 3.3 million unique drug-like molecules. This pretraining corpus contains no protein structures, pocket annotations, docking poses, or target labels, and therefore does not introduce pocket-level information from the downstream structure-conditioned benchmark.

Structure-conditioned fine-tuning. For pocket-aware generation, we curated a high-quality structural dataset from PDBbind^{60,61} v2020 and scPDB,⁶² restricted to co-crystallized protein–ligand complexes with X-ray resolution ≤ 2.5 Å. Complexes with incomplete binding sites or ambiguous ligand definitions were excluded. The final dataset comprised 25,914 protein–ligand pairs spanning diverse target families and binding geometries. To prevent direct target leakage across splits, data partitioning was performed at the protein-identity level⁶³, rather than at the individual-complex level. All complexes associated with the same protein identity were grouped before splitting and assigned to a single split. We verified that no protein identity present in the 100-target benchmark test set appeared in the training or validation set. The 100 independent test targets were reserved exclusively for benchmarking pocket-conditioned generation and were not used during model selection, hyperparameter tuning, or structure-conditioned fine-tuning.

Pocket Embedding

Pocket-level structural context was represented using DeeplyTough-derived geometric embeddings.⁶⁴ For each protein–ligand complex, the binding pocket was defined by protein heavy atoms within 8 Å of the co-crystallized ligand and voxelized into a multi-channel 3D grid V_i . The voxelized pocket was processed by a 3D convolutional encoder adopted from the DeeplyTough framework. Given the pocket grid V_i , the encoder f_θ produced a 128-dimensional pocket embedding:

$$z_i = f_\theta(V_i), z_i \in \mathbb{R}^{128} \quad (1)$$

The encoder was trained with a structural contrastive objective:

$$\mathcal{L}_{\text{pocket}} = y_{ij} d(z_i, z_j)^2 + (1 - y_{ij}) \max(0, m - d(z_i, z_j))^2, \quad (2)$$

where z_i and z_j denote the embeddings of two pockets, $d(\cdot, \cdot)$ denotes the distance between two pocket embeddings, $y_{ij} = 1$ indicates structurally related pockets and $y_{ij} = 0$ indicates unrelated pockets, and m is the contrastive margin. This objective encourages structurally similar binding pockets to occupy nearby regions in the embedding space while separating structurally dissimilar pockets.

The resulting pocket embedding was projected into the shared semantic conditioning space and used as the pocket-level prefix condition for the Transformer decoder. Thus, molecular generation is performed over SMILES tokens, whereas three-dimensional pocket information is introduced through the precomputed geometric embedding. Because the pocket encoder is trained on co-crystallized protein–ligand complexes, the embedding captures cavity shape and local pocket chemistry associated with compatible ligand chemotypes. Pocket–ligand shape complementarity is therefore encouraged implicitly through pocket-conditioned generation and was subsequently assessed by docking-based pose evaluation.

Molecular Property Profiling

Physicochemical and pharmacokinetic descriptors were computed using ADMETlab 2.0³². While a comprehensive ADMET suite was calculated for downstream benchmarking, four key descriptors were selected as conditioning constraints during training: LogP (partition coefficient), QED (drug-likeness), TPSA (topological polar surface area), and SAS (synthetic accessibility). These scalars served as control tokens to steer the generative process toward specific property manifolds.

Fragment Vocabulary Construction

To enable substructure-aware conditioning, molecules were decomposed using the BRICS algorithm⁶⁵. We implemented a hierarchical retention strategy, preserving both terminal leaf fragments and intermediate scaffolds from recursive splits to capture multi-scale structural logic. The final vocabulary comprised 18,988,771 fragments derived from over 56 million decomposition records, indexed in canonical SMILES format.

Constraint-adherence metrics

Fragment retention was defined as the fraction of evaluated molecules containing the queried fragment as an RDKit substructure match. Docking-score compliance was defined as the fraction of molecules with docking scores equal to or lower than a predefined target-specific threshold. Property compliance was calculated as the fraction of molecules satisfying all predefined physicochemical or developability criteria. Interaction-fingerprint alignment was quantified using the F1-score between six-dimensional PLIP-derived interaction-type vectors. Detailed formulas are provided in Supplementary Methods S2.

Benchmark metric calculation

Molecular-quality and docking-based metrics were calculated using a unified evaluation workflow across PocketGPT variants and comparison models whenever generated outputs could be regenerated or reprocessed. Validity, novelty, uniqueness, and Top-10 docking scores were computed after standardized RDKit processing and unified docking evaluation. Detailed benchmark metric calculations are provided in Supplementary Methods S3.

Model Architecture and Generation Strategy

PocketGPT employs a conditional Transformer decoder to autoregressively generate molecular SMILES. Pocket geometry, physicochemical-property objectives, fragment priors, and interaction fingerprints are encoded by modality-specific projectors and fused into a semantic prefix through learned gating. This prefix is prepended to the molecular token sequence and conditions next-token prediction during decoding. Different generation regimes were implemented by activating different subsets of conditioning variables while keeping the same model architecture and parameters. Detailed equations for condition encoding, semantic gating, prefix-conditioned decoding, and the training objective are provided in Supplementary Methods S1.

Chemical synthesis and characterization.

Compounds 1–5 were synthesized and characterized by NMR, HRMS, and HPLC. All compounds used in biological assays had $\geq 95\%$ purity. Detailed synthetic procedures and characterization data are provided in Supplementary Note 8.

Biochemical and cellular assays

PARP1 and PARP2 inhibition was evaluated using fluorescence polarization assays. Direct binding of Compound 2 to PARP1 was measured by ITC. Cellular activity was assessed in MDA-MB-436 cells using viability, clonogenic survival, and western blot assays. Detailed procedures are provided in Supplementary Methods S6.

In vivo tolerability assessment

Female ICR mice were treated with Compound 2 by daily oral gavage for 15 consecutive days. Body weight, clinical signs, and H&E-stained sections of major organs were evaluated. Detailed animal procedures are provided in Supplementary Methods S6.

Molecular dynamics and MM/GBSA analysis

A 200 ns explicit-solvent MD simulation was performed for the predicted Compound 2–PARP1 complex. Protein backbone RMSD, protein–ligand hydrogen bonds, MM/GBSA binding free energy, and per-residue energy decomposition were analyzed to assess the dynamic stability and energetic support of the predicted binding mode. Detailed simulation parameters are provided in Supplementary Methods S5.

REFERENCES

- Dickson, M., and Gagnon, J.P. (2004). Key factors in the rising cost of new drug discovery and development. *Nat Rev Drug Discov* **3**:417–429. DOI:10.1038/nrd1382
- Paul, S.M., Mytelka, D.S., Dunwiddie, C.T., et al. (2010). How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov* **9**:203–214. DOI:10.1038/nrd3078
- Sun, D., Gao, W., Hu, H., and Zhou, S. (2022). Why 90% of clinical drug development fails and how to improve it. *Acta Pharm Sin B* **12**(7):3049–3062. DOI:10.1016/j.apsb.2022.02.002
- Vamathevan, J., Clark, D., Czodrowski, P., et al. (2019). Applications of machine learning in drug discovery and development. *Nat Rev Drug Discov* **18**:463–477. DOI:10.1038/s41573-019-0024-5
- Catacutan, D.B., Alexander, J., Arnold, A., and Stokes, J.M. (2024). Machine learning in preclinical drug discovery. *Nat Chem Biol* **20**:960–973. DOI:10.1038/s41589-024-01679-1
- Du, Y., Jamasb, A.R., Guo, J., et al. (2024). Machine learning-aided generative molecular design. *Nat Mach Intell* **6**:589–604. DOI:10.1038/s42256-024-00843-5
- Vert, J.-P. (2023). How will generative AI disrupt data science in drug discovery. *Nat Biotechnol* **41**:750–751. DOI:10.1038/s41587-023-01789-6
- Khater, T., Alkhatib, S.A., AlShehhi, A., et al. (2025). Generative artificial intelligence based models optimization towards molecule design enhancement. *J Cheminformatics* **17**:116. DOI:10.1186/s13321-025-01059-4
- Tong, X., Liu, X., Tan, X., et al. (2021). Generative Models for De Novo Drug Design. *J Med Chem* **64**:14011–14027. DOI:10.1021/acs.jmedchem.1c00927
- Wang, M., Wang, Z., Sun, H., et al. (2022). Deep learning approaches for *de novo* drug design: An overview. *Curr Opin Struc Biol* **72**:135–144. DOI:10.1016/j.sbi.2021.10.001
- Ivanenko, Y., Zagrebnyy, B., Malyshev, A., et al. (2023). The Hitchhiker's Guide to Deep Learning Driven Generative Chemistry. *ACS Med Chem Lett* **14**:901–915. DOI:10.1021/acsmchemlett.3c00041
- Deng, J., Yang, Z., Ojima, I., et al. (2022). Artificial intelligence in drug discovery: applications and techniques. *Brief Bioinform* **23**:bbab430. DOI:10.1093/bib/bbab430
- Tropsha, A., Isayev, O., Varnek, A., et al. (2024). Integrating QSAR modelling and deep learning in drug discovery: the emergence of deep QSAR. *Nat Rev Drug Discov* **23**:141–155. DOI:10.1038/s41573-023-00832-0
- Grisoni, F., Moret, M., Lingwood, R., and Schneider, G. (2020). Bidirectional Molecule Generation with Recurrent Neural Networks. *J Chem Inf Model* **60**:1175–1183. DOI:10.1021/acs.jcim.9b00943
- Tang, Y., Moretti, R., and Meiler, J. (2024). Recent Advances in Automated Structure-Based De Novo Drug Design. *J Chem Inf Model* **64**:1794–1805. DOI:10.1021/acs.jcim.4c00247
- Jiang, Y., Zhang, G., You, J., et al. (2024). PocketFlow is a data-and-knowledge-driven structure-based molecular generative model. *Nat Mach Intell* **6**:326–337. DOI:10.1038/s42256-024-00808-8
- Schneuing, A., Harris, C., Du, Y., et al. (2024). Structure-based drug design with equivariant diffusion models. *Nat Comput Sci* **4**:899–909. DOI:10.1038/s43588-024-00737-x
- Xu, M., Ran, T., and Chen, H. (2021). De Novo Molecule Design Through the Molecular Generative Model Conditioned by 3D Information of Protein Binding Sites. *J Chem Inf Model* **61**:3240–3254. DOI:10.1021/acs.jcim.0c01494
- Wu, K., Xia, Y., Deng, P., et al. (2024). TamGen: drug design with target-aware molecule generation through a chemical language model. *Nat Commun* **15**:9360. DOI:10.1038/s41467-024-53632-4
- Zhang, W., Zhang, K., and Huang, J. (2023). A Simple Way to Incorporate Target Structural Information in Molecular Generative Models. *J Chem Inf Model* **63**:3719–3730. DOI:10.1021/acs.jcim.3c00293
- Ünlü, A., Çevrim, E., Yiğit, M.G., et al. (2025). Target-specific *de novo* design of drug candidate molecules with graph-transformer-based generative adversarial networks. *Nat Mach Intell* **7**:1524–1540. DOI:10.1038/s42256-025-01082-y
- Peng, X., Luo, S., Guan, J., et al. (2025). Pocket2Mol: Efficient Molecular Sampling Based on 3D Protein Pockets. DOI: 10.48550/arXiv.2205.07249
- Zhang, O., Zhang, J., Jin, J., et al. (2023). ResGen is a pocket-aware 3D molecular generation model based on parallel multiscale modelling. *Nat Mach Intell* **5**:1020–1030. DOI:10.1038/s42256-023-00712-7
- Hu, Q., Sun, C., He, H., et al. (2025). Target-aware 3D molecular generation based on guided equivariant diffusion. *Nat Commun* **16**:7928. DOI:10.1038/s41467-025-63245-0
- Huang, L., Xu, T., Yu, Y., et al. (2024). A dual diffusion model enables 3D molecule generation and lead optimization based on target pockets. *Nat Commun* **15**:2657. DOI:10.1038/s41467-024-46569-1
- Zhung, W., Kim, H., and Kim, W.Y. (2024). 3D molecular generative framework for interaction-guided drug design. *Nat Commun* **15**:2688. DOI:10.1038/s41467-024-47011-2
- Yu, J.-L., Zhou, C., Ning, X.-L., et al. (2025). Knowledge-guided diffusion model for 3D ligand-pharmacophore mapping. *Nat Commun* **16**:2269. DOI:10.1038/s41467-025-57485-3
- Ghorbani, M., Gendele, L., Beroza, P., and Keiser, M.J. (2023). Autoregressive fragment-based diffusion for pocket-aware ligand design. DOI: 10.48550/arXiv.2401.05370
- Peng, J., Yu, J.-L., Yang, Z.-B., et al. (2025). Pharmacophore-oriented 3D molecular generation toward efficient feature-customized drug discovery. *Nat Comput Sci* **5**:898–914. DOI: 10.1038/s43588-025-00850-5
- Yang, Y., Hsieh, C.-Y., Kang, Y., et al. (2023). Deep Generation Model Guided by the Docking Score for Active Molecular Design. *J Chem Inf Model* **63**:2983–2991. DOI:10.1021/acs.jcim.3c00572
- Dong, T., You, L., and Chen, C.Y.-C. (2025). Multi-objective drug design with a scaffold-aware variational autoencoder. *Chem Sci* **16**:13352–13367. DOI:10.1039/D4SC08736D
- Xiong, G., Wu, Z., Yi, J., et al. (2021). ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. *Nucleic Acids Res* **49**:W5–W14. DOI: 10.1093/nar/gkab255
- Kotsias, P.-C., Arús-Pous, J., Chen, H., et al. (2020). Direct steering of *de novo* molecular generation with descriptor conditional recurrent neural networks. *Nat Mach Intell* **2**:254–265. DOI:10.1038/s42256-020-0174-5
- Degen, J., Wegscheid-Gerlach, C., Zaliani, A., and Rarey, M. (2008). On the Art of Compiling and Using 'Drug-Like' Chemical Fragment Spaces. *ChemMedChem* **3**:1503–1507. DOI:10.1002/cmdc.200800178
- Pang, C., Qiao, J., Zeng, X., et al. (2024). Deep Generative Models in De Novo Drug Molecule Generation. *J Chem Inf Model* **64**:2174–2194. DOI:10.1021/acs.jcim.3c01496
- Schake, P., Bolz, S.N., Linnemann, K., and Schroeder, M. (2025). PLIP 2025: introducing protein–protein interactions to the protein–ligand interaction profiler. *Nucleic Acids Res* **53**:W463–W465. DOI:10.1093/nar/gkaf361
- Liu, Z., Su, M., Han, L., et al. (2017). Forging the Basis for Developing Protein–Ligand Interaction Scoring Functions. *Accounts Chem Res* **50**:302–309. DOI:10.1021/acs.accounts.6b00491
- Moret, M., Pachon Angona, I., Cotos, L., et al. (2023). Leveraging molecular structure and bioactivity with chemical language models for *de novo* drug design. *Nat Commun* **14**:114. DOI:10.1038/s41467-022-35692-6
- Bagal, V., Aggarwal, R., Vinod, P.K., and Priyakumar, U.D. (2022). MolGPT: Molecular Generation Using a Transformer-Decoder Model. *J Chem Inf Model* **62**:2064–2076. DOI:10.1021/acs.jcim.1c00600
- Grechishnikova, D. (2021). Transformer neural network for protein-specific *de novo* drug generation as a machine translation problem. *Sci Rep* **11**:321. DOI:10.1038/s41598-020-79682-4
- Chen, Y., Wang, Z., Wang, L., et al. (2023). Deep generative model for drug design from protein target sequence. *J Cheminformatics* **15**:38. DOI:10.1186/s13321-023-00702-2
- Born, J., and Manica, M. (2023). Regression Transformer enables concurrent sequence regression and generation for molecular language modelling. *Nat Mach Intell* **5**:432–444. DOI:10.1038/s42256-023-00639-z
- Zhang, H., Song, H., Li, S., et al. (2023). A Survey of Controllable Text Generation Using Transformer-based Pre-trained Language Models. *ACM Comput Surv* **56**:64:61–64:37. DOI: 10.1145/3617680.
- Atz, K., Cotos, L., Isert, C., et al. (2024). Prospective *de novo* drug design with deep interactome learning. *Nat Commun* **15**:3408. DOI:10.1038/s41467-024-47613-w
- Li, Y., Zhang, L., Wang, Y., et al. (2022). Generative deep learning enables the discovery of a potent and selective RIPK1 inhibitor. *Nat Commun* **13**:6891. DOI:10.1038/s41467-022-34692-w
- Vaswani, A., Shazeer, N., Parmar, N., et al. (2023). Attention Is All You Need. DOI: 10.48550/arXiv.1706.03762.
- Ross, J., Belgodere, B., Chenthamarakshan, V., et al. (2022). Large-scale chemical language representations capture molecular structure and properties. *Nat Mach Intell* **4**:1256–1264. DOI:10.1038/s42256-022-00580-7
- Su, M., Yang, Q., Du, Y., et al. (2019). Comparative Assessment of Scoring Functions: The CASF-2016 Update. *J Chem Inf Model* **59**:895–913. DOI:10.1021/acs.jcim.8b00545
- Yang, Y., Chen, G., Li, J., et al. (2024). Enabling target-aware molecule generation to follow multi objectives with Pareto MCTS. *Commun Biol* **7**:1074. DOI:10.1038/s42003-024-06746-w
- Liu, H., Qin, Y., Niu, Z., et al. (2024). How Good are Current Pocket-Based 3D Generative Models?: The Benchmark Set and Evaluation of Protein Pocket-Based 3D Molecular Generative Models. *J Chem Inf Model* **64**:9260–9275. DOI: 10.1021/acs.jcim.4c01598.
- Jocys, Z., Grundy, J., and Farrahi, K. (2024). DrugPose: benchmarking 3D generative methods for early stage drug discovery. *Digit Discov* **3**:1308–1318. DOI: https://doi.org/10.1039/d4dd00076e.
- Baumgartner, C.K., Ebrahimi-Nik, H., Iracheta-Vellve, A., et al. (2023). The PTPN2/PTPN1 inhibitor ABBV-CLS-484 unleashes potent anti-tumour immunity. *Nature* **622**:850–862. DOI:10.1038/s41586-023-06575-7
- Gong, G.Q., Bilanges, B., Allsop, B., et al. (2023). A small-molecule PI3Kα activator for cardioprotection and neuroregeneration. *Nature* **618**:159–168. DOI:10.1038/s41586-023-05972-2
- Wang, Y., Zhang, J., Wu, X., et al. (2025). The Potential of PARP Inhibitors as Antitumor Drugs and the Perspective of Molecular Design. *J Med Chem* **68**:18–48. DOI:10.1021/acs.jmedchem.4c02642
- Baell, J.B., and Holloway, G.A. (2010). 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. DOI:10.1021/jm901137j
56. Passaro, S., Corso, G., Wohlwend, J., et al. (2025). Boltz-2: Towards Accurate and Efficient Binding Affinity Prediction. DOI: 10.1101/2025.06.14.659707.
 57. Yang, Z., Sun, W., Huang, Q., et al. (2026). Deep learning-assisted discovery of a potent and cell-active inhibitor of RNA N6-methyladenosine recognition protein YTHDC2. *Nat Commun* **17**:46. DOI:10.1038/s41467-025-65542-0
 58. Polykovskiy, D., Zhebrak, A., Sanchez-Lengeling, B., et al. (2020). Frontiers | Molecular Sets (MOSES): A Benchmarking Platform for Molecular Generation Models. *Front Pharmacol* **11**:565644. DOI: 10.3389/fphar.2020.565644.
 59. Brown, N., Fiscato, M., Segler, M.H.S., and Vaucher, A.C. (2019). GuacaMol: Benchmarking Models for de Novo Molecular Design. *J Chem Inf Model* **59**:1096–1108. DOI:10.1021/acs.jcim.8b00839
 60. Wang, R., Fang, X., Lu, Y., et al. (2005). The PDBbind Database: Methodologies and Updates. *J Med Chem* **48**:4111–4119. DOI:10.1021/jm048957q
 61. Liu, Z., Li, Y., Han, L., et al. (2014). PDB-wide collection of binding data: current status of the PDBbind database. *Bioinformatics* **31**:405–412. DOI:10.1093/bioinformatics/btu626
 62. Desaphy, J., Bret, G., Rognan, D., and Kellenberger, E. (2015). sc-PDB: a 3D-database of ligandable binding sites—10 years on. *Nucleic Acids Res* **43**:D399–D404. DOI:10.1093/nar/gku928
 63. Graber, D., Stockinger, P., Meyer, F., et al. (2025). Resolving data bias improves generalization in binding affinity prediction. *Nat Mach Intell* **7**:1713–1725. DOI:10.1038/s42256-025-01124-5
 64. Simonovsky, M., and Meyers, J. (2020). DeeplyTough: Learning Structural Comparison of Protein Binding Sites. *J Chem Inf Model* **60**:2356–2366. DOI:10.1021/acs.jcim.9b00554
 65. Liu, T., Naderi, M., Alvin, C., et al. (2017). Break Down in Order To Build Up: Decomposing Small Molecules for Fragment-Based Drug Design with eMolFrag. *J Chem Inf Model* **57**:627–631. DOI:10.1021/acs.jcim.6b00596

FUNDING AND ACKNOWLEDGMENTS

This work was financially supported by the National Natural Science Foundation of China (82304316, 82221002, 22173118, 22307112), National Science and Technology Major Project of the Ministry of Science and Technology of China [2023ZD0507104]. We

thank the Center for Computational Biology and Bioinformatics, Furong Laboratory, the Bioinformatics Center of Xiangya Hospital, Central South University, Haikun Xu, and the High-Performance Computing Center of Central South University for technical and computational support. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

D.C. designed the study and supervised the overall research, provided funding support, and revised the manuscript. R.Q. conceived the study, designed the overall methodology, performed the main computational analyses, and drafted the manuscript. Y.D. designed the strategy for the experimental validation of PARP1. X.L. performed the PARP1 wet-lab experiments and, together with J.D., wrote the corresponding experimental section. W.W. assisted in structuring the manuscript and contributed through regular discussions. T.Z. contributed through discussions and guidance on molecular docking. D.J., J.L., and A.L. reviewed the manuscript and provided suggestions for revision. All authors contributed to the manuscript and approved the final version.

DECLARATION OF INTERESTS

Dongsheng Cao and Aiping Lyu are Editorial Board members of *The Innovation Drug Discovery* and were blinded from reviewing or making final decisions on the manuscript. Peer review was handled independently of these members and their research groups. The other authors declare no conflicts of interest.

DATA AND CODE AVAILABILITY

The data and source code of this study are freely available on GitHub (<https://github.com/qianrong0709/PocketGPT>) to facilitate replication of the results.

SUPPLEMENTAL INFORMATION

It can be found online at <https://doi.org/10.59717/j.xinn-drugdisc.2026.100028>.