Generative Design of Natural Product-Like Molecules: Constraint-Guided Optimization for Bioactivity, Safety, Synthetic Accessibility, and Developability

Sara Ben Youssef[1]*, Amal Trabelsi[1], Karim Boudiaf[2]

[1]Department of AI for Antidiabetic Cactus and Olive Leaf Extracts, Faculty of Pharmacy, University of Tunis, Tunis, Tunisia

[2]Department of Virtual Screening for PPAR Gamma Agonists, Faculty of Pharmacy, University of Sfax, Sfax, Tunisia.

ABSTRACT

Generative artificial intelligence provides a computational basis for exploring molecular structures beyond conventionally enumerated chemical libraries, but unconstrained generation can produce structures that are chemically unrealistic, biologically unsupported, unsafe, difficult to synthesize, or unsuitable for further development. This article proposes a conceptual generative AI architecture for designing natural product-like molecules under explicit chemical, biological, safety, synthesis, and developability constraints. The approach distinguishes natural product-like molecules from isolated natural products and treats natural product-likeness as a multidimensional design characteristic rather than evidence of origin, bioactivity, or therapeutic value. The architecture integrates provenance-aware data curation, stereochemistry-sensitive molecular representations, scaffold- and analogue-oriented generation, chemical-validity controls, natural product-likeness and diversity constraints, target-context and bioactivity prediction, selectivity considerations, ADMET and toxicity filters, synthetic-accessibility assessment, retrosynthetic plausibility, reaction-feasibility review, uncertainty estimation, and expert curation. Multi-objective scoring and Pareto-style prioritization are proposed to expose trade-offs rather than conceal them within a single composite measure. Developability filters address physicochemical properties, solubility, permeability, stability, metabolic liabilities, and formulation-relevant concerns while preserving explicit decision boundaries. The principal contribution is a structured design–filter–review–validate architecture that positions generated structures as testable molecular hypotheses. Synthesis, analytical characterization, biological evaluation, experimental ADMET assessment, toxicity testing, and subsequent development studies remain necessary before stronger pharmacological, safety, or translational claims can be made.

Key Words: *Generative AI, Natural product-like molecules, Molecular design, Constraint-guided optimization, Bioactivity prediction, Synthetic accessibility*

eIJPPR 2026; 16(2):19-35

HOW TO CITE THIS ARTICLE: Youssef SB, Trabelsi A, Boudiaf K. Generative Design of Natural Product-Like Molecules: Constraint-Guided Optimization for Bioactivity, Safety, Synthetic Accessibility, and Developability. Int J Pharm Phytopharmacol Res. 2026;16(2):19-35. https://doi.org/10.51847/wwbWWkGcFo

INTRODUCTION

Generative artificial intelligence has expanded computational molecular design from the ranking of predefined structures toward inverse design, in which candidate structures are proposed in relation to desired properties. This shift can support the exploration of chemical spaces that are too large or structurally diverse for exhaustive enumeration. Nevertheless, inverse molecular design depends on the quality of its molecular representations, property predictors, optimization objectives, and validation procedures; generated structures are therefore outputs of a computational hypothesis-formulation process rather than automatically useful

Corresponding author: Sara Ben Youssef
Address: Department of AI for Antidiabetic Cactus and Olive Leaf Extracts, Faculty of Pharmacy, University of Tunis, Tunis, Tunisia.
E-mail: ✉ sara.benyoussef@fpharm.u-tunis.tn
Received: 01 December 2025; **Revised:** 14 March 2026; **Accepted:** 15 March 2026

compounds [1]. For natural product-like molecular design, this distinction is especially important because structural complexity, stereochemistry, scaffold novelty, and biological relevance can create apparent chemical attractiveness without establishing synthetic feasibility, pharmacological activity, or safety.

The behaviour of a molecular generator is strongly influenced by how chemical structures are encoded and by whether generation occurs through molecular strings, graphs, fragments, scaffolds, reactions, or continuous latent representations. These choices affect chemical validity, reconstruction fidelity, diversity, novelty, controllability, and the ability to optimize molecular properties [2]. A natural product-like design system must therefore do more than reproduce frequently occurring fragments from a training corpus. It must preserve chemically interpretable connectivity, represent stereochemical information, avoid uncontrolled duplication of training examples, and distinguish meaningful scaffold variation from superficial peripheral modification.

Deep generative molecular design also faces recurrent problems involving invalid structures, unstable optimization, narrow sampling, biased predictive objectives, and inadequate evaluation. Models can exploit weaknesses in property predictors, generate structures outside the predictors' applicability domains, or obtain favourable computational scores through chemically undesirable features [3]. These risks become more consequential when optimization combines natural product-likeness, novelty, target-oriented bioactivity, ADMET predictions, toxicity alerts, synthesis-related criteria, and developability objectives. An architecture intended for this setting must consequently convert these limitations into explicit constraints, quality gates, uncertainty flags, and decision boundaries.

This article proposes a conceptual generative AI architecture for constraint-guided design of natural product-like molecules. The architecture coordinates molecular and scaffold generation with chemical-validity controls, natural product-likeness criteria, multi-objective optimization, bioactivity and target-context prediction, safety filtering, synthetic-accessibility assessment, developability review, uncertainty estimation, expert curation, and experimental-validation planning. The approach is modular because atom-level, string-based, graph-based, fragment-oriented, scaffold-based, and reaction-aware design strategies address different aspects of molecular construction and practical feasibility [4]. It is not presented as a trained or benchmarked model; rather, it defines an evidence-grounded structure for generating and prioritizing molecular hypotheses while preventing computational scores from being misinterpreted as experimental or translational proof.

Natural product-like molecular design

Natural product-like molecular design refers here to the computational creation or modification of structures that reproduce selected characteristics associated with natural-product chemical space without claiming that the resulting molecules have been isolated from nature. A generated natural product-like molecule is therefore distinct from an isolated natural product, whose natural occurrence has been experimentally established; a natural product analogue, which is derived through modification of a known natural product; and a generic drug-like molecule, which may satisfy conventional physicochemical criteria without preserving natural-product-related scaffold, stereochemical, or architectural features. Prospective work on natural-product-inspired receptor modulators demonstrates that domain-focused generative design can guide candidate ideation, but synthesis and biological testing remain necessary before computationally proposed structures can be interpreted as active compounds [5].

Natural product-likeness should be treated as a multidimensional relationship to a reference chemical space rather than as a binary label. Relevant dimensions may include scaffold architecture, ring-system composition, stereochemical density, sp^3 character, heteroatom patterns, molecular shape, functional-group distribution, physicochemical properties, and proximity to known natural-product or natural-product-related neighbourhoods. Mapping natural-product chemical space illustrates that chemically diverse structures can occupy different regions while retaining varying degrees of natural-product-related character [6]. Consequently, no single similarity score should be permitted to define natural product-likeness, and resemblance to known natural products must not be interpreted as evidence of biosynthetic origin, target engagement, safety, or therapeutic value.

Large virtual libraries generated from natural-product-focused molecular languages can broaden the range of structures available for computational exploration. Such libraries may contain valid and novel structures whose calculated distributions resemble selected characteristics of natural-product chemical space, but their scale does not establish that individual members are synthesizable, stable, biologically active, or safe [7]. Within the proposed architecture, generated natural product-like structures are therefore assigned provenance records, nearest-neighbour information, scaffold relationships, representation-validity checks, and uncertainty flags. They remain virtual hypotheses until selected candidates undergo synthesis

planning, production, analytical characterization, and experimental testing.

Conditional generation provides a means of preserving selected natural-product-inspired motifs or scaffolds while exploring structural modifications around them. Scaffold conditioning can support analogue generation, motif conditioning can preserve chemically meaningful substructures, and property conditioning can direct exploration toward project-specific regions of chemical space. A conditional transformer developed for natural-product-inspired generation illustrates how scaffold, motif, and property information can be incorporated into molecular generation [8]. However, the proposed architecture separates structural conditioning from claims of chemical or pharmacological quality: retaining an inspired scaffold does not guarantee plausible stereochemistry, useful novelty, target relevance, synthetic feasibility, or developability. Natural product-like design must consequently combine domain resemblance with chemical realism, portfolio diversity, biological context, synthesis awareness, and explicit validation requirements.

Constraint-guided generative optimization

Constraint-guided generative optimization combines molecular generation with rules, predictors, structural conditions, and decision gates that restrict or rank candidate structures. Continuous molecular representations can support traversal of a learned latent space and optimization toward calculated properties, but decoded molecules may be invalid or fall outside the reliable domain of the associated predictors [9]. Reinforcement learning offers another mechanism by which a molecular generator can be biased toward desired objectives through a reward function [10]. In the proposed architecture, neither latent optimization nor reward maximization is treated as sufficient alone. Every generated candidate must pass independent checks for representation validity, chemical plausibility, stereochemical consistency, natural product-likeness, novelty, scaffold diversity, and applicability of the models used to score it.

Generator–predictor coupling can support target-oriented ideation by using predicted biological properties to guide candidate generation, but the generator may exploit predictor artefacts or overfit to narrow regions of chemical space. Deep reinforcement-learning systems have demonstrated computational optimization toward predicted target-related properties, yet such predictions do not establish pharmacological activity [11]. Multiparameter optimization further allows several molecular objectives to be considered together, although the resulting priorities depend on the mathematical form, weighting, and calibration of the desirability functions [12]. The proposed workflow therefore retains the raw value, uncertainty estimate, applicability-domain status, and rationale for every component objective rather than reporting only an aggregate score.

Multi-objective optimization is necessary because natural product-like molecular design involves competing requirements. Increasing structural novelty may reduce the amount of supporting bioactivity or synthesis precedent, whereas strict similarity constraints may reproduce known scaffolds without meaningful exploration. Improvements in predicted target activity may be accompanied by unfavourable solubility, metabolic, selectivity, or synthesis-related characteristics. Distribution-learning and goal-directed benchmarks assess different generative capabilities and should not be treated as interchangeable indicators of model quality [13]. Where objectives conflict, Pareto-style prioritization can identify non-dominated candidates that represent alternative trade-offs. This approach should be accompanied by sensitivity analysis, uncertainty-aware ranking, and expert review because an unstable Pareto frontier constructed from unreliable predictors offers little practical guidance.

Computational evaluation should distinguish technical generation quality from chemical, biological, and translational relevance. Validity, uniqueness, novelty, and distributional similarity are important for identifying malformed structures, duplication, mode collapse, and divergence from a defined reference space, but benchmark success does not establish bioactivity, safety, synthesizability, or developability [14]. In the proposed architecture, hard constraints exclude technically invalid or explicitly prohibited structures, soft constraints guide search toward desirable regions, caution flags trigger review, and validation gates determine whether a candidate can advance to synthesis planning or experimental testing. **Table 1** summarises constraint-guided optimization components for generative design of natural product-like molecules.

Table 1. Constraint-Guided Generative Optimization for Natural Product-Like Molecular Design: Natural Product-Likeness, Chemical Validity, Scaffold Diversity, Bioactivity Prediction, Target Relevance, Safety Constraints, Synthetic Accessibility, Uncertainty, and Validation Requirements

Optimization component	Core design question	Required input	Constraint type	Potential output	Failure risk	Validation requirement	Architecture role

Natural product-likeness	Does the structure occupy a defensible region of natural-product-related chemical space?	Curated natural-product-related structures, descriptors, scaffolds, motifs, and comparison sets	Soft multidimensional constraint	Natural product-like profile and chemical-space neighbourhood	Superficial resemblance treated as proof of natural origin or activity	Descriptor, scaffold, neighbourhood, and expert comparison	Shapes the design domain without assigning therapeutic value
Chemical validity	Can the representation be parsed into a chemically interpretable structure?	Molecular string or graph, atom and bond rules, sanitization procedures	Hard technical gate	Sanitized structure or rejection flag	Syntax-valid but chemically unreasonable valence, charge, or connectivity	Independent cheminformatics validation	First post-generation quality gate
Stereochemical plausibility	Is stereochemistry represented and internally consistent?	Chiral-centre, double-bond, scaffold, and three-dimensional context	Hard representation gate with soft plausibility review	Stereochemically specified candidate or unresolved-status flag	Lost, contradictory, or implausible stereochemical assignments	Stereochemical parsing, expert review, and eventual analytical confirmation	Prevents stereochemistry-blind advancement
Scaffold diversity	Does the candidate set explore distinct cores rather than repetitive analogues?	Scaffold definitions, cluster assignments, training set, generation batches	Population-level diversity constraint	Scaffold-diverse candidate portfolio	Mode collapse or inflated diversity from minor substitutions	Scaffold clustering and comparison with training and generated sets	Maintains portfolio-level chemical exploration
Novelty	Is the candidate distinct from known and training-set structures without becoming implausible?	Training records, reference collections, fingerprints, scaffold and exact-match searches	Soft novelty objective with anti-memorization checks	Novelty profile and nearest-neighbour record	Rediscovery presented as novelty or excessive extrapolation	Exact-match, scaffold, substructure, and similarity analysis	Balances exploration with evidential support
Bioactivity prediction	Does an endpoint-specific model support experimental prioritization?	Candidate representation, validated predictor, training domain, uncertainty information	Soft predictive objective with abstention conditions	Predicted activity distribution and confidence	Predictor exploitation, confounding, or out-of-domain scoring	External validation and experimental bioactivity testing	Ranks hypotheses without confirming activity
Target relevance	Is the selected target connected to the	Mechanistic, pathway, disease-context, and	Evidence-based inclusion constraint	Target-context evidence profile	Optimizing a tractable but biologically irrelevant endpoint	Biological expert review and context-matched experiments	Connects generation to a defensible biological hypothesis

ISSN (Online) 2249-6084 (Print) 2250-1029

	intended biological context?	assay evidence					
Selectivity consideration	Is predicted activity separable from related targets and known liability targets?	Comparative predictors, homolog information, counter-screen endpoints	Relative objective and negative-design constraint	Predicted selectivity and liability profile	Single-target optimization producing promiscuity	Experimental counter-screening	Prevents potency-only prioritization
ADMET prediction	Are endpoint-specific predicted liabilities acceptable for continued evaluation?	Validated endpoint models, applicability domains, candidate descriptors	Multiple soft constraints with caution flags	Endpoint-level ADMET profile	Opaque aggregation or overconfident extrapolation	Experimental ADME and toxicity studies	Provides early liability triage
Toxicity alert	Does the candidate trigger structural or predictive toxicity concerns?	Alert rules, endpoint models, reactive-metabolite knowledge	Conservative flag or conditional veto	Toxicity-alert dossier	Absence of alerts interpreted as safety or broad alerts causing false rejection	Toxicology review and experimental testing	Provides warning signals without certifying safety
Off-target liability	Could undesirable target interactions be plausible?	Liability-target panels, predictive models, similarity and pharmacology evidence	Negative-design objective and review gate	Prioritized off-target testing plan	Sparse coverage or bias toward well-characterized targets	Experimental profiling	Identifies liabilities not represented by the intended-target score
Synthetic accessibility	Is the structure broadly compatible with established chemical synthesis patterns?	Structural complexity, fragment precedent, heuristic accessibility indicators	Preliminary soft filter	Accessibility estimate and structural concern list	Heuristic estimate treated as proof of synthesis	Medicinal-chemistry review	Low-cost early synthesis triage
Retrosynthetic plausibility	Can credible routes be proposed from reasonable precursors?	Candidate structure, retrosynthesis model, reaction knowledge, precursor sources	Route-based quality gate	Ranked route hypotheses and bottlenecks	Unsupported disconnections, unavailable precursors, or circular routes	Expert route review	Provides stronger feasibility evidence than a heuristic score
Reaction feasibility	Are individual route steps	Proposed reactants, transformations,	Step-level gate with uncertainty	Predicted products, confidence, and	Competing products, missing conditions, or	Experimental reaction testing	Evaluates the internal consistency

	chemically credible?	conditions, and forward reaction model		stereochemical warnings	poor chemoselectivity		of proposed routes
Uncertainty estimation	How reliable is each prediction for the candidate under evaluation?	Predictive distributions, ensembles, calibration and domain information	Risk-control constraint with abstention zone	Confidence interval, calibration record, and out-of-domain flag	Poor calibration or conflation of distinct uncertainty sources	Independent calibration and scaffold-aware assessment	Controls advancement and triggers review or abstention
Multi-objective scoring	How should competing objectives be coordinated without hiding serious liabilities?	Normalized component outputs, desirability functions, hard-gate status, uncertainty	Composite scoring after mandatory gates	Candidate ranking with component-level scorecard	Weight sensitivity, compensation of hard failures, and reward exploitation	Sensitivity analysis and preservation of raw outputs	Coordinates optimization while retaining interpretability
Pareto prioritization	Which candidates provide non-dominated trade-offs among competing objectives?	Objective vectors, uncertainty ranges, and gate status	Non-compensatory comparative prioritization	Pareto candidate set	Unstable or excessively large frontier based on unreliable models	Robustness testing under score perturbation	Presents alternative trade-offs rather than one asserted optimum
Expert review	Does the candidate remain chemically, biologically, and practically defensible?	Complete provenance, prediction, alert, route, and uncertainty dossier	Mandatory human quality gate	Advance, redesign, defer, or reject decision	Automation bias, undocumented preference, or incomplete expertise	Multidisciplinary documented review	Final pre-experimental prioritization gate

Bioactivity, safety, and synthetic accessibility

Bioactivity prediction can help prioritize generated structures for testing, but its interpretation must remain conditional on target relevance, endpoint definition, model applicability, uncertainty, and assay context. A high predicted value is not equivalent to experimentally confirmed binding, functional modulation, cellular activity, or therapeutic effect. Prospective integration of generative design with compound production and biological testing demonstrates how selected computational proposals can progress beyond virtual prioritization, while also showing that only synthesized and experimentally evaluated candidates acquire direct evidence within the measured context [15]. In the proposed architecture, the bioactivity layer therefore records the intended target or phenotype, the biological rationale, the predictor's training domain, the candidate's applicability status, the uncertainty estimate, and the experimental assay needed to evaluate the hypothesis. Selectivity and off-target liabilities are treated as separate considerations because optimization toward one endpoint can create undesirable activity elsewhere.

Synthetic accessibility must likewise be separated from synthetic feasibility. Heuristic accessibility measures can identify structural complexity or unusual features, but analyses of generative-model outputs show that apparently favourable scores may not correspond to plausible routes identified by synthesis-planning systems [16]. The architecture consequently applies synthetic assessment in stages: an initial structural-complexity filter, retrosynthetic planning, precursor and reaction-precedent review, step-level forward prediction, and medicinal-chemistry

evaluation. Generation from retrosynthetically prepared building blocks offers a more synthesis-aware construction strategy by limiting assembly to chemically grounded components, although the existence of building blocks does not establish that every proposed connection can be executed under practical laboratory conditions [17]. A candidate is therefore described as synthesis-prioritized or route-supported, not as synthesizable, until a route has been reviewed and experimentally attempted.

Reaction and safety evaluation require comparable caution. Forward reaction prediction can support step-level assessment and provide confidence estimates for proposed transformations, but confidence remains dependent on the reaction domain represented in the training data [18]. Retrosynthetic search can rank multistep pathway hypotheses, yet route recovery does not prove reagent compatibility, chemoselectivity, stereochemical control, purification feasibility, or laboratory yield [19]. ADMET and toxicity models can identify potential absorption, distribution, metabolism, excretion, and toxicity liabilities, but their outputs remain endpoint-specific predictions rather than safety validation [20]. The proposed safety layer also includes structural alerts, predicted metabolic liabilities, selectivity concerns, and off-target review, with uncertainty-based caution flags preventing unsupported advancement. **Figure 1** illustrates how bioactivity, safety, and synthetic-accessibility constraints can be integrated into a generative optimization workflow for natural product-like molecules.

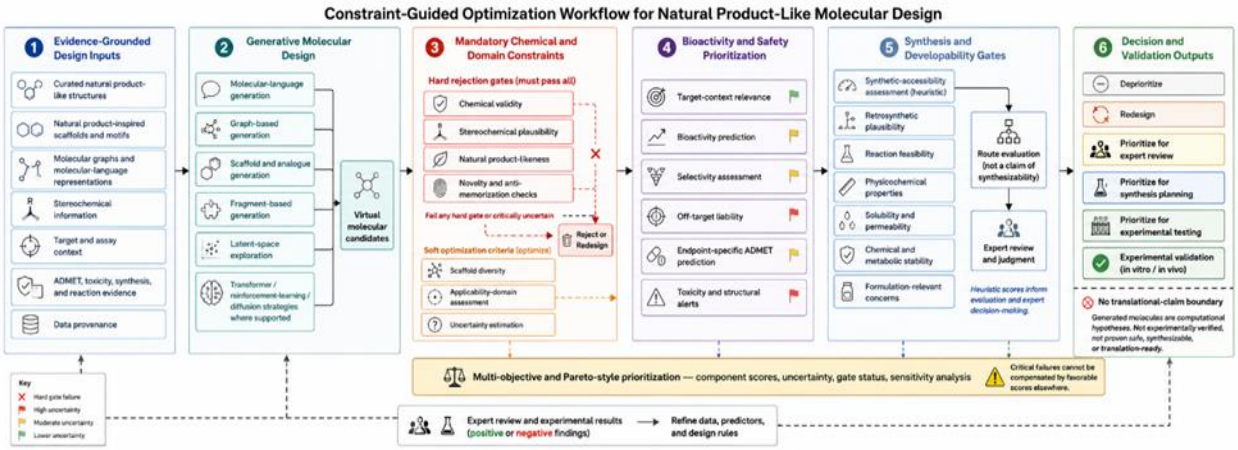

Figure 1. Constraint-Guided Optimization Workflow for Natural Product-Like Molecular Design

Alt-text description

A conceptual workflow showing natural product-like molecule generation passing through constraint layers for chemical validity, natural product-likeness, bioactivity, safety, synthetic accessibility, retrosynthesis, uncertainty, expert review, and validation gates.

An integrated decision rule must prevent one favourable computational dimension from compensating for a critical failure in another. Strong predicted bioactivity cannot override chemical invalidity, unresolved toxicophores, severe uncertainty, or the absence of a defensible synthesis hypothesis. Conversely, a plausible route does not justify claims of activity, safety, or developability. Candidates that pass the computational gates may be prioritized for expert review, synthesis planning, or experimental testing, whereas structures with remediable deficiencies may return to the generator for redesign. Advancement beyond hypothesis generation requires successful synthesis, identity and purity confirmation, biological testing, experimental ADME assessment, and relevant toxicity evaluation. Until those steps are completed, bioactivity, safety, and synthesis-related outputs remain prioritization evidence rather than established molecular properties.

Developability filtering

Developability filtering evaluates whether a generated structure warrants allocation of resources for synthesis and experimental investigation. It extends beyond generic drug-likeness by considering the intended route of administration, biological context, chemical class, scaffold complexity, and stage of discovery. Absorption, distribution, metabolism, excretion, toxicity, and physicochemical properties represent distinct modelling tasks with different datasets, applicability domains, and experimental correlates [21]. The proposed architecture therefore avoids treating developability as a single predicted attribute. Instead, it produces an endpoint-level dossier covering molecular size, lipophilicity, polarity, charge state, hydrogen-bonding capacity, conformational flexibility, solubility, permeability, chemical stability, metabolic liability, and formulation-relevant concerns where supporting information is available.

ISSN (Online) 2249-6084 (Print) 2250-1029

Predictive developability gates must be evaluated using endpoint-appropriate data partitions, metrics, and applicability criteria. Molecular-learning benchmarks demonstrate that performance varies substantially across biochemical, toxicological, pharmacokinetic, and physicochemical tasks and that conclusions depend on the representation, dataset, and split strategy used [22]. Within the proposed workflow, every predictive module is therefore linked to its training provenance, endpoint definition, validation design, domain limitations, and intended decision role. A permeability model cannot substitute for a solubility model, a toxicity classifier cannot establish general safety, and calculated physicochemical properties cannot determine formulation behaviour without experimental evidence. Candidates with missing or unreliable endpoint coverage are not assumed to satisfy the corresponding criterion; the missing evidence is recorded as an unresolved development risk.

Uncertainty estimation is integrated into each developability gate because point predictions can obscure model disagreement, domain extrapolation, and limited evidential support. Comparative evaluations of uncertainty methods for molecular property prediction show that confidence estimation must itself be assessed for calibration and reliability [23]. The architecture consequently distinguishes predicted endpoint value, model confidence, applicability-domain status, and evidence completeness. High uncertainty can trigger abstention, expert review, acquisition of additional data, or redesign rather than automatic rejection. This is particularly important for natural product-like structures containing uncommon scaffolds, dense stereochemistry, unusual ring systems, or physicochemical profiles that are poorly represented in conventional drug-discovery datasets.

Data leakage and benchmark bias can create misleading confidence in predictive filters. Ligand-based evaluations may reward recognition of closely related structures rather than genuine generalization to new scaffolds [24]. The proposed architecture therefore requires exact-structure, analogue, scaffold, and similarity-aware separation between training, validation, and evaluation records, together with explicit checks for duplicated measurements and inconsistent annotations. Even after these controls, retrospective model performance does not establish prospective utility, experimental developability, or translational readiness [25]. Developability filtering is consequently positioned as a prioritization mechanism: it can identify apparent liabilities, compare candidate trade-offs, and define experiments, but it cannot establish that a molecule will possess acceptable exposure, stability, safety, formulation behaviour, or therapeutic utility.

Table 2 maps bioactivity, safety, synthetic accessibility, and developability gates required before generated natural product-like molecules can be prioritized for experimental follow-up.

Table 2. Bioactivity, Safety, Synthetic Accessibility, and Developability Gates for Generated Natural Product-Like Molecules: Target Relevance, Selectivity, ADMET Prediction, Toxicity Alerts, Retrosynthetic Plausibility, Physicochemical Filters, Stability, Solubility, Permeability, and Validation Boundaries

Evaluation gate	Core question	Evidence needed	Architecture input	Architecture output	Caution signal	Validation requirement	Decision boundary
Bioactivity prediction	Is there sufficient computational support to prioritize the candidate for a defined biological test?	Endpoint-specific training evidence, applicability domain, suitable validation, uncertainty	Candidate representation and validated activity predictor	Predicted activity distribution, confidence, and domain flag	High predicted activity with weak applicability or close training analogues	Relevant biochemical, biophysical, cellular, or phenotypic assay	May prioritize testing; must not be described as experimentally active
Target-context relevance	Is the intended target or phenotype connected to a defensible biological mechanism?	Mechanistic, pathway, disease-context, and assay evidence	Target hypothesis and intended biological context	Context-evidence profile and unresolved assumptions	Target selected only because a predictor is available	Biological expert review and context-matched experimental model	Supports a biological hypothesis, not therapeutic confirmation
Selectivity screening	Is the predicted activity distinguishab	Comparative target models, homolog	Candidate and target/off-target panel	Predicted selectivity and liability matrix	Missing homolog coverage or uniformly	Experimental counter-screening at relevant	Guides testing; does not establish selectivity

	le from activity at related or undesirable targets?	information, counter-screen data			high predicted activity	concentrations	
ADMET prediction	Are endpoint-specific pharmacokinetic and toxicity predictions compatible with continued investigation?	Validated endpoint models, applicability information, uncertainty	Candidate descriptors, graph, or molecular representation	Endpoint-level ADMET profile	One favourable aggregate score masking adverse components	Experimental ADME and toxicology studies appropriate to stage	May filter or flag; cannot establish safety or exposure
Toxicity screening	Does the candidate trigger structural or learned toxicity warnings?	Structural alerts, endpoint models, reactive-metabolite knowledge	Candidate structure and predicted metabolites where appropriate	Toxicity-alert dossier with confidence	Absence of alerts interpreted as proof of safety	Expert toxicology review and experimental testing	Serious alerts may pause advancement; no-alert status is not safety clearance
Off-target liability review	Are undesirable pharmacological interactions plausible?	Target-panel predictions, similarity evidence, known liability mechanisms	Candidate and liability-target panel	Prioritized off-target testing plan	Sparse target coverage or disagreement among models	Experimental profiling for advanced candidates	Defines additional tests; cannot confirm absence of liabilities
Synthetic accessibility	Is the structure broadly compatible with established synthetic patterns?	Complexity, ring burden, stereochemical burden, fragment precedent, heuristic measures	Candidate structure	Preliminary accessibility estimate and concern list	Favourable score despite unusual bonds, strained systems, or dense stereochemistry	Medicinal-chemistry inspection	Permits route analysis only; does not prove synthesizability
Retrosynthetic plausibility	Can credible routes be proposed from reasonable precursors?	Ranked routes, reaction precedent, precursor availability, route diversity	Candidate structure and retrosynthesis system	Route hypotheses, bottlenecks, and confidence	Only low-confidence routes or unsupported disconnections	Expert retrosynthetic review	Supports prioritization but not laboratory execution
Reaction feasibility	Are proposed route steps chemically credible under plausible conditions?	Forward prediction, reaction precedent, condition compatibility, chemoselecti	Proposed reactants, products, and transformations	Step-level feasibility and uncertainty record	Competing products, incompatible groups, or low-confidence prediction	Laboratory reaction testing	Computational feasibility is insufficient for a synthesis claim

		vity assessment					
Physicochemical property filter	Are calculated properties compatible with the intended discovery context?	Molecular weight, lipophilicity, polarity, ionization, hydrogen bonding, flexibility, shape	Candidate descriptors	Property profile with contextual deviations	Universal cut-offs applied without regard to scaffold or route	Experimental measurement for selected candidates	Contextual filter rather than absolute drug-likeness verdict
Stability consideration	Could the structure remain chemically and physically stable during handling, storage, and testing?	Labile-group analysis, degradation predictions, anticipated conditions	Candidate structure and use conditions	Stability risks and testing plan	Hydrolysis-prone, oxidation-prone, photolabile, or isomerization-prone features	Forced-degradation and stability studies	Unresolved instability requires testing or redesign
Solubility consideration	Could inadequate solubility limit assay interpretation or exposure?	Calculated solubility, ionization state, solid-state risk, formulation context	Candidate and physicochemical profile	Predicted solubility range with uncertainty	Model disagreement, high lipophilicity, or strong lattice-risk features	Kinetic and thermodynamic solubility measurement	Guides formulation and triage; does not establish usable solubility
Permeability consideration	Is transport across the relevant biological barrier plausible?	Permeability models, ionization, polarity, transporter considerations	Candidate and biological context	Predicted permeability and caution flags	High polarity, high size, efflux risk, or model-domain mismatch	Relevant cell-based or tissue permeability assay	Supports test selection; does not establish exposure
Metabolic liability consideration	Are rapid clearance, unstable sites, or problematic metabolites plausible?	Metabolism models, enzyme interaction information, soft-spot and metabolite analysis	Candidate structure	Metabolic-liability map and testing priorities	Multiple predicted soft spots or uncertain reactive metabolites	Microsomal or hepatocyte stability and metabolite identification	Guides redesign or testing; does not establish metabolic stability
Formulation awareness	Are there early signs that delivery or formulation may be unusually difficult?	Solubility, solid-state, stability, ionization, route-of-administration context	Integrated physicochemical dossier	Formulation-risk flags	Formulation assumptions made without empirical measurements	Stage-appropriate formulation studies	Supports planning only; does not establish dosage-form feasibility
Developability filter	Does the combined evidence justify experimental resource allocation?	All endpoint profiles, uncertainty, synthesis review, safety flags,	Complete candidate dossier	Advance, redesign, defer, or reject recommendation	Composite score concealing a hard failure or missing critical evidence	Multidisciplinary review and stage-appropriate experiments	Developability remains provisional until empirically supported

ISSN (Online) 2249-6084 (Print) 2250-1029

		expert judgment					
Experimental confirmation need	Which computational hypotheses require testing before stronger claims are permitted?	Claim-to-assay mapping, analytical methods, controls, acceptance criteria	Candidate and prediction record	Design–make–test plan	Advancement without synthesis, identity confirmation, or relevant controls	Synthesis, purification, characterization, bioactivity, ADME, and toxicity testing	No experimentally dependent claim before corresponding evidence
No translational-claim boundary	Has the evidence crossed the threshold for a translational or clinical claim?	Reproducible preclinical, exposure, safety, efficacy, manufacturing, and later clinical evidence	Complete evidence chain	Explicit claim limitation	Use of "safe," "clinically useful," or "translation-ready" based on predictions	Appropriate preclinical and clinical development	The conceptual architecture permits no clinical or regulatory-readiness claim

Proposed generative architecture

The proposed architecture begins with a provenance-controlled data foundation and a representation layer capable of preserving complementary aspects of molecular structure. The data foundation includes curated natural product-like structures, natural product-inspired scaffolds, stereochemical annotations, assay and target records, ADMET observations, toxicity alerts, synthetic precedent, reaction information, and explicit source provenance. Molecular representations may include graphs, strings, fingerprints, scaffold encodings, fragment relationships, and stereochemistry-aware features. Robust molecular string representations can reduce syntax-level invalidity, but representation validity does not establish chemical realism, stability, bioactivity, safety, or synthesizability [26]. The architecture therefore separates representation-level acceptance from chemical-validity and downstream evidence gates.

Training-data provenance and coverage are treated as active architecture components rather than passive documentation. Generative exploration depends on which structures, scaffolds, stereochemical patterns, and property distributions are represented in the training corpus. Studies of generative exploration in defined chemical spaces show that training-set composition influences both reproduced regions and apparent extrapolation [27]. The architecture consequently records exact structures, scaffold families, analogue relationships, source collections, assay contexts, missingness patterns, and uncertainty in annotations. It also applies deduplication, standardization, stereochemical quality control, scaffold-aware data partitioning, and bias assessment before model development. Candidates that lie far from the evidence-supported training domain receive explicit extrapolation warnings rather than being interpreted as superior merely because they are novel.

The generative engine is proposed as a modular layer rather than a single mandatory algorithm. It may include molecular-string generation, graph generation, scaffold generation, analogue generation, fragment assembly, latent-space exploration, transformer-based generation, reinforcement-learning optimization, or diffusion-based generation when supported by suitable evidence and representation choices. Property-driven design literature indicates that biological objectives, molecular properties, and liability-related criteria can be incorporated into generative workflows, but these objectives remain dependent on the predictive models used to calculate them [28]. The architecture therefore routes all generated candidates through an independent constraint layer containing chemical validity, stereochemical plausibility, natural product-likeness, novelty, scaffold diversity, target relevance, selectivity considerations, and uncertainty controls before multi-objective prioritization.

The representation and generation modules must also preserve structural controllability. Chemical language models can support conditional generation and efficient exploration, but they require careful treatment of training data, stereochemistry, molecular syntax, optimization objectives, and experimental boundaries [29]. Fragment-aware representations can further support scaffold decoration, fragment linking, and constrained analogue design [30]. In the proposed architecture, these strategies are complementary: a scaffold-preserving generator may be used when an evidence-supported core should be retained, whereas a broader generator may be used when scaffold exploration is the objective. Neither mode bypasses chemical sanitization, novelty analysis, synthesis-related assessment, or expert review.

The final architecture stages integrate multi-objective scoring, safety and synthesis filters, developability gates,

expert curation, validation planning, and feedback. Activity, target relevance, selectivity, ADMET, toxicity, off-target liability, natural product-likeness, novelty, synthesis-related criteria, and developability indicators remain visible as separate evidence channels. Multi-objective or Pareto-style prioritization can expose candidate trade-offs rather than asserting that one composite value defines molecular quality, while expert judgment remains necessary to interpret chemical and project-specific significance [31]. Candidates may be deprioritized, redesigned, advanced for route planning, or nominated for experimental testing, but automated de novo design does not remove the need for medicinal chemistry, synthesis, analytical characterization, biological assays, ADME experiments, or safety evaluation [32]. Validation results, including negative findings, are proposed to feed back into data curation, predictive models, constraints, and subsequent generation cycles.

Table 3 presents the proposed generative AI architecture for constraint-guided design of natural product-like molecules.

Table 3. Proposed Generative AI Architecture for Natural Product-Like Molecule Design: Data Foundation, Molecular Representation, Generative Engine, Constraint Layer, Multi-Objective Scoring, Safety and Synthetic-Accessibility Filters, Developability Gates, Expert Review, Feedback, and Validation Boundaries

Architecture stage	Core purpose	Required input	Computational function	Potential output	Failure risk	Validation requirement	Feedback mechanism	Decision boundary
Data foundation	Establish an evidence-controlled molecular and annotation corpus	Curated natural product-like structures, inspired scaffolds, stereochemistry, assays, targets, ADMET, toxicity, synthesis, reaction, and provenance records	Standardization, deduplication, annotation harmonization, quality control	Provenance-aware training and evaluation corpus	Duplicates, inconsistent structures, missing stereochemistry, biased endpoint coverage	Manual and automated curation audit	Corrected records and experimental data update the corpus	Poorly documented data are excluded or marked uncertain
Natural product-like domain definition	Define the intended design space without assigning biological meaning	Reference chemical space, scaffolds, motifs, descriptors, comparison sets	Multidimensional domain modelling and neighbourhood analysis	Operational natural product-likeness profile	Binary or superficial classification	Chemical-space and expert review	Domain definition is revised when new evidence becomes available	Likeness is never treated as proof of origin, activity, or value
Molecular representation	Encode connectivity, fragments, scaffolds, and stereochemistry	Standardized molecular records	Graph, molecular-language, fingerprint, scaffold, fragment, and stereochemistry-aware encoding	Model-ready molecular representations	Loss of stereochemistry, invalid syntax, representation bias	Reconstruction, parsing, and stereochemical consistency checks	Representation errors return to curation and encoding design	Representation validity does not establish chemical validity
Generative engine	Propose new structures	Molecular representations,	String, graph, transformer,	Virtual candidate structures	Mode collapse, memorizati	Validity, novelty, diversity,	Rejected candidates and	Generated structures are

	or controlled analogues	scaffold priors, design objectives	latent, reinforcement-learning, diffusion, scaffold, or fragment-based generation where supported		on, invalid structures, unsupported extrapolation	and domain checks	validation results refine generator settings	hypotheses only
Chemical-validity layer	Remove malformed or chemically incoherent structures	Generated molecular strings or graphs	Sanitization, valence, charge, bond, ring, and structural-rule evaluation	Chemically interpretable candidates or rejection flags	Syntax-valid but chemically unreasonable outputs	Independent cheminformatics and chemistry review	Failure patterns modify representation or constraints	Invalid structures cannot advance
Stereochemistry layer	Preserve and assess stereochemical specification	Encoded stereochemistry and scaffold context	Chiral-centre, geometric-isomer, and consistency analysis	Specified candidate or unresolved-stereochemistry flag	Missing, contradictory, or implausible stereochemical assignments	Three-dimensional review and eventual analytical confirmation	Experimental stereochemical findings update rules and data	Unresolved critical stereochemistry blocks strong claims
Natural product-likeness and scaffold layer	Maintain relevant domain character while supporting controlled exploration	Domain model, motifs, scaffolds, reference compounds	Similarity, neighbourhood, motif, scaffold, and complexity assessment	Domain profile and scaffold classification	Copying known structures or rewarding superficial resemblance	Multimetric and expert assessment	Accepted and rejected examples refine constraint ranges	Structural resemblance alone cannot justify advancement
Novelty and diversity layer	Prevent memorization and portfolio collapse	Training records, known structures, generated batches	Exact-match, similarity, scaffold, substructure, and clustering analyses	Novelty and diversity dossier	Rediscovery presented as novelty or trivial analogue inflation	Independent comparison with training and reference collections	Recurrent duplication modifies sampling and penalties	Novelty cannot override chemical or evidential weakness
Bioactivity and target-context layer	Prioritize candidates for relevant biological tests	Endpoint models, target rationale, assay context, candidate representation	Activity prediction, applicability assessment, contextual evidence mapping	Predicted activity distribution and biological-testing rationale	Predictor exploitation, confounding, irrelevant target optimization	External model validation and experimental assays	Assay results recalibrate predictors and objectives	Prediction does not establish activity
Selectivity and off-target layer	Identify competing pharmacology and liability risks	Comparative target and liability models	Target-panel scoring and negative-design assessment	Selectivity and off-target testing plan	Sparse coverage and false reassurance	Experimental counter-screening and profiling	Confirmed liabilities update penalties and	No absence-of-liability claim from prediction alone

							training data	
ADMET and toxicity layer	Identify endpoint-specific pharmacokinetic and toxicity concerns	Validated ADMET models, alert rules, applicability data	Endpoint prediction, structural-alert analysis, uncertainty estimation	ADMET profile and toxicity-warning dossier	Opaque aggregation, out-of-domain prediction, missed liabilities	Experimental ADME and toxicology studies	Experimental results update models and thresholds	Prediction cannot establish safety
Synthetic-accessibility layer	Conduct low-cost preliminary synthesis triage	Candidate structure, complexity and precedent information	Heuristic accessibility and structural-complexity analysis	Accessibility estimate and concern list	Heuristic score mistaken for synthesis proof	Medicinal-chemistry review	Chemist feedback adjusts constraints and penalties	Passing permits route analysis only
Retrosynthesis and reaction layer	Evaluate route and step plausibility	Candidate, reaction knowledge, precursor information	Retrosynthetic search, route ranking, forward reaction prediction	Route hypotheses, bottlenecks, and step confidence	Unsupported disconnections, unavailable precursors, poor chemoselectivity	Expert route review and experimental synthesis	Route success and failure update reaction models and constraints	Computational route plausibility is not synthesis confirmation
Developability layer	Compare physicochemical, exposure, stability, and formulation-related risks	Integrated property, ADMET, safety, and synthesis dossier	Endpoint-level filtering, caution flags, uncertainty-aware decision support	Developability risk profile	Composite score hiding a severe liability or missing evidence	Stage-appropriate experimental measurements	Experimental developability data refine prediction modules	Developability remains provisional
Multi-objective prioritization	Expose trade-offs and rank candidates after mandatory gates	Component scores, uncertainty, gate status, project priorities	Desirability analysis, sensitivity assessment, Pareto prioritization where appropriate	Ranked candidates and alternative trade-off sets	Reward hacking, unstable weights, compensation of hard failure	Robustness and sensitivity analysis	Expert choices and validation results refine objectives	No single score establishes candidate quality
Expert curation	Apply multidisciplinary chemical and biological judgment	Complete candidate dossier and decision history	Structured review and rationale recording	Advance, redesign, defer, or reject decision	Automation bias, undocumented preferences, incomplete expertise	Multidisciplinary review	Expert decisions return to constraints and data curation	Expert review supports prioritization, not proof
Experimental validation	Test structure, synthesis, activity, ADME, and safety hypotheses	Prioritized candidate and test plan	Not a generative computation; coordinates evidence acquisition	Experimental evidence or failure record	Selective reporting or inadequate controls	Synthesis, purification, characterization, biological and safety studies	Positive and negative findings update all upstream layers	Stronger claims require corresponding experiment

								al evidence
Decision and claim boundary	Prevent unsupported translational interpretation	Full computational and experimental evidence chain	Rule-based claim control and evidence-status tracking	Permitted terminology and advancement status	Safety, efficacy, or readiness claims based on prediction	Independent scientific review	Claim rules evolve with evidence standards	No clinical, regulatory, or deployment claim from the conceptual architecture

Figure 2 presents the proposed generative AI architecture for natural product-like molecule design, integrating data foundations, molecular representation, constraint-guided optimization, safety and synthetic-accessibility filters, developability gates, expert review, and validation feedback.

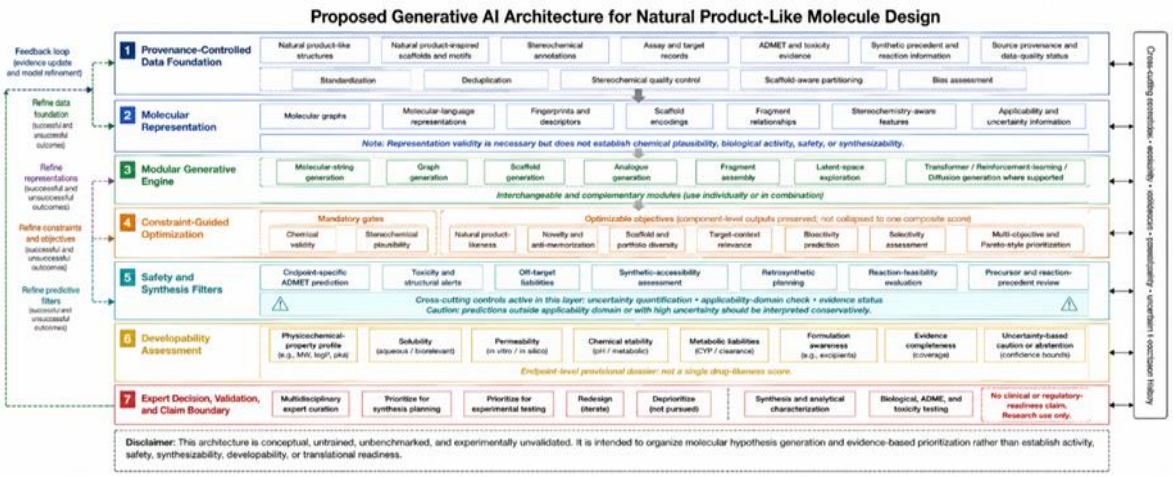

Figure 2. Proposed Generative AI Architecture for Natural Product-Like Molecule Design

Alt-text description

A conceptual architecture diagram showing natural product-like molecular data flowing through representation learning, generative design, constraint layers, multi-objective scoring, bioactivity and safety filters, synthetic accessibility assessment, developability gates, expert review, validation feedback, and decision boundaries.

CONCLUSION

This article proposes a constraint-guided generative AI architecture for natural product-like molecular design that integrates provenance-aware data foundations, stereochemistry-sensitive molecular representation, modular generation, natural product-likeness and chemical-validity controls, bioactivity and target-context prediction, selectivity and safety filtering, synthetic-accessibility and retrosynthetic assessment, developability gates, uncertainty estimation, multi-objective prioritization, expert review, and validation feedback. The architecture is intended to organize molecular ideation and evidence-based prioritization rather than to establish activity, safety, synthesizability, developability, or translational readiness. Generated molecules remain computational hypotheses until they have been synthesized, purified, analytically characterized, biologically tested, experimentally evaluated for ADME and toxicity properties, and reviewed within an appropriate development context. By maintaining explicit distinctions between optimization objectives, predictive outputs, quality gates, experimental evidence, and claim boundaries, the proposed framework provides a cautious and technically structured basis for natural product-like generative design without replacing medicinal chemistry, pharmacology, toxicology, or experimental validation.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

REFERENCES

[1] Sánchez-Lengeling B, Aspuru-Guzik A. Inverse molecular design using machine learning: Generative models for matter engineering. Science. 2018;361(6400):360-5. doi:10.1126/science.aat2663

[2] Elton DC, Boukouvalas Z, Fuge MD, Chung PW. Deep learning for molecular design—a review of the state of the art. Mol Syst Des Eng. 2019;4(4):828-49. doi:10.1039/C9ME00039A

[3] Xue D, Gong Y, Yang Z, Chuai G, Qu S, Shen A, et al. Advances and challenges in deep generative models for de novo molecule generation. Wiley Interdiscip Rev Comput Mol Sci. 2019;9(3). doi:10.1002/wcms.1395

[4] Meyers J, Fabian B, Brown N. De novo molecular design and generative models. Drug Discov Today. 2021;26(11):2707-15. doi:10.1016/j.drudis.2021.05.019

[5] Merk D, Grisoni F, Friedrich L, Schneider G. Tuning artificial intelligence on the de novo design of natural-product-inspired retinoid X receptor modulators. Commun Chem. 2018;1:68. doi:10.1038/s42004-018-0068-1

[6] Zabolotna Y, Ertl P, Horvath D, Bonachera F, Marcou G, Varnek A. NP Navigator: A new look at the natural product chemical space. Mol Inform. 2021;40(9):2100068. doi:10.1002/minf.202100068

[7] Tay DWP, Yeo NZX, Adaikkappan K, Lim YH, Ang SJ. 67 million natural product-like compound database generated via molecular language processing. Sci Data. 2023;10(1):296. doi:10.1038/s41597-023-02207-x

[8] Shen X, Zeng T, Chen N, Li J, Wu R. NIMO: A natural product-inspired molecular generative model based on conditional transformer. Molecules. 2024;29(8):1867. doi:10.3390/molecules29081867

[9] Gómez-Bombarelli R, Wei JN, Duvenaud D, Hernández-Lobato JM, Sánchez-Lengeling B, Sheberla D, et al. Automatic chemical design using a data-driven continuous representation of molecules. ACS Cent Sci. 2018;4(2):268-76. doi:10.1021/acscentsci.7b00572

[10] Olivecrona M, Blaschke T, Engkvist O, Chen H. Molecular de-novo design through deep reinforcement learning. J Cheminform. 2017;9(1):48. doi:10.1186/s13321-017-0235-x

[11] Popova M, Isayev O, Tropsha A. Deep reinforcement learning for de novo drug design. Sci Adv. 2018;4(7). doi:10.1126/sciadv.aap7885

[12] Ståhl N, Falkman G, Karlsson A, Mathiason G, Boström J. Deep reinforcement learning for multiparameter optimization in de novo drug design. J Chem Inf Model. 2019;59(7):3166-76. doi:10.1021/acs.jcim.9b00325

[13] Brown N, Fiscato M, Segler MHS, Vaucher AC. GuacaMol: Benchmarking models for de novo molecular design. J Chem Inf Model. 2019;59(3):1096-108. doi:10.1021/acs.jcim.8b00839

[14] Polykovskiy D, Zhebrak A, Sánchez-Lengeling B, Golovanov S, Tatanov O, Belyaev S, et al. Molecular Sets (MOSES): A benchmarking platform for molecular generation models. Front Pharmacol. 2020;11:565644. doi:10.3389/fphar.2020.565644

[15] Grisoni F, Huisman BJH, Button AL, Moret M, Atz K, Merk D, et al. Combining generative artificial intelligence and on-chip synthesis for de novo drug design. Sci Adv. 2021;7(24). doi:10.1126/sciadv.abg3338

[16] Gao W, Coley CW. The synthesizability of molecules proposed by generative models. J Chem Inf Model. 2020;60(12):5714-23. doi:10.1021/acs.jcim.0c00174

[17] Seo S, Lim J, Kim WY. Molecular generative model via retrosynthetically prepared chemical building block assembly. Adv Sci (Weinh). 2023;10(8):2206674. doi:10.1002/advs.202206674

[18] Schwaller P, Laino T, Gaudin T, Bolgar P, Hunter CA, Bekas C, et al. Molecular Transformer: A model for uncertainty-calibrated chemical reaction prediction. ACS Cent Sci. 2019;5(9):1572-83. doi:10.1021/acscentsci.9b00576

[19] Schwaller P, Petraglia R, Zullo V, Nair VH, Haeuselmann RA, Pisoni R, et al. Predicting retrosynthetic pathways using transformer-based models and a hyper-graph exploration strategy. Chem Sci. 2020;11(12):3316-25. doi:10.1039/C9SC05704H

[20] Yang H, Lou C, Sun L, Li J, Cai Y, Wang Z, et al. admetSAR 2.0: Web-service for prediction and optimization of chemical ADMET properties. Bioinformatics. 2019;35(6):1067-9. doi:10.1093/bioinformatics/bty707

[21] Ferreira LLG, Andricopulo AD. ADMET modeling approaches in drug discovery. Drug Discov Today. 2019;24(5):1157-65. doi:10.1016/j.drudis.2019.03.015

[22] Wu Z, Ramsundar B, Feinberg EN, Gomes J, Geniesse C, Pappu AS, et al. MoleculeNet: A benchmark for molecular machine learning. Chem Sci. 2018;9(2):513-30. doi:10.1039/C7SC02664A

[23] Scalia G, Grambow CA, Pernici B, Li YP, Green WH. Evaluating scalable uncertainty estimation methods for deep learning-based molecular property prediction. J Chem Inf Model. 2020;60(6):2697-717. doi:10.1021/acs.jcim.9b00975

[24] Wallach I, Heifets A. Most ligand-based classification benchmarks reward memorization rather than

generalization. J Chem Inf Model. 2018;58(5):916-32. doi:10.1021/acs.jcim.7b00403

[25] Bender A, Cortés-Ciriano I. Artificial intelligence in drug discovery: What is realistic, what are illusions? Part 1: Ways to make an impact, and why we are not there yet. Drug Discov Today. 2021;26(2):511-24. doi:10.1016/j.drudis.2020.12.009

[26] Krenn M, Häse F, Nigam AK, Friederich P, Aspuru-Guzik A. Self-referencing embedded strings (SELFIES): A 100% robust molecular string representation. Mach Learn Sci Technol. 2020;1(4):045024. doi:10.1088/2632-2153/aba947

[27] Arús-Pous J, Blaschke T, Ulander S, Reymond JL, Chen H, Engkvist O. Exploring the GDB-13 chemical space using deep generative models. J Cheminform. 2019;11(1):20. doi:10.1186/s13321-019-0341-z

[28] Born J, Manica M. Trends in deep learning for property-driven drug design. Curr Med Chem. 2021;28(38):7862-86. doi:10.2174/0929867328666210729115728

[29] Grisoni F. Chemical language models for de novo drug design: Challenges and opportunities. Curr Opin Struct Biol. 2023;79:102527. doi:10.1016/j.sbi.2023.102527

[30] Noutahi E, Gabellini C, Craig M, Lim JSC, Tossou P. Gotta be SAFE: A new framework for molecular design. Digit Discov. 2024;3(4):796-804. doi:10.1039/D4DD00019F

[31] van den Broek RL, Patel S, van Westen GJP, Jespers W, Sherman W. In search of beautiful molecules: A perspective on generative modeling for drug design. J Chem Inf Model. 2025;65(18):9383-97. doi:10.1021/acs.jcim.5c01203

[32] Schneider G, Clark DE. Automated de novo drug design: Are we nearly there yet? Angew Chem Int Ed Engl. 2019;58(32):10792-803. doi:10.1002/anie.201814681