Pharmacophore

ISSN-2229-5402

Journal home page: http://www.pharmacophorejournal.com

DESIGNING MOLECULES UNDER THE FULL WEIGHT OF POTENCY, SELECTIVITY, SYNTHESIZABILITY, SAFETY, AND PHARMACEUTICAL DEVELOPABILITY

Jean-Baptiste Nzouankeu[1]*, Lucy Mwangi[2], Pierre Abega[3], Claire Dupont[4]

1. ***Department of Multidimensional Molecular Design, Faculty of Pharmacy, University of Yaoundé I, Yaoundé, Cameroon.***
2. ***Department of Potency and Selectivity Optimization, Institute of Tropical Medicine, University of Nairobi, Nairobi, Kenya.***
3. ***Department of Synthesizability and Process Chemistry, Faculty of Science, University of Douala, Douala, Cameroon.***
4. ***Department of Safety and Developability Assessment, Faculty of Public Health, Sorbonne University, Paris, France.***

ARTICLE INFO

Received:
12 May 2024
Received in revised form:
25 September 2024
Accepted:
25 September 2024
Available online:
28 October 2024

Keywords: De novo molecular design, Multiparameter optimization, Pharmaceutical developability, Synthesizability, Retrosynthetic accessibility, Uncertainty quantification

ABSTRACT

Computational molecular design can generate chemically novel structures and optimize predicted properties at a scale that exceeds conventional manual ideation. Yet the production of valid or high-scoring structures does not resolve the central pharmaceutical problem: a molecule must satisfy multiple interacting requirements before it can become an experimentally credible candidate. Potency may conflict with selectivity, exposure, safety, solubility, stability, synthetic tractability, route practicality, and material availability. Moreover, confidence in a predicted property does not establish that the underlying model is applicable to the proposed structure or that an acceptable synthesis and development path exists. This article develops an original constraint-aware account of molecular design in which pharmaceutical relevance is determined by non-substitutable feasibility conditions, conditional optimization objectives, explicit uncertainty, and evidence-dependent decision gates. The proposed Full-Weight Pharmaceutical Design Hierarchy distinguishes structural admissibility, pharmacological relevance, selectivity and safety boundaries, synthetic and material feasibility, developability requirements, preference-sensitive optimization, and experimental confirmation. It further separates molecule-level plausibility from route-level accessibility and prevents favorable performance on one objective from compensating for failure of a critical constraint. The contribution is conceptual and methodological rather than empirically validated: it organizes how computational outputs may be evaluated, compared, rejected, or advanced without treating a composite score as evidence of pharmaceutical readiness. Its applicability remains conditional on target biology, therapeutic modality, assay quality, model calibration, product concept, organizational capabilities, and access to experimental expertise. By repositioning generation as the production of testable pharmaceutical hypotheses rather than candidate drugs, the article provides a basis for more defensible model evaluation, interdisciplinary decision-making, and prospective validation of constraint-aware molecular design.

To Cite This Article: Nzouankeu JB, Mwangi L, Abega P, Dupont C. Designing Molecules under the Full Weight of Potency, Selectivity, Synthesizability, Safety, and Pharmaceutical Developability. Pharmacophore. 2024;15(5):37-47. https://doi.org/10.51847/bhejqkbZml

Introduction

Artificial intelligence and machine-learning methods now support activities spanning target analysis, virtual screening, molecular generation, property prediction, and development planning, but their pharmaceutical value remains conditional on suitable data, task-specific validation, and experimental adjudication; a generated molecular structure is therefore a testable hypothesis rather than a drug candidate [1–3]. This distinction is central to the scientific interpretation of computational design. A molecular representation can be formally valid, different from known training examples, and highly ranked by one or more prediction models while remaining pharmacologically irrelevant, synthetically inaccessible, unsafe, unstable, or incompatible with the intended dosage form. The ability to search chemical space more rapidly consequently enlarges the set of hypotheses

Corresponding Author: Jean-Baptiste Nzouankeu; Department of Multidimensional Molecular Design, Faculty of Pharmacy, University of Yaoundé I, Yaoundé, Cameroon. E-mail: jb.nzouankeu@uy1.uninet.cm.

that can be considered, but it does not relax the evidentiary conditions under which any individual structure should be advanced.

Modern computational methods can streamline hypothesis generation and prioritization, yet they remain components of integrated discovery programs rather than substitutes for medicinal chemistry, pharmacology, toxicology, pharmaceutics, process chemistry, and experimental decision-making [4]. Their contribution depends not only on predictive accuracy in an abstract statistical sense but also on whether the prediction target corresponds to the intended pharmaceutical question. A potency model trained on heterogeneous assay data, for example, may not describe the relevant target state, cellular context, exposure condition, or mechanism. A retrosynthetic score may suggest accessibility without establishing reaction success, isolation, purity, scale, or supply continuity. Computational efficiency therefore has pharmaceutical meaning only when the scope of each model, the evidence available for the proposed structure, and the consequences of error are made explicit.

The unresolved problem is particularly acute in generative molecular design. Generators are often assessed through chemical validity, uniqueness, novelty, diversity, similarity to a reference distribution, or improvement in an optimization score. These measures are useful for testing particular computational capabilities, but generative output must ultimately be judged by medicinal-chemistry relevance rather than novelty alone [5]. Novelty can indicate that a model is not merely reproducing its training set, yet unfamiliarity may also move a structure away from reliable property estimates, established structure–activity relationships, known synthetic precedents, and characterized chemical matter. Conversely, a less novel analogue may be more informative if it tests a specific pharmacological hypothesis, resolves a known series liability, or can be synthesized rapidly from available materials. The scientific value of generation thus lies not in distance from prior molecules as such, but in the quality and testability of the pharmaceutical hypothesis produced.

This article proposes the Full-Weight Pharmaceutical Design Hierarchy, a conceptual organization of molecular design under interacting and non-substitutable constraints. The hierarchy treats structural admissibility, pharmacological relevance, selectivity, safety, synthetic feasibility, route accessibility, material availability, and developability as distinct conditions whose meanings cannot be collapsed into a single reward without loss of scientific information. Only after critical feasibility boundaries have been met should preference-sensitive optimization be used to compare acceptable alternatives. Uncertainty and evidence maturity modify every level, while staged decision gates determine whether an output should remain a generated structure, become a synthesis proposal, enter experimental testing, or be considered an experimentally credible candidate. The contribution is not presented as a validated algorithm or universal selection rule. Rather, it specifies an architecture for asking which constraints must be satisfied, which trade-offs may legitimately be negotiated, what evidence is required, and where computational claims must stop.

Why Pharmaceutical Design Cannot Be Reduced to Molecular Novelty

Molecular novelty is a relation between a proposed structure and a defined comparison set; it is not a direct measure of therapeutic value. Its interpretation depends on the molecular representation, similarity function, reference database, stereochemical treatment, and threshold used to define difference. Benchmarking work has appropriately distinguished validity, uniqueness, novelty, and goal achievement as separate dimensions of de novo molecular generation [6]. This separation is scientifically important because each dimension answers a different question. Validity asks whether a representation conforms to specified chemical rules. Uniqueness measures duplication within an output set. Novelty compares outputs with a reference collection. Goal achievement measures performance against a chosen objective. None of these questions alone determines whether a molecule has credible pharmacology, acceptable off-target behavior, a viable route, or a developable product profile.

The limitation becomes more consequential when a model is optimized against a scoring function. A generator does not understand the pharmaceutical intention behind a score; it searches for structures that obtain favorable numerical evaluations under the implemented objective and constraints. Goal-directed optimization can therefore exploit weaknesses in scoring functions and produce apparently strong but chemically or pharmaceutically misleading structures [7]. This behavior should not be interpreted simply as a defective model. It reveals an epistemic mismatch between the measurable proxy and the scientific outcome the proxy is expected to represent. If an activity predictor can be increased through features outside its reliable domain, if a structural alert is absent because the alert library is incomplete, or if a composite reward permits a severe liability to be offset by improvements elsewhere, optimization may amplify rather than correct the weakness.

Distribution-learning evaluations provide additional but still bounded information. Measures of similarity to training distributions, internal diversity, scaffold composition, or physicochemical coverage can identify mode collapse, memorization, or gross departures from a reference chemical space. Benchmark platforms make such comparisons more reproducible, but generative quality metrics characterize model output distributions rather than establish the pharmaceutical usefulness of individual molecules [8]. A distribution can resemble known drug-like molecules while containing no structure appropriate for a particular target or product concept. Alternatively, a useful proposal may occupy a sparsely represented region because the intended mechanism requires unusual chemistry. Distributional fidelity and pharmaceutical relevance must therefore remain analytically distinct: the first concerns model behavior relative to data, whereas the second concerns whether a structure can support a defensible program-specific hypothesis.

The strongest reason not to equate novelty with design success is that generated structures can score favorably while remaining difficult to synthesize [9]. Novelty may be achieved by adding complexity, unusual ring systems, rare functional-group combinations, or substitution patterns with limited precedent, but these same features can increase route length, stereochemical

burden, purification difficulty, reagent dependence, or uncertainty about reaction compatibility. A structure that cannot be connected to an acceptable route and obtainable inputs is not automatically worthless, because it may still reveal a useful conceptual direction. It should, however, remain classified as a molecular hypothesis rather than be promoted as a candidate. Within the proposed hierarchy, novelty is consequently treated as a descriptive or exploratory attribute. It may support chemical-space expansion, intellectual-property differentiation, or hypothesis diversification, but it cannot substitute for pharmacological relevance, safety, synthesis, or developability. **Table 1** organizes the evidence, constructs, and interpretation boundaries needed to develop why pharmaceutical design cannot be reduced to molecular novelty within the article's central argument.

Table 1. Evidence domains, core questions, scientific requirements, and interpretation boundaries for Why pharmaceutical design cannot be reduced to molecular novelty

Construct or evidence domain	Core question	Scientific basis	Role in the article	Failure or overclaiming risk	Boundary statement
Structural validity	Is the molecular representation chemically interpretable and sufficiently specified for evaluation?	Valence rules, charge state, stereochemistry, tautomerism, protonation state, and representation standardization	Establishes the minimum entry condition for downstream analysis	Treating formal validity as evidence of activity or drug-likeness	A valid representation is necessary for evaluation but does not establish pharmaceutical value.
Uniqueness	Does the generated set contain redundant structures?	Within-set identity and similarity analysis	Assesses output redundancy and effective exploration	Presenting non-duplication as meaningful chemical innovation	Uniqueness concerns repetition within an output set, not novelty relative to prior knowledge or usefulness.
Molecular novelty	How different is the proposal from a defined reference set?	Representation-dependent similarity, scaffold comparison, and reference-database coverage	Characterizes chemical-space distance and exploratory potential	Equating greater distance with greater therapeutic value	Novelty is conditional on the comparator and cannot substitute for pharmacological or experimental evidence.
Distributional quality	Does the generated set reproduce or intentionally depart from relevant chemical distributions?	Diversity, scaffold, physicochemical, and distribution-comparison measures	Diagnoses model behavior, memorization, collapse, or uncontrolled divergence	Treating distributional similarity as proof that individual outputs are appropriate drugs	Distribution-level agreement does not establish target-specific or candidate-level credibility.
Goal-directed score	Does the structure perform well under the implemented computational objective?	Property models, similarity functions, docking functions, or composite rewards	Supports prioritization within the scope of the defined proxy	Reward exploitation, out-of-domain extrapolation, and compensation for critical liabilities	A high score indicates success against an implemented objective, not confirmation of the intended pharmaceutical outcome.
Synthetic plausibility	Can the structure be connected to defensible transformations and accessible precursors?	Retrosynthetic reasoning, reaction precedent, chemist assessment, and material context	Prevents computational desirability from obscuring route infeasibility	Treating a molecule-level score as proof of an executable route	Synthetic plausibility requires route-level assessment and remains uncertain until laboratory execution.
Pharmaceutical admissibility	Does the proposal satisfy minimum potency, selectivity, safety, exposure, developability, and feasibility conditions?	Integrated, context-specific evidence across pharmaceutical disciplines	Defines the boundary between an interesting generated structure and a testable pharmaceutical hypothesis	Averaging incompatible properties into a score that hides unacceptable failure	Pharmaceutical admissibility requires multiple non-substitutable conditions; no single novelty or performance measure is sufficient.
Evidence maturity	What claim is justified by the evidence currently available?	Progression from computation through synthesis, analytical verification, and biological and developability testing	Aligns terminology and decision authority with accumulated evidence	Calling an untested output a hit, lead, or candidate	Claim strength must not exceed the stage and independence of supporting evidence.

Potency, Selectivity, Safety, and Developability as Coupled Design Constraints

Potency, selectivity, safety, and developability are coupled because structural changes rarely affect only one pharmaceutical property. Increased molecular complexity may create additional target interactions and improve binding or selectivity, but it can also increase synthetic burden, lipophilicity, conformational uncertainty, metabolic liabilities, and difficulty in controlling solid-state or formulation behavior [10]. The coupling is not uniformly negative or positive. Added three-dimensionality may improve differentiation between related binding sites, while greater rigidity may reduce an entropic penalty yet impair solubility. A polar substituent may support aqueous behavior but reduce passive permeability. The design problem is therefore not to minimize or maximize complexity, polarity, lipophilicity, or flexibility in isolation. It is to determine whether a particular structural configuration satisfies the intended pharmacological and product context without creating unacceptable liabilities elsewhere.

Potency itself should be treated as a context-specific relevance floor rather than a universal scalar objective. A predicted affinity or activity value is meaningful only relative to assay design, target state, mechanism, exposure, and the uncertainty of the supporting model. Selectivity similarly cannot be represented adequately by the intended target score alone; it requires consideration of plausible off-targets, homologous proteins, pathway-level effects, and concentration-dependent margins. Safety adds endpoint-specific constraints that are often non-compensable: a serious reactive, genotoxic, cardiovascular, neurological, or organ-specific liability should not be averaged away by a gain in predicted potency. The multi-causal nature of drug-development failure, involving interacting efficacy, safety, pharmacokinetic, and strategic factors, illustrates why success on one property cannot be interpreted as general candidate quality [11]. The hierarchy proposed here therefore distinguishes properties that may be traded within defined limits from failures that should trigger rejection, redesign, or an explicit hold.

Developability connects molecular properties to the conditions under which a compound could become a reproducible pharmaceutical product. It includes, but is not limited to, solubility, permeability, chemical and metabolic stability, exposure potential, solid-state behavior, formulation compatibility, and the relationship between dose and achievable concentration. These features cannot be interpreted through universal thresholds because their importance changes with indication, route of administration, therapeutic window, dosing frequency, tissue distribution, and modality. Analyses of small-molecule property trends from a developability perspective support evaluating compounds as integrated profiles rather than as collections of independent favorable values [12]. In the proposed hierarchy, the developability envelope is consequently defined by the intended product concept. It asks whether the combined property pattern is compatible with a plausible path to exposure, manufacture, storage, administration, and use—not whether the molecule resembles an abstract average drug.

Early computational profiling can support this assessment by combining predicted absorption, distribution, metabolism, excretion, drug-likeness, physicochemical characteristics, and medicinal-chemistry alerts [13]. Such tools are useful for triage, comparison, hypothesis generation, and the identification of properties that require experimental clarification. They must not be treated as confirmation of pharmacokinetics, safety, formulation performance, or medicinal-chemistry acceptability. A negative alert result may reflect limited rule coverage, and an apparently favorable property estimate may arise from model extrapolation or representation choices. The original constraint-aware contribution of this article is therefore to position these predictions as provisional evidence attached to separate design constraints. Potency, selectivity, safety, and developability should be assessed jointly, but not flattened indiscriminately: each retains its own evidentiary basis, uncertainty, failure conditions, and decision authority. This coupling also creates a direct dependency on synthesis, because a computational trade-off cannot be adjudicated experimentally unless the relevant structures can be produced through accessible routes and materials.

Synthetic Feasibility, Route Accessibility, and Material Availability

Synthetic feasibility is not an intrinsic binary property of a molecular graph. It is a contextual judgment about whether a specified structure can plausibly be produced through chemically defensible transformations under defined constraints. A molecule-level assessment may identify complexity, unusual substructures, or known problematic motifs, but it cannot establish whether a complete route exists. Retrosynthetic planning instead treats synthesis as a route-level search problem in which the target is recursively disconnected into precursors using learned reaction knowledge and symbolic search [14]. This shift is essential to constraint-aware design because two structurally similar molecules may differ substantially in route length, reaction precedent, stereochemical control, protecting-group requirements, purification burden, or dependence on specialized chemistry. Synthetic feasibility should therefore be represented as a structured body of route evidence rather than a single molecular descriptor. The relevant question is not merely whether a structure appears synthesizable, but whether at least one route is sufficiently plausible, interpretable, and compatible with the intended experimental setting to justify further attention.

Route accessibility narrows this inquiry by making the assumptions of the planning environment explicit. Retrosynthetic systems can test whether candidate routes are discoverable under specified reaction-template libraries, search strategies, stopping criteria, and definitions of purchasable stock [15]. A route that is found under one configuration may disappear when the available reaction knowledge, precursor inventory, maximum search depth, or permitted transformations change. Conversely, the absence of a route may reflect incomplete templates or search limitations rather than chemical impossibility. Computational route planning should consequently report the assumptions that generated the route, plausible alternatives, unresolved transformations, precursor provenance, and the degree of precedent supporting critical steps. It should not treat search completion as equivalent to laboratory execution. Route accessibility is best interpreted as evidence that a proposed structure can be connected computationally to a defined starting-material space, subject to the limitations of the planner and its knowledge base.

Rapid synthesizability classifiers can support early triage when full retrosynthetic analysis is computationally expensive. Retrosynthesis-derived accessibility scoring provides an approximate indication of whether an underlying planner is likely to identify a route, enabling implausible proposals to be deprioritized before detailed examination [16]. This function is useful within large generated libraries, but its interpretation must remain bounded. The score inherits the reaction coverage, stock assumptions, molecular representation, and success definition of the planner used to create its labels. It may also obscure why a structure is classified as accessible or inaccessible and may fail to distinguish a short robust route from a nominal route containing an uncertain transformation. Within the proposed hierarchy, a fast accessibility estimate is therefore an early screening signal rather than a feasibility determination. Structures near a decision boundary, structures containing unusual

chemistry, and structures selected for experimental progression require route reconstruction and expert review rather than reliance on the classifier alone.

Material availability is a separate constraint because even a chemically plausible route may depend on precursors, reagents, catalysts, or intermediates that are unavailable, unstable, restricted, prohibitively difficult to obtain, or incompatible with the intended scale and location. Analyses of commercially available building blocks demonstrate that medicinal-chemistry access is shaped by the composition and diversity of supplier inventories [17]. Availability is also time-sensitive and organization-dependent: a listed compound may be back-ordered, supplied only in unsuitable quantities, inadequately characterized, or obtainable only through a route that transfers the original synthesis problem upstream. Constraint-aware design should therefore distinguish catalog presence from verified procurement and distinguish purchasable inputs from materials that must themselves be synthesized. Material constraints may legitimately favor a less novel molecule when it permits rapid experimental learning through accessible analogues. **Table 2** organizes the evidence, constructs, and interpretation boundaries needed to develop synthetic feasibility, route accessibility, and material availability within the article's central argument.

Table 2. Components, relationships, uncertainties, and validation needs within Synthetic feasibility, route accessibility, and material availability

Component or layer	Inputs	Core function	Expected output	Uncertainty or limitation	Validation requirement
Molecular specification	Standardized structure, stereochemistry, charge state, tautomeric form, protecting-group assumptions	Defines the exact synthetic target	Unambiguous target representation	An underspecified structure can produce misleading routes or omit stereochemical burdens	Chemist confirmation that the encoded target matches the intended compound
Rapid accessibility screen	Molecular graph and a model trained on retrosynthetic-planning outcomes	Removes clearly low-priority structures before expensive planning	Approximate accessibility classification or ranking	Inherits planner, reaction-data, stock, and representation limitations	Comparison with full route planning and expert review for selected structures
Retrosynthetic route search	Target structure, reaction templates or learned transformations, search settings, stopping rules	Connects the target to progressively simpler precursors	One or more candidate retrosynthetic trees	Search failure does not prove impossibility; route discovery does not prove laboratory success	Independent route inspection and assessment of critical transformations
Reaction-precedent assessment	Proposed steps, analogous reactions, substrate context, condition information	Evaluates whether each transformation is chemically defensible	Step-level confidence and identification of unsupported disconnections	Literature or training examples may not transfer to the proposed substrate	Experimental testing of uncertain or precedent-poor transformations
Route accessibility	Candidate routes, organizational capabilities, allowed chemistry, equipment, time, scale	Determines whether a route can be attempted in the intended setting	Context-specific route status: accessible, revisable, uncertain, or unsuitable	Accessibility varies across laboratories, platforms, and development stages	Review by synthetic and process chemists under the intended context of use
Building-block availability	Supplier inventories, internal stocks, precursor specifications, geographic and temporal access	Tests whether route inputs can be obtained in suitable form and quantity	Verified, substituted, or unresolved material list	Catalog presence may not equal actual availability, quality, or continuity	Procurement verification and incoming-material characterization
Route robustness	Alternative disconnections, step count, convergence, stereochemical demands, purification and isolation considerations	Assesses resilience to step failure and practical burden	Preferred route with alternatives and identified vulnerabilities	Computational planning often omits yield, impurity, isolation, and scale effects	Laboratory execution, analytical confirmation, and route revision
Material continuity	Expected demand, supplier dependence, custom-synthesis needs, hazardous or restricted materials	Evaluates whether access can be sustained beyond an initial experiment	Supply-risk statement linked to anticipated use	Future demand and supply conditions may be unknown during early design	Reassessment as scale, formulation, and development requirements become clearer

Proposed Constraint Hierarchy for Computational Molecular Design

The proposed Full-Weight Pharmaceutical Design Hierarchy begins from the premise that molecular generation and property-directed search are enabling substrates, not complete pharmaceutical decision systems. Learned continuous representations can organize chemical structures so that movement through a latent space corresponds to changes in predicted molecular properties [18]. Such representations make optimization computationally tractable, but they do not determine which properties should function as exclusions, minimum feasibility conditions, negotiable objectives, or preferences. The hierarchy therefore separates the capacity to search from the authority to advance. At its base lies structural admissibility: the molecular identity must be unambiguous, chemically interpretable, and suitable for the models being applied. Above this lie non-compensable exclusions, including serious safety or reactivity concerns that cannot legitimately be cancelled by improvements in potency or novelty. Structures failing these foundational conditions should be rejected, held for clarification, or redesigned before preference-based optimization begins.

The second level comprises feasibility floors that determine whether a molecular hypothesis can enter an experimentally meaningful design space. Conditional molecular generation demonstrates that output distributions can be steered by multiple requested descriptors [19]. The article extends this logic by arguing that not every condition should be represented as an equally weighted target. Synthetic feasibility, route accessibility, material availability, pharmacological relevance, selectivity, and

product-specific developability should instead be encoded as heterogeneous constraints with different evidentiary meanings. Some may be hard exclusions, some may define minimum acceptable regions, and others may remain provisional until testing. A structure should not receive an overall admissible status merely because a composite objective is favorable. Its status should retain the result for each constraint, the evidence source supporting that result, the uncertainty attached to it, and the consequence of failure.

The third level addresses conditional optimization among structures that have survived the more fundamental constraints. Multiobjective molecular design can benefit from explicit quality assessment that evaluates generated structures beyond raw reward accumulation [20]. In the proposed hierarchy, potency improvement, selectivity margins, physicochemical balance, route efficiency, structural simplicity, novelty, intellectual-property distance, and portfolio diversity may be optimized only within the region that remains pharmaceutically admissible under the available evidence. The hierarchy does not require all objectives to improve simultaneously. Instead, it distinguishes unacceptable deterioration from transparent, context-dependent compromise. It also prohibits unlimited compensation: severe safety concern, lack of an executable route, or incompatibility with the intended product concept cannot be neutralized merely by a sufficiently high score elsewhere. **Figure 1** presents the developability-constrained molecular design crucible, showing how the article's key components and boundaries are connected within proposed constraint hierarchy for computational molecular design.

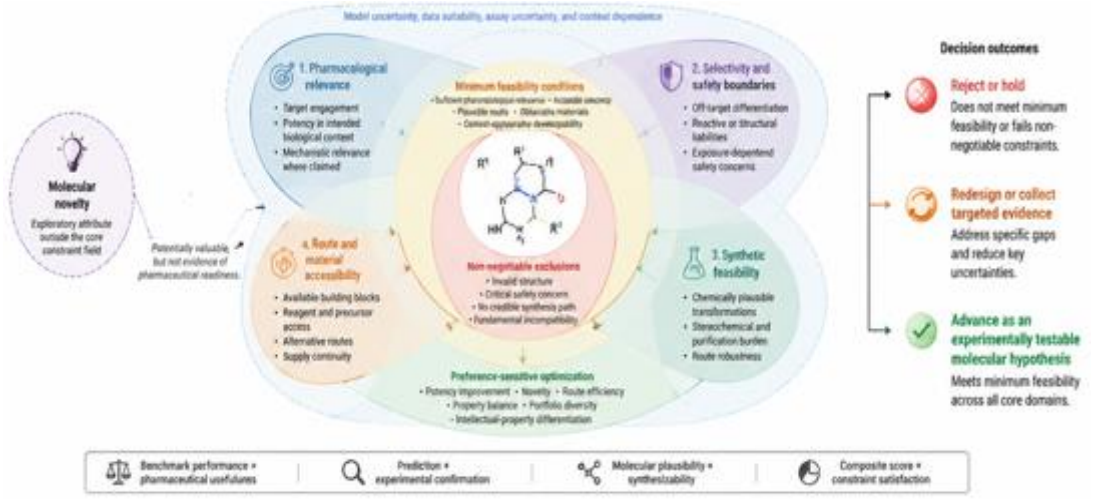

Figure 1. Developability-Constrained Molecular Design Crucible.

The figure is an original conceptual synthesis that organizes the article's central contribution across pharmaceutical design cannot be reduced to molecular novelty, potency, selectivity, safety, and developability as coupled design constraints, developability as coupled design constraints, synthetic feasibility, route accessibility, and material availability, synthetic feasibility, route accessibility. Arrows and grouping indicate proposed or evidence-supported relationships rather than measured effect sizes. The figure does not represent a validated predictive model, regulatory determination, clinical recommendation, or deployment-ready system.

Human judgment forms the final interpretive layer rather than an external correction applied only after generation. Human-in-the-loop molecular design shows that expert preferences and implicit medicinal-chemistry knowledge cannot always be represented adequately by a fixed objective function [21]. Experts are needed to determine whether a predicted liability is credible, whether an apparent trade-off is tolerable, whether an unusual structure is informative despite uncertainty, and whether the next experiment can discriminate among competing hypotheses. Human involvement does not eliminate bias or guarantee correctness; it must itself be documented through decision rationales, disciplinary representation, and predefined escalation rules. The hierarchy is therefore proposed as a shared decision language. It organizes computational evidence and makes non-substitution boundaries visible, but it does not automate candidate nomination or replace accountable multidisciplinary review.

Managing Conflict, Uncertainty, and Trade-Offs Across Objectives

Uncertainty is a transversal property of the hierarchy because every constraint may be estimated from incomplete, heterogeneous, or context-mismatched evidence. Bayesian and neural uncertainty methods can differ substantially in calibration, ranking behavior, computational cost, and robustness across molecular-property tasks; predictive confidence should therefore not be treated as uncertainty without method- and dataset-specific evaluation [22, 23]. A narrow prediction interval may reflect a model's internal behavior while failing to capture assay noise, label inconsistency, target-state ambiguity, distribution shift, or missing mechanisms. Constraint-aware design should preserve at least three distinctions: uncertainty about the model, uncertainty inherent in the measured or predicted endpoint, and uncertainty about whether the endpoint is an adequate proxy for the pharmaceutical decision. These forms of uncertainty have different remedies. Additional training data

may reduce some model uncertainty, whereas a poorly defined assay or an inadequate endpoint requires scientific redesign rather than merely more observations.

Objective conflict arises when improving one attribute worsens another or when the objectives cannot be evaluated on commensurate scales. Molecular-design studies that combine biological activity with blood–brain barrier-related properties illustrate the need to balance activity, permeability, solubility, and associated molecular characteristics rather than optimize a single endpoint [24]. Yet "balance" should not imply that all objectives may compensate freely. The appropriate response depends on the therapeutic context. Reduced permeability may be unacceptable for a central nervous system indication but advantageous when peripheral restriction is desired. Increased potency may permit lower exposure, while increased lipophilicity may simultaneously improve membrane passage and worsen solubility or nonspecific binding. The hierarchy therefore treats trade-offs as conditional propositions: each proposed compromise must state the intended context, direction of change, uncertainty, minimum acceptable floor, and evidence required to determine whether the compromise is real.

Composite scores can facilitate ranking, but they can also conceal the contribution of individual properties and create false precision. Explainable multiparameter optimization supports exposing how property estimates contribute to a design decision rather than returning only an aggregate value [25]. In the proposed approach, every prioritized molecule should be accompanied by a constraint profile showing which conditions are satisfied, which are uncertain, which are close to a boundary, and which are deliberately traded. Weighting functions should be treated as expressions of current priorities rather than natural laws. Sensitivity analysis should test whether rankings change materially under plausible alternative weights, model choices, or uncertainty assumptions. A molecule whose rank is stable only under a narrow weighting scheme should be recognized as preference-sensitive rather than intrinsically superior.

Conflict management also requires portfolio-level reasoning. A single "best" structure may be less scientifically valuable than a small set that spans distinct hypotheses, route options, property trade-offs, and uncertainty profiles. Diversity should therefore be used strategically, not celebrated automatically. One proposal might maximize confidence within an established series, another might address a known selectivity liability, and a third might test whether a novel scaffold can escape a developability limitation. Human review remains necessary when evidence is sparse or objectives are difficult to formalize [21]. The preferred output of the hierarchy is consequently not an unquestioned scalar ranking but a bounded decision set containing admissible structures, rejected structures with explicit reasons, uncertain structures requiring targeted evidence, and alternative structures preserved because they support different learning objectives.

Decision Gates from Generated Structure to Experimentally Credible Candidate

A generated structure acquires stronger pharmaceutical status only through accumulated evidence. Prospective molecular-design studies demonstrate that computational proposals become scientifically credible through synthesis and layered experimental evaluation, including biochemical, cellular, biophysical, selectivity, or structural confirmation appropriate to the program [26–28]. These studies do not establish that any particular generative architecture will generalize across targets, nor do they collapse discovery into an automated sequence. They instead clarify an essential evidence boundary: computation can prioritize and rationalize a proposal, but the existence, identity, activity, and broader behavior of the compound must be established through observation. The proposed decision-gate model therefore begins with restrained terminology. Before synthesis, the object is a generated molecular hypothesis. After a plausible route and materials review, it may become a route-ready proposal. Only after successful production and analytical confirmation should it be called a synthesized compound.

The route-to-experiment gate must evaluate whether the computational synthesis proposal can be executed under actual reaction, equipment, monitoring, and purification constraints. AI-informed robotic flow synthesis shows that planning can be connected to physical execution, but successful synthesis still depends on supported reaction classes, platform capabilities, compatible conditions, and analytical feedback [29]. Automation does not remove chemistry from the process; it makes operational assumptions more consequential. For general molecular-design programs, this gate should require confirmation of target identity, relevant stereochemistry, purity, stability during handling, and sufficient material for the planned assays. A failed synthesis should not automatically invalidate the molecular hypothesis, because the proposed route may be at fault. The decision must distinguish structure failure from route failure and determine whether an alternative disconnection, analogue, or redesigned target can preserve the underlying pharmacological question.

The pharmacology gate should test the claim that justified synthesis rather than merely collect any favorable assay result. Potency must be measured in a relevant system, and selectivity should be assessed against biologically plausible comparators. Safety and liability testing should be proportionate to the anticipated mechanism, structure, exposure, and intended use. Developability evidence should begin early enough to prevent repeated optimization of molecules that cannot support the product concept. Computational predictions may guide assay selection, but they cannot replace identity-confirmed experimental measurements. Disagreement between prediction and experiment should be treated as diagnostic evidence about model applicability, data suitability, assay context, or the molecular hypothesis itself. **Table 3** organizes the evidence, constructs, and interpretation boundaries needed to develop decision gates from generated structure to experimentally credible candidate within the article's central argument.

The integrated review gate determines whether the accumulated evidence justifies further study, redesign, or discontinuation. Passing a gate does not erase uncertainty or establish clinical, regulatory, or commercial viability. It permits only the bounded claim associated with that gate. A compound may be experimentally active yet lack adequate selectivity; it may show a favorable pharmacological profile while remaining unsuitable for the intended formulation; or it may be developable in

principle but dependent on an impractical route. Gate decisions should therefore remain reversible and should record the evidence considered, unresolved risks, human reviewers, and next discriminating experiment. **Figure 2** presents the evidence, validation, and translation boundary observatory for designing molecules under the full weight of, showing how the article's key components and boundaries are connected within decision gates from generated structure to experimentally credible candidate.

Table 3. Evaluation, implementation, and research priorities arising from Designing Molecules under the Full Weight of Potency, Selectivity, Synthesizability, Safety, and Pharmaceutical Developability

Evaluation dimension	What must be tested	Suitable evidence or assessment	Failure signal	Interpretive limitation	Decision relevance
Structural and identity integrity	Whether the proposed molecular identity is fully specified and the synthesized material corresponds to it	Standardized structure review; analytical identity and purity evidence after synthesis	Ambiguous stereochemistry, inconsistent representation, incorrect identity, or unresolved impurities	Computational validity does not establish that the physical compound exists or has been correctly identified	Determines whether downstream predictions and assays refer to the intended object
Constraint admissibility	Whether non-compensable exclusions and minimum feasibility floors are satisfied	Constraint-by-constraint profile with uncertainty and context-of-use statements	Serious liability, absent route, unsuitable product profile, or unresolved critical constraint	Thresholds are context-dependent and may change as evidence matures	Determines rejection, redesign, hold, or entry into preference-based optimization
Route and material credibility	Whether a defensible route and obtainable inputs exist in the intended setting	Retrosynthetic plan, precedent review, alternative routes, verified building-block and reagent access	Unsupported key step, unavailable precursor, excessive dependence on uncertain chemistry, or no viable alternative	A planned route is not evidence of successful execution, yield, purity, or scale	Determines whether synthesis should be attempted and which route carries acceptable learning value
Synthetic execution	Whether the compound can be produced, isolated, and analytically confirmed	Laboratory synthesis, reaction monitoring, purification, identity and stability assessment	Repeated route failure, decomposition, unmanageable mixtures, or inability to obtain assay-quality material	Failure may reflect the route rather than the molecule; initial success may not transfer to scale	Converts a computational proposal into a real compound and identifies route-specific risks
Pharmacological relevance	Whether the compound affects the intended target or biological system under appropriate conditions	Biochemical, biophysical, cellular, mechanistic, or structural assays matched to the claim	Lack of reproducible activity, assay interference, inconsistent mechanism, or context-dependent loss of effect	Activity does not alone establish selectivity, in vivo exposure, mechanism, or therapeutic utility	Determines whether the molecular hypothesis receives experimental support
Selectivity and safety boundaries	Whether activity is sufficiently differentiated and whether critical liabilities emerge	Counter-screening, liability assays, orthogonal testing, reactive or structural alert follow-up	Narrow or absent selectivity margin, concentration-dependent nonspecific effects, or serious liability	Early panels are incomplete and negative results do not prove safety	Determines whether potency can be interpreted as potentially useful rather than merely observable
Developability compatibility	Whether the integrated property profile is compatible with the intended product concept	Solubility, permeability, stability, exposure-relevant, solid-state, formulation, and material assessments as appropriate	Inability to achieve required exposure, unstable material, unsuitable solid form, or incompatible formulation behavior	Early developability evidence is incomplete and highly dependent on route, indication, dose, and formulation	Determines whether continued optimization is scientifically and practically justified
Uncertainty quality	Whether reported uncertainty is calibrated, informative, and responsive to distribution shift	Calibration analysis, applicability-domain assessment, error stratification, prospective prediction tracking	High confidence on repeated errors, unstable rankings, or poor identification of out-of-domain structures	No uncertainty method captures every scientific and operational source of uncertainty	Determines whether predictions can guide action or should trigger additional evidence collection
Trade-off transparency	Whether prioritization remains defensible under alternative preferences and model assumptions	Property-contribution display, weight sensitivity, Pareto analysis, and multidisciplinary review	Ranking reverses under minor changes or a critical failure is hidden by an aggregate score	Transparency does not determine which trade-off is ethically, clinically, or strategically acceptable	Supports accountable selection among admissible alternatives
Integrated evidence readiness	Whether the accumulated evidence supports the precise claim used to advance the compound	Multidisciplinary dossier linking predictions, routes, experiments, limitations, and unresolved risks	Claim language exceeds the last passed evidence gate or important dissent is undocumented	Gate passage does not establish clinical utility, regulatory acceptability, or deployment readiness	Determines the permissible status: hypothesis, route-ready proposal, synthesized compound, active compound, or bounded candidate for further study

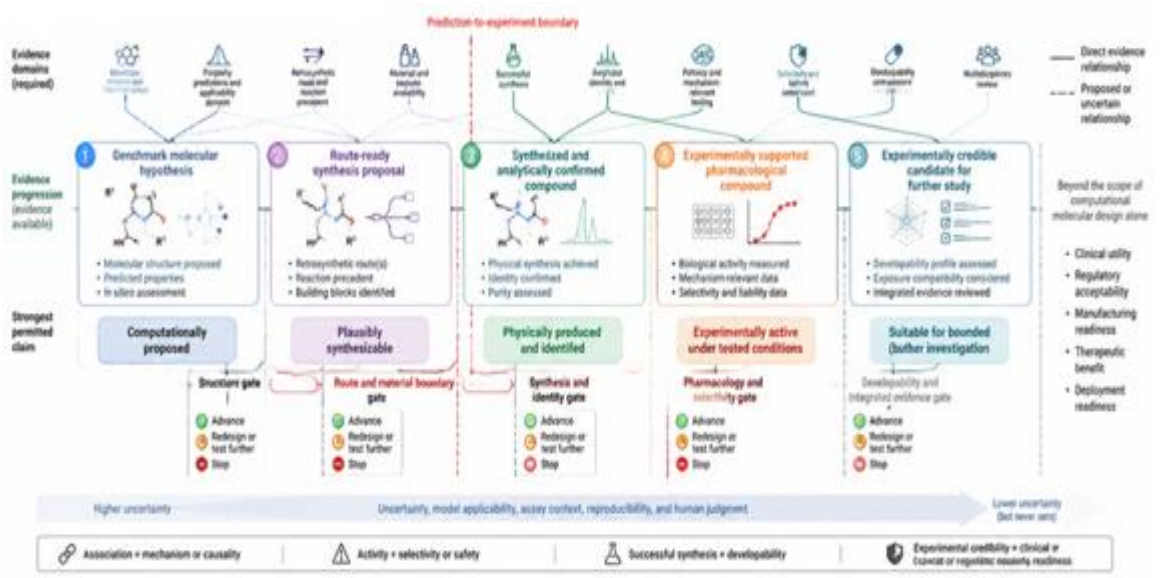

Figure 2. Evidence, Validation, and Translation Boundaries.

The figure is an original conceptual synthesis that shows how claims move from conceptual plausibility through evaluation, uncertainty assessment, and bounded pharmaceutical use. Arrows and grouping indicate proposed or evidence-supported relationships rather than measured effect sizes. The figure does not represent a validated predictive model, regulatory determination, clinical recommendation, or deployment-ready system.

Limitations and Boundary Conditions

The Full-Weight Pharmaceutical Design Hierarchy is an original conceptual contribution and has not been prospectively validated as a selection algorithm, scoring system, or operational workflow. Its levels describe distinctions that should be preserved during molecular design, but they do not provide universal thresholds for potency, selectivity, safety, synthetic accessibility, or developability. Appropriate boundaries depend on the biological target, disease, modality, therapeutic window, route of administration, dose, product concept, and organizational capability. A constraint that functions as a hard exclusion in one program may be a manageable optimization problem in another. The hierarchy should therefore be instantiated for a defined context of use and evaluated against existing multidisciplinary practice rather than applied as a generic template.

The supporting evidence also spans heterogeneous scientific purposes. Benchmark studies characterize model behavior, route-planning studies assess computational access to reactions and stocks, uncertainty studies evaluate predictive methods, and prospective studies demonstrate specific examples of synthesis and testing. These evidence types cannot be treated as interchangeable. Commercial material availability may change over time or differ by region [17]. Predictive uncertainty may fail under distribution shift or may not capture uncertainty in the biological question [22]. Prospective experimental success in one target context does not establish broad generalizability to unrelated targets, chemotypes, or product requirements [26]. The article therefore supports a disciplined structure for interpretation, not a claim that the selected literature collectively validates the hierarchy.

Misuse is possible if the hierarchy is converted into an opaque aggregate score, implemented without pharmaceutical expertise, or presented as authorization for automatic candidate advancement. The design may also underrepresent constraints relevant to particular modalities, including covalent mechanisms, macrocycles, peptides, oligonucleotides, radiopharmaceuticals, or complex drug-delivery systems. Ethical, economic, intellectual-property, environmental, and patient-level considerations may influence real development decisions but are not reducible to molecular properties. Human oversight is itself fallible and may reproduce institutional preferences or discourage unconventional chemistry. Accordingly, the hierarchy cannot establish mechanism, causality, clinical benefit, regulatory acceptability, manufacturing readiness, or therapeutic value. Those claims require evidence beyond the scope of computational molecular design.

Research Agenda and Implementation Priorities

The first research priority is prospective comparative evaluation. Constraint-aware generation should be compared with flat composite optimization and conventional medicinal-chemistry prioritization using matched program questions, predefined decision criteria, and blinded or independently reviewed outputs. Evaluation should test more than the number of high-scoring molecules produced. It should examine whether the hierarchy improves rejection of invalid or infeasible proposals, preserves informative alternatives, identifies uncertainty, supports route selection, and produces experiments that discriminate among hypotheses. Multiobjective methods with explicit quality assessment provide a computational starting point, but they do not establish the practical superiority of a hierarchy without prospective comparison [20]. Studies should report negative outcomes, route failures, model disagreements, and redesign decisions rather than only successful compounds.

The second priority is evidence and reporting infrastructure. Molecular records should link every prediction to model version, training-domain information, uncertainty method, endpoint definition, molecular representation, and context of use. Route records should preserve reaction assumptions, stock definitions, alternative disconnections, material provenance, and expert changes. Human-in-the-loop decisions should document which preferences were elicited, who supplied them, and how they changed prioritization [21]. Uncertainty evaluation should be repeated prospectively because retrospective calibration does not guarantee robustness on newly generated chemistry [23]. Shared reporting standards should distinguish generated structure, route-ready proposal, synthesized compound, experimentally active compound, and candidate for further study so that claim language remains aligned with evidence maturity.

The third priority is staged implementation. Initial use should be decision-supportive and limited to contexts where multidisciplinary teams can inspect each constraint and reverse automated recommendations. Organizations should begin with retrospective shadow evaluation, proceed to prospective non-decisional use, and only then test bounded influence on molecule selection. Integration with synthesis platforms may accelerate learning when route classes and platform constraints are explicit, but automation should not be interpreted as unrestricted chemical executability [29]. Governance should specify stop conditions, escalation pathways, model-update procedures, accountability for adverse recommendations, and criteria for withdrawing a tool when calibration or applicability deteriorates. The practical objective is not to automate judgment completely, but to make evidence, uncertainty, conflicts, and non-substitution boundaries sufficiently visible for more disciplined pharmaceutical decisions.

Conclusion

Designing molecules under the full weight of potency, selectivity, synthesizability, safety, and pharmaceutical developability requires a conceptual shift from producing high-scoring structures to constructing experimentally testable pharmaceutical hypotheses. The proposed Full-Weight Pharmaceutical Design Hierarchy organizes that shift through structural admissibility, non-compensable exclusions, route and material feasibility, pharmacological relevance, selectivity, developability, conditional optimization, explicit uncertainty, and reversible evidence gates. Its principal contribution is to prevent novelty, potency, synthetic-accessibility proxies, or aggregate rewards from standing in for the distinct evidence needed to advance a molecule. It also clarifies that trade-offs are legitimate only within context-specific boundaries and that claim strength must remain limited by the last completed evidentiary gate. The hierarchy is not a validated predictive system, universal development standard, regulatory framework, or automatic candidate-selection mechanism. Its scientific value will depend on prospective testing, transparent reporting, calibrated uncertainty, multidisciplinary oversight, and adaptation to specific targets, modalities, product concepts, and experimental capabilities. Used within these boundaries, constraint-aware design may support more credible prioritization and more informative experiments without confusing computational promise with pharmaceutical readiness.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

References

1. Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov. 2019;18(6):463-77. doi:10.1038/s41573-019-0024-5
2. Schneider P, Walters WP, Plowright AT, Sieroka N, Listgarten J, Goodnow RA Jr, et al. Rethinking drug design in the artificial intelligence era. Nat Rev Drug Discov. 2020;19(5):353-64. doi:10.1038/s41573-019-0050-3
3. Chan HCS, Shan H, Dahoun T, Vogel H, Yuan S. Advancing drug discovery via artificial intelligence. Trends Pharmacol Sci. 2019;40(8):592-604. doi:10.1016/j.tips.2019.06.004
4. Sadybekov AV, Katritch V. Computational approaches streamlining drug discovery. Nature. 2023;616(7958):673-85. doi:10.1038/s41586-023-05905-z
5. Walters WP, Murcko MA. Assessing the impact of generative AI on medicinal chemistry. Nat Biotechnol. 2020;38(2):143-5. doi:10.1038/s41587-020-0418-2
6. 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
7. Langevin M, Vuilleumier R, Bianciotto M. Explaining and avoiding failure modes in goal-directed generation of small molecules. J Cheminform. 2022;14(1):20. doi:10.1186/s13321-022-00601-y

8. Polykovskiy D, Zhebrak A, Sanchez-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
9. 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
10. Méndez-Lucio O, Medina-Franco JL. The many roles of molecular complexity in drug discovery. Drug Discov Today. 2017;22(1):120-6. doi:10.1016/j.drudis.2016.08.009
11. Sun D, Gao W, Hu H, Zhou S. Why 90% of clinical drug development fails and how to improve it? Acta Pharm Sin B. 2022;12(7):3049-62. doi:10.1016/j.apsb.2022.02.002
12. Agarwal P, Huckle J, Newman J, Reid DL. Trends in small molecule drug properties: A developability molecule assessment perspective. Drug Discov Today. 2022;27(12):103366. doi:10.1016/j.drudis.2022.103366
13. Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017;7:42717. doi:10.1038/srep42717
14. Segler MHS, Preuss M, Waller MP. Planning chemical syntheses with deep neural networks and symbolic AI. Nature. 2018;555(7698):604-10. doi:10.1038/nature25978
15. Genheden S, Thakkar A, Chadimová V, Reymond JL, Engkvist O, Bjerrum EJ. AiZynthFinder: A fast, robust and flexible open-source software for retrosynthetic planning. J Cheminform. 2020;12(1):70. doi:10.1186/s13321-020-00472-1
16. Thakkar A, Chadimová V, Bjerrum EJ, Engkvist O, Reymond JL. Retrosynthetic accessibility score (RAscore): Rapid machine learned synthesizability classification from AI driven retrosynthetic planning. Chem Sci. 2021;12(9):3339-49. doi:10.1039/D0SC05401A
17. Zabolotna Y, Volochnyuk DM, Ryabukhin SV, Horvath D, Gavrilenko KS, Marcou G, et al. A close-up look at the chemical space of commercially available building blocks for medicinal chemistry. J Chem Inf Model. 2022;62(9):2171-85. doi:10.1021/acs.jcim.1c00811
18. 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
19. Kotsias PC, Arús-Pous J, Chen H, Engkvist O, Tyrchan C. Direct steering of de novo molecular generation with descriptor conditional recurrent neural networks. Nat Mach Intell. 2020;2(5):254-65. doi:10.1038/s42256-020-0174-5
20. Fang Y, Pan X, Shen HB. De novo drug design by iterative multiobjective deep reinforcement learning with graph-based molecular quality assessment. Bioinformatics. 2023;39(4). doi:10.1093/bioinformatics/btad157
21. Sundin I, Voronov A, Xiao H, Papadopoulos K, Bjerrum EJ, Heinonen M, et al. Human-in-the-loop assisted de novo molecular design. J Cheminform. 2022;14(1):86. doi:10.1186/s13321-022-00667-8
22. Scalia G, Grambow CA, Pernici B, Li YP, Green WH. Evaluating scalable Bayesian deep learning methods for robust computer-aided molecular design. J Chem Inf Model. 2020;60(6):2697-717. doi:10.1021/acs.jcim.9b00975
23. Hirschfeld L, Swanson K, Yang K, Barzilay R, Coley CW. Uncertainty quantification using neural networks for molecular property prediction. J Chem Inf Model. 2020;60(8):3770-80. doi:10.1021/acs.jcim.0c00502
24. Pereira T, Abbasi M, Oliveira JL, Ribeiro B, Arrais J. Optimizing blood–brain barrier permeation through deep reinforcement learning for de novo drug design. Bioinformatics. 2021;37(Suppl 1). doi:10.1093/bioinformatics/btab301
25. Bung N, Krishnan SR, Roy A. An in silico explainable multiparameter optimization approach for de novo drug design against proteins from the central nervous system. J Chem Inf Model. 2022;62(11):2685-95. doi:10.1021/acs.jcim.2c00462
26. Zhavoronkov A, Ivanenkov YA, Aliper A, Veselov MS, Aladinskiy VA, Aladinskaya AV, et al. Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nat Biotechnol. 2019;37(9):1038-40. doi:10.1038/s41587-019-0224-x
27. Ren F, Ding X, Zheng M, Korzinkin M, Cai X, Zhu W, et al. AlphaFold accelerates artificial intelligence powered drug discovery: Efficient discovery of a novel CDK20 small molecule inhibitor. Chem Sci. 2023;14(6):1443-52. doi:10.1039/D2SC05709C
28. Atz K, Cotos L, Isert C, Håkansson M, Focht D, Hilleke M, et al. Prospective de novo drug design with deep interactome learning. Nat Commun. 2024;15(1):3408. doi:10.1038/s41467-024-47613-w
29. Coley CW, Thomas DA III, Lummiss JAM, Jaworski JN, Breen CP, Schultz V, et al. A robotic platform for flow synthesis of organic compounds informed by AI planning. Science. 2019;365(6453). doi:10.1126/science.aax1566