



Why Semisynthetic Accessibility Should Influence Which Natural Product Scaffolds Advance from Computational Prioritization to Medicinal Chemistry Development

Kristian Eriksen^{1*}, Solveig Berg¹, Magnus Dahl²

¹Department of Semisynthetic Accessibility and Scaffold Prioritization, Faculty of Pharmacy, University of Oslo, Oslo, Norway.

²Department of Medicinal Chemistry and Natural Product Optimization, Faculty of Pharmacy, University of Bergen, Bergen, Norway.

ABSTRACT

Computational prioritization can identify natural products with attractive predicted binding, bioactivity, selectivity, or phenotypic signatures, but it does not determine whether medicinal chemistry can interrogate those scaffolds efficiently after nomination. This article develops an original decision framework in which semisynthetic accessibility qualifies, rather than replaces, pharmacological priority. The analysis distinguishes molecular synthetic complexity, retrosynthetic feasibility, locally available reactions and starting materials, parent or precursor supply, chemoselective modifiability, stereochemical obligations, and the experimentally reachable analogue neighborhood. These dimensions answer different questions and should not be collapsed into a single notion of synthesizability. The central proposed contribution is chemical learnability: the practical ability to generate structurally discriminating analogues that test structure–activity, selectivity, physicochemical, and liability hypotheses while preserving the features that made a natural-product scaffold interesting. Semisynthesis is positioned as one route to such learning, alongside late-stage functionalization, biosynthetic engineering, chemoenzymatic synthesis, and total synthesis. The framework assigns no universal weights, thresholds, or pass–fail rules, and chemically difficult scaffolds are not presumed inferior. Exceptional pharmacology, engineered supply, or new selective chemistry may justify continued advancement. Conversely, a highly ranked scaffold may warrant deprioritization when accessible chemistry cannot interrogate the relevant structure. Semisynthetic accessibility is therefore proposed as a context-dependent advancement variable whose value lies in estimating what medicinal chemistry can realistically learn next, not in predicting therapeutic success.

Key Words: *Semisynthetic accessibility, Natural product scaffolds, Chemical learnability, Medicinal chemistry, Computational prioritization, Analogue synthesis*

eIJPPR 2025; 15(3):78-87

HOW TO CITE THIS ARTICLE: Eriksen K, Berg S, Dahl M. Why Semisynthetic Accessibility Should Influence Which Natural Product Scaffolds Advance from Computational Prioritization to Medicinal Chemistry Development. Int J Pharm Phytopharmacol Res. 2025;15(3):78-87. <https://doi.org/10.51847/43QP6wgo8o>

INTRODUCTION

Natural products continue to contribute chemically distinctive starting points to drug discovery, including architectures that are difficult to reproduce through conventional screening collections [1]. Their historical contribution to therapeutics is substantial, but that history

records selected successes rather than evidence that every biologically promising natural-product scaffold is equally suitable for optimization [2]. The critical decision considered here therefore occurs after computational or experimental prioritization: whether a nominated scaffold provides a practical substrate for the iterative chemistry

Corresponding author: Kristian Eriksen

Address: Department of Semisynthetic Accessibility and Scaffold Prioritization, Faculty of Pharmacy, University of Oslo, Oslo, Norway.

E-mail: ✉ kristian.eriksen@farmasi.uio.no

Received: 01 February 2025; **Revised:** 16 May 2025; **Accepted:** 17 May 2025

This is an **open access** journal, and articles are distributed under the terms of the Creative Commons Attribution-Non Commercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.



needed to determine what its activity means and whether liabilities can be modified.

That distinction is easily obscured when ranking is driven by predicted binding, bioactivity, phenotypic similarity, target compatibility, or other compound-centered outputs. Those outputs address pharmacological plausibility. They do not establish whether enough material can be obtained, whether medicinally meaningful positions can be modified selectively, whether stereochemical information can be preserved, or whether a series of analogues can be generated without repeatedly reconstructing the scaffold. Methods developed specifically to interrogate natural-product structure–activity relationships show how substantial these chemical constraints can become [3].

Analogue accessibility provides a relevant computational precedent because it asks a different question from generic molecular synthesizability: how accessible are related compounds around a candidate [4]? For natural products, however, known analogue density may not reveal whether an isolated parent, fermentation product, biosynthetic precursor, or advanced synthetic intermediate can be converted experimentally into the particular analogues needed to test medicinal-chemistry hypotheses.

This article therefore proposes semisynthetic accessibility as an advancement variable distinct from both computational priority and general synthetic accessibility. It refers to the context-dependent opportunity to obtain a natural-product parent or sufficiently advanced precursor and transform it, directly or through hybrid chemical or biological routes, into structurally discriminating analogues. Its intended output is not therapeutic prediction but chemical learnability: whether advancement creates a tractable experimental path for learning which structural features control activity, selectivity, properties, and liabilities.

The missing synthetic dimension in computational prioritization

Synthetic accessibility is already relevant to computational molecular design, but existing approaches encode different

questions. Comparative assessment of accessibility scores demonstrates that they are not interchangeable and may disagree because they embody different assumptions about complexity or retrosynthetic feasibility [5]. Introducing accessibility constraints into generative design can materially change the structures proposed by an algorithm, showing that chemistry-aware prioritization can alter candidate space rather than merely annotate it after generation [6]. Computer-assisted retrosynthesis adds route-level information, yet industrial experience also shows that outcomes depend on reaction knowledge, search procedures, and implementation context [7].

The distinction becomes clearer across specific metrics. SCScore estimates synthetic complexity from reaction-corpus patterns [8]. RAScore rapidly approximates whether a retrosynthetic planner is likely to identify a route [9]. Graph-attention methods provide another molecule-level accessibility representation [10]. None by itself establishes whether an available natural-product scaffold supports an experimentally useful medicinal-chemistry series.

Route-aware analogue generation approaches that question more directly because proposed modifications are conditioned by plausible synthesis rather than structural similarity alone [11]. Retrosynthetic routes can also control where structural modifications enter optimization [12]. Practical feasibility becomes still more contextual when the question changes from whether a route exists somewhere to whether it can be executed with the reactions, reagents, and starting materials actually available to a discovery organization [13].

The proposed consequence is that synthetic accessibility should be decomposed before it influences natural-product advancement. Molecular complexity, route existence, resource-conditioned makeability, analogue-neighborhood reach, and semisynthetic opportunity describe related but non-equivalent constraints. **Table 1** summarizes those distinctions without converting them into a composite score.

Table 1. Accessibility constructs that answer different questions during natural-product scaffold advancement

Construct	Primary object assessed	Information provided	Information still missing	Advancement role
Synthetic complexity	Molecular structure	Relative construction difficulty	Parent supply; informative analogue routes	Coarse chemical triage
Retrosynthetic accessibility	Target and route search space	Plausible route solutions	Experimental execution; diversification value	Route-feasibility qualification
Molecule-level accessibility prediction	Encoded molecular structure	Model-based accessibility estimate	Scaffold-specific chemical reach	Rapid computational filtering
Resource-conditioned makeability	Target plus local reagents and transformations	Feasibility in a defined resource environment	Transferability; information yield	Program-specific planning
Analogue accessibility	Candidate plus represented neighboring chemistry	Relative reach of related compounds	Whether reachable changes are medicinally informative	Neighborhood comparison

Semisynthetic accessibility — proposed	Available parent or advanced precursor plus executable diversification chemistry	Opportunity to generate discriminating analogues from inherited complexity	Pharmacological quality or therapeutic value	Chemical-learnability qualification
--	--	--	--	-------------------------------------

Note: The final row is proposed in this article. The table distinguishes constructs rather than defining a validated numerical hierarchy.

The resulting architecture is broader than a collection of accessibility scores. **Figure 1** shows how computational priority is conditioned by non-equivalent forms of

chemical reachability while remaining distinct from the later pharmacological decision.

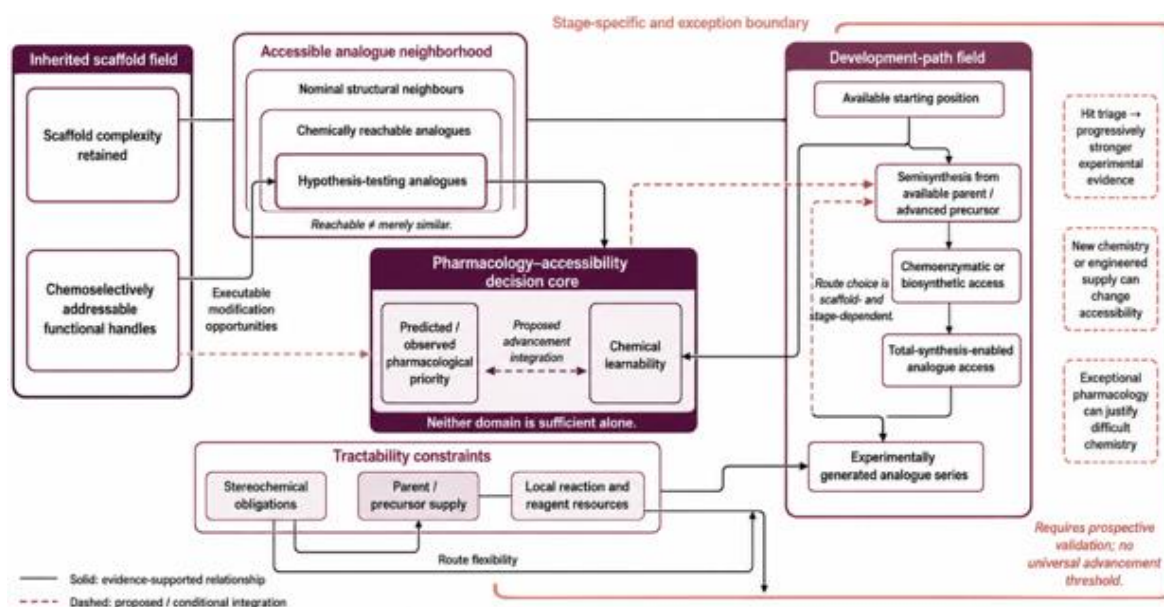


Figure 1. From molecular makeability to chemical learnability: a context-indexed architecture for natural-product scaffold advancement

Natural product complexity as opportunity and constraint

Natural-product complexity cannot be interpreted as either an intrinsic virtue or an intrinsic synthetic defect. Molecular complexity can contribute three-dimensional shape, stereochemical information, and differentiated interaction surfaces while simultaneously increasing the difficulty of synthesis, characterization, and systematic variation [14]. A decision process that penalizes complexity indiscriminately may therefore discard the same architecture that produced unusual biological recognition.

The opposite error is to romanticize complexity while ignoring the development burden it creates. Work addressing sustainable discovery and development in difficult chemical spaces shows that access, production, and downstream constraints can materially shape whether promising chemistry can be pursued [15]. Natural-product-inspired collection design likewise supports preserving useful structural richness while keeping chemical space experimentally interrogable [16].

This motivates a proposed distinction between structural complexity and diversification economy. Structural complexity describes what is already embedded in a scaffold; diversification economy describes how efficiently that inherited information can be perturbed from

a common parent, precursor, or advanced intermediate. A complex natural product may therefore be highly learnable when its core is inherited and several relevant positions can be changed selectively. A simpler scaffold may be expensive to learn from when each analogue requires reconstruction from an early intermediate.

Step count alone is consequently inadequate. Route convergence, the stage at which diversity can be introduced, the stereochemical information inherited from available material, and the proportion of the scaffold that remains experimentally addressable may matter as much as nominal route length. These considerations do not establish a universal complexity penalty; they specify why complexity must be interpreted in relation to the chemistry available around the scaffold.

Functional handles and chemically accessible diversification

The presence of heteroatoms, alkenes, alcohols, carbonyl groups, or other reactive motifs does not by itself make a natural product medicinal-chemistry tractable. A useful functional handle must support a transformation with sufficient site selectivity, chemoselectivity, substrate tolerance, and product stability to create an interpretable perturbation. Late-stage diversification matters because it

can exploit an already assembled natural-product architecture rather than require complete resynthesis for every analogue [17].

Its value extends beyond analogue count. Late-stage functionalization can alter properties relevant to drug discovery while preserving much of an advanced scaffold [18]. Broader molecular-editing strategies similarly show an expanding repertoire of transformations for structurally complex bioactive compounds [19]. The implication is conditional: new chemistry increases development opportunity only when it reaches positions relevant to the hypotheses that need testing.

Synthetic chemistry need not provide every handle alone. GenoChemetics demonstrated that biosynthetic production and synthetic modification can be coupled to diversify complex products [20]. Selective allylic hydroxylation of enones shows how controlled oxidation can create useful derivatives of complex drugs and natural products [21]. P450-catalysed oxyfunctionalization further illustrates enzyme-enabled access to positions difficult to address through conventional reagent control [22]. Site-selective visible-light alkylation of tropones shows more directly that reaction invention can change semisynthetic opportunity by enabling derivatives of complex natural products that were previously difficult to reach [23].

The proposed framework therefore treats chemoselective modifiability, rather than nominal functional-group count, as the advancement variable. A densely functionalized scaffold may remain poorly learnable if transformations require extensive protection, produce inseparable mixtures, destabilize the core, or modify only uninformative positions. A scaffold with fewer obvious handles may become more attractive when one selective transformation permits reproducible exploration of a consequential region.

Stereochemical complexity and the cost of analog generation

Natural-product stereochemistry adds information that medicinal chemistry may need to preserve, interrogate, invert, remove, or simplify. It also adds practical constraints when stereochemical identity depends on biosynthesis, asymmetric synthesis, resolution, or control of transformations that could epimerize sensitive centers. Analyses of naturally occurring racemic products show that assumptions of uniform homochirality are unsafe and that stereochemical state requires explicit chemical interpretation [24].

The relevant advancement question is therefore not how many stereocenters a scaffold contains, but what those stereocenters cost per informative analogue. An available natural product can transfer substantial stereochemical complexity into a semisynthetic series at little additional

construction cost. The same scaffold becomes more burdensome when meaningful modifications require rebuilding stereogenic centers, separating stereoisomers, or accessing configurations unavailable from the natural precursor.

Picrotoxinin provides an instructive hard case. C5 methylation increased synthetic accessibility and enabled a substantial analogue campaign while also changing stability and receptor selectivity [25]. The example prevents the simplistic inference that chemical simplification or easier analogue access preserves pharmacological meaning.

The framework therefore keeps stereochemical burden separate from generic molecular complexity. It asks whether relevant configurations are inherited reliably, whether diversification preserves them, and whether stereochemical alternatives required for SAR can realistically be produced. No numerical penalty is assigned to stereocenter count.

Supply, isolation yield, and precursor availability

Semisynthesis requires material before it requires reaction chemistry. Parent abundance, isolation efficiency, fermentation output, precursor availability, purification stability, and reproducibility of resupply therefore constrain how many analogue cycles can actually be executed. These variables should be measured for the specific program rather than inferred from molecular structure. Importantly, supply is not necessarily fixed: metabolic and cellular engineering provide alternative routes for producing natural products that are difficult to obtain from native sources [26].

The reconstruction of vinblastine biosynthesis in yeast demonstrates that an exceptionally complex plant-derived product can acquire a fundamentally different supply route through pathway engineering [27]. Engineered yeast has also supported de novo biosynthesis of berberine and halogenated benzyloquinoline alkaloids, illustrating that production engineering can intersect with modified chemistry rather than serving only parent supply [28]. Further engineering of protoberberine and benzophenanthridine alkaloid pathways supports the capacity of microbial systems to widen producible natural-product space [29].

Biocatalytic pathway design offers another route. An artificially concise cascade for phenethylisoquinoline manufacture shows that biosynthetic logic can be reorganized into shorter functional sequences [30]. A subsequent concise enzyme cascade produced natural and halogenated protoberberine alkaloids from readily available inputs while identifying remaining enzyme and manufacturing constraints [31]. Synthetic-biology research more broadly confirms that pathway, host, enzyme, and

regulatory engineering can expand production and diversification possibilities [32]. The premarineosin A case makes the connection to medicinal chemistry especially explicit: metabolic engineering improved access to the scaffold, after which late-stage functionalization expanded the chemical space that could be explored [33].

For the framework, supply has two roles. Material sufficiency asks whether reproducible parent or precursor supports repeated chemistry; supply-enabled optionality asks whether the production route also exposes congeners, engineered precursors, or advanced intermediates that enlarge subsequent analogue space. Neither role should be reduced to a universal yield threshold.

Semisynthesis versus total synthesis as development paths

Semisynthesis is most attractive when an available parent or advanced precursor already contains costly structural information and can be diversified at medicinally meaningful positions. Oxidative enzymes used to generate camptothecin analogues illustrate how complex natural-

product architecture can be retained while new variation is introduced [34]. Chemoenzymatic strategies extend this logic by placing enzymatic transformations at early or late stages specifically to support natural-product structure-activity exploration [35].

This does not make total synthesis a default negative comparator. Modern combinations of biocatalytic and radical retrosynthesis demonstrate that difficult natural products can acquire efficient synthetic routes that change an earlier tractability judgment [36]. Semisynthesis likewise has a documented role in pharmacologically informative analogue exploration of opioid-targeting natural products when isolated parents provide productive starting material [37]. The relevant comparison is therefore not short versus long synthesis, but which development path creates the most useful, reproducible access to hypothesis-testing analogues.

Table 2 separates the development paths by where complexity is inherited and where chemical diversity is introduced.

Table 2. Development paths differ in how they convert scaffold access into medicinal-chemistry learning

Path	Starting advantage	Principal learning opportunity	Characteristic constraint	Appropriate interpretation
Semisynthesis	Complex parent already available	Direct modification of inherited architecture	Parent supply and selective chemistry	Strong when relevant positions are addressable
Chemoenzymatic diversification	Enzymes provide otherwise difficult transformations	Modified complex analogues	Enzyme scope and substrate compatibility	Extends chemical synthesis
Chemoenzymatic total synthesis	Hybrid disconnections build difficult scaffolds	Efficient access to scaffold families	Route remains scaffold-specific	Can rescue apparent tractability problems
Biosynthetic or hybrid access	Pathway provides advanced material	Parent, precursor, or congener generation	Host and pathway constraints	May improve supply and diversification
Conventional total synthesis	Full control over construction sequence	Broad redesign when routes are modular	Repeated core construction can be costly	Rational when inherited access is too restrictive

Note: The comparison is conceptual and does not establish a universal route hierarchy.

Figure 2 makes the application logic explicit: neighborhood that can actually be explored, but it does not define intrinsic pharmacological value.

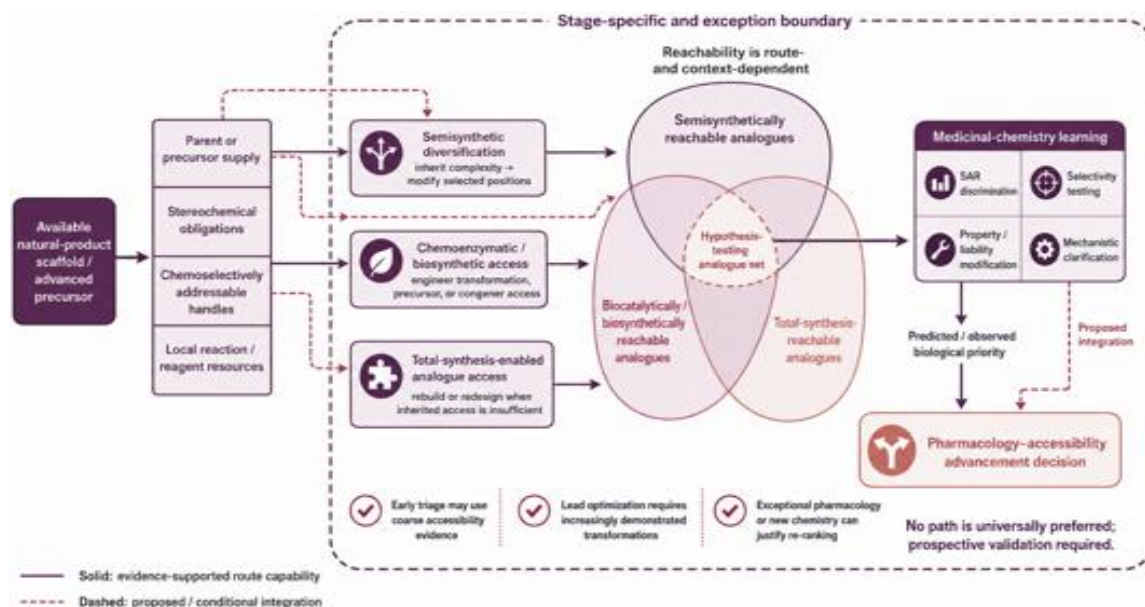


Figure 2. Development-path choice reshapes the experimentally reachable analogue neighborhood without defining a universal hierarchy

Mapping the accessible analog neighborhood around a scaffold

The relevant analogue neighborhood is not every structure that can be imagined around a natural product. Modular biosynthesis of halogenated tryptophan-derived molecules demonstrates that changing pathway inputs and combinations can generate families of related products [38]. Late-stage photo- and electrochemical methods likewise show that modifications available to alkene-containing natural products depend on actual functional-group and reaction compatibility [39].

The framework therefore defines the accessible analogue neighborhood as the subset of analogues reachable from specified parents or precursors through executable chemistry under defined resources. A still smaller subset is medically informative because it perturbs positions capable of distinguishing competing pharmacological or property hypotheses. Neighborhood size alone is consequently insufficient: positional coverage, stereochemical control, structural diversity, route reproducibility, and interpretability matter.

This distinction also separates computational neighborhood density from experimental reach. A scaffold surrounded by many catalogued or literature analogues may still offer poor local medicinal chemistry if those derivatives require routes unavailable to the current program. Conversely, a sparsely documented scaffold may become learnable when one modular semisynthetic or biosynthetic entry opens multiple orthogonal positions. The useful neighborhood is therefore conditional, dynamic, and experimentally anchored rather than a fixed geometric radius in chemical space. This prevents accessibility from becoming a synonym for literature

popularity and keeps the unit of decision aligned with the experiments a medicinal-chemistry team can actually execute.

A semisynthetic accessibility-weighted advancement framework

The proposed framework uses semisynthetic accessibility to qualify candidate advancement, not to generate an independent winning score. Accessibility is conditional on how analogue reach is represented, how synthetic feasibility is measured, and which reactions and starting materials are actually available [4, 5, 13]. These domains should remain visible rather than being compressed into a universal scalar.

For a computationally prioritized natural product, the assessment asks four sequential questions. First, is a reproducible parent, precursor, or advanced intermediate available? Second, can chemically and stereochemically controlled transformations reach positions relevant to the biological hypothesis? Third, does that chemistry generate a sufficiently discriminating analogue set rather than merely numerous derivatives? Fourth, would the resulting series allow medicinal chemistry to test activity, selectivity, properties, or liabilities that currently limit interpretation?

A candidate should rise relative to a similarly promising comparator when these questions indicate greater chemical learnability. It should fall when attractive predictions cannot be followed by experimentally informative perturbation. The framework does not imply that these dimensions are additive or compensatory. Strong supply cannot rescue an analogue neighborhood that interrogates only irrelevant positions, and abundant functional handles

do not compensate for unstable or stereochemically ambiguous products. Equally, difficult chemistry does not automatically outweigh compelling biological evidence. The framework is therefore best interpreted as a structured argument for evidence collection. Its value lies in forcing a program to make explicit what is presently predicted, what is experimentally demonstrated, what can change with new chemistry or supply, and what remains unknowable until analogues are actually made. It is a proposed decision architecture, not a validated predictive model.

Balancing predicted pharmacology against chemical tractability

Chemical tractability should neither dominate nor be ignored. Late-stage diversification can expand structure–activity or property exploration around advanced bioactive scaffolds without rebuilding the entire molecule for every analogue [17-19]. That makes tractability particularly relevant when two candidates offer comparable biological reasons for advancement.

The proposed balancing rule is asymmetric. Accessibility can strengthen or weaken confidence that medicinal chemistry will resolve uncertainty, but it does not convert mediocre biology into a compelling program. Conversely, unusually strong or distinctive pharmacology may justify accepting inefficient chemistry. The relevant question is not which scaffold is easiest to synthesize, but which candidate offers the better combination of biological importance and experimentally reachable information.

This approach also prevents a common inversion of priorities. A tractable scaffold should not advance merely

because it is easy to modify if the resulting series cannot answer an important biological question. Likewise, a high computational rank should not immunize a scaffold from chemical scrutiny when the relevant structure cannot be perturbed. Accessibility is most defensible as a qualifier of expected learning value rather than a substitute for pharmacological evidence.

When a difficult scaffold still deserves priority

A difficult scaffold should remain competitive when its biology is sufficiently distinctive or when its apparent chemical difficulty is changeable. The picrotoxinin case demonstrates that a structural modification can improve accessibility while also altering stability or selectivity, showing why simplification cannot be assumed to preserve pharmacological meaning [25]. Modern chemoenzymatic total synthesis shows that tractability can change as synthetic methodology advances [36], while new site-selective reactions can open semisynthetic positions that were previously inaccessible [23].

The proposed framework therefore includes an explicit exception pathway. A scaffold may remain high priority despite poor present accessibility when it provides unusual target engagement, selectivity, phenotype, mechanistic leverage, or another biological feature that is difficult to substitute. The decision should then acknowledge the extra chemical burden rather than conceal it.

Table 3 makes these re-ranking dimensions explicit without turning them into weights.

Table 3. Proposed dimensions for re-ranking natural-product candidates after initial computational prioritization

Dimension	Evidence sought	Favors advancement when	Caution
Pharmacological distinctiveness	Orthogonal activity, selectivity, or mechanistic evidence	Biology is difficult to replace	Does not erase chemical constraints
Parent or precursor access	Reproducible isolation, fermentation, biosynthesis, or synthesis	Resupply supports repeated cycles	Present supply may change
Chemoselective modifiability	Demonstrated transformations at relevant positions	Analogues test specific hypotheses	Handle count alone is insufficient
Analogue-neighborhood quality	Positional, stereochemical, and property diversity	Series is discriminating rather than merely large	Known analogues may exceed local reach
Route adaptability	Alternative semisynthetic, hybrid, or total-synthesis options	New routes change prior tractability judgments	Feasibility remains context dependent
Exceptional-scaffold override	Compelling biology despite low accessibility	Information value justifies disproportionate effort	Proposed exception, not a default

Note: These dimensions organize re-ranking but are not validated weights or thresholds.

Stage-specific use from hit triage to lead optimization

Accessibility evidence should become more demanding as a program advances. Early triage may reasonably use computational accessibility and route plausibility to identify obvious chemical dead ends. Later stages require

demonstrated resupply, reproducible transformations, and realised analogue series. Route-aware analogue generation is therefore more informative than a coarse score once medicinal-chemistry questions become specific.

Engineered biosynthesis can materially alter this assessment because complex natural-product supply can be reconstructed, broader pathway engineering can change available chemistry, and improved material access can enable subsequent diversification [27, 32, 33]. Accordingly, a scaffold rejected early for poor access may warrant reconsideration when supply, reaction methodology, or route design changes.

Stage-specific use also protects against premature certainty. A prediction that is adequate for portfolio triage should not be carried unchanged into lead optimization. As evidence accumulates, the decision should migrate from predicted makeability toward demonstrated chemical learning: actual material, actual transformations, and actual analogues that resolve uncertainty.

Decision boundaries and practical limitations

The framework has several non-negotiable boundaries. Synthetic-accessibility scores are model dependent [5], retrosynthetic outcomes depend on implementation and reaction knowledge [7], and practical makeability depends on local resources [13]. Analogue-accessibility estimates

can also reflect chemistry represented in databases rather than the complete experimentally reachable neighborhood [4].

Late-stage functionalization remains scaffold and reaction specific [17]. Photo- and electrochemical natural-product modification similarly depends on compatible functionality [39]. Supply engineering introduces enzyme, host, pathway, and manufacturing constraints [31], while synthetic biology does not make every natural product equivalently producible [32].

The framework also risks literature bias. Well-studied scaffolds may appear more accessible because more reactions and analogues have been reported, whereas neglected scaffolds may be chemically richer than their literature record suggests. Conversely, published diversification successes may overrepresent reactions that worked and underrepresent failed routes, unstable intermediates, difficult purifications, or analogue series that yielded little useful SAR.

These limitations are summarized in **Table 4** because the framework is most useful when its failure modes remain explicit.

Table 4. How apparent semisynthetic accessibility can fail to translate into medicinal-chemistry learning

Apparent advantage	Failure mechanism	Boundary condition	Required interpretation
Favorable accessibility score	Metric captures complexity or planner behavior rather than useful analogue reach	Model and training domain	Treat as screening evidence only
Plausible retrosynthetic route	Route is impractical in the local setting	Search space and reaction resources	Verify executable chemistry
Dense known analogue neighborhood	Literature density exceeds program-specific reach	Historical exploration bias	Distinguish known from reachable
Many functional groups	Selectivity or stability prevents informative modification	Scaffold and reaction context	Count executable handles, not motifs
Engineered supply	Host, enzyme, or scale limitations constrain resupply	Production system	Treat supply as conditional
Apparently difficult scaffold	New chemistry changes feasible route space	Methodological progress	Permit explicit re-ranking
Integrated decision framework	No prospective validation	Program and target context	Do not interpret as a predictive score

Note: The failure modes summarize manuscript-derived boundary conditions; they are not quantitative decision thresholds.

CONCLUSION

Computational prioritization identifies molecules worth investigating; medicinal chemistry determines whether those molecules can be interrogated deeply enough to become development programs. For natural products, that transition depends partly on whether inherited structural complexity can be converted into controlled experimental variation.

The framework proposed here treats semisynthetic accessibility as a context-dependent measure of chemical learnability. Parent and precursor supply, stereochemical obligations, chemoselective handles, route architecture, local resources, and the reachable analogue neighborhood

remain separate because each constrains a different part of the learning process. Semisynthesis may provide exceptional leverage when valuable complexity is already available, but chemoenzymatic, biosynthetic, or total-synthesis routes may be equally rational when they generate better experimental access.

The framework is not a universal ranking algorithm and does not establish that tractable scaffolds become better drugs. Its purpose is narrower: to make explicit whether a high-priority candidate can support the next sequence of medicinal-chemistry questions. Advancement should favor neither the easiest molecule nor the most exotic scaffold by default, but the candidate whose biology is sufficiently

important and whose chemistry permits the uncertainties that matter to be tested.

Acknowledgments: None

Conflict of interest: None

Financial support: None

Ethics statement: None

REFERENCES

- [1] Atanasov AG, Zotchev SB, Dirsch VM, Orhan IE, Banach M, Rollinger JM, et al. Natural products in drug discovery: Advances and opportunities. *Nat Rev Drug Discov.* 2021;20(3):200-16. doi:10.1038/s41573-020-00114-z
- [2] Newman DJ, Cragg GM. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J Nat Prod.* 2020;83(3):770-803. doi:10.1021/acs.jnatprod.9b01285
- [3] Ancajas CMF, Oyedele AS, Butt CM, Walker AS. Advances, opportunities, and challenges in methods for interrogating the structure activity relationships of natural products. *Nat Prod Rep.* 2024;41(10):1543-78. doi:10.1039/D4NP00009A
- [4] Ue T, Sato A, Miyao T. Analog accessibility score (AAScore) for rational compound selection. *J Chem Inf Model.* 2024;64(24):9350-60. doi:10.1021/acs.jcim.4c01691
- [5] Skoraczynski G, Kitlas M, Miasojedow B, Gambin A. Critical assessment of synthetic accessibility scores in computer-assisted synthesis planning. *J Cheminform.* 2023;15(1):6. doi:10.1186/s13321-023-00678-z
- [6] Parrot M, Tajmouati H, da Silva VBR, Atwood BR, Fourcade R, Gaston-Mathé Y, et al. Integrating synthetic accessibility with AI-based generative drug design. *J Cheminform.* 2023;15(1):83. doi:10.1186/s13321-023-00742-8
- [7] Saigiridharan L, Hassen AK, Lai H, Torren-Peraire P, Engkvist O, Genheden S. AiZynthFinder 4.0: Developments based on learnings from 3 years of industrial application. *J Cheminform.* 2024;16(1):57. doi:10.1186/s13321-024-00860-x
- [8] Coley CW, Rogers L, Green WH, Jensen KF. SCScore: Synthetic complexity learned from a reaction corpus. *J Chem Inf Model.* 2018;58(2):252-61. doi:10.1021/acs.jcim.7b00622
- [9] Thakkar A, Chadimová V, Bjerrum EJ, Engkvist O, Raymond 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
- [10] Yu J, Wang J, Zhao H, Gao J, Kang Y, Cao D, et al. Organic compound synthetic accessibility prediction based on the graph attention mechanism. *J Chem Inf Model.* 2022;62(12):2973-86. doi:10.1021/acs.jcim.2c00038
- [11] Dolfus U, Briem H, Rarey M. Synthesis-aware generation of structural analogues. *J Chem Inf Model.* 2022;62(15):3565-76. doi:10.1021/acs.jcim.2c00246
- [12] Dolfus U, Briem H, Gutermuth T, Rarey M. Full modification control over retrosynthetic routes for guided optimization of lead structures. *J Chem Inf Model.* 2023;63(21):6587-97. doi:10.1021/acs.jcim.3c01155
- [13] Hassen AK, Šicho M, van Aalst YJ, Huizenga MCW, Reynolds DNR, Luukkonen S, et al. Generate what you can make: Achieving in-house synthesizability with readily available resources in de novo drug design. *J Cheminform.* 2025;17(1):41. doi:10.1186/s13321-024-00910-4
- [14] 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
- [15] Miethke M, Pieroni M, Weber T, Brönstrup M, Hammann P, Halby L, et al. Towards the sustainable discovery and development of new antibiotics. *Nat Rev Chem.* 2021;5(10):726-49. doi:10.1038/s41570-021-00313-1
- [16] Bro FS, Laraia L. Unifying principles for the design and evaluation of natural product-inspired compound collections. *Chem Sci.* 2025;16(7):2961-79. doi:10.1039/D4SC08017C
- [17] Hong B, Luo T, Lei X. Late-stage diversification of natural products. *ACS Cent Sci.* 2020;6(5):622-35. doi:10.1021/acscentsci.9b00916
- [18] Castellino NJ, Montgomery AP, Danon JJ, Kassiou M. Late-stage functionalization for improving drug-like molecular properties. *Chem Rev.* 2023;123(13):8127-53. doi:10.1021/acs.chemrev.2c00797
- [19] Huo T, Zhao X, Cheng Z, Wei J, Zhu M, Dou X, et al. Late-stage modification of bioactive compounds: Improving druggability through efficient molecular editing. *Acta Pharm Sin B.* 2024;14(3):1030-76. doi:10.1016/j.apsb.2023.11.021
- [20] Sharma SV, Tong X, Pubill-Ulledemolins C, Cartmell C, Bogosyan EJA, Rackham EJ, et al. Living GenoChemetics by hyphenating synthetic biology and synthetic chemistry in vivo. *Nat Commun.* 2017;8(1):229. doi:10.1038/s41467-017-00194-3
- [21] Zheng CY, Yue JM. Allylic hydroxylation of enones useful for the functionalization of relevant drugs and

- natural products. *Nat Commun.* 2023;14(1):2399. doi:10.1038/s41467-023-38154-9
- [22] Bortz AR, Bennett JM, Fasan R. A skeletally diverse library of bioactive natural-product-like compounds enabled by late-stage P450-catalyzed oxyfunctionalization. *Chem.* 2024;10(11):3488-502. doi:10.1016/j.chempr.2024.08.003
- [23] Zeng QX, Zheng CY, Ge ZP, Zhao JX, Yue JM. Visible-light-mediated site-selective C(sp²)-H alkylation of tropones facilitates semi-synthesis of cephafortunoids A and B. *Chem Sci.* 2025;16(20):8836-44. doi:10.1039/D5SC01006C
- [24] Bitchagno GTM, Nchiozem-Ngnitedem VA, Melchert D, Fobofou SA. Demystifying racemic natural products in the homochiral world. *Nat Rev Chem.* 2022;6(11):806-22. doi:10.1038/s41570-022-00431-4
- [25] Tong G, Griffin S, Sader A, Crowell AB, Beavers K, Watson J, et al. C5 methylation confers accessibility, stability and selectivity to picrotoxinin. *Nat Commun.* 2023;14(1):8308. doi:10.1038/s41467-023-44030-3
- [26] Yang D, Eun H, Prabowo CPS, Cho S, Lee SY. Metabolic and cellular engineering for the production of natural products. *Curr Opin Biotechnol.* 2022;77:102760. doi:10.1016/j.copbio.2022.102760
- [27] Zhang J, Hansen LG, Gudich O, Viehriig K, Lassen LMM, Schrübbers L, et al. A microbial supply chain for production of the anti-cancer drug vinblastine. *Nature.* 2022;609(7926):341-7. doi:10.1038/s41586-022-05157-3
- [28] Han J, Li S. De novo biosynthesis of berberine and halogenated benzyloquinoline alkaloids in *Saccharomyces cerevisiae*. *Commun Chem.* 2023;6(1):27. doi:10.1038/s42004-023-00821-9
- [29] Jiao X, Fu X, Li Q, Bu J, Liu X, Savolainen O, et al. De novo production of protoberberine and benzophenanthridine alkaloids through metabolic engineering of yeast. *Nat Commun.* 2024;15(1):8759. doi:10.1038/s41467-024-53045-3
- [30] Gao Y, Li F, Luo Z, Deng Z, Zhang Y, Yuan Z, et al. Modular assembly of an artificially concise biocatalytic cascade for the manufacture of phenethylisoquinoline alkaloids. *Nat Commun.* 2024;15(1):30. doi:10.1038/s41467-023-44420-7
- [31] Li F, Yuan Z, Gao Y, Deng Z, Zhang Y, Luo Z, et al. A concise enzyme cascade enables the manufacture of natural and halogenated protoberberine alkaloids. *Nat Commun.* 2025;16(1):1904. doi:10.1038/s41467-025-57280-0
- [32] Seshadri K, Abad AND, Nagasawa KK, Yost KM, Johnson CW, Dror MJ, et al. Synthetic biology in natural product biosynthesis. *Chem Rev.* 2025;125(7):3814-931. doi:10.1021/acs.chemrev.4c00567
- [33] McBride CM, McCauley M, Harris NR, Amin S, Curtis BJ, Verhey-Henke L, et al. Metabolic engineering and late-stage functionalization expand the chemical space of the antimalarial premarineosin A. *Commun Chem.* 2025;8(1):391. doi:10.1038/s42004-025-01779-6
- [34] Nguyen TAM, Nguyen TD, Leung YY, McConnachie M, Sannikov O, Xia Z, et al. Discovering and harnessing oxidative enzymes for chemoenzymatic synthesis and diversification of anticancer camptothecin analogues. *Commun Chem.* 2021;4(1):177. doi:10.1038/s42004-021-00602-2
- [35] Li F, Deng H, Renata H. Chemoenzymatic approaches for exploring structure–activity relationship studies of bioactive natural products. *Nat Synth.* 2023;2(8):708-18. doi:10.1038/s44160-023-00358-8
- [36] Renata H. Combining biocatalytic and radical retrosynthesis for efficient chemoenzymatic synthesis of natural products. *Chem Soc Rev.* 2025;54(17):7913-32. doi:10.1039/D5CS00453E
- [37] Riley AP. Semi-synthesis in the exploration of opioid-targeting natural products. *Nat Prod Rep.* 2025;42(10):1664-75. doi:10.1039/D5NP00029G
- [38] Reed KB, Brooks SM, Wells J, Blake KJ, Zhao M, Placido K, et al. A modular and synthetic biosynthesis platform for de novo production of diverse halogenated tryptophan-derived molecules. *Nat Commun.* 2024;15(1):3188. doi:10.1038/s41467-024-47387-1
- [39] Sang JW, Zhang Y, Hu Z, Wang J, Zhang WD. Photo-/electro-chemical catalysis: A promising toolkit for late-stage functionalization of alkene-containing natural products. *Nat Prod Rep.* 2025;42(10):1717-46. doi:10.1039/D5NP00030K