**Reinforcement learning for retrosynthesis with chemoenzymatic reactions and type I polyketide synthases**

Yash Chainani1,2,3,4, Margaret Guilarte-Silva1,2, Geoffrey Bonnanzio1,2, Aindrila Mukhopadhyay3,4, Jay D. Keasling3,4,5,6,7,8, Hector Garcia Martin3,4,9,10, Tyler W. H. Backman3,4,Keith E. J. Tyo1,2and Linda J. Broadbelt1,2,3,4

1. Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL, USA
2. Center for Synthetic Biology, Evanston, IL, USA
3. Joint BioEnergy Institute, Emeryville, CA, USA
4. Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
5. QB3, University of California, Berkeley, Berkeley, CA, USA
6. Department of Chemical and Biomolecular Engineering, University of California, Berkeley, Berkeley, CA, USA
7. Department of Bioengineering, University of California, Berkeley, Berkeley, CA, USA
8. Center for Biosustainability, Danish Technical University, Lyngby, Denmark
9. BCAM, Basque Center for Applied Mathematics, Bilbao, Spain,
10. DOE Agile BioFoundry, Emeryville, California

# **Abstract**

Retrosynthetic planning tools have traditionally relied on either organic chemistry or monofunctional enzymatic catalysis to decompose target molecules into simpler precursors, leaving a third powerful synthesis modality, multifunctional type I polyketide synthases (PKSs), largely untapped. Here we present TridentSynthRL, a hierarchical multi-agent reinforcement learning framework, built upon the Monte Carlo tree search (MCTS) algorithm, that unifies all three synthesis routes within a single computational platform to explore a broad chemical space of small-molecule targets. Given a target molecule, TridentSynthRL first employs a chemoenzymatic MCTS agent to retrosynthetically fragments this target using our previously published chemical and enzymatic reaction templates. When this search produces an intermediate identified as a polyketide, a separate MCTS agent is subsequently spawned to design chimeric PKS assembly lines capable of synthesizing this intermediate in the forward direction from simple acyl-CoA building blocks. The identification of polyketide intermediates is enabled by a supervised graph neural network classifier trained on approximately one million synthetically generated polyketides and two million structurally similar chemoenzymatically modified variants. The chemoenzymatic MCTS search is further guided by our previously published enzymatic reaction feasibility classifier, DORA-XGB, and by reaction enthalpy estimates for synthetic chemistry transformations, thereby enabling prioritization of thermodynamically favorable pathways. We deployed TridentSynthRL on the kavalactone family from Piper methysticum, designing retrosynthetic pathways to all fifteen members, each accessible via at least one PKS-containing route. We further demonstrate a more efficient hybrid chemoenzymatic pathway to the natural product cryptofolione, improving upon our prior work that combined only mono- and multifunctional enzymes without synthetic chemistry.

# **Introduction**

The scalable synthesis of structurally complex molecules at high titers, rates, and yields is a foundational problem in biological and chemical manufacturing. Natural products and their derivatives, for instance, are particularly valuable manufacturing targets given their widespread use as small-molecule therapeutics, agrochemicals, and commodity chemicals. Yet, their structural diversity and precise stereochemistry make large-scale manufacturing exceptionally difficult. Retrosynthesis, the systematic decomposition of target molecules into simpler, readily available precursors, has long provided a strategic framework for designing synthetic pathways to such complex molecules for eventual scale-up. Over the past several years, computational retrosynthesis tools have advanced rapidly, driven by both automatically-extracted and expert-curated reaction templates, template-free language models, and increasingly, machine learned heuristics and scores. Template-based methods apply explicit reaction rules encoded as SMARTS patterns to recursively deconstruct targets, as exemplified by tools such as ASKCOS, RetroPath, RetroRules, and our group’s own DORAnet, amongst others. Template-free approaches instead use transformer architectures trained on previously published reactions in the literature to predict retrosynthetic transformations without predefined rules. Among the various search algorithms employed to navigate the combinatorial space of multi-step retrosynthesis, Monte Carlo tree search (MCTS) has emerged as a particularly effective strategy, owing to its ability to balance the exploration of novel routes with the exploitation of promising intermediates. First applied to retrosynthetic planning by Segler et al., MCTS draws on the same principled exploration–exploitation trade-off that underpins AlphaGo and AlphaZero, where it was instrumental in achieving superhuman performance in the board game Go.

Despite these algorithmic advances, most retrosynthesis tools have thus far focused exclusively on either synthetic organic chemistry or monofunctional enzymatic transformations to propose and evaluate routes to valuable small molecules. Recognizing the complementary strengths of chemical and enzymatic catalysis, several recent computational platforms have begun integrating both modalities, enabling the exploration of a broader chemical space than would be accessible through either approach alone.

Monofunctional enzymes, defined here as enzymes that catalyze single chemical transformations, such as alcohol dehydrogenases, aminotransferases, or decarboxylases, are highly effective at regio- and stereoselectively modifying specific functional groups on a substrate’s carbon backbone (Figure 1). However, enzymatic catalysis alone cannot support the full range of transformations required for small-molecule synthesis. Enzymes typically operate within narrow ranges of temperature, pressure, and solvent conditions and often lose activity outside physiological environments, thereby limiting the diversity of reactions that can be performed biologically. In contrast, synthetic chemistry can access a wider suite of transformations through either transition-metal catalysis or reaction conditions that fall outside common physiological limits. Examples include palladium- or nickel-catalyzed carbon–carbon bond-forming reactions such as Suzuki, Heck, and Liebeskind–Srogl cross-couplings, as well as olefin metathesis reactions used to construct or remodel carbon scaffolds. These transformations often require non-physiological conditions, metal catalysts that are cytotoxic to cells, and/ or involve intermediates, such as aldehydes, radicals, or long-chain hydrophobic metabolites that are typically unstable or poorly soluble in aqueous environments (Figure 1). Collectively, these constraints highlight the value of integrated retrosynthetic approaches that jointly consider chemical and enzymatic transformations within a single planning framework.

To this end, our group recently released **DORAnet**, a synthesis planning tool that recursively applies 386 synthetic chemistry and 3,604 enzymatic chemistry reaction templates to a given target molecule for a predetermined number of enzymatic and synthetic chemistry steps. Earlier efforts to merge chemistry and biology include the dual-network framework developed by Levin et al., which maintains separate machine learning models trained on chemical and enzymatic reactions and combines their output probability scores during search to balance both modalities. Sankaranarayanan and Jensen similarly coupled ASKCOS with an enzymatic reaction search module using RetroBioCat to support multistep chemoenzymatic pathway design, while Kreutter and Reymond employed transformer models in an iterative loop to propose multistep chemoenzymatic pathways.

While integrating synthetic chemistry and monofunctional enzymatic transformations expands accessible chemical space, incorporating multifunctional enzyme systems such as polyketide synthases (PKSs) can broaden this space even further (Figure 1). Here, we define multifunctional enzymes as enzyme assemblies capable of catalyzing multiple distinct transformations within a single coordinated complex, in contrast to monofunctional enzymes that catalyze individual reactions only. Type I PKSs are modular megaenzymes that function as molecular assembly lines composed of multiple catalytic domains operating in concert to iteratively construct elongated carbon backbones from simple acyl-coenzyme A (acyl-CoA) building blocks (Figure 1). Central to this process is the ketosynthase (KS) domain, which catalyzes iterative carbon–carbon bond-forming Claisen condensation reactions that extend the growing polyketide chain and enable the construction of complex carbon scaffolds with defined stereochemistry. Since these successive reactions occur within the same multienzyme complex, PKSs and related systems can generate complex scaffolds without isolating intermediates after each transformation, potentially reducing downstream separation and purification requirements relative to stepwise chemical or monofunctional enzymatic synthesis.

|  |
| --- |
| **A circular puzzle with different colored circles  AI-generated content may be incorrect.** |
| **Figure 1.** TridentSynthRL leverages the complementary strengths of synthetic chemistry as well as mono- and multi-functional enzymes to access a wide chemical space of molecular targets. Monofunctional enzymes, which catalyze a single reaction, excel at regioselectively and stereoselectively performing small modifications on a substrate’s carbon backbone. Meanwhile, multifunctional enzymes, which catalyze multiple reactions, can construct such carbon backbones through the iterative catalysis of acyl-CoA substrates, as in the case of polyketide synthases (PKSs). Despite their ability to catalyze reactions in a controlled and selective manner, enzymes cannot function under temperatures and pressures outside of typical physiological conditions, nor can their host cells survive cytotoxic intermediates, such as aldehydes or long-chain alcohols. Consequently, the suite of functional group transformations biology can catalyze is relatively narrow, thereby providing an opportunity to couple both forms of enzymatic chemistry with synthetic chemistry to ultimately synthesize more targets than would be possible with either route alone. |

In biological systems, PKSs are encoded within biosynthetic gene clusters and are responsible for the production of numerous therapeutically valuable natural products that are difficult to access through synthetic chemistry alone. Recent advances in PKS engineering and heterologous expression have demonstrated that these assembly lines can be reconfigured to produce not only natural product analogues but also a range of commodity and specialty chemicals, underscoring their potential as a programmable biosynthetic platform for manufacturing. **In our prior work, we developed a computational PKS-based forward-synthesis planning platform that combined mono- and multifunctional enzymatic transformations in a breadth-first-search approach to propose pathways toward both commodity chemicals and structurally complex natural products, such as cryptofolione.**

Here, we have built upon our previous work and harnessed reinforcement learning (RL) to efficiently integrate all three synthesis modalities, i.e., synthetic chemistry, monofunctional enzymatic transformations, and multifunctional enzyme systems within a single unified framework, TridentSynthRL. To the best of our knowledge, this represents the first computational retrosynthesis platform to simultaneously leverage all three synthesis routes. TridentSynthRL employs a hierarchical search strategy based on the Monte Carlo tree search (MCTS) algorithm to navigate the vast combinatorial design space that emerges when these synthesis routes are considered simultaneously. Given a target molecule, a chemoenzymatic MCTS agent first uses chemical and monofunctional enzymatic reaction templates to iteratively fragment this target into simpler upstream intermediates. When a candidate polyketide is detected among these intermediates, a secondary forward-synthesis MCTS agent is invoked to evaluate whether this fragment can be synthesized by assembling modular PKS architectures from simple acyl-coenzyme A building blocks. In order to efficiently detect polyketides, we trained a supervised graph neural network (GNN) classifier on approximately one million synthetically generated polyketides and two million structurally related chemoenzymatically modified molecules. To prioritize experimentally viable routes, TridentSynthRL also incorporates reaction thermodynamic estimates and machine-learning–based enzymatic reaction feasibility scores directly into the search policy, thereby biasing the algorithm towards feasible and thermodynamically favorable pathways. We demonstrate the capabilities of TridentSynthRL through two representative case studies. First, we applied TridentSynthRL to the kavalactone family of natural products from Piper methysticum, a class of polyketide-derived lactones whose structural complexity present challenges for traditional synthetic chemistry-based routes. In this study, TridentSynthRL identified pathways to all fifteen kavalactones, with each target accessible through at least one route incorporating a PKS-derived intermediate. We further applied TridentSynthRL to the structurally complex natural product cryptofolione, for which a hybrid chemoenzymatic–PKS synthesis strategy was found that reduces the number of required PKS modules relative to our previously proposed designs, thereby improving experimental feasibility since shorter PKS chimeras are easier to express heterologously in microbial cell factories. Together, these case studies illustrate how integrating synthetic chemistry, monofunctional enzymes, and multifunctional PKS assembly within a unified search architecture enables the identification of experimentally plausible routes various targets that may be difficult to access using any single synthesis modality alone.

**Results**

### **Formulating chemoenzymatic retrosynthesis as a Markov Decision Process**

In building TridentSynthRL, we formulated the chemoenzymatic retrosynthesis problem as a Markov decision process (MDP) and addressed it using the Monte Carlo tree search (MCTS) algorithm. In RL, an MDP provides a mathematical framework for sequential decision making in which an agent transitions between states by selecting actions and receives rewards that guide it towards desirable outcomes. Within this MDP framework, retrosynthetic planning can be viewed as a sequence of decisions, each corresponding to the application of a reaction template, that recursively transform an input small-molecule target, which defines the initial state, into simpler precursor molecules, which represent subsequent states.

In our implementation, each node (or state) in our search tree represents a molecule encoded as an RDKit Mol object, with the root node corresponding to the target molecule. Child nodes are then generated by applying both chemical and enzymatic reaction templates with DORAnet (available at <https://github.com/wsprague-nu/doranet>) in the retrosynthetic direction, thus producing upstream intermediates. In this way, the search tree expands outward from the target towards simpler building blocks (Figure 2). The objective of our search is to identify sequences of transformations that can convert the target molecule into commercially available or biologically accessible building blocks while maximizing pathway feasibility.

|  |
| --- |
| A screen shot of a computer  AI-generated content may be incorrect. |
| **Figure 2.** Overview of the TridentSynthRL multi-agent Monte Carlo tree search (MCTS) framework for retrosynthetic planning**.** The algorithm proceeds through three iterative phases. (1) Selection: The chemoenzymatic agent traverses the search tree from the target molecule, selecting a leaf node (red circle) for expansion based on learned value estimates. (2) Chemoenzymatic Expansion: The selected node is expanded by applying chemical (blue arrows) and enzymatic (green arrows) retrosynthetic transformations to generate precursor nodes. Expansion terminates when all precursors resolve to commercially available building blocks (blue squares), biological building blocks (green squares), or known polyketide intermediates (orange squares), yielding a terminal state *S*\_T. When a predicted or known polyketide intermediate is encountered during expansion, a specialized PKS agent is launched to explore polyketide synthase assembly line configurations (orange arrows) for that substructure. (3) Update: Upon reaching a terminal state, rewards are backpropagated (red dashed arrows) from the terminal node up to the root, updating value estimates along the traversed path to guide future iterations toward higher-quality retrosynthetic routes. |

Classical Monte Carlo tree search (MCTS) proceeds through four stages during each iteration: (1) selection, (2) expansion, (3) simulation (rollout), and (4) update. In the simulation phase, the MCTS algorithm typically performs a fast forward rollout from a newly expanded node to estimate the long-term value of that state by randomly sampling possible future trajectories. In retrosynthetic planning, such rollouts would entail repeatedly applying reaction templates for many additional steps to determine whether a pathway ultimately reaches viable starting materials. However, realistic synthetic routes may span numerous reaction steps, and the branching factor of chemical reaction networks is large. As a result, rollout simulations over long retrosynthetic horizons can be computationally expensive and prone to accumulating chemically infeasible or unrealistic transformations, introducing noise into node value estimates. To avoid this cost while maintaining efficient search, we omit the rollout phase in TridentSynthRL and instead rely on heuristic reaction feasibility and thermodynamic scores to provide immediate feedback to the search process during node expansion and backpropagation.

Consequently, within TridentSynthRL, each iteration consists of three steps: (1) selection, (2) expansion and (3) update. During the selection step, the existing retrosynthesis tree is traversed beginning from the root node (target molecule) until a leaf node is reached (Figure 2). This selection is guided by the widely-applied upper confidence bound 1 (UCB1) selection policy (see Methods), which balances the exploration of under-visited nodes with the exploitation of promising intermediates. In the expansion step, the selected leaf node is expanded (Figure 2) by applying all 386 chemical and 3604 enzymatic reaction templates within DORAnet in the retrosynthetic direction, generating upstream precursor states (compounds) that are then added to the existing tree as new child nodes. Finally, during the update step, the rewards associated with these newly generated nodes are propagated back through the tree (Figure 2) to update each node’s visit counts and value estimates. Within this framework, the state space of the chemoenzymatic agent consists of individual molecular compounds encountered during retrosynthetic decomposition while the action space comprises the full set of 3990 reaction templates within DORAnet. Each action corresponds to applying one template retrosynthetically to the current molecule to generate precursor fragments, which are then added to the tree as upstream intermediates. Prior to running TridentSynthRL, users can specify both the maximum number of search iterations and the maximum tree depth for which to execute the retrosynthetic search and TridentSynthRL terminates when either of the two stopping criterion is reached. After termination, the ensuing retrosynthesis tree is traversed one final time to extract pathways between the initial root node and any terminal leaf nodes.

A node is considered terminal if it satisfies one of three criteria: (1) it corresponds to a commercially available chemical building block drawn from the Enamine catalog (~278,000 compounds), (2) it matches a biological building block from a curated set of common metabolites across multiple organisms, or (3) it is identified as a polyketide intermediate for which a chimeric PKS design has been successfully generated. Nodes corresponding to prohibited or hazardous compounds (~652 entries) are immediately pruned from the search.

Within RL algorithms, designing reward functions is crucial to attaining good solutions. Here, to guide the retrosynthetic search toward experimentally feasible routes, we employ a modular reward policy that accounts for both the identity of individual intermediate molecules that are encountered and the feasibility of the pathway taken to reach them. At the node level, compounds corresponding to terminal states, i.e., commercial or biological building blocks, or polyketide intermediates verified to be PKS-synthesizable receive a fixed base reward of . For non-terminal intermediates, a dense reward signal is provided using a rescaled synthetic accessibility (SA) score derived from RDKit, which is linearly transformed (see Methods) so that molecules that are more synthetically accessible receive higher rewards.

Path-level feasibility is subsequently incorporated through reaction-level scores derived from both machine learning and thermodynamic estimates. For enzymatic steps, feasibility is evaluated using our previously published DORA-XGB reaction feasibility classifier, which produces dimensionless scores ranging from 0 to 1, with higher scores representing more feasible reactions. For synthetic chemistry transformations, reaction enthalpies are computed using our previously released open-sourced thermodynamics calculator PAthermo (available at <https://github.com/dmdqy/pathermo>). Since reaction enthalpies from PAthermo are in kcal/mol, these are first converted to the same numerical scale as DORA-XGB scores using a sigmoidal transformation that penalizes highly endergonic steps while preserving favorable or near-thermoneutral reactions (see Methods). This normalization ensures that feasibility metrics for chemical and enzymatic reactions are directly comparable and bounded between 0 and 1.

For a given pathway, the cumulative feasibility is computed as the geometric mean of the per-step feasibility scores along that route. Using the geometric mean allows pathways of different lengths to be compared fairly, since it avoids systematically penalizing longer but otherwise plausible sequences. The final reward assigned to a node is then calculated as the product of the state-level reward (e.g., terminal-state or synthetic accessibility score) and this path-level feasibility factor, allowing the search to preferentially expand pathways that are both synthetically accessible and chemically or biologically feasible. During the update phase of TridentSynthRL, these normalized rewards are propagated back through the tree

A distinguishing feature of TridentSynthRL is its hierarchical multi-agent design. When the expansion step produces an intermediate predicted to be compatible with PKS biosynthesis, a specialized MCTS agent is launched in the forward-synthesis direction to evaluate whether the fragment can indeed be constructed by a PKS assembly line. If the PKS agent successfully identifies a chimeric PKS design that yields the target fragment, a terminal reward of +1.0 is assigned and propagated back through the chemoenzymatic search tree. If no PKS design is found, the intermediate instead receives its scaled SA score. This hierarchical coupling enables the coordinated exploration of hybrid pathways that combine chemical, enzymatic, and PKS-based transformations and is described in detail in the following section.

### **Formulating chimeric polyketide synthase design as a Markov Decision Process**

When the chemoenzymatic MCTS agent identifies an intermediate predicted to be a polyketide, TridentSynthRL spawns a second, specialized MCTS agent that operates in the forward-synthesis direction to evaluate whether a chimeric type I PKS design can be found to produce this fragment from simple acyl-CoA building blocks. This PKS MCTS agent is built upon our previously released breadth-first-search PKS design tool RetroTide (available at: <https://github.com/JBEI/RetroTide/tree/main>). We formulated this PKS design problem as a separate MDP in which sequential decisions correspond to the iterative addition of extension modules to a growing polyketide chain, mirroring the biosynthetic logic of type I PKSs.

In this MDP, each state represents a PKS intermediate defined by three components: (1) the chemical structure of the polyketide product bound to the synthase via a thioester linkage to the acyl carrier protein (ACP), encoded as an RDKit Mol object, (2) the complete PKS module architecture specifying the ordered sequence of catalytic domains selected thus far, and (3) the current depth in the search tree, corresponding to the number of extension modules that have been appended. The root state of this MDP corresponds to an empty PKS design with no modules, from which the search builds outward by iteratively adding extension modules in the forward biosynthetic direction.

The action space of the PKS agent comprises the set of possible extension modules that can be appended at each step. Each action corresponds to adding a single extension module to the growing assembly line, where the module is defined by the choice of extender unit, i.e., the acyl-CoA substrate loaded by the acyltransferase (AT) domain (e.g., malonyl-CoA, methylmalonyl-CoA, or ethylmalonyl-CoA), and the combination of optional tailoring domains that determine the reduction state of the newly incorporated unit. Specifically, each extension module minimally contains a ketosynthase (KS) domain, which catalyzes the Claisen condensation that extends the polyketide chain by two carbon units, and an ACP domain, which tethers the growing intermediate. In addition, each module may optionally include a ketoreductase (KR) domain that reduces the β-keto group to a hydroxyl, a dehydratase (DH) domain that eliminates water to form an enoyl intermediate, and an enoylreductase (ER) domain that fully saturates the carbon–carbon double bond. This combinatorial selection of extender units and domain architectures gives rise to four possible reduction states per extension module: (1) no reduction (KS-AT-ACP), (2) partial reduction (KS-AT-KR-ACP), (3) partial reduction with dehydration (KS-AT-KR-DH-ACP), and (4) full reduction (KS-AT-KR-DH-ER-ACP). The first module in the assembly line additionally specifies the loading module, which determines the starter unit (e.g., acetyl-CoA, benzoyl-CoA, or propionyl-CoA, among others) that initiates chain biosynthesis. At each expansion step, our previously released PKS design tool RetroTide enumerates the set of feasible one-step extensions given the current PKS architecture and the target polyketide fragment, returning up to 25 candidate designs ranked by structural similarity to the target.

Like the chemoenzymatic agent, the PKS agent proceeds through three stages per MCTS iteration: (1) selection, (2) expansion, and (3) update. During selection, the agent traverses the existing PKS design tree from the root using a modified UCB1 policy that incorporates subgraph-guided pruning to focus the search on chemically promising branches. Specifically, when evaluating unvisited child nodes, the agent checks whether the polyketide product at that node is a subgraph of the target fragment using RDKit’s substructure matching routines. Children whose products are confirmed subgraphs of the target receive maximal selection priority, forcing their immediate exploration, while children whose products fail the subgraph test are assigned minimal priority and are effectively pruned from the search. This structural filtering ensures that the agent only expands PKS designs whose growing intermediates remain consistent with the target’s carbon skeleton. For previously visited children, standard UCB1 scoring is applied, balancing exploitation of high-value nodes with exploration of under-visited alternatives.

During expansion, the selected leaf node is expanded by calling RetroTide’s one-step PKS design function, which extends the current assembly line by a exactly one module. RetroTide evaluates the set of possible extender units and domain architectures and returns up to 25 candidate designs, each representing a distinct next module choice. These candidates are added to the tree as new child nodes, with each child inheriting the parent’s PKS architecture augmented by the newly appended module.

Following expansion, each newly generated child node is immediately evaluated to determine its reward. For each child, the agent simulates product release from the PKS assembly line through two thioesterase (TE)-catalyzed mechanisms: thiolysis, which cleaves the thioester bond to yield a linear carboxylic acid, and lactonization, which performs an intramolecular cyclization to generate a lactone. If the released product from either mechanism is identical to the target fragment, the node is marked terminal and assigned a reward of +1.0. If no exact match is found, a partial reward is computed using the maximum common substructure (MCS) similarity between the released product and the target, defined as the ratio of atoms in the MCS to the total atom count of the union of the query and reference molecules. This immediate node-level evaluation replaces the simulation (rollout) phase used in classical MCTS, analogous to the chemoenzymatic agent's omission of rollouts described in the previous section.

During the update step, the reward obtained from node evaluation is backpropagated from the expanded leaf node up through all ancestors to the root, incrementing each node's visit count and accumulating its total reward. These statistics are subsequently used by the UCB1 selection policy in future iterations to compute the average value and exploration bonus for each node, thereby directing the search toward PKS architectures that have historically yielded high-similarity intermediates.

The PKS MCTS search terminates when any of three conditions is met: (1) the maximum tree depth is reached, which corresponds to the maximum number of extension modules permitted in the chimeric assembly line (default: 15 modules), (2) an exact target match is identified through either thiolysis or lactonization release, or (3) no valid child nodes remain after subgraph pruning, indicating that all possible extensions have diverged from the target structure. Upon termination, if the PKS agent has identified at least one successful chimeric design, a terminal reward of +1.0 is returned to the parent chemoenzymatic agent and backpropagated through the retrosynthetic search tree, thereby reinforcing the pathway that led to the polyketide intermediate. If no successful design is found, the intermediate instead receives its rescaled synthetic accessibility score, as described in the previous section. This hierarchical coupling between the two agents enables TridentSynthRL to merge chemoenzymatic reactions with PKSs.

**Training a supervised classifier to detect polyketides**

A critical component of TridentSynthRL is the mechanism by which polyketide intermediates are detected during the chemoenzymatic MCTS expansion, thereby triggering the spawning of PKS agents. The simplest approach to polyketide detection is a static database lookup against a precomputed library of known PKS products; in our implementation, this library comprises approximately 106,000 polyketide structures enumerated by RetroTide, and provides a deterministic baseline for validation and benchmarking. However, the accessible chemical space of type I PKS products grows combinatorially with the number of extension modules, extender unit choices, and tailoring domain configurations, meaning that any finite enumeration will inevitably miss novel scaffolds that fall outside the precomputed set. Moreover, the synthetic accessibility (SA) score employed as a dense reward signal for guiding the chemoenzymatic search is inherently biased toward the commercial chemical building blocks on which it was originally calibrated. This score tends to assign low (favorable) synthesizability values to compounds resembling entries in chemical vendor catalogs, while penalizing biomolecules, metabolites, and polyketide natural products with higher (less favorable) scores—precisely the compounds that TridentSynthRL should recognize as biologically accessible starting materials for downstream PKS-based synthesis. We therefore sought to develop a learned classifier capable of generalizing from a representative training set to recognize the structural hallmarks of polyketide backbones, including the characteristic patterns of β-keto, β-hydroxy, enoyl, and fully reduced functionalities that arise from iterative Claisen condensation and optional tailoring domain modifications within PKS assembly lines.

|  |
| --- |
| A screenshot of a cell phone  AI-generated content may be incorrect. |
| **Figure 3.** Polyketide libraries and their chemoenzymatically-modified counterparts were synthetically generated to build a training set for a supervised polyketide classification model. Our previously released PKS design tool, RetroTide, was used to build an initial database of polyketide scaffolds by combinatorially assembling chimeric type I PKS designs *in silico* with 1) a single loading module, 2) up to 3 extension modules (each with all possible domain architectures, i.e., KS, AT, ACP, and optional KR, DH, and ER domains), and a 3) terminal thioesterase domain allowing for both thiolysis and lactonization offloading reactions, yielding approximately 1 million core polyketide structures (labelled 1) for training. These structures were then modified further via DORAnet for a single step using all synthetic chemistry and enzymatic chemistry reaction templates. For each polyketide, the most chemically similar product from the synthetic chemistry modified set as well as that from the from the enzymatic chemistry modified set was selected, yielding approximately 2 million polyketide variants (labelled 0) to be added to the training set. |

We constructed the training set through a two-step synthetic data generation pipeline designed to provide both positive examples of polyketide scaffolds and structurally similar hard negatives (Figure 3). In the first step, RetroTide was used to combinatorially enumerate chimeric type I PKS assembly line configurations *in silico*, systematically varying the choice of starter unit in the loading module, the extender unit and tailoring domain architecture of each extension module, and the thioesterase-catalyzed offloading mechanism to generate approximately one million core polyketide scaffolds, each labelled as class 1 (polyketide). We deliberately restricted the enumeration to a maximum of three extension modules per assembly line for two complementary reasons: first, shorter chimeric PKS constructs are more readily expressed and folded in heterologous hosts such as *Escherichia coli* and *Streptomyces* species, making them more experimentally tractable for downstream validation; second, this constraint keeps the enumerated library to a computationally manageable size while still capturing the core structural diversity of PKS products, including linear carboxylic acids generated by thiolysis and macrolactones generated by lactonization at all feasible ring sizes.

In the second step, we generated structurally similar but non-polyketide molecules to serve as challenging negative examples for the classifier. Each of the approximately one million polyketide scaffolds was subjected to a single step of DORAnet chemoenzymatic expansion using the full complement of 3,604 enzymatic and 386 synthetic reaction templates, producing a large pool of chemoenzymatically modified derivatives. From this pool, we selected two molecules per parent polyketide—the most structurally similar product from the enzymatic modification set and the most structurally similar product from the synthetic modification set, as ranked by Tanimoto similarity on Morgan fingerprints—yielding approximately two million modified variants, each labelled as class 0 (non-polyketide). The rationale for this similarity-based selection strategy is that it produces hard negatives: molecules that retain much of the parent polyketide’s carbon skeleton and functional group topology but have undergone at least one transformation outside the scope of PKS biosynthesis, such as a hydroxylation, oxidation, decarboxylation, or cross-coupling reaction. By training on these near-neighbor negatives rather than randomly sampled non-polyketide molecules, the classifier is forced to learn the precise structural motifs that are diagnostic of PKS products—for example, the specific spacing and stereochemistry of β-oxygenation patterns along the polyketide backbone—rather than relying on gross molecular dissimilarity. This design choice is particularly important for TridentSynthRL’s intended deployment context, where the chemoenzymatic MCTS agent generates intermediates that are, by construction, one or a few reaction steps removed from potential polyketide precursors, and the classifier must therefore discriminate between true polyketides and their close chemoenzymatic derivatives at the decision boundary.

|  |
| --- |
| A diagram of a graph  AI-generated content may be incorrect. |
| **Figure 4.** Architecture of supervised GNN. Input SMILES strings are parsed with RDKit into molecular graphs in which atoms are represented as nodes and bonds as edges. Each atom is encoded as a 40-dimensional one-hot feature vector capturing atomic number, degree, formal charge, hydrogen count, hybridization, aromaticity, and ring membership, while each bond is encoded as a 5-dimensional one-hot vector indicating bond type or self-loop. Node features are projected into a 256-dimensional hidden space via a learned linear transformation, and edge features are incorporated directly at the attention computation step. The backbone consists of 3 stacked graph attention layers, each with 4 multi-head attention heads, residual connections, LayerNorm, ELU activation, and dropout (*p* = 0.1). After the final attention layer, node embeddings are mean-pooled to produce a single graph-level representation, which is passed through a two-layer classification head with sigmoid activation to yield the predicted probability of the input molecule being a polyketide synthase (PKS) product. |

We chose a graph neural network (GNN) architecture for polyketide classification because molecular graphs provide a natural and expressive representation of chemical structure that preserves atom connectivity, bond topology, and local chemical environments without requiring fixed-length fingerprint encodings that can lose structural information through hash collisions. Specifically, we adopted a graph attention network (GAT) backbone, which computes learned attention coefficients over each atom’s local neighborhood during message passing, thereby enabling the model to selectively weight the contributions of different neighboring atoms and bonds when constructing node-level representations. This attention-based mechanism is particularly well suited to polyketide classification because the diagnostic structural features of PKS products—such as the alternating pattern of carbonyl, hydroxyl, and methylene groups along the backbone—are inherently local motifs whose importance varies depending on the surrounding chemical context. The full architecture, including atom and bond featurization details, the number of attention layers and heads, and the classification head design, is described in Figure 4. Importantly, edge features encoding bond type information are incorporated directly into the attention coefficient computation rather than being concatenated to node features, allowing the model to modulate attention weights based on bond characteristics such as single versus double bond connectivity. After three rounds of neighborhood aggregation via the stacked GAT layers, the resulting node embeddings are mean-pooled into a single graph-level representation, which captures a global summary of the molecule’s structural features. This pooled representation is then passed through a two-layer classification head with a sigmoid output to produce the predicted probability that the input molecule is a PKS product. Residual connections around each attention layer ensure stable gradient flow during training and allow the network to retain low-level atomic features alongside the higher-order structural patterns learned by deeper layers.

The complete dataset of approximately three million molecules was partitioned into training, validation, and test splits using an 80/10/10 ratio with stratified sampling to preserve the 2:1 class imbalance across all splits. The model is trained with binary cross-entropy loss (pos\_weight = 2.0 to compensate for the 2:1 class imbalance of negatives to positives) using AdamW optimization (learning rate = 3×10–4, weight decay = 1×10–4) with gradient clipping at norm 1.0. An optional cosine annealing learning rate schedule with linear warmup over the first 5% of training epochs is supported to stabilize early optimization and smoothly decay the learning rate thereafter. We selected area under the precision-recall curve (AUPRC) as the primary evaluation metric rather than area under the receiver operating characteristic (AUROC) because AUPRC is more informative for imbalanced classification tasks where the positive class (polyketides) constitutes only one-third of the dataset, and because high precision is critical in the TridentSynthRL deployment context: a false positive polyketide prediction triggers a computationally expensive PKS MCTS agent, whereas a false negative merely results in the intermediate receiving its default SA score without incurring additional search cost. The model supports distributed multi-GPU training via PyTorch DistributedDataParallel, which enabled us to train on the full three-million-molecule dataset within a practical time frame.

During MCTS deployment, the trained GNN classifier is integrated into the chemoenzymatic agent’s expansion step as the primary polyketide detection mechanism, replacing or supplementing the static database lookup. Specifically, when a new child node is generated during chemoenzymatic expansion, its canonical SMILES string is passed through the GNN in a single forward pass; if the predicted probability exceeds a threshold of 0.5, the node is flagged as a candidate polyketide and the specialized RetroTide PKS MCTS agent is spawned to evaluate whether a chimeric PKS assembly line can be designed to synthesize the flagged intermediate. GNN inference on individual molecules is computationally inexpensive relative to the cost of running the PKS MCTS agent itself, adding negligible overhead to the expansion step. Users retain the option of using the static database lookup for deterministic matching against the precomputed library of approximately 106,000 known PKS products, a mode that is particularly useful for validation, benchmarking, and reproducing known biosynthetic routes. However, the GNN-based approach offers a critical advantage in that it generalizes beyond the boundaries of any finite enumerated library to recognize novel polyketide scaffolds on the basis of learned structural features, a capability that becomes increasingly important as TridentSynthRL is applied to diverse molecular targets whose retrosynthetic intermediates may fall outside previously characterized polyketide chemical space. In practice, we found that the GNN and the static database lookup produce concordant predictions for known polyketide structures while the GNN additionally flags structurally plausible polyketide candidates that are absent from the precomputed library, thereby expanding the set of intermediates for which PKS-based routes can be explored.

## **Deploying TridentSynthRL on the Kavalactone Family**

We deployed TridentSynthRL on the family of kavalactones from *Piper methysticum* (kava), a collection of 15 structurally related natural products whose anxiolytic, analgesic, and neuroprotective properties have made them subjects of considerable pharmacological interest (Figure 5). Kavalactones have historically been extracted from the roots and rhizomes of the kava plant, but the low and variable concentrations obtained through plant extraction, combined with growing demand for standardized pharmaceutical-grade material, have motivated efforts to develop alternative production routes. Traditional total synthesis of kavalactones has proven challenging due to the need to install the α-pyrone lactone ring with the correct substitution pattern and degree of unsaturation, while biological production routes remain underexplored. The kavalactone family is therefore an ideal test case for TridentSynthRL for several reasons. First, all 15 members share a core α-pyrone (six-membered lactone) scaffold that is a natural candidate for construction via PKS-mediated chain elongation followed by thioesterase-catalyzed lactonization, directly mirroring the biosynthetic logic of type I polyketide synthases. Second, the family members differ systematically in their aromatic substitutions (phenyl versus methoxyphenyl versus methylenedioxyphenyl), degree of backbone saturation (unsaturated, partially saturated, and fully saturated variants), and patterns of hydroxylation and methoxylation, collectively requiring the full diversity of chemoenzymatic and synthetic chemistry transformations available within TridentSynthRL’s reaction template library. Third, the existence of both simple (e.g., desmethoxyyangonin) and complex (e.g., methysticin) family members allows us to evaluate how the framework distributes synthetic burden across its three modalities as target complexity increases.

|  |
| --- |
| A black background with a black square  AI-generated content may be incorrect. |
| **Figure 5.** Chemical structures of the kavalactone family from *Piper methysticum* (kava) for which TridentSynthRL was tested against. The six major kavalactones, i.e., kavain, 7,8-dihydrokavain, yangonin, Desmethoxyyangonin, Methysticin, and 7,8-dihydromethysticin are shown, along with their related structural analogs. These compounds share a core -pyrone scaffold and vary in their aromatic substitutions, degrees of saturation, and hydroxylation/ methoxylation patterns, all of which contribute to their various pharmacological effects, including their anxiolytic, analgesic, and neuroprotective properties. TridentSynthRL was able to find thermodynamically feasible pathways to all members of the kavalactone family with at least one pathway for each structure involving a type I PKS. Users also have the option to tune the total number of pathways enumerated by running TridentSynthRL for fewer or more iterations. |

TridentSynthRL successfully identified thermodynamically feasible retrosynthetic pathways to all 15 members of the kavalactone family, achieving complete coverage of the target set. For every compound, at least one pathway involves a type I PKS to construct the core carbon backbone, confirming that the hierarchical multi-agent architecture can consistently identify PKS-compatible intermediates within the retrosynthetic search tree. Across the 15 targets, the framework discovered a diverse portfolio of route architectures: some pathways rely predominantly on PKS biosynthesis with minimal post-PKS modifications, others distribute the synthetic burden across all three modalities by coupling short PKS assembly lines with chemical coupling reactions and enzymatic tailoring steps, and still others proceed entirely through synthetic chemistry when the MCTS search determines that a chemical route is thermodynamically and kinetically more favorable than a biological alternative. This diversity of route types emerged naturally from the reward-guided search without any manual bias toward a particular modality, underscoring the value of integrating all three synthesis routes within a single planning framework. Complete pathway details for all 15 kavalactones, including full module architectures, reaction enthalpies, and DORA-XGB feasibility scores, are provided in the Supporting Information. Below we highlight three representative pathways that illustrate different strategic strengths of TridentSynthRL and the complementary roles of PKS biosynthesis, monofunctional enzymatic catalysis, and synthetic chemistry in accessing the kavalactone family.

**PKS-dominated route: 7,8-dihydrokavain (Figure 6).** As shown in Figure 6, TridentSynthRL identifies a PKS-dominated pathway in which the majority of the molecular scaffold is assembled by a chimeric type I PKS, with only two subsequent post-PKS transformations required to arrive at the final target. The selection of benzoyl-CoA as the starter unit is notable because it directly installs the aromatic phenyl ring that constitutes 7,8-dihydrokavain’s western fragment, eliminating the need for any subsequent aromatic functionalization chemistry. By delegating the backbone construction to the PKS and reserving only the dehydration and *O*-methylation steps for post-PKS modification, TridentSynthRL minimizes the total number of reaction steps while ensuring that each transformation is both thermodynamically favorable and enzymatically feasible. Both post-PKS steps received high feasibility assessments: the dehydration is enthalpically downhill, and the *O*-methylation using *S*-adenosylmethionine (SAM) as the methyl donor received a high DORA-XGB feasibility score, reflecting the well-characterized nature of SAM-dependent methyltransferases and their broad substrate tolerance toward hydroxylated aromatic and aliphatic scaffolds. This route exemplifies a general design paradigm in which the PKS handles the stereochemically demanding construction of the carbon backbone—a task for which the iterative Claisen condensation and tailoring domain machinery of type I PKSs is uniquely suited—while monofunctional enzymes perform the final peripheral modifications that fine-tune the molecule’s functional group pattern. Such a division of labor leverages the complementary strengths of each modality and represents a strategy that TridentSynthRL is well positioned to identify across diverse target families.

|  |
| --- |
| A diagram of different colored circles  AI-generated content may be incorrect. |
| **Figure 6.** TridentSynthRL successfully proposes a pathway to the kavalactone 7,8-dihydrokavain via type I PKSs, monofunctional enzymes, and synthetic organic chemistry. In the select pathway shown, a chimeric PKS with a benzoyl-CoA starter unit is suggested, along with 3 extension modules, each of which incorporate malonyl-CoA as extender units and module-specific tailoring domains (DH, KR, ER). Following a thioesterase (TE)-mediated lactonization offloading reaction, the core -pyrone scaffold for 7,8-dihydrokavain is produced. With this scaffold, an enthalpically downhill (-6.27 kcal/mol) dehydration reaction is subsequently suggested, followed by an enzymatic *O-*methylation reaction. This *O-*methylation reaction is predicted to be enzymatically feasible with a feasibility score of 0.955, as assigned by our previously released supervised reaction feasibility classifier, DORA-XGB. |

**Hybrid PKS + chemical coupling route: desmethoxyyangonin (Figure 7).** In contrast to the PKS-dominated route for 7,8-dihydrokavain, the pathway proposed for desmethoxyyangonin adopts a fundamentally different strategy by distributing the synthetic burden across all three modalities (Figure 7). Rather than using a longer PKS assembly line to construct the entire α-pyrone scaffold, TridentSynthRL identifies that a minimal PKS with only a single extension module is sufficient to generate a polyketide fragment corresponding to the aromatic styryl moiety, while the six-membered lactone ring is instead sourced from a commercially available chemical building block. This convergent strategy—in which the PKS and commercial chemistry each contribute a distinct molecular fragment that are then joined via a synthetic coupling reaction—represents a qualitatively different route architecture from the linear, PKS-first approach seen in the 7,8-dihydrokavain example. The coupling reaction itself is an electrophilic aromatic alkylation that is strongly enthalpically favorable, and the three subsequent enzymatic tailoring steps (*O*-methylation, decarboxylation, and dehydration) all received high DORA-XGB feasibility scores, indicating that each enzymatic transformation is individually well supported by precedent in characterized enzyme families. From an experimental implementation standpoint, this route is attractive because it reduces the PKS engineering challenge to a single-module chimera—the shortest possible extension beyond the loading module—which substantially mitigates the protein misfolding and low titer issues that frequently plague longer chimeric PKS constructs expressed in heterologous hosts. At the same time, the use of a commercially available lactone building block eliminates the need for the PKS to perform the macrolactonization step, which can be a bottleneck for thioesterase-mediated ring closure at certain ring sizes. This route thus illustrates TridentSynthRL’s capacity to identify convergent, modality-spanning strategies that minimize the engineering complexity of each individual step while still achieving the overall synthetic goal.

|  |
| --- |
| A screenshot of a computer  AI-generated content may be incorrect. |
| **Hybrid PKS + chemical coupling route: desmethoxyyangonin (Figure 7).** A shorter chimeric PKS with only one extension module (benzoyl-CoA starter, hydroxymalonyl-CoA extender) produces a polyketide fragment that accounts for the aromatic styryl moiety. TridentSynthRL then proposes coupling this fragment with the commercially available chemical building block 4-hydroxy-2*H*-pyran-2-one—which provides the six-membered lactone ring—via an enthalpically downhill electrophilic aromatic alkylation (Δ*H* = −11.38 kcal/mol). Three subsequent enzymatic transformations complete the synthesis: *O*-methylation (DORA-XGB = 0.977), decarboxylation (DORA-XGB = 0.995), and dehydration (DORA-XGB = 0.964). This route illustrates TridentSynthRL’s capacity to combine a minimal PKS with chemical coupling and enzymatic tailoring, distributing the synthetic burden across all three modalities. |

**Purely chemical route: 5,6,7,8-tetrahydroyangonin (Figure 8).** While the preceding two examples demonstrate TridentSynthRL’s ability to integrate biological and chemical synthesis, the pathway proposed for 5,6,7,8-tetrahydroyangonin reveals an equally important capability: the framework’s willingness to forgo biological routes entirely when a purely synthetic chemistry approach is more efficient (Figure 8). In this case, the fully saturated backbone of 5,6,7,8-tetrahydroyangonin means that the characteristic β-oxygenation pattern produced by PKS tailoring domains (KR, DH, ER) would need to be fully reduced at every position, effectively negating the stereochemical advantages that PKS biosynthesis typically offers for unsaturated or partially reduced polyketide scaffolds. TridentSynthRL’s MCTS search recognized this implicitly through the reward structure: the purely chemical route, which proceeds from commercially available building blocks through a sequence of carbon–carbon bond-forming reactions and functional group manipulations to construct the lactone ring, accumulated higher cumulative rewards than competing PKS-containing alternatives. This outcome is a direct consequence of the unbiased reward design described earlier, in which the geometric mean of per-step feasibility scores allows the algorithm to objectively compare routes of different lengths and modality compositions without systematically favoring biological or chemical transformations. The discovery of this purely chemical pathway is significant from a practical standpoint as well, since it eliminates the need for any biological engineering infrastructure—no PKS expression, no enzyme sourcing, no fermentation optimization—and instead relies entirely on well-established synthetic organic transformations that can be performed in a conventional chemistry laboratory. Together with the PKS-dominated and hybrid routes described above, this example demonstrates that TridentSynthRL does not impose a preferred synthesis modality but rather allows the MCTS search to discover the most efficient pathway architecture for each individual target, adapting its strategy to the specific structural features and synthetic accessibility of the molecule at hand.

|  |
| --- |
| A black screen with text and symbols  AI-generated content may be incorrect. |
| **Figure 8:** Purely chemical retrosynthetic route to 5,6,7,8-tetrahydroyangonin proposed by TridentSynthRL. Starting from two commercially available building blocks — 1-(but-3-yn-1-yl)-4-methoxybenzene and 3-bromoprop-2-ynoic acid — a Suzuki coupling with alkynes forms a methoxyphenyl-substituted bromoacrylic acid intermediate. Subsequent addition of HBr to the alkyne moiety yields a vinyl bromide bearing the full carbon skeleton. Intermolecular esterification with methanol, a third chemical building block, followed by intramolecular addition of the alcohol to the pendant alkene generates the six-membered lactone ring. A final dehydrohalogenation eliminates HBr to furnish 5,6,7,8-tetrahydroyangonin. |

**A shorter pathway to cryptofolione via chemoenzymatic retrosynthesis**

Beyond the kavalactone family, we applied TridentSynthRL to the structurally complex natural product cryptofolione to evaluate whether the integration of synthetic chemistry into a previously enzyme-only retrosynthetic strategy could yield more experimentally tractable pathways. Cryptofolione is a polyketide-derived natural product featuring a six-membered α-pyrone lactone ring fused to an extended aliphatic chain bearing two hydroxyl groups and a styryl moiety, making it a challenging target that spans both the biosynthetic and synthetic chemical domains. In our prior work using RetroTide alone, we proposed a chimeric PKS assembly line spanning five extension modules to construct the full cryptofolione scaffold in a single biosynthetic step. While this design was conceptually valid and demonstrated RetroTide’s capacity for forward PKS design, such long chimeric assembly lines present significant experimental challenges: heterologous expression of engineered PKSs with five or more extension modules in microbial hosts such as *Escherichia coli* or *Streptomyces* species is prone to low expression levels, protein misfolding, and reduced catalytic efficiency as module count increases, all of which diminish the practical feasibility of the proposed route.

TridentSynthRL identifies a fundamentally different and more efficient pathway to cryptofolione by leveraging its ability to reason across all three synthesis modalities simultaneously (Figure 9). Rather than relying on a single long PKS to build the entire scaffold, TridentSynthRL proposes a chimeric PKS with only two extension modules—a cinnamoyl-CoA starter unit paired with two malonyl-CoA extenders, both extension modules equipped with ketoreductase (KR) tailoring domains—to produce a polyketide fragment that captures the aromatic styryl moiety and the six-membered lactone ring of cryptofolione via thioesterase-mediated lactonization. This fragment accounts for the structurally demanding portion of the molecule, including the aromatic ring and the α-pyrone core, while the remaining aliphatic diol-containing chain is introduced through a subsequent synthetic chemistry step: an olefin ring-opening metathesis (ROM) reaction that couples the PKS-derived fragment with cyclohex-3-ene-1,4-diol. The ring-opening metathesis is a well-established transformation in synthetic organic chemistry and, in this case, proceeds in an enthalpically downhill manner. Cyclohex-3-ene-1,4-diol is in turn obtained from an enzymatic monooxygenation of the commercially available building block cyclohex-3-en-1-ol, a transformation that received a DORA-XGB feasibility score of 0.931, indicating high predicted enzymatic feasibility.

The reduction from five to two extension modules represents a substantial improvement in experimental tractability. Shorter chimeric PKS assembly lines are significantly easier to express, fold, and maintain catalytic activity in heterologous hosts, and the two-module design proposed here falls within the range of chimeric PKS constructs that have been successfully expressed in prior experimental studies. The trade-off introduced by this shorter PKS—namely, the addition of a ring-opening metathesis coupling step and an enzymatic monooxygenation—is modest relative to the experimental benefits gained: both the metathesis reaction and the enzymatic hydroxylation are well-precedented transformations with established protocols, and the overall pathway remains thermodynamically favorable. This result exemplifies the central advantage of TridentSynthRL’s multi-modal approach: by allowing the MCTS search to distribute synthetic burden across PKS biosynthesis, monofunctional enzymatic catalysis, and synthetic chemistry simultaneously, the framework can identify globally more efficient pathway architectures that would be inaccessible to any single-modality planner. Where RetroTide alone was constrained to propose an entirely biosynthetic solution requiring a long and experimentally challenging PKS, TridentSynthRL discovers a hybrid strategy that reduces the biological engineering burden while maintaining a feasible and thermodynamically sound overall route.

# **Materials and methods**

### **Chemoenzymatic MCTS agent**

The chemoenzymatic MCTS agent implements a three-phase iterative search (selection, expansion, update) over a retrosynthetic tree rooted at the target molecule. Each node in the tree stores an RDKit Mol object representing a molecular intermediate, along with MCTS statistics (visit count, cumulative reward, selection score) and reaction metadata (SMARTS pattern, reaction name, reactant and product SMILES, provenance as enzymatic or synthetic). The agent applies reaction templates from the DORAnet reaction network (available at https://github.com/wsprague-nu/doranet), which comprises 386 synthetic organic chemistry templates (e.g., esterifications, Suzuki couplings, olefin metathesis, dehydrohalogenations) and 3,604 enzymatic templates (e.g., alcohol dehydrogenases, aminotransferases, decarboxylases, methyltransferases, monooxygenases). All templates are encoded as SMARTS patterns and are applied in the retrosynthetic direction during expansion. Prior to running the search, target molecules are preprocessed by removing stereochemistry information (since DORAnet templates do not preserve stereochemistry), sanitizing the molecule, converting to canonical SMILES, and computing the target molecular weight for downstream fragment filtering.

During the selection phase, the agent traverses the tree from the root using the upper confidence bound 1 (UCB1) policy. For a node s with child s’, the UCB1 score is computed as: Q(s’)/N(s’) + c × sqrt(ln(N(s))/N(s’)), where Q(s’) is the cumulative reward, N(s’) is the visit count of the child, N(s) is the visit count of the parent, and c = sqrt(2) is the exploration constant. Unvisited nodes receive an infinite selection score to ensure breadth-first exploration at each level. We also implemented an optional depth-biased variant that adds a depth bonus term β × d to the UCB1 score, where β is a configurable coefficient (default 2.0) and d is the node depth. For unvisited nodes under the depth-biased policy, the selection score is set to 1000.0 + β × d, allowing depth to differentiate among unvisited candidates and encouraging deeper exploration before exhaustive breadth search. To prevent wasted iterations when all children at a given level are terminal, a frontier fallback mechanism maintains a max-heap of unexpanded non-terminal nodes ordered by depth (deepest first); when standard traversal reaches an all-terminal branch, the agent falls back to the deepest unexpanded node in the frontier.

During expansion, the selected leaf node is expanded by applying all synthetic and enzymatic reaction templates via DORAnet in the retrosynthetic direction, generating a set of upstream precursor fragments. These fragments undergo post-processing and filtering: wildcard-containing SMILES (template artifacts) are removed; duplicate fragments (by canonical SMILES) are eliminated; self-loops (fragments identical to the input) are discarded; known cofactors and small molecules (e.g., H2O, CO2, NAD+, NADPH, SAM, and other biology cofactors and chemistry helpers) are excluded; fragments exceeding 1.5 times the target molecular weight are filtered to prevent unrealistic dimerization products; and prohibited or hazardous chemicals (652 compounds) are pruned. When the number of valid fragments exceeds the maximum children per expansion (default 5), a child downselection strategy is applied. The default strategy, “most\_thermo\_feasible,” ranks fragments by their thermodynamic feasibility score with priority bonuses for sink compounds (+1000) and PKS library matches (+500), ensuring that terminal and near-terminal fragments are preferentially retained.

A node is marked terminal when it matches one of three building block databases: (1) the Enamine commercial chemical catalog comprising approximately 278,000 compounds; (2) a curated set of 334 biological metabolites drawn from BRENDA, KEGG, and MetaCyc across multiple model microorganisms; or (3) a polyketide intermediate verified by the RetroTide PKS agent. Building block databases are stored as canonical SMILES sets and loaded with pickle caching (validated by content hash) for efficient repeated access. Fragment expansion results are also cached to disk using a pickle-based system keyed by the MD5 hash of expansion parameters, enabling deterministic replay and avoiding redundant DORAnet computations across MCTS iterations.

The reward policy employs a modular, composable architecture. The default configuration uses a terminal reward of +1.0 for nodes matching commercial building blocks, biological building blocks, or verified polyketide intermediates. For non-terminal intermediates, a dense reward signal is provided via the RDKit synthetic accessibility (SA) score, linearly transformed as: reward = (10 − SA) / 10, yielding rewards in the range [0.0, 0.9] where lower SA scores (more synthetically accessible) produce higher rewards. Path-level feasibility is incorporated through a thermodynamic scaling wrapper that computes the geometric mean of per-step feasibility scores along the pathway from root to leaf. For enzymatic reactions, feasibility is assessed using our previously published DORA-XGB gradient-boosted tree classifier, which predicts reaction feasibility on a 0–1 probability scale by operating on forward-direction reaction SMILES (retro-reactions are reversed before scoring). For synthetic chemistry transformations, reaction enthalpies (ΔH) are computed using PAthermo, a Benson group additivity method, and converted to a 0–1 feasibility score via sigmoid transformation: score = 1 / (1 + exp(0.2 × (ΔH − 15))), where the threshold of 15 kcal/mol and steepness of 0.2 were chosen to smoothly penalize highly endothermic steps while preserving favorable or near-thermoneutral reactions. The final reward for a node is computed as: reward = base\_reward × ((1 − w) + w × pathway\_feasibility), where w = 0.8 is the default feasibility weight, providing soft thermodynamic biasing that downweights unfavorable pathways without completely zeroing their rewards.

### To accelerate search, the chemoenzymatic agent is parallelized using a virtual-loss–based MCTS implementation. Specifically, in the parallel variant, DORAnet expansion is dispatched to worker processes via a ProcessPoolExecutor, while selection, reward evaluation, and backpropagation remain on the main process. To prevent multiple workers from selecting the same node, a virtual loss is applied at expansion launch (+1 visit, −1.0 value), temporarily reducing that node’s attractiveness; the virtual loss is removed once the expansion completes and the true reward is backpropagated. The number of parallel workers and maximum in-flight expansions are configurable, and SMILES canonicalization is cached with a 50,000-entry LRU cache to avoid redundant RDKit calls.

### **Polyketide synthase MCTS agent**

The PKS MCTS agent operates in the forward-synthesis direction to evaluate whether a chimeric type I PKS assembly line can produce a given target polyketide fragment from acyl-CoA building blocks. This agent is built upon our previously released PKS design tool RetroTide (available at https://github.com/JBEI/RetroTide). Each node in the PKS search tree represents a complete PKS chimera design at a given stage of assembly, encoding: the chemical structure of the polyketide product bound to the synthase via a thioester linkage to the acyl carrier protein (ACP), the ordered sequence of catalytic domains selected thus far, and the current depth (number of extension modules appended). The root state corresponds to an empty PKS design with no modules, from which the search builds outward by iteratively adding extension modules.

The action space comprises the set of possible extension modules, where each module is defined by the choice of acyl-CoA extender unit (6 options, including malonyl-CoA, methylmalonyl-CoA, ethylmalonyl-CoA, methoxymalonyl-CoA, hydroxymalonyl-CoA, and aminomalonyl-CoA) and the combination of optional tailoring domains. Each extension module minimally contains ketosynthase (KS), acyltransferase (AT), and acyl carrier protein (ACP) domains. Optional tailoring domains include: ketoreductase (KR), which reduces the β-keto group to a hydroxyl; dehydratase (DH), which eliminates water to form an enoyl intermediate; and enoylreductase (ER), which fully saturates the resulting double bond. This gives rise to four possible domain architectures per module: KS-AT-ACP (no reduction), KS-AT-KR-ACP (partial reduction), KS-AT-KR-DH-ACP (partial reduction with dehydration), and KS-AT-KR-DH-ER-ACP (full reduction). The loading module additionally specifies one of 29 possible starter units (e.g., acetyl-CoA, benzoyl-CoA, cinnamoyl-CoA, propionyl-CoA). At each expansion step, RetroTide enumerates feasible one-step extensions given the current PKS architecture, returning up to 25 candidate designs ranked by maximum common substructure (MCS) similarity to the target fragment.

A key innovation of the PKS agent is subgraph-guided selection. Before the search begins, a “bag of graphs” is constructed by extracting all submolecules from the target fragment at multiple radii. For lactone targets, the target is additionally ring-opened via a simulated thioesterase reaction to enable matching of linear intermediates against cyclic targets. During selection, unvisited child nodes are checked for subgraph isomorphism (without stereochemistry) against this reference set using RDKit’s substructure matching routines. Children whose PKS products are confirmed subgraphs of the target receive infinite selection scores (forcing immediate exploration), while children failing the subgraph test receive negative-infinity scores (effective pruning). For previously visited children, standard UCB1 scoring balances exploitation and exploration.

Following expansion, each child node is immediately evaluated by simulating product release through two thioesterase (TE)-catalyzed mechanisms: thiolysis, which cleaves the thioester bond to yield a linear carboxylic acid, and lactonization, which performs intramolecular cyclization to generate a lactone. If the released product from either mechanism is graph-isomorphic to the target fragment, the node is assigned a terminal reward of +1.0. Otherwise, a partial reward is computed as the MCS similarity between the released product and the target, defined as the ratio of atoms in the MCS to the total atom count of the union of query and reference molecules. The search terminates when the maximum depth is reached (default: 15 extension modules), an exact target match is identified, or no valid child nodes remain after subgraph pruning. Upon termination, if a successful design is found, a terminal reward of +1.0 is returned to the parent chemoenzymatic agent; otherwise, the intermediate receives its default SA score.

### **Supervised GNN classifier for polyketide detection**

The training set was generated through a two-step synthetic data pipeline. In the first step, RetroTide combinatorially enumerated chimeric type I PKS assembly line configurations comprising a single loading module (29 possible starter units), up to 3 extension modules (each with all four possible domain architectures crossed with 6 extender units), and a terminal thioesterase domain allowing both thiolysis and lactonization offloading. This enumeration yielded approximately one million core polyketide structures, each labelled as class 1 (polyketide). In the second step, each polyketide was subjected to a single step of DORAnet chemoenzymatic expansion using all 3,604 enzymatic and 386 synthetic reaction templates. For each parent polyketide, the most structurally similar product from the enzymatic modification set and the most structurally similar product from the synthetic modification set, as ranked by Tanimoto similarity on Morgan fingerprints, were selected as hard negatives, yielding approximately two million modified variants labelled as class 0 (non-polyketide). The complete dataset was partitioned into training, validation, and test splits using an 80/10/10 ratio with stratified sampling.

Input SMILES strings are parsed via RDKit into molecular graphs. Each atom is encoded as a 40-dimensional feature vector comprising: atomic number (13-dimensional one-hot with unknown bucket), degree (7 dimensions), formal charge (6 dimensions), hydrogen count (6 dimensions), hybridization (6 dimensions), aromaticity (1 binary), and ring membership (1 binary). Each bond is encoded as a 5-dimensional vector: bond type (4-dimensional one-hot for single, double, triple, and aromatic bonds) plus a self-loop indicator (1 binary). Self-loops are added for all atoms to enable self-message passing.

The GNN architecture consists of an input projection layer (40 → 256 dimensions), followed by 3 stacked graph attention layers, each with 4 multi-head attention heads. Each attention layer applies multi-head attention with edge feature projection (edge features are incorporated directly into the attention coefficient computation rather than being concatenated to node features), followed by LayerNorm, ELU activation, dropout (p = 0.1), and a residual connection. After the final attention layer, node embeddings are mean-pooled to produce a single 256-dimensional graph-level representation, which is passed through a two-layer classification head (256 → 128 with ReLU activation, then 128 → 1 with sigmoid activation) to yield the predicted polyketide probability.

The model is trained using binary cross-entropy loss with pos\_weight = 2.0 to compensate for the 2:1 class imbalance of negatives to positives. Optimization uses AdamW (learning rate = 3 × 10⁻⁴, weight decay = 1 × 10⁻⁴) with gradient clipping at norm 1.0. An optional cosine annealing learning rate schedule with linear warmup over the first 5% of training epochs is supported. The primary evaluation metric is area under the precision–recall curve (AUPRC), chosen over AUROC because of its greater sensitivity to performance on the minority positive class and because high precision is critical in the deployment context: a false positive polyketide prediction triggers a computationally expensive PKS MCTS agent, whereas a false negative merely results in the intermediate receiving its default SA score. Distributed multi-GPU training is supported via PyTorch DistributedDataParallel. During MCTS deployment, the trained GNN serves as the primary polyketide detection mechanism: when a new child node is generated during chemoenzymatic expansion, its canonical SMILES is passed through the GNN; if the predicted probability exceeds 0.5, the RetroTide PKS agent is spawned. Users retain the option of using database lookup against the precomputed library of approximately 106,000 known PKS products for deterministic matching.

### **Reaction feasibility scoring**

Enzymatic reaction feasibility is evaluated using DORA-XGB, a previously published gradient-boosted tree classifier trained on curated enzymatic reaction databases. DORA-XGB accepts forward-direction reaction SMILES as input and returns a probability score on a 0–1 scale along with a binary feasibility label (feasible if score ≥ 0.5). Since DORAnet stores reactions in the retrosynthetic direction, each reaction is reversed to the forward direction prior to scoring. DORA-XGB is applied exclusively to enzymatic reactions; synthetic reactions are scored thermodynamically. Cofactor positioning in the reaction SMILES follows a descending molecular weight convention for consistency with the classifier’s training data.

Thermodynamic feasibility for all reactions (both enzymatic and synthetic) is assessed using PAthermo, an open-source Benson group additivity calculator (available at https://github.com/dmdqy/pathermo). PAthermo computes enthalpies of formation (ΔHf) for each molecule in a reaction, from which the reaction enthalpy is calculated as ΔH = ∑(ΔHf of products) − ∑(ΔHf of reactants) in the forward direction. A reaction is labelled thermodynamically feasible if ΔH < 15 kcal/mol. To unify enzymatic and thermodynamic scores on a common 0–1 scale, reaction enthalpies are converted via a sigmoid transformation: score = 1 / (1 + exp(0.2 × (ΔH − 15))), which maps ΔH = −20 kcal/mol to a score of approximately 0.999, ΔH = 0 to approximately 0.953, ΔH = 15 to exactly 0.500, and ΔH = 30 to approximately 0.047. For enzymatic reactions where DORA-XGB scores are available, the DORA-XGB probability is used as the feasibility score; the sigmoid-transformed enthalpy serves as a fallback when DORA-XGB scores are unavailable.

Pathway-level feasibility is aggregated using the geometric mean of per-step feasibility scores along the route from root to leaf node. The geometric mean was chosen because it normalizes for path length, allowing fair comparison of routes with different numbers of steps without systematically penalizing longer but otherwise plausible sequences. The aggregated pathway feasibility is then used to scale the base reward via: final\_reward = base\_reward × ((1 − w) + w × pathway\_feasibility), where w is the feasibility weight (default 0.8). This soft scaling formulation ensures that thermodynamically unfavorable pathways receive reduced but non-zero rewards, allowing the MCTS search to explore them if no better alternatives exist while still preferentially expanding thermodynamically favorable routes.

### **Default MCTS parameters for reported experiments**

For the kavalactone experiments, the chemoenzymatic MCTS agent was configured with the following default parameters: 100 total iterations, a maximum tree depth of 4, a maximum of 5 children per expansion with the “most\_thermo\_feasible” downselection strategy, 1 DORAnet generation per expansion step, and the depth-biased UCB1 selection policy with a depth bonus coefficient of 4.0. Both enzymatic and synthetic reaction modes were enabled, and all three building block databases (chemical, biological, and PKS) were active. The rollout policy used SpawnRetroTideOnDatabaseCheck with a success reward of 1.0 and failure reward of 0.0. The reward policy used SAScore\_and\_TerminalRewardPolicy (terminal reward = 1.0, SA score for non-terminals) wrapped by ThermodynamicScaledRewardPolicy with feasibility weight = 0.8, sigmoid steepness k = 0.2, and sigmoid threshold = 15.0 kcal/mol. The molecular weight filter excluded fragments exceeding 1.5 times the target molecular weight. For the PKS sub-agent, default parameters were: maximum depth of 5 extension modules, 50 MCTS iterations, and up to 500 tracked PKS designs per RetroTide call.

**Discussion**

Here, we have presented TridentSynthRL, which to the best of our knowledge is the first retrosynthetic planning framework to unify synthetic organic chemistry, monofunctional enzymatic catalysis, and type I polyketide synthase biosynthesis within a single hierarchical Monte Carlo tree search. By formulating chemoenzymatic retrosynthesis and chimeric PKS design as separate but coupled Markov decision processes, TridentSynthRL introduces a two-agent architecture in which a chemoenzymatic retrosynthetic agent spawns forward-synthesis PKS agents upon detecting polyketide intermediates, creating a bidirectional feedback loop between retrosynthetic fragmentation and forward biosynthetic verification. A supervised graph neural network trained on approximately one million synthetically generated polyketides and two million structurally similar hard negatives serves as the polyketide detection mechanism, enabling generalization beyond any finite enumerated library to novel polyketide scaffolds encountered during search. Thermodynamic and enzymatic feasibility scoring—via PAthermo group additivity estimates and the DORA-XGB reaction feasibility classifier, respectively—are integrated directly into the reward policy to bias the search toward experimentally executable pathways.

The deployment of TridentSynthRL on the kavalactone family from *Piper methysticum* demonstrates the framework’s capacity to discover diverse route architectures tailored to individual target structures. Across all 15 kavalactone family members, TridentSynthRL identified thermodynamically feasible retrosynthetic pathways, with each compound accessible via at least one PKS-involving route. The three representative pathways highlighted in this work—a PKS-dominated route to 7,8-dihydrokavain, a hybrid PKS-chemical coupling route to desmethoxyyangonin, and a purely chemical route to 5,6,7,8-tetrahydroyangonin—illustrate how the MCTS search naturally distributes synthetic burden across modalities without manual bias, selecting the most efficient combination for each target’s structural features. For cryptofolione, TridentSynthRL identified a pathway using a two-module PKS coupled with ring-opening metathesis and enzymatic monooxygenation, reducing the required PKS modules from five (in our prior RetroTide-only design) to two. This reduction substantially improves the likelihood of successful heterologous expression, as shorter chimeric PKS assembly lines are more readily expressed and folded in microbial hosts.

Several features of TridentSynthRL’s design merit further discussion. The omission of the simulation (rollout) phase from the classical MCTS algorithm was motivated by the high computational cost and noise associated with random rollouts over long retrosynthetic horizons. Instead, the combination of heuristic feasibility scores (SA score, DORA-XGB, thermodynamic estimates) and hierarchical PKS verification provides immediate and informative feedback to the search without the overhead of multi-step random sampling. The modular policy architecture allows users to substitute or compose alternative reward functions, selection strategies, and feasibility metrics as needed for different target classes or experimental constraints. The thermodynamic scaling approach using geometric mean aggregation ensures that pathway-level feasibility comparisons are normalized for path length, avoiding systematic penalization of longer routes that may nonetheless be experimentally viable.

Several avenues for future development could further enhance TridentSynthRL’s capabilities. Expanding the reaction template library to include additional enzyme classes, such as type II and type III PKSs, non-ribosomal peptide synthetases (NRPSs), and engineered enzyme variants, would broaden the accessible chemical space. Incorporating learned value functions—for example, a graph neural network trained to predict node-level synthesis feasibility or pathway completion probability—could replace or augment UCB1 selection with more informed search guidance, potentially reducing the number of MCTS iterations required to discover high-quality pathways. Reaction condition compatibility checking, which would verify that sequential steps use compatible solvents, temperatures, pH ranges, and catalysts, would improve the practical executability of proposed routes. Extension to other natural product families beyond kavalactones and cryptofolione, as well as to commodity chemicals and pharmaceutical intermediates, would further test the generality and scalability of the framework. Finally, experimental validation of the proposed pathways—beginning with the shorter two-module PKS route to cryptofolione and the PKS-dominated route to 7,8-dihydrokavain—remains the critical next step toward establishing TridentSynthRL as a practical tool for synthesis planning at the interface of chemistry and biology.

# **Acknowledgements**

This work conducted at the Joint BioEnergy Institute was supported by the U.S. Department of Energy, Office of Science, Biological and Environmental Research Program, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. Any subjective views or opinions that might be expressed in this paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.. Funding for Y.C for this study was partly provided by the Northwestern University Graduate School Cluster in Biotechnology, Systems, and Synthetic Biology, which is affiliated with the Biotechnology training program. The funding for M.G.S was partly provided by the Northwestern University DeFeo Family Undergraduate Research Fellowship in Synthetic Biology and partly by the Northwestern University Summer Undergraduate Research Award. This research project was supported in part through the computational resources and staff contributions provided by the Quest high performance computing facility at Northwestern University, which is jointly supported by the Office of the Provost, the Office of Research, and Northwestern University Information Technology. This research also used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 using NERSC award BER-ERCAP0037563.

# **Author Contributions Statement**

**Y.C:** Methodology and Software, Data curation, Writing – original draft and subsequent edits. Visualization – generated the figures in this manuscript. **M.G.S:** Methodology and Software, **G.B:** Methodology and Software, **A.M:** Conceptualization – conceptualized the approach of this study. Investigation. Formal analysis. Writing – review & editing, were the principal investigators who directed this project, contributed to the data analysis and interpretation, as well as edited the manuscript. **J.D.K:** Conceptualization – conceptualized the approach of this study. Investigation. Formal analysis. Writing – review & editing, were the principal investigators who directed this project, contributed to the data analysis and interpretation, as well as edited the manuscript. **H.G.M:**  Conceptualization – conceptualized the approach of this study. Investigation. Formal analysis. Writing – review & editing, were the principal investigators who directed this project, contributed to the data analysis and interpretation, as well as edited the manuscript. **T.W.H.B:** Methodology and Software, Data curation, Conceptualization – conceptualized the approach of this study. Investigation. Formal analysis. Writing – review & editing, were the principal investigators who directed this project, contributed to the data analysis and interpretation, as well as edited the manuscript. **K.E.J.T:** Conceptualization – conceptualized the approach of this study. Investigation. Formal analysis. Writing – review & editing, were the principal investigators who directed this project, contributed to the data analysis and interpretation, as well as edited the manuscript. **L.J.B:** Conceptualization – conceptualized the approach of this study. Investigation. Formal analysis. Writing – review & editing, were the principal investigators who directed this project, contributed to the data analysis and interpretation, as well as edited the manuscript.

# 

# C**ompeting Interests Statement**

J.D.K. has financial interests in Amyris, Ansa Biotechnologies, Apertor Pharma, Berkeley Yeast, Cyklos Materials, Demetrix, Lygos, Napigen, ResVita Bio and Zero Acre Farms. The other authors declare no competing interests.

# 