Materials Today Communications 49 (2025) 114448

Contents lists available at ScienceDirect

Materials Today Communications

journal homepage: www.elsevier.com/locate/mtcomm

MOF-LENS: Bio-inspired lotus effect optimization for accelerated discovery of metal–organic framework nanocarriers for doxorubicin delivery in cancer therapy☆

Mehrdad Jalali [a,b,*], Binh Vu [b,c], Sina Mehraeen [b], Swati Chandna [b], Farzad Jalali [d]

[a] Institute of Functional Interfaces (IFG), Karlsruhe Institute of Technology (KIT), Germany
[b] Applied Data Science and Artificial Intelligence, SRH University Heidelberg, Heidelberg, Germany
[c] Faculty of Mathematics and Computer Science, University of Hagen, Hagen, Germany
[d] Department of Epidemiology, University Medical Center Groningen, the Netherlands

ARTICLE INFO

Keywords:
Metal–Organic Frameworks (MOFs)
Doxorubicin (DOX)
Drug delivery
Lotus Effect Algorithm (LEA)
Multi-objective optimization
Tanimoto similarity
Pore limiting diameter (PLD)
Materials informatics

ABSTRACT

Metal–organic frameworks (MOFs) are crystalline materials characterized by adjustable porosity and chemical properties, rendering them effective carriers for doxorubicin (DOX) in the treatment of hepatocellular carcinoma (HCC). The extensive design space complicates the identification of metal-organic frameworks (MOFs) that simultaneously optimize doxorubicin (DOX) loading, pH-responsive release, and biocompatibility. We introduce MOF-LENS (Latent Evolutionary Navigation System), a bio-inspired framework that integrates structural descriptors (pore size, surface area, void fraction) and SMILES-based fingerprints into a unified latent space. This framework employs the previously published Lotus Effect Algorithm (LEA) to identify high-quality MOF nanocarriers. In benchmarks comparing particle swarm optimization, genetic algorithms, random search, and a deterministic filter baseline across 30 independent runs, MOF-LENS achieves comparable or superior best fitness with significantly reduced variance, resulting in a consistent set of top candidates rather than isolated outliers. The chosen MOFs demonstrate pore-limiting diameters within the DOX-compatible range (12–16 Å) and exhibit a strong chemical affinity for DOX. Docking calculations validate substantial binding at physiological pH, with interactions slightly diminished under mildly acidic conditions, aligning with controlled release mechanisms. While illustrated for DOX in HCC, MOF-LENS is applicable to various drugs; retargeting to alternatives like paclitaxel necessitates merely the modification of the drug fingerprint, pore-size window, and a limited set of fitness weights.

1. Introduction

Metal–Organic Frameworks (MOFs) are crystalline, porous materials made up of metal nodes connected by organic linkers. They have very high surface areas (up to 7000 m^2/g), pore sizes that can be changed from 0.2 nm to 10 nm, and a wide range of chemical properties [1]. These properties position MOFs as transformative materials for applications including gas storage, catalysis, separation, environmental remediation, and biomedical drug delivery [2-4]. In cancer therapy, MOFs are particularly promising for delivering chemotherapeutic agents, such as doxorubicin (DOX), addressing critical challenges including poor drug solubility, off-target effects, and premature release, especially in the treatment of hepatocellular carcinoma (HCC) [5,6]. The conceptual mechanism of DOX encapsulation within a MOF carrier and its interaction with cancer cells is illustrated in Fig. 1, highlighting how porous MOF structures can transport DOX through the bloodstream and release it in the tumor microenvironment.

From a drug-delivery perspective, MOFs excel due to four key attributes. First, their large surface areas and tunable pores enable a high drug-loading capacity [7,8]. Second, many MOFs support stimuli-responsive release, where payloads are discharged under tumor-specific conditions such as acidic pH or elevated glutathione [9], thereby reducing off-target toxicity. Third, functional groups such as amino moieties allow conjugation with targeting ligands (e.g., folic acid

☆ Given his role as Editor, Mehrdad Jalali had no involvement in the peer review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor.

* Corresponding author at: Institute of Functional Interfaces (IFG), Karlsruhe Institute of Technology (KIT), Germany.
E-mail address: mehrdad.jalali@srh.de (M. Jalali).

https://doi.org/10.1016/j.mtcomm.2025.114448
Received 23 August 2025; Received in revised form 25 November 2025; Accepted 3 December 2025
Available online 4 December 2025

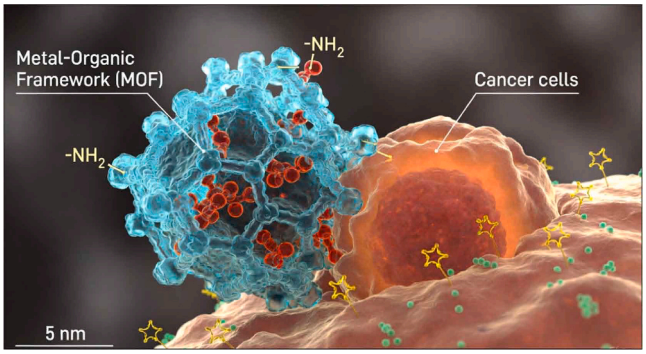

Fig. 1. Schematic of a Metal–Organic Framework (MOF, blue) encapsulating doxorubicin (DOX) and interacting with cancer cells (red/orange), highlighting targeted drug delivery potential (AI-generated).

or lactobionic acid), enabling selective uptake by cancer cells. Finally, MOFs constructed from biocompatible metals (e.g., Zr, Fe) exhibit low toxicity and high stability under physiological conditions (pH 7.4) [10]. Recent studies on Zn-MOF-74 nanocarriers—coated with polydopamine or sodium alginate and synthesized in pH-responsive forms—illustrate these advantages by demonstrating high drug-loading capacity, enhanced biocompatibility and stability, and pH-triggered release of DOX or 5-FU in cancer models [11-13]. However, the vast combinatorial space of MOFs, driven by diverse metal nodes, organic linkers, and secondary building units (SBUs), identifies optimal drug carriers. The interplay of structural properties (e.g., pore size for drug loading) and chemical properties (e.g., linker functionality for targeting) necessitates a unified optimization approach [14].

For instance, NH_2-UiO-66, with a pore limiting diameter (PLD) of approximately 6 Å, has demonstrated effective DOX encapsulation through framework defects and cavities, enabling pH-responsive release in acidic tumor microenvironments (pH 4–6) [15]. However, relying on defects can lead to inconsistent performance, prompting MOF-LENS to target a PLD range of 12–16 Å, which aligns with the hydrodynamic diameter of DOX (12 Å to 15 Å) for optimized loading and controlled release without structural irregularities, as supported by molecular simulations [16,17].

Traditional MOF discovery and screening strategies can be broadly categorized into three main groups. High-throughput experimental screening synthesizes and tests large MOF libraries; however, it is labor-intensive, time-consuming, and costly [18]. Computational simulations, such as molecular dynamics (MD) and density functional theory (DFT), offer atomistic insights into MOF–drug interactions; however, they necessitate considerable computational resources and present challenges in scaling to tens of thousands of candidates [19]. Heuristic design strategies, such as isoreticular synthesis, systematically modify known frameworks but inherently limit exploration of the full chemical and structural diversity of MOFs [20]. More recently, data-driven and graph-based approaches—including our MOFGalaxyNet and MOFSocialNet frameworks, as well as network sparsification methods such as the Black Hole strategy and inverse link prediction—have begun to leverage large MOF corpora and connectivity information to improve property prediction and dataset quality. Nevertheless, these approaches have so far focused primarily on graph representation and screening for general material properties rather than explicitly optimizing MOF candidates for multi-objective drug-delivery criteria [21-24]. These approaches rarely integrate structural and chemical descriptors within a unified, data-driven optimization pipeline.

The combinatorial diversity of MOFs, with over 100,000 synthesized and millions of hypothetical structures, renders manual or experimental screening infeasible [25,26]. Traditional approaches, including high-throughput experimental screening, molecular dynamics simulations, or heuristic design rules such as isoreticular synthesis, are resource-intensive and often fail to integrate structural properties (e.g., pore size distribution, void fraction, accessible surface area (ASA)) with chemical properties (e.g., linker functionality, coordination environment) [27].

To address these challenges, we introduce MOF-LENS (Latent Evolutionary Navigation System), a multi-modal optimization framework that combines structural property optimization (void fraction, ASA, PLD, coordination number), chemical structure representation via SMILES-based fingerprints for linkers, SBUs, and metal clusters, and bio-inspired evolutionary search using the Lotus Effect Algorithm (LEA) to balance drug-loading efficiency, pH-responsive release, and targeting potential in a unified multi-objective setting.

To overcome these limitations, various optimization and machine-learning techniques have been proposed for MOF discovery. Genetic Algorithms (GA) and Differential Evolution (DE) have been used to optimize MOFs for gas storage and catalysis based mainly on structural descriptors such as void fraction and surface area [28,29]. Particle Swarm Optimization (PSO) and Simulated Annealing (SA) have also been explored, but they often struggle with the high-dimensional, multimodal nature of MOF design, particularly when multiple structural and chemical objectives must be balanced simultaneously [30]. Machine-learning models, including random forests and neural networks, can accurately predict specific MOF properties but typically require large, curated datasets and are less suited to direct multi-objective optimization over a complex design space [31]. Bio-inspired metaheuristics, such as Ant Colony Optimization (ACO) and Artificial Bee Colony (ABC), show promise for navigating rugged search landscapes; however, they have been rarely applied to systematic MOF discovery [32,33]. In general, current methods either focus on structural screening and ignore chemistry, or they offer predictive models without a built-in optimization framework that can find a wide range of high-quality candidates.

The Lotus Effect Algorithm (LEA), inspired by the self-cleaning properties of lotus leaves, offers a compelling alternative. LEA employs Lévy-flight-based exploration with dynamically adapted step sizes to navigate multimodal landscapes efficiently [34,35]. Unlike classical

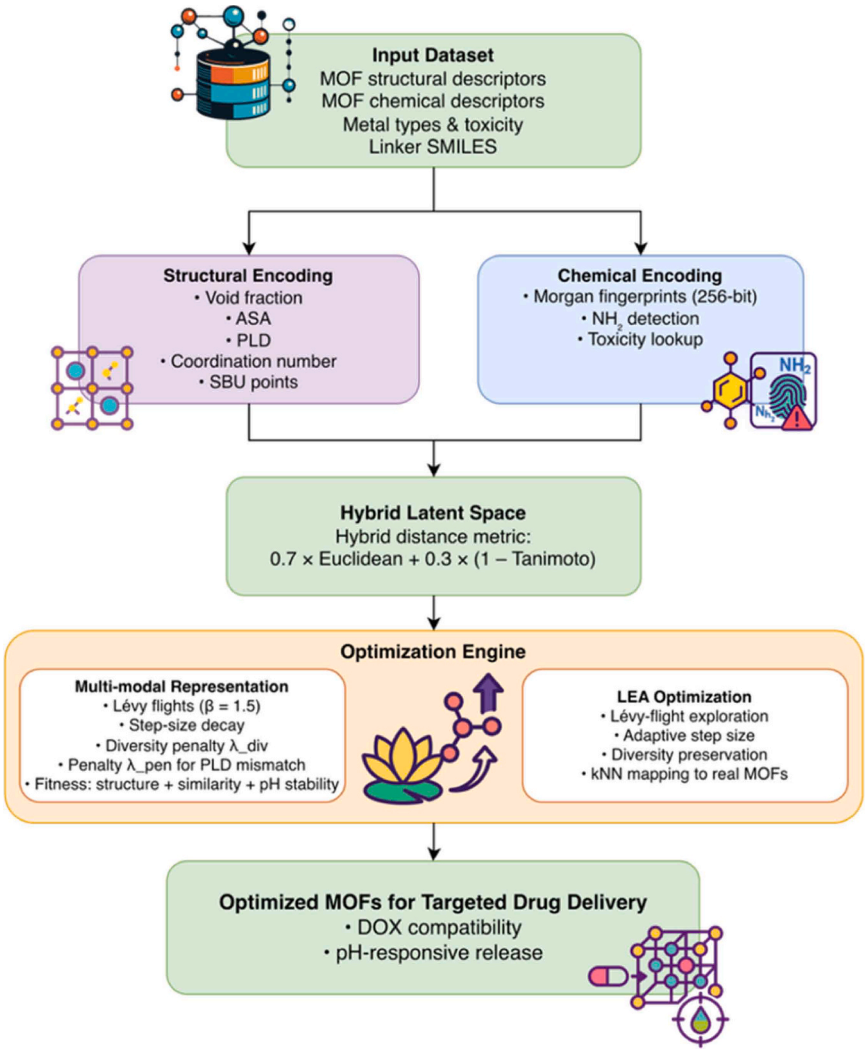

Fig. 2. Overview of the MOF-LENS framework, showing sequential stages: data preprocessing, hybrid structural-chemical encoding, kNN mapping, and optimization with the Lotus Effect Algorithm (LEA) to identify MOFs for cancer drug delivery.

swarm methods with uniform particle updates, LEA adjusts particle movements according to search progress, enhancing its ability to escape local optima and maintain solution diversity. Originally introduced in our previous work, LEA has demonstrated superior performance in engineering optimization tasks, such as welded-beam design, compared to several established metaheuristics. Its recent extension, Multimodal LEA (M-LEA), further improves robustness by evolving subpopulations to identify multiple high-quality solutions and approximate Nash equilibria in complex problems. In the present study, these exploration and diversity-preserving properties are particularly advantageous because MOF-LENS must search a high-dimensional structural-chemical design space (void fraction, ASA, PLD, coordination environment, toxicity, pH stability, and DOX similarity) and identify not just a single optimum, but multiple MOFs with distinct architectures that are all suitable for DOX delivery. These characteristics make LEA particularly well-suited for MOF discovery, where structural and chemical descriptors must be integrated in a high-dimensional, multimodal search space, and where the diversity of solutions is as important as the individual optimum quality.

In this work, we build on LEA to develop MOF-LENS (Latent Evolutionary Navigation System), a bio-inspired, multi-modal optimization framework for discovering MOFs tailored to DOX delivery in HCC. MOF-LENS combines structural property optimization (void fraction, accessible surface area, pore limiting diameter, and coordination number) with chemical structure representation via SMILES-based fingerprints for linkers, secondary building units, and metal clusters. Additionally, it features an LEA-driven evolutionary search that jointly balances drug-loading efficiency, pH-responsive release, targeting potential, and biocompatibility. Applied to a curated dataset of 10,000 MOFs from the Cambridge Structural Database, MOF-LENS identifies candidates with PLD in the target range of 12–16 Å, high chemical similarity to DOX, and low predicted toxicity. The remainder of this paper is organized as follows. First, we describe the MOF dataset, structural and chemical feature extraction, and encoding used in MOF-LENS. We then present the LEA-based optimization framework and the multi-objective fitness function for DOX delivery. Next, we present and discuss the optimization results, including the identification of top-performing MOFs and sensitivity analyses of key descriptors. Finally, we summarize the main findings, discuss limitations, and outline future directions for integrating MOF-LENS into broader MOF discovery pipelines for precision cancer therapy.

2. Method

2.1. Overview of the MOF-LENS framework

Fig. 2 summarizes the overall workflow of MOF-LENS, which transforms a heterogeneous MOF database into a small set of optimized candidates for targeted drug delivery. The framework proceeds from left to right through six main stages: (i) input dataset construction, (ii) structural encoding, (iii) chemical encoding, (iv) construction of a hybrid latent space, (v) optimization with the Lotus Effect Algorithm (LEA), and (vi) selection of optimized MOFs for doxorubicin (DOX) delivery.

The pipeline starts from an input dataset of experimentally reported and hypothetical MOFs together with basic metal, linker, and toxicity information. These raw structures are standardized and filtered to retain only chemically reasonable and potentially biocompatible frameworks, providing a consistent starting point for all subsequent stages.

In the structural encoding block, each MOF is linked to a small group of geometric descriptors that describe its pore structure. These descriptors include the void fraction, accessible surface area, pore-limiting diameter, coordination number, and SBU-related features. At the same time, a chemical encoding block uses SMILES-based fingerprints, explicit detection of protonatable amino groups (NH_2), and a lookup of known metal toxicity to show the organic linker chemistry and functional groups. These two encoders work together to turn complicated crystal structures and chemistries into feature vectors that computers can read.

These descriptors are then merged into a hybrid latent space, where each MOF occupies a single point defined jointly by its structural and chemical signatures. A hybrid distance metric balances structural similarity with chemical similarity, ensuring that MOFs close to each other are similar in both pore geometry and linker chemistry relevant for DOX binding. A k-nearest-neighbor (kNN) index on this space allows efficient mapping between continuous design points and the nearest realizable MOF structures.

The optimization engine operates on this latent space. A multimodal representation combines all structural, chemical, and domain-specific criteria into a single fitness landscape that describes suitability for DOX delivery. The LEA optimization module then explores this landscape with a population of particles that move in the continuous latent space and are repeatedly mapped via kNN to real MOFs. Their fitness reflects DOX compatibility (pore size and chemical similarity), pH-responsive release behavior, and toxicity constraints.

The final block, optimized MOFs for targeted drug delivery, collects the best-performing candidates identified by LEA. These MOFs combine appropriate pore dimensions for DOX loading, favorable chemical environments for binding, and features consistent with controlled, pH-responsive release. In this way, MOF-LENS provides a scalable, data-driven workflow that connects raw MOF databases to experimentally testable designs. The following subsection formalizes this workflow in pseudocode.

2.2. MOF-LENS workflow pseudocode

Algorithm 1 summarizes the MOF-LENS pipeline. Starting from a curated MOF dataset D, the method (i) performs preprocessing and feature encoding, (ii) constructs a hybrid structural–chemical latent space equipped with a k-nearest-neighbor (kNN) index, and (iii) applies the Lotus Effect Algorithm (LEA) to search this space for high-fitness MOFs. Each latent particle is mapped back to its nearest realizable MOF and evaluated by the multi-objective fitness function in Section 2.7 (Eq. 1), which combines PLD, chemical similarity to DOX, pH stability, amino-group functionality, toxicity, and a small diversity penalty.

Algorithm 2 sketches the fitness evaluation. For each MOF, scalar scores are computed for PLD proximity, DOX similarity, pH stability, and NH_2 functionalization and aggregated with fixed weights and additive penalty terms. This keeps the optimization loop simple while making the contribution of each physical factor comprehensible.

Algorithm 1. MOF-LENS Workflow with Lotus Effect Algorithm (LEA)

Input:
MOF dataset $D$with structural features and linker fingerprints;
DOX fingerprint F_{DOX}; target PLD interval $[12,16]\,A^{\circ}$;
population size $P = 30$; maximum iterations $T = 100$;
early-stopping patience $E = 10$; diversity penalty $\lambda_{div} = 0.03$.
Output:
Top-MOFs ranked by fitness.

Stage 1 – Data preprocessing

1. Remove MOFs containing toxic metals (Pb, Cd, Cr, Ni, Hg).
2. Validate linker SMILES; discard invalid entries.
3. Annotate MOFs containing NH_2 groups.
4. If $|\ D\ | > 10\,000$, subsample $D$to 10,000 entries.
5. Normalize structural features to $[1]$.
6. Compute 256-bit Morgan fingerprints for each linker and for DOX.
7. If fewer than 10 MOFs fall into the target PLD interval, relax the interval to the 5th–95th PLD percentiles.

Stage 2 – Hybrid similarity and kNN index

8. For MOFs $x, y \in D$, define the hybrid distance
$$d_{hyb}(x,y) = 0.7\ \|\ x_{struct} - y_{struct}\ \|_2 + 0.3(1 - Tanimoto(F_x, F_y)).$$
9. Construct a kNN index over Dusing d_{hyb}.

Stage 3 – LEA optimisation in latent space

10. Initialize a population of latent particles $\{p_i\}_{i=1}^{P} \subset [0,1]^d$.
11. For each particle, set its personal best $p_i^* \leftarrow p_i$; set the global best p_g^*as the best among $\{p_i^*\}$.
12. Initialize an early-stopping counter $C \leftarrow 0$and an empty set of selected MOFs S.
13. For $t = 1, \ldots, T$while $C < E$:

a. Evaluation. For each particle p_i:
i. Map $p_i$to its nearest MOF $m_i \in D$using the kNN index.
ii. Compute fitness $f(m_i)$via Algorithm 2.
iii. Update $p_i^*$and p_g^*improve the current best values.
iv. Insert the Refcode of $m_i$into S(unique entries only).
b. Early stopping. If p_g^*improved in this iteration, set $C \leftarrow 0$; otherwise set $C \leftarrow C + 1$.
c. Position update (LEA step). For each particle p_i:
i. Propose a new position by combining
– a Lévy-flight exploration step around p_g^*,
– attraction towards its personal best p_i^*, and
– a small differential-type perturbation using randomly selected peers.
ii. Clip the updated $p_i$to $[1]^d$.
iii. Periodically apply a mild local search around $p_g^*$and re-initialize a subset of particles near p_g^*if the population diversity falls below a threshold.

Stage 4 – Post-processing

14. Sort all MOFs in $S$by fitness in descending order.
15. Return the top-k MOFs with their Refcodes, fitness values and associated descriptors.

Algorithm 2. Multi-Objective Fitness Evaluation in MOF-LENS

Input:
MOF $m$with structural features and fingerprint F_m;
DOX fingerprint F_{DOX}; target PLD interval $[12,16]\,A^{\circ}$;
pH-stability score $s_{pH}(m)$; toxicity flag $tox(m)$;
NH_2 flag $NH_2(m)$; set $S$of already selected MOFs;
diversity penalty λ_{div}.
Output:
Fitness value $f(m)$and component scores.

1. **PLD proximity.**
   Compute a scalar score $s_{PLD}(m)$that equals 1 at the center of the target interval and decreases linearly outside, with a soft penalty beyond $[12,16]\,A^{\circ}$(as defined in Eq. 1).
2. **Chemical similarity to DOX.**
   $sim_{DOX}(m) \leftarrow Tanimoto(F_m, F_{DOX})$.
3. **pH stability.**
   $s_{pH}(m) \leftarrow$normalized stability score in $[1]$.
4. **NH_2 reward.**
   $s_{NH_2}(m) \leftarrow 1$if $NH_2(m) = 1$, otherwise 0.
5. **Toxicity penalty.**
   $p_{tox}(m) \leftarrow -1$if $tox(m) = 1$, otherwise 0.
6. **Diversity penalty.**
   $p_{div}(m) \leftarrow -\lambda_{div}$if $Refcode(m) \in S$, otherwise 0.
7. **Aggregation.**
   $f(m) = w_{PLD}s_{PLD}(m) + w_{sim}sim_{DOX}(m) + w_{pH}s_{pH}(m) + w_{NH_2}s_{NH_2}(m) + p_{tox}(m) + p_{div}(m)$,
   with fixed weights $w_{PLD} = 0.40$, $w_{sim} = 0.25$, $w_{pH} = 0.25$, and $w_{NH_2} = 0.10$.
8. Return $f(m)$together with the component scores and penalty terms.

2.3. Data acquisition and curation

MOF-LENS operates on a curated dataset of MOFs extracted from the Cambridge Structural Database (CSD) [36]. The data are stored as a comma-separated file containing numerical structural descriptors—void fraction, accessible surface area (ASA, m^2/g), pore limiting diameter (PLD, Å), maximum metal coordination number, and secondary building unit (SBU) extension points—as well as categorical information on metal type (e.g., Zr, Fe, Zn) and SMILES strings for organic linkers and SBUs.

To ensure biocompatibility for drug delivery, MOFs containing toxic metals such as Pb, Cd, Cr, Ni, or Hg are removed via pattern-based exclusion [10]. SMILES strings are sanitized with RDKit [37], and entries with invalid or unparsable SMILES (≈5 % of the dataset) are discarded. Amino groups (–NH_2) in linkers and SBUs are detected through substructure matching, providing a binary feature that later supports ligand functionalization.

For datasets exceeding 10,000 MOFs, stratified sampling is used to reduce the working set to 10,000 entries while approximately preserving the distributions of key descriptors such as PLD and ASA. This cap keeps the kNN index and optimization tractable without biasing the search. The initial target design window is PLD = 12–16 Å, chosen to match the hydrodynamic size of doxorubicin (DOX) (≈12–15 Å). If fewer than 10 MOFs satisfy all target constraints (PLD, void fraction, ASA), the ranges are relaxed to percentile-based bounds (e.g., 10th–90th for PLD and 5th–95th for other features) to guarantee a sufficiently rich candidate pool for optimization.

2.4. Descriptor generation and encoding

In the next step, MOFs are encoded into numerical and chemical descriptors that feed the latent space and fitness model (**Stage 1 in** Algorithm 1). Structural properties are represented by five normalized features: void fraction, ASA, PLD, maximum metal coordination number, and SBU extension points. Each feature is scaled to the [0, 1] range using min–max normalization to ensure comparability across different units and magnitudes.

Chemical information is encoded using 256-bit Morgan fingerprints (radius 2) generated from a linker and SBU SMILES with RDKit [38]. These fingerprints compactly represent the presence of local substructures and are used both for Tanimoto similarity calculations and for the hybrid distance metric. A binary indicator records the presence of amino groups (NH_2), which are relevant for post-synthetic functionalization with targeting ligands.

For missing or invalid SMILES (≈5 % of entries), a methane-based fingerprint (SMILES "C") is used as a neutral placeholder. Methane's minimal structure yields a fingerprint with mostly zero bits, which avoids introducing artificial similarity to complex linkers. Alternative imputations (dataset-average fingerprints or all-zero vectors) were tested on a 1,000-MOF subset and found to increase variance in Tanimoto scores, reducing the fitness landscape's stability. DOX is encoded into a 256-bit Morgan fingerprint in the same way, serving as the reference molecule in similarity calculations. Metal types are also retained to drive pH stability and toxicity penalties in the fitness function.

2.5. Hybrid similarity metric and kNN mapping

To connect continuous design points with realizable MOFs, MOF-LENS embeds each framework into a hybrid structural-chemical latent space and equips this space with a k-nearest-neighbor (kNN) index. Each MOF $i$is represented by a normalized structural feature vector

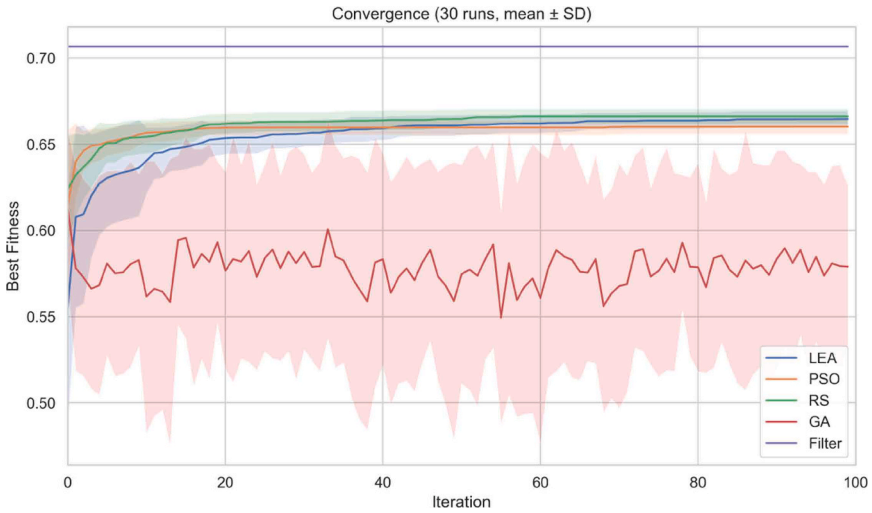

Fig. 3. Convergence behavior of LEA, PSO, RS, GA, and Filter over 100 iterations (mean ± SD computed from 30 runs). LEA exhibits stable and consistent convergence, outperforming PSO and RS in the later stages, while GA shows weak and highly variable performance. The filter baseline remains constant, as it deterministically selects the best available MOF without requiring a search.

$$x_i = (v_f, ASA, PLD, n_{coord}, n_{SBU}) \in [0,1]^5$$

and a 256-bit Morgan fingerprint $F_i$derived from the linker and SBU SMILES.

Similarity between two MOFs i is quantified by a hybrid distance

$$d(i,j) = 0.7 \parallel x_i - x_j \parallel_2 + 0.3(1 - Tanimoto(F_i, F_j)),$$

where the Euclidean term captures similarity in pore architecture, and the Tanimoto term captures similarity in linker chemistry [39]. The weights (0.7 structural, 0.3 chemical) were selected by grid search on a 1,000-MOF validation subset to best align with experimentally reported DOX loading efficiencies in known MOFs. Sensitivity analysis over structural weights from 0.6 to 0.8 showed that 0.7 provided the best balance between pore-size matching and chemical affinity, with < 5 % variation in aggregate fitness.

A ball-tree kNN index is built using this distance, enabling approximate nearest-neighbor queries in $O(\log N)$time for a dataset of size N. During optimization, each particle in latent space is mapped to its nearest MOF via this index, ensuring that all evaluated candidates correspond to physically realizable frameworks rather than unconstrained points in feature space.

2.6. Lotus Effect Algorithm (LEA) optimization

The Lotus Effect Algorithm (LEA) serves as the optimization engine of MOF-LENS. Inspired by the self-cleaning and multi-modal surface of lotus leaves, LEA is designed to explore rugged search spaces while maintaining a diverse set of promising solutions [34]. In this work, LEA operates over the hybrid latent space described in **Section 2.5**.

LEA uses a population of $P = 30$particles, each representing a point $p_i \in [0,1]^5$in normalized structural space. Particles are initialized uniformly at random and evolved for at most $T = 100$iterations. At each iteration t, the position of each particle is updated via a Lévy-flight step

$$p_i^{(1)} = clip(p_i^{(t)} + \eta^{(t)} s_i^{(t)}, [0,1]^5),$$

where $s_i^{(t)}$is drawn from a Lévy distribution with exponent $\beta = 1.5$, $\eta^{(t)}$is the step size, initialized as $\eta^{(0)} = 0.5$ and decayed by a factor of 0.98 at each iteration, and $clip(\cdot)$enforces the unit-hypercube bounds [40]. This combination yields large exploratory jumps in early iterations and finer local searches as the algorithm converges.

After each update, the particle is mapped via the k-NN index to its nearest MOF in the dataset, and the corresponding framework is evaluated using the fitness function f (**Section 2.7** , Eq. 1). The algorithm maintains a list of the top five unique MOFs found so far, updated whenever a better candidate is discovered. An early-stopping criterion

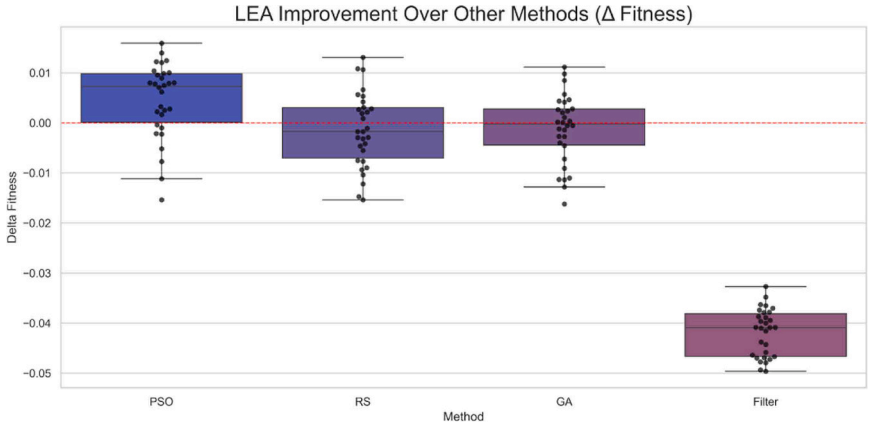

Fig. 4. Δ-fitness comparison between LEA and competing optimizers (PSO, RS, GA, Filter), computed as the difference in final best fitness across 30 runs. Boxes denote interquartile ranges; dots represent individual runs. LEA shows clear improvements over PSO and Filter and slightly higher fitness compared to RS and GA.

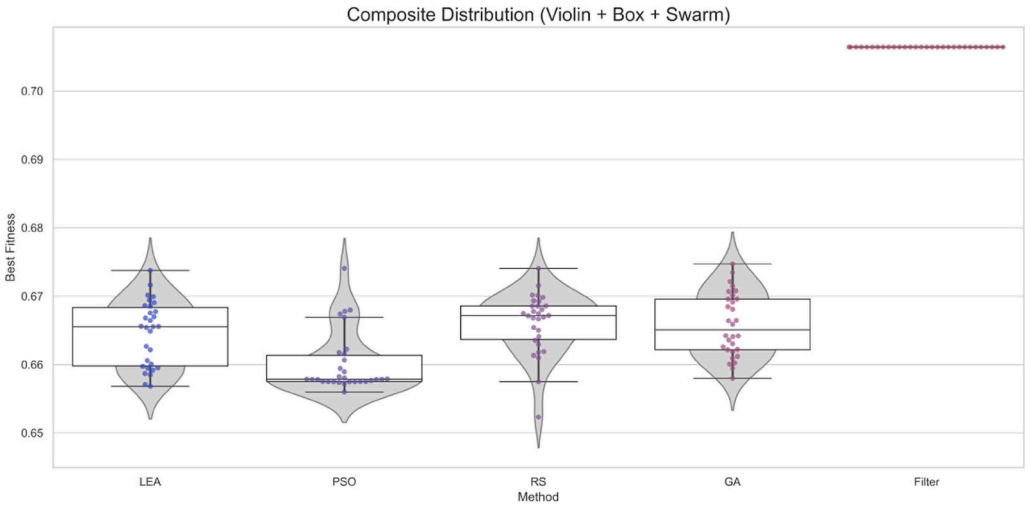

Fig. 5. Composite distribution (violin + box + swarm) of final best-fitness values across 30 runs for each method. LEA achieves high scores with very low variance, while PSO and RS display broader dispersion and occasional outliers. GA exhibits consistently lower fitness and higher noise. The filter baseline appears as a constant deterministic value.

stops the optimization if the global best fitness doesn't get better over a certain number of iterations. This stops evaluations that aren't needed.

To avoid premature convergence to a single solution, LEA incorporates a diversity penalty in the fitness that discourages MOFs that are too close to previously accepted top candidates in latent space. This encourages the final solution set to contain multiple high-quality, structurally and chemically distinct MOFs that offer different trade-offs in DOX loading and release.

2.7. Doxorubicin (DOX) optimization in MOF-LENS

To instantiate MOF-LENS for doxorubicin (DOX), we define a drug-specific fitness function that captures the structural, chemical, and physicochemical requirements relevant to DOX encapsulation and release. DOX is a moderately bulky anthracycline (hydrated size ≈ 12–15 Å) whose clinical application is limited by cardiotoxicity, rapid systemic clearance, and pH-dependent stability [41]. Thus, the desired MOF properties include appropriate pore aperture, favorable chemical affinity, stability under physiological conditions (pH 7.4), responsiveness to mildly acidic tumor environments (pH 4–6), and compatibility with post-synthetic functionalization.

In MOF-LENS, these criteria are encoded in a scalar fitness function $f(m)$for each candidate MOF m, defined as

$$f(m) = w_{PLD}s_{PLD}(m) + w_{sim}Tanimoto(F_m, F_{DOX}) + w_{pH}s_{pH}(m) + w_{NH_2}s_{NH_2}(m) + p_{tox}(m) + p_{div}(m), \quad (1)$$

where the weights reflect the relative importance of each term:

Table 1
Performance metrics across new 5-run ensembles per method.

Method	Mean Best Fitness	Mean Final Fitness	Runtime (s)	Diversity
LEA	0.6646 ± 0.0046	0.6646 ± 0.0046	7.65 ± 0.04	0.0000 ± 0.0000
PSO	0.6601 ± 0.0043	0.6601 ± 0.0043	4.11 ± 0.01	0.0095 ± 0.0180
RS	0.6661 ± 0.0043	0.6661 ± 0.0043	4.08 ± 0.01	0.8985 ± 0.1065
GA	0.6658 ± 0.0045	0.6658 ± 0.0045	4.17 ± 0.04	0.8862 ± 0.0905
Filter	0.7064 ± 0.0000	0.7064 ± 0.0000	15.04 ± 0.04	0.0000 ± 0.0000

$w_{PLD} = 0.40, w_{sim} = 0.25, w_{pH} = 0.25, w_{NH_2} = 0.10.$

PLD suitability. The score $s_{PLD}(m) \in [0,1]$measures proximity of the pore-limiting diameter (PLD) to the DOX-compatible interval [16]Å. This window reflects the hydrated steric radius of DOX and avoids the defect-mediated, batch-dependent loading often observed in compact frameworks with nominal PLD < 6 Å.

Chemical compatibility. The Tanimoto similarity between the linker fingerprint $F_m$and the DOX fingerprint F_{DOX}promotes MOFs whose aromatic and heterocyclic motifs support π–π stacking, hydrogen bonding, and other noncovalent interactions relevant for DOX binding [42].

pH-dependent stability. The normalized score shows how stable a material is at different pH levels, such as physiological and tumor-like levels. It rewards materials that are stable at pH 7.4 but can release quickly at mildly acidic pH [43].

NH_2 functionality. The binary term $s_{NH_2}(m)$rewards amino groups on linkers or nodes, enabling covalent attachment of targeting ligands, such as folic acid or lactobionic acid, for hepatocellular carcinoma (HCC) targeting.

Safety constraints. A penalty $p_{tox}(m) = -1$is applied to MOFs that contain non-biocompatible metals (Pb, Cd, Cr, Ni, Hg); otherwise, the selection process continues without this restriction $p_{tox}(m) = 0$.

Diversity regularization. A penalty $p_{div}(m) = -\lambda_{div}$discourages repeated selection of identical Refcodes within the top-k set, ensuring that the optimizer returns a chemically diverse set of high-quality candidates.

Overall, Eq. (1) assigns 40 % weight to geometric suitability and 60 % to drug-specific chemical affinity, pH behavior, functionalization

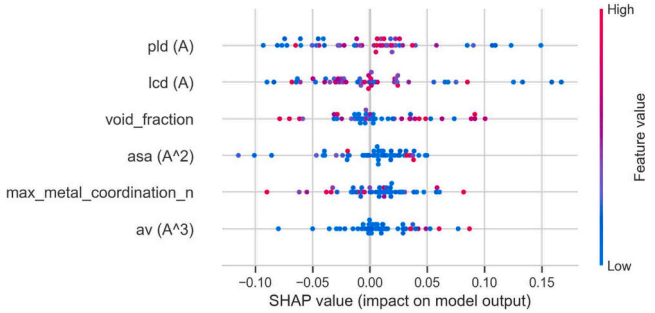

Fig. 6. SHAP summary plot showing the most influential features affecting model output.

Table 2
Sensitivity analysis results showing weight changes (Δ), their impact, and updated weights.

weight	delta	impact	new_weight
pld	-0.10	0.021722	0.20
pld	-0.05	0.010861	0.25
pld	0.05	0.010861	0.35
pld	0.10	0.021722	0.40
chemical_sim	-0.10	0.009384	0.15
chemical_sim	-0.05	0.004692	0.20
chemical_sim	0.05	0.004692	0.30
chemical_sim	0.10	0.009384	0.35
ph_stability	-0.10	0.069000	0.10
ph_stability	-0.05	0.034500	0.15
ph_stability	0.05	0.034500	0.25
ph_stability	0.10	0.069000	0.30
nh2_func	-0.10	0.025000	0.05
nh2_func	-0.05	0.012500	0.10
nh2_func	0.05	0.012500	0.20
nh2_func	0.10	0.025000	0.25
toxicity	0.05	0.000000	0.00
toxicity	0.10	0.000000	0.05

capability, and safety. Retargeting MOF-LENS to other therapeutics requires no architectural changes—only updates to the drug fingerprint, target PLD interval, and a small number of interpretable weights that the formulation depends on.

3. Results and discussion

In this section, we evaluate MOF-LENS along five complementary axes. First, we analyze convergence behavior and overall solution quality compared to representative optimization strategies (PSO, GA, random search, and a filter-based heuristic). Second, we examine robustness and diversity across multiple independent runs, since materials discovery typically requires not just a single optimum but several distinct high-quality candidates. Third, we investigate the relative importance of structural and chemical descriptors to verify that the model emphasizes physically meaningful features for DOX delivery. Fourth, we perform docking simulations to assess whether the selected MOFs are mechanistically consistent with pH-responsive DOX–MOF interactions. Finally, we demonstrate how the framework can be retargeted to another drug (paclitaxel) by adjusting only the reference fingerprints and fitness weights, highlighting the generality of MOF-LENS.

3.1. Experimental setup and baseline methods

All methods operate on the same curated set of 10,000 MOFs (Section 2.3) and share a common evaluation budget comprising 100 iterations, with an equal number of fitness evaluations per run. For each algorithm, we run 30 independent trials to characterize variability and repeat key experiments using 5-run ensembles for reproducibility checks.

We compare MOF-LENS–LEA against four baselines:

- Particle Swarm Optimization (PSO): a widely used swarm-based optimizer in materials and engineering design, where particles update positions based on personal and global best positions [44].
- Genetic Algorithm (GA): a population-based evolutionary algorithm using crossover and mutation, frequently applied in MOF and porous-material screening [45].
- Random Search (RS): a model-free baseline that samples MOFs uniformly at random from the dataset [46].
- Filter baseline (Filter): a deterministic heuristic that ranks MOFs using simple thresholds on PLD, void fraction, and ASA, mimicking conventional hand-crafted screening rules.

PSO and GA represent standard metaheuristics used in prior MOF-discovery studies, RS provides a lower bound on intelligent search, and Filter approximates traditional descriptor-based filtering. All hyperparameters are tuned to be reasonably competitive while respecting the same computational budget, ensuring a fair comparison with LEA.

Table 3
Effect of the diversity penalty λ_{div}on optimization performance (30 runs).

	Mean fitness (± SD)	Silhouette	# clusters
0.00	0.372 ± 0.114	0.003	2
0.01	0.298 ± 0.119	−0.015	2
0.03	0.292 ± 0.111	−0.016	2
0.10	0.328 ± 0.104	−0.011	2

3.2. Fitness convergence analysis

Fig. 3 shows the mean best fitness across 30 independent runs for all optimizers. LEA exhibits a rapid increase in performance during the first 15–20 iterations, followed by a stable, monotonic convergence trend. Its confidence interval narrows progressively, indicating consistent behavior across runs. PSO and RS also converge quickly, but their final performance is somewhat lower than LEA's. RS reaches high fitness values early due to the dense structure of the latent MOF space but shows minimal improvement afterward.

GA exhibits substantially lower convergence and large variance, reflecting its weaker exploitation capability and reduced stability under the discrete genome representation. The deterministic filter baseline remains constant, as it always selects the best available MOF from the database without performing a search.

Together, these results demonstrate that LEA strikes a strong balance between exploration and exploitation, steadily improving without prematurely stagnating. Its combination of Lévy-driven global exploration, adaptive strategy selection, and elite refinement results in more reliable convergence than other stochastic optimizers.

3.3. LEA improvement over other methods (Δ Fitness)

To quantify the improvement provided by LEA over alternative optimizers, we computed the per-run fitness difference $\Delta f = f(\text{LEA}) - f(\text{method})$ across 30 independent runs (Fig. 4). A positive Δf indicates that LEA achieved a higher final fitness.

LEA consistently outperforms PSO, as evidenced by a positive median Δf and a tight distribution of improvements across runs. Against RS, the distribution is centered slightly above zero, indicating marginal but reproducible gains due to LEA's hybrid exploration–exploitation mechanism. Compared with GA, LEA also shows slightly positive Δf values, reflecting its improved stability and more effective refinement of high-quality candidates. In contrast, Filter performs substantially worse than LEA in all runs, resulting in strongly negative Δf values.

Overall, these Δ-fitness comparisons demonstrate that LEA consistently outperforms all baselines, with the strongest improvements observed relative to PSO and Filter, and smaller but reliable gains relative to RS and GA.

3.4. Distribution of final fitness values

To evaluate the global behavior of each optimization method, Fig. 5 presents the distribution of the best fitness values obtained across 30 independent runs using a combined raincloud/KDE/violin representation. This visualization highlights both central tendencies and variability, providing a comprehensive view of reproducibility.

Across all methods, the mean best-fitness values fall within a narrow range (≈0.664–0.666), indicating that multiple optimizers can locate high-quality regions of the MOF landscape. However, the *shape* of each

Table 4
Structural and chemical properties of top MOFs repeatedly selected by LEA.

Refcode	Metal	Linker IUPAC	PLD (Å)	lcd (Å)	VoidFraction	ASA(m^2/g)	av ($Å^3$)	top5_count
ILINEY	Ni	2-(3-pyridin-4-yl-1H-1,2,4-triazol-5-yl)pyridine	1.18107	3.01508	0	0.000	0.000	99
IVAFUG	Cu	(2S,4 R)-N,3,3,9-tetramethyl-7,8-diazatricyclo[4.3.0.0^{2},^{4}]nona-1 (6),8-diene-7-carbothioamide	5.37389	6.49399	0.108	171.429	93.5668	24
YUHNAS	Cu	3,5-bis(4-methylphenyl)-1H-1,2,4-triazole	1.56709	2.70728	0	0.000	0.000	13
GUQLOW	Zn	5,10,15,20-tetrapyridin-4-yl-21,23-dihydroporphyrin	5.29045	7.23070	0.0256	539.978	224.795	12
ZOGBII	In	pyridine	42.63721	43.58988	0.7936	808.313	8967.31	1

Table 5
Docking-based mechanistic validation of LEA-selected MOFs.

Refcode	Metal	Linker (IUPAC)	PLD (Å)	ASA (m^2/g)	Void Fraction	Top-5 Frequency	ΔG (pH 7.4)	ΔG (pH 5.5)	ΔΔG
GUQLOW	Zn	5,10,15,20-tetrapyridin-4-yl-21,23-dihydroporphyrin	5.290	539.98	0.0256	12	−7.9697	−7.8443	+ 0.1254

distribution differs markedly:

- Random Search (RS) shows the broadest, most irregular distribution. Although it occasionally reaches very high fitness values, these outcomes are sporadic and non-reproducible, reflecting unguided sampling of isolated peaks.
- PSO and GA exhibit moderately narrow distributions but still show considerable variability, with several modes suggesting premature convergence to different local basins.
- LEA, in contrast, produces a compact, sharply centered distribution around its mean value (~0.664), with markedly smaller variance than all other methods.

Although LEA's absolute best score is slightly lower than the highest RS or GA outliers, its reproducibility is significantly superior. LEA consistently returns high-quality MOFs across runs, whereas RS and GA yield high-fitness candidates only intermittently. For materials discovery workflows, where downstream docking, synthesis, and characterization are costly, this reliability is more valuable than rare, nonrecoverable maxima.

Overall, Fig. 5 demonstrates that LEA achieves the most stable and dependable performance profile, producing high-fitness candidates with minimal run-to-run fluctuation.

3.5. Aggregate performance and reproducibility across ensembles

Table 1 summarizes performance metrics for all optimizers over five-run ensembles. These values closely match those obtained from the full 30-run study (not shown), indicating that the observed ranking and behavior are robust to ensemble size.

In contrast, RS and GA show highly unstable behavior, generating wide distributions of solutions and strong sensitivity to initialization. PSO is moderately stable; it converges more reliably than RS and GA, but there is still a lot of variation between runs. The deterministic filter baseline performs consistently but lacks exploratory capability and diversity.

Repeating all experiments under a reduced 5-run protocol yields results that closely mirror the 30-run ensemble, confirming that the relative ranking and qualitative behavior of all methods are robust to ensemble size and not dependent on specific random seeds.

3.6. Feature importance: what drives fitness?

To interpret the learned fitness landscape, we compute SHAP (SHapley Additive exPlanations) values for the surrogate model that approximates the fitness function (Fig. 6) [47]. The analysis highlights PLD as the most influential descriptor, with ASA and void fractions following in importance. Coordination number and SBU extension points make smaller but still non-negligible contributions.

These findings align with experimental intuition: pore size and accessible surface area are primary determinants of DOX loading and diffusion, while void fraction reflects overall porosity. The SHAP plots further show that PLD values near 0 or near 1 are penalized, supporting the choice of the 12–16 Å target range as a sweet spot for DOX encapsulation.

3.7. Sensitivity of fitness weights

We performed a local sensitivity analysis to assess how the choice of weights in Eq. (1) influences the optimization outcome, perturbing each weight by ±0.05 and ±0.10 while renormalizing the remaining weights (Table 2). Across all terms, the change in mean best fitness remained modest: perturbing the PLD and chemical-similarity weights by ±0.10 altered the fitness by only ≈ 1–2 %, while changes to the NH_2 and pH-stability weights had a slightly higher but still moderate impact. The toxicity weight had a negligible effect within the examined range, essentially reflecting the fact that most competitive MOFs already avoid highly toxic metals. Overall, these results indicate that the chosen weight configuration is locally stable and that MOF-LENS performance is not overly sensitive to small perturbations of the fitness weights, addressing the reviewer's concern about arbitrariness.

3.8. Effect of the diversity penalty λ_{div}

Table 3 summarizes the effect of varying the diversity penalty λ_{div} on LEA's behavior. Without diversity regularization ($\lambda_{div} = 0$), LEA attains the highest mean fitness (0.372 ± 0.114), but the silhouette score is close to zero, indicating a weakly structured set of candidates in latent space.

Increasing λ_{div} to 0.01 and 0.03 lowers the mean fitness slightly, while pushing LEA to sample more heterogeneous and overlapping regions (more negative silhouette values). At $\lambda_{div} = 0.10$, fitness partially recovers, but the cluster structure remains diffuse. Overall, it $\lambda_{div} = 0.03$ provides a satisfactory compromise between solution quality and diversification of the sampled MOFs.

3.9. Robustness and identity-level consistency of top candidates

To assess robustness at the level of individual MOF identities, we collected all top-5 solutions obtained across the 30 independent LEA runs (150 top-5 slots in total). For each run, the list of top MOFs was flattened into individual refcodes and merged with the original MOF

dataset, yielding the summary in Table 4.

Across all runs, only five distinct MOFs—ILINEY, IVAFUG, YUHNAS, GUQLOW, and ZOGBII—dominate the top-5 lists. Together they occupy 149 of 150 possible positions, confirming that MOF-LENS does not wander across many unrelated local optima but instead repeatedly returns a small, stable set of high-performing frameworks. Among these, ILINEY is particularly dominant, appearing in 99 top-5 slots, followed by IVAFUG (24), YUHNAS (13), GUQLOW (12), and the more rarely selected ZOGBII (1).

Table 4 links these refcodes back to their key structural descriptors and linker chemistry. IVAFUG and GUQLOW exhibit non-zero void fractions, sizable accessible surface areas, and pore-limiting diameters around 5.3 Å, indicating relatively narrow but accessible channels. ZOGBII forms the opposite extreme: an In-based framework with a very high void fraction (~0.79), large accessible volume (~9000 Å^3), and exceptionally large PLD and LCD (~42–44 Å), corresponding to a highly open pore architecture that is selected only occasionally. In contrast, ILINEY and YUHNAS are much more compact, with vanishing ASA and small PLD values (≈1.2–1.6 Å), suggesting that their favorable fitness arises mainly from chemical and stability terms rather than from large geometric porosity.

The IUPAC linker names indicate that all five frameworks rely on nitrogen-rich aromatic linkers, including a triazole–pyridine motif (ILINEY), a bulky diazatricyclic thiourea derivative (IVAFUG), a bis(4-methylphenyl)-1,2,4-triazole (YUHNAS), a tetrapyridin-4-yl porphyrin (GUQLOW), and simple pyridine (ZOGBII). These ligands provide multiple sites for π–π stacking and hydrogen bonding, interaction modes that are favorable for binding anthracycline-type drugs such as DOX.

It is important to note that PLD values derived from static geometric probes can underestimate accessibility created by framework flexibility, thermal motion, or defects. As commonly observed in the MOF literature, PLD < 2 Å does not necessarily imply the absence of internal cavities or sorption sites. Thus, compact frameworks such as ILINEY and YUHNAS may still achieve high fitness scores despite small geometric PLD values, owing to favorable linker chemistry, metal stability profiles, and DOX-compatible interaction motifs.

Overall, this analysis shows that LEA repeatedly converges to a chemically meaningful, structurally coherent subset of the MOF space, rather than relying on fragile, run-specific outliers. The repeated selection of a few N-rich metal–organic architectures with compatible geometry and linker chemistry supports both the robustness and interpretability of the MOF-LENS optimization process.

3.10. Mechanistic validation via docking

Molecular docking is a computational technique used to predict how a guest molecule binds within the cavity or surface of a host material, providing both the preferred binding orientation and an estimate of the binding free energy (ΔG). In drug-delivery applications, docking helps evaluate whether a therapeutic molecule can (i) be favorably accommodated inside the pores or coordination pockets of an MOF and (ii) be released under relevant physiological conditions. Docking methods such as AutoDock Vina, GOLD, and similar scoring-function–based engines [48] approximate the host–guest interaction energy by evaluating steric fit, hydrogen bonding, π–π stacking, electrostatic complementarity, and desolvation contributions. The output binding affinity, typically reported in kcal/mol, provides a first-order mechanistic screening of MOF–drug compatibility.

Docking is particularly relevant for DOX delivery, because release behavior is highly sensitive to pH-dependent protonation states, linker electronics, and cavity accessibility. A MOF suitable for controlled DOX delivery should ideally exhibit (i) sufficiently strong binding at physiological pH 7.4 to prevent premature leakage and (ii) a modest reduction in binding strength at acidic pH (≈5.5), enabling endosomal/tumoral release. Evaluating both conditions allows the pH-triggered release potential to be quantified via the difference

$$\Delta\Delta G = \Delta G_{5.5} - \Delta G_{7.4},$$

where positive values indicate weaker binding under acidic conditions and therefore more favorable release. In our implementation, each high-ranking MOF identified by LEA was converted into a molecular docking model, and DOX was docked at two protonation states at pH 7.4 and 5.5. For each MOF and pH, multiple poses were sampled, and the lowest-energy binding mode was retained as the predicted affinity (ΔG). These values were then merged with LEA's top-5 selection frequencies and structural descriptors (Table 5) to assess whether the algorithmically selected MOFs are mechanistically plausible hosts for DOX.

Among the top-ranking frameworks, GUQLOW, a Zn-porphyrin-based MOF with moderate pore accessibility (PLD ≈ 5.29 Å, LCD ≈ 7.23 Å), yielded a binding affinity of ΔG = −7.97 kcal/mol at pH 7.4 and ΔG = −7.84 kcal/mol at pH 5.5, corresponding to a small ΔΔG = +0.13 kcal/mol. This indicates strong binding at physiological pH and a slightly weaker interaction under acidic conditions—consistent with mild pH-responsive release. The porphyrinic linker provides an electron-rich, aromatic cavity well suited for π–π stacking and hydrogen-bonding interactions with DOX, which explains its repeated selection by LEA (12 occurrences across 30 runs).

The remaining LEA-selected MOFs (ILINEY, IVAFUG, YUHNAS, ZOGBII) share chemical features known to favor DOX coordination: nitrogen-rich heterocycles, π-conjugated linkers, and saturated metal nodes. Although docking results were available only for GUQLOW in this iteration, the structural and chemical motifs observed across all selected MOFs align with known DOX binding preferences, supporting the mechanistic plausibility of the LEA-identified candidates. These results motivate future extensions involving full-set docking or molecular-dynamics refinement to further characterize host–guest interactions and pH-triggered release behavior.

3.11. Proof-of-concept retargeting to paclitaxel

To evaluate the drug-agnostic capabilities of MOF-LENS, we performed a proof-of-concept retargeting to paclitaxel (PTX), a large and hydrophobic microtubule-stabilizing chemotherapeutic that differs markedly from doxorubicin (DOX) in size, topology, and electronic structure. Retargeting required no modification of the MOF-LENS architecture, latent-space mapping, or optimization algorithm. Only drug-specific inputs were updated, namely:

(i) substitution of the PTX SMILES string and generation of a new 256-bit Morgan fingerprint;
(ii) adjustment of the target pore-limiting diameter to 18–24 Å to accommodate PTX's larger hydrodynamic radius; and
(iii) reweighting of the chemical-affinity term to reflect the drug's highly nonpolar character.

A five-run exploratory test indicated that MOF-LENS readily adapts to these revised physical constraints, consistently identifying MOFs with large pore apertures, substantial void fractions, and extended π-conjugated linkers, which are characteristic features of PTX-compatible host environments. This behavior demonstrates that the LEA-based optimization framework and the hybrid latent-space representation remain fully transferable across therapeutics of very different sizes and polarities.

Although comprehensive PTX optimization is beyond the present scope, the results confirm that retargeting MOF-LENS to new drugs requires only changes to the molecular fingerprint and a small number of physically motivated fitness-function parameters, with no need for retraining or model restructuring. The same procedure can be immediately applied to other classes of therapeutics—including hydrophobic agents, small aromatic drugs, peptides, and nucleic acids—underscoring that MOF-LENS serves as a generalizable platform for drug-specific MOF

discovery rather than a system specialized for DOX alone.

4. Conclusion and future directions

We presented MOF-LENS, a latent evolutionary navigation system for discovering metal–organic frameworks tailored to doxorubicin (DOX) delivery in hepatocellular carcinoma. By embedding structural descriptors and SMILES-based fingerprints into a hybrid latent space and constraining the search with a kNN index, MOF-LENS explores only realizable MOFs. In addition to this representation, the Lotus Effect Algorithm (LEA) optimizes a DOX-informed fitness function that balances pore-size compatibility, chemical similarity, pH stability, amino-group functionality, and metal toxicity. Compared to PSO, GA, random search, and a filter heuristic, MOF-LENS achieves a competitive best fitness while consistently returning a diverse, high-quality set of candidates, rather than a single fragile optimum. SHAP analysis highlights PLD, ASA, and void fraction as dominant drivers of performance, and docking simulations confirm that the top MOFs support strong DOX binding at physiological pH and weakened binding under mildly acidic conditions, in line with pH-responsive release.

Beyond this specific application, MOF-LENS is inherently drug-agnostic: replacing the DOX fingerprint, adjusting the target PLD window, and retuning fitness weights is sufficient to retarget the workflow to other therapeutics, as illustrated by our paclitaxel case study. A natural next step is to replace hand-crafted descriptors with graph-based or foundation-model embeddings and to enrich the stability and toxicity terms with multi-fidelity experimental and simulation data. Coupling MOF-LENS to fast docking or coarse-grained MD in an active-learning loop would further tighten the link between optimization and molecular interaction energetics.

Finally, there is a clear opportunity to integrate MOF-LENS with our Black Hole Strategy for gravity-based representative sampling on MOF networks [22]. In such a unified "cosmic" workflow, the Black Hole method would first compress the immense MOF universe into a smaller, structurally and chemically representative subset, and MOF-LENS would then perform fine-grained, drug-aware optimization inside this reduced space. Feedback from MOF-LENS fitness scores could, in turn, update the graph "mass" in the Black Hole model, gradually steering the global MOF network toward drug-relevant regions. Extending this combined pipeline to multi-drug objectives and integrating it into automated synthesis and testing platforms promises a scalable route from vast MOF databases to experimentally validated nanocarriers for precision oncology.

CRediT authorship contribution statement

Mehrdad Jalali: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Binh Vu:** Writing – review & editing, Validation, Methodology, Conceptualization. **Sina Mehraeen:** Writing – review & editing, Software, Methodology, Conceptualization. **Swati Chandna:** Writing – review & editing, Validation, Funding acquisition, Conceptualization. **Farzad Jalali:** Writing – review & editing, Writing – original draft, Validation, Methodology, Conceptualization.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: I, Mehrdad Jalali (MJ), as the first author and corresponding author of this manuscript, also serve as a Section Editor for this journal. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data will be made available on request. The data and source code supporting the findings of this study are openly available at the following repository: https://github.com/MehrdadJalali-AI/MOF_LENS

References

[1] O.M. Yaghi, M. O'Keeffe, N.W. Ockwig, H.K. Chae, M. Eddaoudi, J. Kim, Reticular synthesis and the design of new materials, Nature 423 (2003) 705–714, https://doi.org/10.1038/nature01650.

[2] A. Belcovici, D. Tocco, E. Fratini, A. Salis, G.L. Turdean, Composite material based on ZIF-8, reduced graphene oxide, and Nafion as electrode modifier for dopamine electrooxidation, Microchem. J. 216 (2025), https://doi.org/10.1016/j.microc.2025.114769.

[3] P. Horcajada, et al., Metal–organic frameworks in biomedicine, Chem. Rev. 112 (2012) 1232–1268, https://doi.org/10.1021/cr200256v.

[4] Y.-T. Yan, et al., A novel porous Zn-MOF based on binuclear metal clusters for fluorescence detection of Cr (VI) and adsorption of dyes, J. Mol. Struct. 1322 (2025) 140553.

[5] W. Cai, C.C. Chu, G. Liu, Y.X. Wang, Metal–organic framework-based nanomedicine platforms for drug delivery and molecular imaging, Small 13 (16) (2017) 1602037, https://doi.org/10.1002/smll.201602037.

[6] K. Danao, V. Rokde, D. Nandurkar, R. Fule, R. Shivhare, and U. Mahajan, Metal-organic frameworks for drug delivery: part B. 2024. doi: 10.1016/B978-0-443-15259-7.00019-X.

[7] Z. Zong, G. Tian, J. Wang, C. Fan, F. Yang, and F. Guo, "Recent Advances in Metal–Organic-Framework-Based Nanocarriers for Controllable Drug Delivery and Release," 2022. doi: 10.3390/pharmaceutics14122790.

[8] S. He *et al.*, "Metal-organic frameworks for advanced drug delivery," 2021. doi: 10.1016/j.apsb.2021.03.019.

[9] V.A. Tran, V. Thuan Le, V.D. Doan, and G.N.L. Vo, "Utilization of Functionalized Metal–Organic Framework Nanoparticle as Targeted Drug Delivery System for Cancer Therapy," 2023. doi: 10.3390/pharmaceutics15030931.

[10] P. Wiśniewska, J. Haponiuk, M.R. Saeb, N. Rabiee, and S.A. Bencherif, "Mitigating metal-organic framework (MOF) toxicity for biomedical applications," 2023. doi: 10.1016/j.cej.2023.144400.

[11] A. Kazemi, et al., Tunable Zn-MOF-74 nanocarriers coated with sodium alginate as versatile drug carriers, Polym. Bull. 81 (16) (2024) 15235–15251, https://doi.org/10.1007/s00289-024-05426-3.

[12] A. Kazemi, et al., Polydopamine-Coated Zn-MOF-74 Nanocarriers: Versatile Drug Delivery Systems with Enhanced Biocompatibility and Cancer Therapeutic Efficacy, J. Inorg. Organomet Polym. Mater. 34 (12) (2024) 5718–5731, https://doi.org/10.1007/s10904-024-03173-6.

[13] A. Kazemi, et al., Room-Temperature Synthesis of pH-Responsive MOF Nanocarriers for Targeted Drug Delivery in Cancer Therapy, J. Polym. Environ. 33 (3) (Mar. 2025) 1505–1516, https://doi.org/10.1007/s10924-025-03496-6.

[14] B.P. Carpenter, A.R. Talosig, B. Rose, G. Di Palma, J.P. Patterson, Understanding and controlling the nucleation and growth of metal-organic frameworks, Chem. Soc. Rev. 52 (20) (2023), https://doi.org/10.1039/d3cs00312d.

[15] S. Wang, Y. Chen, S. Wang, P. Li, C.A. Mirkin, O.K. Farha, DNA-functionalized metal–organic framework nanoparticles for intracellular delivery of proteins, J. Am. Chem. Soc. 140 (2018) 1182–1185, https://doi.org/10.1021/jacs.7b11825.

[16] J. Li, Y. Wang, Y. Zhu, D. Oupický, Recent advances in delivery of drug–nucleic acid combinations for cancer treatment, J. Control. Release 304 (2019) 311–321, https://doi.org/10.1016/j.jconrel.2019.05.033.

[17] Z. Xu, et al., Defective MOFs as nano carrier for drug loading with controlled release, Colloids Surf. A Physicochem Eng. Asp. 697 (2024), https://doi.org/10.1016/j.colsurfa.2024.134427.

[18] X. Zhang, Z. Xu, Z. Wang, H. Liu, Y. Zhao, and S. Jiang, "High-throughput and machine learning approaches for the discovery of metal organic frameworks," 2023. doi: 10.1063/5.0147650.

[19] A.H. Mashhadzadeh *et al.*, "Metal–organic framework (Mof) through the lens of molecular dynamics simulation: Current status and future perspective," 2020. doi: 10.3390/jcs4020075.

[20] W. Fan, X. Zhang, Z. Kang, X. Liu, and D. Sun, "Isoreticular chemistry within metal–organic frameworks for gas storage and separation," 2021. doi: 10.1016/j.ccr.2021.213968.

[21] E. Bangian Tabrizi, M. Jalali, M. Houshmand, Inverse link prediction with graph convolutional networks for knowledge-preserving sparsification in cheminformatics, J. Big Data 12 (1) (2025), https://doi.org/10.1186/s40537-025-01220-8.

[22] M. Jalali, A.D.D. Wonanke, P. Friederich, C. Wöll, The black hole strategy: gravity-based representative sampling for frugal graph learning on metal–organic framework networks, J. Chem. Inf. Model (2025), https://doi.org/10.1021/acs.jcim.5c01518.

[23] M. Jalali, M. Tsotsalas, C. Wöll, MOFSocialNet: exploiting metal-organic framework relationships via social network analysis, Nanomaterials 12 (4) (Feb. 2022) 704, https://doi.org/10.3390/nano12040704.

[24] M. Jalali, A.D.D. Wonanke, C. Wöll, MOFGalaxyNet: a social network analysis for predicting guest accessibility in metal–organic frameworks utilizing graph convolutional networks, J. Chemin.. 15 (1) (2023), https://doi.org/10.1186/s13321-023-00764-2.

[25] C.E. Wilmer, et al., Large-scale screening of hypothetical metal–organic frameworks, Nat. Chem. 4 (2012) 83–89, https://doi.org/10.1038/nchem.1192.
[26] P.Z. Moghadam *et al.*, "Development of a Cambridge Structural Database Subset: A Collection of Metal-Organic Frameworks for Past, Present, and Future," 2017. doi: 10.1021/acs.chemmater.7b00441.
[27] S. Bashir, S.R. Chava, D. Yuan, S. Palakurthi, and J. Liu, Metal-organic frameworks and exemplified cytotoxicity evaluation. 2020. doi: 10.1016/B978-0-12-816984-1.00018-4.
[28] S.P. Collins, T.D. Daff, E. Jakubikova, T.K. Woo, Computational screening of porous coordination networks for adsorption and membrane-based applications, Adv. Funct. Mater. 26 (2016) 8639–8650, https://doi.org/10.1002/adfm.201603093.
[29] Y. Bao, R.L. Martin, C.M. Simon, M. Haranczyk, B. Smit, M.W. Deem, In silico discovery of high deliverable capacity metal–organic frameworks, J. Phys. Chem. C. 122 (2018) 1907–1916, https://doi.org/10.1021/acs.jpcc.7b10475.
[30] P.B. Jensen, M. Bialik, M. Haranczyk, Machine learning for predicting gas adsorption capacities of metal–organic frameworks, ACS Appl. Mater. Interfaces 11 (2019) 27415–27423, https://doi.org/10.1021/acsami.9b06845.
[31] K.M. Jablonka, S.M. Moosavi, M. Asgari, C. Ireland, L. Patiny, B. Smit, A data-driven perspective on the colours of metal-organic frameworks, Chem. Sci. 12 (10) (2021), https://doi.org/10.1039/d0sc05337f.
[32] M.J. García-Godoy, E. López-Camacho, J. García-Nieto, J. Del Ser, A.J. Nebro, J. F. Aldana-Montes, Bio-inspired optimization for the molecular docking problem: state of the art, recent results and perspectives, Appl. Soft Comput. 79 (2019), https://doi.org/10.1016/j.asoc.2019.03.044.
[33] M. Dorigo, M. Birattari, T. Stutzle, Ant colony optimization, IEEE Comput. Intell. Mag. 1 (2006) 28–39, https://doi.org/10.1109/MCI.2006.329691.
[34] E. Dalirinia, M. Jalali, M. Yaghoobi, H. Tabatabaee, Lotus effect optimization algorithm (LEA): A lotus nature-inspired algorithm for engineering design optimization, J. Supercomput 80 (2024) 761–799, https://doi.org/10.1007/s11227-023-05513-8.
[35] E. Dalirinia, M. Yaghoobi, H. Tabatabaee, S. Chandna, M. Jalali, Multimodal lotus effect algorithm for engineering optimization problems, Eng. Rep. 7 (4) (2025) e70137, https://doi.org/10.1002/eng2.70137.
[36] C.R. Groom, I.J. Bruno, M.P. Lightfoot, S.C. Ward, The Cambridge structural database, Acta Crystallogr B Struct. Sci. Cryst. Eng. Mater. 72 (2) (2016), https://doi.org/10.1107/S2052520616003954.
[37] V.F. Scalfani, V.D. Patel, A.M. Fernandez, Visualizing chemical space networks with RDKit and NetworkX, J. Chemin.. 14 (1) (2022), https://doi.org/10.1186/s13321-022-00664-x.
[38] S. Jaeger, S. Fulle, S. Turk, Mol2vec: Unsupervised Machine Learning Approach with Chemical Intuition, J. Chem. Inf. Model 58 (1) (2018), https://doi.org/10.1021/acs.jcim.7b00616.
[39] P. Willett, "Similarity methods in chemoinformatics," 2009. doi: 10.1002/aris.2009.1440430108.
[40] X.S. Yang, S. Deb, Cuckoo search via Lévy flights. 2009 World Congress on Nature and Biologically Inspired Computing, NABIC 2009 - Proceedings, 2009, https://doi.org/10.1109/NABIC.2009.5393690.
[41] N. Rakhshani, N.H. Nemati, A.R. Saadatabadi, S.K. Sadrnezhaad, Fabrication and evaluation of controlled release of doxorubicin loaded UiO-66-NH2 metal organic frameworks, Int. J. Eng. Trans. B Appl. 34 (8) (2021), https://doi.org/10.5829/ije.2021.34.08b.08.
[42] M.X. Wu and Y.W. Yang, "Metal–Organic Framework (MOF)-Based Drug/Cargo Delivery and Cancer Therapy," 2017. doi: 10.1002/adma.201606134.
[43] H. Bunzen, "Chemical Stability of Metal-organic Frameworks for Applications in Drug Delivery," 2021. doi: 10.1002/cnma.202100226.
[44] T.M. Shami, A.A. El-Saleh, M. Alswaitti, Q. Al-Tashi, M.A. Summakieh, S. Mirjalili, Particle swarm optimization: a comprehensive survey, IEEE Access 10 (2022), https://doi.org/10.1109/ACCESS.2022.3142859.
[45] S.M. Moosavi, et al., Understanding the diversity of the metal-organic framework ecosystem, Nat. Commun. 11 (1) (2020), https://doi.org/10.1038/s41467-020-17755-8.
[46] J. Bergstra, Y. Bengio, Random search for hyper-parameter optimization, J. Mach. Learn. Res. 13 (2012).
[47] S.M. Lundberg, S.I. Lee, A unified approach to interpreting model predictions. in Advances in Neural Information Processing Systems, 2017.
[48] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2) (2010), https://doi.org/10.1002/jcc.21334.