C4: EVIDENCE-AWARE SEARCH FOR LLM-GUIDED MATERIALS DISCOVERY

Anonymous authors
Paper under double-blind review

ABSTRACT

Historical results guide candidate generation in materials discovery, but a limited evaluation budget leaves many directions in the vast composition and structure space rarely or never tried. Search must therefore use successful experience while distinguishing insufficient observation from low potential. We introduce C4, a framework that coordinates four complementary tools: Open expands candidate support, Element identifies promising chemical directions, Explore probes underexplored alternatives, and Refine develops promising parents. C4 combines evaluated outcomes and shared visit records as search evidence, drawing on the exploration–exploitation principle of considering observed returns together with the evidence behind them. An LLM proposes tool-call proportions and branch-scoring weights from structured feedback. The Harness turns these proposals into budgeted calls and concrete substitution choices. Across fourteen multi-property tasks and three seeds with Qwen3-8B, C4 averages 448.33 new feasible compositions, compared with 105.67 for LLEMA, and leads all five baselines in each seed's total. Component studies and parent reuse illustrate the roles of the tools and their coordination. The advantage also extends to GPT-4o and longer search, while knowledge-guided cases show how additional context redirects the same search process.

1 INTRODUCTION

Discovering crystalline materials with several desired properties is central to advances in energy, electronics, and structural applications. A useful candidate must satisfy interacting requirements on composition, structure, stability, and function. Experimental exploration and physics-based computation provide essential evidence, but the number of possible candidates makes exhaustive evaluation impractical. Machine-learning property predictors accelerate screening (Xie & Grossman, 2018; Choudhary & DeCost, 2021), while generative models increasingly propose structures directly (Antunes et al., 2024). Large language models (LLMs) extend this approach by connecting task descriptions, chemical knowledge, and evaluation feedback within an iterative discovery process.

Recent frameworks use LLMs to modify materials, evolve candidate populations, or retrieve relevant examples (Jia et al., 2024; Gan et al., 2025; Abhyankar et al., 2026; Ding et al., 2024). Memory preserves successful candidates and useful failures to guide later proposals, as in broader reflective and experience-based agents (Shinn et al., 2023; Ouyang et al., 2025). However, an affordable search can evaluate only a small fraction of the possible compositions and substitutions. Many directions remain untried or poorly sampled, making their potential difficult to judge. Memory makes observed outcomes reusable, but storing them does not by itself determine which underexplored actions should receive the next attempts. LLEMA can already propose beyond previous compounds (Abhyankar et al., 2026). The remaining question is how to use limited experience without overlooking promising alternatives that lack observations.

Here, *search evidence* means evaluated outcomes together with the visits associated with attempted actions. Property feedback indicates how well a parent meets the task, while substitution outcomes identify useful chemical directions. Visit counts show how much exploration supports those judgments. A poorly sampled direction differs from one repeatedly tried without success, even when both have few discoveries. Connecting these records allows search to pursue promising alternatives without distributing its budget indiscriminately across all unknown materials.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

We introduce C4, which coordinates Open, Element, Explore, and Refine within one generation–evaluation loop. Open expands the candidate support, Element identifies promising substitution directions, Explore probes underexplored alternatives, and Refine focuses on promising parents. For a structured attempt, the system records the source and target elements, visits to the selected substitution, and the evaluated outcome. These records are shared across parents and tools. The Harness presents this evidence in structured prompts, and the LLM proposes both tool-call proportions and branch-scoring weights. The Harness translates these proposals into calls and substitutions within the remaining budget, then makes the resulting parents and observations available to later decisions.

Across fourteen multi-constraint tasks and three search seeds, C4 averages 4.24 times as many new feasible compositions as LLEMA and 72.2% more than MCTS. It leads all five baselines in each seed's total. Component studies, GPT-4o transfer, knowledge-guided search, and longer trajectories further examine the benefits of coordinated search. Our contributions are threefold:

- **Evidence-aware search organization.** Evaluated substitution outcomes and shared visit records connect observed promise to insufficiently explored alternatives.
- **Complementary search tools.** Open proposals, chemical direction selection, visit-aware exploration, and parent refinement operate on shared parents and evidence.
- **Budget-aware coordination.** The Harness combines LLM-proposed tool allocation and branch scoring with executable selection and scheduling rules.

2 Related Work

LLM-guided materials discovery. LLM-based frameworks improve proposals through explicit edits, evolutionary selection, and retrieval. LLMatDesign uses self-reflection to guide material edits (Jia et al., 2024). MatLLMSearch embeds proposals in population evolution (Gan et al., 2025), while LLEMA combines chemistry-informed evolution, surrogate evaluation, and success/failure memories (Abhyankar et al., 2026). MatExpert connects retrieval, transition reasoning, and structure generation (Ding et al., 2024). C4 builds on the use of knowledge and past candidates by sharing substitution outcomes and visits across tools, so observations guide both new proposals and explicit choices among available modifications.

Exploration–exploitation and structured search. UCB uses a visit-dependent bonus to balance observed returns and exploration (Auer et al., 2002). Simulated annealing and tree search allocate exploration through temperatures and branch visits (Kirkpatrick et al., 1983; Kocsis & Szepesvári, 2006). Materials applications include CASTING and CrystaLLM (Banik et al., 2023; Antunes et al., 2024). Tree of Thoughts and Language Agent Tree Search apply structured search to LLM reasoning and interaction (Yao et al., 2023a; Zhou et al., 2024). Adaptive operator selection and dynamic configuration adjust search behavior from feedback (Li et al., 2014; Adriaensen et al., 2022). C4 addresses a materials-specific coordination problem: new parents continually create substitution choices, and observations gathered by one tool must inform the others' selections.

Memory and execution control in LLM agents. Reflection, memory tiers, linked notes, and evolving playbooks make agent experience reusable (Shinn et al., 2023; Packer et al., 2023; Xu et al., 2025; Ouyang et al., 2025; Zhang et al., 2025b). Other systems organize reasoning, interaction, and tool workflows (Yao et al., 2023b; Wu et al., 2024; Yang et al., 2024; Zhang et al., 2025a). C4 connects these two concerns through a materials-search Harness. It supplies structured feedback to the LLM, executes the resulting allocation and scoring controls, and updates shared parents and visits after each attempt.

3 C4: Evidence-Aware Coordinated Search

3.1 Problem formulation

Task and material space. A task specifies admissible chemistry, property constraints, initial structures, and B candidate attempts. The material space $\mathcal{X}$ comprises crystal representations $X = (L, \mathbf{e}, \mathbf{r})$: lattice, atomic species, and fractional coordinates. The goal is to discover distinct feasible compositions.

Validity and feasible solutions. Representation validity $v(X)$ is one for an executable formula and parsable crystallographic information file (CIF) with recognized elements, and zero otherwise. Task feasibility $z(X)$ is one when all fixed property and composition predicates pass; invalid outputs or missing required values fail. The feasible structures $\mathcal{F} = \{X \in \mathcal{X} : v(X) = z(X) = 1\}$ define the composition solution set $\mathcal{C} = \{c(X) : X \in \mathcal{F}\}$, where $c(X)$ is the canonical reduced composition of the actual CIF. The target set is fixed but unknown.

State and actions. Before attempt t, state S_t contains parents P_t, outcome memory, a composition archive D_t, evaluated identities, search statistics, controls, scheduling counts, and remaining budget. Parents are valid structures available for modification, not necessarily feasible solutions. An action selects a tool and its intention: an unbound proposal or a legal parent substitution, realized as X_t by a shared LLM generator.

Reward and objective. The initial archive D_1 contains feasible initial compositions. An attempt earns

$$Y_t = \mathbf{1}\{v(X_t) = z(X_t) = 1,\ c(X_t) \notin D_t\}, \qquad J_B(\pi) = \mathbb{E}_\pi\left[\sum_{t=1}^{B} Y_t\right], \tag{1}$$

where π denotes the search policy. When $Y_t = 1$, $c(X_t)$ is added to D_t; otherwise the archive is unchanged. Invalid outputs receive zero without evaluating c. Every attempt consumes one budget unit, including invalid and repeated outputs.

3.2 Theoretical motivation: exploration under insufficient evidence

A direction may have few successes because it has low potential, or simply because it has rarely been tried. Classical upper confidence bound (UCB) selection separates observed return from observation count. After initially sampling each option, UCB1 selects the largest index

$$U_i(n) = \bar{r}_i(n) + \sqrt{2 \log n \,/\, n_i(n)}, \tag{2}$$

where n is the total number of observations, $n_i(n) > 0$ counts observations of option i, and $\bar{r}_i(n)$ is its mean reward (Auer et al., 2002). The first term favors observed success. The second gives less-observed options a larger bonus at the same total budget.

The material space is too large to explore every unknown direction equally. Knowledge and evaluations identify promising directions, while visits identify insufficiently tested alternatives. Together they form a revisable *search view*, not a known feasible subset of $\mathcal{X}$. Failures can change preferences without permanently excluding a region, and an infeasible structure may still provide a useful parent. C4 borrows this decision principle to combine observed promise with directed exploration through complementary tools, while keeping the discovery reward unchanged.

3.3 Overview of the shared search process

Figure 1 connects generation, evaluation, and subsequent decisions. The Harness supplies task constraints, memory examples, selected-action guidance, and structured search feedback to the LLM. Optional materials reference text adds context. The LLM returns candidates and control proposals. The Harness executes current actions, collects results, and applies validated controls next iteration.

Islands organize candidate histories and parent records, following LLEMA (Abhyankar et al., 2026). A parent can support several substitutions. A chemical direction can span parents and islands; a *family* denotes an element set, not an island or a crystal-structure class. The Harness enumerates allowed substitutions from P_t, removes unbuildable, unchanged, or previously evaluated intended structures, and merges duplicates into the legal catalogue $\mathcal{B}_t$. Open can propose beyond this catalogue. The other tools select target formulas from it, after which the common generator constructs fresh geometry and Evaluate tests the resulting structures. New parents and observations change later choices through the shared records, rather than through a prescribed order of tools.

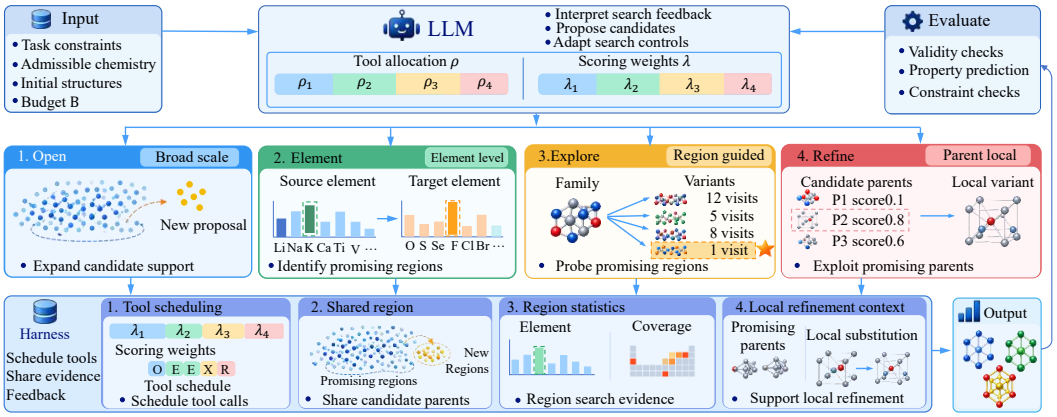


Figure 1: C4's LLM–Harness loop. Evaluated structures, element outcomes, and visits support both tool allocation and branch selection. The Harness turns allocation ρ into a call sequence and applies feature weights λ within structured tools. The control bars are schematic: ρ has four components and λ has seven.

3.4 Four complementary search tools

Finite-budget search needs both new starting points and ways to pursue them selectively. C4 separates four needs: expand the current search view, identify promising chemical directions, investigate their underexplored alternatives, and refine promising parents. The tools differ in how they choose an intention, while sharing generated parents and search evidence.

Open: expand candidate support. Open uses task-conditioned LLM proposals and memory without binding a particular parent or substitution. This leaves a route beyond the chemical directions represented in the current catalogue. Evaluate tests these proposals, so new structures can both yield discoveries and provide starting points for the structured tools. Open thus broadens and corrects the current search view through actual outcomes, rather than asking the LLM to specify all valuable materials in advance. Its valid outputs join the common parent pool beyond the examples shown in a single memory prompt. A later tool can therefore develop a new starting point even when its own proposal context differs.

Element: identify promising regions. Element asks which chemical substitutions deserve further search. It pools outcomes of structured attempts across parents, maintaining separate source-element and target-element statistics. It selects a source represented in $\mathcal{B}_t$, then an eligible target, using smoothed success estimates with an exploration bonus. The shared branch score in Section 3.5 selects the remaining parent and substitution scope. Because Explore and Refine also contribute structured outcomes, a successful direction found by either can guide a later Element call. The source and target summaries are separate marginals, allowing a target to reuse experience from several source elements. Element's guidance transfers across parents without requiring every complete substitution to accumulate its own success history.

Explore: probe promising regions. Explore searches for alternatives that warrant further attempts within the evolving search view. It ranks the whole legal catalogue using shared visit counts, parent potential, and search-exposure features under LLM-proposed weights. Visitation and potential answer different questions: how often an action type has been selected, and how close its parent is to meeting the task. The common score balances these signals, preserving opportunities to test poorly observed variations of promising chemistry instead of seeking every unvisited branch indiscriminately. Element contributes to this context through its generated parents, shared records, and subsequent control feedback. Independently scheduled Explore is not restricted to Element's most recently selected subset.

Refine: exploit promising parents. Refine concentrates on structures with the highest parent potential Q, defined by their smallest normalized margin over the task constraints. This favors parents

whose weakest required property is closest to, or furthest within, the acceptable range. Among their legal branches, Refine keeps the smallest available substitution scope: one site, half a source-element symmetry orbit, then a full orbit. The shared score breaks the remaining competition. Refine can use parents introduced by any tool. Its outcomes in turn inform later direction selection through the shared evidence. This is local chemical refinement; the common generator still constructs fresh geometry from the selected formula.

3.5 Shared evidence, control, and budgeted execution

Evidence that changes the next choice. Evaluations produce more than a discovery count. Valid structures and their property values enter the parent records and can appear in memory. Every structured selection increments a shared visit count, including selections that later fail. Its evaluated outcome also updates the source and target statistics used by Element. Elemental success credit requires a new feasible composition, agreement with the intended element family, and the recorded tool/route checks (Appendix A.5). The composition archive and tool-outcome summary retain the discovery reward in Eq. (1). These separate records distinguish a useful parent, a tried direction, and a new discovery. For example, an Explore substitution can supply a parent for Refine while also updating Element's estimate for the replacement element. A failed attempt still increases visitation, whereas a feasible composition earns reward only when first added to the archive.

Shared branch ranking. For each legal branch b, a coverage key $k(b)$ groups its source orbit, target element, substitution scope, route, and replacement-count category. Its visit count $n_{k(b)}$ is shared across the three structured tools. Eligible branches are ranked using the feature vector $\phi(b) \in [0,1]^7$:

$$H_\lambda(b) = \lambda^\top \phi(b), \qquad \phi_1(b) = \frac{1}{1+n_{k(b)}}, \quad \phi_2(b) = \frac{1+\tanh Q(p(b))}{2}, \tag{3}$$

where $p(b)$ is the parent and λ contains seven nonnegative weights summing to one. The first two features express insufficient visitation and observed parent promise. The remaining features describe family exposure, family–route exposure, validity, repeat avoidance, and island exposure (Appendix A.4). Element and Refine apply this score after their own filters; Explore applies it to $\mathcal{B}_t$. A visit by one tool consequently changes the priorities faced by the others.

Prompt-guided controls. The structured prompt presents current controls, scheduling counts, cumulative tool outcomes, coverage summaries, and previous validation feedback alongside task and memory context. The LLM proposes the four tool-call proportions ρ and seven branch-feature weights λ in the same response as its candidate descriptions, without an extra control call. The Harness checks permitted fields and normalized, finite values. Valid changes take effect next iteration; missing or invalid proposals leave the previous controls in force. Task predicates and rewards remain fixed. Appendix A.9 specifies the prompt blocks and response format.

Scheduling within budget. For each tool g, the Harness maintains a cumulative target q_g and scheduled-call count e_g, initially zero. Before each attempt it adds ρ_g to q_g, selects the largest deficit $q_g - e_g$, and increments the selected tool's count. The counters persist when ρ changes, so a proportion determines call order rather than requiring a second scheduling decision from the LLM. Each attempted candidate consumes one of the B slots, whatever its outcome. The search stops when these slots are exhausted and returns the new compositions in the archive with their evaluated structures.

4 Experiments

4.1 Experimental setup

Tasks and evaluation. We adopt LLEMA's fourteen multi-property task templates (Abhyankar et al., 2026), covering mechanical, electronic, dielectric, and related materials objectives. All local methods use Qwen3-8B (Yang et al., 2025) and the same frozen task-specific CGCNN and ALIGNN evaluators (Xie & Grossman, 2018; Choudhary & DeCost, 2021). We count new feasible compositions: distinct task-local reduced compositions with a valid structure satisfying the proxy

predicates, excluding feasible initial compositions. Appendix A.1 gives the executed predicates and evaluator mappings.

Search budget and baselines. The main budget is 100 iterations with two candidate slots each ($B = 200$); invalid and repeated attempts consume slots. We compare C4 with LLEMA (Abhyankar et al., 2026), a tree-based MCTS baseline (Kocsis & Szepesvári, 2006), simulated annealing (SA; Kirkpatrick et al., 1983), LLMatDesign (Jia et al., 2024), and MatLLMSearch (Gan et al., 2025), at seeds 17001/17002/17003. Open and the shared generator retain LLEMA's interfaces and island-organized candidate histories (Abhyankar et al., 2026). Appendix B specifies comparator implementations and budgets.

Evaluation metrics. For task j, $C_j(b)$ is the number of new feasible compositions added after b slots, i.e., the accumulated reward in Eq. (1). We report the mean and sample variance of this count across the three seeds. Discrete slot-AUC, $\sum_{b=1}^{B} C_j(b)$, also rewards earlier discovery and is reported as a supplementary measure.

4.2 Main comparison

We first compare material discovery under the same candidate budget. Table 1 summarizes the fourteen tasks across three search seeds.

Table 1: Qwen3-8B results on fourteen tasks, $B = 200$ slots per task. Each entry is the mean count of new feasible compositions with sample variance in superscript, $\bar{D}^{s^2_D}$, across three seeds. Bold marks positive task-wise maximum means, including ties.

Method	Stable wide-gap semicond. (06)	SAW/BAW acoustic substrates (04)	High-k dielectrics (03)	Solid-state electrolytes (05)	Piezo energy harvesters (08)	Transparent conductors (11)	Insulating dielectrics (10)
LLEMA	$15.0^{31.0}$	$2.3^{1.3}$	$6.3^{0.3}$	$12.3^{9.3}$	$21.0^{43.0}$	$7.0^{16.0}$	$9.0^{1.0}$
MCTS	$22.7^{30.3}$	$17.3^{32.3}$	$9.3^{2.3}$	$37.0^{291.0}$	$43.3^{286.3}$	$42.3^{149.3}$	$17.7^{100.3}$
SA	$1.7^{1.3}$	$\mathbf{19.3}^{44.3}$	$19.3^{65.3}$	$37.7^{9.3}$	$0.0^{0.0}$	$53.3^{52.3}$	$11.7^{10.3}$
LLMatDesign	$0.0^{0.0}$	$0.0^{0.0}$	$0.0^{0.0}$	$1.0^{0.0}$	$0.0^{0.0}$	$0.0^{0.0}$	$1.0^{0.0}$
MatLLMSearch	$0.7^{0.3}$	$2.7^{14.3}$	$4.7^{21.3}$	$16.0^{36.0}$	$2.3^{2.3}$	$0.3^{0.3}$	$0.0^{0.0}$
C4	$\mathbf{70.0}^{139.0}$	$18.7^{242.3}$	$\mathbf{22.7}^{4.3}$	$\mathbf{59.3}^{226.3}$	$\mathbf{55.3}^{60.3}$	$\mathbf{54.7}^{82.3}$	$\mathbf{30.7}^{65.3}$

Method	Photovoltaic absorbers (07)	Hard coating materials (01)	Hard, stiff ceramics (00)	Structural aerospace materials (02)	Acousto-optic hybrids (09)	Low-density aerospace structures (12)	Toxic-free perovskite oxides (13)
LLEMA	$9.3^{40.3}$	$0.0^{0.0}$	$3.7^{40.3}$	$10.0^{19.0}$	$9.7^{1.3}$	$0.0^{0.0}$	$0.0^{0.0}$
MCTS	$23.3^{25.3}$	$0.0^{0.0}$	$10.0^{151.0}$	$15.7^{1.3}$	$21.7^{30.3}$	$0.0^{0.0}$	$0.0^{0.0}$
SA	$0.7^{1.3}$	$0.0^{0.0}$	$4.0^{19.0}$	$13.0^{3.0}$	$0.0^{0.0}$	$1.0^{0.0}$	$0.0^{0.0}$
LLMatDesign	$0.0^{0.0}$	$0.0^{0.0}$	$0.0^{0.0}$	$0.0^{0.0}$	$0.0^{0.0}$	$0.0^{0.0}$	$0.0^{0.0}$
MatLLMSearch	$4.0^{7.0}$	$0.0^{0.0}$	$0.0^{0.0}$	$5.7^{6.3}$	$3.0^{1.0}$	$0.3^{0.3}$	$0.0^{0.0}$
C4	$\mathbf{37.0}^{144.0}$	$0.0^{0.0}$	$\mathbf{34.0}^{73.0}$	$\mathbf{27.3}^{30.3}$	$\mathbf{22.3}^{32.3}$	$\mathbf{16.0}^{109.0}$	$\mathbf{0.3}^{0.3}$

Total	LLEMA	MCTS	SA	LLMatDesign	MatLLMSearch	**C4**
All 14 tasks	$105.67^{524.33}$	$260.33^{201.33}$	$161.67^{200.33}$	$2.00^{0.00}$	$39.67^{202.33}$	$\mathbf{448.33}^{670.33}$

Substantial gains across the benchmark. C4's fourteen-task discovery total averages 448.33 across three seeds, compared with 105.67 for LLEMA, 260.33 for MCTS, and 161.67 for SA (Table 1): $4.24\times$ LLEMA, 72.2% more than MCTS, and 177.3% more than SA. It leads all five baselines in each seed, with totals 458/468/419. The improvement is broadly distributed: C4 exceeds both LLEMA and MCTS on thirteen task means and ties on the remaining task, while attaining the highest mean among all six methods on twelve tasks.

The gains span distinct property regimes rather than concentrating on one productive task. For wide-gap semiconductors, C4 averages 70 new compositions versus MCTS's 22.67. For hard, stiff ceramics, it reaches 34 versus 10. It also improves on MCTS from 37 to 59.33 for electrolytes and from 17.67 to 30.67 for insulating dielectrics. On low-density structural materials, C4 finds an average of 16 compositions where LLEMA and MCTS find none. Acoustic substrates are a close comparison, with 18.67 for C4 and 19.33 for SA. Overall, the benefit spans several property regimes, although its magnitude depends on the task.

4.3 Component contributions

Tool contributions. To isolate the impact of each search role, we perform ablation studies. The search roles make complementary contributions (Table 2). Removing broad generation (Without Open) reduces new compositions from 446 to 334 (25.1%), while removing parent-level refinement (Without Refine) reduces them to 289 (35.2%). Open supplies new search regions and parents, which Refine can develop into further feasible compositions. Open alone yields only 136 new compositions, confirming that its primary value lies in seeding subsequent refinement rather than producing final candidates directly.

Table 2: Component contributions: Qwen3-8B, all fourteen tasks, seed 505, $B = 200$. Fixed C4 freezes allocation and scoring (Appendix C).

Variant	Count	Slot-AUC
Open only	136	16,793
Without Open	334	36,331
Without Element	356	38,532
Without Explore	407	41,373
Open + Refine	204	19,285
Without Refine	289	28,093
Fixed C4	372	34,567
Full C4	**446**	**42,269**

Open+Refine removes both the direction-selection and visit-aware exploration channels supplied explicitly by Element and Explore. It yields 204 new compositions and 19,285 Slot-AUC, losses of 54.3% and 54.4% relative to Full C4. Removing only Element leaves Explore available and gives 356 compositions; removing only Explore leaves Element and gives 407. These smaller losses are consistent with the overlapping roles in Section 3.4. Element combines shared substitution outcomes with an exploration bonus, while Explore weighs visits and parent promise through the common score. Their interaction is mediated by shared evidence, not serial filtering. Open+Refine still explores through open proposals and local variations, but loses the two explicit channels connecting chemical direction selection with insufficient visitation. These results underscore that both direction selection and visitation awareness are critical for peak performance. Appendix C specifies the interventions.

Adaptive coordination. Having established the importance of individual tools, we next examine whether dynamic coordination further enhances performance. To assess how the tools are used, Fixed C4 keeps their allocation and scoring weights constant while retaining the same Harness and shared records. Full C4 improves on it by 19.9% in new compositions and 22.3% in Slot-AUC. This comparison supports adapting the two controls jointly, beyond choosing a fixed tool mixture. In other words, adjusting emphasis across search stages yields substantial gains over any static configuration.

Cross-tool cooperation. Beyond ablation, we investigate inter-tool dependencies through a detailed case study. We next examine how one tool supplies parents to another, using a separate long-run case on solid-state electrolytes (task05).

Of 300 structured calls in Figure 2(a), 212 reuse parents introduced by another tool and produce 104 new compositions. Open-origin parents support 130 later calls and 62 discoveries, in addition to Open's own direct output shown in panel (b). Reuse also runs between structured tools: Explore-origin parents support 11 discoveries in 24 Element calls and ten in 13 Refine calls. All nine cross-tool relations are productive in this selected case. Thus a tool's immediate yield captures only part of its role in the shared search. The later discoveries belong to the consuming calls and are not counted twice. This cross-fertilization reveals that even tools with modest direct outputs can serve as critical catalysts, highlighting the need for holistic evaluation of component contributions.

4.4 Knowledge-guided mechanisms, longer search, and model transfer

Knowledge and tool allocation. Does additional materials knowledge change how C4 allocates search, and can the same tools use it productively? We cross LLEMA/C4 with the absence or

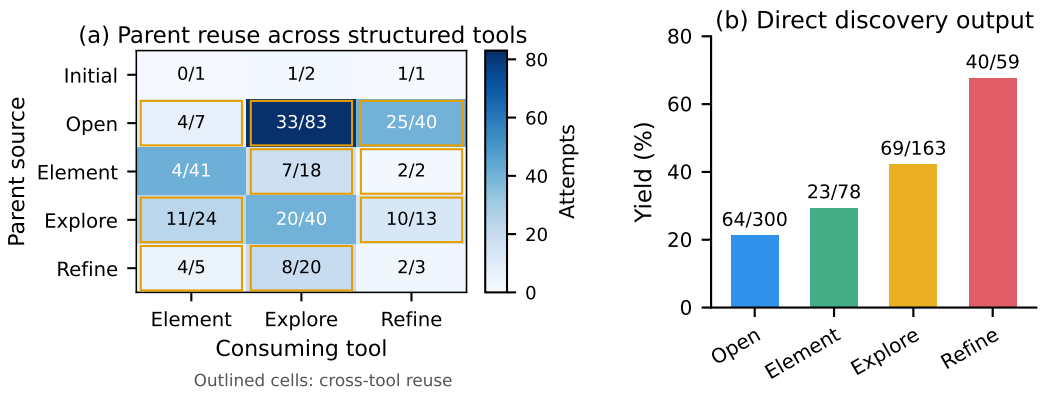


Figure 2: Long-run cooperation on task05 (Solid-State Electrolytes), Qwen3-8B, seed 17002, T=300 (B=600). (a) Parent sources (rows) and consuming tools (columns); color shows attempts, cells discoveries/attempts. (b) Direct discoveries per 100 calls, labeled discoveries/calls. Appendix C defines first-introduction attribution and duplicate handling.

presence of task-specific reference text from MatKG 1.4, which links compositions to properties and applications (Venugopal & Olivetti, 2024; 2023). The text is added to the generation and control context. Task predicates, tool definitions, and base instructions remain unchanged, with no instruction to increase or decrease any tool's share. Table 3 reports three successful cases whose retrieved references concern density and moduli, or dielectric and band-gap properties. These cases show that the existing framework can use external knowledge without redesigning its tools.

Table 3: Illustrative knowledge gains on three property-relevant tasks, Qwen3-8B, seed 17001, $B = 200$. Each task shares its founder across the four conditions. Complete paired results are in Appendix D.1.

Task	LLEMA	+MatKG	C4	+MatKG
02: Structural aerospace	15	14	22	38
03: High-k dielectric	6	5	22	35
10: Insulating dielectric	10	3	26	49

On task10, adding knowledge raises new compositions from 26 to 49 (88.5%). Figure 3 shows the accompanying shift from Open toward Element and Explore. Their combined calls increase from 58 to 104, and their direct discoveries from eight to 34. Knowledge thus helps the search concentrate on productive chemical directions within the same legal substitution rules.

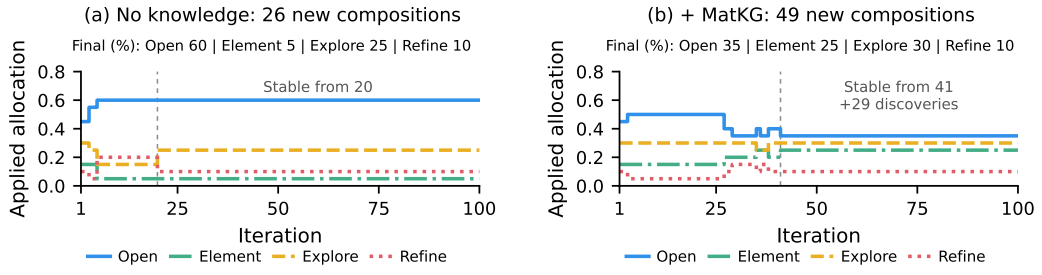


Figure 3: Knowledge-responsive allocation on task10, Qwen3-8B, seed 17001, $B = 200$. Applied proportions use identical axes. Final proportions follow Open/Element/Explore/Refine order. In the knowledge condition, 29 further compositions are discovered after the last adjustment.

Stable allocation does not end exploration. After the last adjustment at iteration 41, the knowledge-guided search keeps the same tool proportions but discovers 29 further compositions. Parents, visits, and branch priorities continue to change within that division of work. The two conditions settle at different mixtures, showing that stable execution can coexist with a response to context. The

intervention changes the information available to the coupled search, not the allowable material space or a predefined scheduling target.

Longer budgets continue to produce new feasible compositions. To test whether the advantage persists beyond early search, we continue all fourteen tasks at seed 17002 to 300 iterations. C4 reaches 468, 968, and 1,452 compositions at iterations 100, 200, and 300 (Figure 4). The three equal windows add 468, 500, and 484 compositions, maintaining discovery well beyond the main budget. At iteration 300, MCTS reaches 918 and LLEMA 331. C4 exceeds MCTS by 58.2% in count and 64.0% in Slot-AUC.

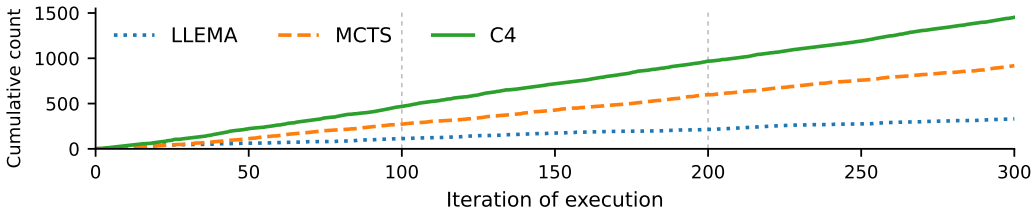


Figure 4: All fourteen continuation tasks, Qwen3-8B, seed 17002. Curves sum new compositions with iteration = slot/2 and $B = 600$. Appendix D.2 lists every endpoint.

The benefit persists with GPT-4o. We test whether coordination remains useful with another generator by repeating the fourteen-task comparison with GPT-4o (OpenAI, 2024), keeping the search methods, evaluators, and candidate budgets fixed (Appendix D.3). C4 finds 756 compositions, compared with 565 for LLEMA and 542 for MCTS (Table 4). The respective gains are 33.8% and 39.5% in count, and 35.2% and 39.1% in Slot-AUC.

LLEMA also improves markedly over its same-seed Qwen3-8B total of 80, consistent with memory-guided evolution benefiting from a more capable generator. Better generation can broaden proposals and improve the use of memory, but changing the backbone does not increase the observations supplied by a fixed search budget or specify how untried actions share evidence. Many possible substitutions therefore remain sparsely assessed. C4 complements stronger generation by organizing those observations across parents and tools, retaining attention to both observed promise and insufficient visitation. The gain with GPT-4o supports the value of this search organization beyond the particular local generator.

Table 4: GPT-4o transfer: all fourteen tasks, seed 17001, $B = 200$. The search methods and evaluators are fixed. Corrected C4 endpoints are specified in Appendix D.3.

Method	Count	Slot-AUC
LLEMA	565	55,106
MCTS	542	53,580
C4	**756**	**74,507**

Scope and limitations. Collectively, the experiments support the coordinated search framework with both Qwen3-8B and GPT-4o, demonstrating initial model generality. The present model coverage is admittedly limited, so broader backbone comparisons across more diverse architectures remain to be examined. The discovered compositions satisfy fixed proxy predicates derived from existing databases; their thermodynamic stability and experimental utility require further physical validation via additional computational screening or synthesis efforts.

5 Conclusion

C4 coordinates proposing, exploiting promising directions, exploring alternatives, and refining parents to make finite search evidence actionable. Classical exploration compensation motivates explicit attention to insufficiently visited choices, implemented through complementary tools and shared feedback. The fourteen-task three-seed study shows broad gains in new feasible compositions over five baselines, with the advantage extending to GPT-4o and longer budgets. Knowledge cases show responsive, stable multi-tool coordination. Choosing how to act on experience is therefore an important part of budget-limited materials discovery, alongside choosing what experience to store.

Ethics Statement

This study concerns computational materials discovery and reports candidates assessed by proxy models. It does not establish their safety, synthesizability, or suitability for deployment. Any experimental follow-up should include appropriate hazard assessment and physical validation.

AI Use Statement

AI-assisted tools contributed to analysis, experimental code, manuscript text, LaTeX, figures, and verification scripts. The human authors remain responsible for the submitted work.

Reproducibility Statement

Appendix A specifies the task predicates, evaluator mappings, tool scores, control rules, generation settings, and prompt templates. Appendix B describes the comparison protocol and statistical conventions; Appendix C defines the component interventions. Appendix D.1 details knowledge retrieval and the paired comparisons, followed by the continuation and GPT-4o protocols. Complete task-level results accompany these descriptions.

References

Nikhil Abhyankar, Sanchit Kabra, Saaketh Desai, and Chandan K. Reddy. LLEMA: Evolutionary search with LLMs for multi-objective materials discovery. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=TIqzhBvCNB.

Steven Adriaensen, André Biedenkapp, Gresa Shala, Noor Awad, Theresa Eimer, Marius Lindauer, and Frank Hutter. Automated dynamic algorithm configuration. *Journal of Artificial Intelligence Research*, 75, 2022. doi: 10.1613/jair.1.13922.

Luis M. Antunes, Keith T. Butler, and Ricardo Grau-Crespo. Crystal structure generation with autoregressive large language modeling. *Nature Communications*, 15:10570, 2024. doi: 10.1038/s41467-024-54639-7.

Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47(2–3):235–256, 2002. doi: 10.1023/A:1013689704352.

Suvo Banik, Troy Loefller, Sukriti Manna, Henry Chan, Srilok Srinivasan, Pierre Darancet, Alexander Hexemer, and Subramanian K. R. S. Sankaranarayanan. A continuous action space tree search for inverse design (CASTING) framework for materials discovery. *npj Computational Materials*, 9:177, 2023. doi: 10.1038/s41524-023-01128-y.

Kamal Choudhary and Brian DeCost. Atomistic line graph neural network for improved materials property predictions. *npj Computational Materials*, 7:185, 2021. doi: 10.1038/s41524-021-00650-1.

Qianggang Ding, Santiago Miret, and Bang Liu. MatExpert: Decomposing materials discovery by mimicking human experts. *arXiv preprint arXiv:2410.21317*, 2024. URL https://arxiv.org/abs/2410.21317.

Jingru Gan, Peichen Zhong, Yuanqi Du, Yanqiao Zhu, Chenru Duan, Haorui Wang, Daniel Schwalbe-Koda, Carla P. Gomes, Kristin A. Persson, and Wei Wang. MatLLMSearch: Crystal structure discovery with evolution-guided large language models. *arXiv preprint arXiv:2502.20933*, 2025.

Shuyi Jia, Chao Zhang, and Victor Fung. LLMatDesign: Autonomous materials discovery with large language models. *arXiv preprint arXiv:2406.13163*, 2024. URL https://arxiv.org/abs/2406.13163.

S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220 (4598):671–680, 1983. doi: 10.1126/science.220.4598.671.

Levente Kocsis and Csaba Szepesvári. Bandit based monte-carlo planning. In *Machine Learning: ECML 2006*, volume 4212 of *Lecture Notes in Computer Science*, pp. 282–293, 2006. doi: 10.1007/11871842_29.

Ke Li, Alvaro Fialho, Sam Kwong, and Qingfu Zhang. Adaptive operator selection with bandits for a multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions on Evolutionary Computation*, 18(1):114–130, 2014. doi: 10.1109/TEVC.2013.2239648.

OpenAI. Hello GPT-4o, 2024. URL https://openai.com/index/hello-gpt-4o/.

Siru Ouyang, Jun Yan, I-Hung Hsu, Yanfei Chen, Ke Jiang, Zifeng Wang, Rujun Han, Long T. Le, Samira Daruki, Xiangru Tang, Vishy Tirumalashetty, George Lee, Mahsan Rofouei, Hangfei Lin, Jiawei Han, Chen-Yu Lee, and Tomas Pfister. ReasoningBank: Scaling agent self-evolving with reasoning memory. *arXiv preprint arXiv:2509.25140*, 2025. URL https://arxiv.org/abs/2509.25140.

Charles Packer, Sarah Wooders, Kevin Lin, Vivian Fang, Shishir G. Patil, Ion Stoica, and Joseph E. Gonzalez. MemGPT: Towards LLMs as operating systems. *arXiv preprint arXiv:2310.08560*, 2023. URL https://arxiv.org/abs/2310.08560.

Noah Shinn, Federico Cassano, Edward Berman, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. Reflexion: Language agents with verbal reinforcement learning. *arXiv preprint arXiv:2303.11366*, 2023. URL https://arxiv.org/abs/2303.11366.

V. Venugopal and Elsa A. Olivetti. MatKG 1.4, 2023. URL https://zenodo.org/records/10144972.

V. Venugopal and Elsa A. Olivetti. MatKG: An autonomously generated knowledge graph in material science. *Scientific Data*, 11:217, 2024. doi: 10.1038/s41597-024-03039-z.

Qingyun Wu, Gagan Bansal, Jieyu Zhang, Yiran Wu, Beibin Li, Erkang Zhu, Li Jiang, Xiaoyun Zhang, Shaokun Zhang, Jiale Liu, Ahmed Hassan Awadallah, Ryen W. White, Doug Burger, and Chi Wang. AutoGen: Enabling next-gen LLM applications via multi-agent conversation. In *Conference on Language Modeling*, 2024. URL https://arxiv.org/abs/2308.08155.

Tian Xie and Jeffrey C. Grossman. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical Review Letters*, 120(14):145301, 2018. doi: 10.1103/PhysRevLett.120.145301.

Wujiang Xu, Zujie Liang, Kai Mei, Hang Gao, Juntao Tan, and Yongfeng Zhang. A-MEM: Agentic memory for LLM agents. *arXiv preprint arXiv:2502.12110*, 2025. URL https://arxiv.org/abs/2502.12110.

An Yang, Anfeng Li, Baosong Yang, et al. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*, 2025.

John Yang, Carlos E. Jimenez, Alexander Wettig, Kilian Lieret, Shunyu Yao, Karthik Narasimhan, and Ofir Press. SWE-agent: Agent-computer interfaces enable automated software engineering. *arXiv preprint arXiv:2405.15793*, 2024. URL https://arxiv.org/abs/2405.15793.

Shunyu Yao, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L. Griffiths, Yuan Cao, and Karthik Narasimhan. Tree of thoughts: Deliberate problem solving with large language models. In *Advances in Neural Information Processing Systems*, volume 36, pp. 11809–11822, 2023a.

Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. ReAct: Synergizing reasoning and acting in language models. In *International Conference on Learning Representations*, 2023b. URL https://arxiv.org/abs/2210.03629.

Jiayi Zhang, Jinyu Xiang, Zhaoyang Yu, Fengwei Teng, Xionghui Chen, Jiaqi Chen, Mingchen Zhuge, Xin Cheng, Sirui Hong, Jinlin Wang, Bingnan Zheng, Bang Liu, Yuyu Luo, and Chenglin Wu. AFlow: Automating agentic workflow generation. In *International Conference on Learning Representations*, 2025a. URL https://arxiv.org/abs/2410.10762.

Qizheng Zhang, Changran Hu, Shubhangi Upasani, Boyuan Ma, Fenglu Hong, Vamsidhar Kamanuru, Jay Rainton, Chen Wu, Mengmeng Ji, Hanchen Li, Urmish Thakker, James Zou, and Kunle Olukotun. Agentic context engineering: Evolving contexts for self-improving language models. *arXiv preprint arXiv:2510.04618*, 2025b. URL https://arxiv.org/abs/2510.04618.

Andy Zhou, Kai Yan, Michal Shlapentokh-Rothman, Haohan Wang, and Yu-Xiong Wang. Language agent tree search unifies reasoning, acting, and planning in language models. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 62138–62160, 2024.

A Task Definitions and Supporting Method Details

This appendix supplies the task predicates and numerical details supporting the tools and shared-evidence mechanism in Section 3.

A.1 Tasks, units, and surrogate identities

The learned evaluators use ALIGNN (Choudhary & DeCost, 2021) and CGCNN (Xie & Grossman, 2018) through the frozen local adapters. In Table 5, K denotes bulk modulus, G_s shear modulus, E_g band gap, E_f formation energy per atom, ϵ_x the static dielectric x component, $\epsilon_{\max}$ the configured maximum dielectric coefficient, $d_{\max}$ the maximum piezoelectric coefficient, ρ_m mass density, and σ electrical conductivity. A candidate's element set is $\mathcal{E}(X)$. Numeric predicates are inclusive. Missing or nonfinite required values cannot satisfy the nonnegative-potential feasibility test. The frozen

Table 5: Executed proxy-feasibility predicates, exported from the frozen scorer (all inequalities inclusive). Units: K, G_s in GPa; E_g in eV; E_f in eV/atom; $d_{\max}$ in pC/N; ρ_m in g/cm^3; σ in S/m; dielectric coefficients dimensionless. Categorical sets are defined below.

ID	Task label	Executed predicates
00	Ceramics	$100 \le K \le 300$; $60 \le G_s \le 200$
01	Hard coating	$K \ge 200$; $E_f \le -1$; $E_g \ge 3$
02	Structural aerospace	$\rho_m \le 5$; $K \ge 100$; $G_s \ge 40$
03	High-k dielectric	$10 \le \epsilon_x \le 90$; $2.5 \le E_g \le 6.5$
04	Acoustic substrate	$25 \le G_s \le 150$; $3.7 \le \epsilon_x \le 95$
05	Electrolyte	$E_f \le -1$; $E_g \ge 2$; at least one of Li, Na, K, Mg, Ca, Al
06	Wide-bandgap	$E_g \ge 2.5$; $E_f \le -1$
07	Photovoltaic	$0.7 \le E_g \le 2$; $E_f \le 0$; $\mathcal{E}(X) \subseteq \mathcal{E}_{\text{allow}}$; $\mathcal{E}(X) \cap \mathcal{E}_{\text{ban}} = \varnothing$
08	Piezoelectric	$d_{\max} \ge 8$; $10 \le \epsilon_{\max} \le 8000$
09	Acousto-optic	$2 \le d_{\max} \le 9$; $8 \le \epsilon_{\max} \le 95$
10	Insulating dielectric	$E_g \ge 2.5$; $\epsilon_x \ge 8$
11	Transparent conductor	$E_g \ge 1.5$; $500 \le \sigma \le 30000$
12	Low-density aerospace	$\rho_m \le 3.5$; $65 \le G_s \le 195$
13	Toxic-free perovskite	$E_g \ge 2$; $90 \le K \le 135$; $\mathcal{E}(X) \subseteq \mathcal{E}_{\text{allow}}$; $\mathcal{E}(X) \cap \mathcal{E}_{\text{ban}} = \varnothing$

composition allow-list is

> H, Li, B, C, O, F, Na, Mg, Al, Si, P, S, Cl, K, Ca, Sc, Ti, V, Mn, Fe, Co, Ni, Cu, Zn, Ga, Rb, Sr, Y, Zr, Nb, La, Ce, Nd.

This list defines $\mathcal{E}_{\text{allow}}$ operationally. The banned set is $\mathcal{E}_{\text{ban}} = \{\text{Hg}, \text{Cd}, \text{Pb}, \text{Tl}, \text{As}, \text{Be}, \text{Se}, \text{U}, \text{Th}, \text{Sb}\}$. The electrolyte condition accepts at least one of Li, Na, K, Mg, Ca, Al. All categorical checks use the candidate species (or its parsed formula when necessary).

Tables 5 and 6 specify the executed property/composition predicates and their nine evaluator routes. Application labels add no extra geometric or thermodynamic tests.

A.2 Parent potential and memory

For a predicted property x_p with threshold τ_p or interval $[l_p, u_p]$, the normalized margin is

$$q_p(x_p) = \begin{cases} (x_p - \tau_p)/\max(|\tau_p|, 1), & x_p \ge \tau_p \text{ is required,} \\ (\tau_p - x_p)/\max(|\tau_p|, 1), & x_p \le \tau_p \text{ is required,} \\ \min(x_p - l_p, u_p - x_p)/\max((u_p - l_p)/2, 10^{-6}), & x_p \in [l_p, u_p] \text{ is required.} \end{cases} \tag{4}$$

Margins are clipped to $[-5, 5]$, with -1 for a missing or nonfinite value. A categorical margin is 0.05 if all categorical constraints pass and -1 otherwise. The parent potential is $Q(X) = \min_p q_p(x_p)$. A valid structure is feasible when $Q(X) \ge 0$. This score ranks observed constraint satisfaction, not a calibrated probability of future discovery.

Table 6: Active evaluator mappings in the executed protocol. Density is calculated from the structure; the other routes use frozen learned models or their wrappers.

Property	Backend	Frozen evaluator
K	ALIGNN	Pretrained bulk-modulus model
G_s	ALIGNN	Pretrained shear-modulus model
E_g	CGCNN	Pretrained band-gap model
E_f	CGCNN	Pretrained formation-energy model
ϵ_x	ALIGNN	Pretrained static dielectric x-component model
$d_{\max}$	ALIGNN	Pretrained maximum piezoelectric-coefficient model
$\epsilon_{\max}$	ALIGNN	Pretrained maximum dielectric-coefficient model
ρ_m	Jarvis Atoms	Mass divided by crystal-cell volume
σ	ALIGNN wrapper	Power factor divided by squared Seebeck coefficient, output in S/m

Memory ranks examples by a separate score. It returns -10 when all constraints fail, zero when only some fail, and otherwise the mean property contribution capped above at ten. An interval contributes $\min(x - l, u - x)$; a one-sided constraint contributes its direction-consistent margin. Conductivity values and thresholds are scaled by 10^4 for this score. The prompt initially requests three successes and three failures from a selected island. Action-conditioned memory then incorporates selected parents and their property records (Appendix A.9). The full parent pool also includes usable island-history and cluster candidates, deduplicated by island and exact structure identity. It is not restricted to the examples shown in the prompt.

The initial structure is stored in each island. Its valid identity is recorded as evaluated, and its composition enters D_1 if feasible. Initialization does not count as a paid discovery or as an attempted elemental substitution. Parents retain their evaluated potential. Where that value is unavailable, the selectors use the stored memory score as a fallback.

A.3 Legal substitutions and realization

A structural representation has a finite atom count $m \geq 1$, a nonsingular lattice $L \in \mathbb{R}^{3\times3}$, m supported species, and fractional coordinates in $[0, 1)^{m\times3}$. A structured intention specifies a parent and island, source and target elements, source symmetry orbit, replacement scope, selected sites, and resulting formula. Allowed target elements come from a fixed substitution graph augmented by chemical-family and periodic-row neighbors, capped at six targets per source.

Symmetry detection uses positional tolerance 10^{-3} and angular tolerance five degrees. SINGLE substitutes one site, HALF substitutes $\lceil m_o/2 \rceil$ sites of an orbit containing $m_o \geq 2$ sites, and FULL substitutes its entire orbit. At most 32 symmetry-distinct HALF configurations are retained. Enumeration removes composition-illegal, unbuildable, unchanged, and previously evaluated intended structures. Within a parent, equivalent intended structures retain their first representative, with FULL enumerated before SINGLE. Consequently, coincident scopes can retain a FULL label even for a one-site change.

The resulting $\mathcal{B}_t$ is a catalogue of legal intentions, not verified feasible materials. The shared generator receives the target formula and produces fresh lattice and site coordinates. Task constraints guide the evolution context and remain fixed in evaluation. Element-family and recorded route/tool checks determine internal substitution credit. These checks do not independently verify a geometric edit of the parent. A valid, feasible, compositionally new realization still receives discovery credit even if it differs from the intended substitution.

A.4 Additional branch features and shared visits

Equation (3) gives the common score and its visitation and parent-potential features. The coverage key combines source-orbit information, target element, scope, route, and the replacement-count/orbit-size category. Parent identity is not a separate key field, so different branches can share a visit count. Counts increase when a structured action is selected, before its outcome is known.

For branch b, let $F(b)$ denote its element-family set, $r(b)$ its substitution route, $\ell(b)$ its island, and V the evaluated structure-identity set. Let N_F and $N_{F,r}$ be the recorded content-evaluation counts for a family and a family–route. For the eligible branch list $\mathcal{B}$, define $I_\ell = \sum_{b\in\mathcal{B}:\ell(b)=\ell} n_{k(b)}$. The five remaining features are

$$\begin{aligned} &\phi_3(b) = (1+N_{F(b)})^{-1}, \qquad \phi_4(b) = (1+N_{F(b),r(b)})^{-1}, \\ &\phi_5(b) = 1, \qquad \phi_6(b) = \mathbf{1}\{\mathrm{id}(b) \notin V\}, \qquad \phi_7(b) = (1+I_{\ell(b)})^{-1}. \end{aligned} \tag{5}$$

Here $\mathrm{id}(b)$ is the intended structure identity. The island exposure sums branch rows, including repeated coverage keys, and is recomputed after a tool's filtering. All seven features lie in $[0,1]$. Family and route counts are distinct from Element's source/target outcome statistics. Validity and repeat-avoidance features may be constant after legal filtering. A reproducible tie-break selects among scores within 10^{-12} of the maximum.

A.5 Element outcome scores

For source or target role r and element e, let $n_e^{(r)}$ count completed structured attempts and $s_e^{(r)}$ sum their internal success credit. With global structured-attempt totals n_0, s_0, set $m_0 = (1+s_0)/(2+n_0)$ and

$$\alpha_e^{(r)} = 2m_0 + s_e^{(r)}, \qquad \beta_e^{(r)} = 2(1-m_0) + n_e^{(r)} - s_e^{(r)}. \tag{6}$$

Writing α, β for these parameters, Element uses

$$\mu = \frac{\alpha}{\alpha+\beta}, \qquad V = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}, \qquad U = \mu + 1.25\sqrt{V\log(n_0+2)}. \tag{7}$$

The score combines a globally smoothed mean and a variance-based exploration bonus. Source and target statistics are separate global marginals: target evidence is pooled across sources, not estimated separately for every source–target pair. Element selects the highest-scoring eligible source, then its highest-scoring eligible target, before applying the shared branch score. All counts start at zero, giving finite initial scores.

A structured attempt increments its selected path and source/target counts regardless of success. Internal credit is Y_t gated by agreement of the realized element family and recorded route/tool contract. The contract also requires scheduled and realized tools to match and no adapter fallback. A route label is inherited from the intention, so this is not an independent test of generated site geometry or target stoichiometry. Tool-level outcomes instead assign Y_t to the scheduled tool. Open has no selected elemental path; its evaluated parents can support later structured attempts, but its own calls do not increment these marginal counts.

A.6 Control validation and execution details

Refine first keeps branches from parents with maximum potential and then the smallest available scope in the order SINGLE, HALF, FULL. The common score resolves the remaining choices. The cumulative deficit scheduler follows Section 3.5, retaining its targets and call counts across changes of allocation and continuation.

An allocation proposal must contain all four tools, values in $[0,1]$, and a unit sum within 10^{-6}. Scoring proposals use the seven supported features. Missing feature weights become zero; nonfinite or negative values and nonpositive total mass are rejected, and valid weights are normalized. Missing or invalid proposals preserve the current controls. A structured tool with no executable action may recover through Open within the same charged slot, with scheduled and realized tool identities retained separately. No new candidate opportunity is created by that recovery.

A.7 Generation settings

Qwen3-8B uses bfloat16, maximum input length 16,384, maximum output length 3,072, temperature 0.7, top-p 0.8, and top-k 20, with thinking disabled and stripped. The main study uses seeds 17001/17002/17003 and the component study seed 505. Each run uses five islands, random island selection, score-ranked memory, and two candidate slots per iteration, without timed island resets. Initial allocation in Open/Element/Explore/Refine order is $(.45, .15, .30, .10)$, and the initial branch score uses coverage only. Continuation preserves the search state, random streams, visits, outcome statistics, and scheduling counters. GPT-4o settings are specified in Appendix D.3.

A.8 Illustration of tool cooperation

Figure 5 illustrates the changing search view. Open reaches beyond currently favored chemistry. Element identifies promising directions using accumulated results, Explore investigates underexplored alternatives in that context, and Refine develops promising parents. Solid and dashed arrows indicate greater and lower attention. The low-priority regions represent limited current interest, not established infeasibility. The regions are not islands: islands organize stored histories, whereas chemical directions can overlap across those histories.



Figure 5: Conceptual illustration of complementary search roles. The four tools share parents and evidence, and their calls can be interleaved. The landscape depicts revisable search preferences rather than measured distances or a partition into known feasible and infeasible materials.

A.9 Harness organization and prompts

The Harness combines structured prompts with fixed selection and accounting rules. Table 7 summarizes the full cycle. The following prompt templates specify the inputs, instructions, and output fields for each stage. The quoted blocks are templates, not sampled model responses.

Table 7: Harness responsibilities and information flow.

Component	Input and operation
Parent and action records	Collect valid parents across islands, construct legal substitutions, and select current actions using the committed controls.
Memory context	Combine success/failure examples with the selected parents' formulas and property values.
Evidence summary	Present current controls, scheduling targets and counts, tool outcomes, under-visited actions, catalogue summaries, and previous validation feedback.
Evolution and control	Ask the LLM for candidate descriptions and proposed allocation/scoring controls in one response.
Control validation	Accept permitted, valid controls for the next iteration; otherwise preserve the previous values.
Structure generation	Realize each selected formula with the shared structure prompt.
Evaluation and feedback	Evaluate the actual structure and update parents, visits, elemental results, tool outcomes, and composition novelty.
Budget accounting	Charge each candidate attempt, including invalid or repeated outputs, and stop at the allotted budget.

Evolution context. The prompt contains the materials task and constraints, six selected chemistry rules, and examples with available property values. The initial island memory supplies up to three successes and three failures. After adding action-selected parents and removing duplicates, at most ten examples per outcome group enter the action-conditioned context. The rule library includes elemental substitution, stoichiometric and oxidation-state changes, prototype and coordination changes, functional analogs, and property-guided design. Its advanced variant emphasizes the more elaborate rules but retains the same task, memory, and output fields.

> **Role:** You are an expert materials discovery assistant with deep knowledge of crystal chemistry and materials design.
> **Task and inputs:** Generate materials for the supplied task and constraints, using the selected rules and success/failure examples.
> **Candidate output:** Return the requested number of candidates, each with a compound formula, the rules used, and a chemical justification.
> **Action order:** Return one candidate for each scheduled action in its given order. Do not merge actions or add candidates.

Tool-specific action guidance. The Harness appends each selected action's family, parent formula, source and target elements, and substitution scope when applicable. Open has no bound parent. Each tool receives the corresponding directive below. They guide the evolution response; the actual structured branch has already been selected by the Harness.

> **Open.** Use open evolution. Explore any chemically plausible new family that satisfies the task without repeating known formulas.
>
> **Element.** Realize the selected source/target intent, then use coverage below the chosen roles. Do not substitute an unrelated family.
>
> **Explore.** Realize the selected symmetry-aware coverage intent. Prefer the least-covered legal cell and keep the requested element roles and scope while producing a new canonical composition.
>
> **Refine.** Perform a conservative local refinement around the selected parent. Keep its useful chemical motif, realize the requested substitution and scope, and avoid unrelated global redesign.

Control prompt and response. The appended control block overrides the base prompt's candidate-list-only instruction. The evidence view contains the current ρ and λ, cumulative targets and calls, successes and failures for each scheduled tool, up to twelve least-visited recorded coverage keys, and summaries of available actions and their visitation. The LLM sees these summaries rather than every stored structure.

> The evidence in the following view is read-only and contains only information available before this iteration. The strategy proposal itself is active after Harness validation.
>
> Choose a continuous allocation using only the supplied past evidence. Do not define rewards, reward code, thresholds, gates, or evaluator logic.
>
> For every permitted control dimension, independently choose its value. The current value is a prior, not a required answer. Retaining it is allowed when supported by evidence. The proposal controls the next iteration only after the Harness validates it.
>
> Return one object containing (1) a strategy proposal with all four tool allocations and the seven nonnegative branch-feature weights, and (2) the requested candidate list.
>
> Optional hypotheses, memory queries, rationale, and confidence are descriptive fields, not additional execution permissions.

The current iteration's actions are selected before this response. The allocation and score proposal therefore affects only later actions. The next iteration receives the results of the current candidates along with the accepted controls. Allocation determines tool frequency and call order; feature weights determine branch ranking. Neither vector changes model parameters or the task's evaluation criteria.

Shared structure prompt. An Open slot uses its proposed formula, while a structured slot uses the target formula derived from the selected substitution. The default structure prompt receives this formula, not parent coordinates or the complete search history. The knowledge intervention additionally supplies its frozen reference block.

> **Role and goal:** You are a materials discovery assistant. Generate a realistic crystallographic configuration for the given compound.
> **Input:** Compound formula.
> **Required output:** Return only one JSON object with exactly four fields: formula, lattice, species, and coordinates.
> **Representation rules:** The lattice consists of three vectors of three numbers in angstroms. Species and coordinates have equal lengths, each coordinate is a fractional three-vector in $[0, 1)^3$, and the formula agrees with the atom counts. Use valid numeric values and no surrounding prose or markup.

After generation, executable checks establish representation validity and the fixed task predicates. The prompt's request to preserve a motif or substitution scope is not a geometric guarantee. The recorded family and route/tool checks govern elemental credit as specified in Appendix A.5, while all new feasible compositions retain their discovery reward.

B Complete Local Results and Comparison Protocol

B.1 Three-seed statistics and scoring

Table 8 gives every task's three new-composition counts, and Table 9 gives supplementary AUC statistics. For values x_1, x_2, x_3, the mean is $\bar{x} = \sum_s x_s/3$ and the sample variance is $s_x^2 = \sum_s (x_s - \bar{x})^2/2$. Task-total statistics are computed from the three within-seed sums, preserving task covariation. Table 1 rounds task statistics to one decimal after computing them from the unrounded counts.

The comparison uses the following initialization and scoring conventions. LLEMA, C4, and MCTS share a founder within each task and seed. Task08 uses its inherited founder at seed 17001 and a legal ZnO founder at seeds 17002/17003. Technically prepared founders are used for LLMatDesign and MatLLMSearch on tasks04/06/07/08/09/10/11/13, and for SA on tasks08/09/11/13; preparation checks structural validity and finite evaluator outputs, not later search performance. Failed SA runs on tasks08/09 at all three seeds and MatLLMSearch runs on task12 at seeds 17002/17003 receive zero counts; their AUCs are unavailable. The completed seed-17001 MatLLMSearch/task12 run resumes its original prefix, with a final count of one and AUC 109. In Tables 8–9, † marks prepared founders, $*$ the task08 seed-specific founder convention, ‡ a resumed run, and § zero-scored failures.

Table 8: All 252 new-composition counts. Each cell lists seeds 17001/17002/17003. Scoring conventions are defined in Appendix B.

Task	LLEMA	MCTS	SA	LLMatDesign	MatLLMSearch	C4
00	0/0/11	1/5/24	1/2/9	0/0/0	0/0/0	42/35/25
01	0/0/0	0/0/0	0/0/0	0/0/0	0/0/0	0/0/0
02	15/8/7	17/15/15	14/11/14	0/0/0	3/8/6	22/27/33
03	6/6/7	11/8/9	28/18/12	0/0/0	2/2/10	22/21/25
04	3/3/1	11/22/19	21/25/12	0/0/0†	1/0/7†	35/17/4
05	9/13/15	32/56/23	41/35/37	1/1/1	16/10/22	58/45/75
06	9/20/16	17/28/23	1/3/1	0/0/0†	1/1/0†	57/80/73
07	2/13/13	18/28/24	0/0/2	0/0/0†	2/3/7†	49/37/25
08	14/22/27*	62/29/39*	0/0/0†§	0/0/0†	4/2/1†	53/64/49*
09	9/9/11	18/19/28	0/0/0†§	0/0/0†	3/4/2†	27/24/16
10	10/8/9	29/10/14	14/13/8	1/1/1†	0/0/0†	26/26/40
11	3/11/7	29/53/45	57/45/58†	0/0/0†	0/0/1†	56/63/45
12	0/0/0	0/0/0	1/1/1	0/0/0	1/0/0‡§	11/28/9
13	0/0/0	0/0/0	0/0/0†	0/0/0†	0/0/0†	0/1/0

Failure-score and initialization annotations are defined in Appendix B; the seed-17001 MatLLMSearch/task12 result retains its timeout-recovery annotation.

Table 9: Supplementary slot-AUC, mean / sample variance across the same three seeds. Mean is in 10^3 AUC units; variance is in 10^6 squared AUC units.

Task	LLEMA	MCTS	SA	LMD	MLS	C4
00	0.39 / 0.46	0.98 / 2.44	0.16 / 0.04	0.00 / 0.00	0.00 / 0.00	2.69 / 2.48
01	0.00 / 0.00	0.00 / 0.00	0.00 / 0.00	0.00 / 0.00	0.00 / 0.00	0.00 / 0.00
02	1.15 / 0.18	1.79 / 0.12	1.63 / 0.07	0.00 / 0.00	0.85 / 0.11	2.50 / 0.62
03	0.97 / 0.04	1.10 / 0.04	2.10 / 0.42	0.00 / 0.00	0.55 / 0.21	2.31 / 0.13
04	0.28 / 0.01	1.77 / 0.43	1.82 / 0.04	0.00 / 0.00†	0.41 / 0.39†	1.58 / 1.63
05	1.35 / 0.30	3.26 / 1.89	3.92 / 0.00	0.20 / 0.00	1.99 / 0.72	5.90 / 0.94
06	1.76 / 0.38	2.03 / 0.25	0.03 / 0.00	0.00 / 0.00†	0.01 / 0.00†	6.20 / 4.04
07	0.93 / 0.31	2.67 / 0.42	0.04 / 0.00	0.00 / 0.00†	0.46 / 0.09†	3.70 / 2.51
08	2.32 / 0.22*	4.34 / 4.58*	—†§	0.00 / 0.00†	0.36 / 0.02†	4.88 / 1.73*
09	1.40 / 0.01	2.08 / 0.06	—†§	0.00 / 0.00†	0.35 / 0.01†	2.22 / 0.06
10	1.38 / 0.06	1.73 / 1.03	1.38 / 0.06	0.20 / 0.00†	0.00 / 0.00†	3.32 / 0.91
11	0.72 / 0.14	3.30 / 0.92	4.27 / 0.65†	0.00 / 0.00†	0.01 / 0.00†	4.70 / 0.44
12	0.00 / 0.00	0.00 / 0.00	0.20 / 0.00	0.00 / 0.00	—‡§	1.57 / 0.35
13	0.00 / 0.00	0.00 / 0.00	0.00 / 0.00†	0.00 / 0.00†	0.00 / 0.00†	0.03 / 0.00

LMD/MLS denote LLMatDesign/MatLLMSearch. A dash is an unavailable three-seed AUC statistic, not zero. Marks are defined in Appendix B.

B.2 Algorithm implementations and evaluation budgets

LLEMA retains its evolutionary proposal logic, island organization, and success/failure memory (Abhyankar et al., 2026). Our tree-based comparator adapts MCTS to the same generator and evaluators, using binary discovery reward and a tree over source element, target element, substitution scope, route, and realized structure identity (Kocsis & Szepesvári, 2006). Each iteration contains one open proposal and one element-directed proposal, with seeded alternation of their positions. Tree backup credits the realized material. Its PUCT coefficient is 1.25; progressive widening uses coefficient 2 and exponent 0.5, with chemical priors enabled.

SA uses five local-substitution chains with stochastic acceptance and a temperature schedule fixed over 600 candidate slots (Kirkpatrick et al., 1983); the main $B = 200$ endpoint is its prefix. LLMatDesign uses explicit edits and subsequent reflection (Jia et al., 2024). MatLLMSearch uses two-parent, five-child generation and population selection (Gan et al., 2025). The latter two use their search operators in local, unrelaxed, shared-evaluator adaptations.

The methods receive equal candidate-evaluation opportunities, while native model usage differs. LLEMA, C4, and MCTS use 301 logical model calls per fresh $B = 200$ case: one initialization, 100 evolution, and 200 structure-generation calls. SA uses 200 geometry calls; LLMatDesign uses 400 edit/reflection calls; MatLLMSearch uses 40 five-child batch calls. The respective output caps are 3,072 tokens for the backbone, 1,024 per LLMatDesign role, and 8,192 per MatLLMSearch batch.

C Component Interventions and Task Results

The seed-505 study uses Qwen3-8B, all fourteen tasks, and $B = 200$ per task. Full C4 initializes allocation at $(0.45, 0.30, 0.15, 0.10)$ in Open/Explore/Element/Refine order and auxiliary weights at coverage-only. It adapts both controls; Fixed C4 freezes both while retaining the same harness, memory, and feedback records. Open only fixes allocation to $(1, 0, 0, 0)$. Removing Refine transfers its mass to Open; removing Explore transfers its mass to Element; removing Element transfers its mass to Explore; removing Open transfers its mass to Explore. Open+Refine transfers Explore's mass to Open and Element's mass to Refine, giving initial allocation $(0.75, 0, 0, 0.25)$. These transfers apply initially and to accepted proposals; the non-fixed variants retain adaptation.

Element and Explore provide complementary direction-selection and visit-aware exploration channels. Element identifies substitution directions using pooled outcomes and applies an elemental filter to its own calls. Independently scheduled Explore ranks the current legal branches using shared coverage and auxiliary scores; it is not restricted to the last subset selected by Element. Their interaction is

mediated by shared parents, counts, and controller feedback. The visit-aware selector also remains inside Refine. Accordingly, removing scheduled Explore calls preserves both shared branch ranking and Element's direction selection, while Open+Refine removes both scheduled channels.

Table 10 gives the available task-level results; the Open+Refine and Without Element aggregates appear in Table 2.

Figure 2 uses a separate task05 trajectory at seed 17002 and $B = 600$, selected for its productive cross-tool reuse among the fourteen completed C4 continuations. A parent's source is the tool that first validly introduced its exact structure identity; the frozen founder is labeled Initial. Ten parent-use slots have identical structures previously generated by multiple tools, and all retain first-introduction attribution. The matrix covers 300 structured calls and the bars cover all 600 calls, which yield 196 new compositions. Heatmap discoveries are the consuming calls' rewards, not additional rewards assigned to the source tool.

Table 10: Task-level new-composition counts for the single-tool removals and fixed/full controls, seed 505, $B = 200$. The Open+Refine aggregate appears in Table 2.

Task	Open only	No Open	No Explore	No Refine	Fixed	Full
00	10	0	0	8	14	47
01	0	0	0	0	0	0
02	11	21	18	31	24	29
03	12	29	20	16	23	25
04	2	28	16	8	12	7
05	16	40	65	53	83	55
06	30	7	69	9	10	81
07	13	9	24	26	27	38
08	9	113	41	50	70	44
09	9	22	20	12	21	10
10	11	19	22	33	6	29
11	13	43	93	37	71	70
12	0	1	16	6	3	11
13	0	2	3	0	8	0
Total	136	334	407	289	372	446

D KNOWLEDGE, CONTINUATION, AND MODEL-TRANSFER DETAILS

D.1 KNOWLEDGE RETRIEVAL AND PAIRED EVALUATION

The conditions are L0 (LLEMA), L1 (LLEMA+MatKG), C0 (C4), and C1 (C4+MatKG). Both enhanced methods receive the same frozen task-specific MatKG 1.4 reference text (Venugopal & Olivetti, 2024; 2023), supplying literature-derived guidance rather than task-oracle observations. The original cohort comprises all thirteen tasks except task08, with four conditions per task at seed 17001 and $B = 200$. A separate task08 supplement uses one frozen legal ZnO founder across its four conditions, distinct from the inherited seed-17001 main-table founder. Table 12 reports the complete paired results.

Retrieval and injection. Four fixed property/application queries per task retrieve composition-linked records from MatKG 1.4. For each query, the top twenty matches are ranked by the published occurrence count, with deterministic ties. Duplicate triples are removed and the four lists are interleaved round-robin. The frozen visible prefix contains at most twelve records and 1,024 tokens. No new retrieval occurs during search. The same text is inserted before the output-format instructions of the evolution and structure-generation prompts. Initialization is unchanged. Every task shares its founder and fixed evaluators across the four conditions. The separate task08 supplement is not substituted into the original main-table results.

Property relevance and allocation response. The registered task02 queries are density, bulk modulus, shear modulus, and aerospace. Task03 queries dielectric, band gap, bandgap, and high

k; task10 queries band gap, dielectric, insulator, and insulating. These match the corresponding task predicates, while the same property-based retrieval design also covers the other tasks. The three main-text cases illustrate successful use of this guidance, rather than an exclusive domain of applicability. Across all thirteen original-founder tasks, LLEMA changes from 66 to 112 new compositions and C4 from 405 to 448; the separate task08 values are 18/9 and 83/127. Knowledge effects depend on the task, as the complete paired results show.

Table 11 connects the task10 allocation curves to actual tool outcomes. Refine's mean allocation decreases from 11.4% to 9.2%, so the increased share belongs to Explore and Element rather than every more specific tool. The knowledge intervention adds a reference block without an allocation directive; its observed response reflects the subsequent coupled search trajectory. The legal substitution rules and task predicates remain unchanged. Without knowledge, task10's applied allocation changes at iterations 3, 5, and 20; its largest ℓ_1 change is 0.40 at iteration 5. The final mixture $(.60, .05, .25, .10)$ in Open/Element/Explore/Refine order is used at iterations 20–100. With MatKG, changes occur at iterations 3, 27, 29, 35, 36, 38, and 41. The largest changes are 0.20, at iterations 27, 35, and 38, followed by stable execution. The final mixture $(.35, .25, .30, .10)$ is used at iterations 41–100. These stable segments characterize the finite, coupled search trajectories shown in Figure 3.

Table 11: Actual tool use in the task10 case. Mean allocation is the committed target; calls/new compositions count realized tool outcomes, not independent tool effects.

	Mean allocation (%)		Calls / new compositions	
	No knowledge	+MatKG	No knowledge	+MatKG
Open	59.60	39.10	119/14	78/9
Explore	23.60	29.85	47/7	60/23
Element	5.40	21.85	11/1	44/11
Refine	11.40	9.20	23/4	18/6

Table 12: Complete paired knowledge results: new-composition count / discrete slot-AUC. All thirteen original tasks and the separate four-condition task08 legal-founder supplement are retained. Qwen3-8B, seed 17001, $B = 200$.

Task / stratum	LLEMA	LLEMA+MatKG	C4	C4+MatKG
00	0 / 0	27 / 3,221	42 / 4,336	45 / 5,798
01	0 / 0	0 / 0	0 / 0	0 / 0
02	15 / 1,635	14 / 1,294	22 / 2,611	38 / 4,334
03	6 / 856	5 / 854	22 / 2,243	35 / 2,982
04	3 / 312	8 / 1,027	35 / 2,777	12 / 1,125
05	9 / 1,014	12 / 1,694	58 / 5,361	54 / 5,378
06	9 / 1,143	6 / 777	57 / 4,933	60 / 5,261
07	2 / 296	13 / 1,231	49 / 5,265	46 / 4,636
09	9 / 1,316	5 / 846	27 / 2,349	25 / 2,468
10	10 / 1,614	3 / 516	26 / 2,751	49 / 4,402
11	3 / 284	14 / 1,608	56 / 5,127	61 / 5,493
12	0 / 0	5 / 686	11 / 1,218	20 / 2,284
13	0 / 0	0 / 0	0 / 0	3 / 195
Original 13	66 / 8,470	112 / 13,754	405 / 38,971	448 / 44,356
08 (legal ZnO)	18 / 2,346	9 / 721	83 / 7,040	127 / 17,153

D.2 Second-seed same-trajectory continuation

The complete continuation cohort comprises all fourteen tasks and all three methods at seed 17002: 42 trajectories with 600 candidate slots each. Every $B = 600$ endpoint continues its own $B = 200$ search state, retaining the parent pool, memory, random streams, statistics, and archive. Figure 4 sums all fourteen slot-wise curves per method and displays slot index divided by two as iteration. At iterations 100/200/300, the totals are 468/968/1,452 for C4, 273/597/918 for MCTS, and

113/214/331 for LLEMA. Their three disjoint 100-iteration windows yield 468/500/484, 273/324/321, and 113/101/117 new compositions, respectively. Table 13 gives every final task result; total discrete slot-AUCs are 428,810/261,518/101,900 for C4/MCTS/LLEMA, obtained by summing all 600 cumulative slot counts. The $B = 200$ prefix counts and areas agree with the corresponding seed-17002 main results for every method and task.

Table 13: Same-trajectory long search: all fourteen tasks at iteration 300 ($B = 600$), seed 17002. Cells are new-composition count / slot-AUC.

Task	LLEMA	MCTS	C4
00	0 / 0	31 / 7,197	85 / 27,797
01	1 / 61	0 / 0	0 / 0
02	53 / 11,237	32 / 10,384	96 / 28,006
03	9 / 3,939	25 / 7,652	78 / 22,656
04	14 / 4,008	52 / 15,957	46 / 14,192
05	24 / 8,218	189 / 54,378	196 / 50,641
06	50 / 17,157	128 / 31,408	225 / 67,401
07	23 / 8,920	81 / 25,164	105 / 33,560
08	70 / 21,136	73 / 25,856	177 / 55,865
09	12 / 5,999	66 / 19,390	73 / 21,844
10	17 / 6,188	48 / 11,832	60 / 19,209
11	54 / 14,263	189 / 51,366	217 / 61,694
12	4 / 774	4 / 934	91 / 24,869
13	0 / 0	0 / 0	3 / 1,076
Total	331 / 101,900	918 / 261,518	1,452 / 428,810

D.3 GPT-4o COMPARISON PROTOCOL

The transfer comparison covers 42 endpoints: fourteen tasks, LLEMA/C4/MCTS, seed 17001, and $B = 200$. The protocol uses a common task-specific founder across the three methods. Table 4 reports the aggregate, and Table 14 gives every task. The same-seed Qwen3-8B totals are 80/245/458 for LLEMA/MCTS/C4; backbone comparisons also retain their recorded sampling settings.

The authors corrected the C4 endpoints for tasks00/02/10 after reporting API-credit interruptions, identifying EF1R1 as the restart label for these runs. The corrected counts are 84/41/35 and slot-AUCs are 9,094/4,731/3,747. The remaining eleven C4 endpoints are retained, including task12 at 18/1,612, giving totals of 756 new compositions and 74,507 slot-AUC.

Generation settings. GPT-4o uses temperature 0.8, top-p 1, maximum completion length 3,072, and zero frequency/presence penalties. The request preserves the original user prompt without adding a system instruction or JSON-response constraint. Evolution and structure-generation requests have separate seed streams derived from search seed 17001 and their call indices. Continuation preserves their offsets.

Table 14: GPT-4o task results, seed 17001, $B = 200$: new-composition count / slot-AUC.

Task	LLEMA	MCTS	C4
00	96 / 9,436	57 / 5,840	84 / 9,094
01	1 / 200	1 / 103	0 / 0
02	50 / 5,076	31 / 3,696	41 / 4,731
03	19 / 2,109	15 / 1,075	23 / 2,754
04	19 / 2,224	21 / 2,736	21 / 2,401
05	71 / 6,385	70 / 6,662	126 / 11,143
06	34 / 3,343	69 / 5,670	82 / 8,000
07	47 / 4,454	52 / 4,630	86 / 8,311
08	99 / 8,495	79 / 8,143	105 / 10,049
09	18 / 1,976	29 / 3,270	23 / 2,599
10	12 / 1,381	29 / 2,812	35 / 3,747
11	73 / 7,258	70 / 6,903	111 / 10,008
12	24 / 2,557	14 / 1,814	18 / 1,612
13	2 / 212	5 / 226	1 / 58