nature machine intelligence

Supplementary information

https://doi.org/10.1038/s42256-026-01298-6

A collaborative agent with two lightweight synergistic models for autonomous crystal materials research

In the format provided by the authors and unedited

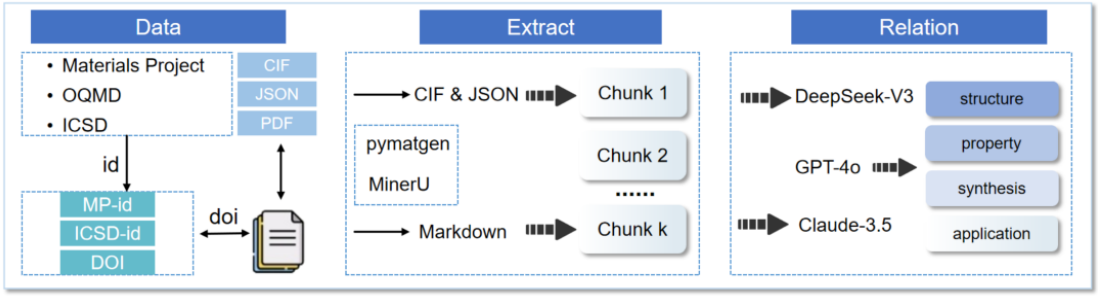

Fig. S1 Schematic overview of the heterogeneous data processing pipeline. Structured CIF and JSON records were processed separately from literature-derived text. Literature chunks were jointly classified by DeepSeek-V3, GPT-4o and Claude-3.5 using majority voting into structure, property, synthesis and application categories. Detailed identifier resolution and data-linkage procedures are described in the Methods.

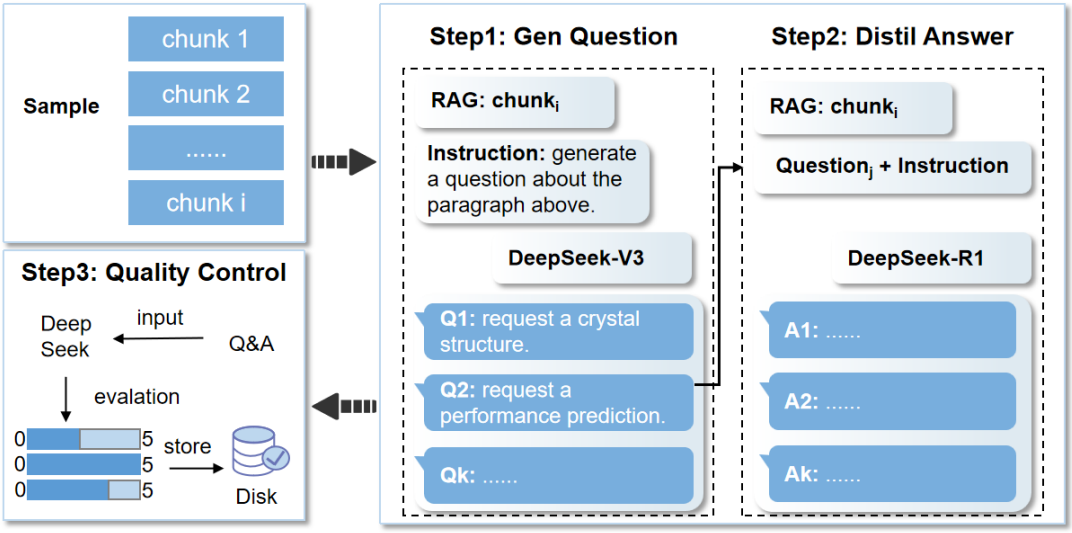

Fig. S2 The generation-distillation strategy. Detailed workflow for constructing the dataset, showing how raw data pairs are processed through generation and distillation steps to produce high-quality instruction-tuning samples.

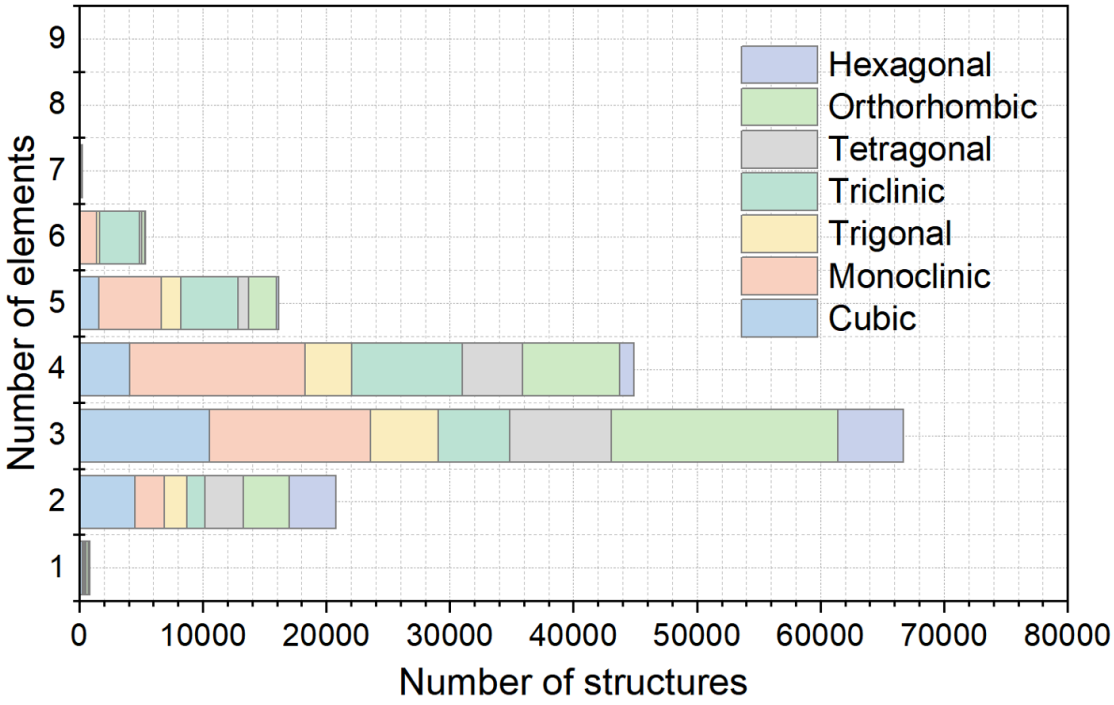

Fig. S3 Distribution of the crystal structures underlying Mat-252K-SFT. The stacked bars show the number of deduplicated crystal structures associated with the Mat-252K-SFT instruction–response pairs, grouped by the number of constituent elements and coloured by crystal system. The counts represent underlying crystal structures rather than the total number of instruction–response pairs.

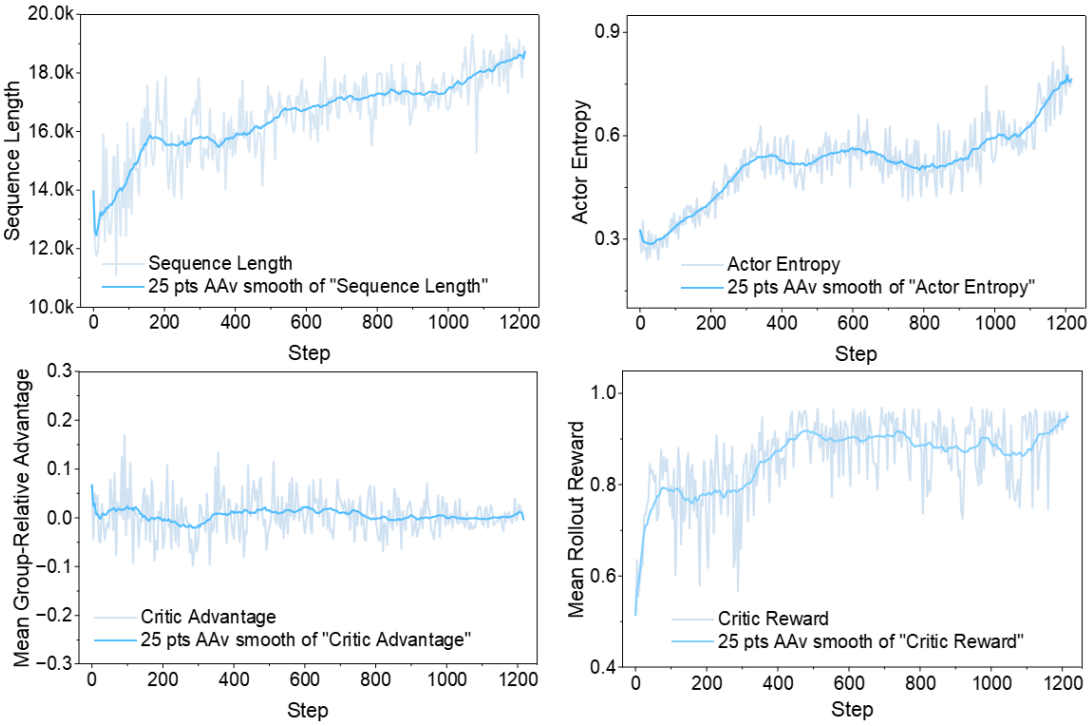

Fig. S4 Training dynamics of Mat-T1 reinforcement learning. Evolution of four logged quantities during DAPO training: mean sequence length, Actor entropy, mean rollout reward and mean group-relative advantage. DAPO uses group-relative advantage estimation without a separately trained critic or value model. The near-zero mean advantage reflects within-group normalization and is therefore not interpreted as an independent indicator of training stability. Light curves show the raw values, and dark curves show the corresponding 25-point adjacent averages.

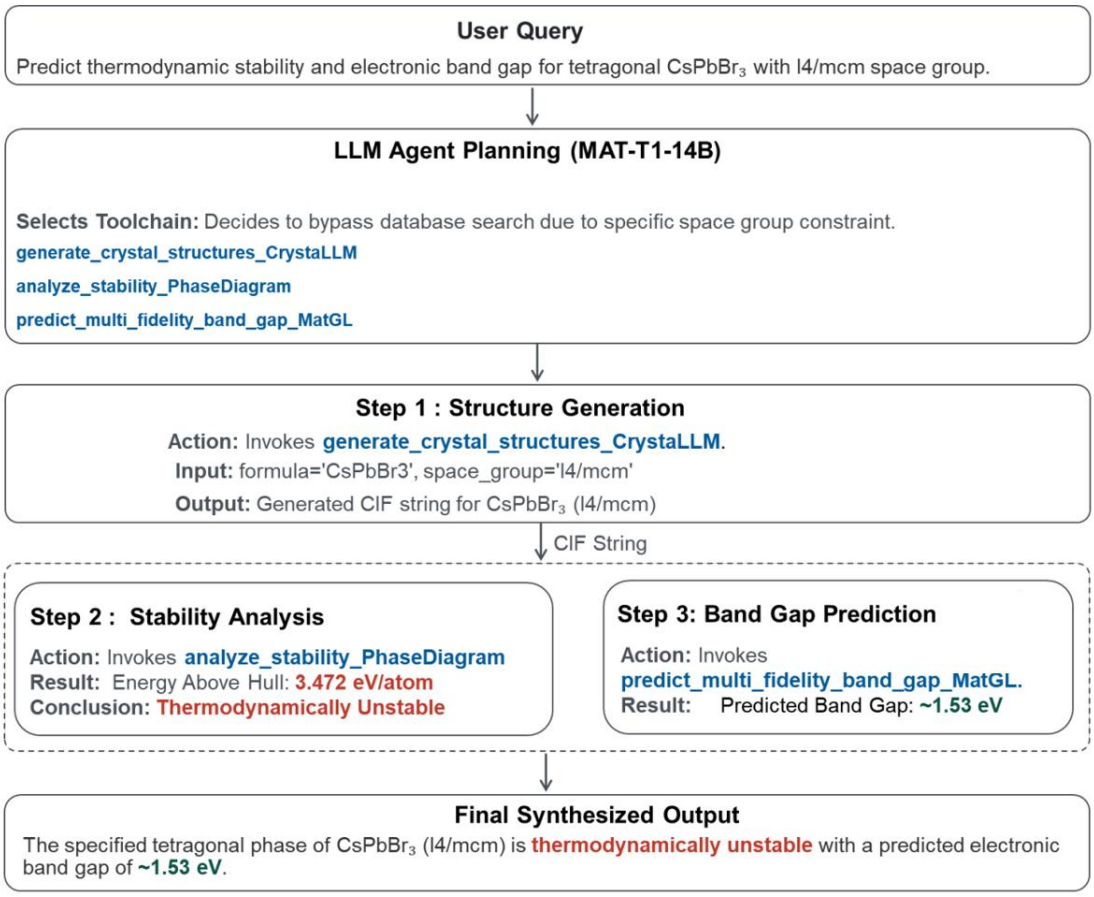

Fig. S5 Schematic of Mat-T1 intent analysis and tool orchestration. Visualization of the executive workflow, showing how Mat-T1 parses user queries, selects tools from the Mat-MCP server, and executes commands.

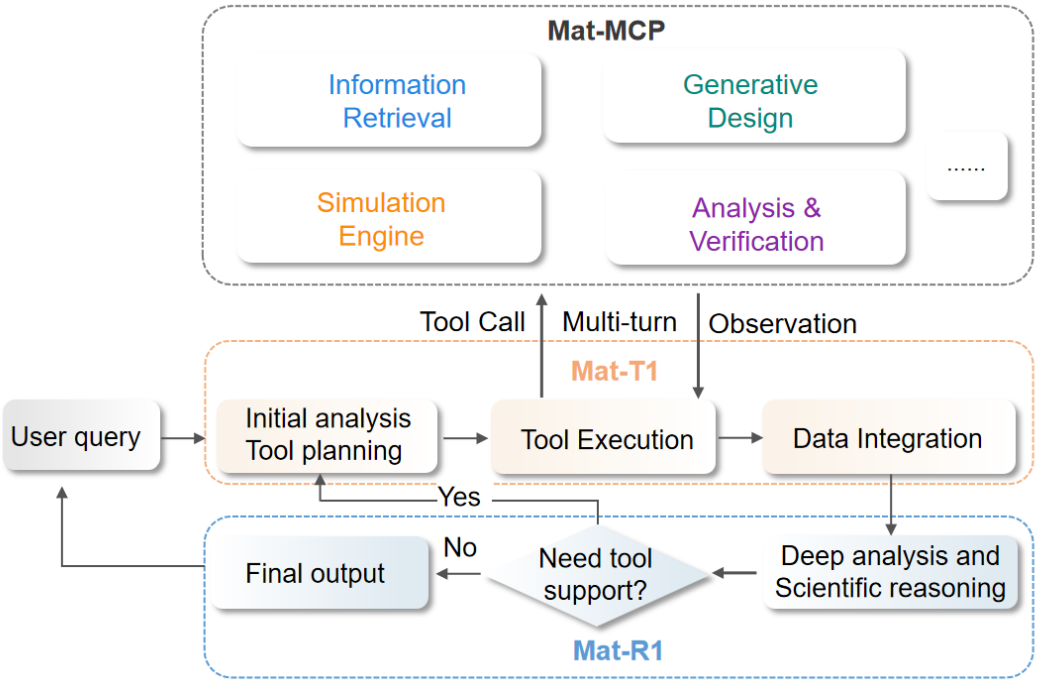

Fig. S6 Architecture of the MatBrain agent workflow. System-level diagram illustrating the dual-model collaborative framework, detailing the information flow between the analytical (Mat-R1) and executive (Mat-T1) models.

Fig. S7 Example of the iterative process. A case study demonstrating the "Generation-Verification-Feedback" mechanism, where Mat-R1 evaluates intermediate results and triggers subsequent tool calls.

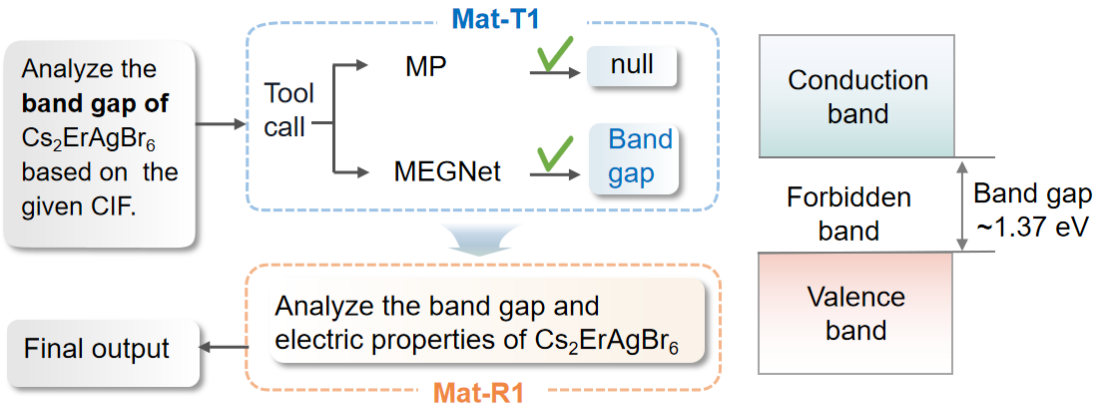

Fig. S8 Band-gap analysis of $Cs_2ErAgBr_6$. Mat-T1 invokes database and MEGNet-based tools, and Mat-R1 interprets the predicted band gap and electronic characteristics.

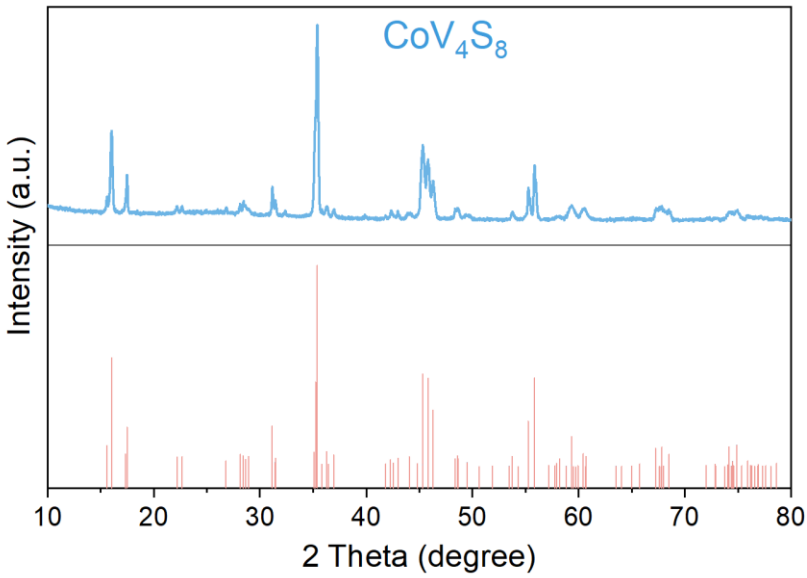

Fig. S9 Theoretical vs. experimental XRD patterns. Comparison of the simulated X-ray diffraction pattern generated by MatBrain and the experimental pattern of synthesized CoV_4S_8, confirming structural agreement.

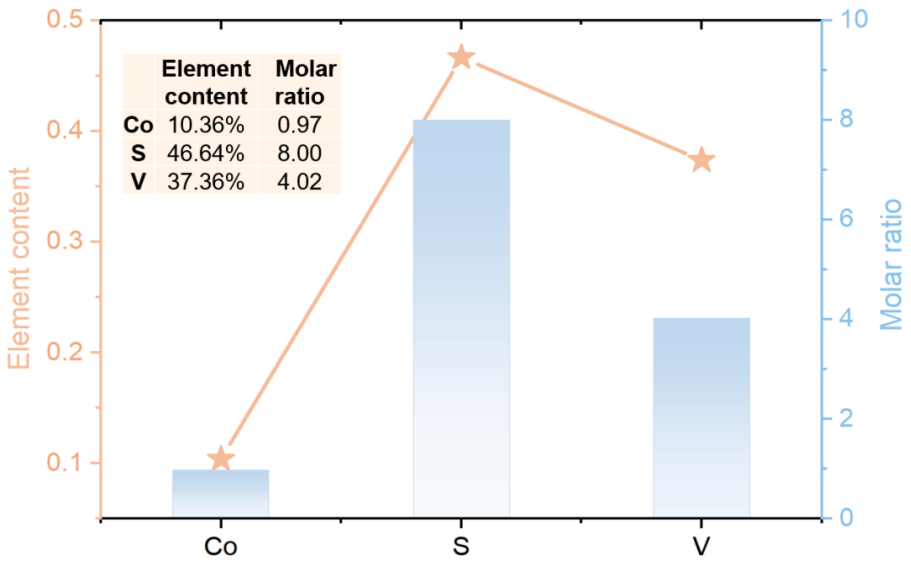

Fig. S10 Inductively coupled plasma (ICP) analysis. Elemental analysis results verifying the Co:V:S molar ratio of 1:4:8 in the synthesized nanosheets.

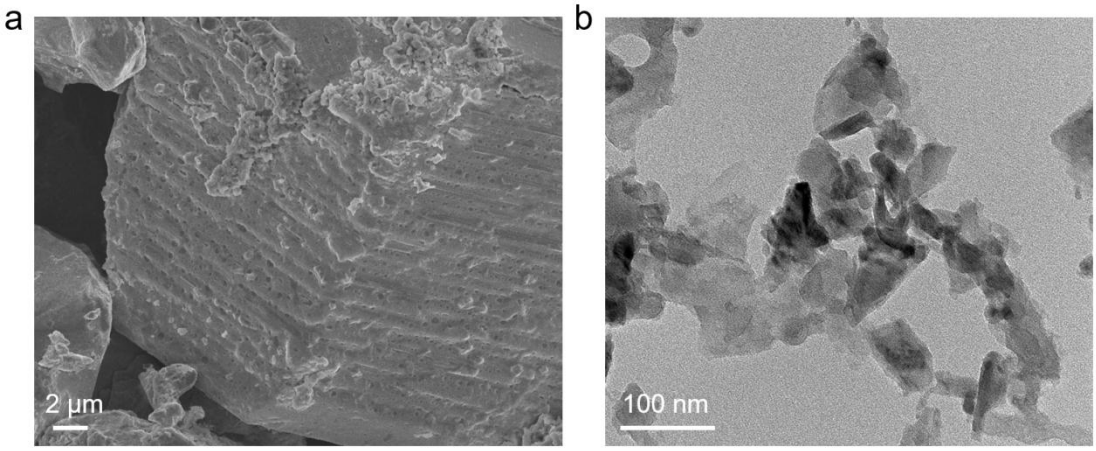

Fig. S11 Morphological characterization. (a) Scanning Electron Microscopy (SEM) image of bulk CoV_4S_8 crystals, showing a layered structure. (b) Transmission Electron Microscopy (TEM) image of exfoliated CoV_4S_8 nanosheets.

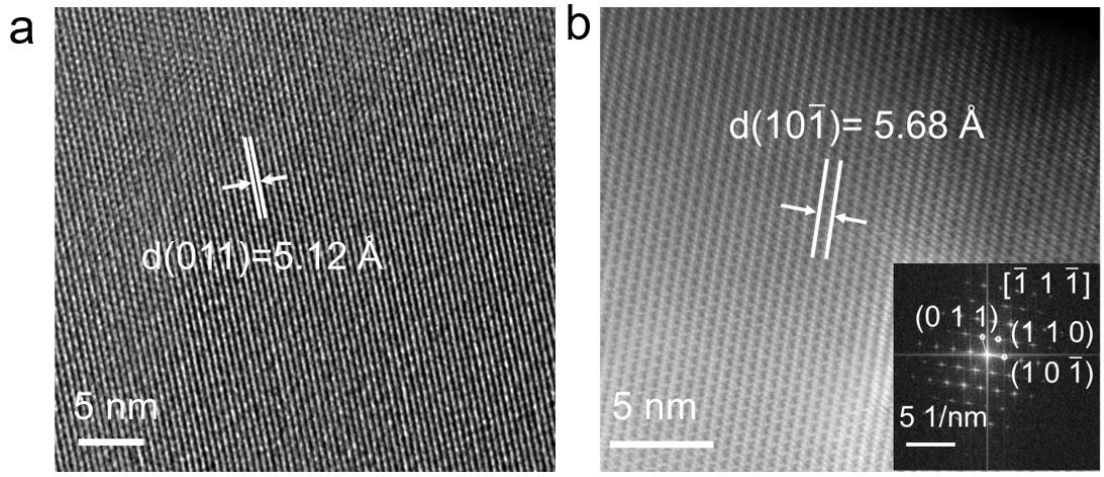

Fig. S12 Electron microscopy analysis. (a) High-resolution Transmission Electron Microscopy (HRTEM) image showing lattice fringes. (b) Aberration-corrected Scanning Transmission Electron Microscopy (STEM) image revealing the atomic arrangement.

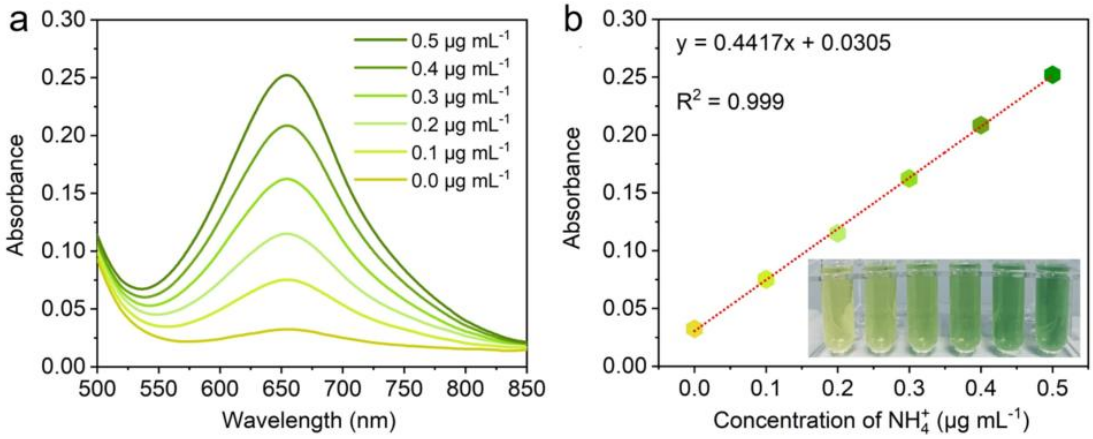

Fig. S13 UV-Vis calibration curves for Ammonia (NH_3). (a) UV-Vis absorption spectra of indophenol blue-stained standard solutions. (b) Calibration curve for NH_3 quantification.

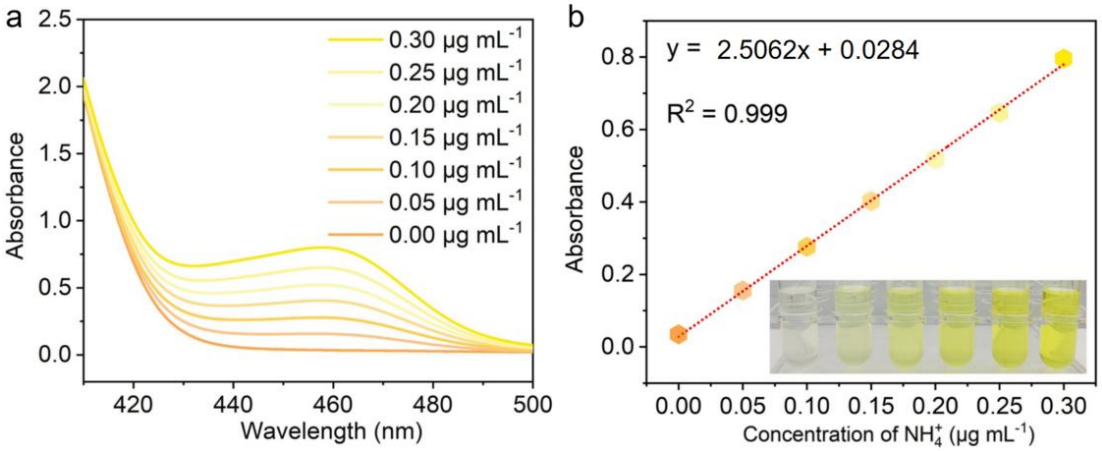

Fig. S14 UV-Vis calibration curves for Hydrazine (N_2H_4). (a) UV-Vis absorption spectra of Watt and Chrisp reagent-stained standard solutions. (b) Calibration curve for N_2H_4 quantification.

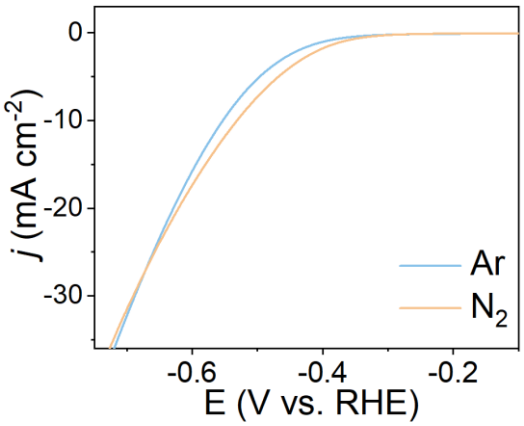

Fig. S15 Linear sweep voltammetry (LSV) curves. Comparison of current density-potential curves for CoV_4S_8 NS in Ar-saturated and N_2-saturated 0.1 M HCl electrolytes.

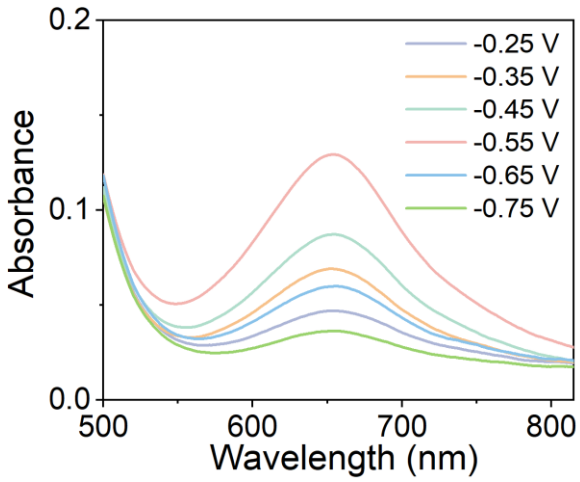

Fig. S16 UV-Vis spectra of catholytes. UV-Vis absorption spectra of stained catholytes collected after 2-hour electrolysis at various potentials, used for yield calculation.

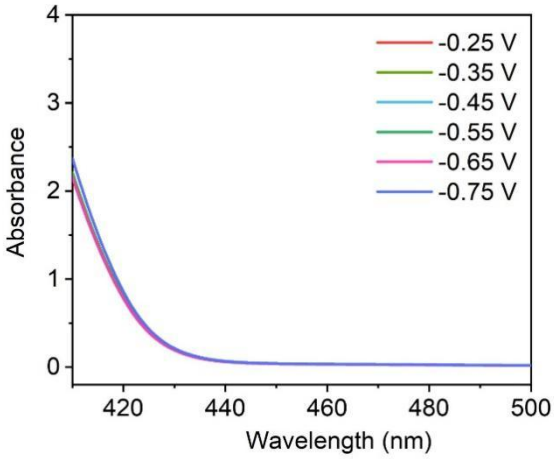

Fig. S17 Hydrazine detection tests. UV-Vis absorption spectra of electrolytes after 2 hours of electrolysis estimated by the Watt and Chrisp method, showing no detection of N_2H_4 by-products.

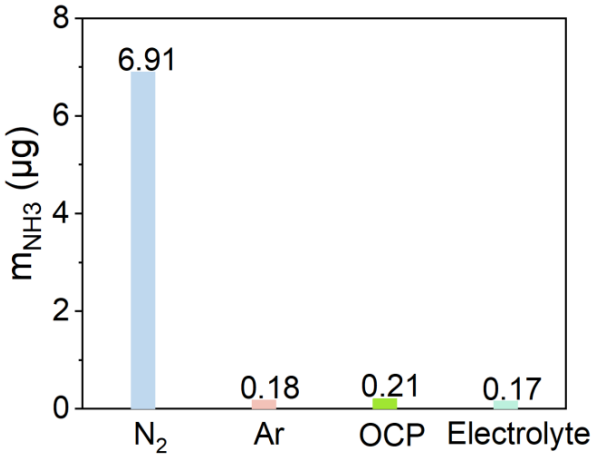

Fig.S18 Control experiment quantification. Comparison of ammonia yields under different conditions (Ar-saturated, Open Circuit Potential) to verify the nitrogen source.

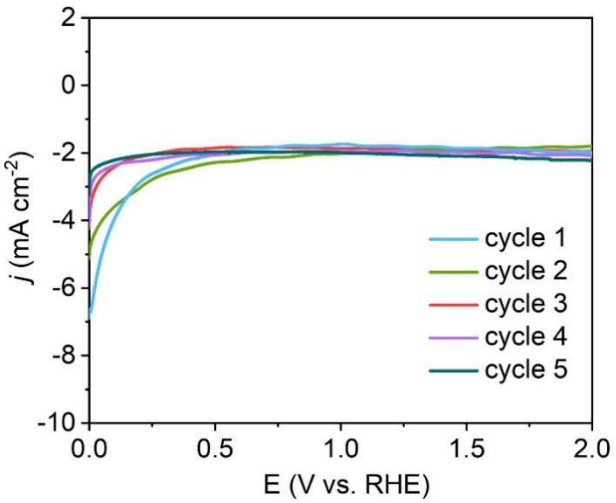

Fig. S19 Cycling stability evaluated by linear sweep voltammetry. Current density–potential curves of CoV_4S_8 nanosheets recorded over five consecutive LSV cycles in N_2-saturated 0.1 M HCl electrolyte.

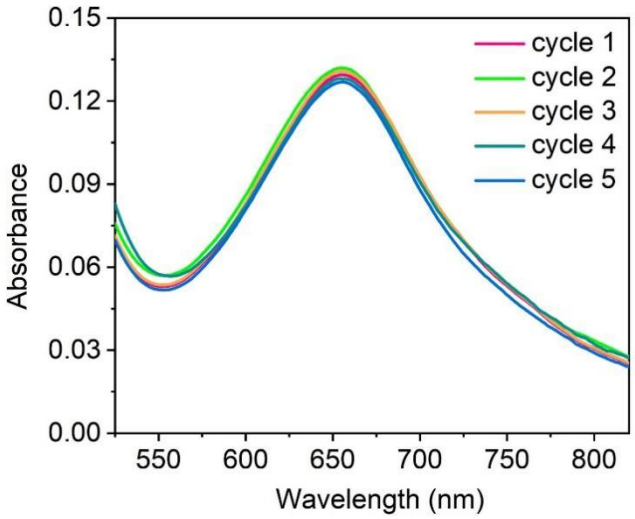

Fig. S20 UV-Vis spectra during cycling tests. Absorption spectra of stained catholytes collected from each of the 5 consecutive cycling tests at −0.55 V, indicating stable ammonia production.

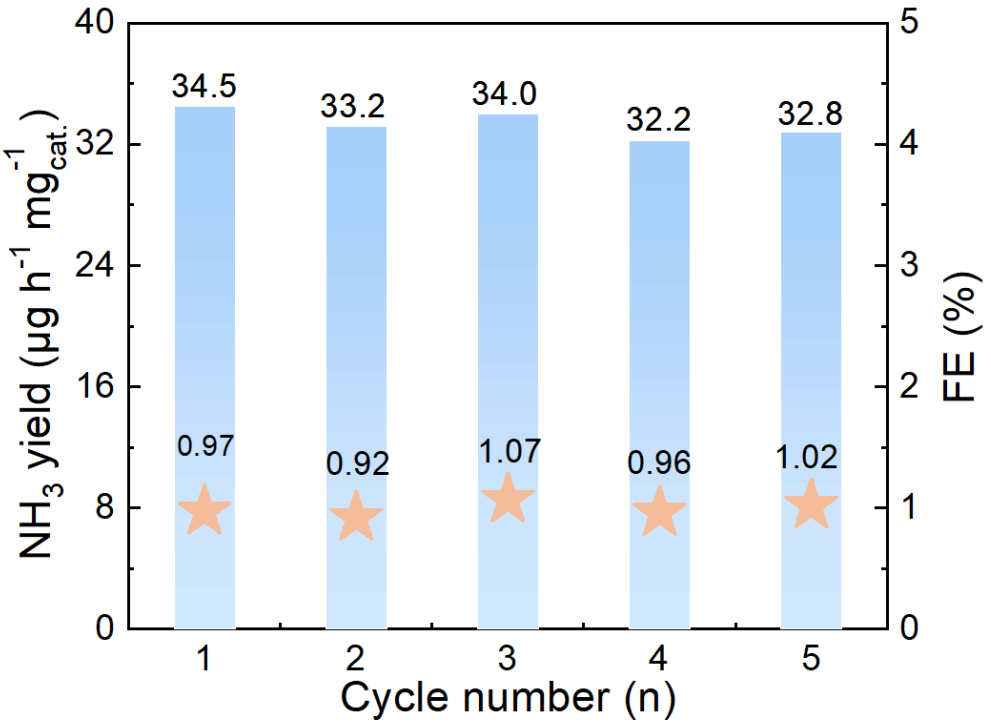

Fig. S21 Cycling stability summary. Calculated NH_3 yield and Faradaic efficiency over 5 consecutive cycles, demonstrating catalyst durability.

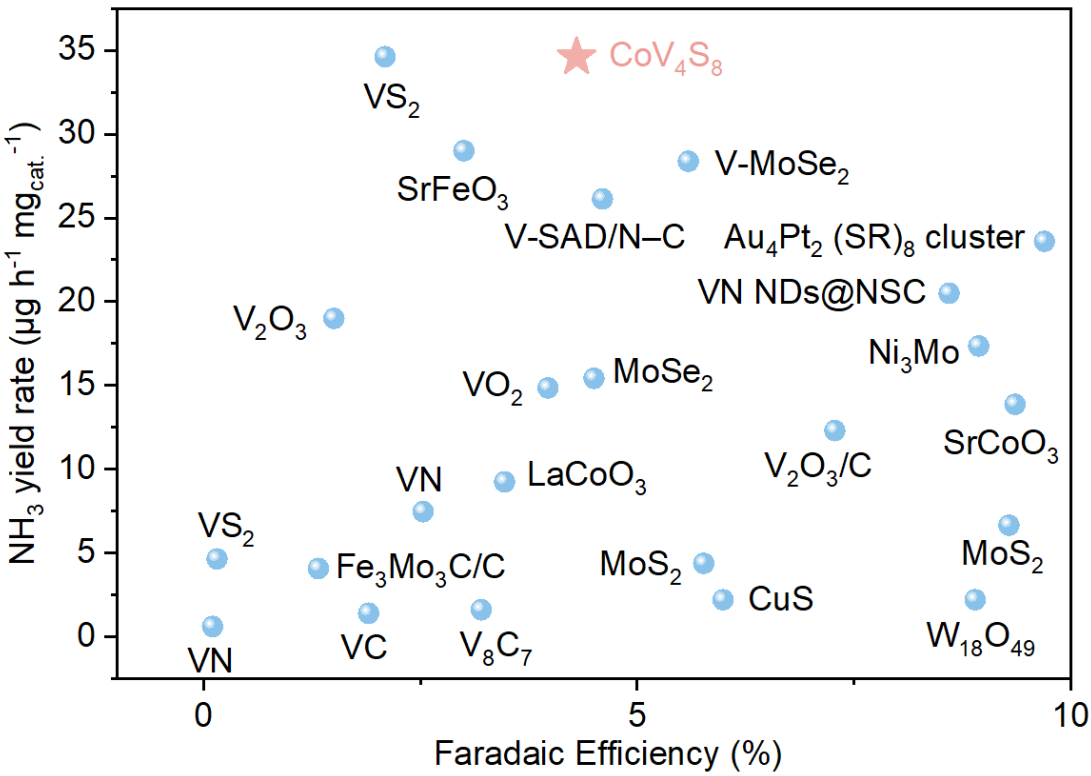

Fig. S22 Performance comparison of CoV_4S_8 with representative previously reported NRR electrocatalysts. [1-16].

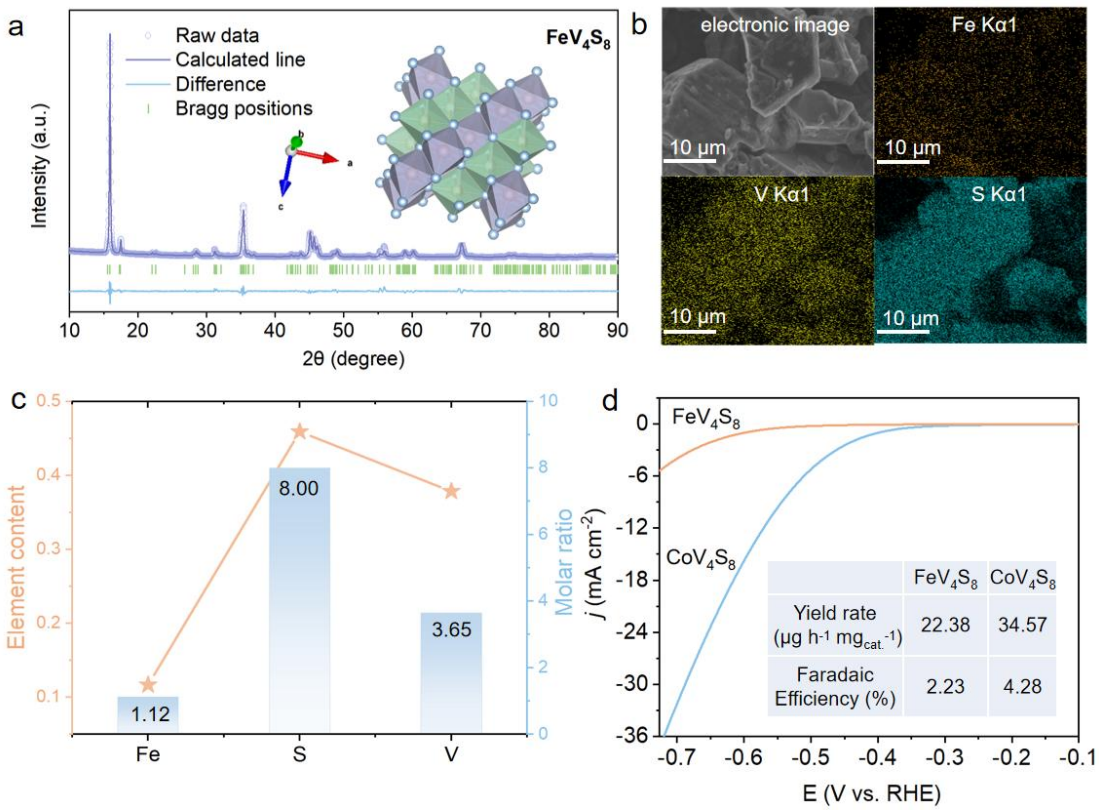

Fig. S23 Experimental validation of the second-priority candidate FeV_4S_8. (a) XRD Rietveld refinement confirming the C2/m space group of FeV_4S_8. (b) SEM image and EDS mapping show the uniform distribution of Fe, V, and S elements. (c) Element content achieved from ICP and molar ratio of Fe, V, and S in FeV_4S_8. (d) LSV curves, ammonia yield rates, and Faradaic efficiencies for CoV_4S_8 and FeV_4S_8.

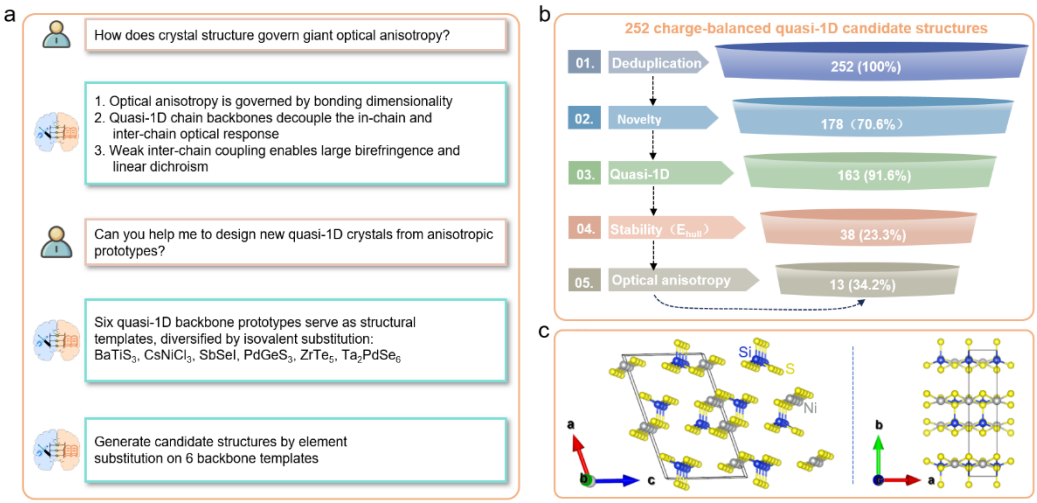

Fig. S24 The discovery of crystalline materials with strong optical anisotropy. (a) Interactive design process from strong optical anisotropy analysis to crystal structure generation. (b) A five-step screening funnel reduces 252 candidates to 13 novel materials through sequential filters. (c) Crystal structure of a typical material $SiNiS_3$.

Table S1 | Structure-design tool pool used in the evaluation experiments.

Function category	Tool	Backend	Function
Structure generation	generate_crystal_structures_crystallm	CrystaLLM	Generate candidate crystal structures from a chemical formula
Structure generation	generate_crystal_pyxtal	PyXtal	Generate crystal structures under space-group symmetry constraints
Symmetry/ feasibility	check_wyckoff_feasibility_pyxtal	PyXtal	Check Wyckoff feasibility for a formula under a target space group
Symmetry/ feasibility	extract_symmetry_and_wyckoff_pymatgen	pymatgen	Extract symmetry and Wyckoff information
Chemical feasibility	screen_smact_oxidation_states	SMACT	Screen feasible oxidation-state combinations
Chemical feasibility	check_smact_charge_neutrality	SMACT	Check charge-neutrality feasibility
Chemical feasibility	check_smact_pauling_test	SMACT	Perform Pauling electronegativity test
Energy/ relaxation	relax_crystal_structure_MatGL	MatGL	Relax crystal structure geometry
Energy/ relaxation	predict_formation_energy_MatGL*	MatGL	Predict formation energy per atom
Validation/ scoring	check_structure_atomic_geometry_pymatgen*	pymatgen	Check atomic geometry, overlap, and structural legality
Validation/ scoring	score_generated_cif_against_target_pymatgen	pymatgen	Score generated CIF against target formula, space group, and site count
Validation/ scoring	validate_formula_verifier	Verifier	Validate chemical formula legality
Validation/ scoring	validate_structure_verifier	Verifier	Validate generated structure legality
Validation/ scoring	score_candidate_structure_verifier	Verifier	Score candidate structures comprehensively
Artifact handling	register_cif_artifact*	Artifact manager	Register a CIF as a reusable cif_ref handle for downstream tools

*Shared between the structure-design and property-prediction tool pools.

Table S2 | Property-prediction tool pool used in the evaluation experiments.

Function category	Tool	Backend	Function
Energy/ stability	predict_formation_energy_MatGL*	MatGL	Predict formation energy per atom
Energy/ stability	estimate_energy_above_hull_pymatgen	pymatgen	Estimate energy above hull and stability label
Electronic structure	predict_multi_fidelity_band_gap_MatGL	MatGL	Predict multi-fidelity band gap
Electronic structure	classify_metallicity_from_band_gap	MatGL	Classify metallicity from band gap
Magnetism	screen_magnetic_moments_CHGNet	CHGNet	Predict magnetic moment and magnetic sites
Magnetism	analyze_collinear_magnetism_pymatgen	pymatgen	Analyze collinear magnetic order, including NM/FM/AFM/FIM
Structure parsing	analyze_cif_structure_pymatgen	pymatgen	Parse CIF for composition, formula, site count, volume, density, and symmetry
Structure parsing	standardize_structure_pymatgen	pymatgen	Standardize crystal structure
Structure parsing	check_structure_atomic_geometry_pymatgen*	pymatgen	Check atomic geometry, overlap, and structural legality
Chemistry/ coordination	check_chemical_formula_valence_pymatgen	pymatgen	Check formula valence consistency
Chemistry/ coordination	estimate_oxidation_states_pymatgen	pymatgen	Estimate oxidation states
Chemistry/ coordination	compute_bond_valence_pymatgen	pymatgen	Compute bond valence
Chemistry/ coordination	analyze_coordination_environment_pymatgen	pymatgen	Analyze the local coordination environment
Featurization	simulate_xrd_pattern_pymatgen	pymatgen	Simulate XRD pattern
Featurization	featurize_structure_matminer	matminer	Generate structure descriptors
Artifact handling	register_cif_artifact*	Artifact manager	Register a CIF as a reusable cif_ref for downstream tools

*Shared between the structure-design and property-prediction tool pools.

Table S3 | Execution-process metrics. (Higher completion rate is better; lower tool-failure rate is better)

System	Structure completion	Structure failure	Structure calls/task	Property completion	Property failure	Property calls/task
Decision tree (scripted pipeline)	61.0%	1.8%	8.5	100%	21.4%	8.0
Mat-T1	94.5%	3.2%	11.2	100%	0.0%	10.0

Table S4 | Deployment cost analysis and hardware requirements. Comparison of estimated memory usage and hardware costs between MatBrain and baseline large-scale models. The calculation assumes FP16 quantization utilizing the vLLM inference engine. To account for KV cache optimization and runtime overhead, the required VRAM is estimated at approximately $2 \times$ the model weight size. The results demonstrate that MatBrain can be deployed on a standard workstation (e.g., $4 \times$ RTX 4090), reducing hardware costs by over 95% compared to the cluster-level requirements of DeepSeek-R1.

	Model Size (FP8)	Min. VRAM	# of 4090 24G	# of H100 80G	Hardware Cost
DeepSeek-R1	671B	~1400 GB	-	~18	~$600 K
Qwen3-235B-A22B	235B	~480 GB	~20	~6	~$150 K
Mat-R1 (30B)	30B	~60 GB	~3	~1	~$11 K
Mat-T1 (14B)	14B	~30 GB	~2	~1	~$ 7K
MatBrain	44B	~90 GB	~4	~2	~$15 K

Table S5 | Extended comparison of property prediction classification tasks under the tool-free setting. Balanced accuracy was reported for metallicity and magnetism. Macro-F1 was reported for magnetic ordering. (Higher values indicate better performance)

Setting	Model	Metallic	Magnetic	Magnetic order
Tool-free	Gemini-2.5-Pro	0.580	0.783	0.342
Tool-free	Gemini-3-Pro	0.6149	0.8127	0.3609
Tool-free	GPT-5	0.597	0.779	0.198
Tool-free	GPT-5.5	0.6013	0.7548	0.1781
Tool-free	Claude-Opus-4.7	0.6621	0.8057	0.4388
Tool-free	Qwen3-30B-A3B (base)	0.525	0.738	0.274
Tool-free	Qwen3.6-35B-A3B	0.5431	0.6501	0.1767
Tool-free	Mat-R1	0.807	0.778	0.355
Tools	Gemini-2.5-Pro	0.527	0.528	0.059
Tools	Gemini-3-Pro	0.5118	0.6093	0.2285
Tools	GPT-5	0.492	0.506	0.020
Tools	GPT-5.5	0.5215	0.5945	0.1377
Tools	Claude-Opus-4.7	0.5513	0.6365	0.2993
Tools	Qwen3-14B-A3B (base)	0.554	0.470	0.071
Tools	Mat-T1	0.541	0.652	0.308
Agent	MatBrain	0.865	0.839	0.460

Table S6 | Extended comparison of property prediction regression tasks under the tool-free setting. (MAE was reported. Lower values indicate better performance)

Setting	Model	Formation energy	Energy above hull	Fermi energy
Tool-free	Gemini-2.5-Pro	0.545	0.223	1.559
Tool-free	Gemini-3-Pro	0.3784	0.2126	1.6368
Tool-free	GPT-5	1.090	0.235	3.025
Tool-free	GPT-5.5	0.6894	0.1939	2.1673
Tool-free	Claude-Opus-4.7	0.2682	0.1991	1.2528
Tool-free	Qwen3-30B-A3B (base)	1.740	0.229	3.454
Tool-free	Qwen3.6-35B-A3B	1.5576	0.2297	3.4439
Tool-free	Mat-R1	0.319	0.193	1.322
Tools	Gemini-2.5-Pro	0.451	0.425	3.569
Tools	Gemini-3-Pro	0.4437	0.4276	3.5691
Tools	GPT-5	0.452	0.411	3.526
Tools	GPT-5.5	0.4555	0.4366	3.5691
Tools	Claude-Opus-4.7	0.4338	0.4325	3.3157
Tools	Qwen3-14B-A3B	0.955	0.328	3.594
Tools	Mat-T1	0.429	0.436	3.569
Agent	MatBrain	0.184	0.119	0.772

Table S7 | Extended comparison of structure design tasks under the tool-free setting. (Values report criterion-satisfaction rates. Higher values indicate better performance.)

Setting	Model	Valid structure	Charge-balanced	Complete success	Formula	Atom count	Crystal system	Space group
Tool-free	Gemini-2.5-Pro	0.880	0.735	0.350	0.610	0.465	0.715	0.685
Tool-free	Gemini-3-Pro	0.875	0.740	0.485	0.665	0.650	0.840	0.755
Tool-free	GPT-5	0.670	0.595	0.135	0.470	0.320	0.625	0.575
Tool-free	GPT-5.5	0.855	0.785	0.415	0.755	0.645	0.800	0.780
Tool-free	Claude-Opus-4.7	0.630	0.485	0.220	0.360	0.350	0.550	0.455
Tool-free	Qwen3-30B-A3B (base)	0.365	0.265	0.010	0.170	0.130	0.335	0.280
Tool-free	Qwen3.6-35B-A3B	0.350	0.275	0.030	0.210	0.075	0.330	0.245
Tool-free	Mat-R1	0.850	0.700	0.475	0.620	0.610	0.800	0.755
Tools	Gemini-2.5-Pro	0.513	0.462	0.271	0.397	0.362	0.437	0.402
Tools	Gemini-3-Pro	0.810	0.760	0.465	0.730	0.670	0.745	0.700
Tools	GPT-5	0.621	0.591	0.298	0.540	0.450	0.571	0.515
Tools	GPT-5.5	0.835	0.800	0.580	0.760	0.700	0.770	0.740
Tools	Claude-Opus-4.7	0.740	0.670	0.440	0.560	0.550	0.680	0.660
Tools	Qwen3-14B-A3B	0.653	0.568	0.236	0.472	0.322	0.427	0.387
Tools	Mat-T1	0.917	0.790	0.441	0.673	0.620	0.837	0.797
Agent	MatBrain	0.945	0.865	0.465	0.710	0.635	0.890	0.860

Table S8 | Crystallographic validation checks of this generated structure.

Validation item	**Expected value**	**MatBrain output**
Parsed structure type	ABX_3 mixed-halide perovskite	pass
X-site occupancy	Cl 0.2 / Br 0.4 / I 0.4	pass
Occupancy sum per X site	1.000	pass
Target total formula	$CsPbCl_{0.6}Br_{1.2}I_{1.2}$	pass
CIF parseability	readable by pymatgen	pass
Abnormal short contacts	absent	pass
Lattice parameter	between $CsPbCl_3$ and $CsPbI_3$ end members	pass
Structure representation	disordered fractional-occupancy CIF	pass
Goldschmidt tolerance factor $t = ((r_A + r_X) / [\sqrt{2}(r_B + r_X)])$	Typical perovskite range, ~0.8–1.0	0.859, Pass
Octahedral factor $\mu = r_B / r_X$	> 0.414	0.587, Pass

Table S9 | Seven-step high-throughput screening workflow. Detailed parameters and criteria for each step of the bio-inspired catalyst screening process, including chemical composition, valence balance, structural deduplication, geometric validity, thermodynamic stability, electronic properties, and novelty checks.

Step	Description	Module	Number of removed structures	The filtered sample
1	Screening for pure M-V-S ternary systems (M=Fe/Co/Ni)	Element Composition Filter	3,027	e.g. V12S4 (No transition metal), Fe1V1S4N1 (Contains disallowed element N)
2	Verification of valence stability and charge balance	Pymatgen Valence Check	5,019	e.g. V2Fe6S4, Fe2V2S2 (No valid oxidation states),
3	Structural clustering to eliminate redundant crystals	Pymatgen Structure Matcher	11,502	site tolerance <0.3, lattice tolerance <0.2, angle tolerance <5°
4	Validation of atomic arrangement and bond length stability	CrystalNN Atomic Overlap Filter	324	e.g. V1S4Fe1 (atom overlap, S-S distance 0.005 Å)
5	Assessment of thermodynamic stability (energy above hull)	CHGNet Phase Stability	10,086	energy above hull ⩽ 0.025 eV/atom
6	Prediction of electronic band gap properties	MEGNet Bandgap Predictor	0	All materials have moderate bandgaps (1.0 -2.0 eV)
7	Verification of novelty against known materials database	Novelty Assessment	4	e.g. V2Co1S4, S4Ni1V2 (Matches MP structure)

Table S10 | Stage-resolved compute budget and wall-clock time for the representative reproduction

Stage	Tool / module	Hardware	Wall-clock time
Structure generation	MatterGen	3 × RTX 4090 (24 GB)	33 h 52 min (90.2%)
Element filtering	Composition parser	CPU	1.8 s
Valence-state filtering	Oxidation-state checker	CPU	45.7 s
Deduplication	pymatgen StructureMatcher	CPU	2 h 24 min (6.4%)
Structure validity check	Geometry validator, multiprocessing	CPU	76.7 s
Stability screening	CHGNet + phase diagram	3 × RTX 4090	1 h 15 min (3.3%)
Electronic-property screening, bandgap	MEGNet (MatGL)	1 × RTX 4090	4.2 s
Novelty checking	MP API + StructureMatcher	CPU	15.7 s
Total	—	3 × RTX 4090	37 h 34 min

Table S11 | Per-task human-active time comparison between a human-expert-operated workflow and the MatBrain-assisted workflow.

For the MatBrain-assisted workflow, human-active time was measured after deployment of the MatBrain and Mat-MCP system. One-time platform development, tool integration and system deployment costs were not included in the per-task comparison. Computational wall-clock time is reported separately in Table S10.

Workflow stage	**Human-expert-operated baseline**	**MatBrain workflow**
Define objective and search space	2–4 weeks	5–10 min
Configure structure generation	~3.7 h	No additional human intervention
Monitor / organize generated results	~0.2 h	No additional human intervention
Element + valence filtering setup	~0.4 h	Automatic
Deduplication parameter setup/debugging	~1.6 h	Automatic
Geometry validation setup/debugging	~1.1 h	Automatic
Stability screening setup/debugging	~2.8 h	Automatic
Electronic-property screening	~0.3 h	Automatic
Novelty check and database comparison	~0.1 h	Automatic
MatBrain-automated screening stages	~10.2 h	~10 min instruction

Table S12 | V-richness-based classification of the 38 candidates.

Category	V/(V+M)	Number of candidates	Representative materials
1	0.80	2	CoV_4S_8, FeV_4S_8
2	0.75	9	CoV_3S_6, NiV_3S_6, FeV_3S_6
3	0.667	9	FeV_2S_4, NiV_2S_4, ...
4	0.50	4	$Co_2V_2S_6$, $Ni_2V_2S_6$, ...
5	0.40	1	$Fe_3V_2S_4$
6	0.333	5	$Fe_4V_2S_4$, $Fe_4V_2S_6$, $Ni_4V_2S_6$, ...
7	0.25	7	Ni_3VS_4, Fe_3VS_4, $V_2Ni_6S_8$, ...
8	0.20	1	Ni_4VS_4

Table S13 | Leakage audit between the training set and the initial held-out test set.

Audit level	**Criterion**	**Overlapping test samples**	**Fraction**
Material identity	MP-ID overlap	0 / 2000	0.00%
Exact structure identity	Exact CIF-hash overlap	0 / 2000	0.00%
Prototype-level overlap	Full AFLOW-style protostructure-label overlap	217 / 2000	10.85%
Structure-similarity overlap	Fingerprint retrieval followed by StructureMatcher confirmation	144 / 2000	7.20%
Composition-level overlap	Raw-formula composition overlap	692 / 2000	34.60%

Note S1. Additional entropy analyses and interpretation scope

Token-level Shannon entropy was used to characterize the predictive distributions of Mat-R1, Mat-T1 and their corresponding Qwen backbones. For a generated token t with predictive distribution p over the vocabulary X, entropy was computed as $H(X) = -\sum_i^n p(x_i)\ log_2\, p(x_i)$, and reported values correspond to mean token-level entropy over the analyzed response segments. This analysis requires access to token log-probabilities or logits, and was therefore restricted to models for which such probabilities were available under the same definition.

The main entropy profiles in Figure 4 are model-side entropy measurements over generated output-token distributions. To clarify their interpretation, we performed an additional segment-level intrinsic entropy analysis on real held-out trajectories using Qwen3-30B-A3B as a fixed reference model. This analysis compares the variability of the tool-observation segments and the final analytical-answer segments, and is distinct from the model-side entropy profiles in Figure 4.

Data segment	**Property prediction**	**Structure design**
Tool observation | Mat-T1	0.94 bit	0.53 bit
Analytical answer | Mat-R1	0.39 bit	0.25 bit

We further paired the model-side entropy shifts in Figure 4 with held-out functional metrics for each trained model and its own backbone. This comparison connects the entropy profiles to independently measured task behavior, without treating entropy magnitude alone as a correctness metric.

Backbone → Trained model	**Entropy change**	**Performance change**
Qwen3-30B-A3B → Mat-R1	0.673 → 0.483	Metallicity classification: 0.525 → 0.807; formation-energy MAE: 1.740 → 0.319; structure-design complete-success rate: 0.010 → 0.475
Qwen3-14B → Mat-T1	0.878 → 0.974	Tool-failure rate in structure design: 20.0% → 3.2%; tool-failure rate in property prediction: 47.8% → 0.0%; structure-design complete-success rate: 0.236 → 0.441

In addition, the Mat-T1 training dynamics in Figure S4 show that actor entropy increased together with the mean rollout reward during RL optimization. The group-

relative advantages were centred within each prompt group by construction; their near-zero batch mean was therefore not treated as an indicator of training stability. This training-time trend indicates that the policy did not collapse into a narrow low-entropy behavior while reward improved. This training-dynamics analysis is distinct from Figure 4, where entropy was measured at test time from the generated output-token distributions of Mat-R1, Mat-T1 and their corresponding backbones. Together, these supplementary analyses clarify the interpretation of the entropy profiles in Figure 4, while task performance is assessed independently by the controlled benchmarks in Figure 3.

References

[1] Wang, X. et al. Optimizing Electrocatalytic nitrogen reduction via interfacial electric field modulation: elevating d-band center in WS_2-WO_3 for enhanced intermediate adsorption. *Angew Chem Int Ed Engl* **62**, e202303794, doi:10.1002/anie.202303794 (2023).

[2] Lv, X.-W., Liu, Y., Wang, Y.-S., Liu, X.-L. & Yuan, Z.-Y. Encapsulating vanadium nitride nanodots into N,S-codoped graphitized carbon for synergistic electrocatalytic nitrogen reduction and aqueous Zn-N_2 battery. *Applied Catalysis B: Environmental* **280**, 119434, doi:10.1016/j.apcatb.2020.119434 (2021).

[3] Zhao, L. et al. Efficient N_2 reduction with the VS_2 electrocatalyst: identifying the active sites and unraveling the reaction pathway. *Journal of Materials Chemistry A* **9**, 24985–24992, doi:10.1039/d1ta05237c (2021).

[4] Zhang, C. et al. Vanadium carbide with periodic anionic vacancies for effective electrocatalytic nitrogen reduction. *Materials Today* **40**, 18–25, doi:10.1016/j.mattod.2020.04.031 (2020).

[5] Zhang, R. et al. Metal–organic framework-derived shuttle-like V_2O_3/C for electrocatalytic N_2 reduction under ambient conditions. *Inorganic Chemistry Frontiers* **6**, 391–395, doi:10.1039/C8QI01145A (2019).

[6] Lv, S. et al. Multidimensional manipulation of $MoSe_2$ enabled by vanadium doping for efficient nitrogen reduction. *ACS Applied Materials & Interfaces* **18**, 21978–21986, doi:10.1021/acsami.6c00415 (2026).

[7] Zhang, R. et al. Electrocatalytic N_2 fixation over hollow VO_2 microspheres at ambient conditions. *ChemElectroChem* **6**, 1014–1018, doi:10.1002/celc.201801484 (2019).

[8] An, T. Y. et al. V_2O_3/VN electrocatalysts with coherent heterogeneous interfaces for selecting low-energy nitrogen reduction pathways. *SusMat* **4**, e226, doi:10.1002/sus2.226 (2024).

[9] Liu, Y. et al. Interface engineering of CoS/MoS_2 heterostructure for the electrocatalytic reduction of N_2 to NH_3. *Inorganic Chemistry Frontiers* 10, 5700–5709, doi:10.1039/d3qi01139a (2023).

[10] Cheng, H., Cui, P., Wang, F., Ding, L. X. & Wang, H. High efficiency electrochemical nitrogen fixation achieved with a lower pressure reaction system by changing the chemical equilibrium. *Angewandte Chemie International Edition* **58**, 15541–15547, doi:10.1002/anie.201910658 (2019).

[11] Cheng, K. et al. The electron-rich interface regulated MBene by S-bridge guided to enhance nitrogen fixation under environmental conditions. A*dvanced Functional Materials* **35**, 2417914, doi:10.1002/adfm.202417914 (2025).

[12] Tong, Y. et al. Vacancy engineering of iron-doped $W_{18}O_{49}$ nanoreactors for low-barrier electrochemical nitrogen reduction. *Angewandte Chemie International Edition* **59**, 7356–7361, doi:10.1002/anie.202002029 (2020).

[13] Yao, C. et al. Atomically-precise dopant-controlled single cluster catalysis for electrochemical nitrogen reduction. *Nature Communications* **11**, 4389, doi:10.1038/s41467-020-18080-w (2020).

[14] Han, N. et al. Tuning spin polarization of iron in oxides to boost electrocatalytic ammonia production. *Advanced Materials* **37**, e00900, doi:10.1002/adma.202500900 (2025).

[15] Peng, G. et al. Crystal structures of molybdenum borides dictate electrocatalytic ammonia synthesis efficiency. *Applied Catalysis B: Environmental* **338**, 123020, doi:10.1016/j.apcatb.2023.123020 (2023).

[16] Zhou, H. Y. et al. Multi-site intermetallic Ni_3Mo effectively boosts selective ammonia synthesis. *Applied Catalysis B: Environmental* **339**, 123133, doi:10.1016/j.apcatb.2023.123133 (2023).