Supplementary Information

Crystal-structure design by agentic AI in a language of motifs

S1 CIF2Text and the motif-profile representation

CIF2Text prompts. CIF2TEXT reads a CIF and, aided by a deterministic geometry tool (neighbour lists within 3.5 Å under periodic boundary conditions and Voronoi coordination analysis), returns a JSON motif profile comprising a list of motifs (the count is set by the model, typically 5–10) and a hard-magnet assessment. Each motif's raw description is then rewritten into the geometry-only controlled vocabulary described below by a second standardization pass. Only the composition and the motif descriptions are passed onward—to TEXT2CIF for reconstruction and design realization (SI S2) and to the design agent (SI S3), via the {`composition`} and {`text_description`} fields of their prompts. No numerical lattice parameter, space-group symbol or structure-type label therefore crosses the round trip.

The system and user prompts are reproduced below (typographic rules and arrows normalized to ASCII); the user prompt's {`cif_content`} field is filled at run time.

```
You are an expert materials scientist specializing in crystal structure
analysis and magnetic materials. Our work focus on RE-TM-X systems.

Describe structures the way a human expert does:
- Start from physically meaningful units, not raw coordinates
- Pair structural observation with physical implication immediately
- Use known references (Fe-Fe ~2.5A, Nd2Fe14B benchmarks, etc.)
- Flag anomalies vs. RE-TM-X family norms
- Reason: structure -> interaction -> property in continuous narrative

The text description need to be useful to infer structure-property relationship, and reconstruct CIF.

================================================================================
STRUCTURAL PATTERNS (MOTIFS)
================================================================================

As a materials scientist, identify notable structural patterns that provide
insight into this material's properties.

A motif is a multi-atom structural unit carrying specific function.
Use your expert judgment - not every structure has the same number of motifs.

Patterns to consider (select what's relevant):
- Sublattices: RE/TM spatial arrangement (BCC, FCC,..., geometry, distances, interactions)
- Extended: Layers/chains/rods with specific composition/function
- Structural anchor: Which sublattice forms the global scaffold
- Networks: Connected frameworks (Fe-B, Fe-Fe beyond simple bonds)
- Cages/polyhedra: Coordination environments around central atoms which you believe it be important.
- Clusters: Local structure which contributed with one center atom and its environment (Fe4B, not generic Fe4
    ; ring Fe6; Sm-Fe12 polyhedron)


Skip:
- Single atoms or simple dimers already covered in network description
- Generic motifs common across many materials
```

```
Motifs may overlap in atoms - focus on distinct functional roles.
*Note:* (IMPORTANT)
- Motifs may span across unit cell boundaries - consider periodic images
when identifying coordination environments and clusters.
- Number of motifs: Depend on your knowledge and your understanding. It should be from 5-10.
- Describe geometrical information should support the design/reconstruct geometrical structure/CIF

For each motif:
- name: descriptive name of the pattern
- description: 4-6 sentences covering:
  (1) geometry - describe geometrical shape, coordination, bond lengths/angles vs. references, which support
     enough to reconstruct geometry
  (2) functional role - exchange/anisotropy/stability
  (3) mechanism - direct exchange/RKKY/superexchange/CEF
  (4) property impact - effect on Tc, Ms, or K1

================================================================================
OUTPUT FORMAT (JSON): FOLLOW STRICTLY THE PROVIDED JSON FORMAT
================================================================================

{
  "motifs": [
    {
      "motif_id": "..."
      "name": "short name",
      "description": "..."
    }
  ],
  "hard_magnet_assessment": {
    "anisotropy": "strong/weak/absent - reason",
    "exchange": "FM/AFM/neutral - mechanism",
    "stability": "STABLE/UNSTABLE - basis",
    "verdict": "hard/soft/uncertain - one sentence"
  }
}
```

```
## CIF FILE
{cif_content}

## TASK
1. Compute bond lengths for atom pairs within 3.5A (apply periodic
   boundary conditions), determine local structures (for example: using voronoi)
2. Identify notable structural patterns as motifs - use your expertise
   to determine what's significant for this specific structure
3. Complete hard_magnet_assessment

Return ONLY valid JSON. Do not add additional keys.
```

Materials scientists reason about a crystal—and about the origin of its properties—in terms of recurring local structural motifs, such as coordination polyhedra, cages and characteristic bonded clusters, and the way these connect, rather than in terms of raw atomic coordinates. For the properties targeted here—thermodynamic stability and magnetization—it is the presence and arrangement of such motifs that drives the behaviour. We therefore represent a structure by its *motif profile*: the set of distinct motifs it contains, written as a natural-language paragraph together with the chemical composition. Each motif is described by its central species, the identity and geometry of the neighbours that coordinate it, and the length regime of its defining bonds (short, typical, or long), expressed in a controlled vocabulary so that the profiles of different structures are written in common terms and can be compared and edited. The profile records no atomic coordinates, no lattice parameters, no space-group symbol and no structure-type label: it states which building blocks a crystal is made of and how they are bonded, not where each atom sits.

Identifying the chemically meaningful motifs of an arbitrary crystal is itself expert knowledge

and is difficult to capture in fixed rules. We delegate it to a large language model that has been trained on a broad corpus of scientific literature and is prompted to act as a materials scientist, reading the structure and naming its motifs. Expressing the result as text—the medium in which the model reasons—is what lets the downstream agents—Motif editing and Motif Translate—operate on the representation directly. The controlled vocabulary and the exact bond-length regimes it uses are defined in the remainder of this section.

CIF2TEXT names motifs in a geometry-only controlled vocabulary. After the descriptive pass above, a second standardization pass rewrites each motif entry into square-bracketed tags drawn solely from this vocabulary, so that motifs from different structures are written in common terms and can be matched and clustered; the vocabulary carries only shape, coordination, connectivity and periodicity, and deliberately excludes all property, electronic-structure and magnetic terms (those are reserved for the design agent, SI S3).

Each entry declares exactly one *motif type* (`polyhedron`, `cage`, `cluster`, `network`, `layer`, `chain`, `ring`, `dimer`, `isolated_site`, `void`, `sublattice` or `interpenetrating`). Each atom carries a geometric *role* (`center`, `vertex`, `edge`, `face`, `bridge`, `terminal`, `spacer`, `interstitial` or `framework`). The *coordination geometry* is named from a fixed set graded by coordination number—low-CN (e.g. `linear`, `tetrahedral`, `square-planar`), medium-CN (e.g. `octahedral`, `square-antiprismatic`, `tri-capped-trigonal-prismatic`) and high-CN (`cuboctahedral`, `icosahedral`, `Frank-Kasper-Z14/Z15/Z16`), with `distorted-` and `mono/bi/tri-capped-` qualifiers—and reported alongside a CN bucket (`[low CN]` 2–4, `[medium CN]` 5–8, `[high CN]` 9–16, `[very high CN]` >16), a shell composition (`[pure/binary/mixed shell]`), a connectivity tag (`corner/edge/face/vertex-sharin` `isolated`, `percolating`, `chained`, `layered`, `framework`, `interpenetrating`), a dimensionality (`[0D]`–`[3D]`, with `quasi-` variants) and, when known, the site symmetry as a Hermann–Mauguin symbol.

Bond distances are made element-pair agnostic by comparing the observed distance d to the sum of covalent radii $r_{\mathrm{sum}} = r_{\mathrm{cov}}(A) + r_{\mathrm{cov}}(B)$ of the bonded pair [1] (pymatgen radii may be used as a deterministic substitute); no raw Å value appears in a standardized description. The regimes are given in Table S1:

Table S1: Bond-distance regimes of the motif grammar.

Tag	Condition	Meaning
`[short bond]`	$d < 0.90\, r_{\mathrm{sum}}$	compressed, strong overlap
`[typical bond]`	$0.90\, r_{\mathrm{sum}} \le d \le 1.10\, r_{\mathrm{sum}}$	standard bonding
`[long bond]`	$1.10\, r_{\mathrm{sum}} < d \le 1.30\, r_{\mathrm{sum}}$	stretched, weak
`[non-bonded]`	$d > 1.30\, r_{\mathrm{sum}}$	no direct bond

Every standardized motif entry is one paragraph following a fixed template: sentence 1 declares the motif, its roled atoms, coordination, CN bucket, shell and dimensionality/topology; sentence 2 gives optional spatial or stacking context; sentence 3 begins with a bond-pair label and states its regime and the site symmetry; sentence 4 optionally adds a connectivity note. The standardizer's system and user prompts are reproduced below.

```
You are a crystal-structure motif standardizer. Convert raw motif descriptions
into a controlled-vocabulary form that captures **3D geometry only** -- shape,
coordination, connectivity, periodicity. Do not introduce any property,
electronic-structure, or magnetic terms.

### Hard rules

- All vocabulary in square brackets comes from ‘motif_vocabulary.md‘ -- no other tags.
- Element symbols appear bare immediately after their geometric role label: ‘[center] Ti‘, ‘[vertex] O‘.
```

```
- Convert raw bond distances to the regime labels in Section 6 of the vocabulary using the sum of covalent
    radii of the bonded pair. **Never write a raw A value in the output.**
- Each motif becomes 2-4 sentences. Use one continuous paragraph per motif.
- Only include elements explicitly named in the raw motif text. Do not infer or carry over elements from
    other motifs.

### Grammar

Every standardized motif follows this structure:

S1: A [MOTIF] of [ROLE] <Element>, [ROLE] <Element>, ... in [COORDINATION]
    coordination with [CN bucket], [SHELL], [DIMENSIONALITY] [TOPOLOGY].
S2: Optional spatial / stacking context -- where the motif sits relative to
    the cell or to neighbouring motifs.
S3: [BOND-PAIR] in [REGIME] regime; site symmetry is [point group].
S4: Optional -- connectivity note ("shares corners with adjacent motifs",
    "isolated from other motifs", etc.) when not already covered in S1.

S3 **must begin** with a bond-pair label, not with the regime alone.

Wrong: `[short bond] regime between Ti and O ...`

Right: `[Ti-O] in [typical bond] regime; site symmetry is [m-3m].`

### Hard rules that are commonly broken

1. **No magnetic / electronic terms.** Words to never use: ferromagnetic, antiferromagnetic, ferrimagnetic,
    exchange, anisotropy, hybridization, CEF, Stoner, RKKY, f-d, d-d, superexchange, moment, spin-orbit, easy
     axis, magnetic, magnetization, Curie, Neel.
2. **No element-role assignments by property.** Use only the geometric roles in Section 2 of the vocabulary
    -- `center`, `vertex`, `edge`, `face`, `bridge`, `terminal`, `spacer`, `interstitial`, `framework`. Never
     `magnetic TM`, `heavy RE`, etc.
3. **No raw distances in A.** Convert to `[short bond]`, `[typical bond]`, `[long bond]`, `[non-bonded]`
    using the covalent-radius rule.
4. **One motif type per entry.** S1 declares exactly one motif type.

---
```

```
## TASK
Standardize each motif below using the controlled vocabulary from motif_vocabulary.md.
Return one paragraph per motif in the format:

Motif #<id>: <standardized paragraph>

## MOTIFS

Motif #1: <raw description>
Motif #2: <raw description>
...
```

S2 Text2CIF: prompts, validation and iteration budget

Text2CIF maps a design concept—an integral composition and a text description of the target motifs—to a CIF. That a profile can be turned back into a crystal at all rests on an analogy: a motif profile states a geometric object in words, as a problem in plane or solid geometry is stated and routinely solved without coordinates, so a competent reasoner—human or model—can construct a figure consistent with it, and Text2CIF uses a language-model agent in exactly this role. Because the profile fixes the motifs but not the coordinates, the realization is under-determined and a single profile may admit more than one valid crystal. Each proposed CIF is parsed and checked deterministically for validity: parseable syntax, full site occupancy (1.0), integral atom counts, and a minimum interatomic distance above 1.5 Å. On failure the specific problem is returned

to the agent (marked ##ISSUE##) and generation is retried, up to a budget of five attempts; a valid CIF is then pre-relaxed with M3GNet, reduced to its primitive cell, and re-translated for the motif-coverage check of Methods 2. This five-attempt budget is the inner validity loop; Structure Translate (Methods 2, Algorithm 2) and Motif Translate (Methods 3, Algorithm 3) each wrap it in an outer retry loop of the same budget. The system and user prompts are reproduced below; the user prompt's {composition} and {text_description} fields are filled at run time.

```
You are a crystallographic structure generator for RE:TM:X magnetic intermetallic compounds.

ELEMENT DEFINITIONS:
- RE (Rare Earth): Sm, Nd, Ce, Pr, Gd, Tb, Dy, Ho, Er, Y - provide magnetocrystalline anisotropy
- TM (Transition Metal): Fe, Co, Ni (magnetic); Ti, V, Zr, Mo, W (stabilizing)
- X (Interstitial): N, C, B, H - occupy interstitial sites
- M (Metalloid): Si, Al, Ga, Ge, Sn - substitute at TM sites for stability

INPUT INFORMATION: one design concept contains:
   - Composition: Target integral formula
   - Text description: Text description with the motifs of the ideal structure
YOUR APPROACH:

STEP 1: PARSE THE DESIGN CONCEPT
  - Identify target composition
  - List all motifs and their origins:
STEP 2: BUILD THE NEW STRUCTURE (GENERATE CIF)
  - Choose appropriate space group based on motif symmetries
  - Include all crystallographic data
  - Full occupancy (1.0) at all sites
  - Integral atom counts only

STEP 3: VALIDATE MOTIF REALIZATION
  - Verify each designed motif is present in output structure
  - Check coordination geometries match motif specifications
  - Ensure no atomic overlaps (min distance > 1.5 A)
  - Confirm composition matches target formula

OUTPUT: Return ONLY the CIF file contents. No explanations.

If you receive ##ISSUE## feedback, analyze the problem and regenerate.
```

```
##DESIGN CONCEPT##:

Composition: {composition}

Text description: {text_description}
```

S3 Evolutionary operator over motif profiles

New candidates are produced by editing motif profiles. The operator (EvoOperator, Methods 3) takes the profiles of one or two parent structures and generates an offspring profile T_{off} through one of three moves: crossover, combining motifs drawn from two parents; element mutation, substituting the species within a parent's motifs; or geometry mutation, altering a motif's coordination or connectivity. This is the only component that reasons about properties, and it does so only to *propose* edits—the ERS (Methods 5) later ranks the resulting candidates statistically—while the motif profiles that the translation agents pass onward are strictly geometry-only (Section S1). The resulting offspring profile is emitted as a target for Text2CIF. The population is initialized from $\mathcal{D}_{\text{seed}}$ and advanced for four generations. In the reported run only the *crossover* operator is used: each design step selects two parents from the current pool—one drawn from the 200 most stable structures and one from the 200 of highest saturation magnetization—and produces a

single offspring, so that every candidate combines a stability-favoured and a magnetization-favoured parent; the element- and geometry-mutation operators are provided by the framework but were not applied here. Design steps are repeated to populate each generation; the ERS ranks the offspring, and every DFT-validated candidate (Methods 5–6) joins the pool from which the next generation's parents are drawn.

Design-agent prompts. The design agent returns, as strict JSON, one offspring design—its motif set (`motifs` in the JSON), an integral composition, and a self-contained text description that TEXT2CIF can realize without parent context. The system prompt and the two task templates (crossover, which receives two parents; mutation, which receives one) are reproduced below; {`operator`} and {`description_*`} are filled at run time.

```
You are a materials design assistant for discovering novel RE:TM:X magnetic intermetallic compounds.

PRIMARY OBJECTIVE:
Design novel magnetic material structures that achieve:
  1. HIGH STABILITY - Thermodynamically stable, resistant to decomposition
  2. SYNTHESIZABILITY - Can be practically synthesized in laboratory
  3. HIGH MAGNETIZATION - Strong magnetic properties (high Ms, high Tc, good anisotropy)

All design decisions must be justified in terms of these three objectives.

INPUT:
For each parent structure, you will receive TEXT REPRESENTATION containing:

   - ### MOTIFS: Local structural patterns with functional descriptions

DESIGN OPERATORS:

You have THREE operators to design improved structures.

================================================================================
1) CROSSOVER - Combine motifs from multiple parents
================================================================================

   PURPOSE: Create superior motif sets by selecting from parent motif pools.

   GUIDANCE:
   - Selecting SUBSETS of motifs (not all motifs need to be in every skeleton); REORGANIZING motif
     combinations (different structural arrangements); MODIFYING motif connections (different attachment
     patterns)
   - Ensure selected motifs have compatible coordination environments
   - Balance functional roles: exchange, anisotropy, moment, stability
   - Explore different motif combinations to create diverse stoichiometries

================================================================================
2) ELEMENT MUTATION - Modify elemental composition
================================================================================

   PURPOSE: Improve properties by changing elements in motifs.

   GUIDANCE:
   - Substitute elements to enhance stability and/or magnetization
   - Add or remove elements to optimize the objectives
   - Slight geometry adjustment allowed if needed

================================================================================
3) GEOMETRY MUTATION - Modify coordination geometry
================================================================================

   PURPOSE: Improve properties by changing local geometry of motifs.

   GUIDANCE:
   - Change geometry to enhance functional contributions
```

```
   - Modify coordination to improve stability and/or magnetization
   - Slight element adjustment allowed if needed

DESIGN PROCESS:

STEP 1: ANALYZE PARENT MOTIFS
  - For each motif, identify its functional role (exchange, anisotropy, moment, stability)
  - Evaluate how each motif contributes to or limits the objectives
  - Determine which motifs to preserve, modify, or replace

STEP 2: DESIGN MOTIF SET AND COMPOSITION
  - Apply operators (CROSSOVER, ELEMENT MUTATION, GEOMETRY MUTATION)
  - Ensure designed motifs can coexist in a single structure
  - Derive target composition from the designed motif set
  - Write comprehensive reasoning that guides the generator:
    * How motifs were selected/modified from parent(s)
    * Key structural requirements (coordination, sites, distances)
    * Why this design achieves the objectives

CONSTRAINTS:
- Composition MUST be integral (no fractional occupancies)
- NO micro-tuning (no "slight adjustments", "strain", "pressure")
- All motifs must be physically reasonable and coexist in a single structure

OUTPUT FORMAT (STRICT JSON ONLY):
{
  "designs": [
    {
      "motifs": [
         {"motif #<ID of motif. The ID strictly starts from 1>": "<Text description of the motif>"}
      ],
      "composition": "<INTEGRAL formula>",
      "text_description": "Text description of the ideal structure for the idea design"
    }
  ],
}


RULES:
- Text description must be detailed enough for generator to create the correct structure
- NO code fences. Return JSON only.
- *Strict note*: motifs and text_description must be *self-contained* for a generator with no parent context.
      Do not define motifs by parent indices (e.g., A#2, B#1).
- All motifs must be physically reasonable and coexist in a single structure that improve STABILITY and/or
     MAGNETIZATION over the parent.
```

```
## TASK:
{operator}

Design one novel structure that improve STABILITY and/or MAGNETIZATION over both parents.


OUTPUT: Return STRICT JSON.
---INPUT PARENTS---

## PARENT A: {description_1}
## PARENT B: {description_2}
```

```
## TASK:
{operator}

Design one novel structure that improve STABILITY and/or MAGNETIZATION over the parent.


OUTPUT: Return STRICT JSON.
---INPUT PARENTS---
```

```
## PARENT A: {description_1}
```

S4 Evaluation metrics

S4.1 Benchmark rates

All generation methods are scored on a common footing. For a method producing a set $\mathcal{C}$ of DFT-relaxed candidate structures ($N = |\mathcal{C}| = 400$ for every method), every rate below is a fraction of N. In the head-to-head benchmark each method's 400 DFT relaxations are spent on candidates selected by its own ranking procedure from a disclosed raw-proposal pool (baselines: top 400 of 1,000 by CHGNet-estimated formation energy and magnetic moment; MatEvolve: top 100 of 200 per generation over four generations by the ERS), so the reported rates are conditional on each method's selection step—each method evaluated in its intended mode of use. The full proposal pool comprises all 800 MatEvolve proposals, selected and unselected, every one relaxed with DFT.

We fix a stability threshold $\varepsilon = 0.1$ eV/atom, a design-target magnetization threshold (for this study) $\tau = 1$ T, and a reference database $\mathcal{D}_{\text{ref}}$ = Materials Project ∪ Alexandria. For a candidate c we define the predicates: stable(c) iff $E_{\text{hull}}(c) \leq \varepsilon$; unique($c$) iff c is not a `StructureMatcher` duplicate of another candidate in $\mathcal{C}$ (one representative per duplicate class is counted); novel(c) iff c matches no entry of $\mathcal{D}_{\text{ref}}$ under `StructureMatcher`; $\mathcal{P}(c)$ iff c satisfies the design-target property criterion (here $\mathcal{P} : \mu_0 M_s(c) > \tau$); and newproto($c$) iff c realises a structural prototype absent from $\mathcal{D}_{\text{ref}}$ (defined below).

Following the crystal-structure generation literature [2, 3], a candidate is **S.U.N.** when it is simultaneously *stable*, *unique* and *novel*. On this established base, the design target—a rare-earth-lean permanent-magnet compound beyond the known structure types—sets two further requirements, which we assess separately and together:

$$\text{S.U.N.} = \big\{c \in \mathcal{C} : \text{stable}(c) \wedge \text{unique}(c) \wedge \text{novel}(c)\big\}, \tag{S1}$$

$$\text{magnetic} = \big\{c \in \text{S.U.N.} \ : \ \mathcal{P}(c)\big\}, \tag{S2}$$

$$\text{new prototype} = \big\{c \in \text{S.U.N.} \ : \ \text{newproto}(c)\big\}, \tag{S3}$$

$$\text{joint} = \big\{c \in \text{S.U.N.} \ : \ \text{newproto}(c) \wedge \mathcal{P}(c)\big\}. \tag{S4}$$

The magnetic and new-prototype sets are subsets of S.U.N., but neither contains the other: a structure may be magnetic only, a new prototype only, both, or neither; the joint set is their intersection. The four *rates* reported throughout are the corresponding fractions |S.U.N.|/N, |magnetic|/N, |new prototype|/N and |joint|/N: the base S.U.N. rate asks whether a method finds stable new materials at all; the *magnetic* rate asks whether they exhibit the target property; the *new-prototype* rate asks whether they are new arrangements rather than substitutions on known lattices, the requirement for extending the known magnet-compound families; and the *joint* rate asks whether a single candidate meets both requirements, as the design target demands. For the permanent-magnet target of this study, the design-target property is $\mathcal{P} : \mu_0 M_s > 1$ T; the 1 T threshold marks an appreciable ferromagnetic moment of permanent-magnet relevance ($Nd_2Fe_{14}B$ reaches $\mu_0 M_s \approx 1.6$ T).

The prototype predicate uses a label Λ computed in two steps, identically for candidates and for every reference entry. (i) The structure is symmetrized (pymatgen `SpacegroupAnalyzer`, `symprec` = 0.1 Å). (ii) It is encoded as an AFLOW prototype string [4, 5]: anonymized composition, Pearson symbol, space-group number and per-species Wyckoff positions, with species lettered by ascending stoichiometry (ties broken by Wyckoff signature, never element identity).

Because species are anonymized, an elemental substitution that preserves stoichiometry, symmetry and site occupation shares its host's label ($SmFe_{12}$ and $DyFe_{12}$ share a label), whereas a change in any of the three yields a new one. Then newproto(c) holds iff $\Lambda(c) \notin \{\Lambda(r) : r \in \mathcal{D}_{\text{ref}}\}$. This single rule is applied to every method and chemistry, so the cross-method comparison is not biased by a magnet-specific taxonomy. The same reference $\mathcal{D}_{\text{ref}}$ and prototype rule apply to the full proposal pool (all 800) and to the benchmark set (the 400 selected by the ERS), so the two are directly comparable.

S4.2 Reconstruction fidelity measures

Experiment 1 (main text, Section 2.3) compares each reconstruction returned by TEXT2CIF with the original Materials Project structure from which its profile was written. The comparison uses each reconstruction as Motif Translate emitted it, without any DFT relaxation. Four measures are used.

Match rate. A reconstruction is counted as matching its original when the pair is accepted by pymatgen's `StructureMatcher` at the tolerances under which generative models report match rates: `ltol` $= 0.3$, `stol` $= 0.5$ and `angle_tol` $= 10^\circ$.

Root-mean-square displacement (RMSD). For every pair, the RMSD is the root-mean-square distance, in Ångströms, between corresponding atomic positions after optimal superposition of the two structures; it is computed for all 948 pairs.

Continuous shape measure (CShM). Shape fidelity is quantified with the continuous shape measure [6, 7],

$$\mathcal{S} = 100 \, \min \frac{\sum_i |q_i - p_i|^2}{\sum_i |q_i - q_0|^2}, \tag{S5}$$

where q_i are the atomic positions of the reference (the original structure), p_i those of the reconstruction, q_0 the reference centroid, and the minimization runs over translation, rotation, uniform scaling and atom correspondence. The measure is dimensionless and scores the deviation of a structure from its reference on a scale of 0–100: $\mathcal{S} = 0$ denotes an identical match and larger values greater distortion. Values were computed with SHAPE and cosymlib [8, 9].

Two remarks govern how the CShM is read in the main text. First, because the denominator is the second moment of the point set about its centroid, a given atomic displacement yields a smaller value in a more extended point set; a unit cell is more extended than a single coordination polyhedron, so cell-level values are systematically smaller than polyhedron-level ones for the same displacement. Second, the interpretive bands quoted in Section 2.3 (essentially ideal below about 0.1, slightly distorted up to 3, notably distorted from 3 to 7.5, and beyond 7.5 a breakdown of the target geometry [6, 10–12]) were calibrated on single coordination polyhedra. They are therefore used only to order the reconstructions and describe the distribution, not as an acceptance criterion; acceptance within the pipeline is decided by the round-trip tests of Methods 2.

Local-environment fidelity (OFM). The orbital-field matrix (OFM) [13, 14] encodes the coordination of every atom in a structure—the valence-orbital types of each atom and of its neighbors, weighted by the solid angles of its Voronoi polyhedron—and averages these local descriptors into one fixed-length vector per structure. Because the descriptor is an average over sites, it needs no atom correspondence and applies to all 948 original–reconstruction pairs. The distance between two structures is the L_1 norm of the difference of their OFM vectors. Two quantities are reported in the main text (Section 2.3, Fig. 3c): the distance between each original and its own reconstruction, compared with the distances from that original to its 100 nearest neighbors among the other originals of $\mathcal{D}_{\text{seed}}$; and, for each original, the rank of its own reconstruction among all other

structures (the remaining originals and every reconstruction), from which the fraction of originals whose reconstruction is the nearest neighbor, or lies among the 10 nearest, is obtained.

S5 Structure map by motif-group substitutability

To relate structures through their motif content and to quantify combinatorial coverage (Fig. 6 of the main text), each structure is represented by the *set* of the 17 motif groups it contains. The distance between two structures S_i and S_j combines their set overlap with the substitutability of the groups by which they differ, following the construction we introduced for alloy compositions [14, 15]: $\overline{M}[i,j] = d_{\mathrm{sub}}(S_i, S_j)\,\big(1 - J(S_i, S_j)\big)$, where J is the Jaccard overlap of the two group sets [16] and d_{sub} is a substitutability distance between the differing groups—small when those groups are readily interchangeable—derived from the Dempster–Shafer substitutability belief M of Methods 5, whose evidence is the labelled seed structures augmented each generation by the DFT-validated candidates, discounted by the same uncertainty parameter $\alpha = 0.1$ used there. Structures are embedded from $\overline{M}$ using t-SNE [17]. Combination coverage (Fig. 6b of the main text) is obtained by enumerating, for each order $k \in \{2, 3, 4\}$, the $\binom{17}{k}$ possible group combinations and labelling each as observed in the seed set, the generated set, both, or neither. Pair stability rates (Fig. 6c of the main text) are the fraction of structures containing a given group pair that lie within 0.1 eV/atom of the convex hull; pairs whose rate is at least 0.75 are highlighted. Because all $\binom{17}{2}$ pairs are examined at varying and sometimes small support, this threshold identifies a robustly supported class of compatible pairings rather than individually significance-tested pairs.

S6 Motif-group word clouds

Figure 5a of the main text labels each of the 17 motif groups by a word cloud of the controlled-vocabulary tags (SI S1) that recur among its motifs; because the terms are the same bracketed geometry tags that name the motifs, each word cloud reads in the grammar of the representation rather than in free text. The construction (Algorithm S1) has three stages. First, five independent agents extract, from each group's standardized motif descriptions, the vocabulary tags present—reusing the standardizer definitions of SI S1. Second, the five tag sets of a group are pooled and deduplicated, each tag normalized to its canonical vocabulary form. Third, the pooled tags are weighted by term frequency–inverse document frequency (TF–IDF)—treating a group's tags as one document and the 17 groups as the corpus, so that a tag's weight rises with its frequency within a group and falls with the number of groups sharing it—and the 50 highest-weighted tags of each group are rendered with type size proportional to TF–IDF weight.

S7 Relation to text-based crystal representations

Table S2 situates the motif profile among prior text-based treatments of crystals, ordered by the direction of the crystal–text translation. Both endpoint directions are established: automated description turns a crystal into text, and text-conditioned generation turns a specification back into a crystal. The motif profile differs from both. Against description, it is written in a geometry-only controlled vocabulary that withholds the coordinates, lattice parameters, space group, and structure-type label that a description reports, and it is certified generative—accepted only if a matching structure can be rebuilt from it (Methods 2). Against text-conditioned generation, its realization step starts not from a composition-and-symmetry prompt but from the detailed motif profile itself, and is one leg of a design loop rather than the goal.

Algorithm S1 Motif-group word-cloud construction

Input: motif groups $\mathcal{G}$ ($|\mathcal{G}| = 17$); standardized motif descriptions (SI S1); controlled vocabulary $\mathcal{V}$; workers $R = 5$
Output: the 50 highest-weighted vocabulary tags per group

for each group $g \in \mathcal{G}$ **do**
  for $r = 1$ **to** R **do**
    $K_{g,r} \leftarrow$ ExtractTags(descriptions of g, $\mathcal{V}$) ▷ *independent agent; SI S1 standardizer*
  end for
  $K_g \leftarrow \textsc{Dedup}\big(\bigcup_r K_{g,r}\big)$ ▷ *pool the 5 agents; normalize to canonical* $\mathcal{V}$
end for
for each tag t **do** $\mathrm{df}(t) \leftarrow |\{g : t \in K_g\}|$ ▷ *groups containing* t
for each group g and each tag $t \in K_g$ **do**
  $\mathrm{tf}(t, g) \leftarrow$ occurrences of t among group g's motif tags
  $w(t, g) \leftarrow \mathrm{tf}(t, g) \cdot \log\big(|\mathcal{G}|/\mathrm{df}(t)\big)$ ▷ *TF–IDF weight*
end for
render the 50 highest-$w(\cdot, g)$ tags of each group g, type size $\propto w$

Table S2: Prior text-based treatments of crystals and their relation to MatEvolve.

Work	**Direction**	**Purpose**	**Relation to MatEvolve**
Robocrystallographer [18]	Crystal → text	Structure description	Direct precedent of CIF2Text; its text reports space group and prototype and is an end product, not an editable object
LLM-Prop [19]	Text → property	Property prediction	Establishes text as a machine-readable crystal representation
Lattice Lingo [20]	Crystal → text → property	Multimodal property prediction	Establishes the value of semiglobal textual information
Ozawa et al. [21]	Representation → property	Input-representation comparison	Establishes that representation design governs model performance
CrysText [22]	Text → crystal	Crystal generation	Direct analogue of Text2CIF, conditioned on composition and symmetry rather than a motif profile

Two of these results bear directly on the profile's design. First, Lattice Lingo reports that semiglobal textual information—connectivity and structural arrangement—contributes most to property prediction beyond graph-encoded local structure [20]. The motif vocabulary (Section S1) deliberately spans this range, from local coordination units through cages, nets, layers, and chains to connectivity and dimensionality, so the finding supports the vocabulary's reach as a design choice rather than an assertion. Second, Ozawa et al. report that input representations carrying space-group information outperform composition-only inputs for property prediction [21]; the motif profile nonetheless withholds the space group, deliberately. A space-group label would be invalidated by almost any motif edit, making the profile brittle as an editing medium, and stating it would hand the round-trip test part of its answer, letting the generator reconstruct the lattice without using the motifs. The cost of this omission is what Experiment 1 measures.

S8 First-principles settings and reference databases

All relaxations used one set of VASP input parameters: PREC = Accurate, ENCUT = 520 eV, EDIFF = 10^{-4} eV, ALGO = Fast, NELM = 100, ISPIN = 2, IBRION = 2, ISIF = 3, NSW = 50, EDIFFG at its default (10^{-3} eV, energy-based), ISMEAR = -5, KSPACING = 0.25 Å^{-1} with Γ-centered meshes (KGAMMA = .TRUE.), LASPH = .TRUE., LMAXMIX = 6, LREAL = Auto and LORBIT = 11. Table S3 lists the PAW potential and the initial magnetic moment m_0 assigned to each element. With the _3 potentials (Pr, Nd, Pm, Sm, Tb, Dy, Ho, Er, Tm, Lu) the $4f$ electrons are kept in the core; Ce, Eu and Gd are treated with their standard potentials and Yb with Yb_2.

Table S3: PAW potentials (POTCAR library version 5.4) and initial magnetic moments m_0 (μ_B) used in every spin-polarized relaxation; "–": no initial moment defined.

El.	PAW	m_0	El.	PAW	m_0	El.	PAW	m_0	El.	PAW	m_0
Ac	Ac	–	F	F	–	Nb	Nb_pv	0.6	Si	Si	0.6
Ag	Ag	0.6	Fe	Fe_pv	5.0	Nd	Nd_3	0.6	Sm	Sm_3	0.6
Al	Al	0.6	Ga	Ga_d	0.6	Ne	Ne	–	Sn	Sn_d	0.6
Ar	Ar	–	Gd	Gd	0.6	Ni	Ni_pv	5.0	Sr	Sr_sv	0.6
As	As	0.6	Ge	Ge_d	0.6	Np	Np	–	Ta	Ta_pv	0.6
Au	Au	0.6	H	H	0.6	O	O	–	Tb	Tb_3	0.6
B	B	0.6	He	He	–	Os	Os_pv	0.6	Tc	Tc_pv	–
Ba	Ba_sv	0.6	Hf	Hf_pv	0.6	P	P	0.6	Te	Te	0.6
Be	Be_sv	0.6	Hg	Hg	0.6	Pa	Pa	–	Th	Th	0.6
Bi	Bi	0.6	Ho	Ho_3	0.6	Pb	Pb_d	0.6	Ti	Ti_pv	0.6
Br	Br	0.6	I	I	0.6	Pd	Pd	0.6	Tl	Tl_d	0.6
C	C	0.6	In	In_d	0.6	Pm	Pm_3	–	Tm	Tm_3	0.6
Ca	Ca_sv	0.6	Ir	Ir	0.6	Pr	Pr_3	0.6	U	U	0.6
Cd	Cd	0.6	K	K_sv	0.6	Pt	Pt	0.6	V	V_pv	5.0
Ce	Ce	5.0	Kr	Kr	–	Pu	Pu	0.6	W	W_sv	5.0
Cl	Cl	0.6	La	La	0.6	Rb	Rb_sv	0.6	Xe	Xe	–
Co	Co	0.6	Li	Li_sv	0.6	Re	Re_pv	0.6	Y	Y_sv	0.6
Cr	Cr_pv	5.0	Lu	Lu_3	0.6	Rh	Rh_pv	0.6	Yb	Yb_2	0.6
Cs	Cs_sv	0.6	Mg	Mg_pv	0.6	Ru	Ru_pv	0.6	Zn	Zn	0.6
Cu	Cu_pv	0.6	Mn	Mn_pv	5.0	S	S	0.6	Zr	Zr_sv	0.6
Dy	Dy_3	0.6	Mo	Mo_pv	5.0	Sb	Sb	0.6			
Er	Er_3	0.6	N	N	0.6	Sc	Sc_sv	0.6			
Eu	Eu	10.0	Na	Na_pv	0.6	Se	Se	0.6			

Reference databases. Novelty and prototype status are assessed against $\mathcal{D}_{\text{ref}}$ = Materials Project (database version v2026.04.13) ∪ Alexandria (PBE release 2025.07.02), and the convex hulls are

built from the same Materials Project version.

References

[1] B. Cordero, V. Gómez, A. E. Platero-Prats, M. Revés, J. Echeverría, E. Cremades, F. Barragán, S. Alvarez, Covalent radii revisited, Dalton Transactions (2008) 2832–2838.

[2] T. Xie, X. Fu, O.-E. Ganea, R. Barzilay, T. Jaakkola, Crystal diffusion variational autoencoder for periodic material generation, in: International Conference on Learning Representations (ICLR), 2022. ArXiv:2110.06197.

[3] C. Zeni, R. Pinsler, D. Zügner, A. Fowler, M. Horton, X. Fu, S. Shysheya, J. Crabbé, L. Sun, J. Smith, B. Nguyen, H. Schulz, S. Lewis, C.-W. Huang, Z. Lu, Y. Zhou, H. Yang, H. Hao, J. Li, C. Yang, W. Li, R. Tomioka, T. Xie, A generative model for inorganic materials design, Nature 639 (2025) 624–632. doi:10.1038/s41586-025-08628-5.

[4] M. J. Mehl, D. Hicks, C. Toher, O. Levy, G. L. W. Hart, S. Curtarolo, et al., The AFLOW library of crystallographic prototypes: Part 1, Computational Materials Science 136 (2017) S1–S828. doi:10.1016/j.commatsci.2017.01.017.

[5] D. Hicks, M. J. Mehl, M. Esters, C. Oses, O. Levy, G. L. W. Hart, C. Toher, S. Curtarolo, The AFLOW library of crystallographic prototypes: Part 3, Computational Materials Science 199 (2021) 110450. doi:10.1016/j.commatsci.2021.110450.

[6] M. Pinsky, D. Avnir, Continuous symmetry measures. 5. the classical polyhedra, Inorganic Chemistry 37 (1998) 5575–5582. doi:10.1021/ic9804925.

[7] H. Zabrodsky, S. Peleg, D. Avnir, Continuous symmetry measures, Journal of the American Chemical Society 114 (1992) 7843–7851. doi:10.1021/ja00046a033.

[8] M. Llunell, D. Casanova, J. Cirera, P. Alemany, S. Alvarez, SHAPE, version 2.1, Universitat de Barcelona, Barcelona, 2013.

[9] A. Carreras, D. Casanova, P. Alemany, S. Alvarez, Cosymlib: A python library for continuous symmetry measures, Journal of Chemical Information and Modeling (2024). doi:10.1021/acs.jcim.4c00609.

[10] S. Alvarez, P. Alemany, D. Casanova, J. Cirera, M. Llunell, D. Avnir, Shape maps and polyhedral interconversion paths in transition metal chemistry, Coordination Chemistry Reviews 249 (2005) 1693–1708. doi:10.1016/j.ccr.2005.03.031.

[11] L. Shao, Y. Li, J. Yu, R. Xu, Divalent-metal-stabilized aluminophosphates exhibiting a new zeolite framework topology, Inorganic Chemistry 51 (2012) 225–229. doi:10.1021/ic201515z.

[12] A. V. Funes, L. Carrella, Y. Rechkemmer, J. van Slageren, E. Rentschler, P. Alborés, Synthesis, structural characterization and magnetic behaviour of a family of $[Co_2^{III}Ln_2^{III}]$ butterfly compounds, Dalton Transactions 46 (2017) 3400–3409. doi:10.1039/c6dt04713k.

[13] T. L. Pham, H. Kino, K. Terakura, T. Miyake, K. Tsuda, I. Takigawa, H. C. Dam, Machine learning reveals orbital interaction in materials, Science and Technology of Advanced Materials 18 (2017) 756–765. doi:10.1080/14686996.2017.1378060.

[14] T.-L. Pham, D.-N. Nguyen, M.-Q. Ha, H. Kino, T. Miyake, H.-C. Dam, Explainable machine learning for materials discovery: predicting the potentially formable Nd–Fe–B crystal structures and extracting the structure–stability relationship, IUCrJ 7 (2020) 1036–1047. URL: https://doi.org/10.1107/S2052252520010088. doi:10.1107/S2052252520010088.

[15] M.-Q. Ha, D.-N. Nguyen, V.-C. Nguyen, T. Nagata, T. Chikyow, H. Kino, T. Miyake, T. Denœux, V.-N. Huynh, H.-C. Dam, Evidence-based recommender system for high-entropy alloys, Nature Computational Science 1 (2021) 470–478. URL: https://doi.org/10.1038/s43588-021-00097-w. doi:10.1038/s43588-021-00097-w.

[16] P. Jaccard, The distribution of the flora in the alpine zone, New Phytologist 11 (1912) 37–50. doi:10.1111/j.1469-8137.1912.tb05611.x.

[17] L. van der Maaten, G. Hinton, Visualizing data using t-SNE, Journal of Machine Learning Research 9 (2008) 2579–2605.

[18] A. M. Ganose, A. Jain, Robocrystallographer: automated crystal structure text descriptions and analysis, MRS Communications 9 (2019) 874–881. doi:10.1557/mrc.2019.94.

[19] A. Niyongabo Rubungo, C. Arnold, B. P. Rand, A. B. Dieng, Llm-prop: predicting the properties of crystalline materials using large language models, npj Computational Materials 11 (2025) 186. doi:10.1038/s41524-025-01536-2.

[20] M. Munjal, J. Lee, C. Park, S. Han, Lattice lingo: Effect of textual detail on multimodal learning for property prediction of crystals, arXiv:2412.04670, 2024. doi:10.48550/arXiv.2412.04670.

[21] S. Ozawa, I. Takahara, T. Mizoguchi, Scale-dependent input representation and confidence estimation for LLMs in materials property prediction, arXiv:2605.03515, 2026. doi:10.48550/arXiv.2605.03515.

[22] T. Mohanty, M. Mehta, H. M. Sayeed, B.-E. Oded, I. Pitussi, A. Borenstein, V. Srikumar, T. D. Sparks, Crystext: A generative ai approach for text-conditioned crystal structure generation using llm, Integrating Materials and Manufacturing Innovation 15 (2026) 416–430. doi:10.1007/s40192-026-00451-8.