PRICIN: PRINCIPLE-CENTERED INORGANIC RETROSYNTHESIS

Anonymous authors
Paper under double-blind review

ABSTRACT

Bridging the gap between what is designable by computational discovery and what is synthesizable in the lab remains a central obstacle for closed-loop materials science. We tackle single-step inorganic retrosynthesis and show that explicit chemical principles are potent inductive biases for learning to plan syntheses. We introduce **PRICIN**, a principle-centered approach that reformulates precursor planning around two laws: elemental conservation and electron balance. PRICIN embeds stoichiometry and oxidation-state semantics directly into the target representation via two pretraining objectives, including an auxiliary oxidation-state supervision that injects charge awareness. At inference, a lightweight element-wise filter first predicts the required number of precursors and then prunes candidates that violate conservation constraints, yielding explainable, chemically consistent precursor sets without external retrieval or rigid templates. Across the Retrieval-Retro (year-split) and Ceder benchmarks, PRICIN attains state-of-the-art performance on Top-k and combination Top-k metrics, improving over the previous best by **+5.59** Top-1 and by up to **+19.2** percentages on Top-k. Ablations confirm that oxidation-state supervision and conservation-aware filtering are both necessary and complementary, substantially reducing early-rank errors. The code will be released upon acceptance.

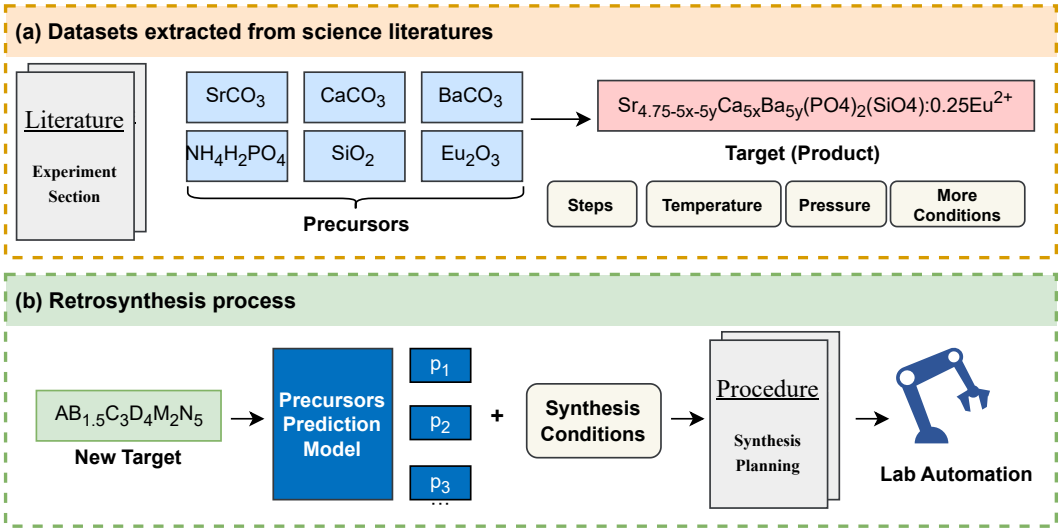

Figure 1: **The inorganic retrosynthesis task.** (a) Prior work has focused on extracting synthesis recipes from scientific literature, where the precursors and conditions for a known target are documented. (b) In contrast, the retrosynthesis process inverts this problem: given a novel target compound, the goal is to predict a set of viable precursors and synthesis conditions. This automated synthesis planning is a key step toward the ultimate vision of using lab automation to accelerate the discovery and synthesis of new materials.

1 INTRODUCTION

The core objective of materials science is to discover and deploy new materials with superior properties for critical applications, including semiconductors Choubisa et al. (2023), energy storage Yao et al. (2023) and bio-materials McDonald et al. (2023). While traditional materials exploration relies heavily on trial and error, the advent of theoretical and first-principles calculations has significantly accelerated the design phase. Systems like GNoME Merchant et al. (2023) and MatterGen Zeni et al. (2023) can now efficiently generate promising compositions and structures from vast candidate spaces. However, with the rise of the closed-loop paradigms such as A-Lab Szymanski et al. (2023b), the primary bottleneck in the full Design-Make-Analyze-Test (DMAT) cycle has become increasingly apparent: many existing models excel at screening and design but struggle to provide actionable synthesis routes, creating a significant gap between what is designable and what is synthesizable.

Similar to organic chemistry, the synthesis of inorganic materials can be framed as a retrosynthesis problem. However, the two fields have significant differences that make direct translation of methods unfeasible: (1) the lack of large-scale, standardized datasets comparable to the USPTO database Somnath et al. (2021); (2) the greater difficulty in calculating the properties and structures of inorganic crystals Ratcliff et al. (2017), which involve larger numbers of atoms and periodic boundary conditions; (3) the critical role of stoichiometry in determining both the target crystal structure and the feasible precursor combinations; and (4) the absence of universally transferable reaction centers Lan et al. (2024) or transition state mechanisms Zhang et al. (2016); Ucak et al. (2022) in inorganic solid-state synthesis, making data-driven modeling more challenging. Consequently, inorganic retrosynthesis has emerged as a key unsolved problem for achieving closed-loop materials discovery.

In recent years, automated methods for inorganic retrosynthesis have begun to emerge. Thermodynamics-based pathfinding approaches frame solid-state synthesis as a search for a feasible path on a reaction network, where edge weights encode proxies for reaction energy and phase competition, successfully reproducing literature routes and proposing candidates for new targets McDermott et al. (2021); Miura et al. (2021). For precursor set generation, early work used text mining and generative models to learn synthesis planning from literature Kim et al. (2020). More recent methods like SynthesisSimilarity He et al. (2023) learn material similarity from historical recipes to recommend precursors by analogy, ElementwiseRetro Kim et al. (2022) uses a template-based GNN ranker, and RetrievalRetro Noh et al. (2024) incorporates a reaction energy-based retriever. While these methods have advanced the field, they primarily rely on precedent-based learning, with limited explicit modeling of chemical principles. This can lead to early-ranking errors, inflexibility in handling multi-source precursors.

In this work, we propose a principle-centered approach (PRICIN) to inorganic retrosynthesis. We reformulate the precursor planning task around two core chemical reaction laws: **elemental conservation** and **electron balance** (the total increase in oxidation numbers equals the total decrease). Instead of relying solely on templates or retrieval, we design two tasks that embed elemental stoichiometry and oxidation-state directly into the target material's representation. One objective uses auxiliary supervision on oxidation states to implicitly model valence changes during reaction, while the other models compositional ratios to ensure elemental conservation. We also introduce a lightweight element-wise filter that first predicts the required number of precursors and then filters out candidates that violate elemental conservation, ensuring that elements of every precursor are sourced validly. Guided by these principles, PRICIN generates explainable and synthesizable precursor sets without relying on external retrieval and significantly increases top-k accuracy.

We systematically evaluate our method on the Retrieval-Retro Noh et al. (2024) and Ceder Kononova et al. (2019) benchmark datasets, with the former using a year-based split to test for temporal generalization. Our results demonstrate state-of-the-art performance across multiple metrics, validating the effectiveness and practical potential of a principle-centered approach to inorganic retrosynthesis.

Our contributions are summarized as follows:

- **Explicit Oxidation-State Supervision.** To the best of our knowledge, we are the first to propose that explicit oxidation-state supervision should be a core component of modeling inorganic retrosynthesis, providing a direct chemical signal that significantly improves prediction accuracy.

Table 1: **Capability comparison of retrosynthesis methods.** Our approach is benchmarked against prior works, highlighting key features for successful precursor prediction: the ability to retrieve from a database of known reactions, the integration of explicit chemical domain knowledge, and the capacity to extrapolate predictions to novel materials.

Model	Retrieval capability	Chemical domain knowledge	Extrapolation to new systems
ElemwiseRetro Kim et al. (2022)	✗	Low	Medium
Synthesis Similarity He et al. (2023)	✓	Low	Low
Retrieval-Retro Noh et al. (2024)	✓	Medium	Medium
Retro-Rank-In Prein et al. (2025)	✗	Low	High
Ours	✓	High	High

- **Principle-Centered Formulation.** We develop a comprehensive modeling framework that reconstructs the retrosynthesis task around two key chemical principles: elemental conservation and electron balance, integrating these rules into both the learning process and inference constraints.
- **Effective Element-Wise Filter.** We employ a simple and highly efficient element-wise filter at inference time that first predicts the number of precursors and then prunes illegal candidates, leading to consistent and substantial accuracy improvements with minimal computational overhead.

2 Related Work

Literature mining and datasets. Recent literature-mining efforts have demonstrated that large-scale, automated extraction of inorganic synthesis knowledge is feasible and directly enables recipe-level datasets. (Kononova et al., 2019) builds a large-scale, automatically text-mined corpus of solid-state synthesis "recipes" by scraping articles, detecting synthesis paragraphs and converting text into structured JSON records that capture targets, precursors, operations. (Huo et al., 2019) uses semi-supervised topic modeling to discover interpretable step-topics and classifies synthesis modalities while reconstructing procedural order. (He et al., 2020) develops a two-step pipeline that masks material mentions and infers roles with a BiLSTM-CRF, assembling a large corpus of precursors and targets and proposing a precursor-similarity metric that supports reactant substitution. (Wang et al., 2022b) extends mining to solution-based syntheses with a publisher-scale pipeline combining a BERT-based paragraph classifier. (Wang et al., 2022a) introduces a unified ontology (ULSA) and a learned mapping from text to standardized action graphs, providing a common procedural vocabulary, allowing operation prediction and full-step synthesis planning.

Precursor recommendation. (McDermott et al., 2021) cast solid-state synthesis planning as pathfinding on a thermochemistry-derived reaction network whose edge weights encode thermodynamic proxies, recovering literature routes (e.g., $YMnO_3$, $Y_2Mn_2O_7$, Fe_2SiS_4, $YBa_2Cu_3O_{6.5}$) and proposing routes to unseen targets. (Miura et al., 2021) recast ceramic synthesis as a sequence of pairwise interfacial reactions, rank interface reactivity via ab-initio thermodynamics, and predict the earliest nonequilibrium intermediates that steer phase evolution. (Kim et al., 2020) mine the materials-science literature with an NLP pipeline (ELMo/FastText embeddings and NER) and train an unsupervised conditional VAE to model synthesis actions and precursors conditioned on a target compound, retrospectively proposing plausible precursors for unseen perovskites such as $InWO_3$ and $PbMoO_3$ while providing literature-trained representations that complement thermodynamic checks. He et al. (2023) propose a data-driven strategy (SynthesisSimilarity) that learns a neural notion of chemical similarity from 29,900 literature recipes to recommend precursor sets for novel targets by analogy to historically synthesized materials. Kim et al. (2022) develop a graph neural framework (ElementwiseRetro) that ranks precursor sets under a probabilistic template requiring each target element be sourced from exactly one precursor, a constraint that can limit flexibility for multi-source routes. Noh et al. (2024) introduce a retrieval-based approach (RetrievalRetro) that implicitly extracts precursor information from reference materials and injects thermodynamic priors via a Neural Reaction Energy retriever, yet can propagate early ranking errors to higher-k lists.

Applications and LLMs. (Chen et al., 2024) formalizes a thermodynamic strategy for solid-state synthesis and validates these principles robotically across hundreds of reactions with routinely higher phase purity than traditional recipes. Song et al. (2025) introduces CSLLM, a domain-adapted LLM that predicts synthesizability, methods, and precursors from a crystal-text representation and a large curated corpus, though without explicitly enforcing charge balance. Several other domain-specialized language models have been developed for materials science, such as MatBERT Walker et al. (2021) and MatSciBERT Gupta et al. (2022). Szymanski et al. (2023a) presents A-Lab, a closed-loop platform that fuses ab initio phase-stability priors, literature-trained recipe models, robotics, ML-based XRD, and active learning to reliably realize computationally proposed oxides and phosphates.

3 PRELIMINARY

We consider the single–step inorganic retrosynthesis problem: given a target compound x and synthesis conditions $(\mathcal{T}, \mathcal{P})$, predict a multiset of precursors $C = \{p_i\}_{i=1}^{m}$ together with stoichiometric coefficients $\mathbf{s} = (s_1, \ldots, s_m)$ and optional byproducts $B = \{b_j\}_{j=1}^{r}$ with coefficients $\mathbf{t} = (t_1, \ldots, t_r)$ such that the reaction is chemically consistent and thermodynamically driven.

Elemental conservation. Let $n(E, y)$ denote the count of element E in material y per formula unit. Element balance requires that there exist non–negative integers $\{s_x\}, \{s_i\}$ and $\{t_j\}$ satisfying, for all E in the element set $\mathcal{E}$,

$$\sum_{i=1}^{m} s_i \, n(E, p_i) \;=\; s_x \, n(E, x) \;+\; \sum_{j=1}^{r} t_j \, n(E, b_j) \, . \tag{1}$$

In practice, when multiple precursor sets may synthesize the same target, we relax per-atom balance to the following element-level formulation: We write $M(y)$ for "the set of elements that appear in material y." The coverage constraint is simply

$$\bigcup_{i=1}^{m} M(p_i) \;=\; M(x) \;\cup\; \bigcup_{j=1}^{r} M(b_j). \tag{2}$$

Electron balance. Let $z(E, \cdot)$ denote admissible oxidation states. Chemical principles dictate that each compound must be charge–neutral and that the overall reaction must be redox-balanced. To formalize charge neutrality, we account for elements with multiple oxidation states by defining the net charge of a compound y as a sum over its constituent species σ (element-oxidation pairs): $Q(y) = \sum_\sigma z(\sigma, y)\, n(\sigma, y)$, where $z(\sigma, y)$ is the charge of species σ and $n(\sigma, y)$ is its count. Consequently, for any valid compound: $Q(x) = 0, Q(p_i) = 0, Q(b_j) = 0$. Beyond the neutrality of individual compounds, a valid chemical reaction must also maintain redox balance. This principle requires that the sum of oxidation state changes across all elements in the reaction be zero, meaning any oxidation is precisely balanced by a reduction. Therefore, the valence states of the precursors are fundamentally linked to those of the target material, motivating our subsequent use of precursor oxidation states as an informative signal for our model.

Problem statement. Inorganic retrosynthesis seeks to find and rank precursor sets $(C, \mathbf{s})$ (and optional $B, \mathbf{t}$) that satisfy elemental conservation and charge consistency. Downstream sections instantiate this definition with learnable embeddings, retrieval evidence, and decision–time constraint checks.

4 METHODS

Overview. Our proposed pipeline, illustrated in Figure 2, consists of a multitask training stage followed by a constrained ranking stage. Given a target compound x, we first apply a fixed feature extraction module $h(\cdot)$ to obtain its compositional representation, $\mathbf{x}_{\text{feat}} = h(x)$. This feature vector is then passed to a trainable encoder $f_\theta(\cdot)$ to produce a chemically-aware target embedding in the latent space, $\mathbf{t} = f_\theta(\mathbf{x}_{\text{feat}})$. This embedding is jointly optimized against three objectives using

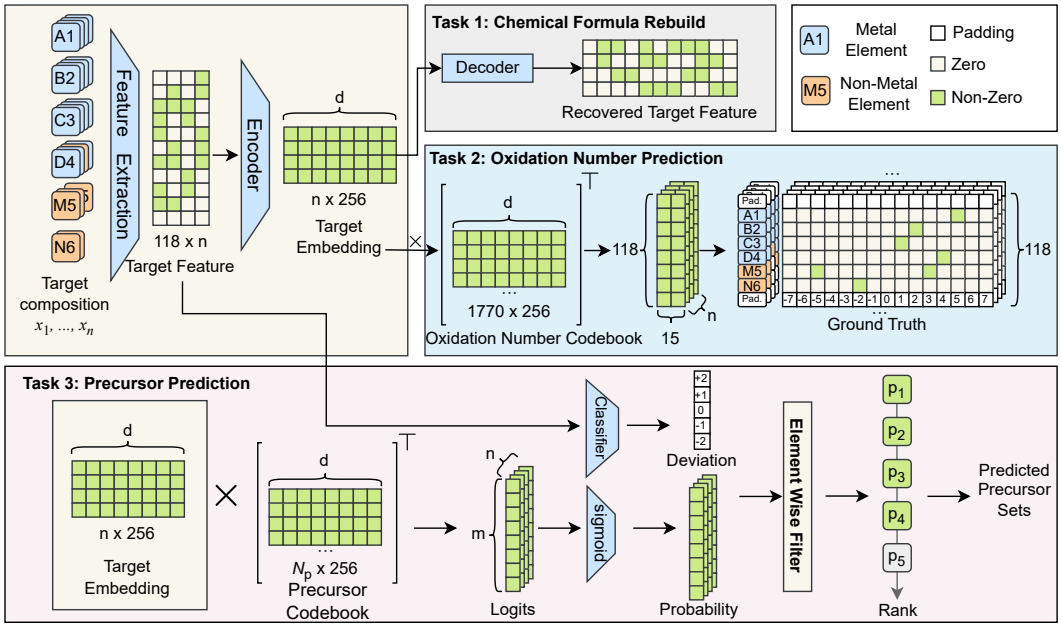

Figure 2: **Overview of our chemical principle-centered pipeline.** The model takes a target composition, extracts features, and encodes it into a fixed-dimensional target embedding. This embedding is trained via three auxiliary tasks: (1) **Chemical Formula Rebuild**, where a decoder reconstructs the elemental fractions of the target to preserve stoichiometric information; (2) **Oxidation Number Prediction**, where the model learns to predict the distribution of oxidation states for each element by comparing the target embedding to a species codebook, thereby injecting oxidation-awareness; and (3) **Precursor Prediction**, where the target embedding is used to rank candidates from a precursor codebook. In the final step, an **Element-wise Filter** is applied to the ranked precursor probabilities to enforce elemental conservation, yielding a list of chemically coherent precursor sets.

two distinct learnable codebooks: an **Oxidation State Codebook**, $\mathcal{C}_{\text{ox}} \in \mathbb{R}^{N_{ox} \times d}$, and a **Precursor Codebook**, $\mathcal{C}_{\text{prec}} \in \mathbb{R}^{N_p \times d}$. The objectives are: reconstructing the initial compositional features, predicting oxidation states, and predicting the precursor set. Finally, the ranking stage uses the optimized target embedding to generate precursor recommendations, which are refined by a element wise filter that enforces elemental conservation.

4.1 Task 1: Chemical Formula Rebuild.

The goal of this task is to train the encoder f_θ to produce a robust target embedding $\mathbf{t}$ that retains the essential stoichiometric information of the original compound. To achieve this, we employ a decoder, $d_\phi(\cdot)$, which attempts to reconstruct the initial target feature vector from the embedding, i.e., $\hat{\mathbf{x}}_{\text{feat}} = d_\phi(\mathbf{t})$. This autoencoding structure ensures that the learned embedding is a compressed but faithful representation of the compound's composition. The reconstruction loss is formulated as a binary cross-entropy between the original and reconstructed feature vectors:

$$\mathcal{L}_{\text{frac}} = -\sum_{i=1}^{|\mathcal{E}|} \Big[x_{\text{feat},i} \log(\hat{x}_{\text{feat},i}) + (1 - x_{\text{feat},i}) \log(1 - \hat{x}_{\text{feat},i}) \Big], \tag{3}$$

where $\mathbf{x}_{\text{feat}}$ is the ground-truth feature vector and $\hat{\mathbf{x}}_{\text{feat}}$ is the reconstructed vector. This task serves as a regularization objective that enhances the stability and generalization of downstream tasks.

4.2 Task 2: Oxidation Number prediction.

Chemical reactions are governed by charge conservation, but the oxidation states in a novel target compound are often unknown (e.g., $Ba_{0.5}Sr_{0.5}Co_xFe_{1-x}O_{3-\delta}$). We address this by predicting the target's oxidation state distribution, using the known states of its precursors as the ground truth.

Specifically, we use the learned target embedding $\mathbf{t}$ to query a learnable **Oxidation State Codebook** $\mathcal{C}_{\text{ox}} \in \mathbb{R}^{(N_{\text{elem}} \times N_{\text{ox\_states}}) \times d}$. This codebook stores an embedding for each possible element-oxidation state pair. We compute a logit for each pair via a matrix-vector product: $\mathbf{z}_{\text{ox}} = \mathcal{C}_{\text{ox}}^{\top}\mathbf{t}$. The logits are then passed through a sigmoid function to yield the predicted probability distribution, $\hat{\boldsymbol{\pi}} = \sigma(\mathbf{z}_{\text{ox}})$. The model is trained to match this prediction to the ground-truth distribution $\boldsymbol{\pi}^{\star}$, which is derived from the aggregated oxidation states of the precursors. The loss is:

$$\mathcal{L}_{\text{ox}} = -\sum_{E}\sum_{z}\Big[\pi^{\star}_{E,z}\log(\hat{\pi}_{E,z}) + (1-\pi^{\star}_{E,z})\log(1-\hat{\pi}_{E,z})\Big]. \tag{4}$$

This task injects charge awareness into the target embedding, encouraging a latent-space representation that reflects valence chemistry.

4.3 Task 3: Precursor Prediction

Limitations of implicit count prediction. A fundamental challenge in retrosynthesis is determining the number of precursors, m, required for a target material x. Prior methods often address this implicitly, for instance, by applying a fixed threshold (e.g., 0.5) to a sigmoid output layer to select a precursor set. This approach is suboptimal: if the number of precursors selected via thresholding does not match the ground truth m, the model is trained on an incorrect premise, forcing it to merely redistribute probabilities over an erroneously sized set rather than correcting the count itself.

Explicit count prediction as classification. A more principled approach is to first predict the number of precursors m, and subsequently predict their identities. As illustrated in Figure 2 (Task 3), we introduce a dedicated precursor count predictor. This module frames the task as a classification problem, guided by the chemical heuristic that m often correlates with the number of non-metallic elements, E_{non}, in the target (e.g., O, N, F, Cl, Br). Instead of predicting m directly, the classifier learns to predict the deviation of m from E_{non}. Specifically, it outputs a value from the discrete set $\{-2, -1, 0, +1, +2\}$, which represents the predicted difference $m - E_{\text{non}}$. This transforms the prediction of an arbitrary integer into a constrained, chemically-informed classification problem, allowing the model to learn a robust prior over the size of the precursor set.

Precursor prediction Loss. We use the target embedding $\mathbf{t} = f_{\theta}(\mathbf{x}_{\text{feat}})$ to predict the precursor set. We compute a logit vector $\mathbf{z} \in \mathbb{R}^{N_p}$ by taking a matrix-vector product between the target embedding and the **Precursor Codebook**, $\mathbf{z}_{\text{prec}} = \mathcal{C}_{\text{prec}}^{\top}\mathbf{t}$. The selection probabilities for all N_p library precursors are then obtained by an elementwise sigmoid function, $\hat{\mathbf{y}} = \sigma(\mathbf{z}_{\text{prec}})$. Training uses a multi-label binary cross-entropy over the full library:

$$\mathcal{L}_{\text{prec}} = -\sum_{i=1}^{N_p}\Big[\, y_i\log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)\Big], \tag{5}$$

where $y_i = 1$ if the i-th precursor is in the ground-truth set C, and $y_i = 0$ otherwise.

Total objective. The final training loss is a weighted sum:

$$\mathcal{L} = \lambda_{\text{frac}}\mathcal{L}_{\text{frac}} + \lambda_{\text{ox}}\mathcal{L}_{\text{ox}} + \lambda_{\text{prec}}\mathcal{L}_{\text{prec}}. \tag{6}$$

5 Experiments

5.1 Experimental Setup

Datasets. We evaluate on two corpora of inorganic solid-state synthesis datasets. (1) The large-scale dataset from the Ceder group Kononova et al. (2019). We use a train/validation/test split of 44,736 / 2,254 / 2,934 recipes. (2) Following Noh et al. (2024), we adopt a chronological split by publication year: training/validation $\leq$ 2017 and test $\geq$ 2018. In our experiments, we also enforce a closed-vocabulary condition for evaluation by restricting the val/test set to precursors that appear in the training set. The curated subset containing 33,343 recipes (train 24,034 / val 1,842 / test 2,558). We consider Retrieval-Retro dataset the more challenging and realistic benchmark because it is smaller in size and uses a temporal split that mirrors how materials scientists propose new syntheses from prior literature.

Table 2: **Performance comparison on Retrieval-Retro Dataset.** Models are evaluated on two metrics: (a) Top-k accuracy ↑ and (b) Combination Top-k accuracy ↑. Bold values indicate the best performance and underline the second best. Results of ElemwiseRetro* and SynthesisSimilarity* are copied from Retrieval-Retro Noh et al. (2024), and details of procesing Retrieval-Retro Dataset can be found in Section 5.1.

	(a) Top-k accuracy ↑					(b) Combination Top-k accuracy ↑				
Model	Top-1	Top-3	Top-5	Top-10	Top-20	Top-1	Top-3	Top-5	Top-10	Top-20
Matminer (Magpie) (Ward et al., 2018)	33.58	35.54	49.57	68.92	82.06	33.58	46.17	51.02	58.21	64.11
Roost Bartel et al. (2020)	44.53	45.82	56.41	70.37	81.94	44.53	54.81	57.51	63.45	67.32
CrabNet Wang et al. (2021)	57.47	57.58	62.39	78.54	85.89	57.47	60.13	62.59	67.24	71.93
ElemwiseRetro* Kim et al. (2022)	-	-	-	-	-	53.45	57.07	58.19	60.84	-
SynthesisSimilarity* He et al. (2023)	-	-	-	-	-	45.03	48.02	49.11	51.09	-
Retrieval-Retro Noh et al. (2024)	61.02	61.77	66.30	70.72	72.87	61.02	64.82	65.83	67.20	69.23
Ours (Precursor Number Not Given)	66.61	67.86	76.09	85.84	89.14	66.61	74.50	76.62	81.06	83.45
Ours	**68.65**	**69.94**	**78.42**	**88.47**	**91.87**	**68.65**	**76.78**	**78.97**	**83.54**	**86.00**

Baseline Methods. We fist compare against three popular composition-only baselines: **Matminer** Ward et al. (2018) constructs fixed-length composition-based feature vectors from elemental attributes (e.g., atomic number, electronegativity, covalent radius) using the featurization toolkit. **Roost** Bartel et al. (2020) treats a chemical formula as a fully connected graph whose nodes are elements and whose weights reflect stoichiometric fractions. Learned element embeddings with attention-based message passing enable end-to-end inference of composition descriptors without structural inputs. **CrabNet** Wang et al. (2021). CrabNet is a transformer-style architecture that applies compositionally restricted self-attention over element tokens to model inter-element context and predict material properties from composition alone.

Beyond composition-only models, we also compare against methods tailored for precursor recommendation. **ElemwiseRetro** Kim et al. (2022) represents a target composition via a fully connected graph over its constituent elements and infers precursor candidates through element-level interactions. Its element-wise matching scheme encourages near one-to-one correspondences between elements and selected precursors. **SynthesisSimilarity** He et al. (2023) introduces a masked precursor completion to improve supervision for precursor selection. By expressing target materials in the space of precursor tokens, the model naturally supports retrieval-augmented inference from a precursor library. **Retrieval-Retro** Noh et al. (2024) combines a learned retriever informed by reaction energetics with a graph-based encoder for composition, yielding a retrieval-augmented pipeline that improves the top-k accuracy and ranking of plausible precursors.

Metrics. We report two complementary metrics in this work:(a) Top-k accuracy: For each target material, we rank precursor candidates by the model's scores. If all ground-truth precursors appear within the top k candidates, we count a hit. (b) Combination Top-k accuracy For each target, we first take the top 20 precursors by predicted probability. Given the known number of true precursors (n), we consider all size-n combinations from these 20 candidates and rank the combinations by their joint probability. If the ground-truth set appears among the top k combinations, we count a hit.

5.2 Implementation Details

Our model is trained using the AdamW Loshchilov & Hutter (2019) optimizer with a learning rate of 1e-2, with a decay weight of 1e-5. The weights for the multi-task loss function were set to $\lambda_{\text{frac}} = 0.1$, $\lambda_{\text{ox}} = 0.1$, and $\lambda_{\text{prec}} = 1.0$. The model is trained for a maximum of 2000 epochs, employing an early stopping mechanism that halts training if the validation loss fails to improve for 100 consecutive iterations. All experiments were conducted on a single NVIDIA RTX 5090 GPU.

5.3 Quantitative Results

Among three prior systems tailored to inorganic retrosynthesis (ElemwiseRetro, SynthesisSimilarity and Retrieval-Retro), the masked precursor completion paradigm yields weak SynthesisSimilarity. Early retrieval-based approaches improved by modeling inter-reaction relations but did not sufficiently encode intra-reaction structure. Retrieval-Retro strengthened inter-reaction similarities and added a neural reaction-energy module, setting the previous state of the art. Compared to Retrieval-

Table 3: **Performance comparison on Ceder Dataset.** Models are evaluated on two metrics: (a) Top-k accuracy ↑ and (b) Combination Top-k accuracy ↑. Bold values indicate the best performance and underline the second best.

	(a) Top-k accuracy ↑					(b) Combination Top-k accuracy ↑				
Model	Top-1	Top-3	Top-5	Top-10	Top-20	Top-1	Top-3	Top-5	Top-10	Top-20
Matminer (Magpie) (Ward et al., 2018)	30.64	32.58	46.11	65.78	79.75	30.64	42.91	49.18	55.15	61.42
Roost Bartel et al. (2020)	51.87	53.14	62.27	76.11	82.89	51.87	61.11	64.18	67.79	73.35
CrabNet Wang et al. (2021)	61.86	<u>63.02</u>	69.70	79.31	<u>85.38</u>	61.86	68.64	70.86	73.93	76.96
Retrieval-Retro Noh et al. (2024)	56.78	57.16	62.58	68.34	70.11	56.78	61.35	62.54	65.24	67.11
Ours (Precursor Number Not Given)	<u>61.87</u>	62.90	<u>71.46</u>	<u>81.34</u>	84.98	<u>61.87</u>	<u>70.22</u>	<u>72.17</u>	<u>75.86</u>	<u>78.96</u>
Ours	**65.03**	**66.12**	**75.12**	**85.51**	**89.33**	**65.03**	**73.82**	**75.87**	**79.75**	**83.00**

Table 4: **Ablation on chemical constraints.** We toggle Rebuild evidence, oxidation-state constraint, and element-wise filtering.

Setting	**Rebuild**	**Oxidation-state**	**Filter**	**Top-K accuracy ↑**					**Combination Top-K accuracy ↑**				
				Top-1	Top-3	Top-5	Top-10	Top-20	Top-1	Top-3	Top-5	Top-10	Top-20
Base	✗	✗	✗	61.7	63.1	73.7	84.2	90.5	61.7	72.3	75.2	79.8	83.3
+ Rebuild	✓	✗	✗	61.5	62.8	73.7	84.4	90.7	61.5	72.2	75.3	79.7	83.3
+ Oxidation	✗	✓	✗	61.8	62.9	73.5	84.4	90.3	61.8	72.1	75.3	80.2	82.8
+ Oxidation & Rebuild	✓	✓	✗	64.0	65.1	75.7	85.8	91.0	64.0	73.9	76.7	81.2	84.2
Ours (Full)	✓	✓	✓	**68.6**	**69.9**	**78.4**	**88.5**	**91.9**	**68.6**	**76.8**	**79.0**	**83.5**	**86.0**

Retro, our model lifts Top-1 accuracy by +5.59, which means our model jointly captures inter-reaction relations via a precursor codebook that conditions on the product/target, and intra-reaction constraints via atom and charge conservation. This combination delivers substantial, consistent gains in both Top-k and Combination Top-k accuracy over all baselines on both datasets. Despite the harder setting, our model improves over Retrieval-Retro by +8.3 / +9.0 / +12.5 / +17.2 / +19.2 percentage on Top-k.

Effective inorganic retrosynthesis planning benefits from modeling both relations across reactions and physical conservation within reactions. In contrast, formula-level enhancements alone are insufficient, and naively scaling data without addressing stoichiometry redundancy can hurt performance.

5.4 Ablation Study

We analyze the impact of three key components of our model: (i) the element-wise filter (Filter), (ii) the oxidation-state prediction task (Oxidation-state), and (iii) the compositional fraction reconstruction task, which provides rebuild evidence (Rebuild). The results of this ablation study are presented in Table 4.

Our base model, with all three components disabled, establishes a solid performance baseline. However, the results show that adding either the **+ Rebuild** evidence or the **+ Oxidation** constraint in isolation yields only marginal improvements. This indicates that neither task alone is sufficient to substantially enhance the model's predictive power. A significant performance gain is observed when the **+ Oxidation & Rebuild** tasks are combined. This synergistic effect, which boosts Top-1 accuracy from 61.7 to 64.0, underscores the importance of learning representations that are concurrently aware of both chemical valence and compositional integrity.

Finally, the **Ours (Full)** model, which applies the element-wise **Filter** on top of the combined tasks, achieves the best performance across all metrics by a significant margin (e.g., a leap to 68.6 in Top-1 accuracy). This demonstrates that while the learned representations are powerful, the constraint-based filter is crucial for pruning chemically implausible candidates and refining the final predictions to a state-of-the-art level.

5.5 Case Study

Case Study I: Element-wise Filter The element-wise filter is a critical component for ensuring chemical plausibility by enforcing elemental conservation. As illustrated in Figure 3(a), the filter prunes the list of initial precursor candidates by removing any that contain elements not present in the target compound. For instance, in predicting precursors for the com-

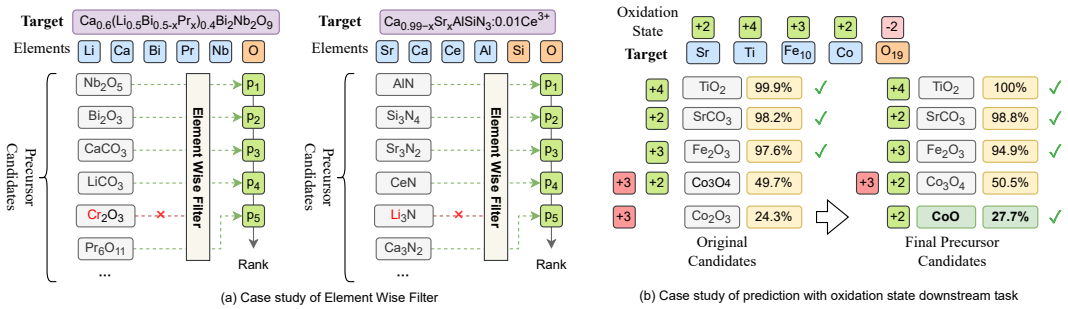

Figure 3: (a) Illustration of the Element-wise Filter. For two distinct targets, an oxide and a nitride (both with doping elements), the filter correctly rejects wrong precursor candidates that introduce extraneous elements not present in the target, such as Cr_2O_3 for the oxide target and Li_3N for the nitride target. (b) Case study of the oxidation state prediction task. For a target requiring Co^{2+}, a baseline model (left) might propose common precursors like Co_3O_4 or Co_2O_3 where cobalt is in a +3 state. Our model, guided by the oxidation state auxiliary task, correctly identifies CoO as the appropriate precursor that provides the required Co^{2+} valence state (right).

plex oxide target $Ca_{0.6}(Li_{0.5}Bi_{0.5-x}Pr_x)_{0.4}Bi_2Nb_2O_9$, the filter correctly identifies and discards Cr_2O_3 because Chromium (Cr) is an extraneous element. Similarly, for the nitride target $Ca_{0.99-x}Sr_xAlSiN_3$:0.01Ce^{3+}, a candidate like Li_3N is rejected because Lithium (Li) is not a constituent of the target. This simple yet effective screening step significantly reduces the search space, allowing the downstream ranking model to focus only on stoichiometrically valid candidates.

Case Study II: Oxidation State Embedding The oxidation state prediction task is critical for selecting chemically plausible precursors, especially for elements that can exist in multiple valence states. We examine the synthesis of a complex oxide target containing Strontium (Sr^{2+}), Titanium (Ti^{4+}), Iron (Fe^{3+}), and Cobalt (Co^{2+}). While precursors for Sr, Ti, and Fe are relatively straightforward ($SrCO_3$, TiO_2, Fe_2O_3), selecting the correct cobalt source is challenging.

As shown in Figure 3, a model without explicit oxidation modeling may propose common cobalt oxides like Co_3O_4 or Co_2O_3 based on their frequency in the training data. However, these precursors primarily contain Co^{3+}, which is inconsistent with the target's requirement for Co^{2+}. Our model, equipped with the oxidation state auxiliary task, learns to correlate the required valence state in the target with the valence state in the precursors. Consequently, it correctly upweights CoO, which provides the necessary Co^{2+}, and downweights the incorrect alternatives. This case study demonstrates that explicit oxidation-aware learning is crucial for refining precursor selection beyond simple co-occurrence statistics and towards chemically coherent predictions.

6 Conclusion

We addressed the gap between compositional design and actionable synthesis by reformulating inorganic retrosynthesis around two chemical reaction rules: elemental conservation and electron balance. Our principle–centered framework, PRICIN, injects these chemical laws into both learning and inference via (i) explicit oxidation–state supervision that embeds redox semantics into the target representation, and (ii) a lightweight element–wise filter that predicts precursor count and prunes candidates violating conservation. Comprehensive experiments on the Ceder corpus and the more realistic, temporally split Retrieval–Retro benchmark demonstrate consistent state-of-the-art performance, including substantial gains in Top-k and combination Top-k accuracy. Ablations show that oxidation–state supervision and the composition rebuild objective are complementary: together they markedly reduce early-rank errors, while the inference-time element filter delivers additional, low-cost improvements by enforcing feasibility. These capabilities hold the potential to advance closed-loop scientific discovery and self-driven laboratories in inorganic materials design.

REFERENCES

Christopher J Bartel, Amalie Trewartha, Qi Wang, Alexander Dunn, Anubhav Jain, and Gerbrand Ceder. A critical examination of compound stability predictions from machine-learned formation energies. *npj computational materials*, 6(1):97, 2020.

Jiadong Chen, Samuel R. Cross, Lincoln J. Miara, Jeong-Ju Cho, Yan Wang, and Wenhao Sun. Navigating phase diagram complexity to guide robotic inorganic materials synthesis. *Nature Synthesis*, 3(5):606–614, May 2024. ISSN 2731-0582. doi: 10.1038/s44160-024-00502-y. URL https://www.nature.com/articles/s44160-024-00502-y. Publisher: Nature Publishing Group.

Hitarth Choubisa, Petar Todorović, Joao M. Pina, Darshan H. Parmar, Ziliang Li, Oleksandr Voznyy, Isaac Tamblyn, and Edward H. Sargent. Interpretable discovery of semiconductors with machine learning. *npj Computational Materials*, 9(1):117, June 2023. ISSN 2057-3960. doi: 10.1038/s41524-023-01066-9. URL https://www.nature.com/articles/s41524-023-01066-9. Publisher: Nature Publishing Group.

Tanishq Gupta, Mohd Zaki, N. M. Anoop Krishnan, and Mausam. MatSciBERT: A materials domain language model for text mining and information extraction. *npj Computational Materials*, 8(1):102, May 2022. ISSN 2057-3960. doi: 10.1038/s41524-022-00784-w. URL https://doi.org/10.1038/s41524-022-00784-w.

Tanjin He, Wenhao Sun, Haoyan Huo, Olga Kononova, Ziqin Rong, Vahe Tshitoyan, Tiago Botari, and Gerbrand Ceder. Similarity of precursors in solid-state synthesis as text-mined from scientific literature. *Chemistry of Materials*, 32(18):7861–7873, 2020.

Tanjin He, Haoyan Huo, Christopher J Bartel, Zheren Wang, Kevin Cruse, and Gerbrand Ceder. Precursor recommendation for inorganic synthesis by machine learning materials similarity from scientific literature. *Science advances*, 9(23):eadg8180, 2023.

Haoyan Huo, Ziqin Rong, Olga Kononova, Wenhao Sun, Tiago Botari, Tanjin He, Vahe Tshitoyan, and Gerbrand Ceder. Semi-supervised machine-learning classification of materials synthesis procedures. *npj Computational Materials*, 5(1):62, 2019.

Edward Kim, Zach Jensen, Alexander van Grootel, Kevin Huang, Matthew Staib, Sheshera Mysore, Haw-Shiuan Chang, Emma Strubell, Andrew McCallum, Stefanie Jegelka, et al. Inorganic materials synthesis planning with literature-trained neural networks. *Journal of chemical information and modeling*, 60(3):1194–1201, 2020.

Seongmin Kim, Juhwan Noh, Geun Ho Gu, Shuan Chen, and Yousung Jung. Element-wise formulation of inorganic retrosynthesis. In *AI for Accelerated Materials Design NeurIPS 2022 Workshop*, 2022.

Olga Kononova, Haoyan Huo, Tanjin He, Ziqin Rong, Tiago Botari, Wenhao Sun, Vahe Tshitoyan, and Gerbrand Ceder. Text-mined dataset of inorganic materials synthesis recipes. *Scientific Data*, 6(1):203, 2019.

Zixun Lan, Zuo Zeng, Binjie Hong, Zhenfu Liu, and Fei Ma. RCsearcher: Reaction center identification in retrosynthesis via deep Q-learning. *Pattern Recognition*, 150:110318, June 2024. ISSN 0031-3203. doi: 10.1016/j.patcog.2024.110318. URL https://www.sciencedirect.com/science/article/pii/S0031320324000694.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization, 2019. URL https://arxiv.org/abs/1711.05101.

Matthew J McDermott, Shyam S Dwaraknath, and Kristin A Persson. A graph-based network for predicting chemical reaction pathways in solid-state materials synthesis. *Nature Communications*, 12(1):3097, 2021.

Samantha M. McDonald, Emily K. Augustine, Quinn Lanners, Cynthia Rudin, L. Catherine Brinson, and Matthew L. Becker. Applied machine learning as a driver for polymeric biomaterials design. *Nature Communications*, 14(1):4838, August 2023. ISSN 2041-1723. doi: 10.1038/s41467-023-40459-8. URL https://www.nature.com/articles/s41467-023-40459-8. Publisher: Nature Publishing Group.

Amil Merchant, Simon Batzner, Samuel S. Schoenholz, Muratahan Aykol, Gowoon Cheon, and Ekin Dogus Cubuk. Scaling deep learning for materials discovery. *Nature*, 624(7990):80–85, December 2023. ISSN 1476-4687. doi: 10.1038/s41586-023-06735-9. URL https://www.nature.com/articles/s41586-023-06735-9. Publisher: Nature Publishing Group.

Akira Miura, Christopher J. Bartel, Yosuke Goto, Yoshikazu Mizuguchi, Chikako Moriyoshi, Yoshihiro Kuroiwa, Yongming Wang, Toshie Yaguchi, Manabu Shirai, Masanori Nagao, Nataly Carolina Rosero-Navarro, Kiyoharu Tadanaga, Gerbrand Ceder, and Wenhao Sun. Observing and modeling the sequential pairwise reactions that drive solid-state ceramic synthesis. *Advanced Materials*, 33(24):2100312, 2021. ISSN 1521-4095. doi: 10.1002/adma.202100312. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.202100312.

Heewoong Noh, Namkyeong Lee, Gyoung S. Na, and Chanyoung Park. Retrieval-retro: Retrieval-based inorganic retrosynthesis with expert knowledge. *arXiv preprint arXiv:2410.21341*, 2024.

Tobias Prein, Anke Hamm, Arghya Basu, Tobias Jägle, Aleksandar Lekic, Tutku Karadayilar, Xiaolong Hu, Dev Ashish Jha, Rishin Sharma, Loan Talamini, et al. Retro-rank-in: A ranking-based approach for inorganic materials synthesis planning. *arXiv preprint arXiv:2502.04289*, 2025.

Laura E. Ratcliff, Stephan Mohr, Georg Huhs, Thierry Deutsch, Michel Masella, and Luigi Genovese. Challenges in large scale quantum mechanical calculations. *WIREs Computational Molecular Science*, 7(1):e1290, 2017. ISSN 1759-0884. doi: 10.1002/wcms.1290. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/wcms.1290. _eprint: https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wcms.1290.

Vignesh Ram Somnath, Charlotte Bunne, Connor W. Coley, Andreas Krause, and Regina Barzilay. Learning graph models for retrosynthesis prediction. November 2021. URL https://openreview.net/forum?id=SnONpXZ_uQ_.

Zhilong Song, Shuaihua Lu, Minggang Ju, Qionghua Zhou, and Jinlan Wang. Accurate prediction of synthesizability and precursors of 3d crystal structures via large language models. *Nature Communications*, 16(1):6530, 2025.

Nathan J. Szymanski, Bernardus Rendy, Yuxing Fei, Rishi E. Kumar, Tanjin He, David Milsted, Matthew J. McDermott, Max Gallant, Ekin Dogus Cubuk, Amil Merchant, Haegyeom Kim, Anubhav Jain, Christopher J. Bartel, Kristin Persson, Yan Zeng, and Gerbrand Ceder. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature*, 624(7990):86–91, December 2023a. ISSN 1476-4687. doi: 10.1038/s41586-023-06734-w. URL https://www.nature.com/articles/s41586-023-06734-w.

Nathan J Szymanski, Bernardus Rendy, Yuxing Fei, Rishi E Kumar, Tanjin He, David Milsted, Matthew J McDermott, Max Gallant, Ekin Dogus Cubuk, Amil Merchant, et al. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature*, 624(7990):86–91, 2023b.

Umit V. Ucak, Islambek Ashyrmamatov, Junsu Ko, and Juyong Lee. Retrosynthetic reaction pathway prediction through neural machine translation of atomic environments. *Nature Communications*, 13(1):1186, March 2022. ISSN 2041-1723. doi: 10.1038/s41467-022-28857-w. URL https://www.nature.com/articles/s41467-022-28857-w. Publisher: Nature Publishing Group.

Nicholas Walker, Amalie Trewartha, Haoyan Huo, Sanghoon Lee, Kevin Cruse, John Dagdelen, Alexander Dunn, Kristin Persson, Gerbrand Ceder, and Anubhav Jain. The impact of domain-specific pre-training on named entity recognition tasks in materials science. *Available at SSRN 3950755*, 2021.

Anthony Yu-Tung Wang, Steven K Kauwe, Ryan J Murdock, and Taylor D Sparks. Compositionally restricted attention-based network for materials property predictions. npj computational materials 7 (1): 77, 2021.

Zheren Wang, Kevin Cruse, Yuxing Fei, Ann Chia, Yan Zeng, Haoyan Huo, Tanjin He, Bowen Deng, Olga Kononova, and Gerbrand Ceder. Ulsa: Unified language of synthesis actions for the representation of inorganic synthesis protocols. *Digital Discovery*, 1(2):313–324, 2022a.

Zheren Wang, Olga Kononova, Kevin Cruse, Tanjin He, Haoyan Huo, Yuxing Fei, Yan Zeng, Yingzhi Sun, Zijian Cai, Wenhao Sun, and Gerbrand Ceder. Dataset of solution-based inorganic materials synthesis procedures extracted from the scientific literature. *Scientific Data*, 9(1):231, 2022b.

Logan Ward, Alexander Dunn, Alireza Faghaninia, Nils ER Zimmermann, Saurabh Bajaj, Qi Wang, Joseph Montoya, Jiming Chen, Kyle Bystrom, Maxwell Dylla, et al. Matminer: An open source toolkit for materials data mining. *Computational Materials Science*, 152:60–69, 2018.

Zhenpeng Yao, Yanwei Lum, Andrew Johnston, Luis Martin Mejia-Mendoza, Xin Zhou, Yonggang Wen, Alán Aspuru-Guzik, Edward H. Sargent, and Zhi Wei Seh. Machine learning for a sustainable energy future. *Nature Reviews Materials*, 8(3):202–215, March 2023. ISSN 2058-8437. doi: 10.1038/s41578-022-00490-5. URL https://www.nature.com/articles/s41578-022-00490-5.

Claudio Zeni, Robert Pinsler, Daniel Zügner, Andrew Fowler, Matthew Horton, Xiang Fu, Sasha Shysheya, Jonathan Crabbé, Lixin Sun, Jake Smith, et al. Mattergen: a generative model for inorganic materials design. *arXiv preprint arXiv:2312.03687*, 2023.

Xinhao Zhang, Lung Wa Chung, and Yun-Dong Wu. New mechanistic insights on the selectivity of transition-metal-catalyzed organic reactions: The role of computational chemistry. *Accounts of Chemical Research*, 49(6):1302–1310, June 2016. ISSN 0001-4842. doi: 10.1021/acs.accounts.6b00093. URL https://doi.org/10.1021/acs.accounts.6b00093. Publisher: American Chemical Society.

A Appendix

Ethics Statement

This research is dedicated to advancing materials science through computational methods and does not involve any ethical concerns related to human subjects, animal welfare, or data privacy. The datasets used are from publicly available sources, and our work aims to accelerate scientific discovery in a responsible manner.

Reproducibility Statement

The code and data required to reproduce our results will be made publicly available upon publication. Detailed instructions for setting up the environment and running the experiments will be provided in the supplementary materials and a public repository. We have taken care to document our methodology and experimental setup to ensure that our work is transparent and reproducible.

The Use of Large Language Models (LLMs)

During the preparation of this manuscript, we utilized LLMs as a general-purpose writing assistant. The primary role of LLMs was to assist with polishing the text, including improving grammar, clarity, and readability. All authors have reviewed, edited, and take full responsibility for the final content of this paper.

A.1 Precursor-count Prediction Model Architecture

We formulate precursor-count prediction as a multi-class classification problem and instantiate a compact encoder–attention–classifier architecture. Given $\mathbf{x} \in \mathbb{R}^{118}$, an MLP encoder with ReLU activations and Dropout produces a 128-D representation. This representation is refined by a self-attention block (multihead attention with 4 heads and output dimension 128) followed by LayerNorm ($\varepsilon = 10^{-5}$), capturing non-local dependencies among feature dimensions. We concatenate the attention-refined embedding with an auxiliary scalar which is fed to a classifier MLP. The model is trained with a cross-entropy objective, and the predicted count is obtained by taking argmax over the logits. This design emphasizes parameter efficiency and stable optimization, while the self-attention module consistently improves early-rank accuracy relative to a pure MLP baseline.

A.2 Performance on Retrieve-Retro Dataset

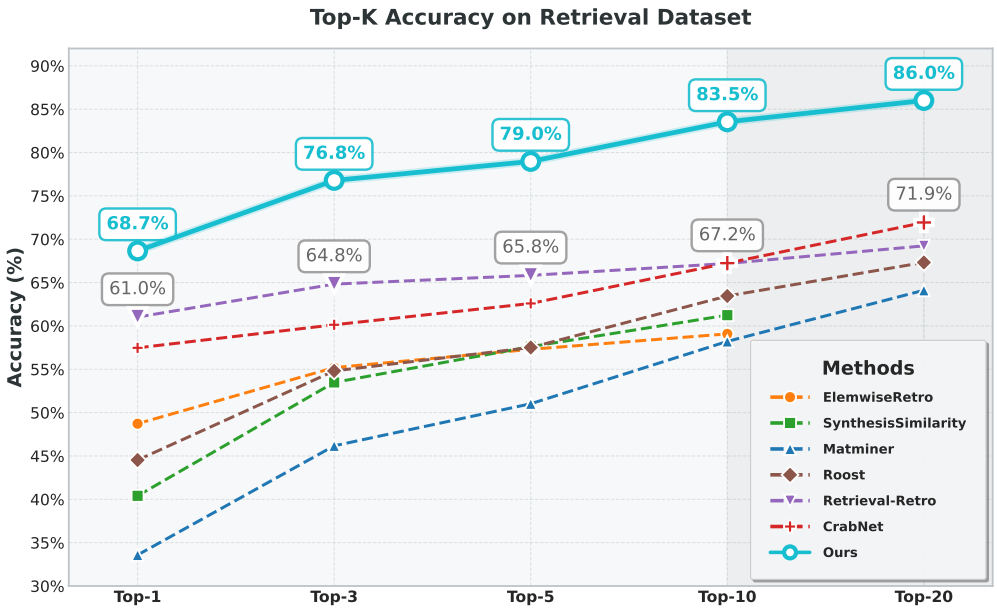

Figure 4: **Comparison of Top-K accuracy on the Retrieval-Retro Dataset (Combination setting).** Our method outperforms all baselines in Top-K accuracy, with only our approach and a few others achieving Top-20 results. This highlights the effectiveness of our chemically-informed pretraining and constrained retrieval pipeline.

A.3 Limitations and Future Work

Our study focuses on inorganic retrosynthesis planning under two datasets and does not model operating conditions or kinetics explicitly. Extending PRICIN to (i) multi-step planning with byproducts, (ii) joint prediction of temperature, atmosphere, and time, and (iii) calibrated uncertainty for active learning in autonomous labs are promising directions.

In summary, enforcing chemical constraints provides a robust inductive bias for inorganic retrosynthesis, advancing the DMAT loop toward reliable, closed-loop materials discovery.