General-purpose LLMs as Constrained Crystal Composition Generators

Hedda Oschinski†, Maximilian L. Ach†, Konstantin S. Jakob, Christian Carbogno†, Karsten Reuter

Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-14195 Berlin, Germany
†author's email: oschinski@fhi.mpg.de, ach@fhi.mpg.de, carbogno@fhi.mpg.de

Targeted discovery of inorganic materials is challenging due to the vastness of design spaces and the high (computational) cost of exhaustive screening. This motivates the development of more flexible approaches for navigating these spaces efficiently. General-purpose large language models (LLMs) offer one such alternative to task-specific generative models, generating candidates without specific training. We investigate how LLMs can be used to systematically explore large chemical spaces, rather than merely suggesting individual promising candidates. Using a well-characterized Elpasolite composition space of 2 million possible ABC_2D_6 compositions as a benchmark, we develop an iterative workflow (cf. Figure 1) in which an LLM proposes compositions, receives formation-energy feedback, and refines its search strategy. Targeting compositions with formation energies below -2.26 eV/atom, our approach identifies more low-energy Elpasolites within 5,000 proposals (96% coverage) than prior generative models (GAN, VAE, and reinforcement-learning) trained specifically for this task. The performance is based on iterative in-context learning, where the LLM adapts its sampling based on prior suggestions. Model parameters, such as size, reasoning effort, and feedback frequency influence performance; however, we expect efficacy to improve further as general-purpose LLMs continue to advance. Chemically meaningful compositions are generated even when material class and structural information are withheld, indicating the model infers relevant chemical relationships from generation history. Our results establish general-purpose LLMs as flexible, accessible tools for constrained materials composition generation, highlighting their potential as components of scalable, agent-driven inverse materials design workflows. [1]

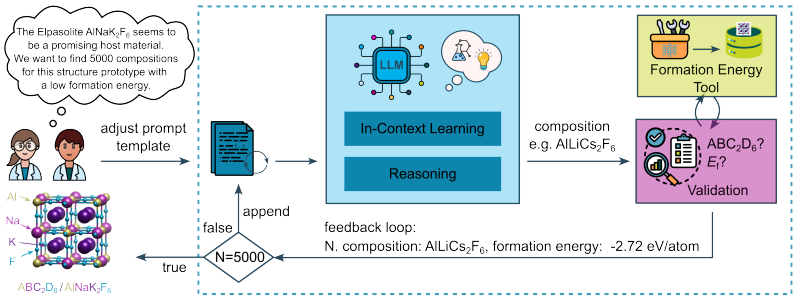

Figure 1: LLM-based workflow for generating low-formation-energy Elpasolite compositions. The LLM is prompted with generation instructions and draws on chemical knowledge encoded in its learned parameters to generate candidate Elpasolite compositions. Compositions are validated for correct format, and a formation energy tool (in this case a database of pre-computed formation energies) determines the stability of the material. Finally, feedback is provided to the LLM depending on the formation energy, and the next composition is generated until the desired amount is met.

References

[1] H. Oschinski, M. L. Ach, K. S. Jakob, C. Carbogno, K. Reuter, *npj Comp. Mat.* **12(1)**, 337 (2026).