Generative AI Meets Phonon Validation: A Multi-Stage Workflow for Reliable Discovery of Hydrogen-Storage Hydrides

M. Kaufman[1]
[1]Regional Technological Institute, Faculty of Mechanical Engineering, University of West Bohemia in Pilsen, Univerzitni 8, 301 00 Pilsen, Czech Republic

mkaufman@fst.zcu.cz

Lightweight complex hydrides are among the most promising solid-state hydrogen-storage materials for clean energy technologies, yet discovering new stable phases remains a major challenge. Generative models for inorganic materials can now propose tens of thousands of hypothetical crystal structures at scale, but thermodynamically favorable candidates frequently lack dynamic lattice stability — a failure mode that enthalpy-based screening alone cannot detect. Here we present a multi-stage generative-AI–MLIP–DFT workflow that addresses this reliability gap by coupling diffusion-based crystal generation (MatterGen [1]) with rapid relaxation and ranking using the universal machine-learning interatomic potential CHGNet [2], followed by ab initio density-functional perturbation theory (DFPT) phonon validation as an acceptance criterion for physical realizability. Starting from over 50,000 MatterGen-generated structures, CHGNet screening for negative formation enthalpy relative to elemental reference states (a necessary first-pass filter, though not equivalent to full convex-hull stability) and high gravimetric hydrogen content yielded 41 candidate hydrides. Benchmarking CHGNet formation enthalpies against full DFT (PBE, Quantum ESPRESSO) for all 41 candidates gives a Spearman rank correlation ρ_s = 0.695, confirming that the MLIP preserves relative stability ordering with sufficient fidelity for coarse-grained screening while motivating subsequent ab initio refinement. Five chemically diverse top-ranked candidates — $LiBH_3$, $Li_3Mg_2H_7$, $TiVH_7$, $LiCa_2(AlH_4)_3$, and Li_2BeH_4 — were then subjected to DFPT phonon calculations (4×4×4 q-mesh). Despite all five exhibiting favorable formation enthalpies, four display imaginary phonon modes — indicating that enthalpy-only selection of AI-generated structures can produce a high proportion of dynamically unstable false positives. $TiVH_7$ shows only a shallow soft mode (−11.75 cm^{-1}), consistent with possible anharmonic stabilization at finite temperatures, whereas Li_2BeH_4 is strongly unstable (−560.30 cm^{-1}). The sole fully validated candidate, $LiBH_3$ (ΔH_f = −0.58 eV/atom, ~14.5 wt% H), adopts a monoclinic C2/m structure (Z = 2, density ~1.0 g/cm^3) that, to the best of our knowledge, has not been reported previously and is distinct from the cubic perovskite Pm-3m model considered in earlier theoretical studies [3]. Our results demonstrate that integrating generative design with MLIP pre-screening and mandatory DFPT phonon validation is essential for trustworthy AI-driven discovery of dynamically stable candidate hydrides for renewable-energy hydrogen storage.

References

[1] C. Zeni et al., Nature, 639 (2025) 624–632.
[2] B. Deng et al., Nat. Mach. Intell., 5 (2023) 1031–1041.
[3] X.-J. San et al., Chin. Phys. B, 17 (2008) 2222–2228.

Figures

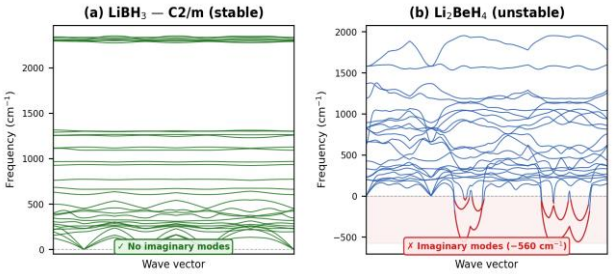

Figure 1. Phonon dispersion curves of (a) the dynamically stable $LiBH_3$ (C2/m) and (b) a thermodynamically favorable but dynamically unstable AI-generated candidate, Li_2BeH_4. Red curves and shading indicate imaginary phonon frequencies, underscoring the necessity of DFPT validation