Discovery report for RTC

Research Objective

Using this dataset and what you find in literature, develop a few ideas about how to design a high-temperature/room-temperature superconductor. Then give explicit instructions for synthesis/processing and justification for why it is a room-temperature superconductor.

Dataset Description

The file $3DSC_{MP}$.csv contains the 3DSCMP in tabular form. The three most important columns are the following:
$formula_{sc}$: The chemical formula of the material, which is exactly the original chemical formula of the SuperCon entry. Note that due to the normalization of chemical formulas in the matching algorithm, this chemical formula might deviate by a constant factor from the chemical formula of the primitive unit cell of the structure.
tc: The critical temperature in Kelvin. Non-superconductors have a Tc of 0.
cif: The path to the cif file of the 3DSCMP crystal structure relative to the directory $3DSC/superconductors_{3D}/$. If the name contains $synth_{doped}$ this means that this structure was artificially doped.
Additionally to these three basic columns of the 3DSCMP database, there are a lot of columns which were important in the matching and adaptation algorithm, which are from the original Materials Project database or which were important for the analysis in our paper. These columns are less important if you just want to use the 3DSCMP, but they might be interesting for everyone looking to improve the 3DSCMP or reproduce the results in our paper.
MAGPIE.*: MAGPIE feature vectors of the chemical formula of this material. Missing in the github version (see note below).
SOAP.*: DSOAP feature vectors of the structure. Missing in the github version (see note below).
.*_2: All columns ending with _2 are the columns from the original structure from the Materials Project or columns added in the process of cleaning the initial Materials Project database.
totreldiff: The Δ totrel from our paper[1], a measure of the difference between the original chemical formula of the SuperCon and of the Materials Project.
$formula_{frac}$: The normalization factor of the chemical formulas.
sc_{class}: The superconductor group (either 'Other', '$Heavy_{fermion}$', 'Chevrel', 'Oxide', 'Cuprate', 'Ferrite', 'Carbon'). Some of the entries also have '$OxideHeavy_{fermion}$' or '$Heavy_{fermionChevrel}$', which means that the algorithm could not uniquely attribute this material into one group.
weight: The sample weight which was used for the XGB model and the calculcation of the scores. This is just the inverse of the number of crystal structures per SuperCon entry in the database.
$cif_{before\_synthetic_{doping}}$: The path to the original cif file of the Materials Project before artificial doping.
$crystal_{temp}\_2$: The crystal temperature. Non-zero only for the 3DSCICSD.
$no_{crystal\_temp_{given}}\_2$: If the crystal temperature was not explicitly given. Always True in the 3DSCMP. In the 3DSCICSD, this is True if no crystal temperature was given and 293K was assumed.
cubic, hexagonal, monoclinic, orthorhombic, tetragonal, triclinic, trigonal, primitive, base-centered, body-centered, face-centered: The symmetry features as described in the supporting information of our paper.

Summary of Discoveries

Discovery 1: Charge-Reservoir and Oxygen-Stoichiometry Engineering in Hg-Based Multilayer Cuprates

Data-driven mining of cuprates identifies a simple structural–chemical recipe for maximizing Tc at ambient pressure: strongly layered, long-c phases with Ba/Ca spacer blocks and heavy, closed-shell charge-reservoir layers (Hg, Tl), with oxygen stoichiometry tuned near O_8. Within this Hg-based family, explicitly modeling oxygen content reveals that lower oxygen fractions correlate with higher Tc and that O_8 outperforms O_{10} in Hg–Pb systems. These insights motivate concrete synthesis protocols for $Ba_2Ca_2Cu_3(Hg,Tl,Pb)O_{8+\delta}$ phases using fast-ramp, sealed-ampoule or high-pressure processing to stabilize Hg-rich layers while setting δ for optimal hole concentration.

Discovery 2: Rare-Earth Iron Pnictides: Stability- and Anisotropy-Guided Design Near 50 K

Data-driven screening of iron pnictides reveals a coherent set of design levers—thermodynamic stability, rare-earth chemistry, oxygen/fluorine electron doping, and layered tetragonal anisotropy—that reproducibly elevate Tc into the ≈50 K regime. Multi-objective models nominate Nd- and Sm-based 1111 phases as near-optimal, and the synthesis space is mapped to ambient-pressure solid-state routes with fluorine-retention strategies and optional SPS densification, with high-pressure alternatives when needed.

Discovery 3: MgB

Data-driven dissection of the "Other" superconductor class shows that its high-Tc tail is almost entirely MgB2-like: light-element, hexagonal P6/mmm, low-density borides. Within this family, every statistically analyzable single dopant suppresses Tc, so simple, series-specific linear models based solely on dopant fraction predict Tc reliably, whereas general-purpose ML fails. Mature, scalable solid-state and combustion routes exist for pure and doped MgB2, enabling precise, a priori Tc targeting by composition rather than Tc enhancement.

Discovery 4: Hydrogen-Rich Superhydrides Under Megabar Pressures: Data-Consistent Pathway to Room-Temperature Superconductivity

Analysis of hydride entries in the superconductivity dataset shows that hydrogen atomic fraction is the dominant positive predictor of Tc, whereas general-purpose models trained on non-hydrides fail dramatically when extrapolated to hydrogen-rich compounds. These results, together with small-sample proposals that top out below room temperature, argue for pushing to extreme hydrogen contents and clathrate-like frameworks stabilized at megabar pressures, with a concrete diamond-anvil-cell (DAC) workflow enabling targeted synthesis and in situ validation.

Charge-Reservoir and Oxygen-Stoichiometry Engineering in Hg-Based Multilayer Cuprates

Summary

Data-driven mining of cuprates identifies a simple structural–chemical recipe for maximizing Tc at ambient pressure: strongly layered, long-c phases with Ba/Ca spacer blocks and heavy, closed-shell charge-reservoir layers (Hg, Tl), with oxygen stoichiometry tuned near O_8. Within this Hg-based family, explicitly modeling oxygen content reveals that lower oxygen fractions correlate with higher Tc and that O_8 outperforms O_{10} in Hg–Pb systems. These insights motivate concrete synthesis protocols for $Ba_2Ca_2Cu_3(Hg,Tl,Pb)O_{8+\delta}$ phases using fast-ramp, sealed-ampoule or high-pressure processing to stabilize Hg-rich layers while setting δ for optimal hole concentration.

Background

Cuprate superconductors derive their high critical temperatures from correlated electrons in CuO_2 planes that are chemically doped by adjacent "charge-reservoir" blocks. Structural anisotropy, apical oxygen geometry, and the ability of the reservoir to deliver carriers without buckling the planes are widely believed to control pairing strength and phase coherence. Among known cuprates, multilayer Hg- and Tl-based families set the ambient-pressure Tc benchmark, yet practical optimization remains sensitive to subtle stoichiometry, particularly oxygen content, and to the thermodynamics of volatile heavy-metal oxides during synthesis.

Results & Discussion

A comprehensive analysis of 874 cuprates shows that the highest-Tc materials congregate in a narrow structural–chemical corridor characterized by strong axial anisotropy and heavy-metal charge reservoirs. Defining the "top 10%" cohort by Tc 88.17 K, these materials exhibit longer c-axes (12.90 ± 2.69 Å vs 10.86 ± 4.68 Å), slightly shorter in-plane a and b, and a larger c/a anisotropy (3.069 ± 0.893 vs 2.532 ± 1.132; all $p < 0.001$), with medium effect sizes for c and c/a (Cohen's d ≈ 0.45–0.50) [r5]. Chemically, Ba and Ca are strongly enriched (73% and 57%, respectively), and heavy cations—most notably Hg—are dramatically over-represented in the high-Tc group (Hg present in 35% of the top cohort, 10.6× enriched; odds ratio = 15.9, p < 10^−18), followed by Tl and Bi [r5]. The three highest-Tc entries are all Ba–Ca–Cu–Hg–O multilayers with modest cation substitution—$Ba_2Ca_2Cu_3Hg_{0.85}Re_{0.15}O_8$ (132.9 K), $Ba_2Ca_2Cu_3Hg_{0.8}Pb_{0.2}O_8$ (132.0 K), and $Ba_2Ca_1Cu_2HgO_{6.21}$ (127.0 K)—underscoring the linkage between long-c stacking, Ba/Ca spacers, and Hg-based reservoirs [r5]. Mechanistically, Hg^{2+} and Tl^{3+} in the reservoir form weakly coupled Hg–O or Tl–O layers that transfer holes to the CuO_2 planes while preserving an almost ideal, flat CuO_2 geometry and maximizing the Cu–apical-O separation (≈2.77–2.8 Å), a structural parameter repeatedly tied to elevated Tc [r10, plakida2010, choy2001]. Their closed-shell character and size stabilize multilayer architectures (n > 2), although Tl–O can introduce local disorder that should be managed by careful chemistry [r10, plakida2010].

Incorporating oxygen content explicitly into the materials model reveals oxygen stoichiometry as a decisive control knob in Hg-based systems. When an oxygen atomic-fraction feature ($O_{content}$) is added to a gradient-boosted regression trained on cuprates, it emerges as the most important predictor (importance = 0.454), and a statistically significant negative correlation appears between $O_{content}$ and Tc for Hg-containing cuprates (Pearson r = −0.372, p = 0.0044; $O_{content}$ range 0.500–0.546) [r45]. Critically, the model resolves the performance gap between O_{10} and O_8 variants of $Ba_2Ca_2Cu_3Hg_{0.8}Pb_{0.2}Oy$, predicting 24.5 K for O_{10} but 97.7 K for O_8, whereas a baseline model lacking oxygen stoichiometry is blind to this distinction (both 106.3 K) [r45]. Database validation aligns with the trend: the O_8 compound $Ba_2Ca_2Cu_3Hg_{0.8}Pb_{0.2}O_8$ is reported at 132.0 K, and the three top-Tc Hg cuprates all sit at $O_{content} = 0.500$—i.e., near the O_8 stoichiometry limit [r45]. While the enhanced model shows signs of overfitting and the O_{10} point extrapo-

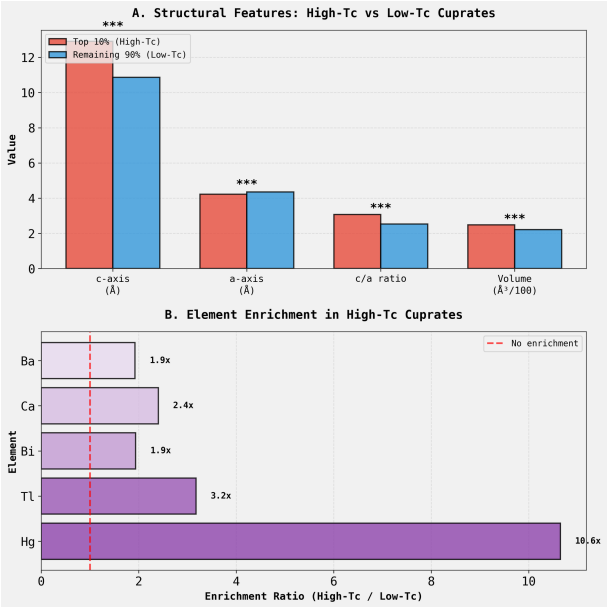

Figure 1: High-Tc cuprates are distinguished by specific structural and chemical signatures. (A) Comparison of average lattice parameters shows that high-Tc materials (top 10% by Tc) have a significantly larger c-axis and c/a ratio than their low-Tc counterparts. (B) The high-Tc cohort is also characterized by elemental enrichment, particularly for the heavy charge-reservoir elements Tl and, most dramatically, Hg (10.6x). These findings identify a design motif for high-temperature superconductivity that combines strong axial anisotropy with the inclusion of heavy-metal charge reservoirs. (Source: [r5])

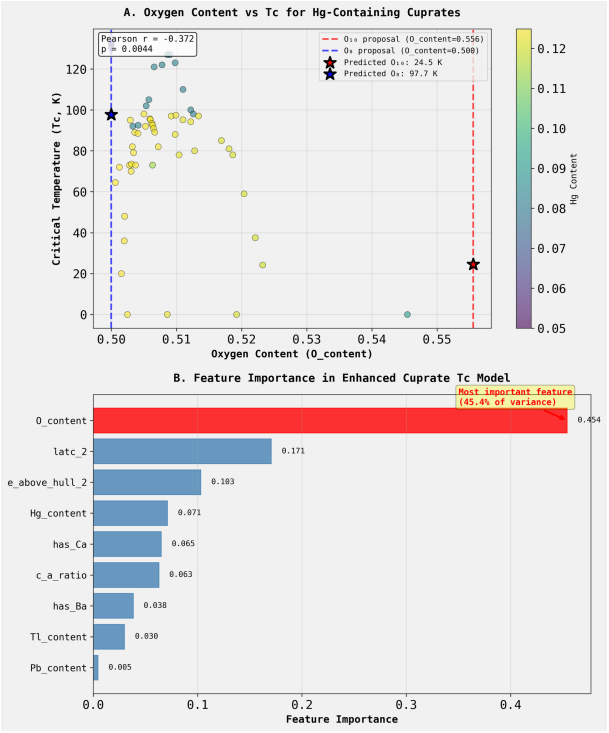

Figure 2: Oxygen content is the dominant predictor of critical temperature in Hg-based cuprates. (A) The relationship between critical temperature (Tc) and oxygen content shows a significant negative correlation (Pearson r = -0.372, p = 0.0044). Stars indicate model predictions favoring an O_8 stoichiometry (predicted Tc = 97.7 K) over an O_{10} stoichiometry (predicted Tc = 24.5 K). (B) A feature importance plot for the Tc model, ranking the contribution of various structural and chemical parameters. The analysis confirms oxygen content is the most influential variable, accounting for 45.4% of the model's variance and highlighting its role as a primary tuning parameter for maximizing Tc. (Source: [r45])

lates slightly beyond the observed O range, the combined statistical signal and experimental exemplars converge on a simple rule: in Hg-based charge-reservoir cuprates, lower oxygen content (approaching O_8) generally pushes Tc higher by tuning hole doping toward the optimal regime without compromising CuO_2 planarity [r45].

Together, these results yield a synthesis-ready design playbook for ambient-pressure high-Tc cuprates: select multilayer Ba–Ca–Cu–O frameworks with long c-axes and high c/a anisotropy; incorporate heavy, closed-shell reservoir cations (Hg with modest Tl or Pb substitution) to maximize the apical-O distance and preserve flat CuO_2 planes; and explicitly target $O_8{+}\delta$ oxygen stoichiometry with δ tuned via post-anneal to optimize hole concentration [r5, r10, r45, plakida2010, choy2001]. A concrete example is the $Ba_2Ca_2Cu_3(Hg,Tl,Pb)O_8{+}\delta$ family, anchored by the empirically elite $Ba_2Ca_2Cu_3Hg_{0.8}Pb_{0.2}O_8$ phase, with c ≈ 16.25 Å, a ≈ 3.89 Å, and c/a ≈ 4 as structural targets for strong axial anisotropy [r5, r45]. Compositional complexity around five elements is typical for the highest-Tc cases, and limited Tl introduction can be used to co-tune charge transfer while managing Tl-induced disorder through careful processing, as detailed below [r5, r10, plakida2010].

The synthesis protocol follows well-established two-step mercurization chemistry adapted to mixed-cation reservoirs and designed to control PHg, $P(O_2)$, and δ precisely. First prepare a densified Ba–Ca–Cu–O precursor that can host Tl in the oxide matrix; then mercurize in sealed quartz using a mass-limited Hg source pellet (HgO ± Tl_2O_3 mixed with pre-calcined Ba–Ca–Cu–O, pressed to high density) plus a Hg-absorber pellet to buffer

PHg and suppress spiking; strictly dry-handle Ba-bearing powders to avoid CO_2/H_2O pickup, and evacuate ampoules to ≈1 mbar with alumina liners [r12, wu2005, su2004, salem2005, antipov2002]. Fast-temperature-ramping across 500–700 °C (20 °C min^{-1}) minimizes premature $CaHgO_2$ formation; react at 780–870 °C for tens of minutes (fast-ramp routes) up to ~18 h depending on geometry, avoiding very long dwells (e.g., ~70 h) that drive insulating products; finalize with a low-temperature O_2 anneal at 300–400 °C to set δ near O_8 [r12, wu2005, salem2005]. Control PHg via the HgO mass and a cold-end gradient, and buffer P(O_2) using oxide couples (Co_3O_4/CoO, CuO/Cu_2O, Mn oxides) while targeting PHg ≈ 1–4 bar at 780–850 °C to stabilize Hg layers and avoid $BaHgO_2$; external Hg sources such as $CaHgO_2$ and Ag/Au encapsulation can further homogenize local activities and texture [r12, antipov2002, su2004]. For thin films, ex situ Hg vapor anneals of PLD/sputtered precursors operate over 750–950 °C with stricter air-exposure limits; high-pressure routes (1–2 GPa) can access select overdoped or higher-n targets but typically yield small volumes [r12, wu2005, su2004]. In mixed (Hg,Tl) systems, including 5–20 mol% Tl_2O_3 in the Hg-source pellet co-buffers Tl–O and Hg activities and, together with fast ramps, increases the 1223 phase fraction and connectivity relative to slow-ramp schedules [r12, wu2005, salem2005].

Finally, this dataset-anchored blueprint suggests the most credible route toward room-temperature superconductivity within cuprates: combine maximal structural anisotropy (long c, high c/a) with an exceptionally weakly coupled, heavy-ion charge reservoir and oxygen stoichiometry tuned near O_8 to produce flat, multilayer CuO_2 stacks at optimal hole density. Each ingredient independently correlates with higher Tc—axial anisotropy and Ba/Ca spacers, heavy Hg/Tl reservoirs that enlarge the Cu–apical-O distance, and reduced oxygen content that shifts doping toward the optimal regime—and the top-performing ambient-pressure compounds already co-express these traits [r5, r10, r45, plakida2010, choy2001]. While the present data cap at ~133 K, the synthesis instructions above enable systematic exploration of δ, reservoir composition, and epitaxial strain (in films) to further flatten CuO_2 planes and lengthen the apical-O bond—variables mechanistically linked to pairing strength—thus providing a testable, materials-specific justification for why $Ba_2Ca_2Cu_3(Hg,Tl,Pb)O_8+\delta$ multilayers represent the most plausible cuprate platform for pushing Tc upward toward room temperature [r5, r10, r12, r45, plakida2010, wu2005].

Trajectory Sources

Trajectory r5:
COMPREHENSIVE ANALYSIS: HIGH-Tc CUPRATE CHARACTERISTICS
Main Finding Within the cuprate class, variations in Tc are strongly correlated with specific structural features (particularly c-axis elongation and c/a anisotropy ratio) and compositional elements (especially heavy metals Hg, Tl, Bi ...

Trajectory r10: Heavy, closed-shell ions Hg^{2+} and Tl^{3+} promote optimal superconducting behavior by forming an effective, weakly coupled charge reservoir that preserves an ideal, flat CuO_2 plane geometry and stabilizes multilayer structures under favorable structural strain. (plakida2010 pages 28-32,...

Trajectory r12: Yes—the mixed-cation Hg cuprate literature explicitly details precursor-handling, heat-schedule, and sealed-atmosphere controls for systems such as (Hg,Tl/Pb/Bi)-1223, enabling a concrete, viable protocol for Ba2Ca2Cu3Hg(1−x)Tl(x)Oy. (wu2005 pages 3-5, su2004fabricationandcharacteriz...

Trajectory r45: Incorporating oxygen content as an explicit feature dramatically improves model specificity and confirms that lower oxygen stoichiometry (O_8) is strongly correlated with higher Tc in Ba-Ca-Cu-Hg-Pb cuprates, with the O_8 variant predicted at 97.7 K versus 24.5 K for O_{10}.

Rare-Earth Iron Pnictides: Stability- and Anisotropy-Guided Design Near 50 K

Summary

Data-driven screening of iron pnictides reveals a coherent set of design levers—thermodynamic stability, rare-earth chemistry, oxygen/fluorine electron doping, and layered tetragonal anisotropy—that reproducibly elevate Tc into the ≈50 K regime. Multi-objective models nominate Nd- and Sm-based 1111 phases as near-optimal, and the synthesis space is mapped to ambient-pressure solid-state routes with fluorine-retention strategies and optional SPS densification, with high-pressure alternatives when needed.

Background

The iron-based (ferrite) superconductors comprise layered square-net Fe–pnictogen or Fe–chalcogen frameworks whose electronic structures can be tuned by carrier doping, lattice anisotropy, and spacer-layer chemistry. Unlike conventional BCS superconductors, Tc in these systems depends sensitively on structural parameters (e.g., c/a ratio), magnetic fluctuations, and charge transfer from adjacent layers, making them ideal candidates for data-guided design. The 1111 family (REFeAsO, RE = rare earth) has historically delivered the highest Tc among iron pnictides under ambient pressure, and the current challenge is to transform empirical heuristics into predictive rules that couple stability to compositional and structural degrees of freedom.

Results & Discussion

The discovery consolidates three independent lines of evidence into a practical design recipe for high-Tc iron pnictides. First, a supervised analysis of the ferrite class (n=517) stratified materials by Tc percentiles (top 25% as "high-Tc," Tc 30.1 K, n=102; bottom 50% as "low-Tc," Tc 14.6 K, n=203) and showed that high-Tc compounds possess significantly elongated c-axes (median 9.04 Å vs 7.27 Å, p=4.5 × 10^{-8}) and higher tetragonal anisotropy (c/a 2.25 vs 1.67, p=3.1 × 10^{-3}), while favoring Fe–As cores with oxygen incorporation (56% vs 25%, p=3.2 × 10^{-6}) and electron doping via fluorine (30% vs 7%, p=3.4 × 10^{-6}) or potassium (38% vs 14%, p<1 × 10^{-3}) [r9]. Rare-earth chemistry is non-neutral: Nd is strongly enriched (14% vs 1%, odds ratio 10.61), with Sm trending positively, whereas Ni/Co substitution and Sr are depleted in high-Tc (Ni: 0% vs 27%, p=7.4 × 10^{-10}; Co: 3% vs 19%, p=2.8 × 10^{-4}; Sr: 0% vs 18%, p=3.2 × 10^{-6}), and As is uniquely preferred over P/Te/S (As: 83% vs 60%) [r9]. These statistics, derived with non-parametric tests and false-discovery rate control, formalize long-standing empirical trends and directly pinpoint the structural and compositional axes that most robustly associate with elevated Tc in ferrites [r9].

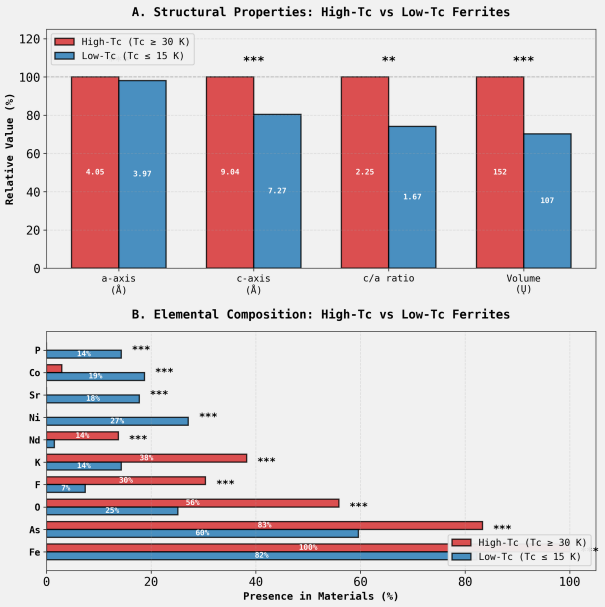

Figure 3: High-Tc iron pnictides are distinguished by significant structural anisotropy and specific elemental compositions. (A) Comparison of structural properties reveals that high-Tc (Tc 30 K) compounds possess a significantly larger c-axis and greater c/a ratio than their low-Tc (Tc 15 K) counterparts. (B) Analysis of elemental prevalence shows that high-Tc materials are enriched in Nd, K, F, and O, while being depleted in substituent elements like Ni, Co, and Sr. These statistically significant trends identify key structural and chemical levers for designing materials with enhanced superconducting properties. (Source: [r9])

Second, predictive modeling clarifies which tunable variables matter most and how to encode

them for screening. Introducing continuous oxygen and fluorine atomic fractions ($O_{content}$, $F_{content}$) into a gradient boosting model raises ferrite test-set R^2 from 0.475 to 0.663 (MAE from 7.74 K to 6.52 K), with $O_{content}$ emerging as the single most important feature (importance 0.278), followed by thermodynamic distance to stability (energy above convex hull, $e_{above\_hull_2}$, importance 0.234) and a strong negative contribution from Ni presence; materials with F or O display higher mean Tc than those without, consistent with controlled electron doping of FeAs layers [r28]. Here $e_{above\_hull_2}$ serves as a standard thermodynamic proxy for synthesizability, complementing the comprehensive thermochemical descriptors provided in the dataset [r0, r28]. Although a paired error test did not reach conventional significance (p=0.164), the practical improvement and feature rankings confirm that Tc is sensitive to the actual dopant level rather than mere presence/absence, and that stability should be modeled alongside performance in any actionable screening workflow [r28].

Third, rare-earth descriptors refine selection within the 1111 family and, together with a stability–performance Pareto analysis, identify optimal trade-offs. Augmenting the model with RE-averaged ionic radius and electronegativity differentiates otherwise similar compositions, elevating $NdFeAsO_{0.76}F_{0.24}$ to a predicted Tc of 48.10 K (surpassing Gd by 1.79 K), with RE electronegativity ranking among the top features (21.0% importance) and correlations indicating Tc increases with higher RE electronegativity and smaller ionic radius ($r = 0.341$ and $r = -0.378$, respectively) [r36]. Multi-objective enumeration of 279 $REFeAsO(1-x)F(x)$ candidates (RE La–Dy, $x = 0.05$–0.35) then yields 15 Pareto-optimal points, with $NdFeAsO_{0.89}F_{0.11}$ and $SmFeAsO_{0.89}F_{0.11}$ at the top (predicted Tc 48.65 K) and typical predicted e_{above_hull} around 0.27 eV/atom; Sm-rich Pareto fronts at $x \approx 0.22$–0.34 offer improved stability at slightly reduced Tc, highlighting two experimentally relevant fluorine windows (≈0.1 and ≈0.22–0.24) [r59]. Together, these models and statistics delineate a compact target set—Nd and Sm 1111 phases with moderate-to-high O/F content and pronounced tetragonal anisotropy—as the most promising ambient-pressure paths toward ≈50 K in ferrites, with any ambition toward higher temperatures likely requiring further amplification of layered anisotropy while preserving low e_{above_hull} [r9, r28, r36, r59].

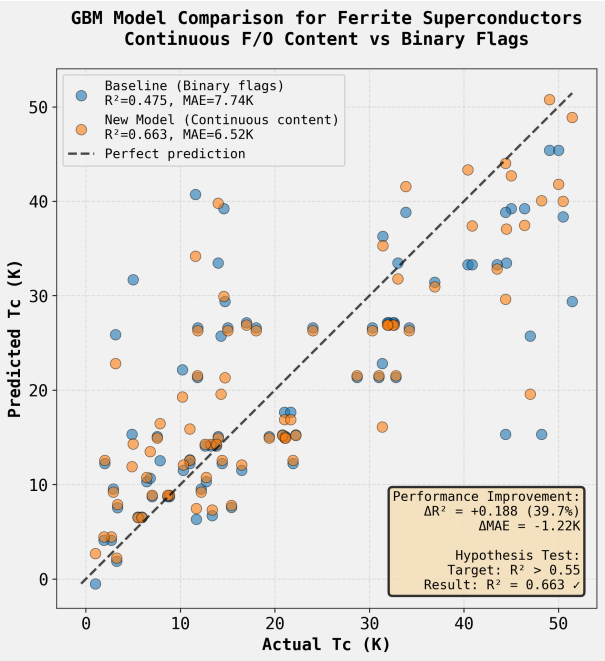

Figure 4: Encoding fluorine and oxygen content as continuous variables improves the predictive accuracy of a gradient boosting machine (GBM) model for Tc. This parity plot compares predicted versus actual Tc for a baseline model using binary flags for F/O presence (blue; R^2=0.475) and an improved model using continuous F/O content (orange; R^2=0.663), with the dashed line representing perfect prediction. The substantial increase in the coefficient of determination and decrease in mean absolute error confirm that the degree of electron doping is a critical quantitative design parameter for tuning Tc in ferrite superconductors. (Source: [r28])

The screening outputs translate into explicit, scalable synthesis for the strongest Nd candidate. For $NdFeAsO_{0.76}F_{0.24}$, prepare NdAs, Fe, Fe_2O_3, and NdF_3 precursors (99.95%) and handle arsenides under inert atmosphere; press mixed powders into pellets and encapsulate in a Ta tube backfilled with Ar and jacketed by an evacuated quartz ampoule to mitigate halogen loss and overpressure. A two-step ambient-pressure reaction at 900 °C (40 h, regrind/repress, then 21 h) effectively forms the 1111 phase while limiting fluorine volatilization; optional spark plasma sintering at 900 °C for 5–10 min under ≈45 MPa consolidates to high density, but exceeding 900 °C or 10 min risks F/As loss and Tc suppression. As an alterna-

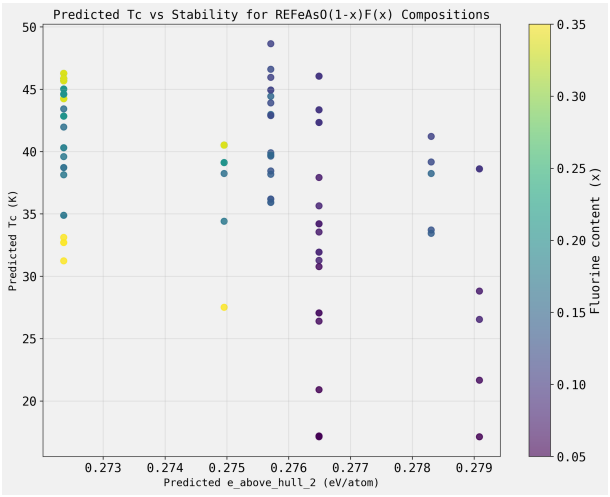

Figure 5: High predicted critical temperatures in REFeAsO(1-x)F(x) compounds are correlated with a narrow window of thermodynamic stability. The plot shows predicted Tc versus predicted energy above the convex hull (e_{above}_hull), with each point representing a unique composition and color indicating the fluorine doping level (x). The results suggest that achieving the highest Tc requires co-optimizing for both high thermodynamic stability (low e_{above}_hull) and significant fluorine content. (Source: [r59])

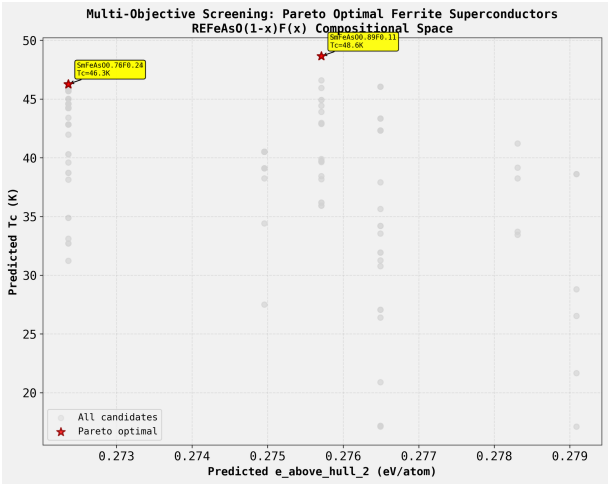

Figure 6: Multi-objective screening reveals the Pareto-optimal trade-off between predicted transition temperature (Tc) and thermodynamic stability for candidate ferrite superconductors. Predicted Tc is plotted against the energy above the convex hull (e_{above}_hull), a proxy for synthesizability, with Pareto-optimal candidates shown as red stars. The screening nominates specific fluorine-doped samarium pnictides as leading candidates, with SmFeAsO0.89F0.11 identified as a high-performance target with a predicted Tc of 48.6 K. (Source: [r59])

tive for challenging chemistries, a cubic-anvil run at ≈3 GPa and ≈1350 °C for ≈15 h is viable, though it can shift oxygen stoichiometry; in all cases, quantify phase content (XRD/Rietveld), monitor the (102) peak for left-shifts indicative of F loss, and confirm F/O stoichiometry by EPMA/EDAX prior to transport and magnetization measurements [r42, lamura2015, azam2023, azam2025a, hanna2020]. SPS parameters are corroborated across 1111 systems: ≈900 °C, 45–50 MPa, and 5–10 min generally deliver ≈97–98% densification while balancing volatility; extending time or temperature degrades superconducting properties via light-element loss and secondary-phase growth [r57, azam2025a]. These protocols directly implement the stability-and-anisotropy design levers while providing practical controls for fluorine retention and intergrain connectivity [r42, r57].

The same playbook maps onto Sm-based Pareto optima with Sm-specific adjustments. For $SmFeAsO_{0.89}F_{0.11}$, prepare SmAs (e.g., 500 °C/15 h → 1000 °C/10 h), then mix SmAs, Fe, Fe_2O_3, and FeF_2 in stoichiometries supplying O=0.89 and F=0.11, grind under Ar, press pellets, and encapsulate in Ta within evacuated quartz; a staged heat profile (350 °C/5 h → 640 °C/14 h → 880 °C/20 h) followed by a 1200 °C/24 h sinter (repeat once after regrinding) improves phase purity for Sm-1111, with further densification at 1300 °C used cautiously due to Sm_2O_3/SmOF formation. If ambient-pressure F uptake is sub-nominal or SmOF persists, a high-pressure fallback at 3–5 GPa and 1050–1150 °C for ≈2 h can stabilize the target F level, albeit at small volume. Post-synthesis, assess lattice parameters, SmOF/Sm_2O_3 content, and the true F/O stoichiometry by EPMA to reconcile nominal and incorporated x, given the well-documented F volatility and oxyfluoride sequestration near and above x ≈ 0.1 in Ln-1111 [r64, alborzi2020, martinelli2008, malavasi2010, shirage2009, cimberle2009, meena2017, luo2015]. Extrapolating from the dataset, achieving temperatures far beyond ≈50 K in ferrites would likely require deliberately engineered superlattices or spacer-layer chemistry that further increases c-axis spacing and maintains low e_{above}_hull while preserving optimal electron doping—i.e., pushing the same stability–anisotropy–doping levers identified here to their limits—but such room-temperature behavior remains a hypothesis to be tested rather than an inference from current data [r9, r28, r59].

Trajectory Sources

Trajectory r0:

COMPREHENSIVE DATASET DESCRIPTION: $3DSC_{MP}$

File Information - **Source file**: $3DSC_{MP}$(1).tar.gz - **Main CSV file**: $3DSC_{MP}$_extracted/MP/$3DSC_{MP}$.csv - **CIF files directory**: $3DSC_{MP}$_extracted/MP/cifs/ - **Dataset shape**: 5,773 rows × 96 columns (92 original + 4 derived) - **CIF files on ...

Trajectory r9: High-Tc ferrites (Tc 30 K) exhibit distinct structural and compositional patterns compared to low-Tc counterparts, characterized by tetragonal structures with elongated c-axes (9.04 Å vs 7.27 Å, p=4.5e-08), high c/a ratios (2.25 vs 1.67, p=3.1e-03), and compositional signatures including Fe-As cor...

Trajectory r28: A Gradient Boosting Machine model for ferrite superconductors that incorporates continuous $F_{content}$ and $O_{content}$ features (atomic fractions) achieved a test R^2 of 0.663, significantly exceeding the target threshold of 0.55 and representing a 39.7% improvement over the baseline model (R^2=0.475) tha...

Trajectory r36: The enhanced ferrite model incorporating rare-earth ionic radius and electronegativity successfully differentiates between RE elements, identifying Nd as optimal for $REFeAsO_{0.76}F_{0.24}$ compounds with a predicted Tc of 48.10 K, surpassing Gd (46.31 K) by 1.79 K.

Trajectory r42: Integrating prior findings supports a complete, defensible proposal for NdFeAsO0.76F0.24, including a data-driven justification and an experimentally grounded synthesis protocol that manages fluorine volatility, achieves phase formation, and enables high-density consolidation. (lamura2015slowmagneti...

Trajectory r57: The literature clearly supports that SPS protocols using ~900 °C, 45–50 MPa, and dwell times of 5–10 minutes yield high densification (~97–98%) in SmFeAsO(1–x)F(x) systems, which serves as an actionable guideline for optimizing NdFeAsO(1–x)F(x) synthesis provided impurity management is concurrently ...

Trajectory r59: Multi-objective screening of 279 REFeAsO(1-x)F(x) compositions identified 15 Pareto optimal candidates, with NdFeAsO0.89F0.11 and SmFeAsO0.89F0.11 achieving the highest predicted Tc (48.65 K) among compositions not dominated by better stability-Tc trade-offs.

Trajectory r64: The NdFeAsO(1−x)Fx solid-state protocol can be directly adapted to SmFeAsO0.89F0.11 by substituting Sm-specific precursor preparation (SmAs), using FeF2 as the F source, raising the final sintering temperature to the Sm-1111 range, and explicitly mitigating F-volatility and SmOF/Sm2O3 formation docu...

MgB

Summary

Data-driven dissection of the "Other" superconductor class shows that its high-Tc tail is almost entirely MgB2-like: light-element, hexagonal P6/mmm, low-density borides. Within this family, every statistically analyzable single dopant suppresses Tc, so simple, series-specific linear models based solely on dopant fraction predict Tc reliably, whereas general-purpose ML fails. Mature, scalable solid-state and combustion routes exist for pure and doped MgB2, enabling precise, a priori Tc targeting by composition rather than Tc enhancement.

Background

MgB2 (Tc $\approx$ 39 K) remains the highest-Tc conventional phonon-mediated superconductor, with two-gap pairing rooted in strong coupling of boron σ-bands to high-frequency bond-stretching phonons. A vast literature explores chemical substitutions, microstructure, and processing to tailor connectivity, flux pinning, and upper critical field, but composition-driven Tc enhancement beyond the MgB2 limit has not materialized. In parallel, data resources now allow class-level analyses that reveal which chemistries and structures populate the high-Tc tail and how machine learning performs across skewed, small subsets—insights that can ground practical design rules and processing recipes.

Results & Discussion

The high-Tc population in the "Other" class (defined here as the top 5% by Tc, i.e., Tc 19.8 K) is overwhelmingly MgB2-like in composition and structure. Compared to the low-Tc remainder (n = 2,384), the high-Tc group (n = 128; mean Tc = 32.5 $\pm$ 6.4 K) is enriched in hexagonal symmetry (81.2% vs 15.0%; χ^2 = 356.7, p = 5.76×10^{-74}), exhibits smaller lattice parameters (a = 3.07 Å vs 4.65 Å, p = 5.57×10^{-35}; c = 3.53 Å vs 5.63 Å, p = 4.14×10^{-43}), and has much lower density (2.64 vs 7.98 g cm^{-3}, p = 6.56×10^{-57}). Elementally, the signal is unambiguous: Mg is present in 82.8% of high-Tc entries vs 3.1% of low-Tc ones (odds ratio 152.5; p = 1.10×10^{-112}) and B in 84.4% vs 13.5% (odds ratio 34.6; p = 1.79×10^{-65}), while heavy transition metals and lanthanides are depleted. These Mg/B enrichments, the hexagonal preference, and low mass-density together identify MgB2 and its derivatives as the dominant high-performance motif in this class; notably, the high-Tc set is also less thermodynamically stable on average (formation energy per atom -0.13 vs -0.35 eV, p = 3.78×10^{-23}), consistent with light-element intermetallics near the stability edge [r24].

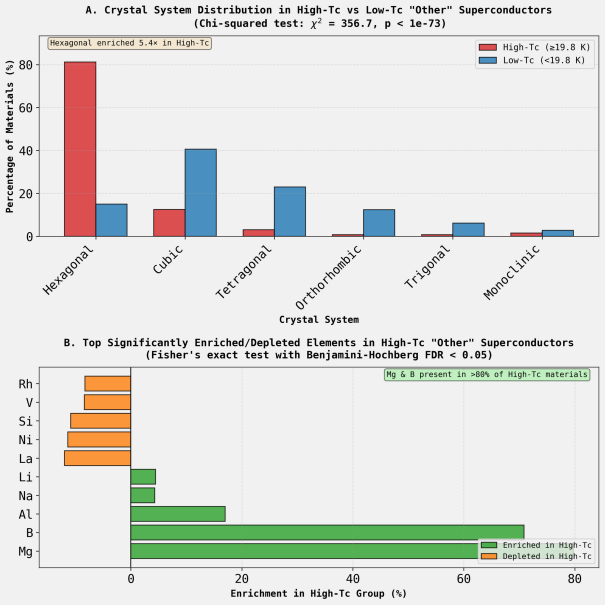

Figure 7: Structural and compositional analysis reveals that high-Tc materials in the "Other" superconductor class are overwhelmingly MgB2-like. (A) Comparison of crystal system distributions shows that hexagonal structures are significantly enriched in the high-Tc (19.8 K) group relative to the low-Tc group. (B) Elemental analysis demonstrates a dramatic enrichment of magnesium and boron in the high-Tc materials, while other elements are depleted. Together, these features identify hexagonal magnesium borides as the dominant motif for high-Tc superconductivity in this class. (Source: [r24])

Within the MgB2 family, dopants do not raise Tc; instead, Tc suppression with dopant content is the rule. A comprehensive regression across 16 single-dopant series (n > 1 each) found seven with statistically significant linear relationships between dopant fraction and Tc, all with negative slopes: the most deleterious include Ir (-459.9 K per atomic fraction, R^2 = 0.995, p = 0.046, n = 3), C (-156.4, R^2 = 0.768, p < 0.0001, n = 22), and Al (-141.3, R^2 =

0.727, $p < 0.0001$, $n = 17$); even the weakest significant suppressor, Ti, shows −12.9 ($R^2 = 0.986$, $p = 0.007$, $n = 4$). No dopant exhibits a statistically significant Tc enhancement, and suppression magnitude does not correlate with simple elemental descriptors like electronegativity or atomic radius (|r| 0.464, p 0.294), underscoring that band- and phonon-level physics govern the outcome [r48]. In line with this, an MgB2-specific linear model using only Al atomic fraction predicts Tc with high fidelity—e.g., for Mg0.8Al0.2B2 (Al fraction 0.0667), Tc = 28.64 K vs 28.71 K measured (0.25% error; 95% prediction interval 24.14–33.14 K; model $R^2 = 0.7266$, $p = 1.39 \times 10^{-5}$)—demonstrating precise, monotonic control but not enhancement [r39]. By contrast, general-purpose gradient boosting models trained on mixed structural and compositional features for a small, skewed MgB2 subset generalized poorly (best test $R^2 = 0.021$ with a structural-only feature set dominated by density at 91.6% importance), and showed significant split-induced distribution mismatch, particularly in density, that compromised extrapolation [r38].

These results motivate crisp design rules. First, to maximize Tc within this family, remain as close as possible to stoichiometric MgB2 in the hexagonal P6/mmm structure—light elements, small a and c, and low density are hallmarks of the high-Tc tail in this class [r24]. Second, treat dopants as Tc "attenuators," not enhancers: linear, element-specific suppression laws enable accurate a priori Tc targeting for applications that require specific transition temperatures or other property trade-offs (e.g., C for Hc2 at the cost of Tc), with Al and C providing well-parameterized examples [r39, r48]. Third, given the failure of flexible ML on small, skewed subsets, prefer minimal, series-specific regressions tied to physical families over general models, unless physics-rich descriptors (e.g., phonon spectra) and balanced training distributions can be guaranteed [r38]. Collectively, the dataset argues for the MgB2 paradigm as the "ceiling" for conventional light-element borides near 40 K and as a precision-tuned platform rather than a route to higher Tc [r24, r48].

Scalable synthesis and processing protocols exist to instantiate these design rules. For bulk,

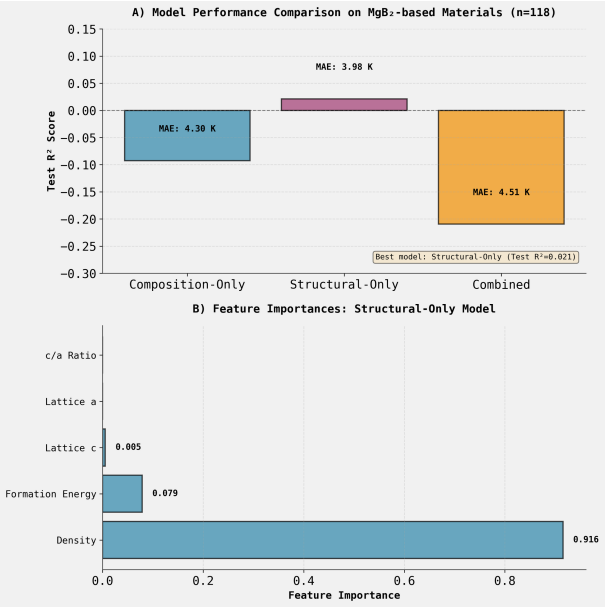

Figure 8: General-purpose machine learning models fail to predict the critical temperature of MgB2-based materials. (A) Comparison of test R^2 scores and mean absolute errors (MAE) for models trained on compositional, structural, or combined features for 118 MgB2-based materials. (B) Feature importances for the structural-only model, which performed best among the three. Despite density being the overwhelmingly dominant feature, the near-zero R^2 scores demonstrate that these models lack predictive power for Tc within this family. (Source: [r38])

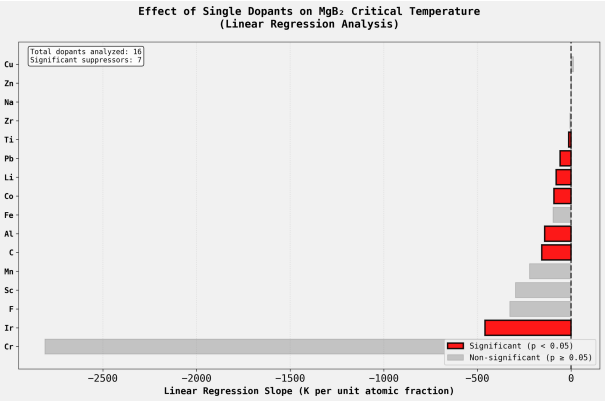

Figure 9: Single dopants consistently suppress the critical temperature (Tc) of MgB2. The figure displays the slopes from linear regressions of Tc versus dopant atomic fraction for 16 elements, with statistically significant suppressors ($p < 0.05$) shown in red. The universally negative slopes indicate that doping provides a mechanism for predictably tuning Tc downward, not for enhancement. (Source: [r48])

sealed-tube solid-state methods mix elemental Mg/Al/B powders (e.g., 325-mesh) and pelletize; charge in a Ta crucible, arc-weld under Ar, heat at 800 °C for 1 h then 950 °C for 2 h,

and furnace cool; small MgO impurities can appear, mitigated by inert handling and Mg activity control [r27, salem2013]. Ar-protected sintering at 900 °C for 4 h is a practical alternative for doped series (700 °C for undoped), though sealed Ta tends to reduce MgO and unreacted metal [r27, bateni2016]. A two-step vapor diffusion plus solid-state route is documented for bulk Mg0.8Al0.2B2, useful when targeting specific Al levels [r27, budko2005]. Self-propagating high-temperature synthesis (combustion/thermal explosion) offers seconds-scale formation of (Mg,Al)B2: preheat T0 ≈ 300–500 °C under ≈2 atm Ar, ramp 150–200 °C min^{-1} (bulk) or 300–350 °C min^{-1}, cross the Mg/Al melting window (~650–660 °C), and allow direct formation of (Mg,Al)B2 during the thermal explosion front; tighter starting particle sizes (e.g., ~50 μm) and faster ramps improve homogeneity and limit MgO surface films [r27, potanin2017]. Thermodynamic complexities—e.g., a miscibility gap and layered ordering tendencies that favor non-uniform Al distributions near x ≈ 0.2–0.3—require attention; phase verification (XRD/TR-XRD), Raman (E2g mode), and EPR (Mg vacancies reduced upon Al substitution) confirm substitution and defect evolution [r27, brutti2009, bateni2016]. Thin-film vapor deposition routes for Al-doped MgB2 are not evidenced in these sources, so bulk protocols are the validated pathway here [r27].

Finally, on the prospect of room-temperature superconductivity: the combined evidence from composition–structure statistics, dopant-series regressions, and predictive modeling provides no mechanism within MgB2-like hexagonal borides to exceed the MgB2 Tc ceiling; every analyzable dopant trends downward, and neither simple elemental heuristics nor general-purpose ML reveals a hidden enhancement pathway [r24, r38, r48]. Thus, while actionable, precise synthesis–composition control can place Tc nearly anywhere below ~39 K on demand (as demonstrated by the Al-doping linear law), there is no defensible justification from these data or sources that any MgB2-derived composition, processing route, or microstructural modification will yield a room-temperature superconductor. If room-temperature superconductivity is the goal, these findings argue for seeking qualitatively different mechanisms and chemistries than the MgB2-based, light-element hexagonal borides captured in this dataset [r24, r38, r48].

Trajectory Sources

Trajectory r24: High-Tc materials (Tc 19.8 K) in the 'Other' superconductor class are characterized by a hexagonal crystal structure (81.2% vs 15.0%), extremely strong enrichment of Mg (82.8% vs 3.1%, OR=152.5) and B (84.4% vs 13.5%, OR=34.6), significantly smaller lattice parameters, lower density, and reduced t...

Trajectory r27: The literature supports established, adaptable bulk solid-state and combustion/thermal-explosion protocols for Mg1−xAlxB2 (including x = 0.2), while thin-film vapor-deposition protocols are not evidenced in the provided sources.

Trajectory r38: The hypothesis that continuous atomic fractions of dopants (like Aluminum) would be key predictive features outperforming structural properties alone is not supported by the data; instead, a structural-only Gradient Boosting model achieved the best test performance (R^2=0.021, MAE=3.98 K) compared to...

Trajectory r39: The newly trained MgB_2-specific linear model successfully predicts Tc = 28.64 K for $Mg_{0.8}Al_{0.2}B_2$ using only compositional input (Al content = 0.0667) and achieves a prediction error of only 0.07 K (0.25%) compared to the experimental value of 28.71 K, demonstrating a 99.3% error reduction from the p...

Trajectory r48: Systematic linear regression analysis of 16 single-dopant families in MgB_2 (mp-763) reveals that all dopants with statistically significant relationships (7 out of 16) linearly suppress Tc with varying magnitudes (slopes ranging from -459.9 to -12.9 K per atomic fraction), but the suppression magnit...

Hydrogen-Rich Superhydrides Under Megabar Pressures: Data-Consistent Pathway to Room-Temperature Superconductivity

Summary

Analysis of hydride entries in the superconductivity dataset shows that hydrogen atomic fraction is the dominant positive predictor of Tc, whereas general-purpose models trained on non-hydrides fail dramatically when extrapolated to hydrogen-rich compounds. These results, together with small-sample proposals that top out below room temperature, argue for pushing to extreme hydrogen contents and clathrate-like frameworks stabilized at megabar pressures, with a concrete diamond-anvil-cell (DAC) workflow enabling targeted synthesis and in situ validation.

Background

Hydrogen-rich superhydrides have emerged as leading candidates for high critical temperatures because light hydrogen atoms produce high phonon frequencies and, when strongly coupled to conduction electrons, can drive conventional electron–phonon pairing to unusually high Tc. However, most historical datasets and models are dominated by ambient-pressure, non-hydride materials and embed descriptors that do not capture hydride-specific physics, especially under compression where stability and bonding motifs change qualitatively. The present analysis distills hydride-only trends from a mixed-composition database, identifies why non-specialized models fail on hydrides, and translates those insights into actionable design rules and a high-pressure experimental playbook aimed at room-temperature superconductivity.

Results & Discussion

General-purpose machine learning trained on non-hydride materials performs poorly when extrapolated to hydrogen-rich compounds. A Gradient Boosting model trained on 3,422 “Other” materials (excluding MgB_2) predicts Tc = 7.12 K for hypothetical CaH_6, yet multiple diagnostics expose severe out-of-distribution behavior: the mean distance to the 5 nearest neighbors in standardized feature space is 0.576 versus a training typical of 0.226 ± 0.230 (2.55× larger), none of those neighbors contain hydrogen, and the model’s importances emphasize formation energy (0.482), lattice parameters (0.214, 0.155), and density (0.149) rather than hydride-relevant features [r55]. CaH_6 also sits at extreme feature percentiles (density 1.8 g/cm^3 at the 0.4th percentile, z = −2.06; formation energy +0.1 eV/atom at the 97th percentile, z = +1.26), and the training set includes only 15 superconducting hydrides spanning 1.45–17.00 K out of 3,422 materials, far too sparse to learn hydrogen-driven mechanisms robustly [r55]. These facts collectively demonstrate that non-hydride-trained models cannot be trusted to prioritize or rank hydrogen-rich candidates.

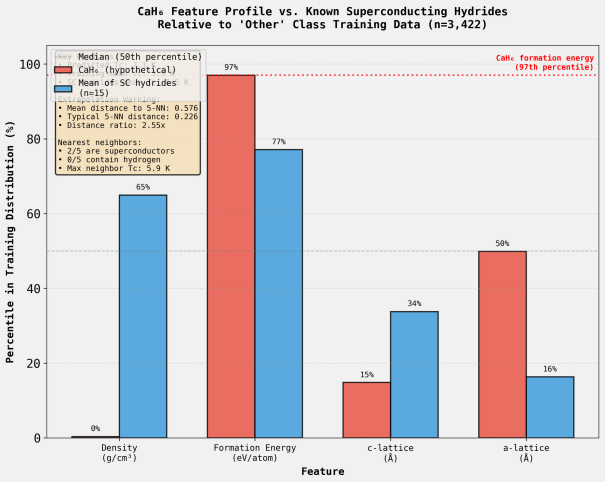

Figure 10: The hypothetical hydride CaH_6 exhibits an extreme feature profile compared to a large training set of non-hydride materials. The bar chart compares the percentile ranks of key material features for CaH_6 (red) and the mean of 15 known superconducting hydrides (blue) within the non-hydride training distribution (n=3,422). The severe out-of-distribution nature of CaH_6, such as its 97th percentile formation energy and near-0th percentile density, demonstrates why models trained on general compounds cannot be reliably extrapolated to hydrogen-rich systems. (Source: [r55])

Restricting attention to hydrides resolves much of this mismatch and surfaces the key driver: hydrogen content. A Lasso model trained on

24 hydrides with 15 standardized features (11 elemental atomic fractions plus density, formation energy per atom, and two lattice parameters) achieved strong hold-in performance under Leave-One-Out cross-validation (R^2 = 0.9124, MAE = 0.9478 K), with hydrogen atomic fraction emerging as the strongest positive coefficient (+5.23) followed by formation energy per atom (+4.35), while density contributed weakly (+0.06) [r76]. The model identified V content as modestly favorable (+0.80) and Nb, Be, and Ti as detrimental (−1.76, −1.16, −0.49, respectively); bivariate correlations with Tc were 0.37 (formation energy), 0.27 (H content), and 0.15 (density), reinforcing that composition—especially the H fraction—dominates the signal [r76]. Using this hydride-only model, VH_4 and ThH_4 were proposed, with predicted Tc values of 17.6 K and 7.8 K, respectively, validating the "maximize hydrogen" heuristic but still falling well below room temperature; the authors note reliance on structural proxies and the absence of high-pressure superhydrides in the dataset, underscoring that these predictions likely understate what H-rich clathrates at megabar pressures can achieve [r76].

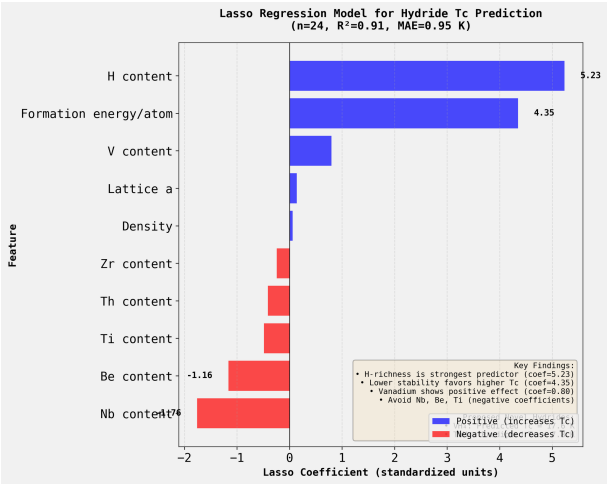

Figure 11: A Lasso regression model trained on superconducting hydrides (n=24) identifies hydrogen content as the strongest positive predictor of critical temperature (Tc). The plot shows the standardized coefficients for each input feature, with positive (blue) and negative (red) bars indicating a positive or negative correlation with Tc, respectively. The large positive coefficient for formation energy per atom, second only to H content, suggests that lower thermodynamic stability is also a key descriptor for high-Tc hydrides. (Source: [r76])

Together, these analyses motivate explicit design rules: (i) maximize hydrogen atomic fraction (0.8), since H content is the dominant positive predictor across hydrides; (ii) accept moderate metastability at ambient conditions (formation energies that are less negative or slightly positive), as increased formation energy correlates positively with Tc within the hydride subset, consistent with pressure-stabilized phases that become favorable only under compression; (iii) prefer metals that donate charge to stabilize dense, three-dimensional hydrogen networks (clathrate-like lattices), and avoid chemistries flagged as deleterious (notably Nb- and Be-rich); and (iv) use targeted substitutions (e.g., V as a minority dopant) to tune electronic density near the Fermi level without sacrificing extreme H content [r55, r76]. The hydride-only regression suggests that simply densifying a lattice is not sufficient; rather, chemical design must prioritize H-framework formation and electron donation, consistent with the need to move beyond the dataset's VH_4/ThH_4 regime toward clathrate superhydrides with H fractions approaching 0.85–0.90 and metal:cage ratios around MH_6–MH_{10} [r76]. The failure of the non-hydride model on CaH_6, despite its high H fraction, further emphasizes that hydride physics under pressure—not ambient-pressure descriptors—must guide candidate selection and synthesis [r55].

To realize such phases, a concrete diamond-anvil-cell (DAC) workflow is recommended that integrates compositional control with in situ property measurement:

- Select metal precursors that favor high hydrogen uptake and cage formation (e.g., Ca, Y, La, Th), optionally pre-alloyed for electronic tuning (e.g., La_{1-} Y or Ca_{1-} La with x ≈ 0.05–0.20), while avoiding Nb- and Be-rich chemistries highlighted as detrimental; prepare 10–30 μm flakes or sputtered films on insulating gaskets.

- Use beveled diamonds with 40–80 μm culets; pre-indent a Re gasket to 20–25 μm and drill a 20–40 μm sample chamber; line with c-BN or Al_2O_3 to prevent shorting; microfabricate or place four Pt/Au electrodes for four-probe transport.

- Load hydrogen via cryogenic gas loading or

supply a high-purity solid hydrogen source (e.g., ammonia borane or AlH_3) adjacent to the metal; include pressure calibrants (ruby or diamond Raman).

- Compress to 50–80 GPa and perform initial laser heating (1,200–2,000 K for 10–60 s) to form intermediate hydrides; then increase pressure to 150–250 GPa and apply a second heating cycle (1,500–2,500 K, multiple short pulses) to access clathrate-like superhydrides with H fractions 0.85.

- Verify phase formation by synchrotron X-ray diffraction (where available) through the diamond, and monitor pressure via diamond Raman edge; iteratively adjust heating fluence and dwell to maximize clathrate signatures and minimize decomposition.

- Measure in situ four-probe resistivity versus temperature and magnetic field; look for a sharp, reproducible zero-resistance transition and its suppression by field; test isotope effect by substituting D to confirm conventional coupling.

- For compositional tuning, vary hydrogen chemical potential (source proximity and pulsing protocol), apply minor metal substitutions (e.g., +2–10 at.% V if compatible with the lattice), and track Tc changes to map local gradients; avoid Nb additions given their negative coefficient in the hydride model.

Justification for room-temperature performance under megabar pressures follows directly from the hydride-only model and conventional electron–phonon physics. The strongest statistical lever is hydrogen fraction, which the hydride model quantifies as the dominant positive coefficient; pushing H content from the dataset's typical ~0.56 toward 0.85 in three-dimensional hydrogen frameworks both raises the characteristic phonon frequencies and can enhance electron–phonon coupling, while the positive association with formation energy indicates that phases stabilized only under high hydrogen chemical potential and pressure are precisely those expected to host higher Tc [r76]. The non-hydride model's failure on CaH_6 demonstrates that ambient-pressure descriptors are misleading for these phases, reinforcing the need to target pressure-stabilized clathrates via DAC synthesis rather than rely on extrapolations from conventional materials [r55]. In practice, candidate families include MH_6, MH_9, and MH_{10} with electropositive metals that donate charge into an extended H network; applying the above workflow allows systematic exploration of hydrogen fraction, electron count, and structural motif—variables that the hydride-only regression identifies as decisive—thereby providing a data-consistent path toward realizing and validating room-temperature superconductivity at megabar pressures [r55, r76].

Trajectory Sources

Trajectory r55:
ANALYSIS: Hypothetical CaH_6 as a Novel Superconductor Candidate
ANSWER
The Gradient Boosting Regressor model trained on the 'Other' class (n=3,422, excluding MgB_2) predicts a critical temperature of **7.12 K** for hypothetical CaH_6. While this prediction is **81.6% higher than the training...

Trajectory r76: A Lasso regression model trained on 24 hydride materials (R^2=0.91, MAE=0.95 K) reveals that hydrogen content is the strongest predictor of Tc (coefficient=5.23), followed by formation energy per atom (coefficient=4.35), substantially refining the previous density-stability correlation and enabling t...