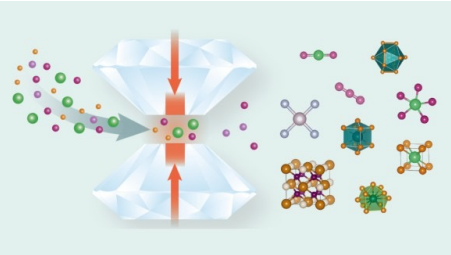
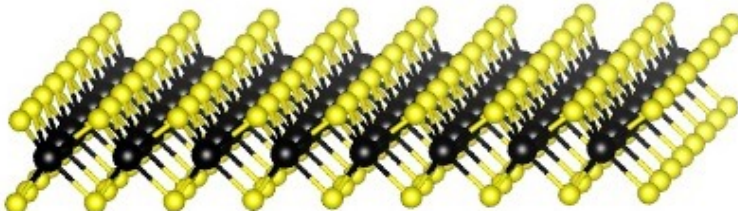
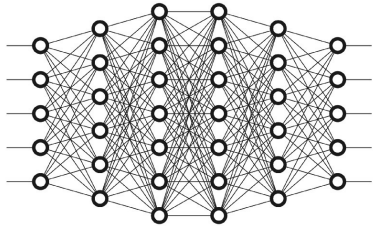
From New Chemistry to New Materials: Computational and Data-Driven Discovery

Austin Ellis
Department of Chemistry & Biochemistry, CSUN







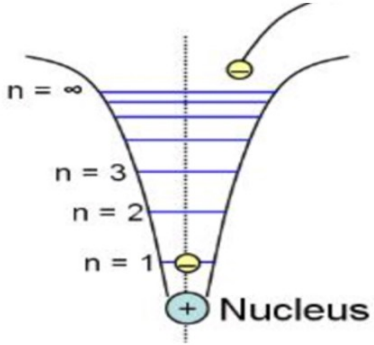
May 1, 2026

Outline

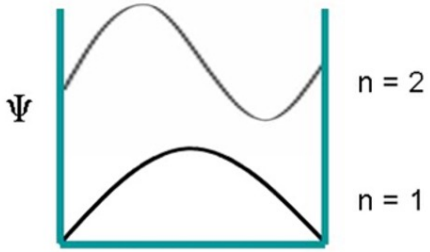
- **Background**
  - Quasi-atoms and chemical templates
  - Density Functional Theory and Crystal Structure Search
  - Machine Learning and Deep Neural Networks
- **Topic 1. Designing Tunable 2D Electrides**
  - TMX_2 and the role of chemical templates.
  - Tunable 2D electrides through defect engineering
- **Topic 2. High Pressure Superconductivity**
  - Superconductivity in Molecular Hydrides
  - Metallization in High-Pressure H_2 molecular crystals
- **Topic 3. Machine-learning Accelerated Materials Discovery**
  - Chemical-template Guided Superhydride Discovery
  - Seitz-invariant ELF Prediction via a 3D CNN
- **Conclusions**

Quasi-atom: a missing concept in solid compounds

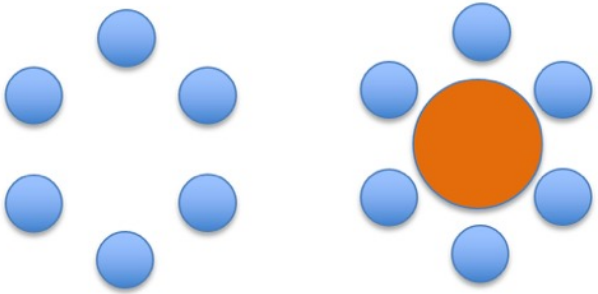
Atom



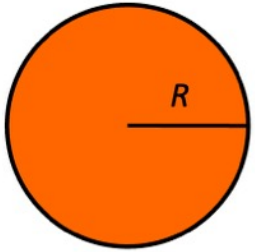
Particle-in-box



Quantum confinement from boundary condition



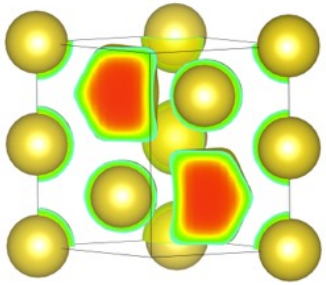
Quasi-atom



In solid compounds, all local orbitals are born equal.

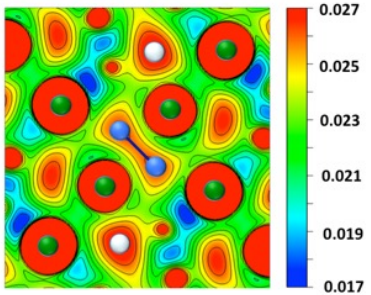
Quasi-atom Chemistry in materials

Quasi-anions



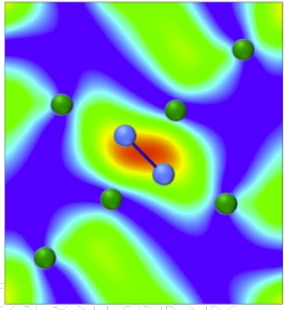
Na 200 GPa

Quasi-molecules

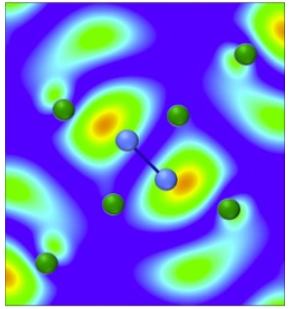


Li 80 GPa

Quasi-atom bonds



Bonding



Anti-bonding

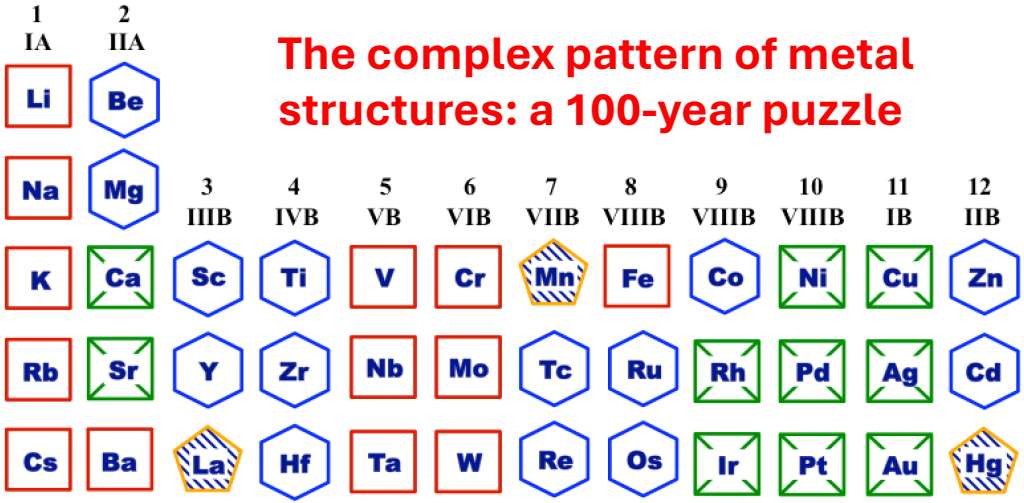
Once occupied, interstitial states become chemically active sites

Overlap creates distinct motifs: quasi-anions, quasi-molecules, and bonding/anti-bonding quasi-atom pairs

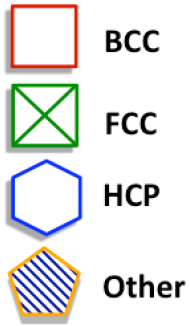
Motifs help explain stability, band-gap behavior, and electron transport in dense materials

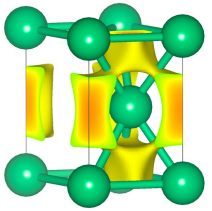
Miao et al. ***Angew. Chem.*** **2016,** 56, 972-975

Quasi-atom bonds govern metal structures

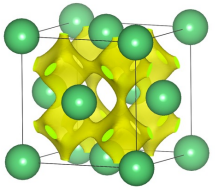


The complex pattern of metal structures: a 100-year puzzle

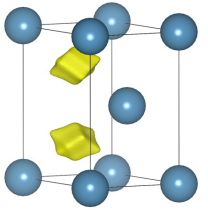




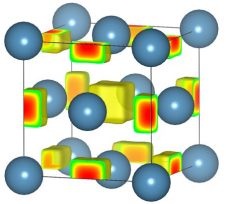
Be HCP



Be FCC



Ca HCP

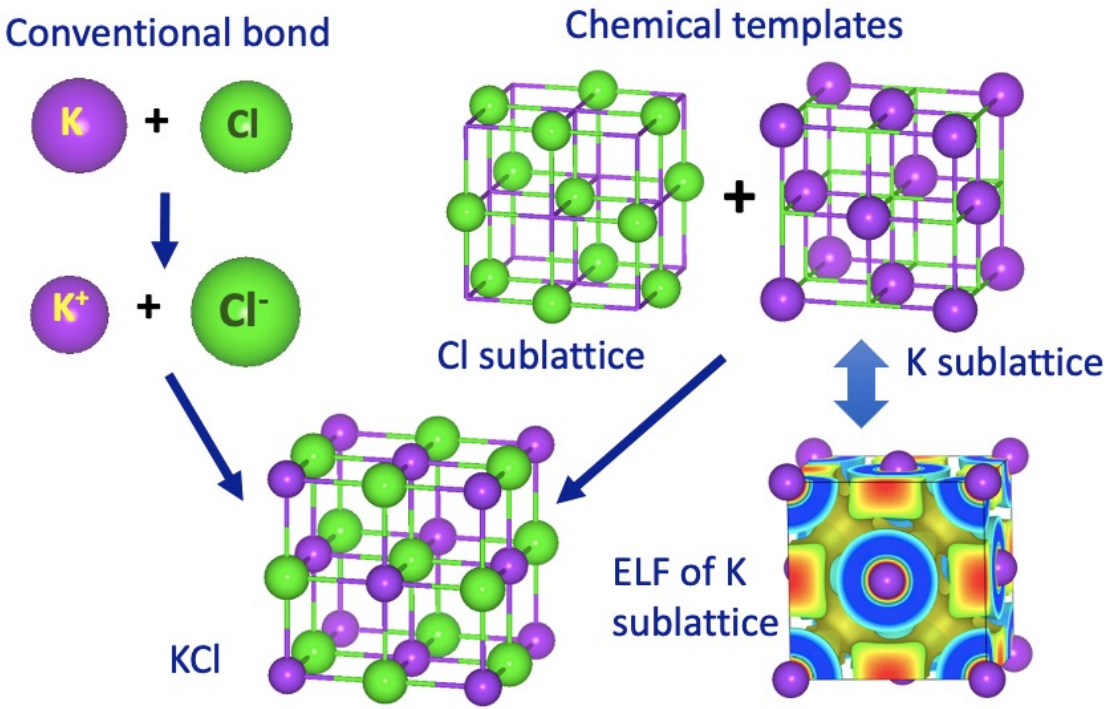


Ca FCC

- ***Alkali metals*** adopt BCC
- ***Alkaline metals*** adopt close packed except Ba (BCC)
- ***Transition metals***
  - Grps 3&4 HCP → Grps 5&6 BCC →
  - Grps 7&8 HCP → Grps 9,10,11 FCC → Grps 12 HCP

Sun & Miao, ***PNAS* 2023**

Chemical template theory



Sun & Miao, ***Chem* 2023**, 9, 443–459.

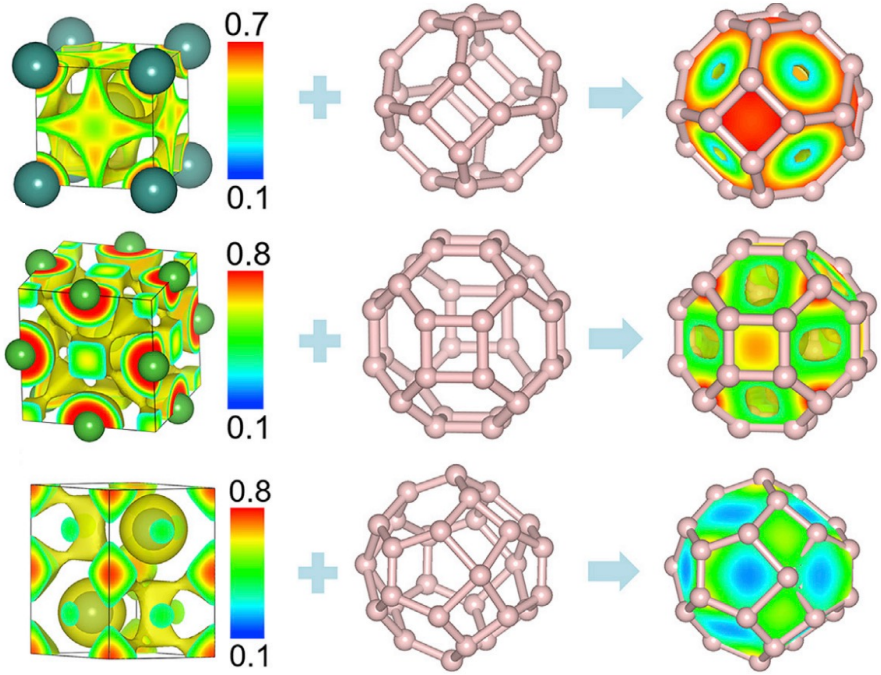
The ELF and the chemical templates in superhydrides

Electron localization function (ELF) compares local kinetic energy density to a homogeneous gas

$$\mathrm{ELF}(\mathbf{r}) = \frac{1}{1 + \chi_\sigma^2(\mathbf{r})}$$

High value = localized electrons
Low value = delocalization

Hydrogen atoms assemble in interstitial sites with high electron localization



Sun & Miao, ***Chem* 2023**, *9*, 443–459.

Density Functional Theory and Structure Search

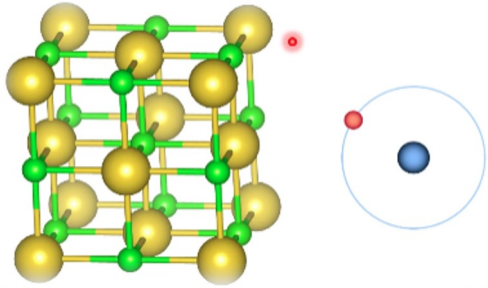
Provides energies, forces, and electronic structure by solving for electron density

Allows relaxation of structures at target pressures via enthalpy comparisons

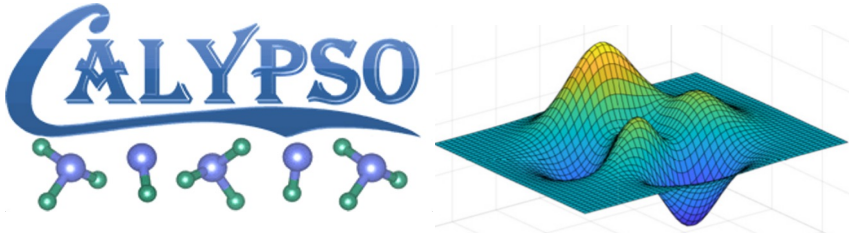
CALYPSO searches over lattice vectors and atomic positions to find low-enthalpy structures

Best candidates undergo further calculations and analysis

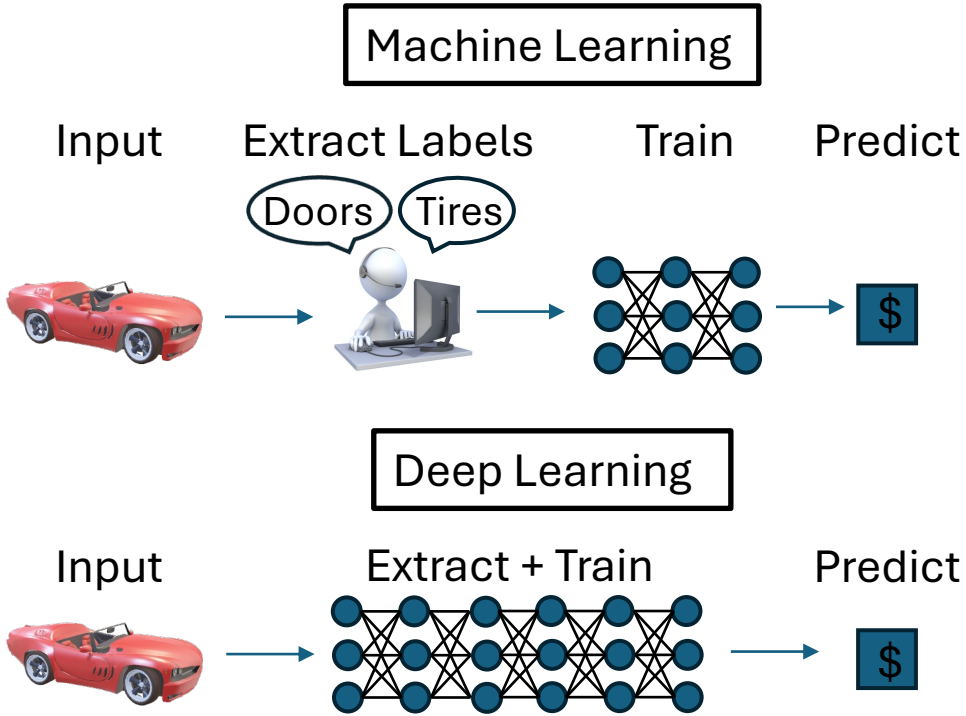
First principles (QM)



$$-\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi = E\Psi$$



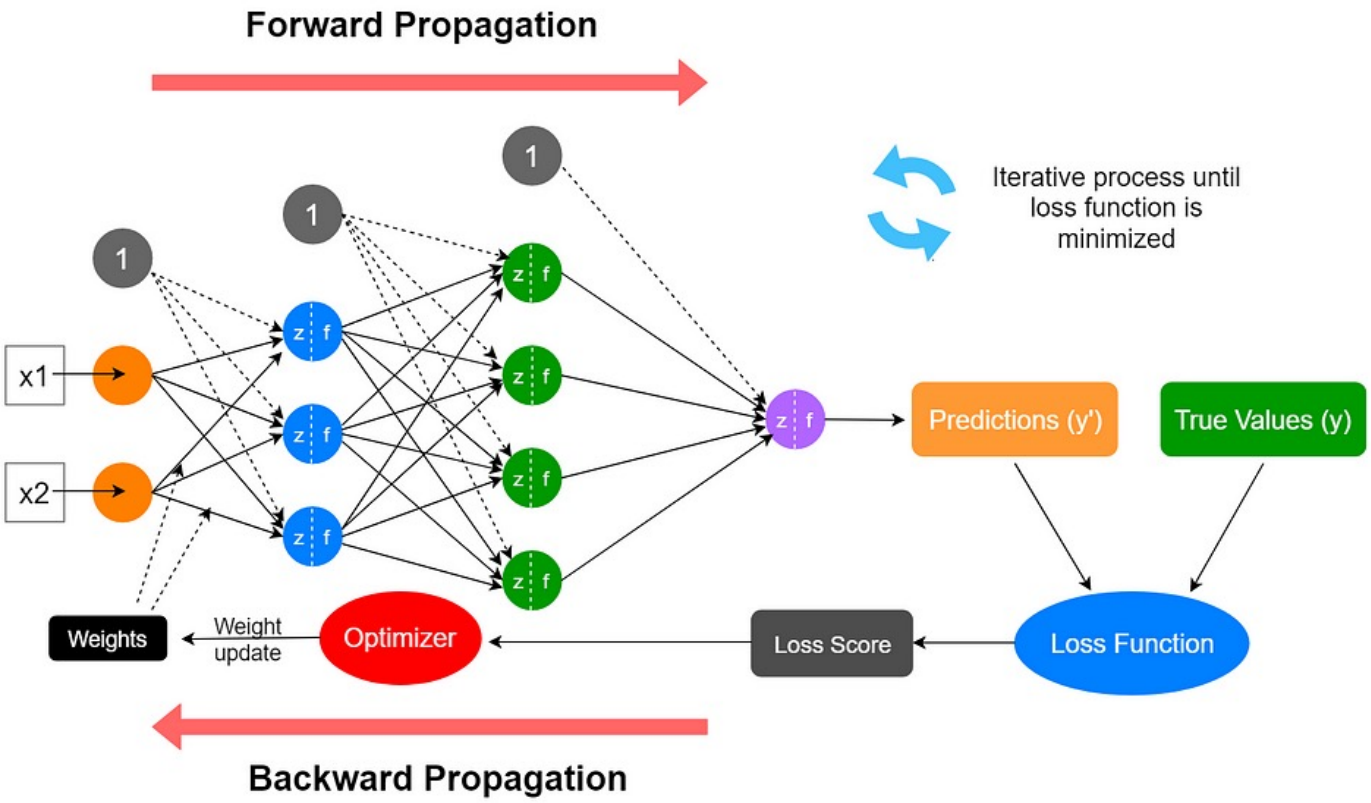
Machine Learning vs Deep Learning



Traditional machine learning usually depends on hand-crafted descriptors extracted from inputs

Deep learning maps those representations automatically during training

Anatomy of a neural network

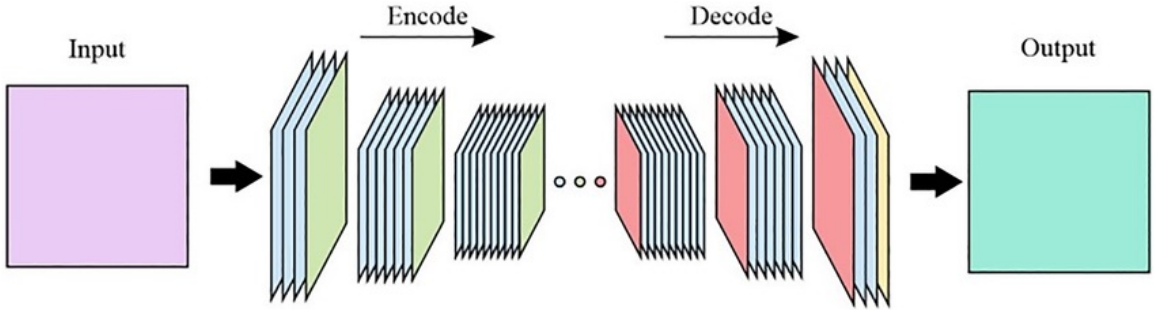


AutoGluon and Convolutional Neural Networks

AutoGluon

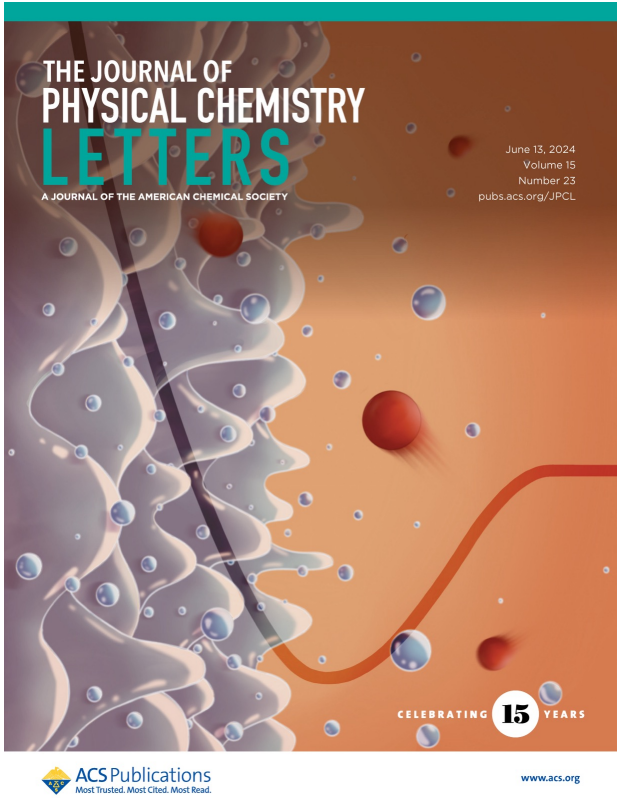
Automated machine learning framework that trains, tunes, and combines multiple models

Useful for feature-based data where input is a set of descriptors



Convolutional neural networks (CNNs) learn spatial patterns from input

In encoder-decoder form, it can transform one spatial field to another

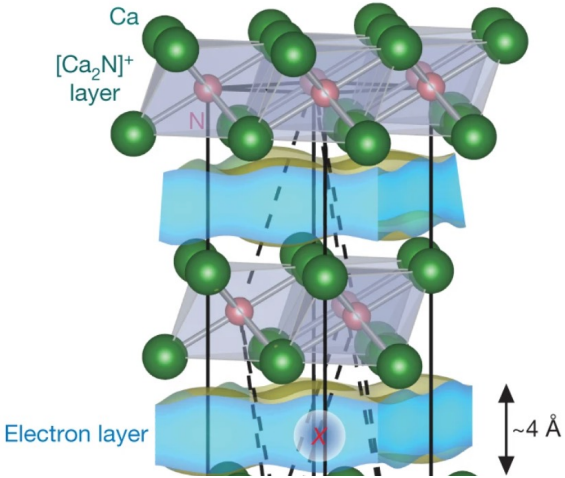


Topic 1: Designing tunable TMX_2 electrides

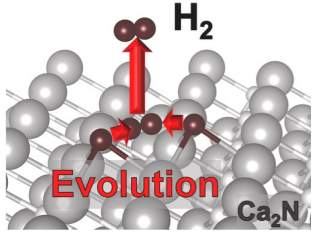
Sun, Ellis, et al. J. Phys. Chem. Lett. 2024, 15, 6174–6182.

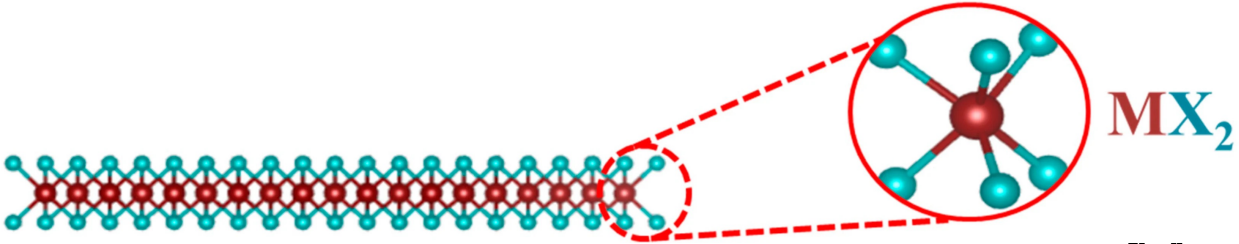
TMX_2 and Chemical Templates

From Ca_2N Electrides to TMDC Catalysts

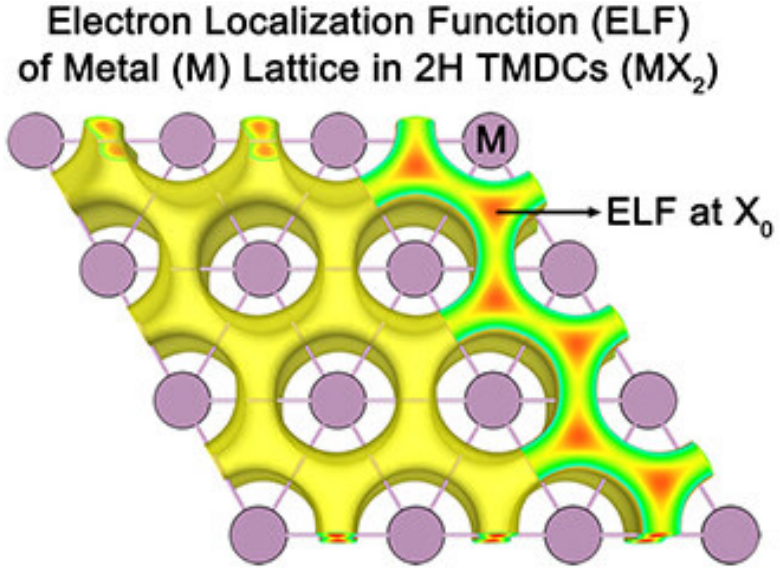


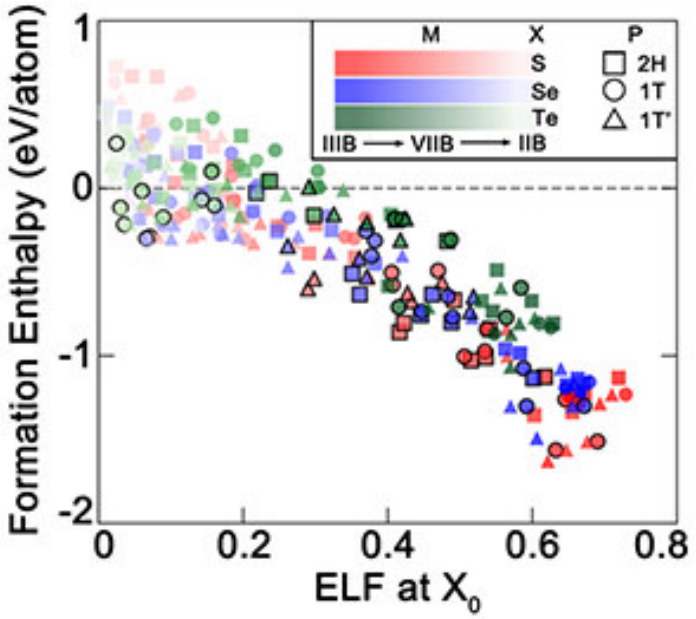
- Electride: material where trapped electrons act as anions
- Electrides enable strong H_2 activation
- Major limitation: Ca_2N is highly sensitive to air and water
- Solution: Can we engineer electride states in stable monolayer compounds like TMDCs?





Chemical-template impacts TMDC stability



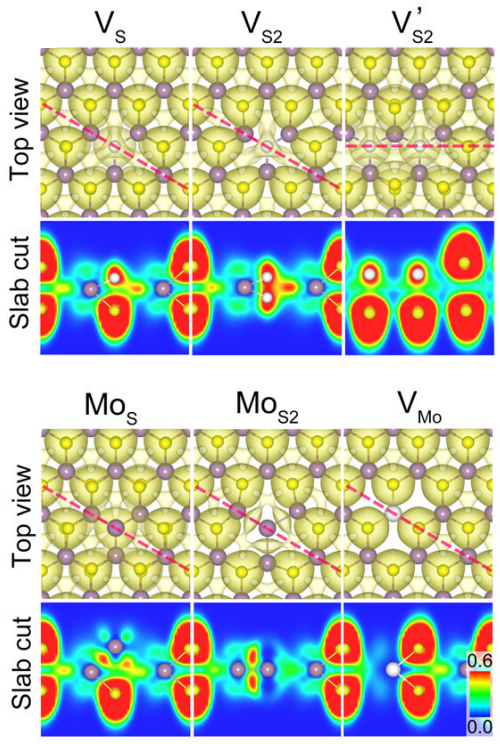


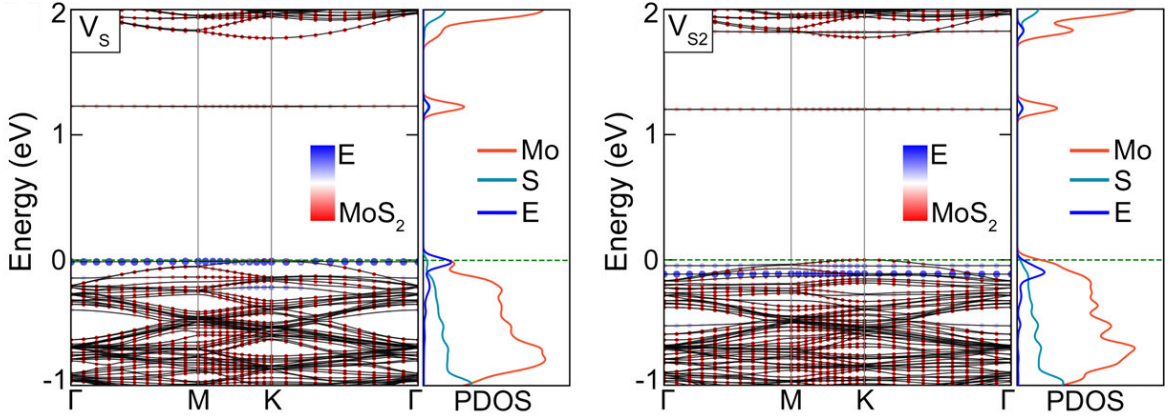
Metal sublattice retains distinct interstitial localization across 2H, 1T, and 1T’ phases

Larger ELF value at X_0 correlates with lower formation enthalpy

Tunable electrides through defect engineering

Vacancy engineering creates strong electrides



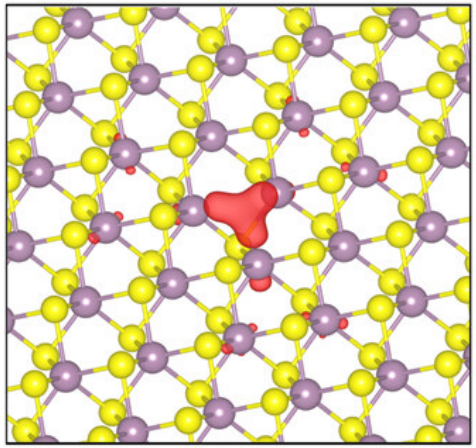


Only chalcogen vacancies generated interstitial electrons at the defect sites

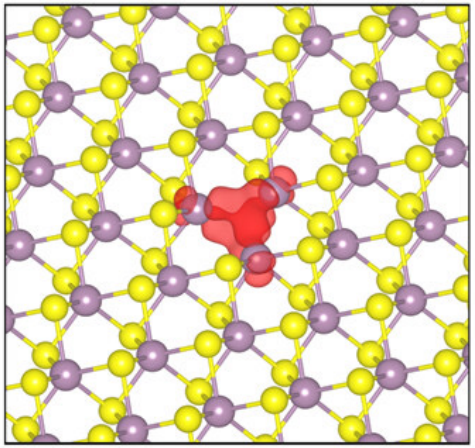
Quasi-atom states in V_S and $V_{S2\text{-}MoS2}$ contribute strongly near the top of the valence bands

Localized quasi-atom states at S vacancies

HOMO at Γ of V_S MoS_2



HOMO-1 at Γ of V_{S2} MoS_2



Frontier states in both V_S and V_{S2} are spatially localized at the vacancy-centered quasi-atom site

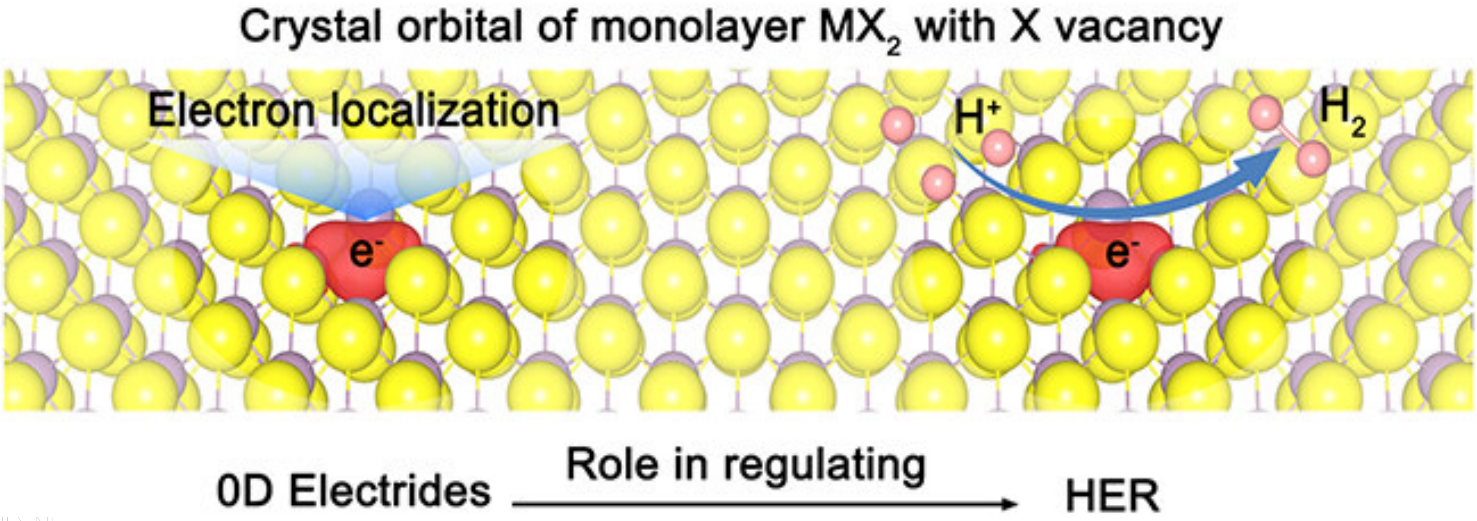
Localized electride states carry substantial charge at vacancy cite

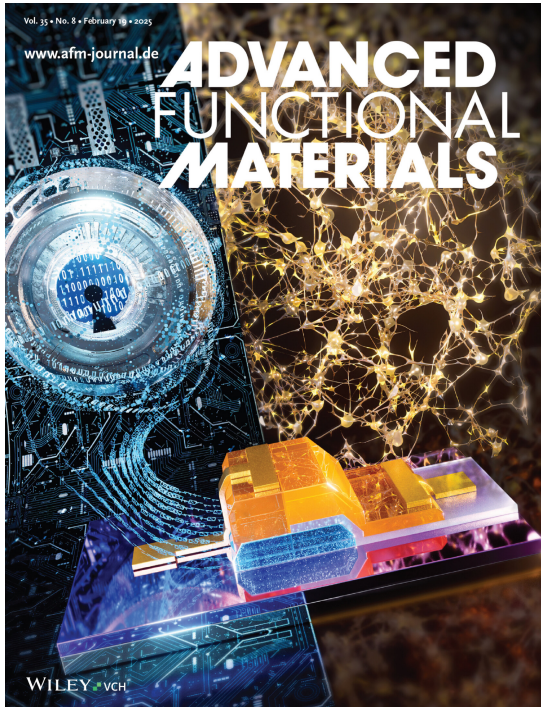
Vacancy-induced electrides regulate H adsorption

Chalcogen vacancies trap excess electrons

Localized electrons behave as 0D electride states

Electride states tune H binding toward optimal HER





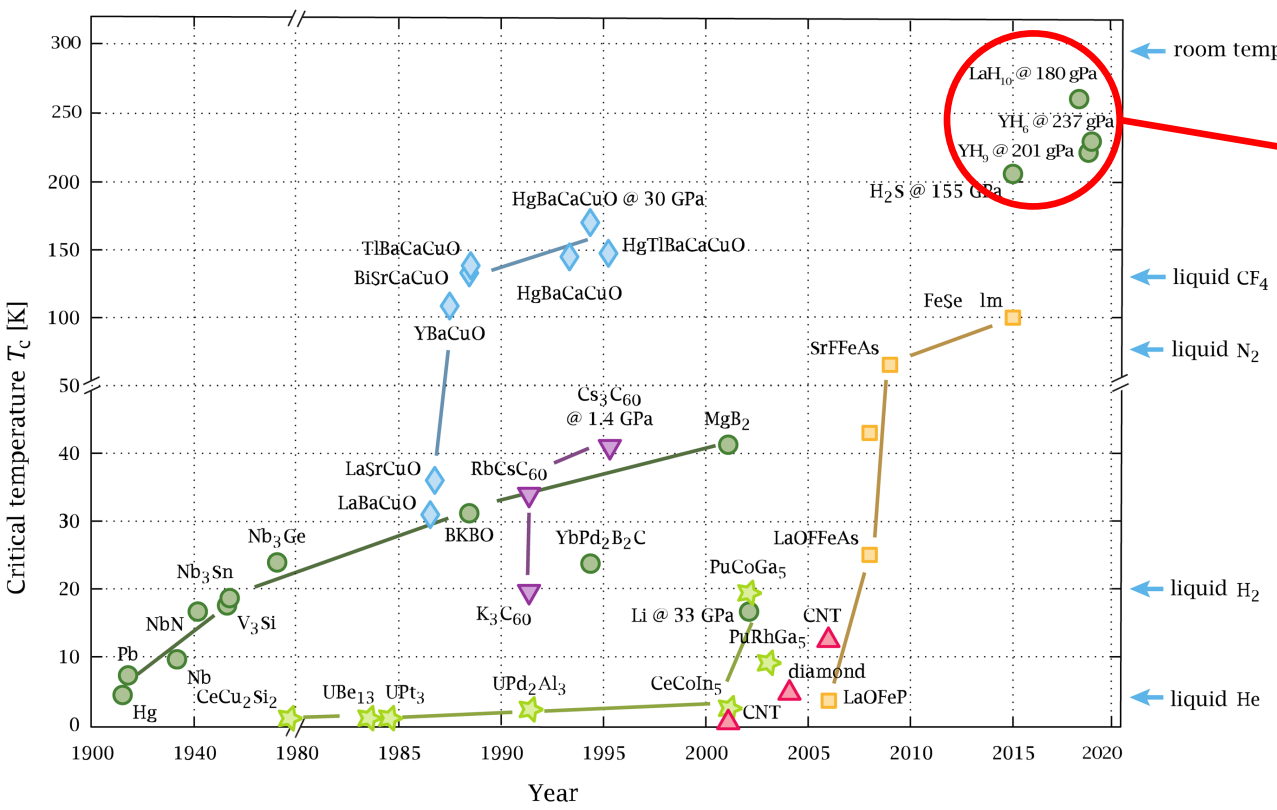
Topic 2: High Pressure Superconductivity

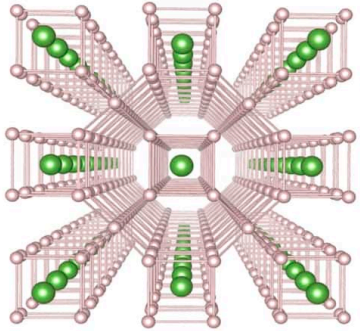
Zhao, Ellis, et al. Adv. Funct. Mater. 2025, 35, 2415910.

Ellis, et al. *J. Phys. Chem. Lett.*, submitted.

Superconductivity in molecular hydrides

Superhydrides hold record superconducting T_c





High-pressure chemistry unveils unique material structures and unusual stoichiometries leading to discoveries of high T_c superhydrides

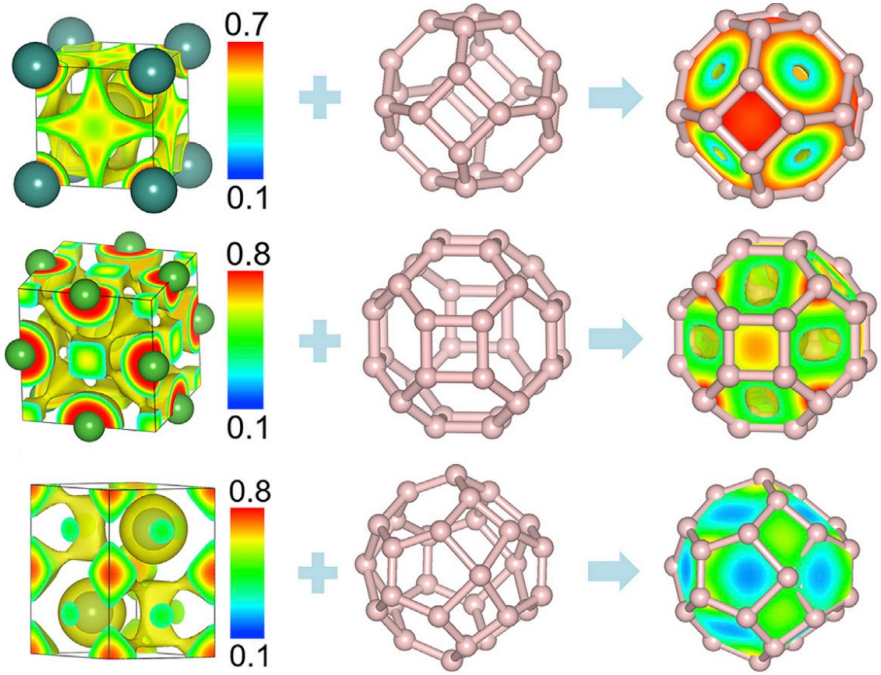
The ELF predicts superhydride formation

Electron localization function (ELF) compares local kinetic energy density to a homogeneous gas

$$\mathrm{ELF}(\mathbf{r}) = \frac{1}{1 + \chi_\sigma^2(\mathbf{r})}$$

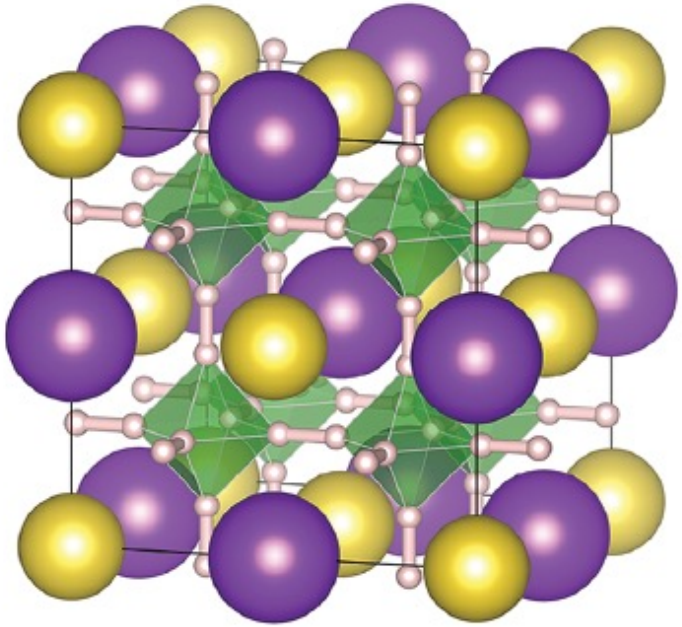
High value = localized electrons
Low value = delocalization

Hydrogen atoms assemble in interstitial sites with high electron localization

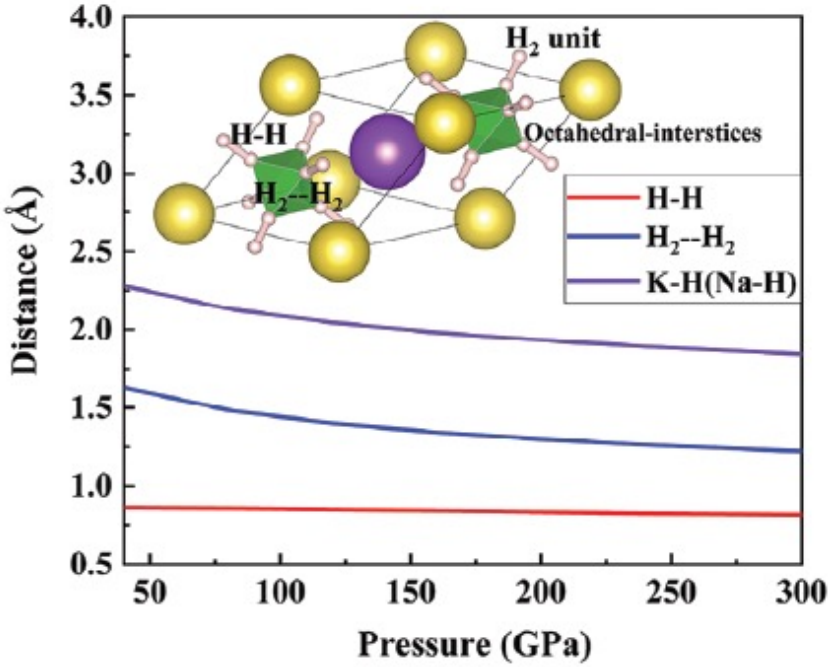


Sun & Miao, *Chem* **2023**, *9*, 443–459.

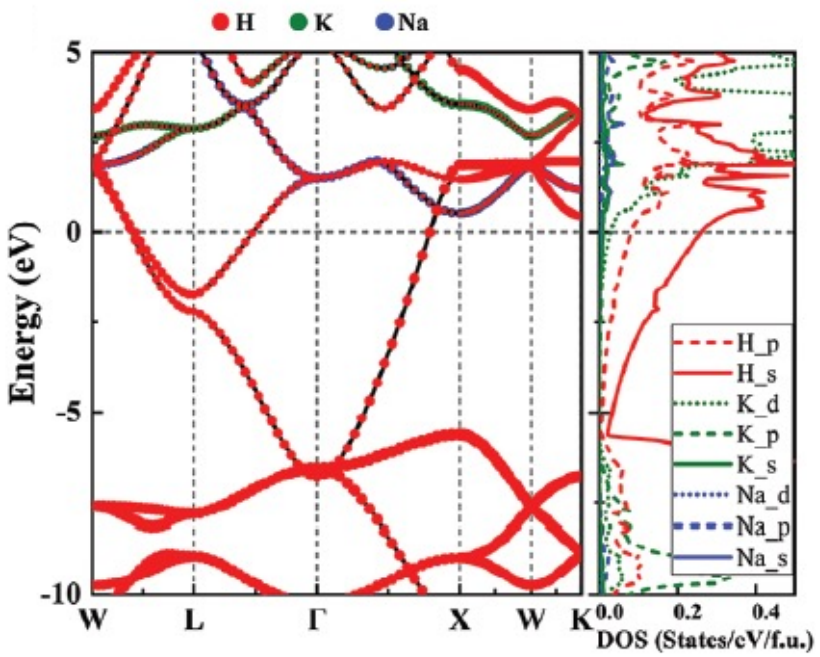
Structure and bonding in $NaKH_{12}$

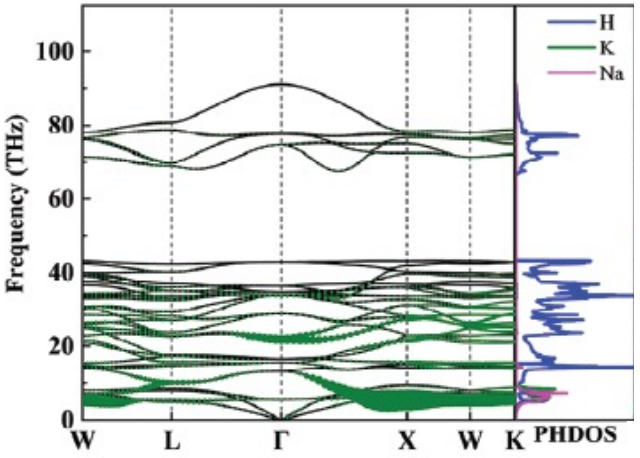


Electron localization at the interstitial sites surrounded by 6 H_2 molecules



Origin of high T_c Superconductivity

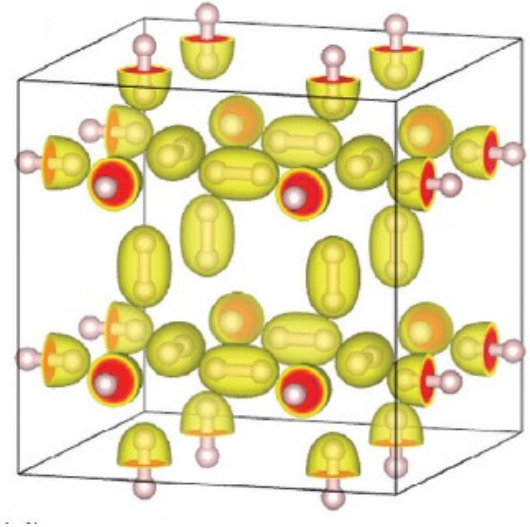




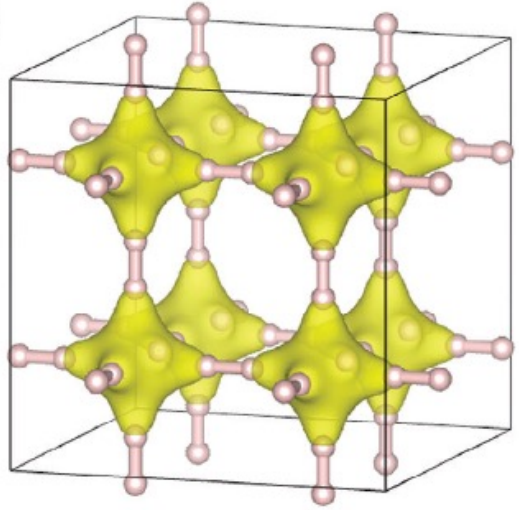
Predicted T_c of 245K at 60 GPa

- Higher density of H states near E_F
- strong electron-phonon coupling

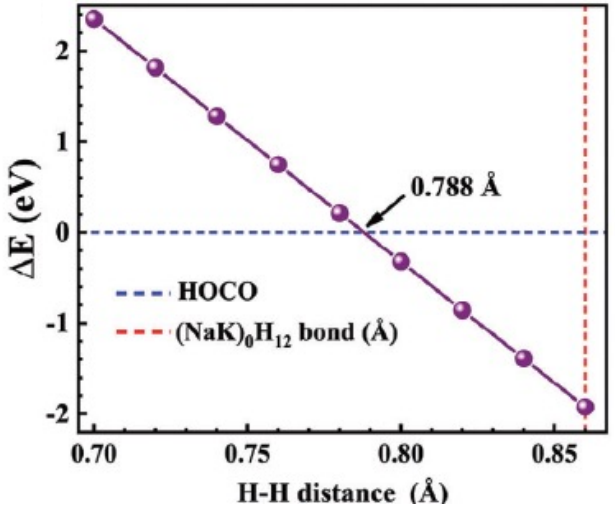
Elongated H_2 bonds create interstitial electronic state



Conventional H_2 bonding state

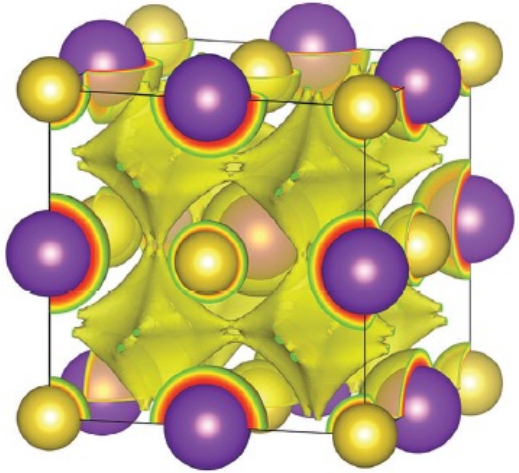


Interstitial orbital between H_2 units

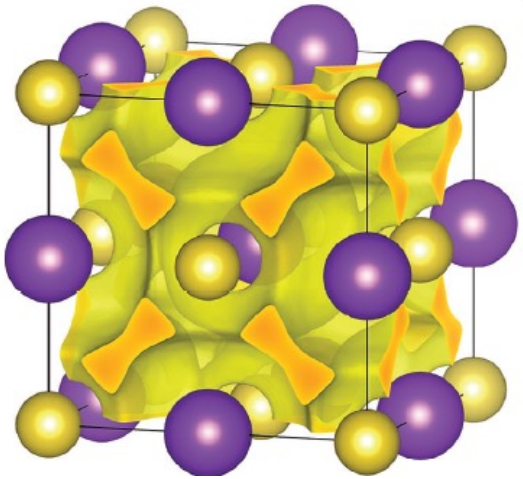


Interstitial state becomes occupied above ~0.79 Å

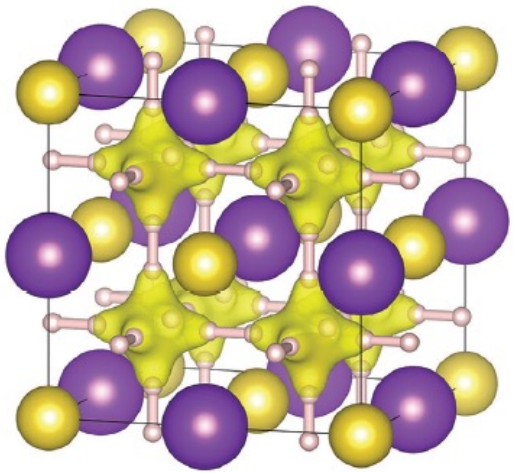
NaK metal sublattice acts as a chemical template



ELF of NaK metal sublattice shows strong interstitial localization



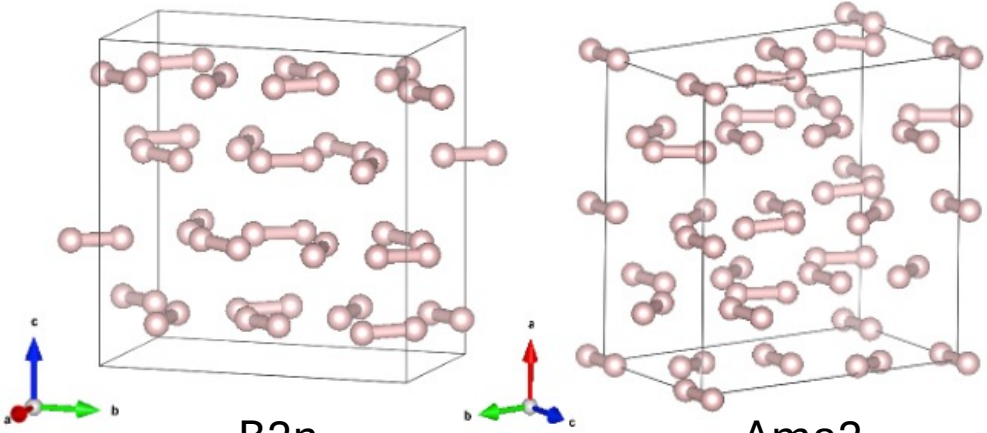
Occupied orbital of sublattice is centered at the same interstitial sites



Orbital persists in full $NaKH_{12}$ linking neighboring H_2 molecules

Metallization in high pressure H_2 molecular crystals

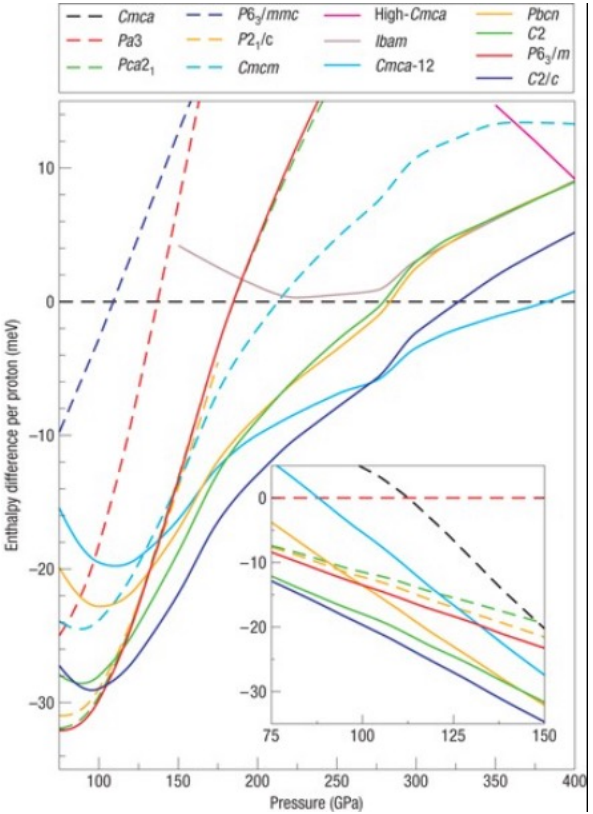
Competing phases for dense H_2 under pressure



B2n Ama2

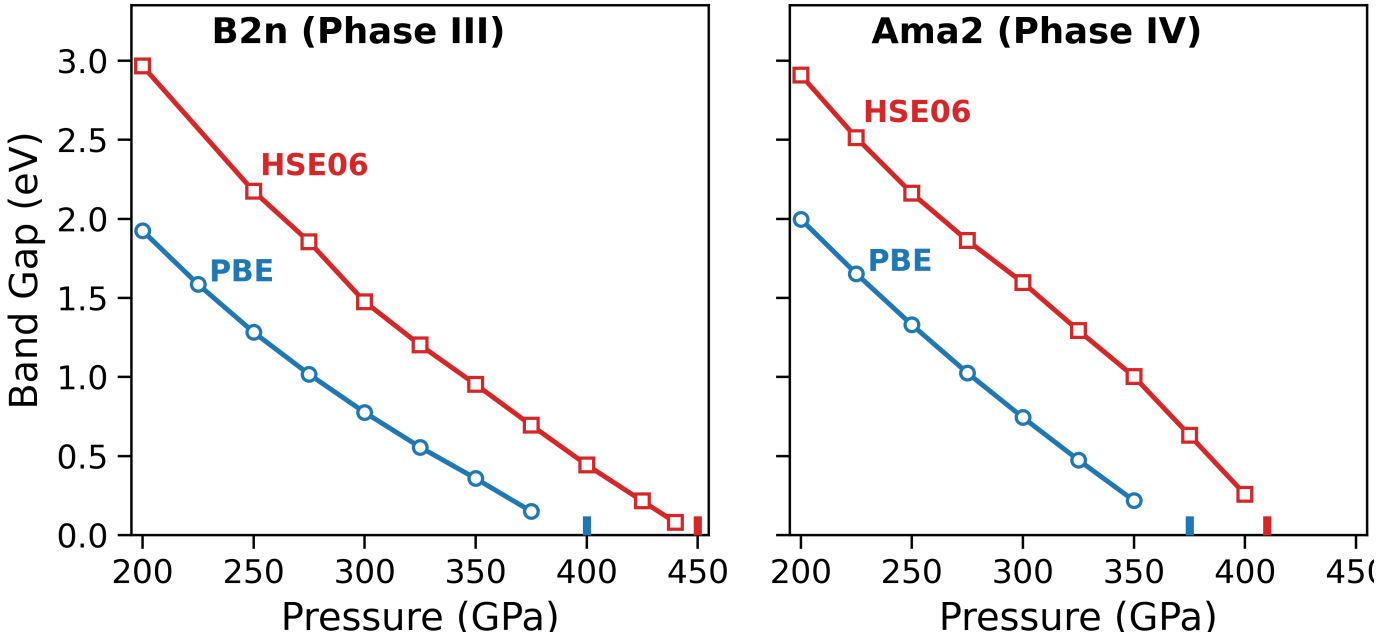
Many proposed phases are very close in enthalpy, assigning a “correct” structure is difficult

Representative structures chosen due to agreement with experimental results



Pickard & Needs, **Nat. Phys. 2007**, 3, 473–476.

Pressure-driven gap closure in molecular hydrogen

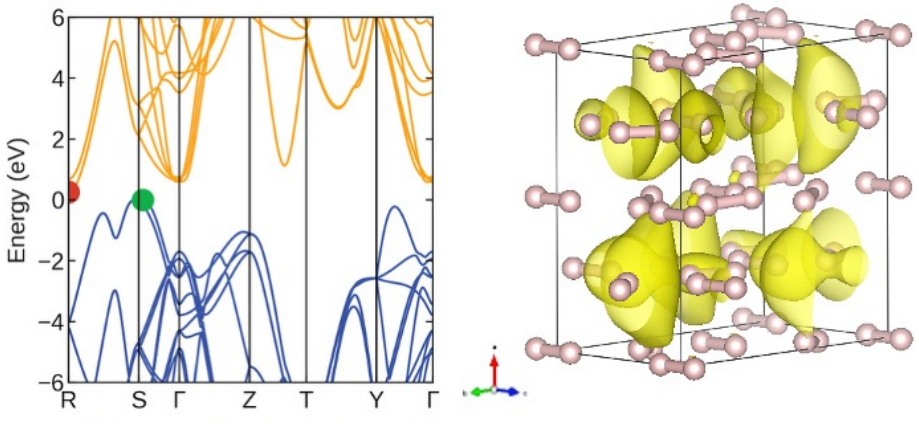


PBE is known to underestimate energy gaps

HSE06 functional is more consistent with experimental metallic transition near 425-500GPa

PBE – standard DFT functional that approximates electron interactions
HSE06 – more accurate, but more expensive functional (exact exchange)

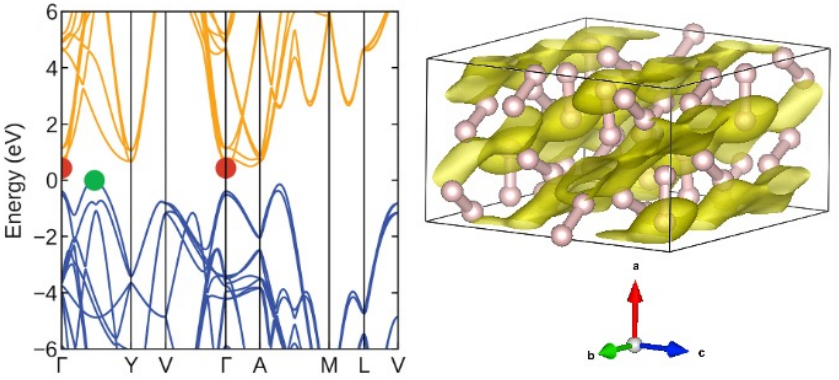
Interstitial band-edge states drive metallization



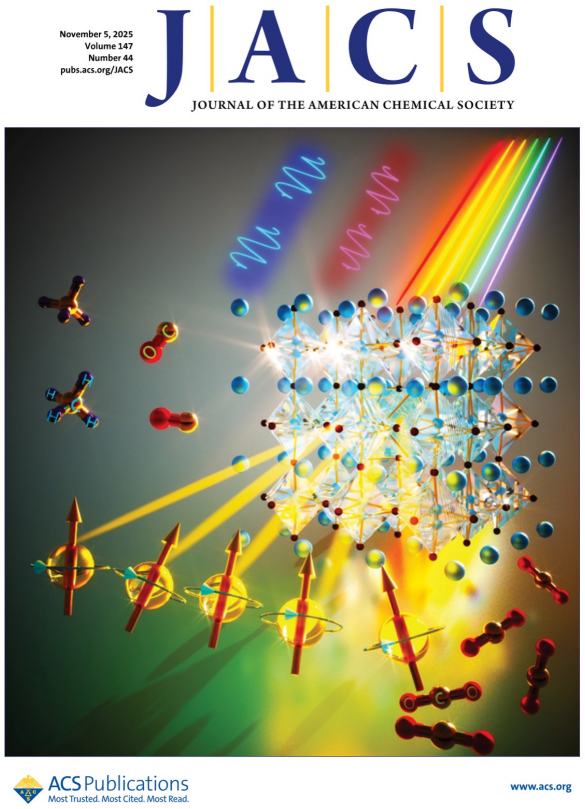
B2n 425 GPa

In B2n, conduction band minimum forms lobes that bridge neighboring molecules through dense 3D lattice

In Ama2, layered geometry channels charge into sheet-like pathways between molecular layers



Ama2 400 GPa



Topic 3: Machine learning assisted materials discovery

Sun, Ellis, et al. JACS 2025, 147, 40407–40419.

Ellis, Miao. *PRX Intelligence*, submitted.

Chemical-template guided superhydride discovery

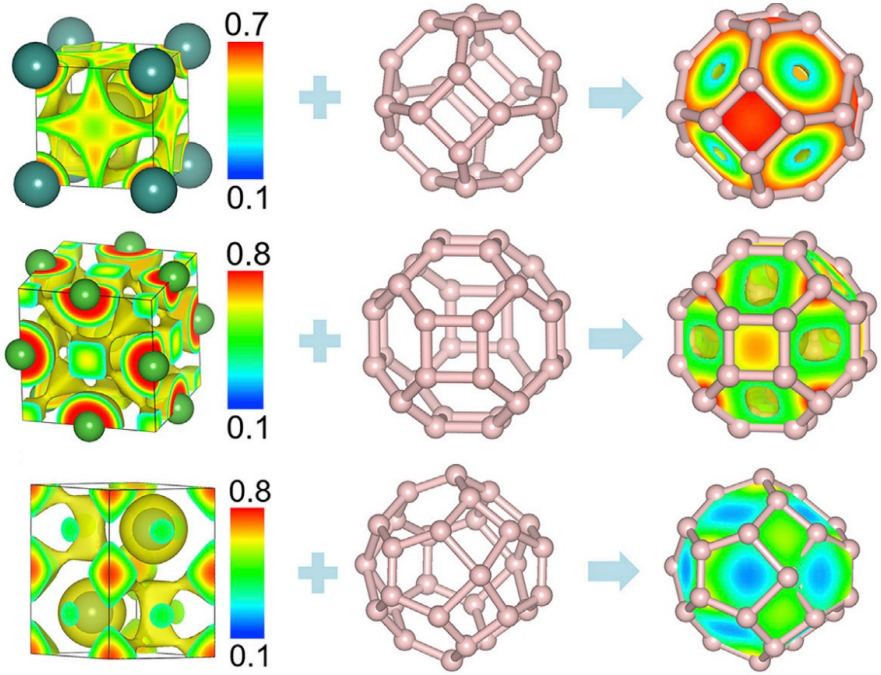
The ELF predicts superhydride formation

Electron localization function (ELF) compares local kinetic energy density to a homogeneous gas

$$\mathrm{ELF}(\mathbf{r}) = \frac{1}{1 + \chi_\sigma^2(\mathbf{r})}$$

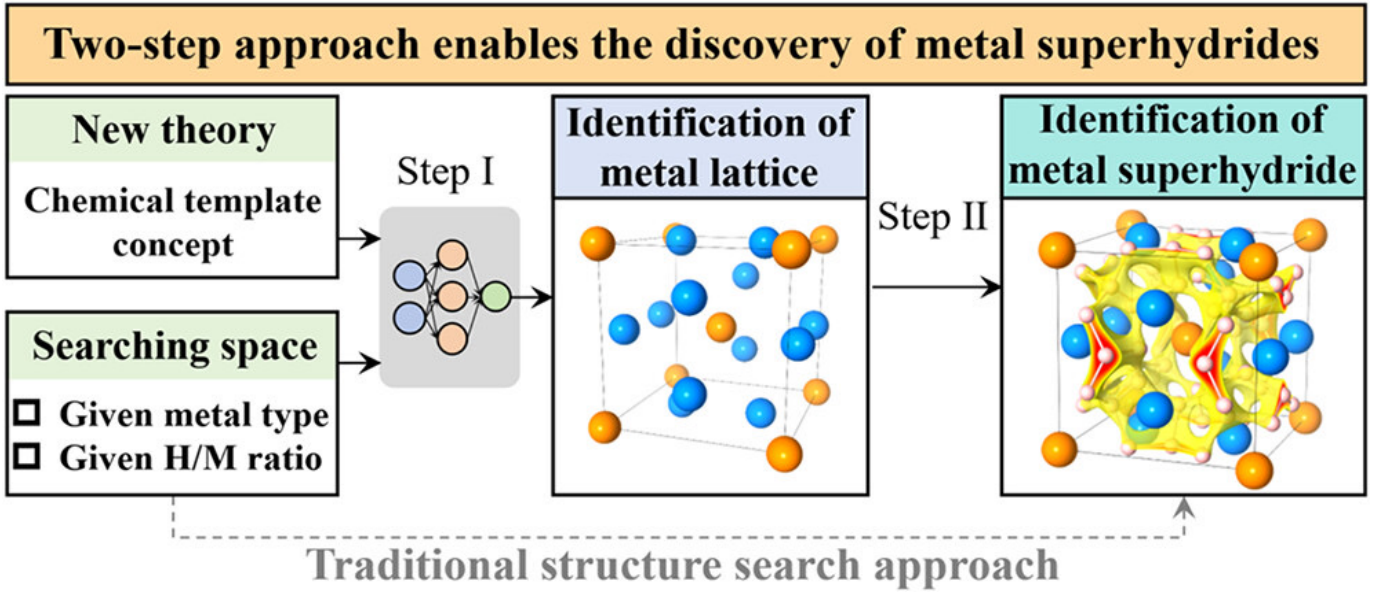
High value = localized electrons
Low value = delocalization

Hydrogen atoms assemble in interstitial sites with high electron localization



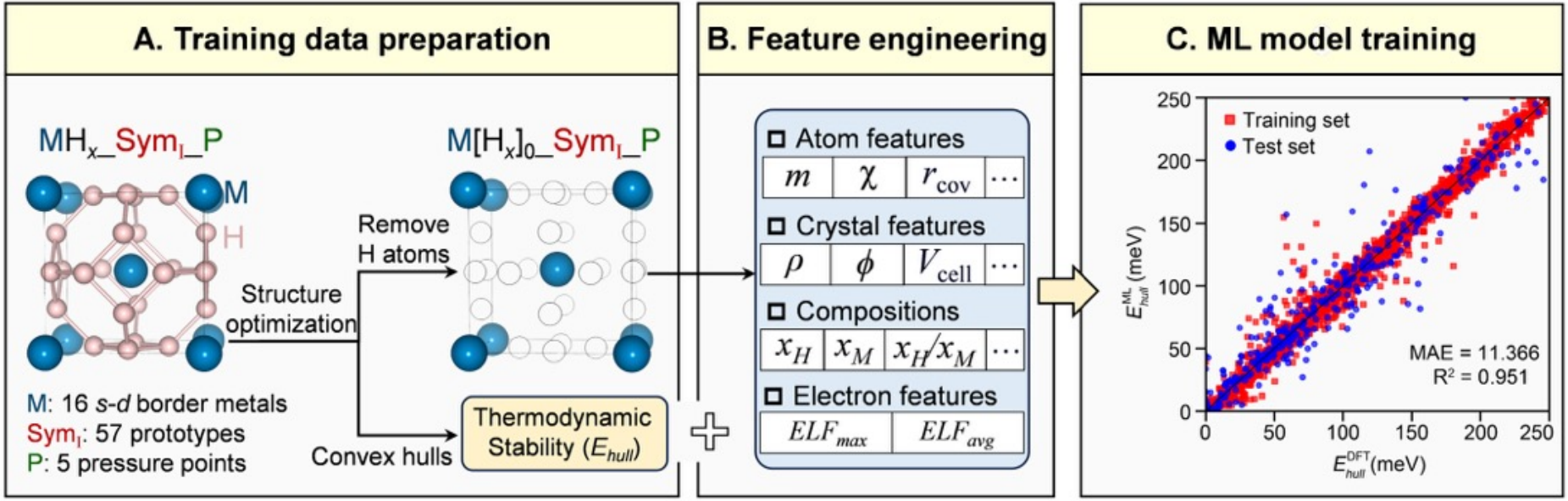
Sun & Miao, *Chem* **2023**, *9*, 443–459.

Chemical template-based discovery workflow



Using the chemical template theory and corresponding 2-step approach greatly reduces the search space and makes superhydride discovery more efficient.

Predicting formation energy from metal templates

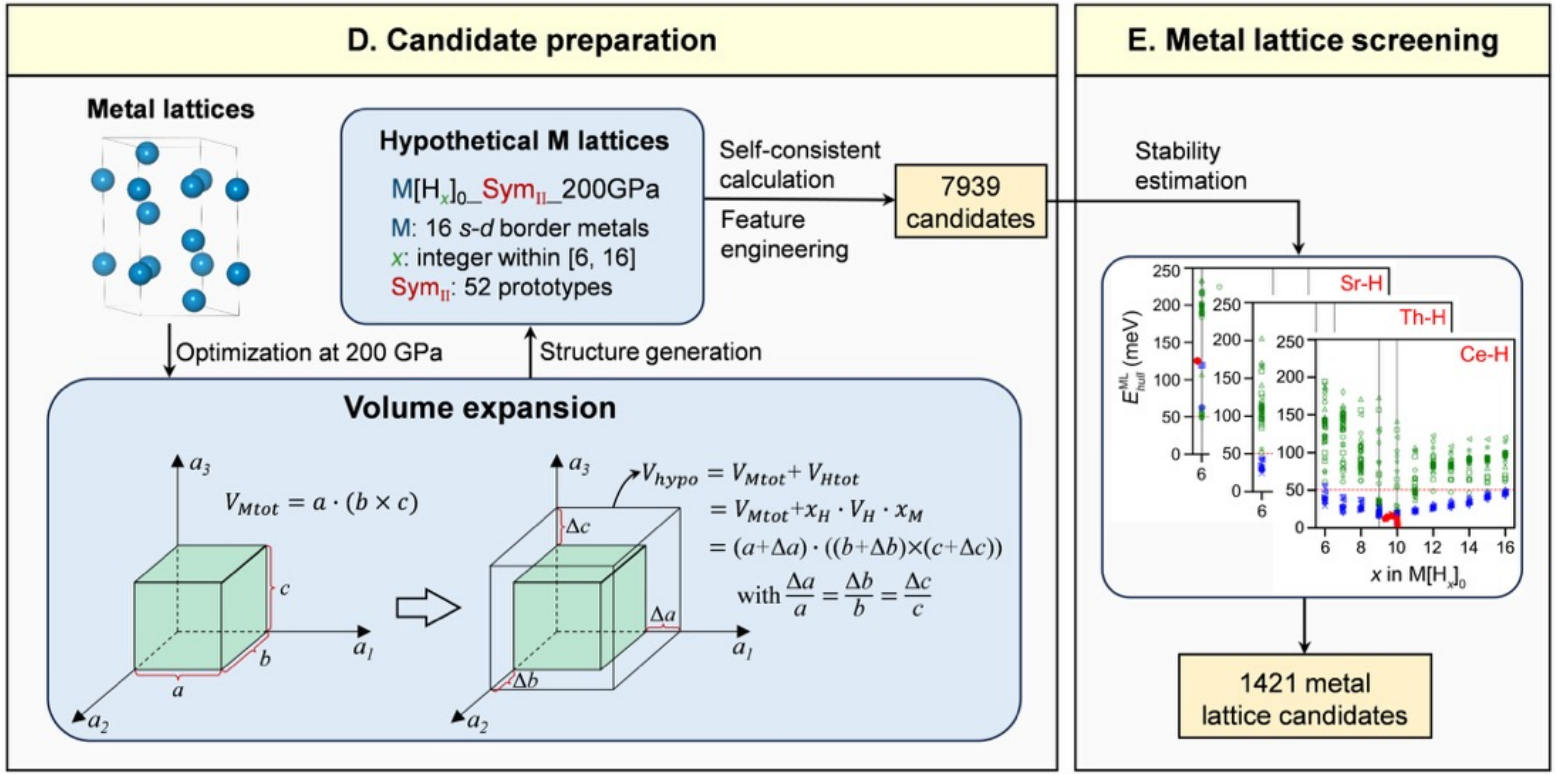


Step 1. training set preparation (DFT)

Step 2. feature engineering

Step 3. ML model training

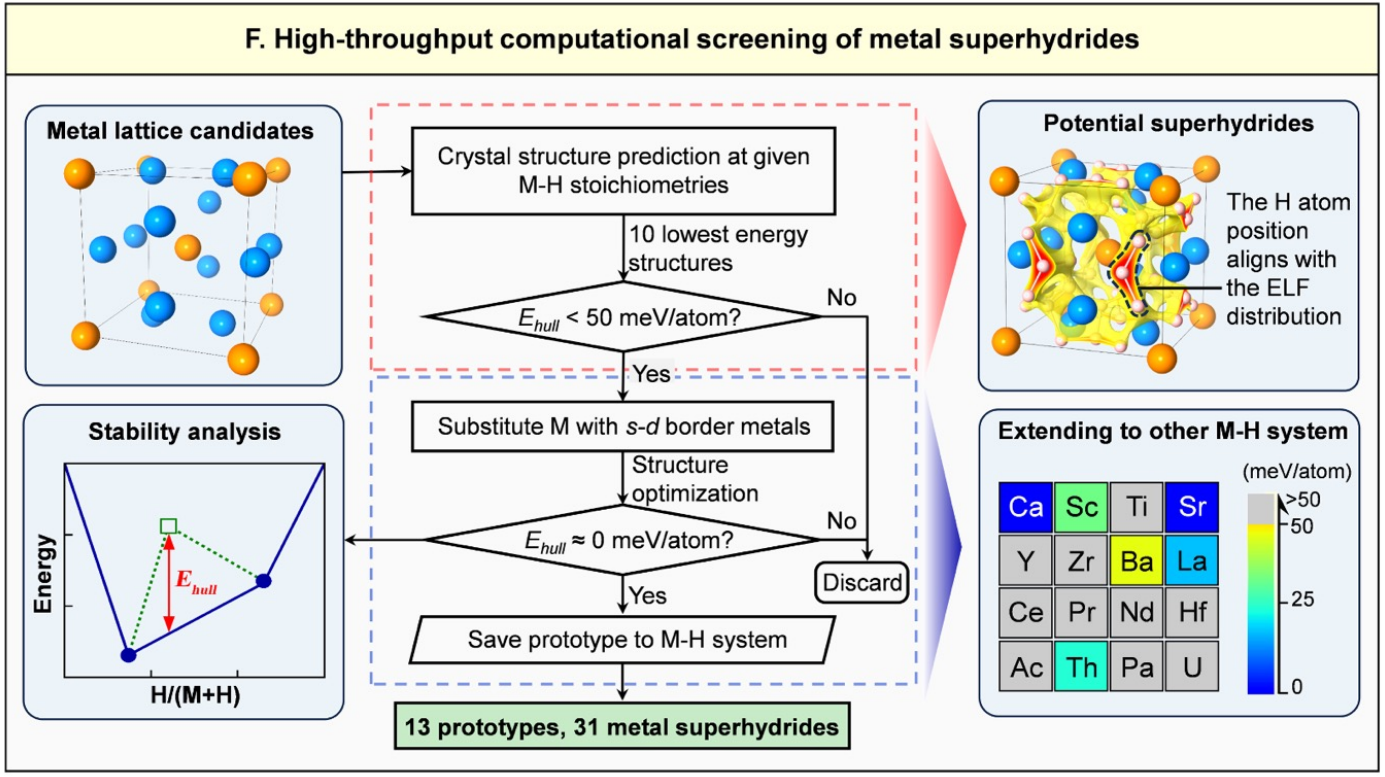
Screening metal lattices for promising candidates



Step 4. metal lattice candidates

Step 5. metal lattice screening

Predicting new superhydride structures



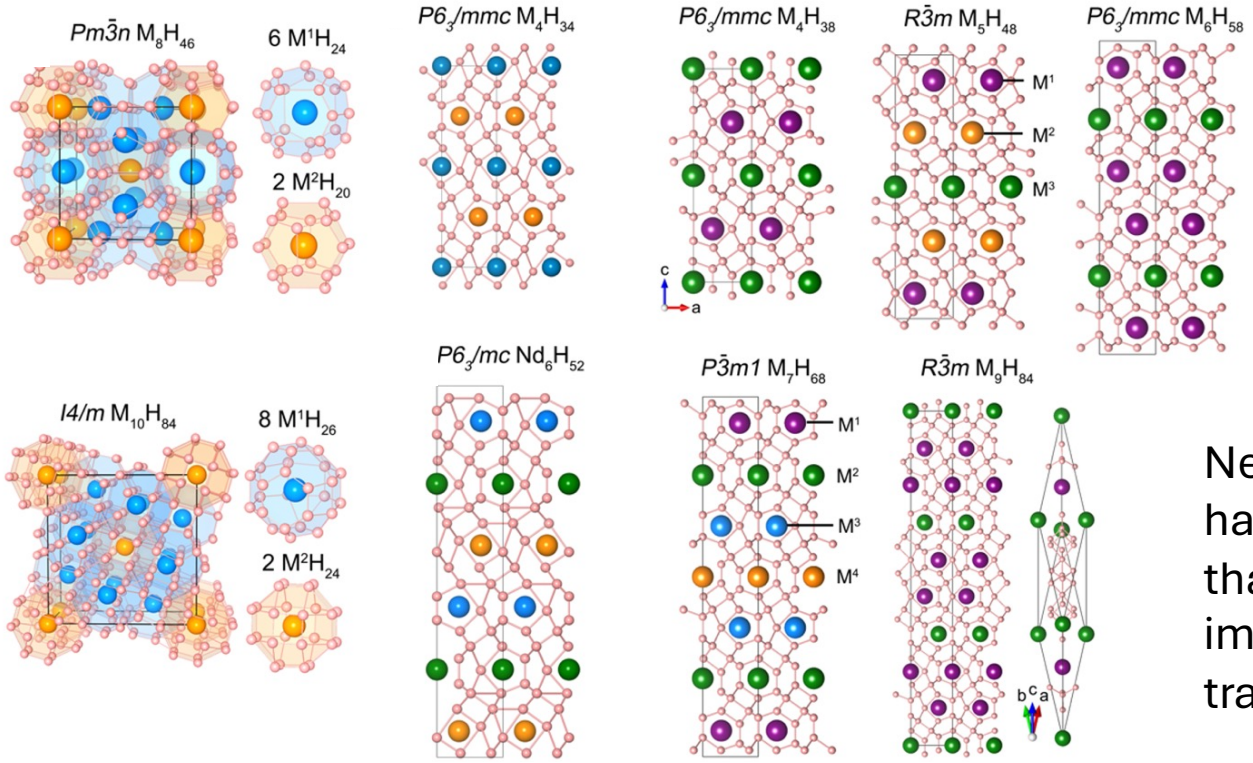
Step 6. Search MH_n

Workflow based on the chemical template

New structures will be transferred to other MH_n and SCC

Allows us to search complex MH_n with very large cells

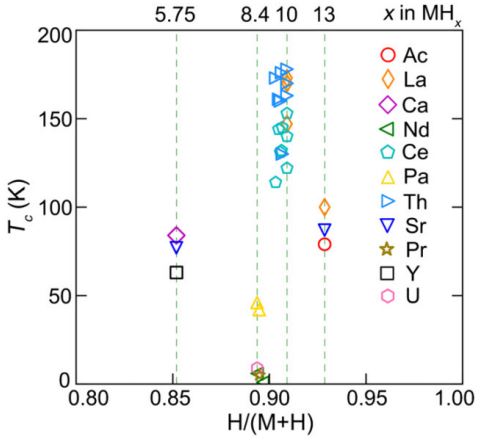
Newly identified metal superhydrides

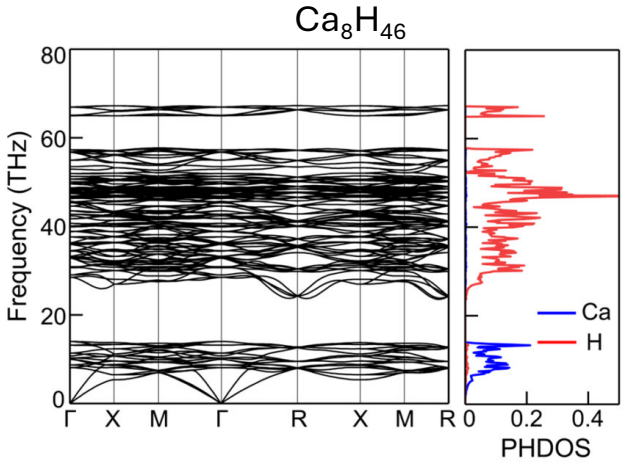


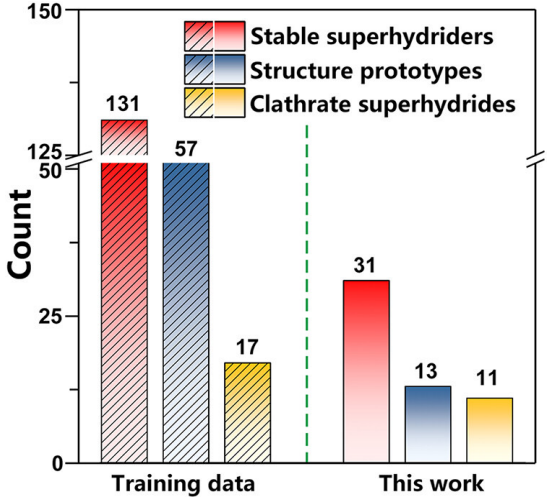
Many new phases appear at non-integer H/M ratios

Newly identified phases have large primitive cells that would've been nearly impossible to search with traditional methods

Superconductivity of novel superhydrides







Estimated T_c values suggest that many of the newly discovered superhydrides are strong high T_c candidates, with 19 above 100 K

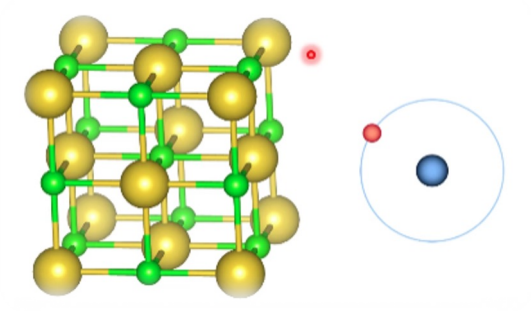
Superconductivity is driven mainly by H-dominated high-frequency phonons

Workflow identified 31 new stable metal superhydrides and 13 new structural prototypes

Seitz-invariant ELF prediction via 3D CNN

The ELF can be expensive to calculate

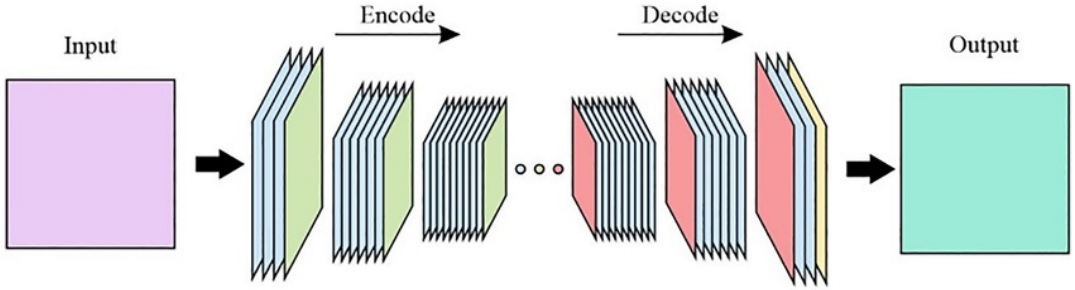
First Principles (QM)



$$-\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi = E\Psi$$

The time required to run a quantum mechanics-based calculation scales cubically with electron count (N)

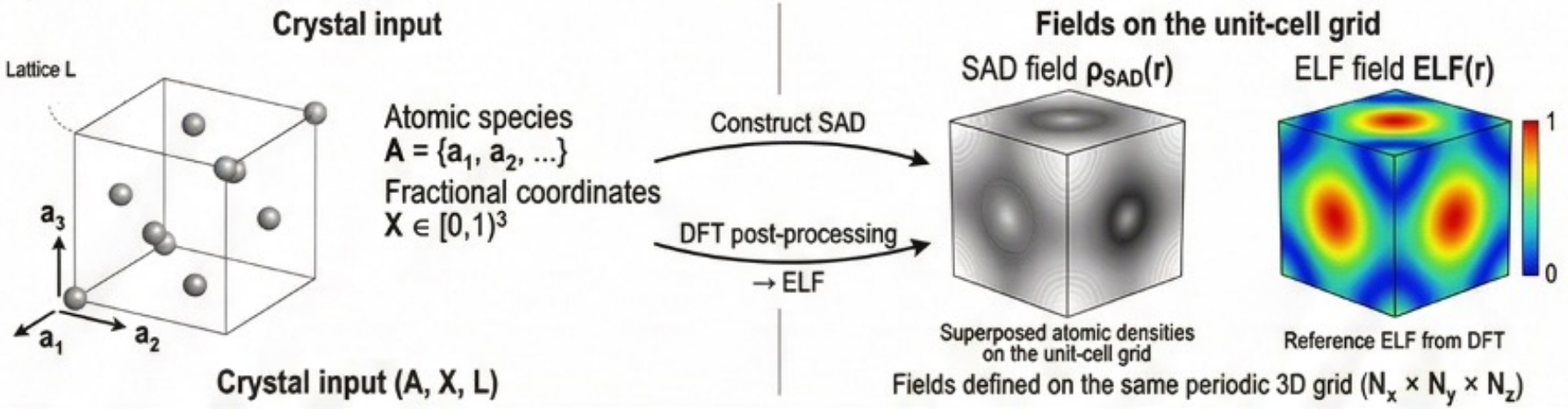
Machine Learning (ML)



Solution: We can **treat the ELF as a 3D image** and train a model to predict it directly

Problem: How do we represent our structure?

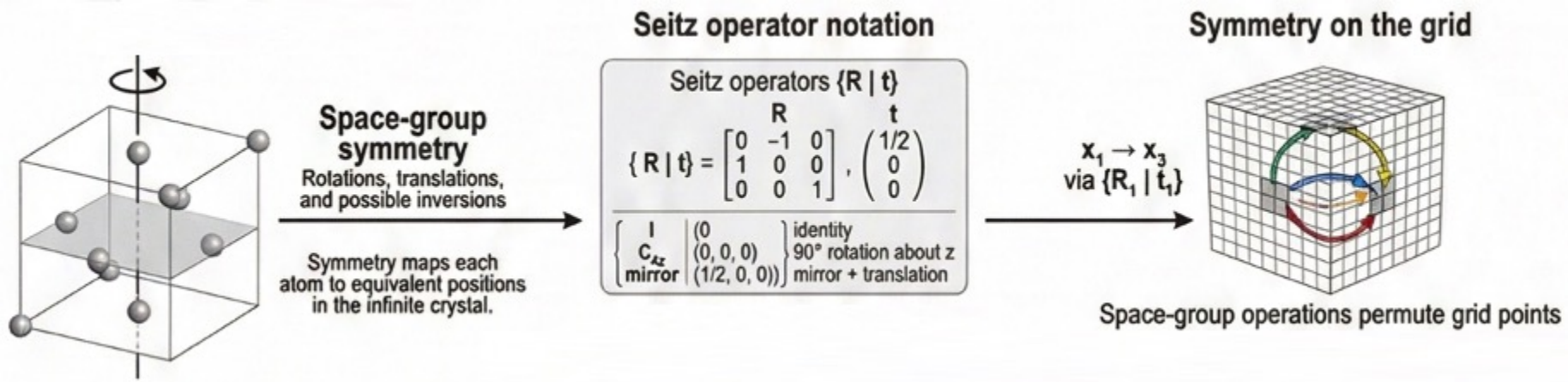
Crystal structure representation



SAD is the electron density obtained by freezing atomic charge density at each atom center without allowing their electron clouds to hybridize or relax into bonded states

Since the SAD and ELF have the same grid size, the **problem is now an image transformation**

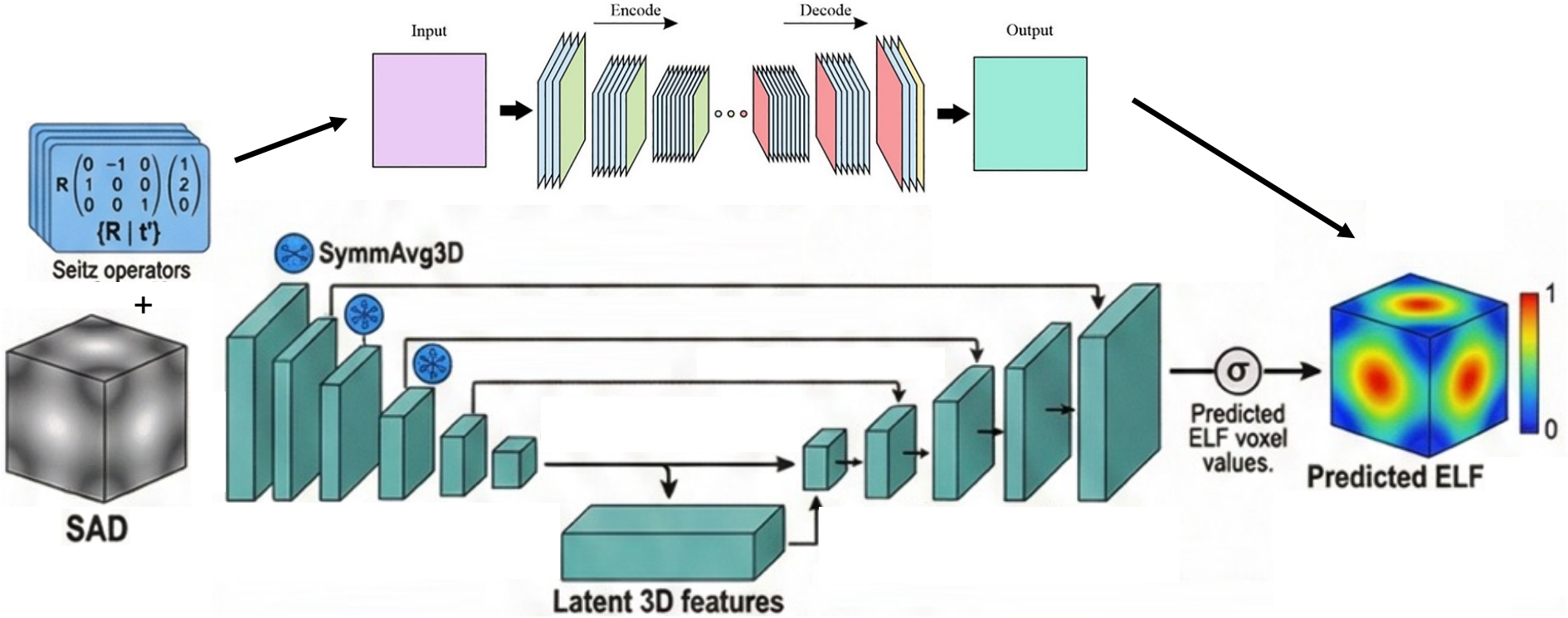
Making the model Seitz-Invariant



Seitz symmetry – Standard notation system for describing the symmetry operations of crystal structures in 3D space groups

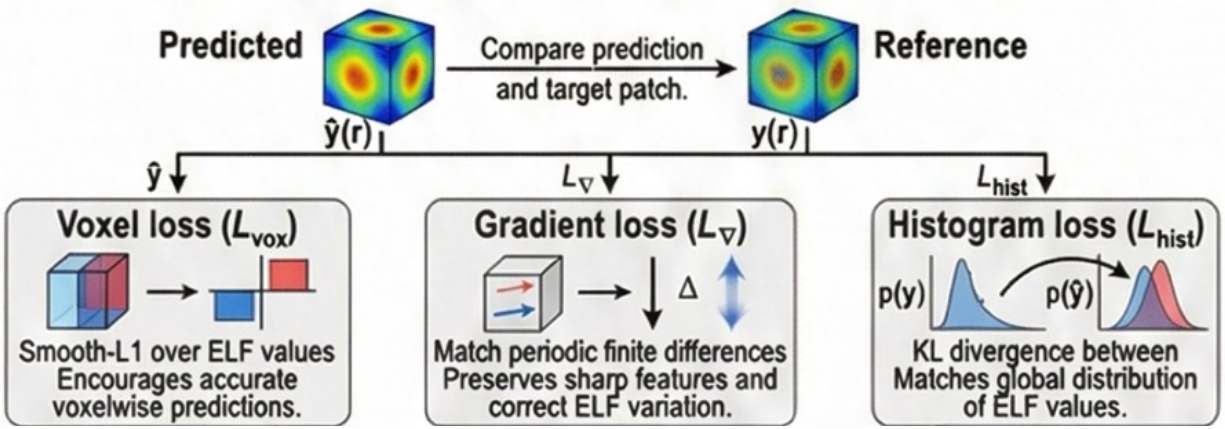
Invariance – Property that guarantees the model makes the same prediction even if the crystal structure is rotated or translated according to its symmetry group

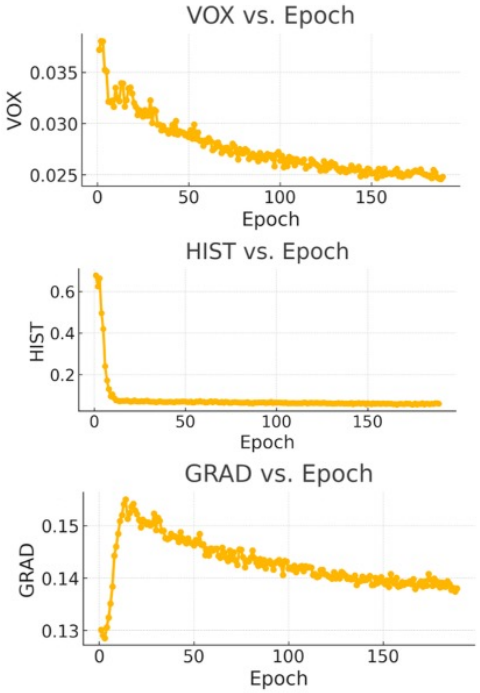
Architecture for a symmetry-aware image transform



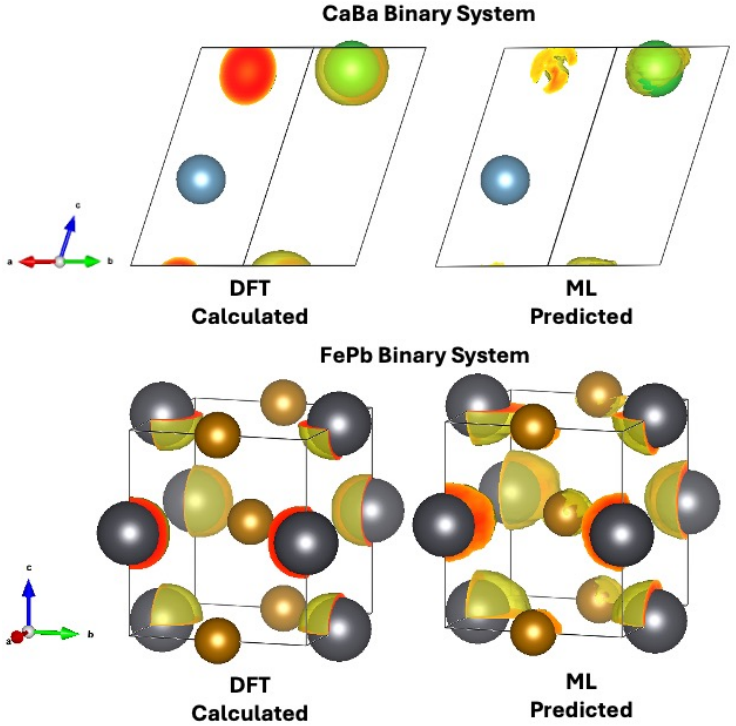
Physics-aware loss function

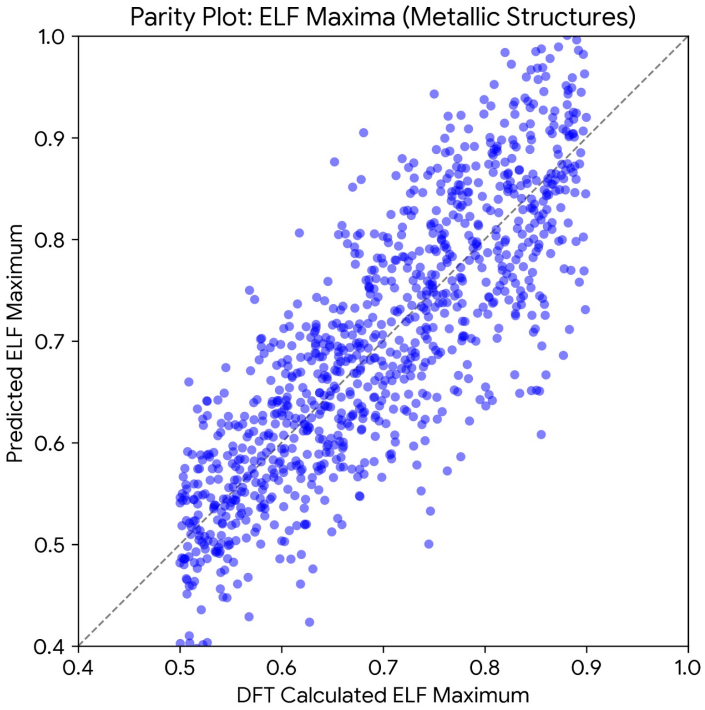
Kendall Loss: An adaptive weighting scheme that automatically learns how to balance the voxel, gradient, and histogram objectives during training, eliminating the need to manually tune loss weights.





Electron localization function prediction



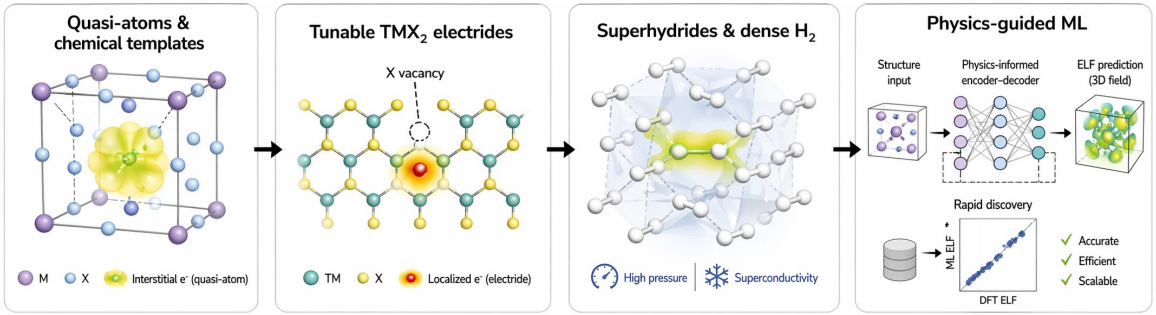


Model tested on subset of metal lattices

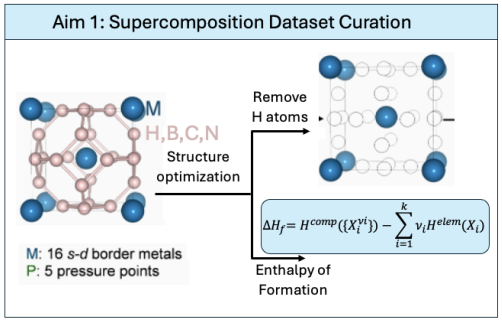
Maxima most relevant to prediction of superhydride structures

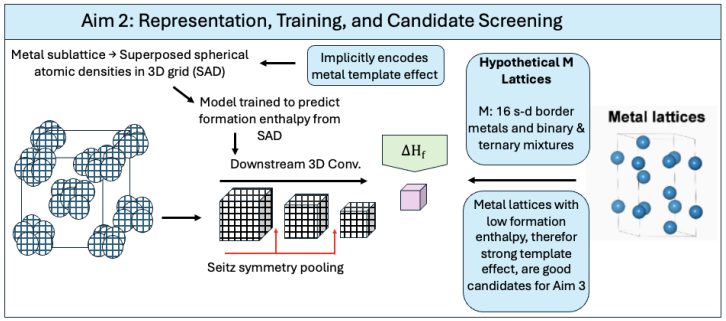
Conclusions

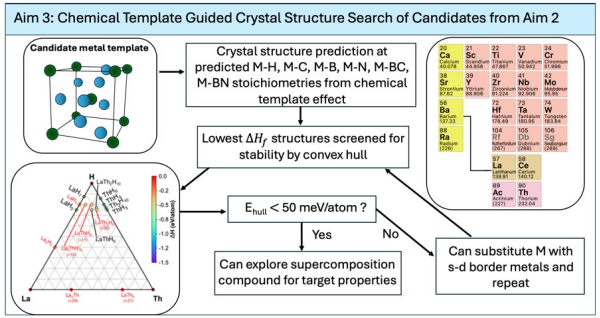
- Quasi-atom states and chemical templates provide a framework for understanding the formation mechanism and structure preference of many compounds
- In TMX_2 monolayers, chemical templates and vacancies create tunable electride behavior and influence H adsorption
- Interstitial states help explain strong superconductivity in molecular hydrogen and pressure driven metallization of H_2
- Template-guided machine learning accelerates discovery, enabling efficient screening of superhydrides and rapid ELF prediction



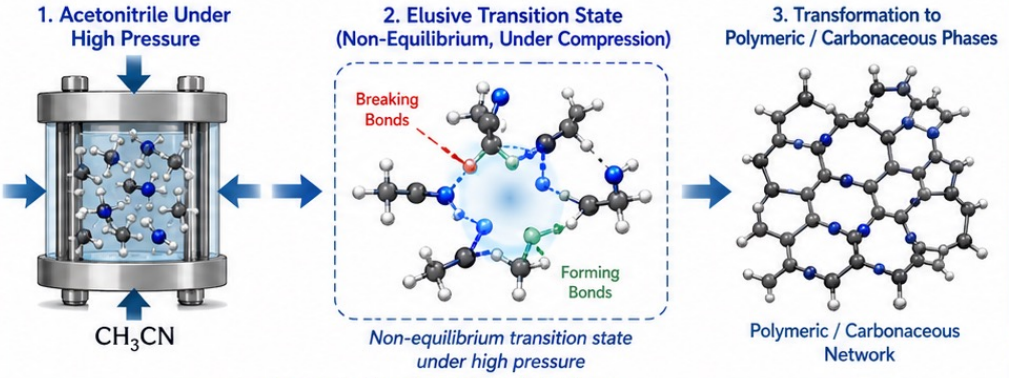
Future work


Aim 1: Supercomposition Dataset Curation
M
H,B,C,N
Structure optimization
Remove H atoms
$\Delta H_f = H^{comp}(\{X_i^{vi}\}) - \sum_{i=1}^{k} v_i H^{elem}(X_i)$
M: 16 s-d border metals
P: 5 pressure points
Enthalpy of Formation


Aim 2: Representation, Training, and Candidate Screening
Metal sublattice → Superposed spherical atomic densities in 3D grid (SAD)
Implicitly encodes metal template effect
Model trained to predict formation enthalpy from SAD
Downstream 3D Conv.
Seitz symmetry pooling
ΔH_f
Hypothetical M Lattices
M: 16 s-d border metals and binary & ternary mixtures
Metal lattices with low formation enthalpy, therefor strong template effect, are good candidates for Aim 3
Metal lattices


Aim 3: Chemical Template Guided Crystal Structure Search of Candidates from Aim 2
Candidate metal template
Crystal structure prediction at predicted M-H, M-C, M-B, M-N, M-BC, M-BN stoichiometries from chemical template effect
Lowest ΔH_f structures screened for stability by convex hull
E_{hull} < 50 meV/atom ?
Yes
No
Can explore supercomposition compound for target properties
Can substitute M with s-d border metals and repeat


OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE
FOR U.S. DEPARTMENT OF ENERGY


1. Acetonitrile Under High Pressure
CH_3CN
2. Elusive Transition State (Non-Equilibrium, Under Compression)
Breaking Bonds
Forming Bonds
Non-equilibrium transition state under high pressure
3. Transformation to Polymeric / Carbonaceous Phases
Polymeric / Carbonaceous Network

Acknowledgements

Dr. Miao

Dr. Teprovich

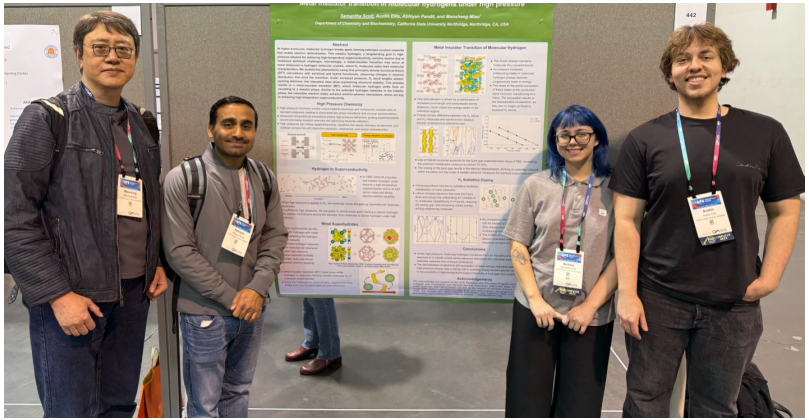
Dr. Lau

Dr. Sun

Dr. Pandit

Samantha Ellis

Friends at CSUN



Department of Chemistry and Biochemistry

College of Science and Math

U.S. Department of War

Department of War HBCU/MI

NSF MAJOR RESEARCH INSTRUMENTATION

California State University, Northridge 1958