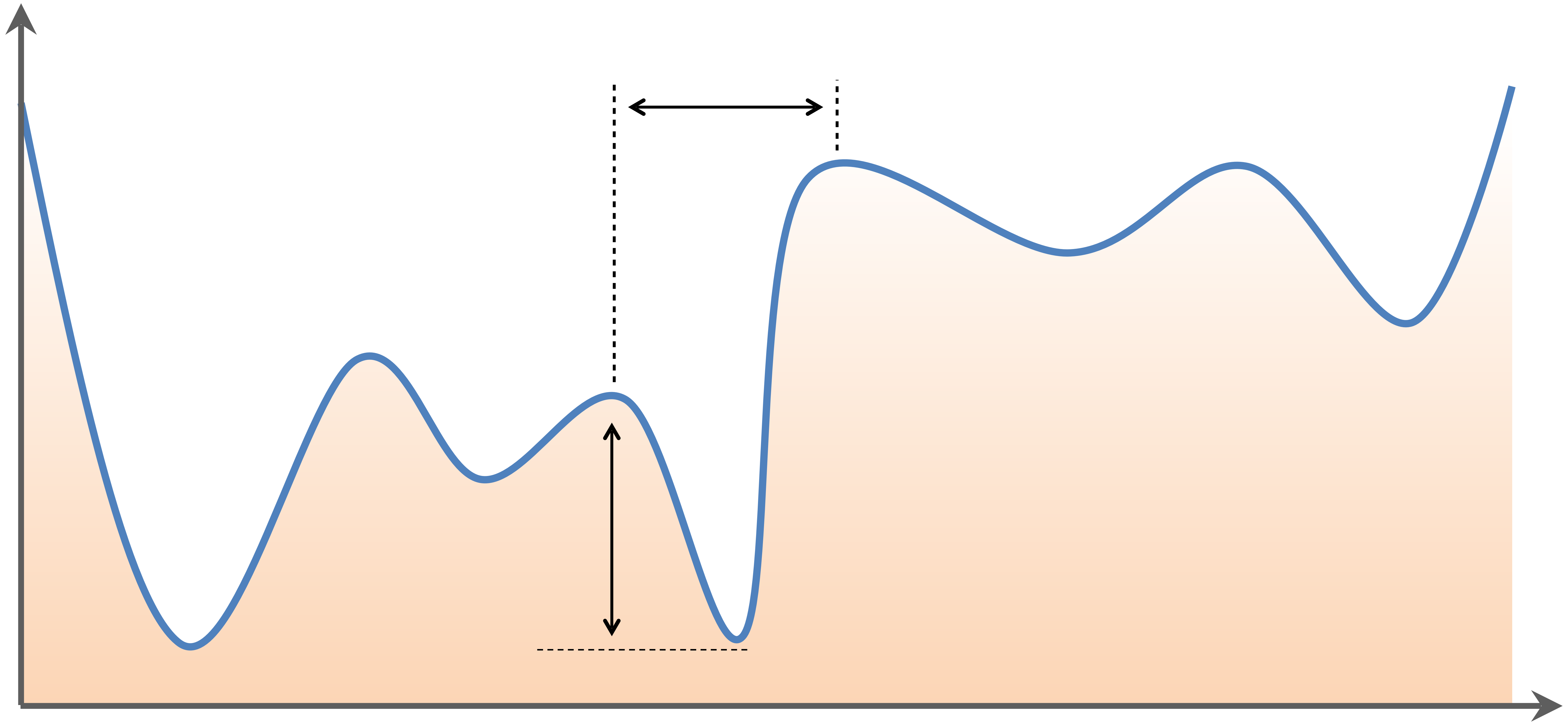
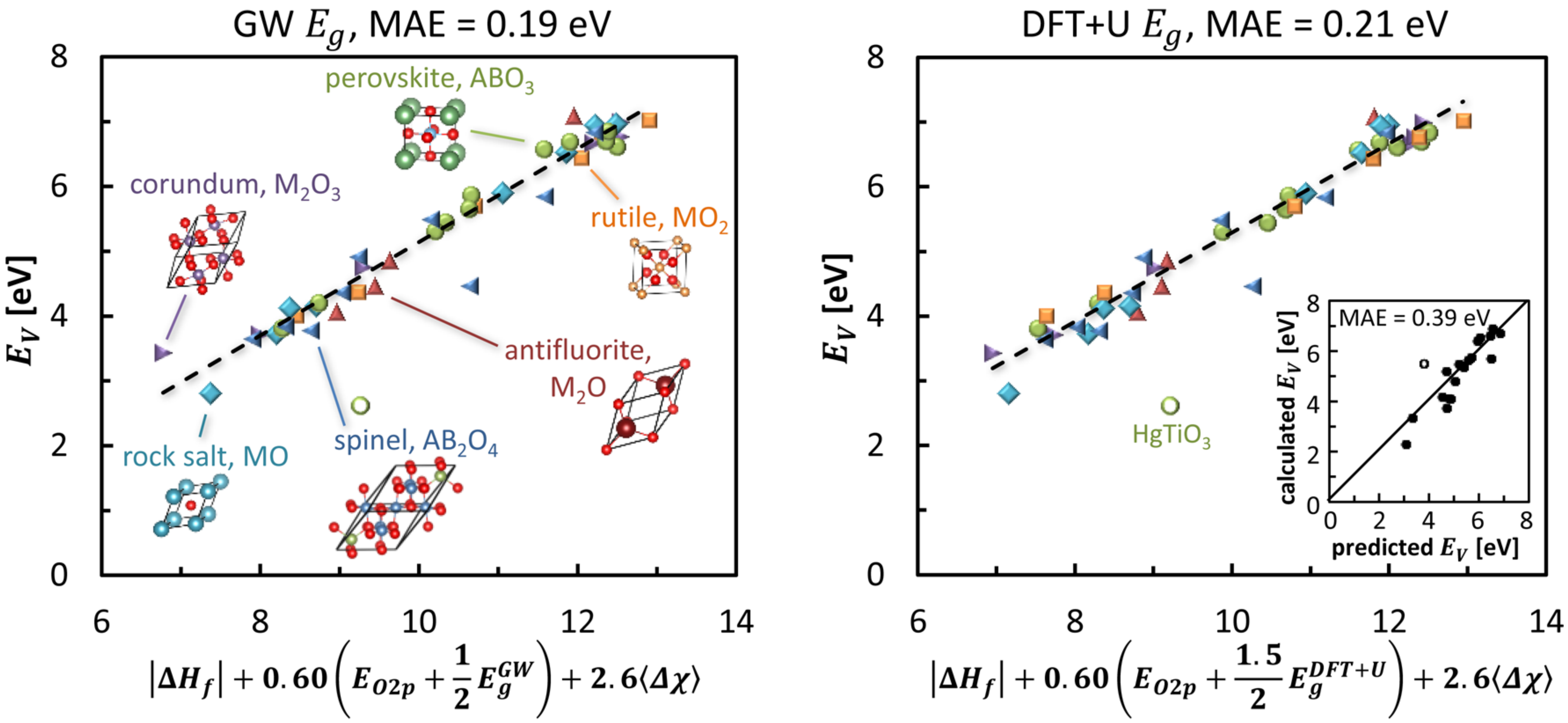
Sampling the Potential Energy of Solids
- Relevance to Predictive Synthesis and an Opportunity for ML -

Vladan Stevanović

Colorado School of Mines & National Renewable Energy Laboratory



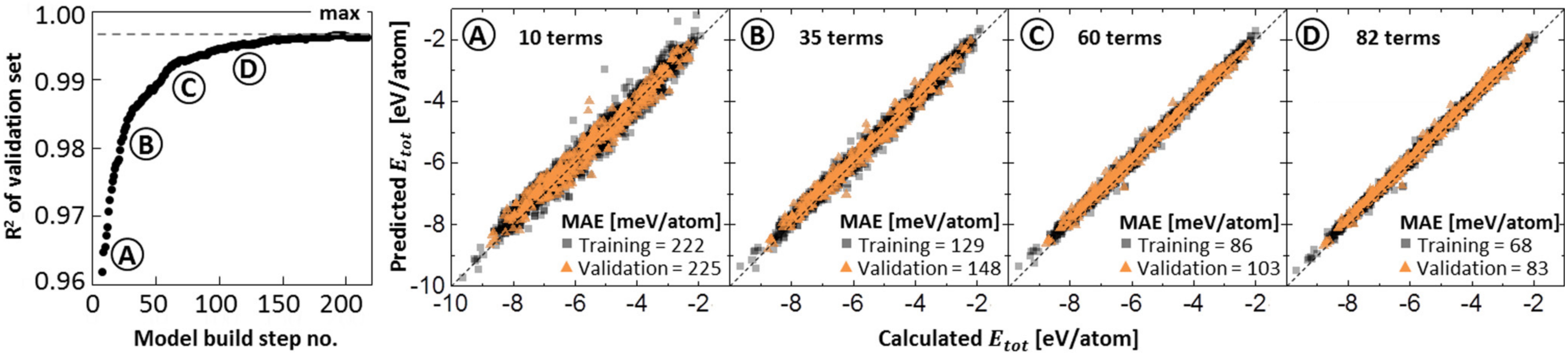
Regression model for oxygen vacancy formation energy



- We used the stepwise linear regression to investigate possible models for O vacancy formation entry and to select the most statistically significant subset of candidate descriptors
- Candidate descriptors - bulk properties plus quantities characterizing elemental constituents (13 total)

A. M. Deml, A. M. Holder, R. O'Hayre, C. B. Musgrave, and V. Stevanovic, J. Phys. Chem. Lett. 6, 1948 (2015)

Modeling DFT total energy - Also regression



- Stepwise linear regression again to model DFT total energies of compounds from the available descriptors of the chemical composition and elemental constituents (no structural information)
- A total of ~2,000 compounds (oxides, sulfides,…) formed the training set with additional ~300 as the test set
- Candidate descriptors - properties describing the compound composition and the physical and chemical properties of its elemental constituents
- We included the: maximum, minimum, range, standard deviation, stoichiometric weighted means, sqrt and inverse of each term as well as their products - a total of ~5,000 starting descriptors

A. M. Deml, R. O'Hayre, C. Wolverton, V. Stevanovic, Phys. Rev. B 93, 085142 (2016)

Phonon contributions to the free energy

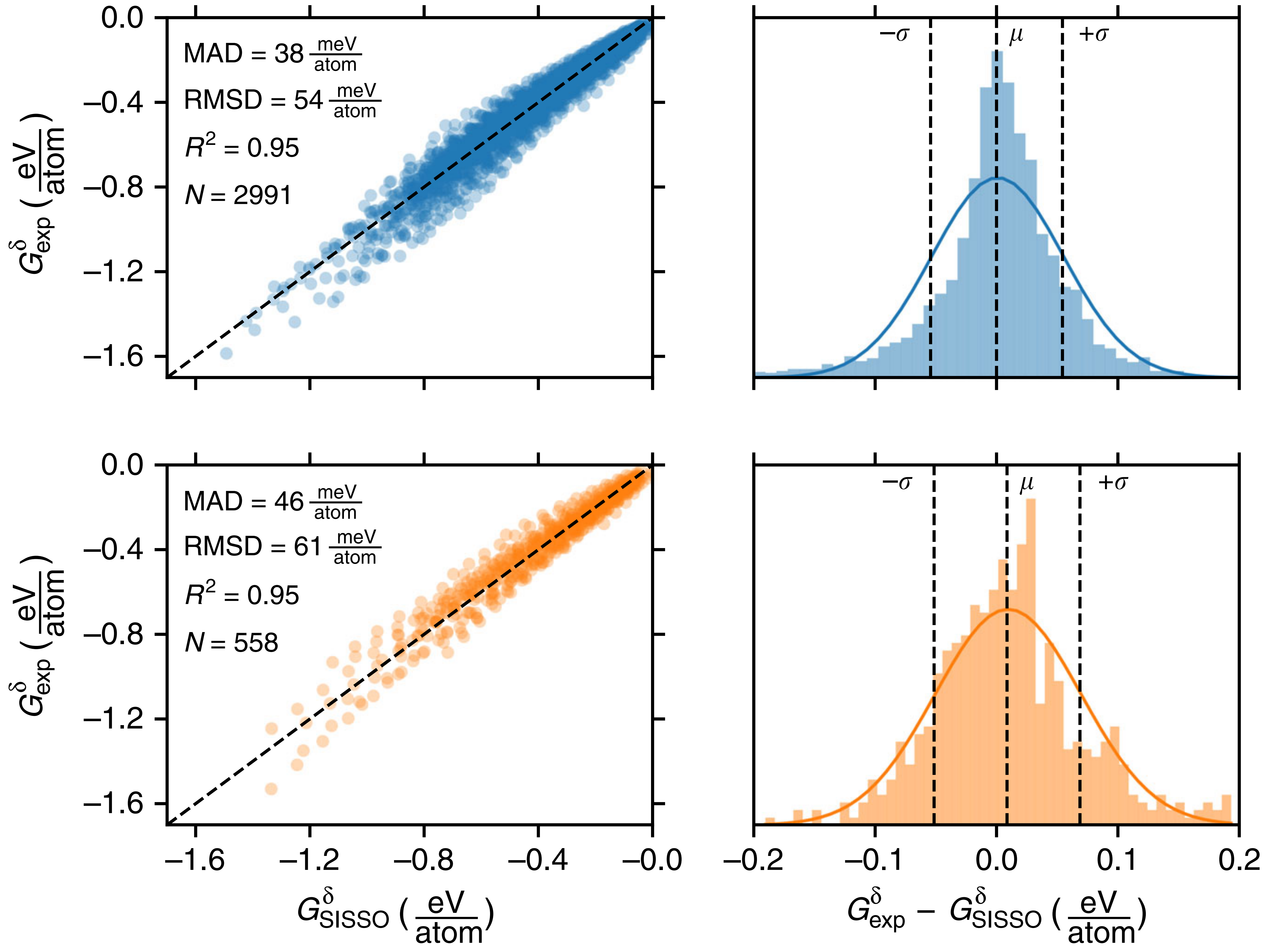
- Phonon contributions to the compounds free energy of formation:

$$\Delta G_\mathrm{f}(T) = \Delta H_\mathrm{f}(298\,K) + G^\delta(T) - \sum_{i=1}^{N} \alpha_i G_i(T)$$

- SISSO (sure independence screening and sparsifying operator) approach was used to identify the following model for $G^\delta(T)$:

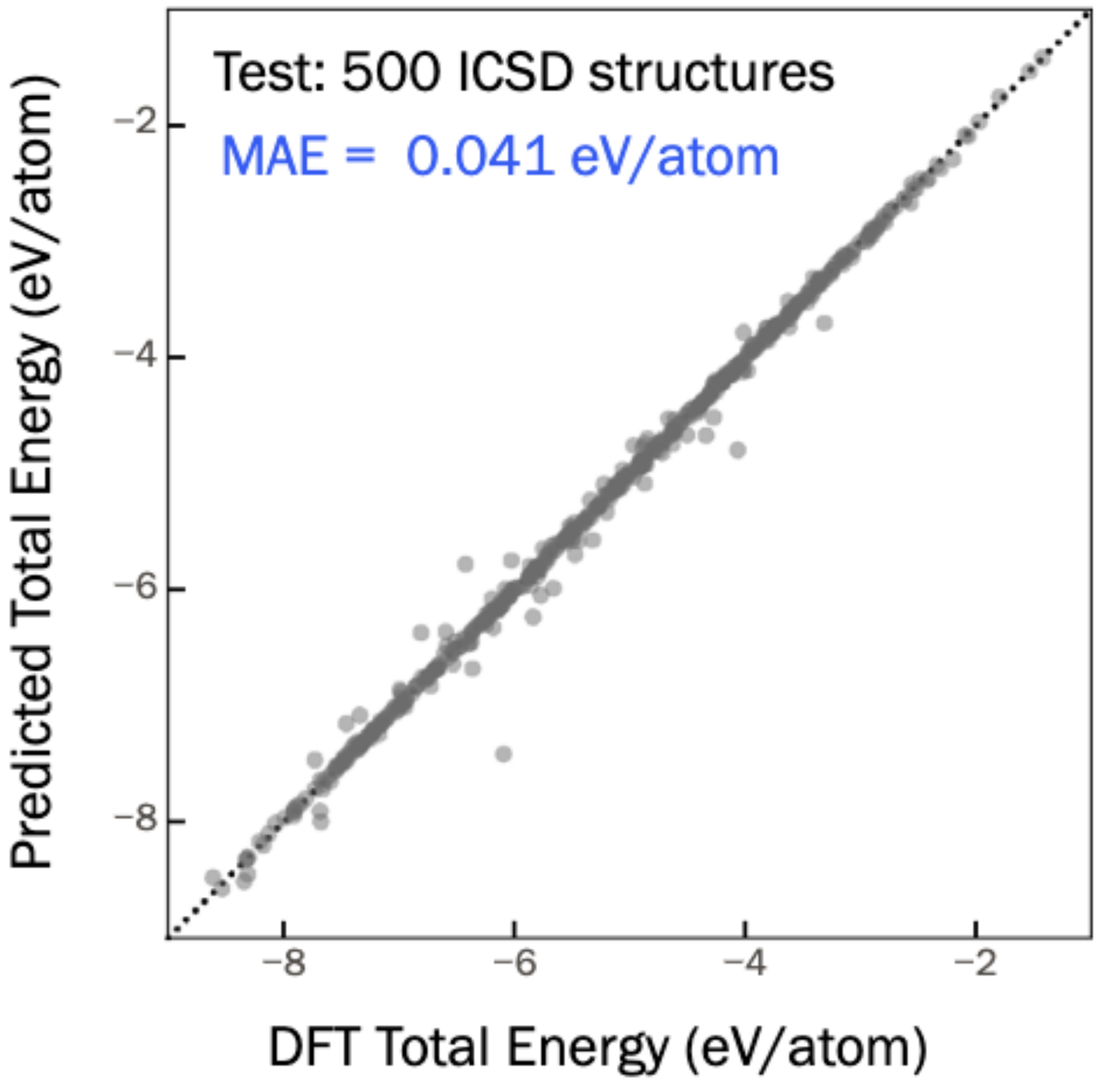
$$G^\delta_{\mathrm{SISSO}}(T)\left[\frac{\mathrm{eV}}{\mathrm{atom}}\right] = \left(-2.48*10^{-4}*\ln(V) - 8.94*10^{-5}\mathrm{mV}^{-1}\right)T + 0.181*\ln(T) - 0.882$$

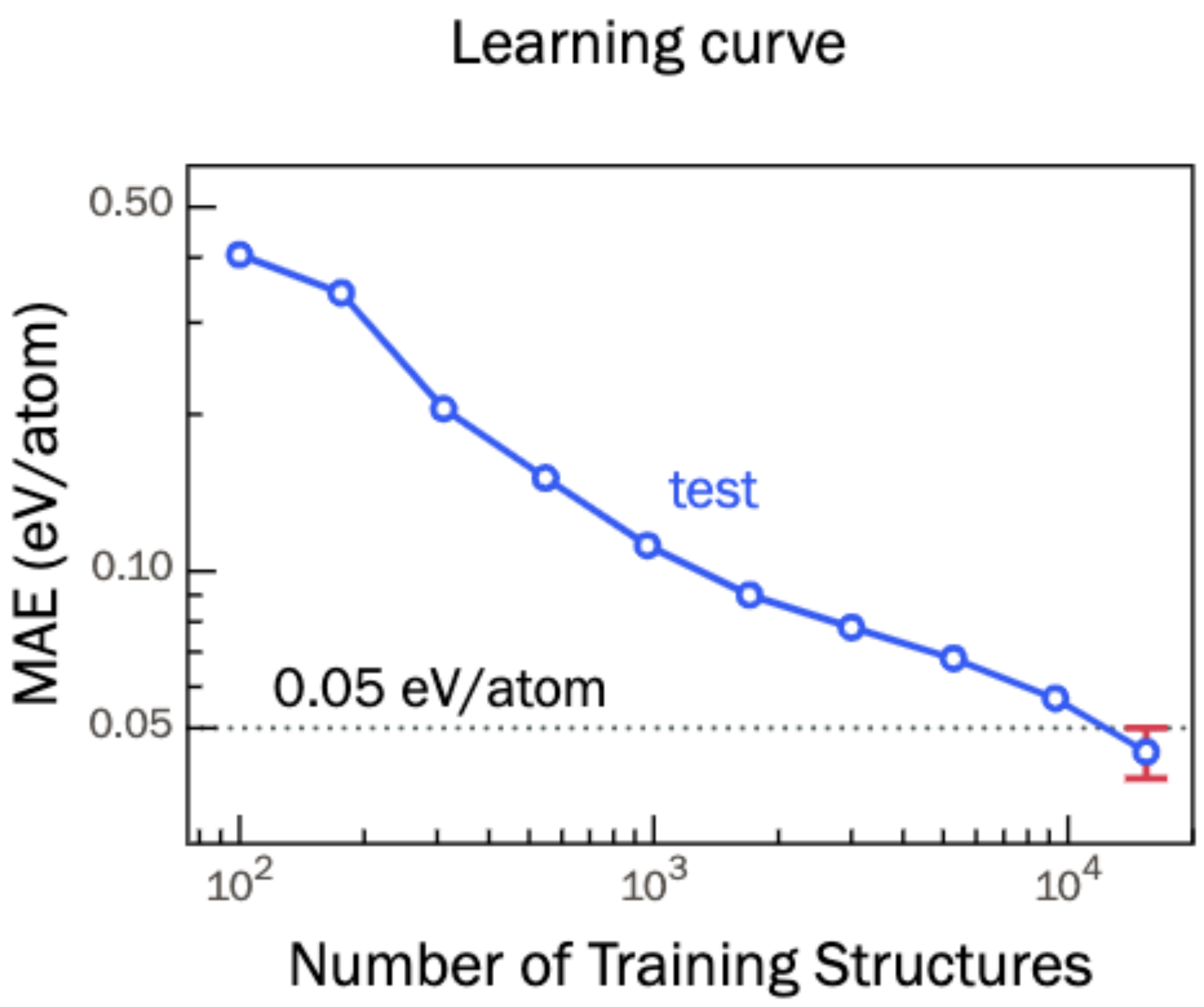
Training set - 262 compounds with 2,991 (T, G^δ) points
Test set - 47 compounds with 558 (T, G^δ) points
Large set (millions) of candidate descriptors

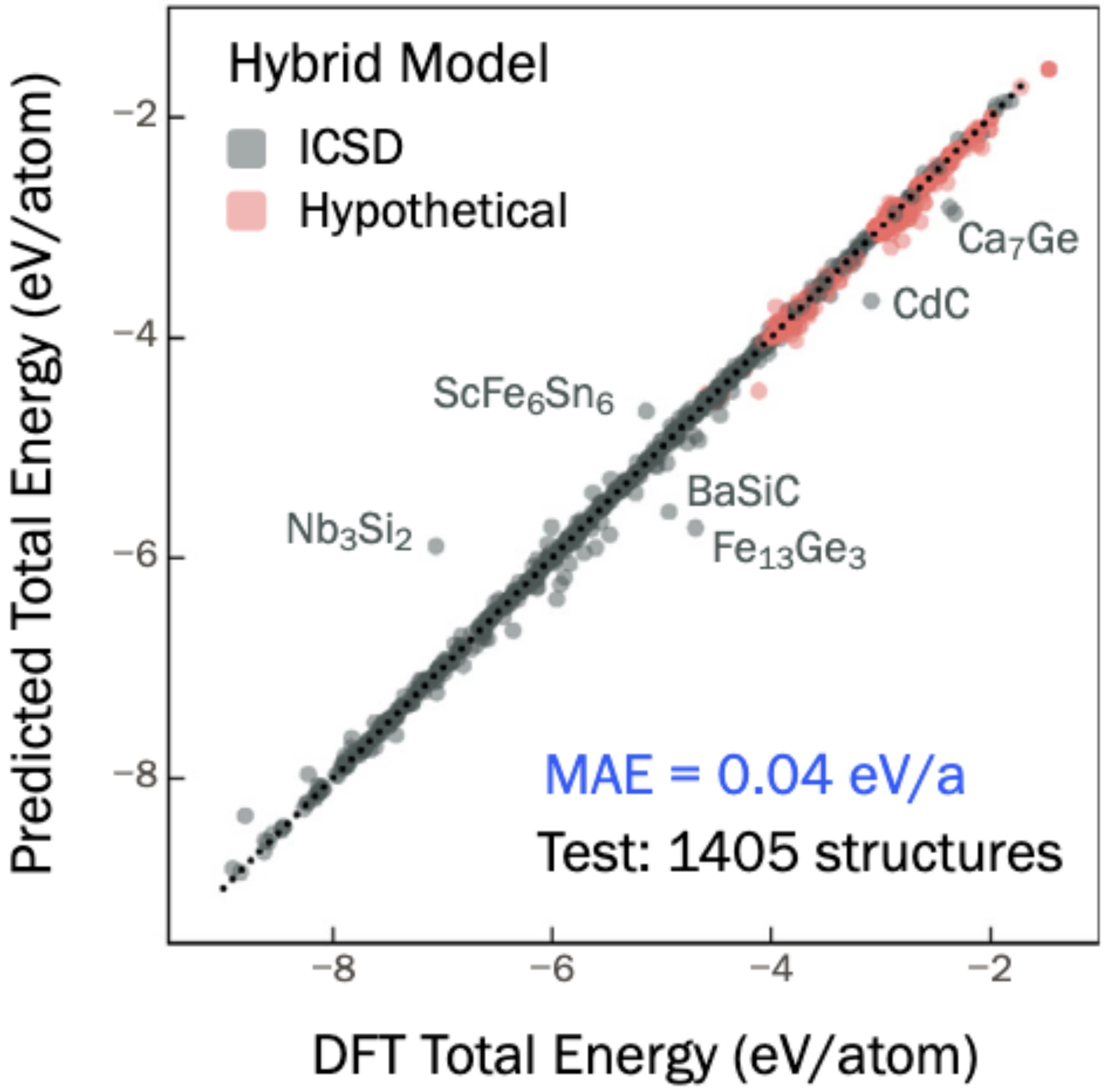


C. J. Bartel et al., Nature Communications 9, 4168 (2018)

DFT total energy from neural networks (GNN) - Ground state vs. higher energy structures

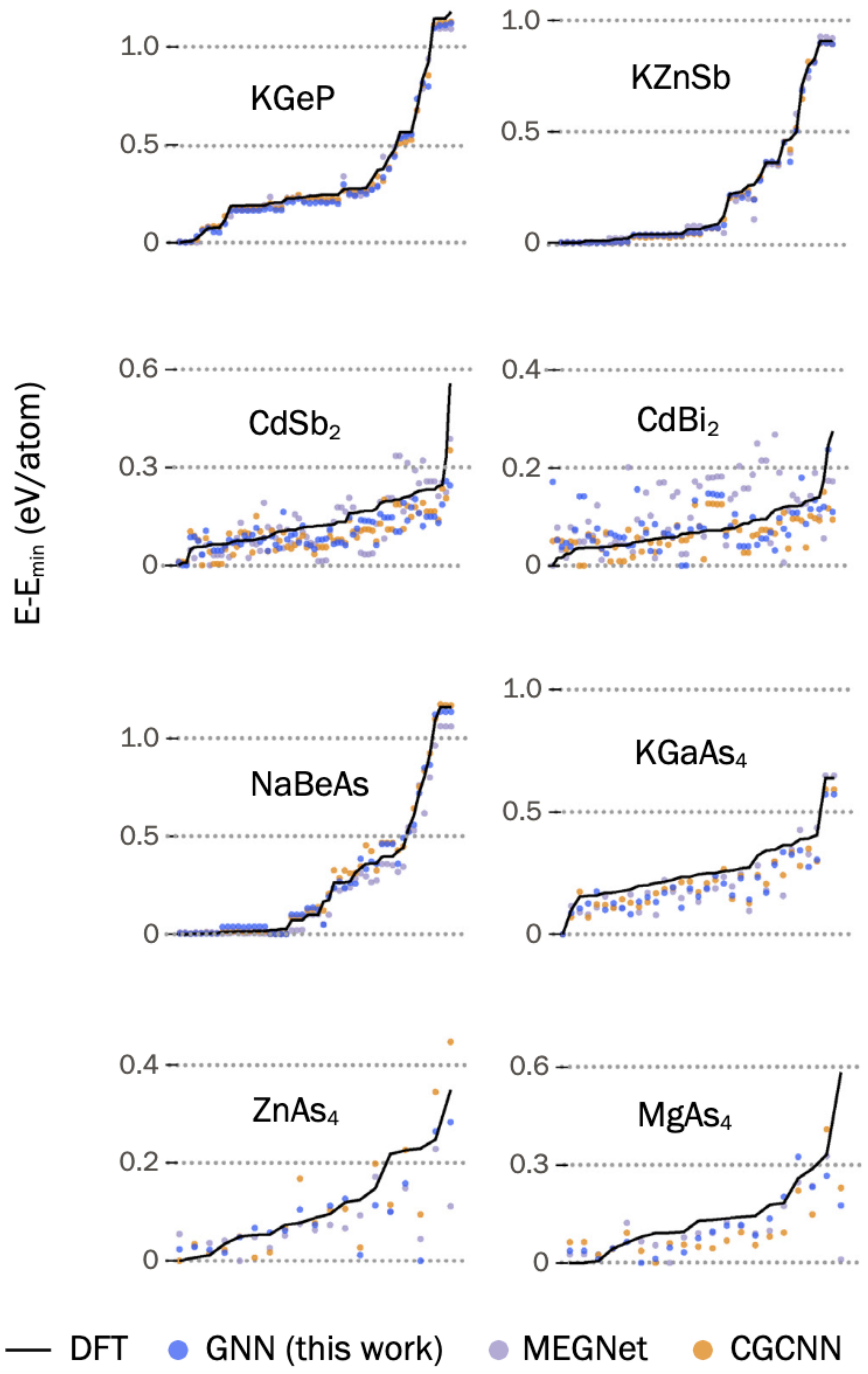






- Graph neural networks (GNN) used to model DFT total energies for both ground-state and higher energy (hypothetical) structures with the goal to facilitate structure predictions
- Structure information now explicitly used (in the form of graphs)
- Training set: DFT energies of 14,845 ICSD and 9,980 hypothetical structures (171 compositions)
- Test set: 1405 structures across 10 compositions

S. Pandey, J. Qu, V. Stevanovic, P. St. John, and P. Gorai, Patterns 2, 100361 (2021)



Challenges/Opportunities:

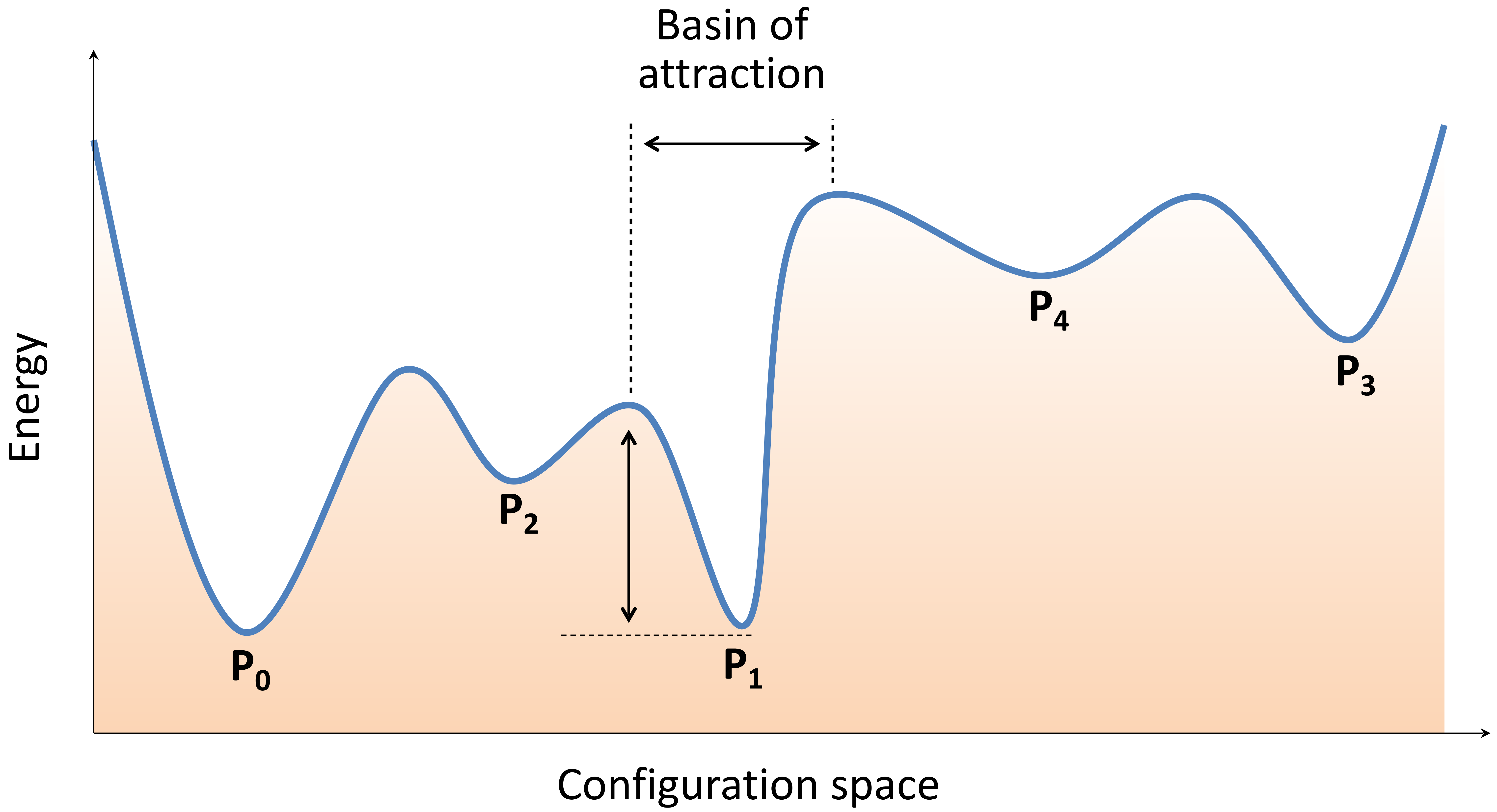
- The set of higher energy structures is in principle infinite
- Chemical substitution generates only a subset of possible higher energy structures
- Experimental data scarce
- Also, to make the problem tractable we restricted our training database to ~170 compositions for higher energy
- As usual, extrapolation does not work very well
- Good news: extrapolating to chemistries that are well-represented in the training set works better
- Obvious, but not so simple solution - increase the size of the training set
- Return on investment - accurate structure predictions, a necessary step in predicting synthesis outcomes

S. Pandey, J. Qu, V. Stevanovic, P. St. John, and P. Gorai, Patterns 2, 100361 (2021)

Predicting synthesis outcomes - The holy grail of materials discovery

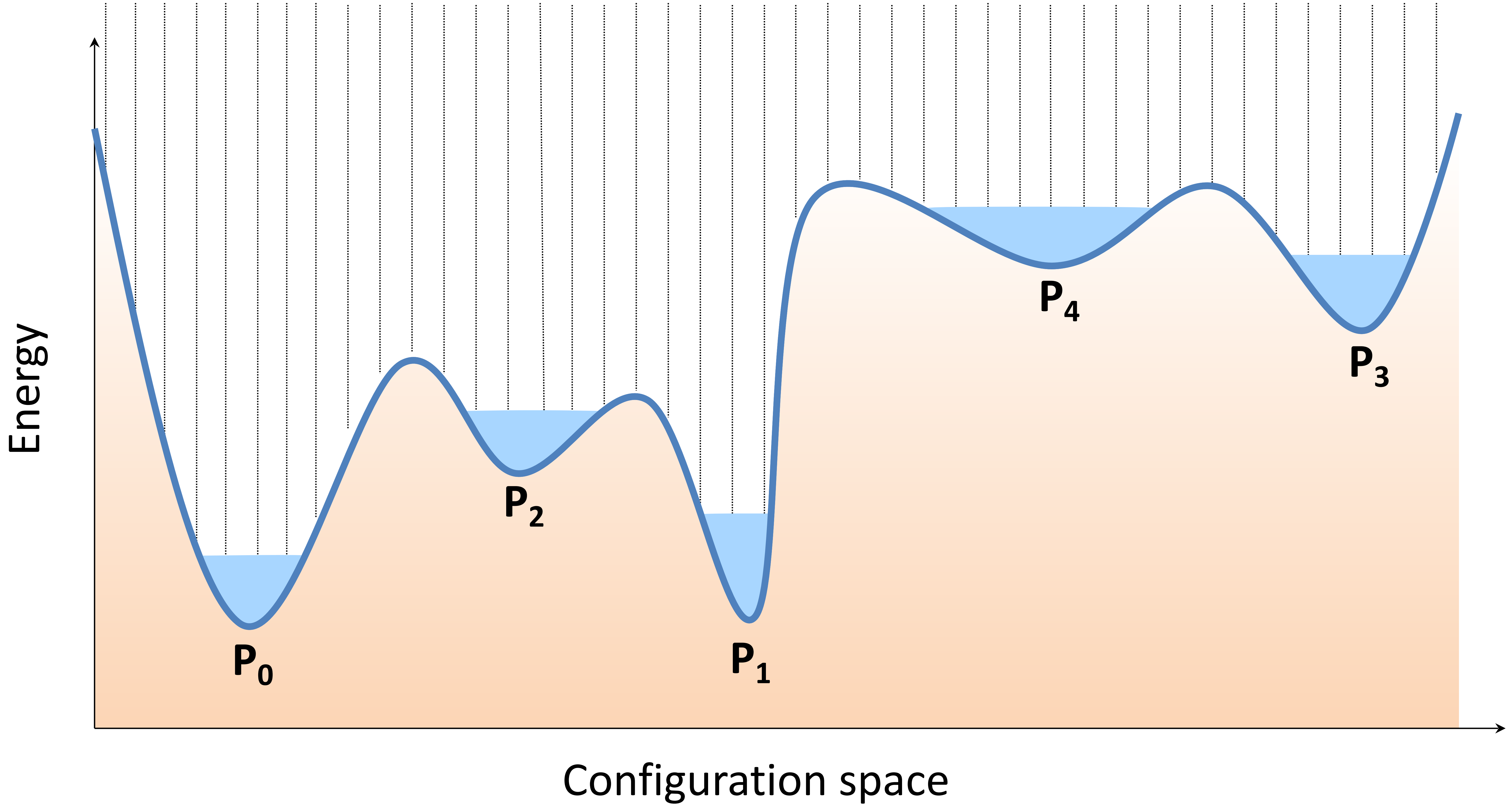
- Reliable structure predictions are necessary for materials discovery
- They require a sufficiently accurate description of the entire potential energy surface
- However, that is not sufficient
- Physical picture and principles governing synthesis outcomes still under investigation

Hypothesized governing principle - Size of local minima on the PES



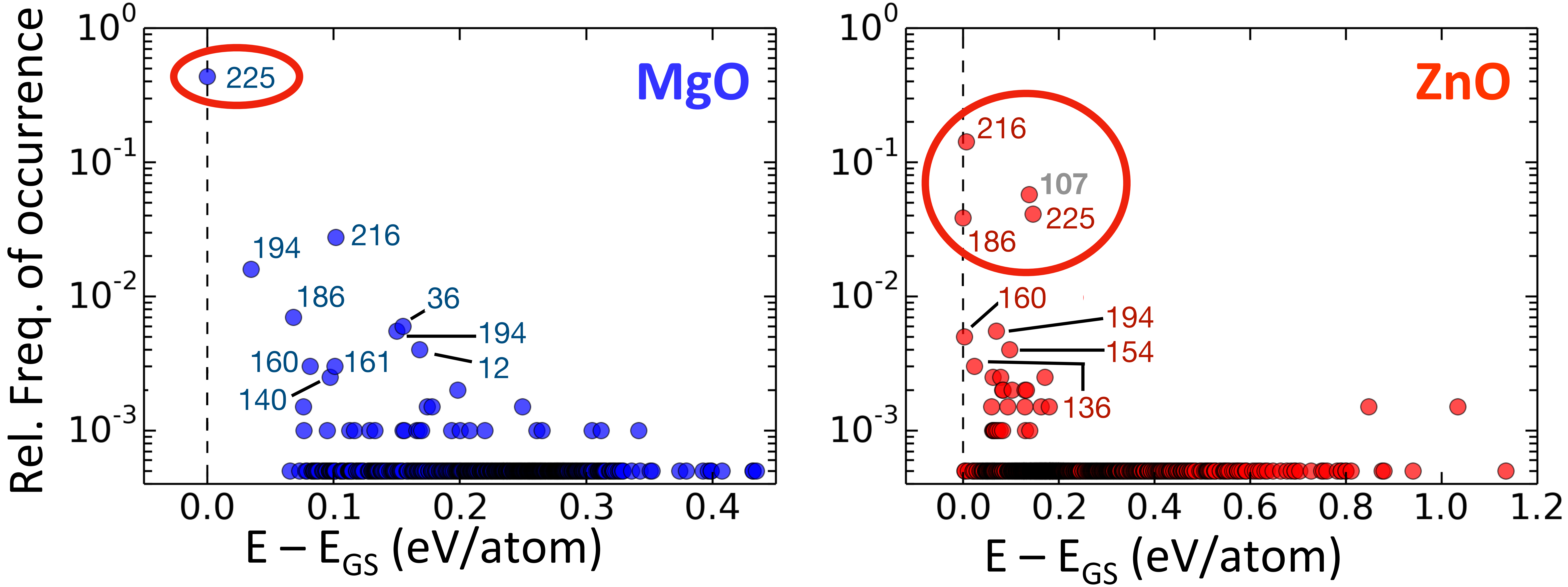
- “Widths” of attraction basins determine the probability to fall into a given local minimum
- Previously discussed by S. Goedecker et al. (PRL 112, 083401, 2014) in the context of metallic clusters

First-principles random structure sampling



- Random structure sampling (à la Pickard and Needs) followed by DFT relaxations
- Frequencies of occurrence in random sampling can measure the “widths” of individual basins

Well "widths" and observed polymorphs

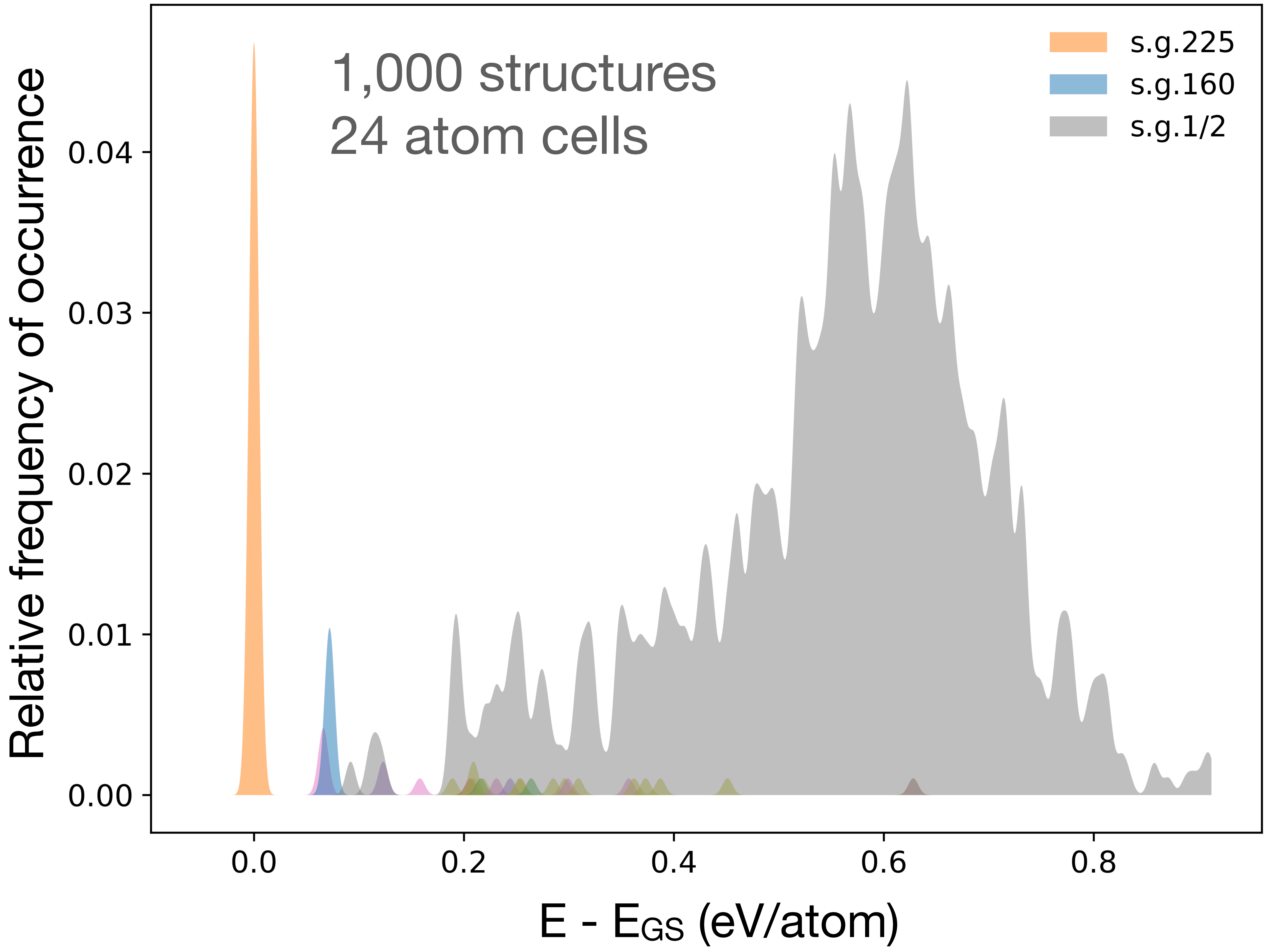


- 2,000 random structures per system
- All experimentally known metastable polymorphs appear as high freq. structures

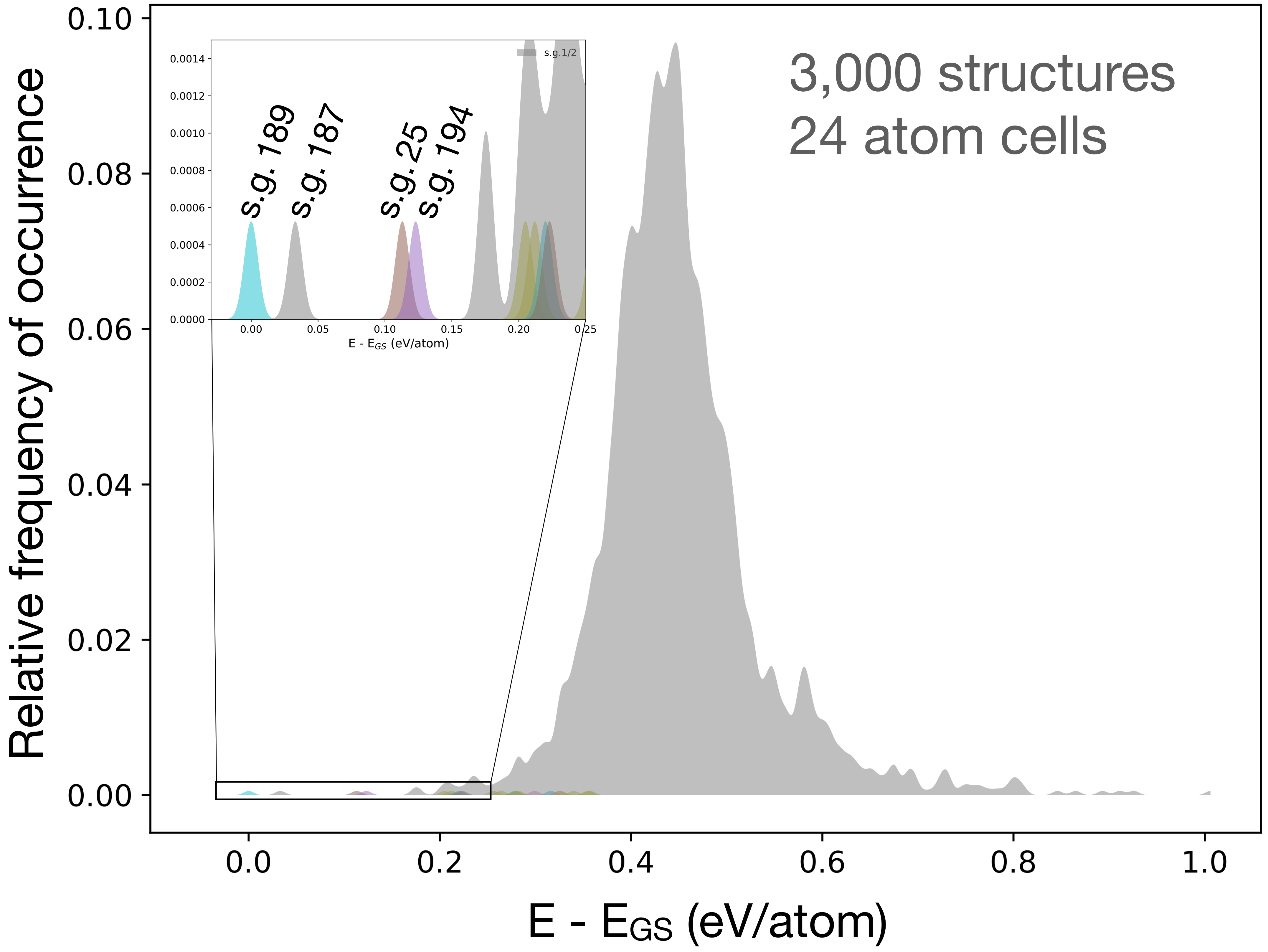
V. Stevanovic, Phys. Rev. Lett. **116**, 075503 (2016)

Narrow local minima and the impact on synthesis

TaC - large (“wide”) ground state



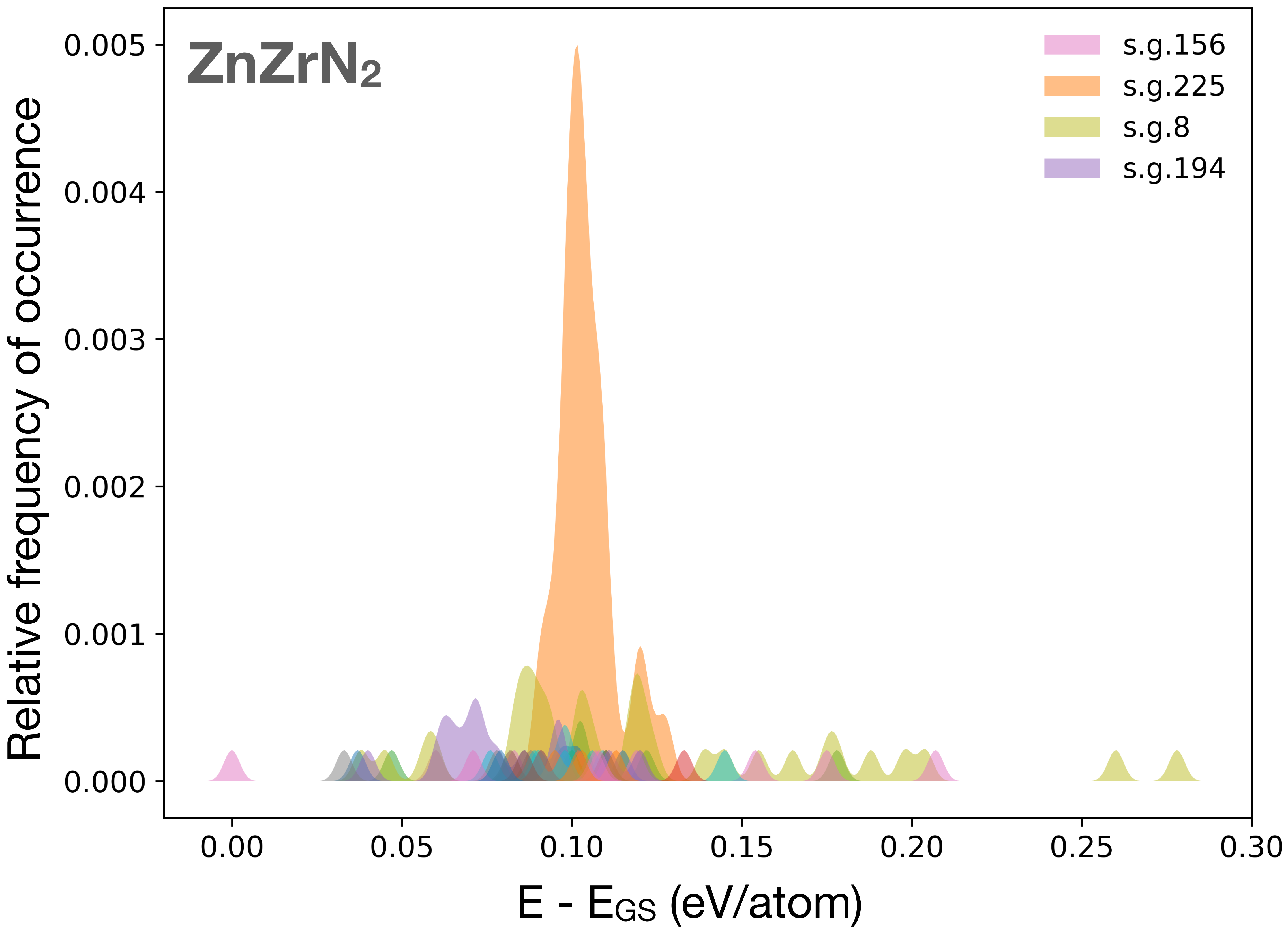
TaN - narrow ground state

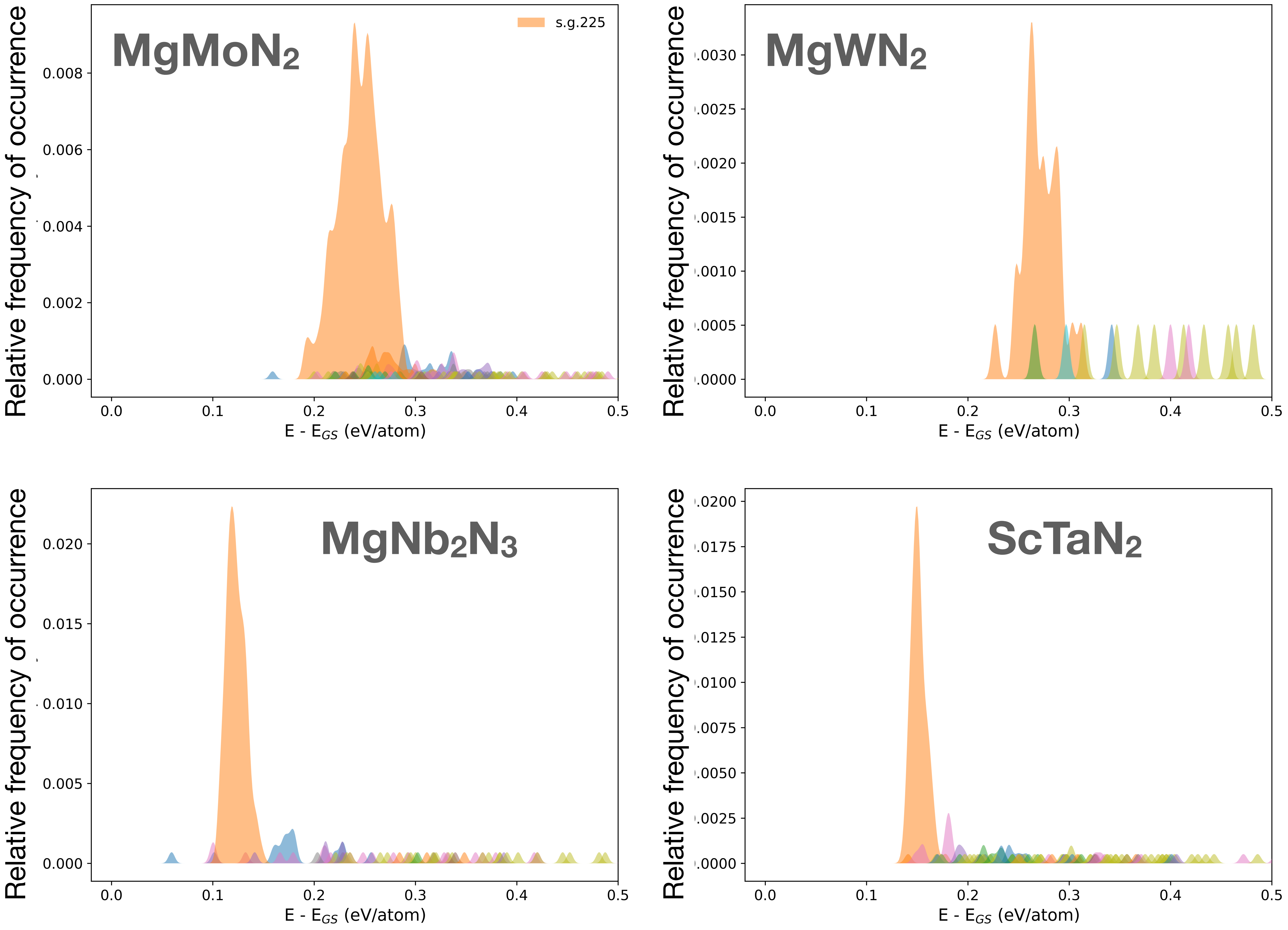


- Important consequence is that narrow local minima are hard to synthesize
- The example of TaN and the contrast to TaC provide some evidence that this is true

A. Novick et al., J. Am. Chem. Soc. 147, 4419 (2025)

Predicting outcomes of thin-film synthesis - ternary nitrides



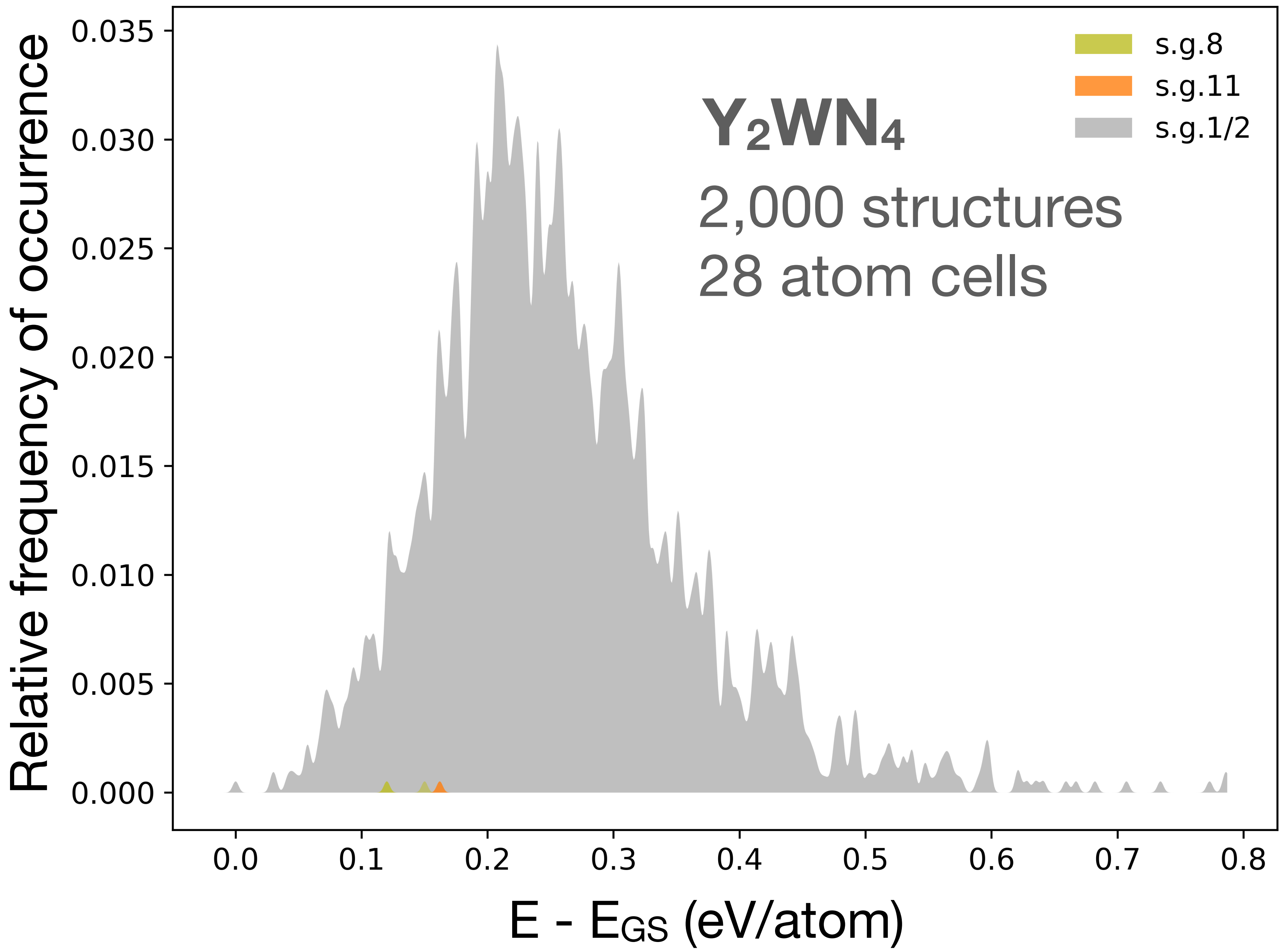


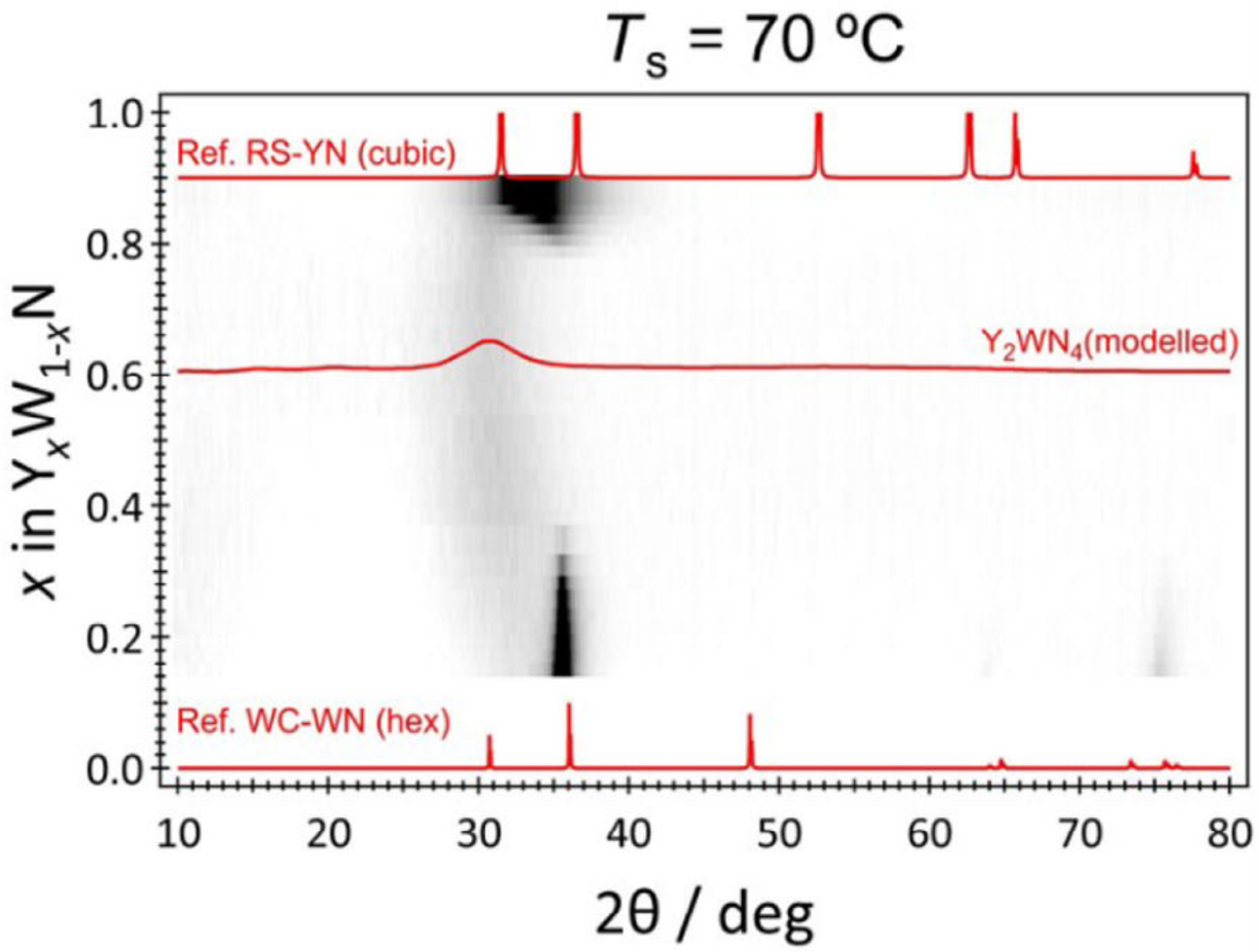
- A large group of ternary nitrides grown at NREL further supports these conclusions
- They are all grown in the cation disordered rocksalt structure and not in their ground states
- Disordered RS consistently appears as the most frequent structure in random sampling

R. Woods-Robinson et al., Phys. Rev. Materials 6, 043804 (2022)

A. Zakutayev et al., Nature Synthesis 3, 1471 (2024)

Amorphous phases and glasses

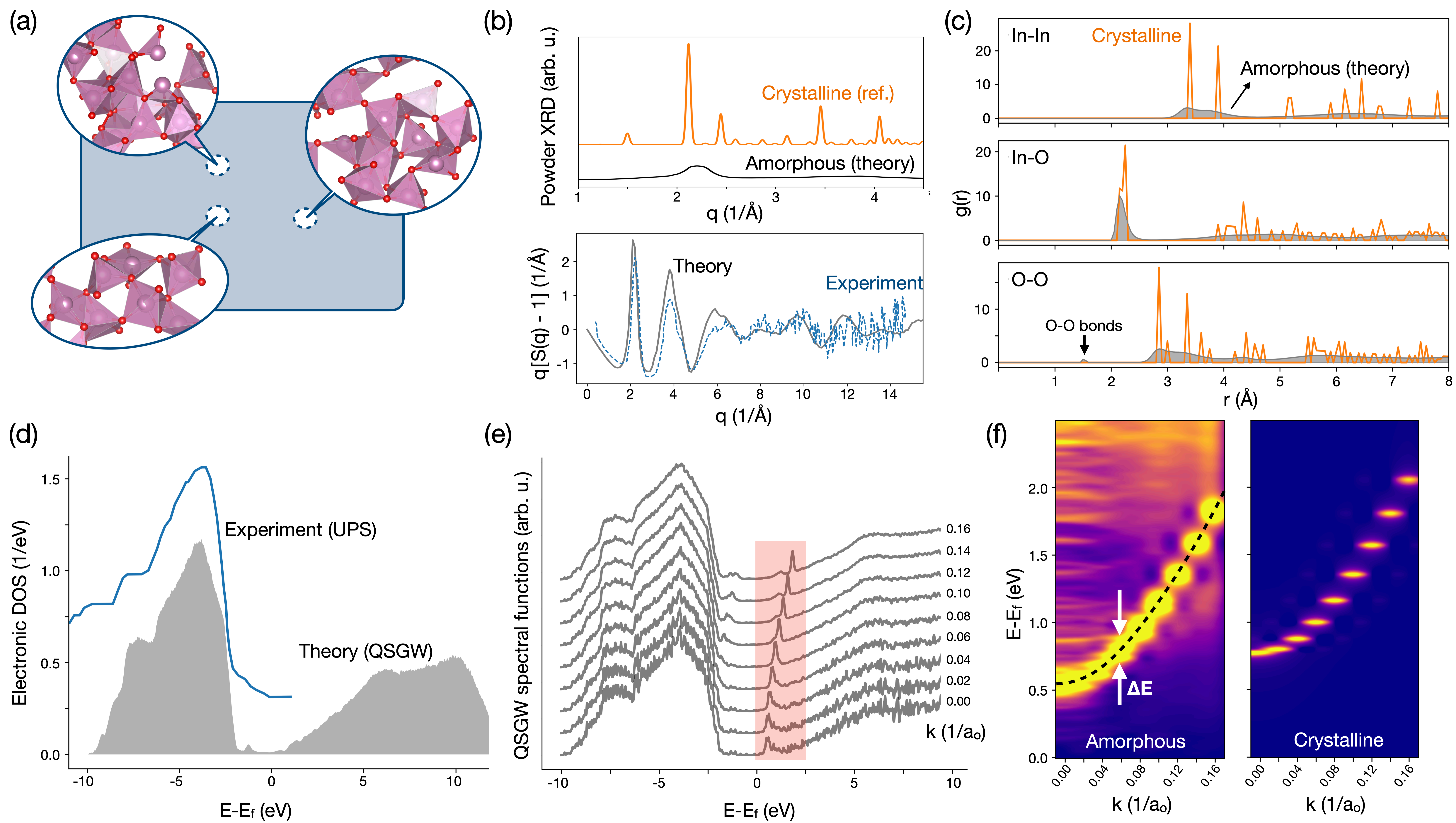




- In the extreme, very few (if any) symmetric structures are found
- The case of Y_2WN_4 indicates that this may be the feature of systems that form very stable amorphous phases
- Additionally, we discovered that by averaging over the properties of individual “random” structures we can reproduce measured properties of the amorphous phase

O. V. Pshyk et al., Advanced Materials 2501074 (2025)

Functional properties of amorphous phases - band like transport in In_2O_3

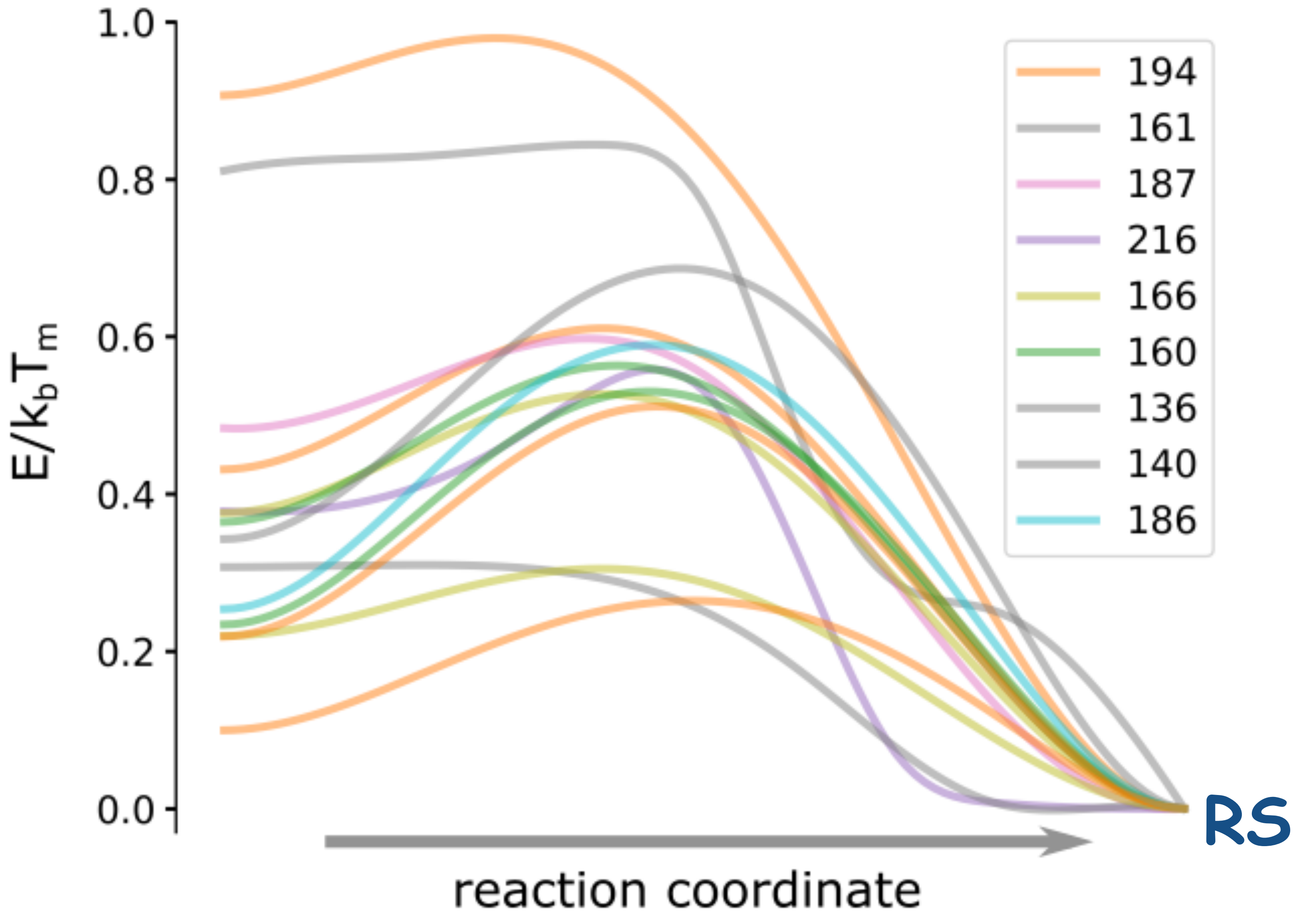


M. Jankousky et al., in review (arXiv:2505.07707)

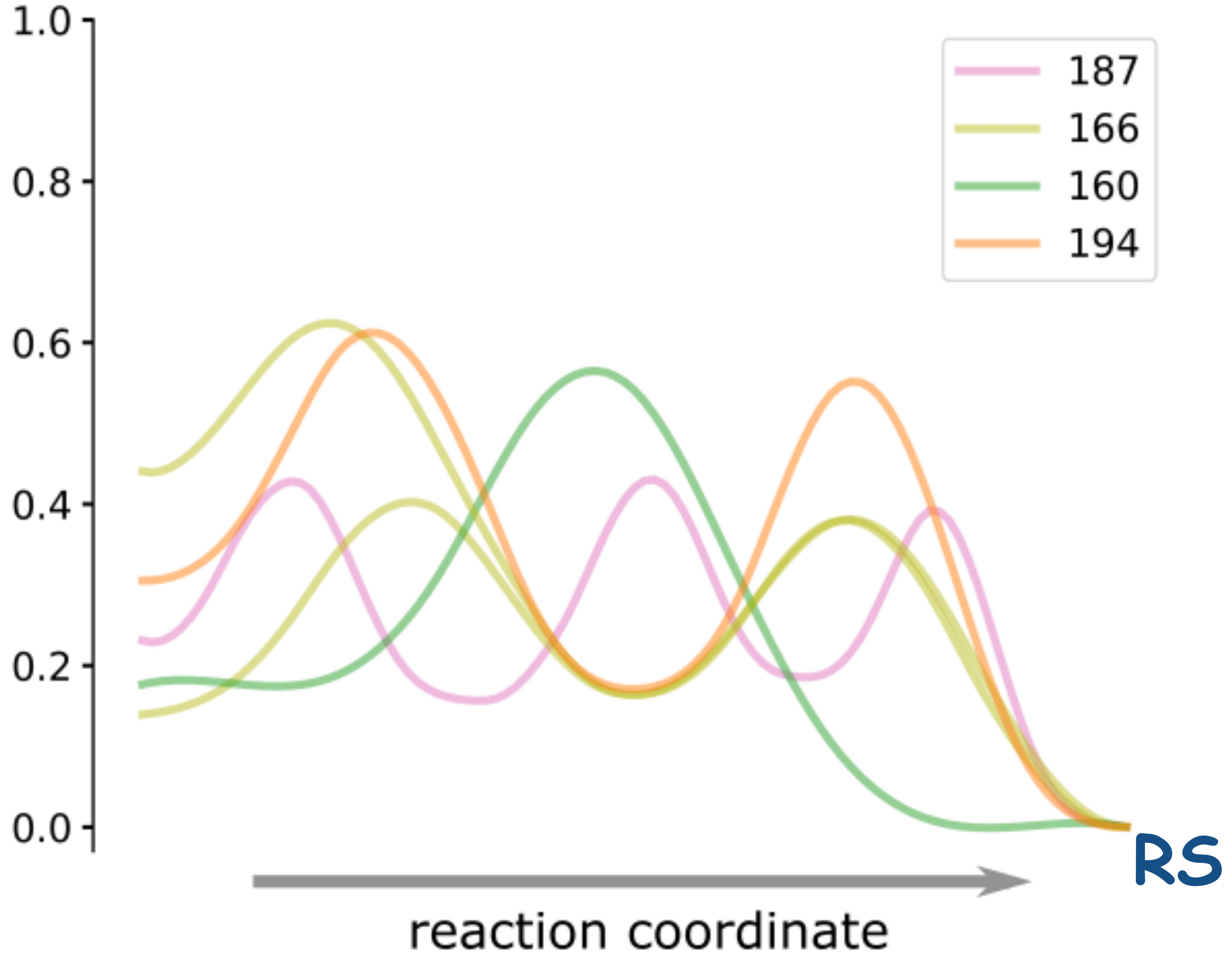
Depths of local minima

- We developed a geometric algorithm to map 3D periodic structures onto each other by minimizing displacements of atoms (F. Therrien, P. Graf, and V. Stevanovic, J. Chem. Phys. 152, 074106, 2020)
- Allows to create an initial guess of the solid-solid transformation mechanisms
- NEB method can then provide information about the bulk kinetic barriers, which measure the depths of local minima
- Successfully reproduces known phase transformations and explains absence of polymorphism in binary rock salts

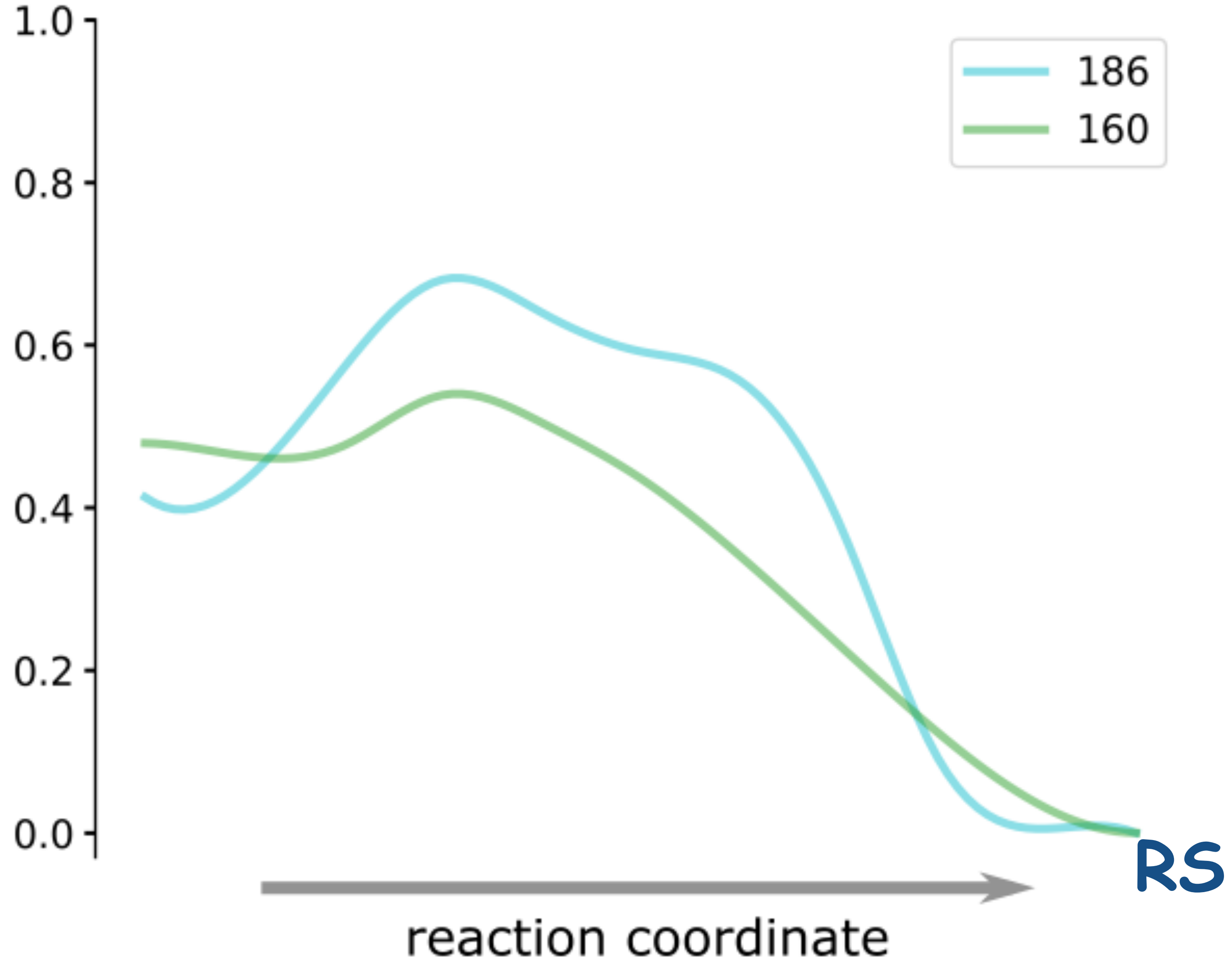
MgO



TaC

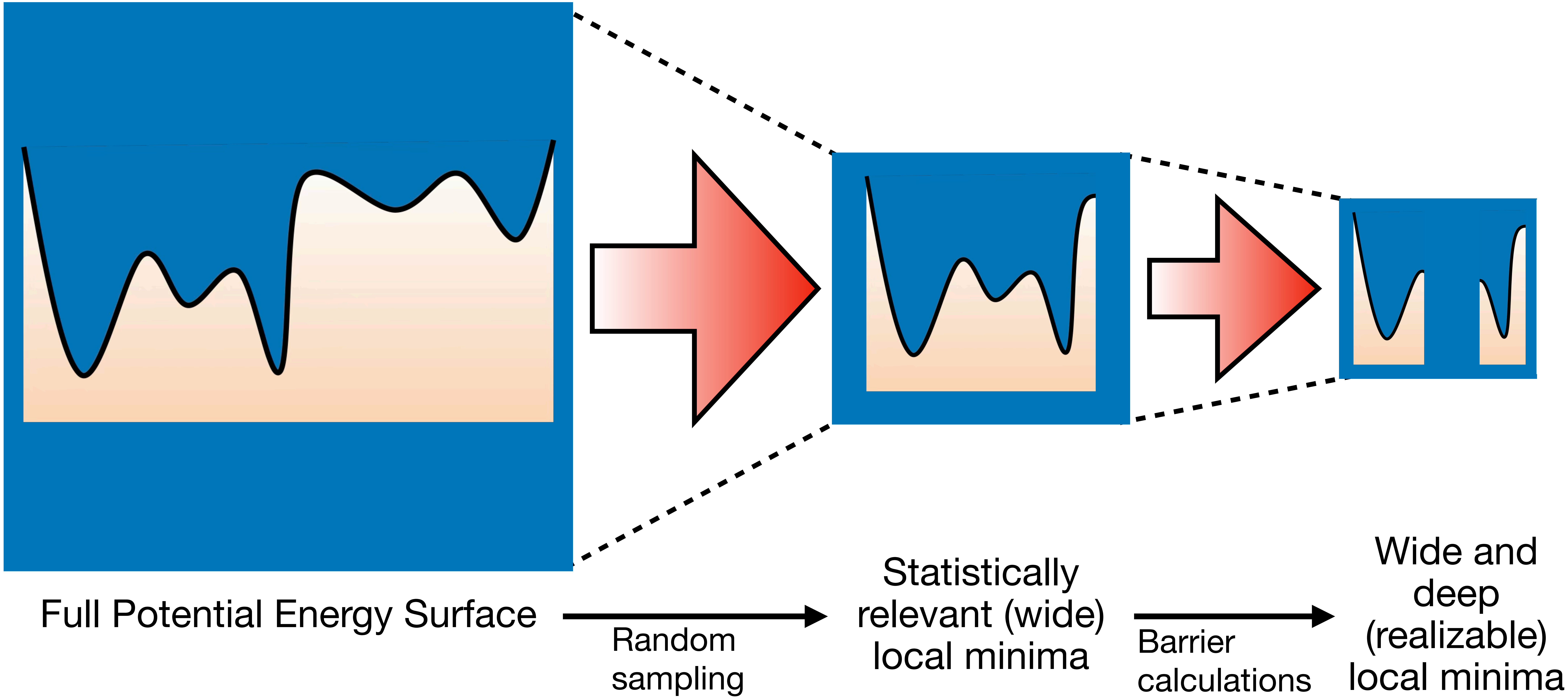


PbTe



M. Jankousky, H. Chen, A. Novick, V. Stevanovic, J. Am. Chem. Soc. 146, 23004 (2024)

A type of PES coarse graining that allows to predict the “ease” of synthesis



Full Potential Energy Surface

Random sampling

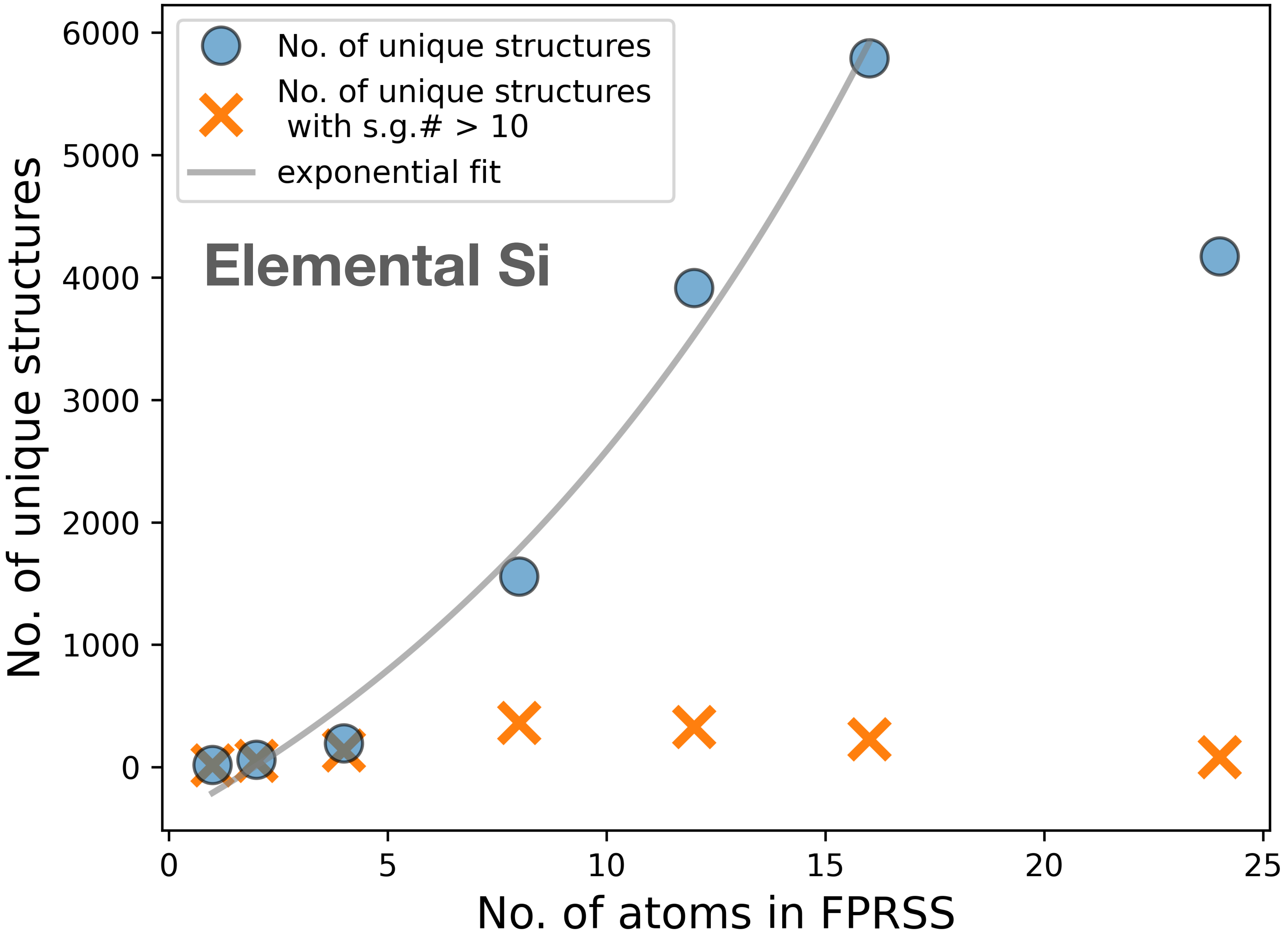
Statistically relevant (wide) local minima

Barrier calculations

Wide and deep (realizable) local minima

- *Chemistries for which this type of coarse graining does not work are prone to forming stable glassy/amorphous states*

Outstanding challenges and an opportunity for ML

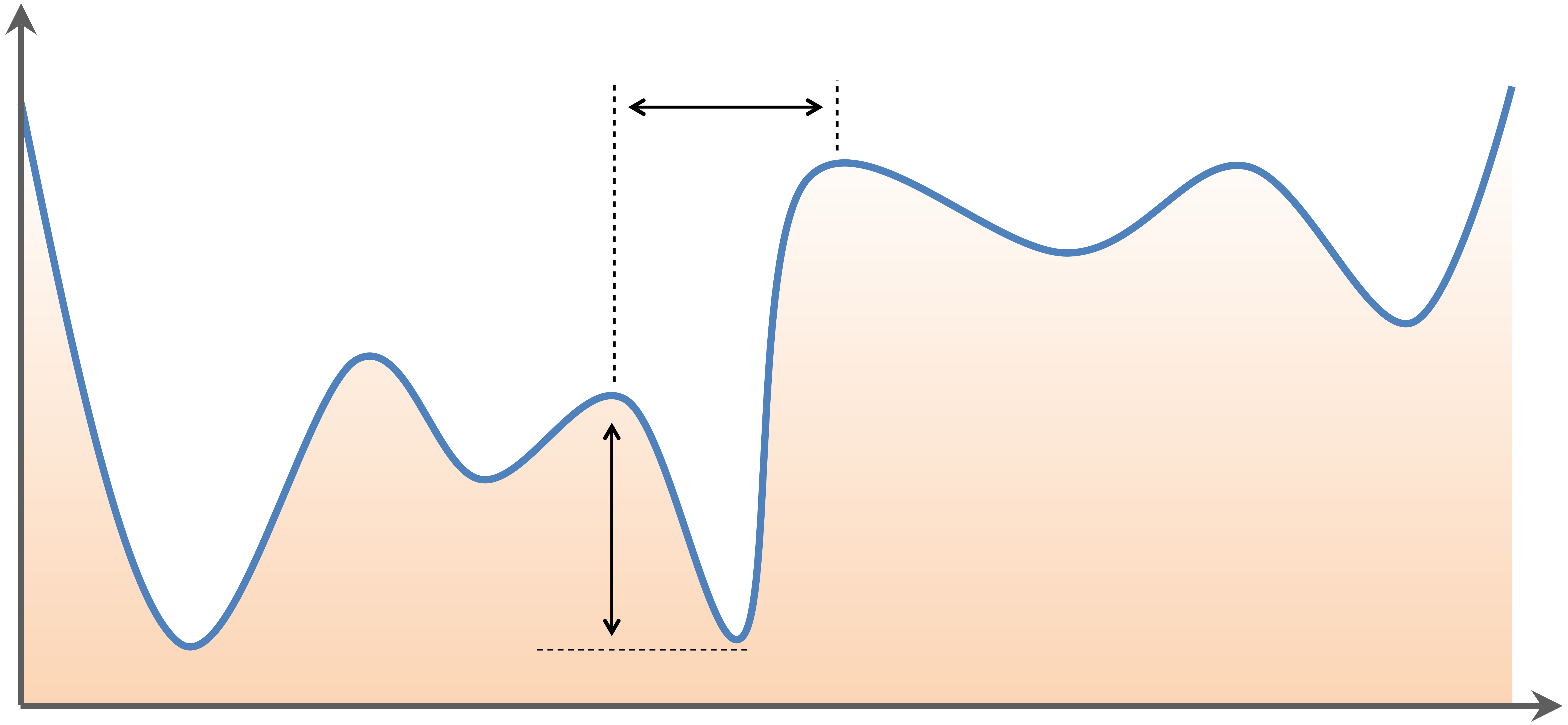


- Exponential increase of the number of PES local minima with the number of atoms (Stilinger 1999)
- Randoms sampling result on elemental Si
- ~32,000 randoms structures relaxed

- Plenty of evidence accumulated over the years in support of our hypothesis
- Wide local minima more likely to synthesize
- **Challenge 1**: the exponential growth of local minima with the number of atoms
- To complete the coarse graining evaluating the depths is also important
- **Challenge 2**: quadratic scaling of the number of pathways that need to be evaluated
- ML force fields necessary to advance these ideas to nucleation length scales (several nm)
- **Challenge 3**: Extrapolation of ML force fields to structures outside the training set

Conclusions

- Machine learning provides ample opportunities to model physical quantities that are hard (or even impossible) to calculate
- Problem is not solved though, we need a lot of data and extrapolation remains a challenge
- Herein I discussed a novel approach to predicting the "ease" of synthesis of solid phases that critically relies on structure prediction
- ML force fields are crucial to make this concept more quantitative and able to rationalize nucleation of different phases and their phase transformations



Acknowledgments

The Group (Oct. 2024)

HPC Resources

NREL Transforming ENERGY

COLORADO SCHOOL OF MINES

PATh

Funding

NSF

U.S. DEPARTMENT OF ENERGY
Office of Science

NREL Transforming ENERGY

Thank you!