

# From *Crystal Structure Prediction* to Materials Discovery : the case of metastable (oxy)sulfide compounds

[Isabelle.braems@cnsr-imn.fr](mailto:Isabelle.braems@cnsr-imn.fr)



L. Cario  
M. Caldes  
B. Corraze  
C. Guillot-Deudon  
E. Janod  
S. Jobic  
S. Sasaki  
L-B. Mvélé

G. Frapper  
B. Wang

S. Clarke  
S.Giri



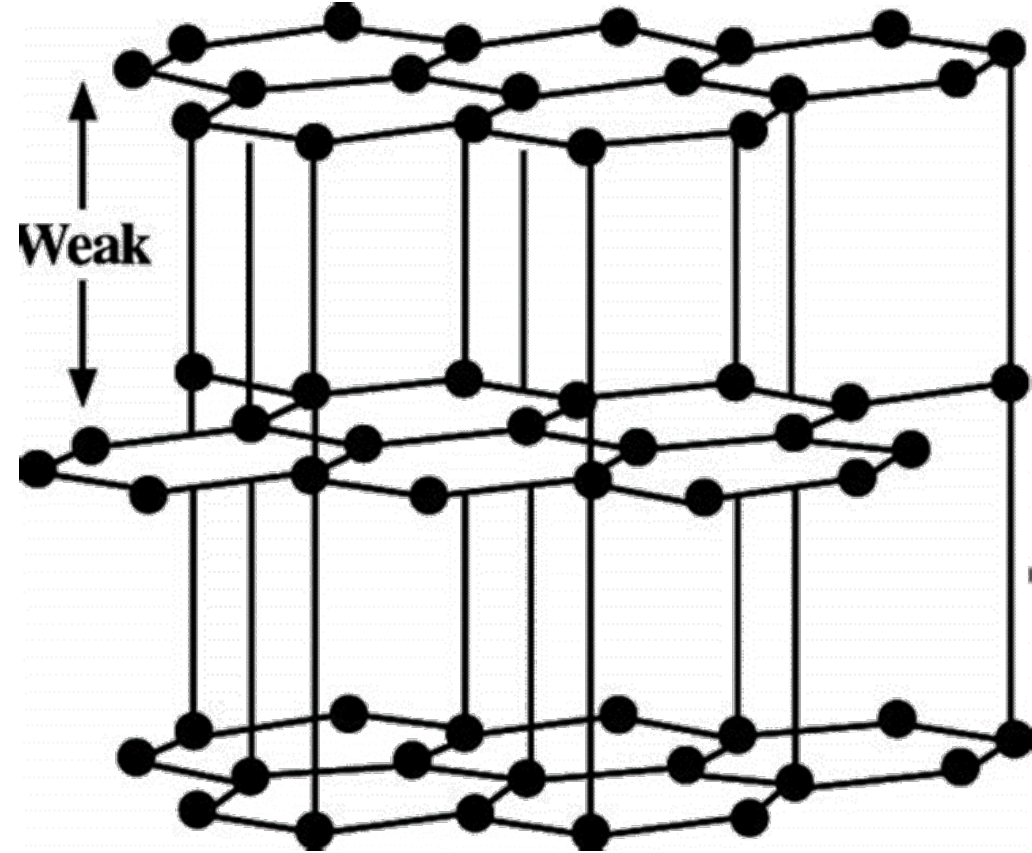
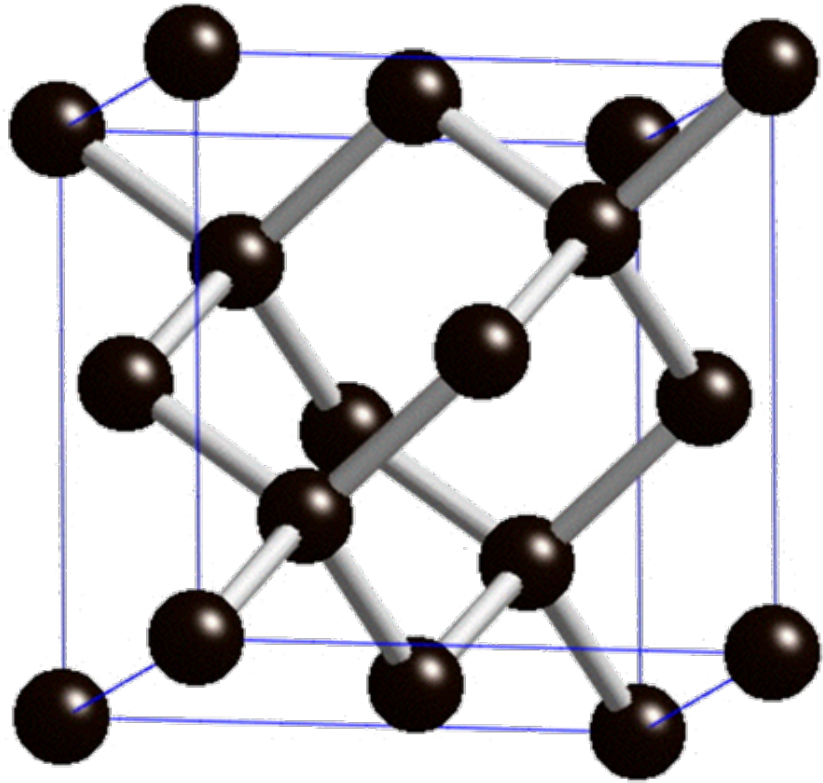
G. Steciuk  
L. Palatinus



***GDR MEETICC Conférence plénière 2023***  
*Latresne - Bordeaux, 30 mai- 2 juin 2023*

?

# Crystal structure explains crystal properties

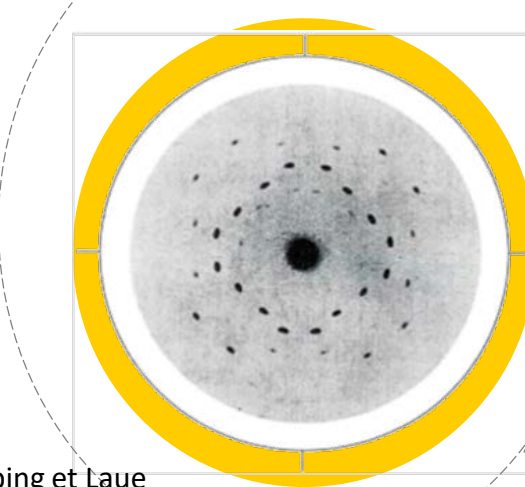




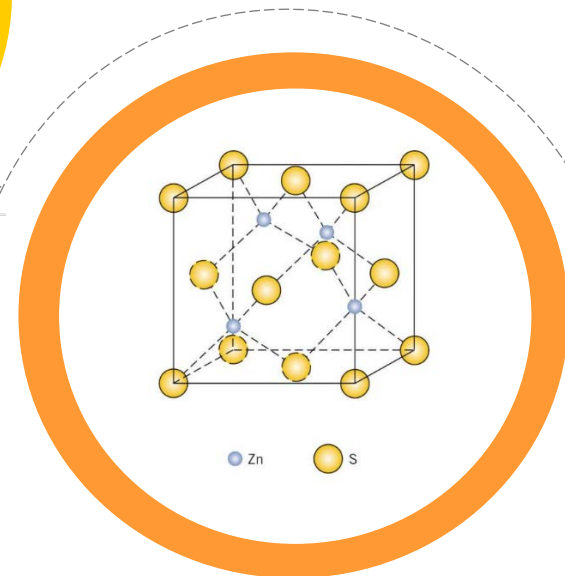
# « Seeing is believing » V. Ramakrishnan (NP Chemistry, 2009)

## The Nobel Prize in Physics 1914

Max von Laue “for his discovery of the diffraction of X-rays by crystals”

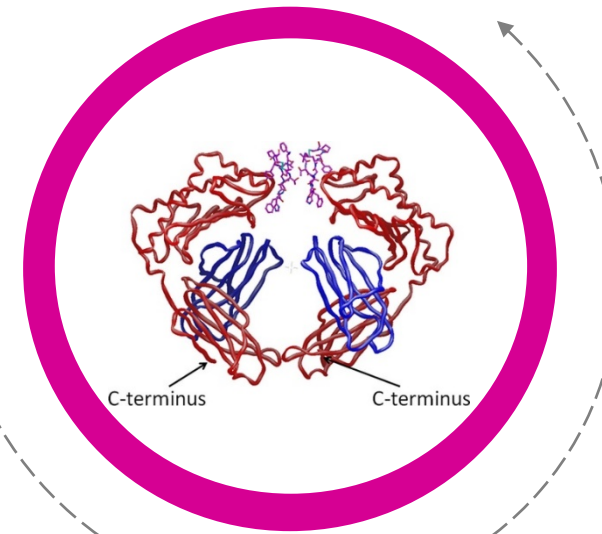


Friedrich, Knipping et Laue  
Sitzungsberichte der Kgl. Bayer.  
Akad. der Wiss. 1912



## The Nobel Prize in Chemistry 2017

Jacques Dubochet, Joachim Frank and Richard Henderson “for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution”



## The Nobel Prize in Physics 1915

Sir William Henry Bragg and William Lawrence Bragg “for their services in the analysis of crystal structure by means of X-rays”



# Crystal Structure Prediction (CSP)

Predict the crystal structure based only on structural formula  $A_xB_yC_z$



# Who would guess ?

*Acc. Chem. Res.* **1994**, *27*, 309–314

## Are Crystal Structures Predictable?

ANGELO GAVEZZOTTI\*

*Dipartimento di Chimica Fisica ed Elettrochimica, Università di Milano, Milano, Italy*

*Received May 16, 1994*

“No”: by just writing down this concise statement, in what would be the first one-word paper in the chemical literature, one could safely summarize the

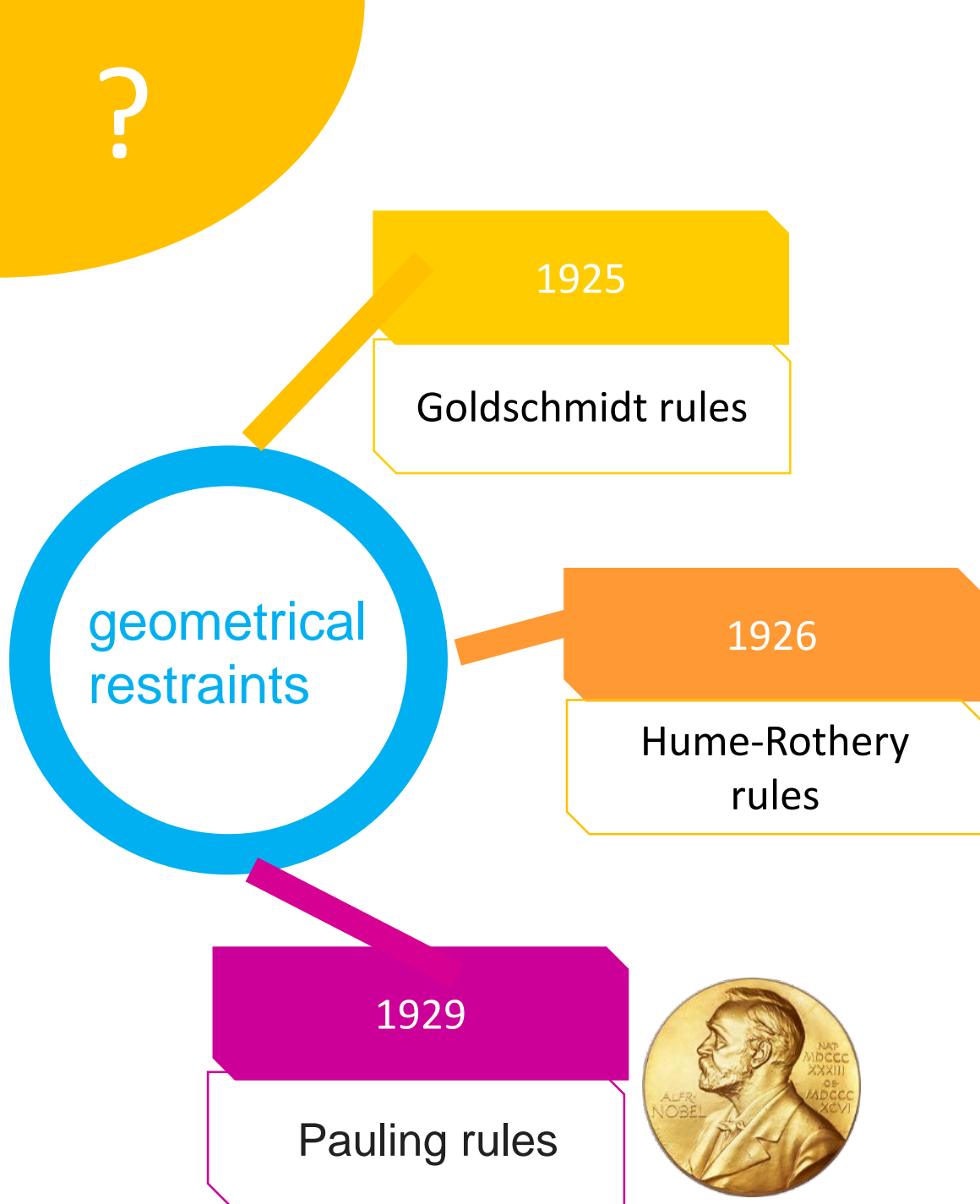
NATURE VOL. 335 15 SEPTEMBER 1988

NEWS AND VIEWS

201

## Crystals from first principles

ONE of the continuing scandals in the physical sciences is that it remains in general impossible to predict the structure of even the simplest crystalline solids from a knowledge of their chemical composition. Who, for example, would guess that





# Computational material design to **compress time...**

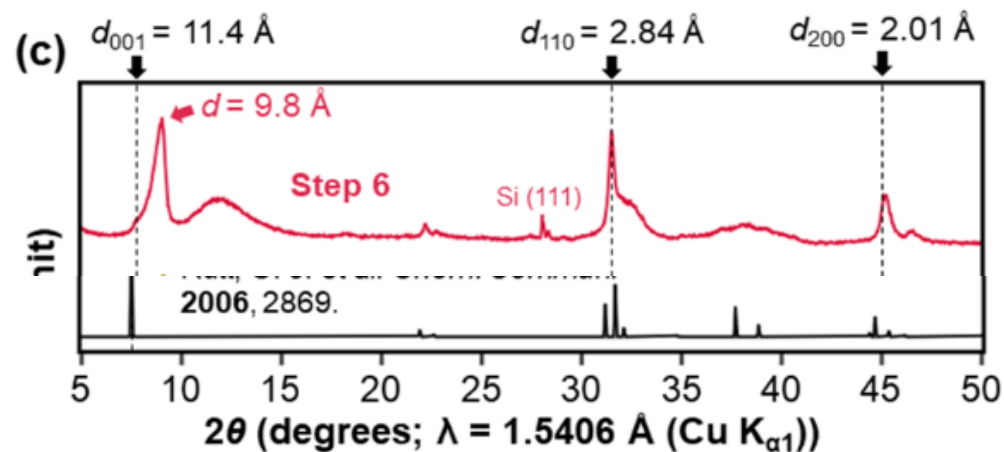


“The process of material design and discovery is traditionally long and complex. ***It takes roughly 10 years and upwards of \$10 to \$100 million on average*** to discover one new material with specific properties.

**A. Curioni** (IBM Fellow, Vice President of IBM Europe & Africa and Director of the IBM Research Lab in Zurich, Switzerland)

Only about 72% of binary, 16% of ternary, 0.6% of quaternary, and less than 0.05% of more complex systems have ever been studied experimentally.

... key to help **experimental** structure characterization



?

# CSP ladder

COMPUTING TIME

INVERSE METHODS

Global Optimisation  
(2006)

Data Mining  
(2003)

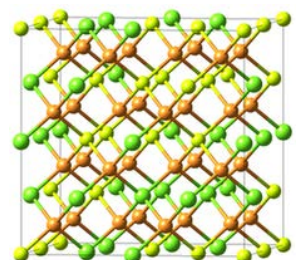
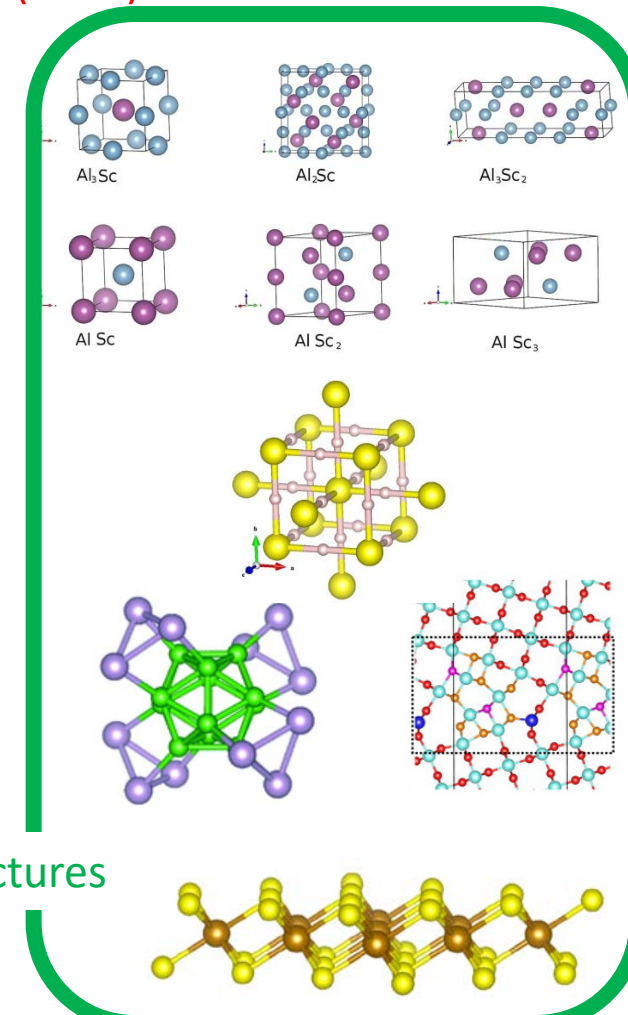
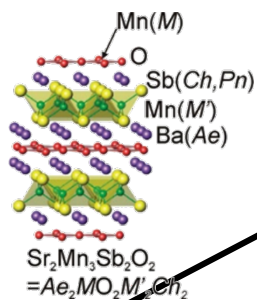
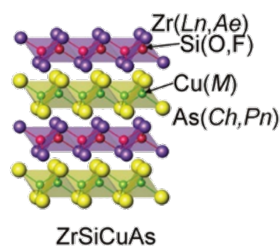
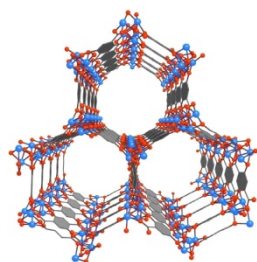
Assembling  
Building units  
(2000)

KNOWLEDGE-BASED  
(DIRECT)

Chemical decoration  
(1980)

Metals  
Close-packed oxides

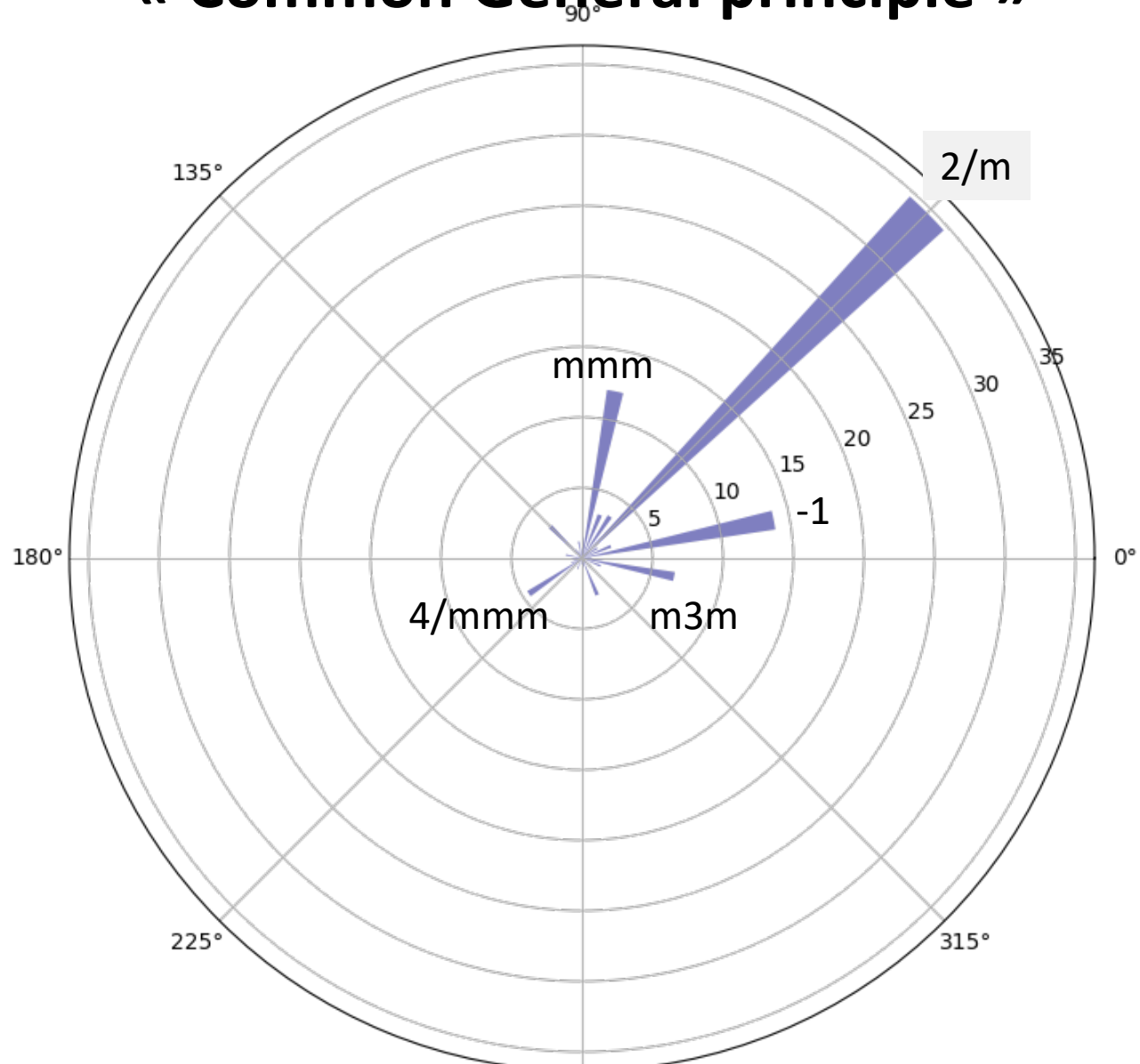
PRIOR  
KNOWLEDGE





280 000  
Inorganic  
Chemical compounds

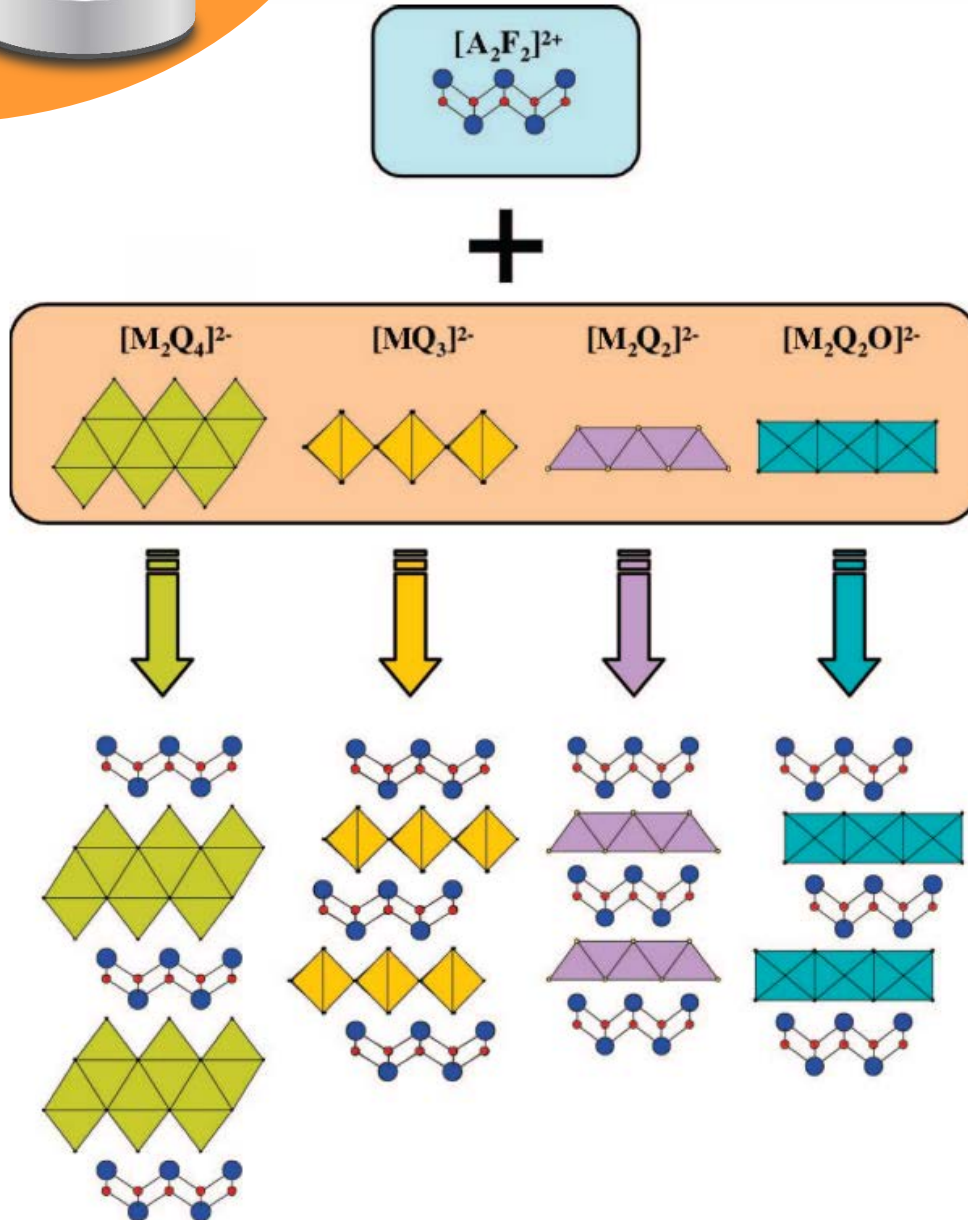
# « Common General principle »







# Assembling building units

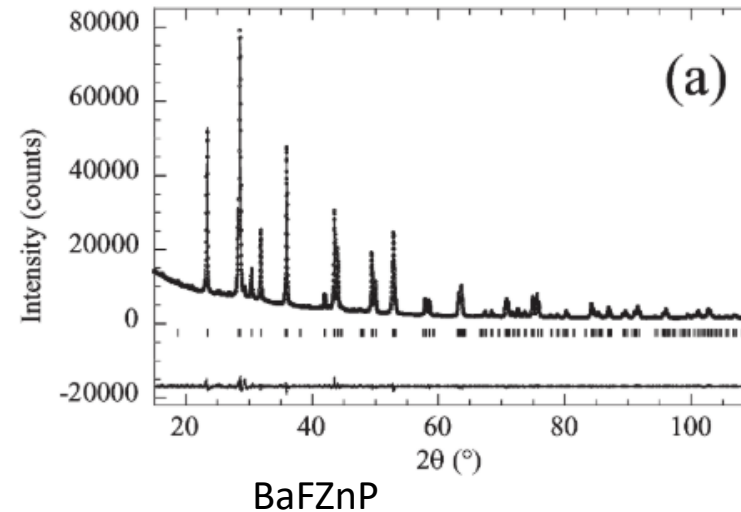


- (a) Ferey, G. J. Sol. State Chem. 2000, 152, 37.
- (b) Yaghi, O. M.; O'Keefe, M.; Ockwig, N. W.; Chae, H. K.; Eddaoudi, M.; Kim, J. Nature 2003, 423, 705.
- (c) Mellot-Draznieks, C.; Ferey, G. Curr. Opin. Solid State Mater. Sci. 2003, 7, 13

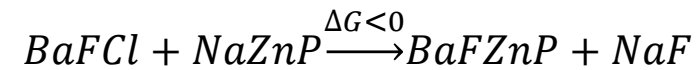
H. Kabbour, E. Janod, B. Corraze, M. Danot, C. Lee, M. Whangbo, and L. Cario, JACS, 2008, 130, 8261–8270

L. Cario, H. Kabbour, A. Meerschaut, Chem Mater. 17,2005, 234-236

H Kabbour, L. Cario, S. Jobic and B. Corraze J. Mater. Chem., 2006, 16, 4165–4169 | 4165



**Direct method**





# Correlation-based Data Mining

« The best way to have a good idea is to have lots of ideas »

Linus Pauling

Large 3D databases (3 530 330 in AFLOWlib)



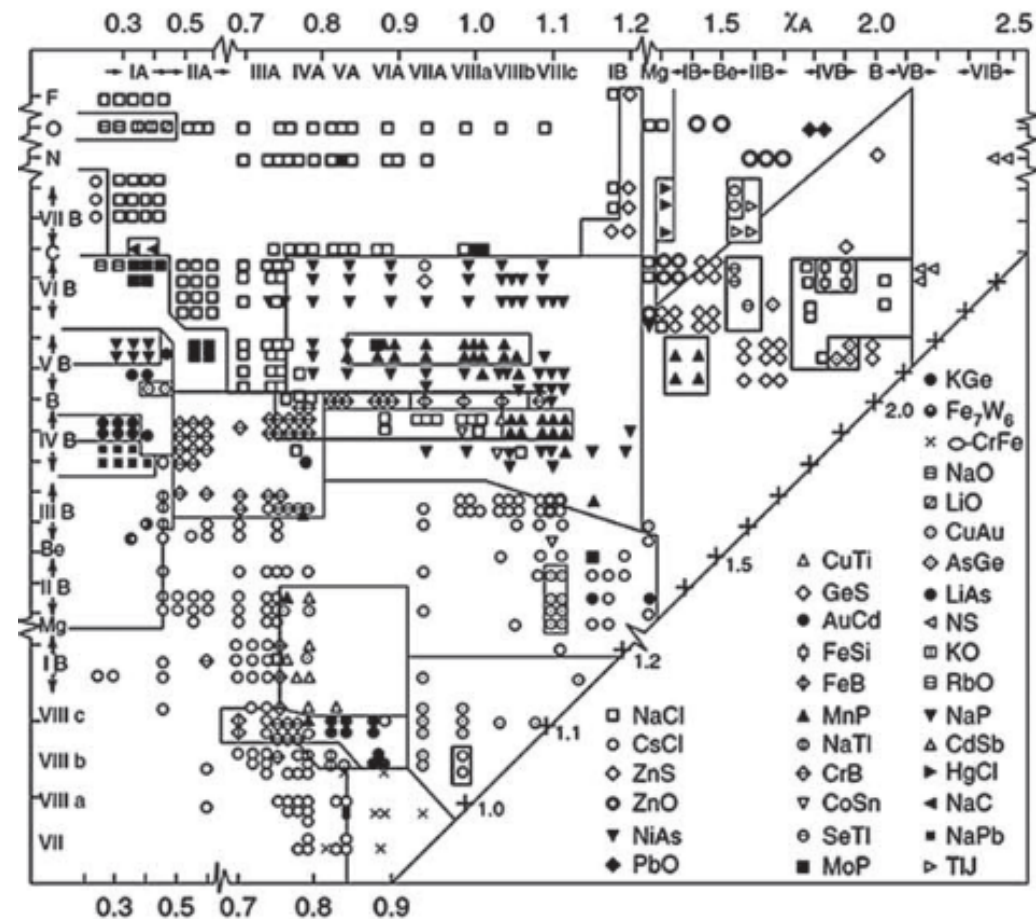
Small 2D DataBases (6351 materials, 2dmatpedia)

Materials Cloud two-dimensional crystals database (MC2D)

2D Materials Encyclopedia

MAGICS  
Materials Genome Innovation for Computational Software

UNREAL ENGINE 5

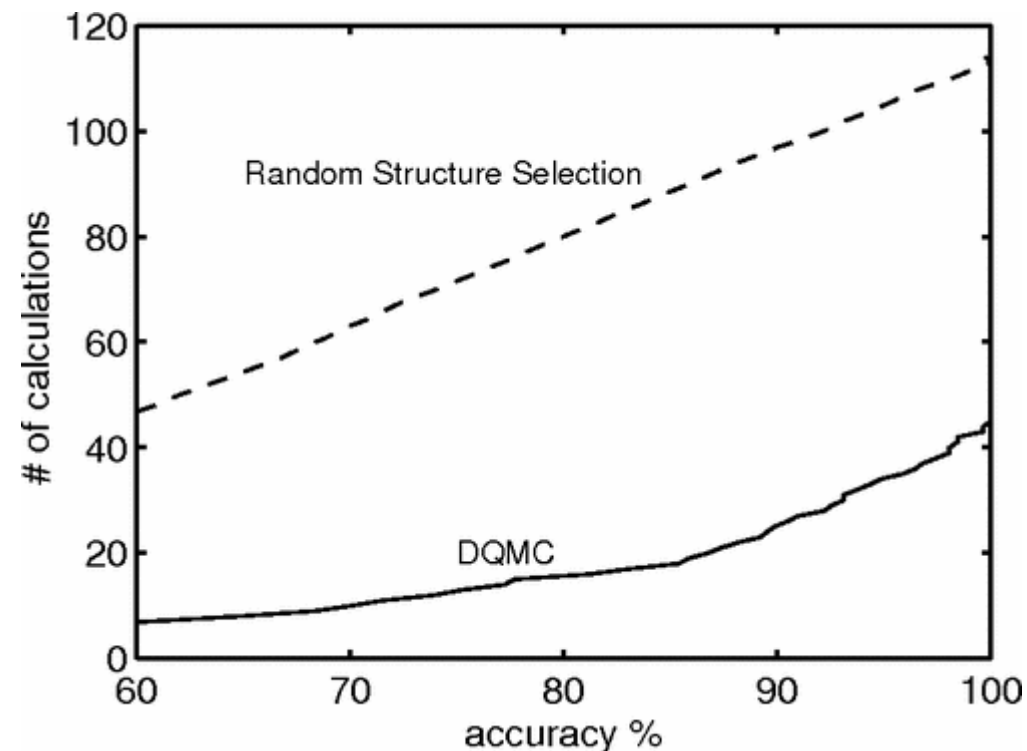
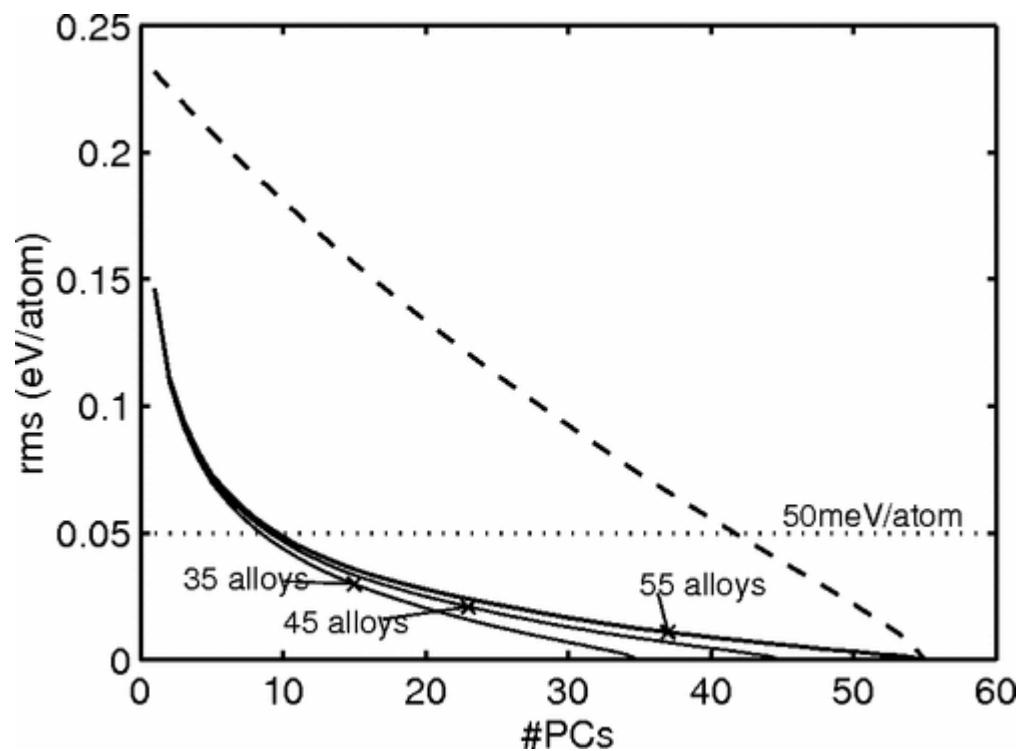


D. G. Pettifor, A chemical scale for crystal structure maps, *Solid State Commun.*, 1984, 51, 31–34.



# Data Mining approaches

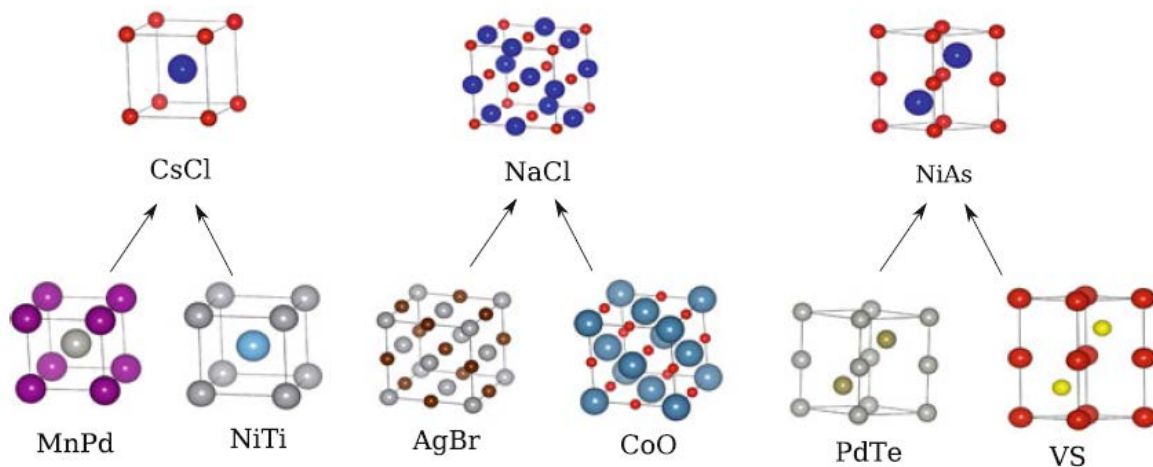
## ① Virtual *computed* prototypes: Principal component analysis: LiB





# Data Mining approaches

- ① **Virtual *computed*** prototypes: Principal component analysis: LiB
- ② **Classification based on *correlation*** in in ICSD prototypes:



LaMn<sub>2</sub>O<sub>5</sub> :  
- PrMn<sub>2</sub>O<sub>5</sub>  
- NdMn<sub>2</sub>O<sub>5</sub>  
- EuMn<sub>2</sub>O<sub>5</sub>  
- SmMn<sub>2</sub>O<sub>5</sub>  
- GdMn<sub>2</sub>O<sub>5</sub>

(see [Materialsproject.org](http://Materialsproject.org))

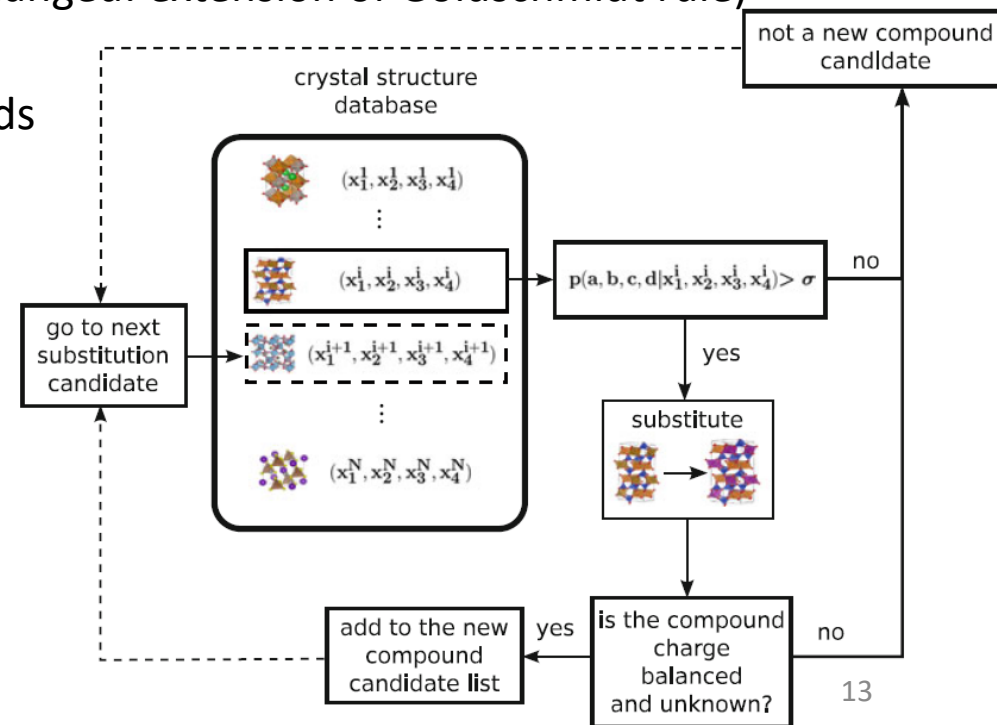


# Data Mining approaches

- ① **Virtual *computed*** prototypes: Principal component analysis : LiB
- ② **Classification** based on **correlation** in in ICSD prototypes
- ③ Data Mined Ionic **Substitution** Model: a probabilistic model (likelihood) that predicts ionic substitutions which keep the crystal structure of a compound unchanged: extension of Goldschmidt rule)

first superconducting pnictide oxide  
 $\text{LaFeAsO}_{1-x}\text{F}_x$

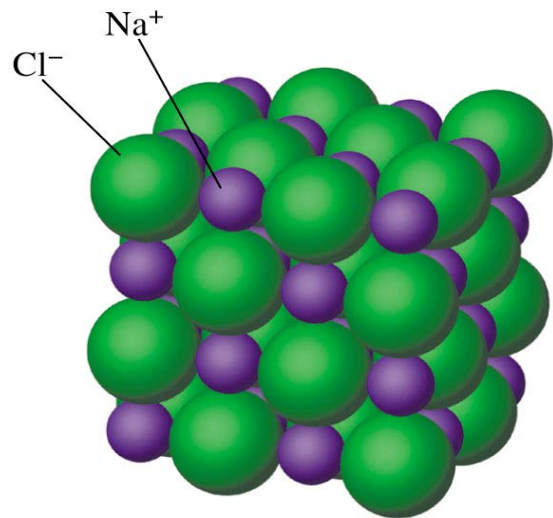
Isostructural new compounds  
 $\text{SmLaFeAsO}_{1-x}\text{F}_x$



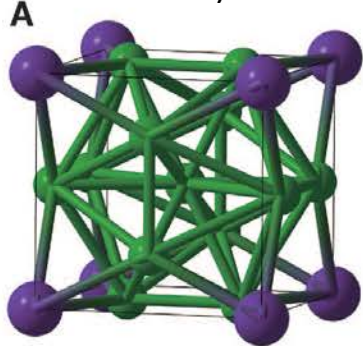


# What if the underlying lattice is not known?

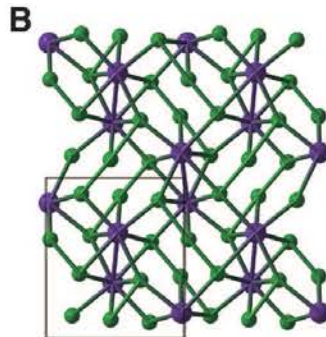
Pressure induced phases: computationally predicted and (some) later confirmed experimentally



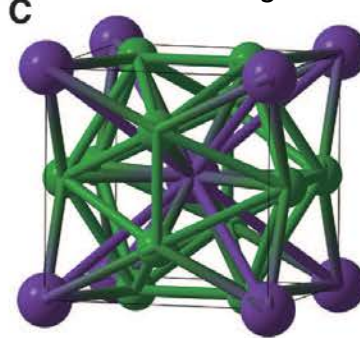
A Pm3-NaCl<sub>7</sub>.



B Pnma-NaCl<sub>3</sub>



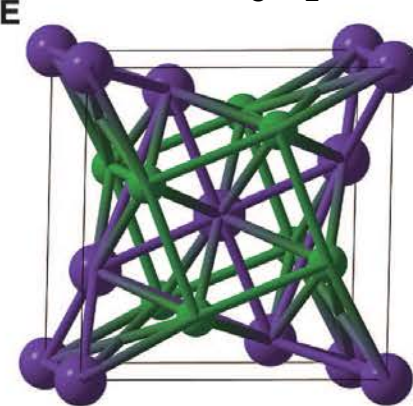
C Pm3n-NaCl<sub>3</sub>.



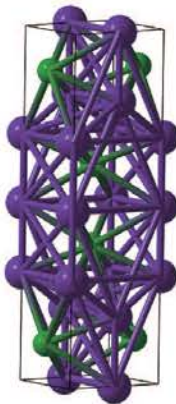
D P4/mmm-Na<sub>3</sub>Cl.



E P4/m-Na<sub>3</sub>Cl<sub>2</sub>.

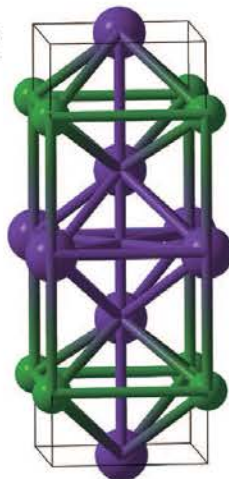


F



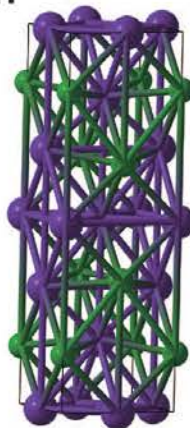
Cmmm-Na<sub>3</sub>Cl<sub>2</sub>.

G



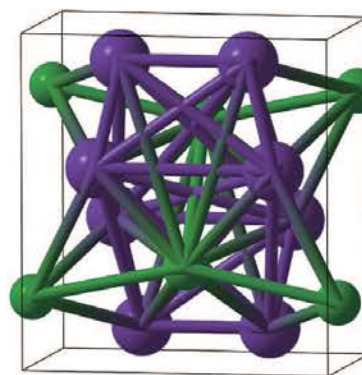
P4/mmm-Na<sub>2</sub>Cl

H

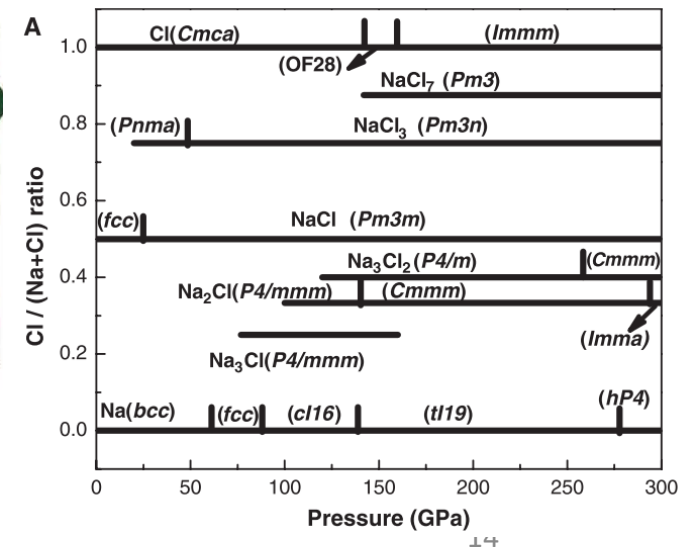


Cmmm-Na<sub>2</sub>Cl.

I



Imma-Na<sub>2</sub>Cl

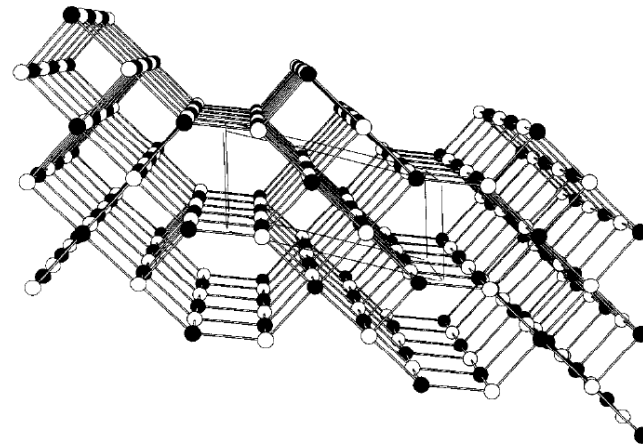
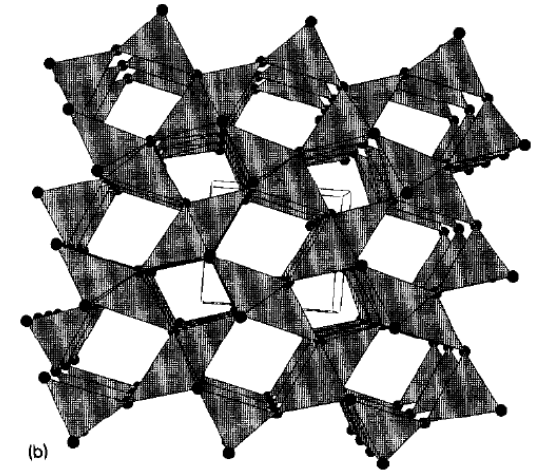
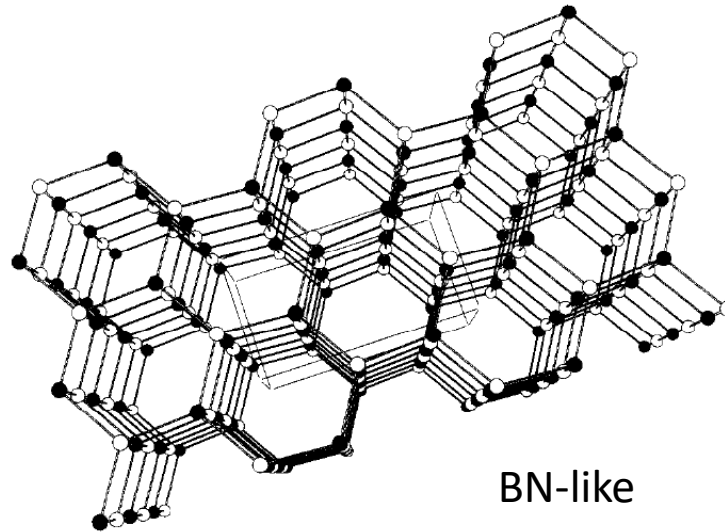
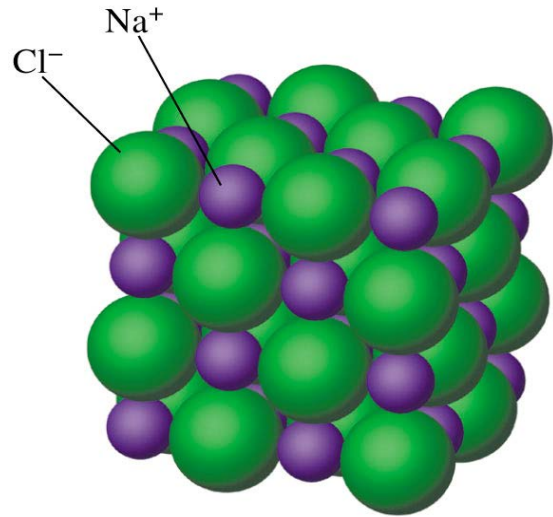




# What if the underlying lattice is not known?

*Metastable phases computationally predicted and (some) later confirmed experimentally*

## **Metastable compounds**



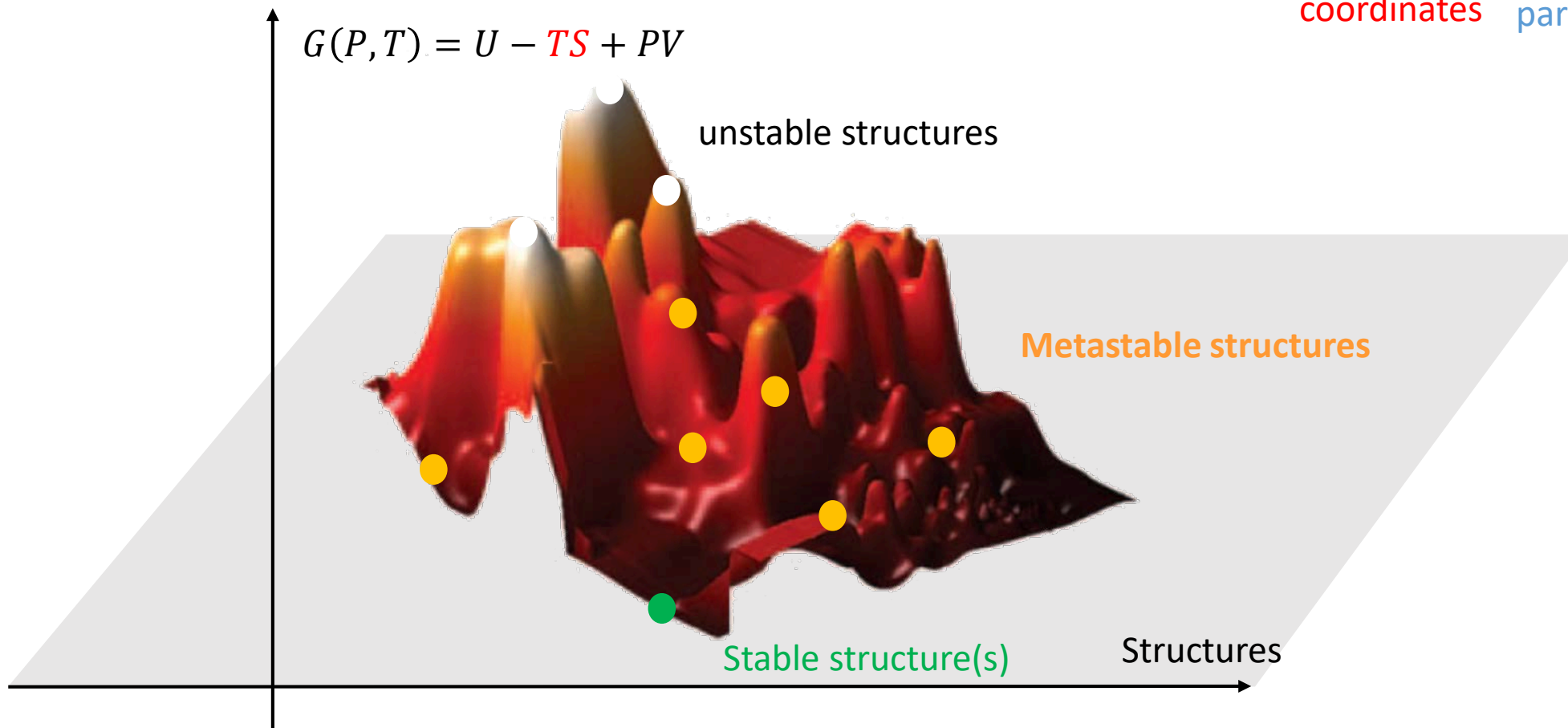
Schön, JC, Jansen, M. *Comp. Mater. Sci* **4** (1995) 43-58

Requires a method functioning with **no prior knowledge** to allow unexpected discoveries

# Exploration of the rugged Potential Energy Surface (PES)

$$d = 3N - 3 + 6$$

atom coordinates      Cell parameters



2D-projection of the reduced landscape of  $\text{Au}_8\text{Pd}_4$  (Oganov, Modern Methods of CSP, Wiley, 2011)

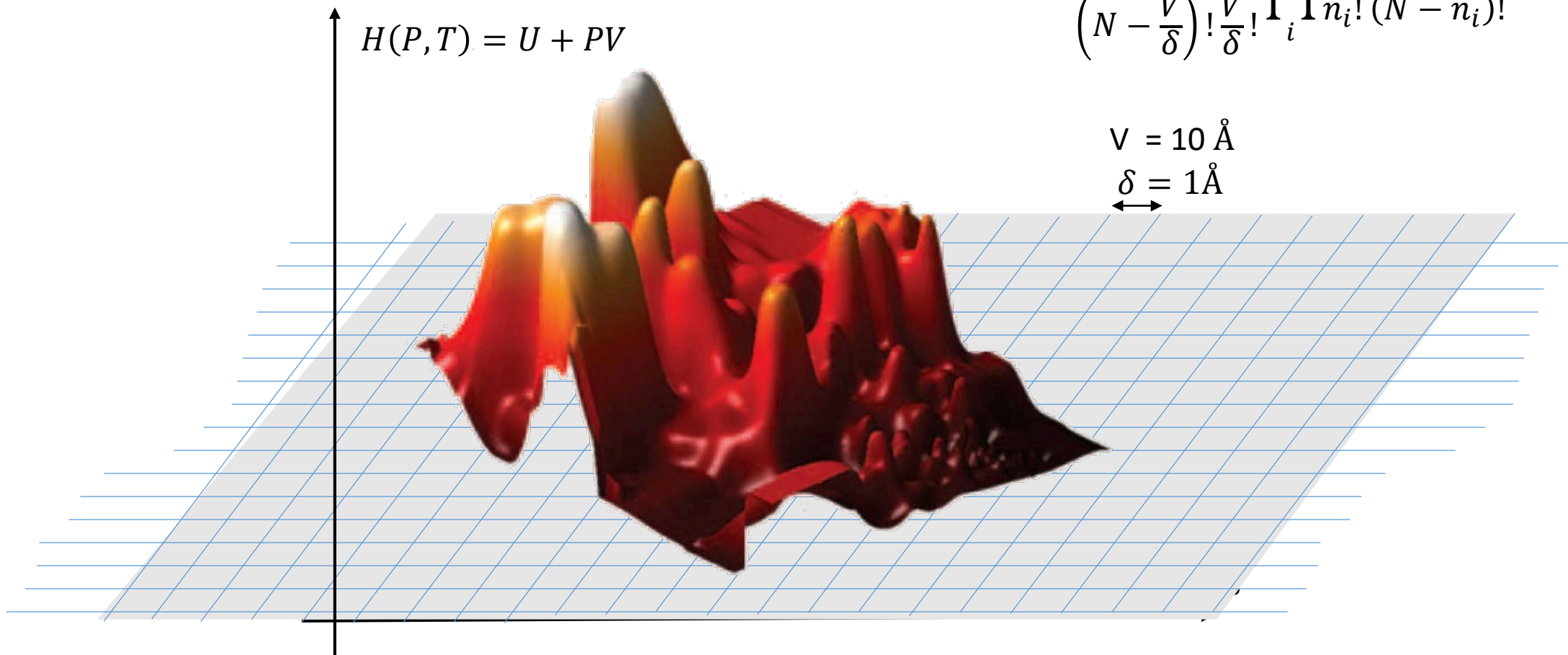


# Exhaustive exploration is a NP-hard problem

$$C = \frac{N!}{\left(N - \frac{V}{\delta}\right)! \frac{V}{\delta}!} \prod_i \frac{N!}{n_i! (N - n_i)!} = 10^N$$

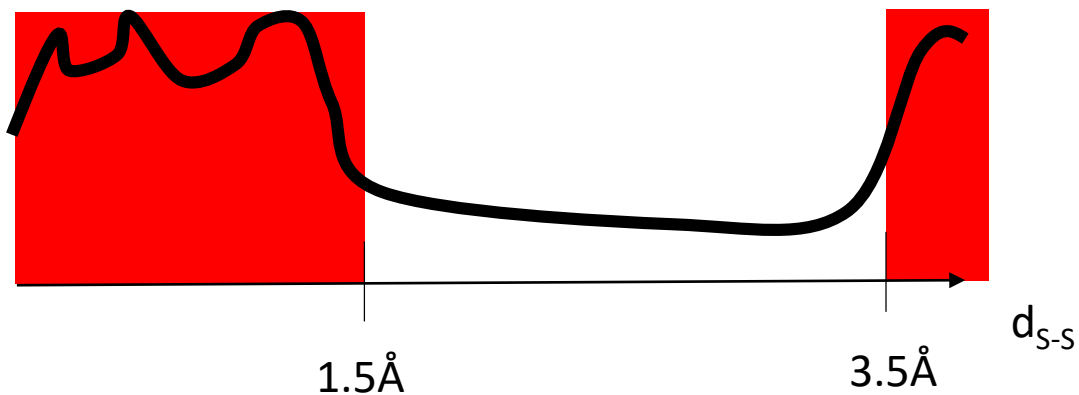
$$H(P, T) = U + PV$$

$$v = 10 \text{ \AA}$$
$$\delta = 1 \text{ \AA}$$

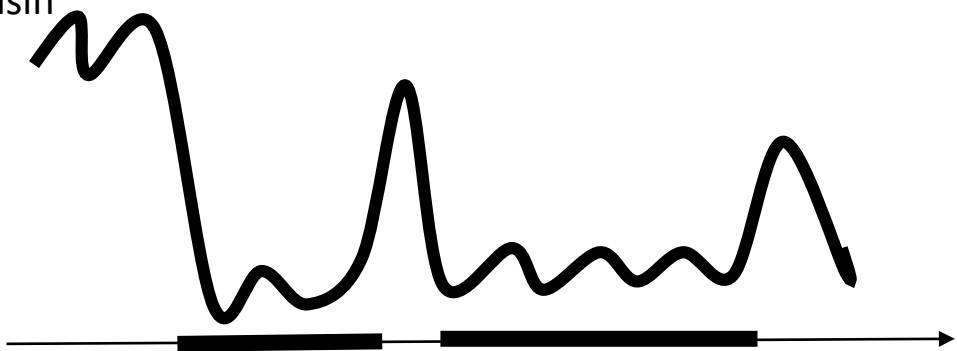


# Why does it work ?

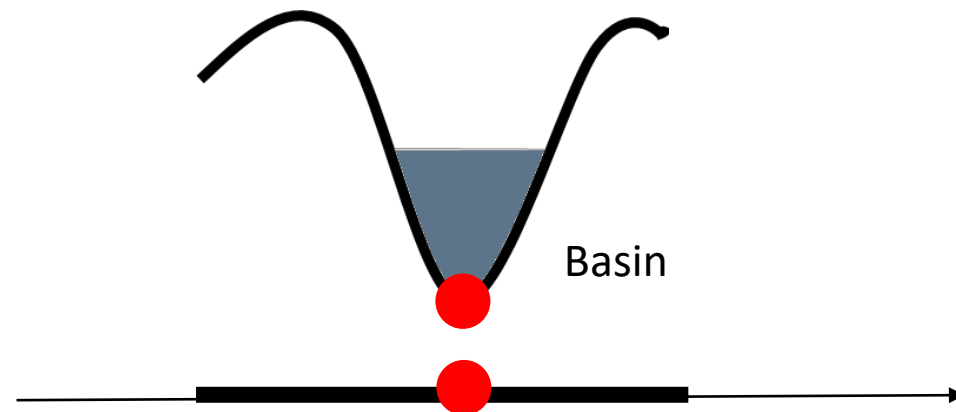
- Only certain regions of the PES are **chemically relevant** (bond types, lengths..) No need to explore chemically unreasonable regions.



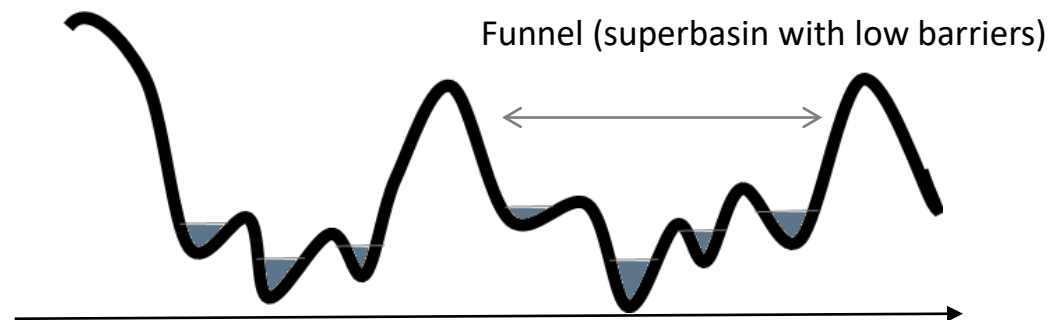
- Low-energy basins occupy **the largest amount of** the PES : a randomly generated structure has a high probability to fail in the basin



- **Relaxation** decreases the dimensionality of the PES

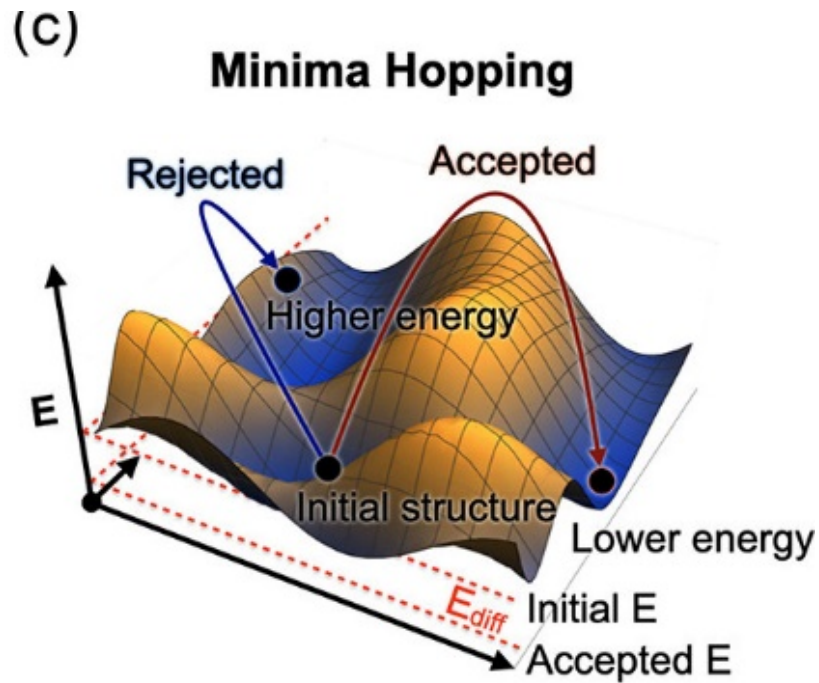


- **Bell Evans Polanyi principle**: barriers between low-lying minima in a PES are expected to be small. Low-energy basins are likely to be close to one another.

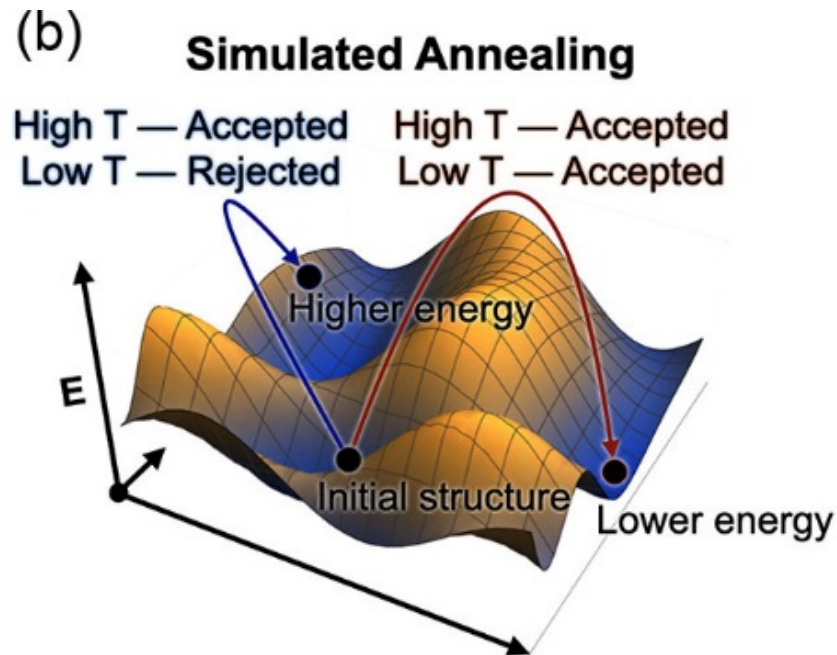


# Global optimization techniques

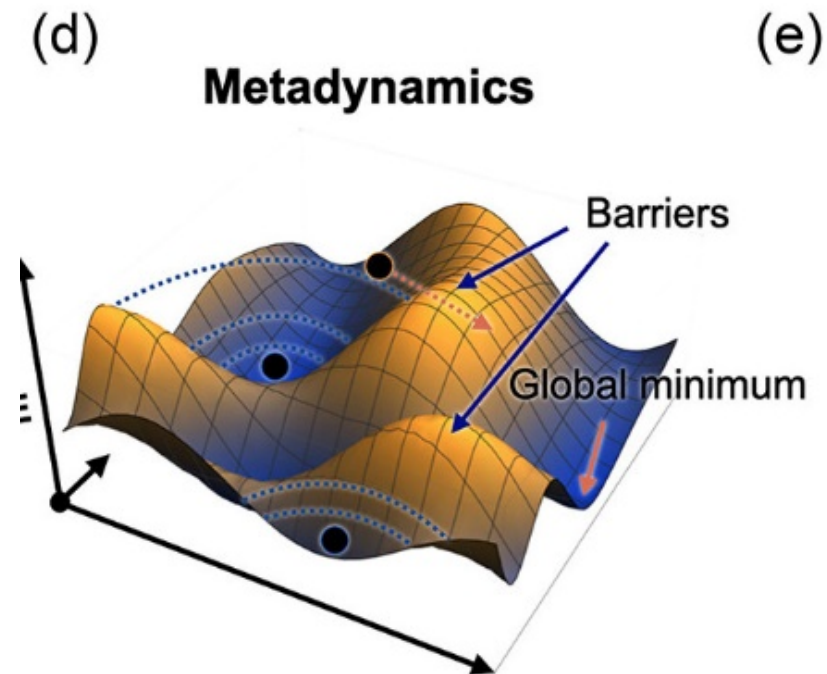
When a good guess of the structure is known



superconducting  $S_xSe_{(1-x)}H_3$  phases



metal pernitrides, BN,  $GeF_2$



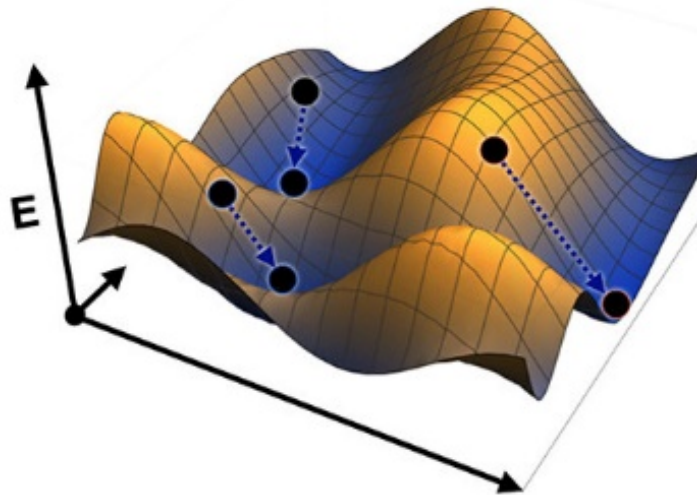
Carbon under pressure

*Mainly used as accelerating tools in current softwares*

# Global optimization techniques

When nothing or little about the structure is known

(a) **Random Searching**



Battery materials

```
.o.  oooo ooooooo.  .ooooo.  .oooooo.  .o
.ooo.  *ooo* *oo *oo.  ooo.  *oo *oo
.oooo.  ooo  ooo  ooo*  ooooo.  ooooo.
*o*  ooo.  ooo  ooooooo*  *oooo.  *oooo.
.ooooooo.  ooo  ooo*  ooo.  *oooo  *oooo
oT  ooo.  ooo  ooo  *ooo  oo  ooo  oo  ooo
oT  oooo  oooo  oooo  oooo  o:ooooo*  o:ooooo*
oT  ooo  oooo  oooo  oooo  o:ooooo*  o:ooooo*
  An Initial Random Structure Searching
```

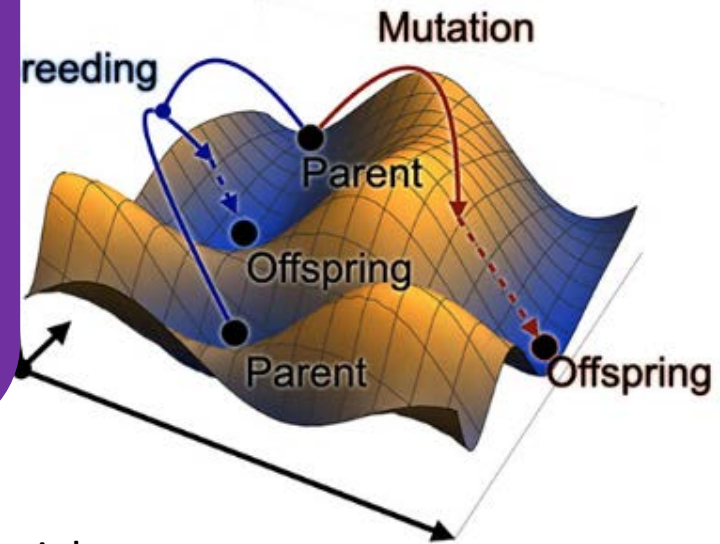
(e)

**No Free Lunch Theorem** : any searching algorithm that performs well on one class of problems will perform poorly on another class.

No way to confirm the **global minimum** has been found unless all local minima have been explored.

(f)

**Genetic Algorithm**



clusters,  
2D layered materials,  
solids, electrides  
and superhard  
compounds.



# Evolutionary Algorithm Variation Operators

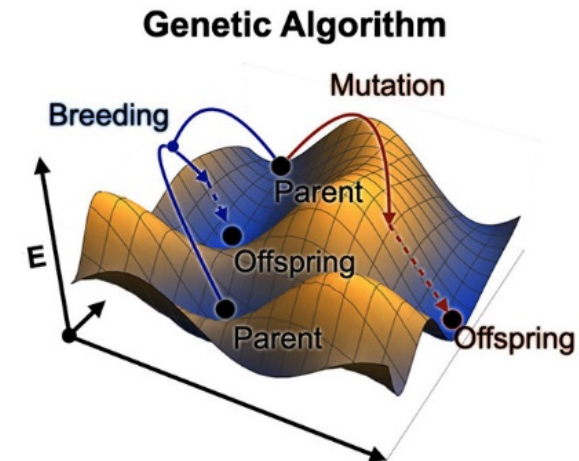
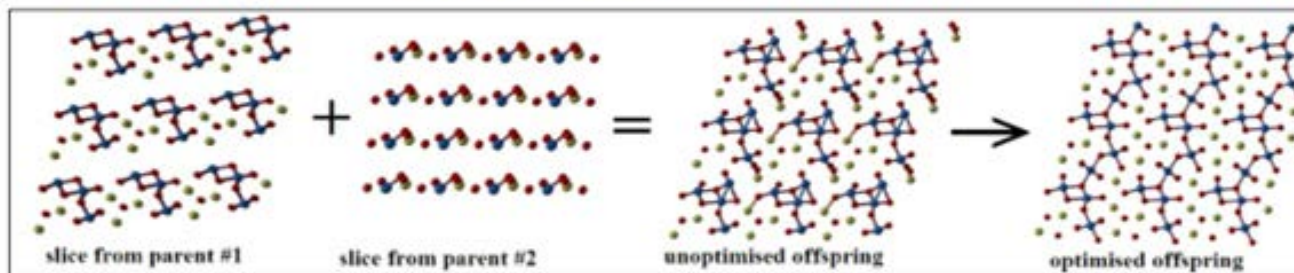


Generation of a **Diverse** Population of Structures



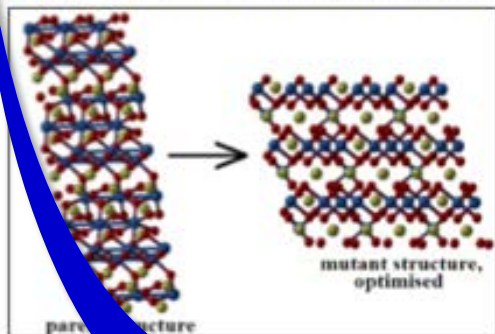
« **Self-Learning** » tool

## (1) Heredity (crossover)

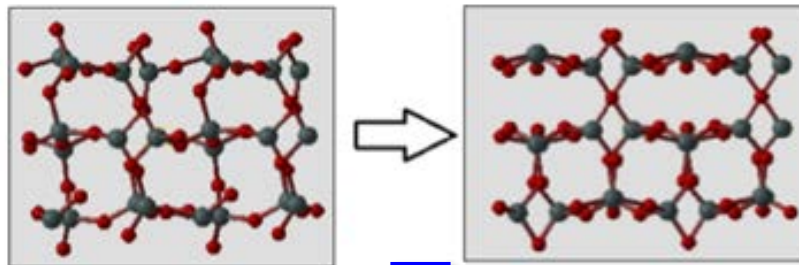


## Mutation operators - Enhanced Search

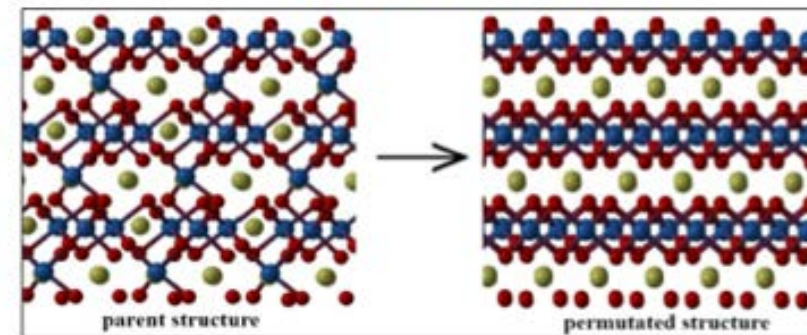
### (2) Lattice mutation



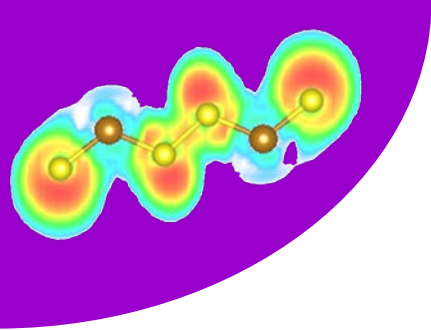
### (3) Softmutation



### (4) Permutation



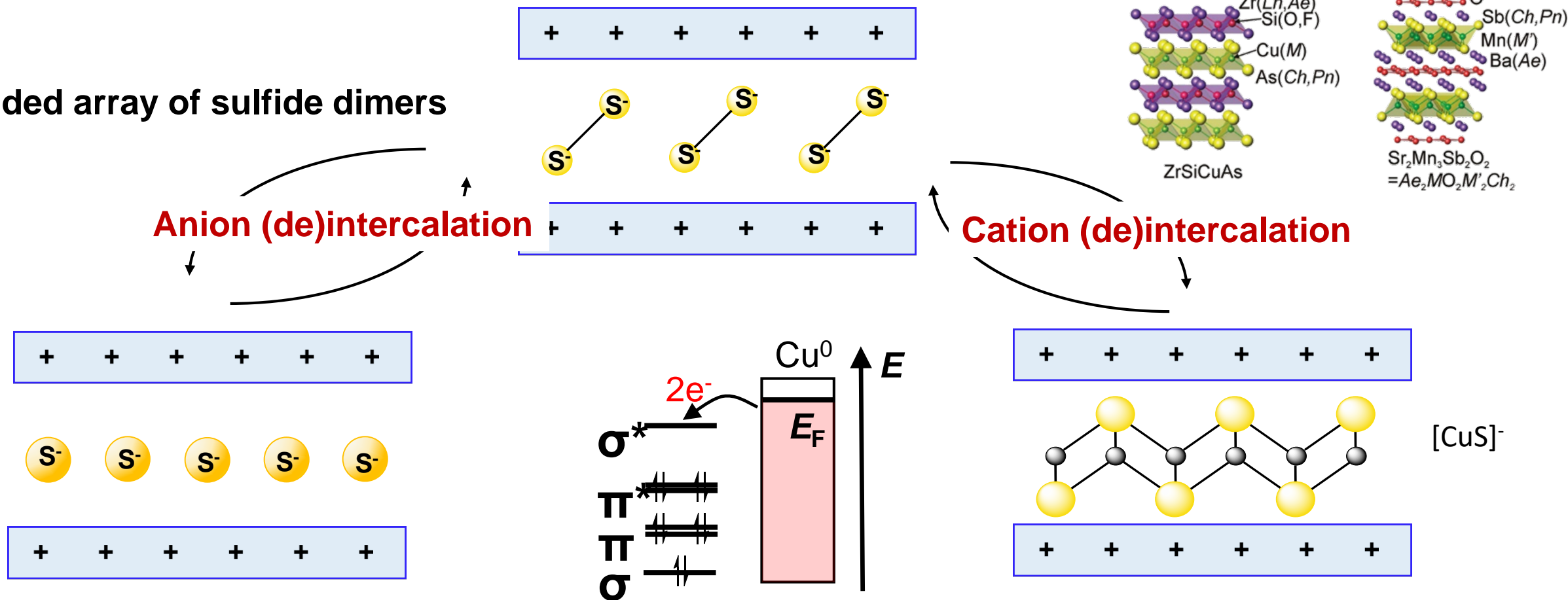
Eugenic selection of the fittest (survivors) and elimination of the others



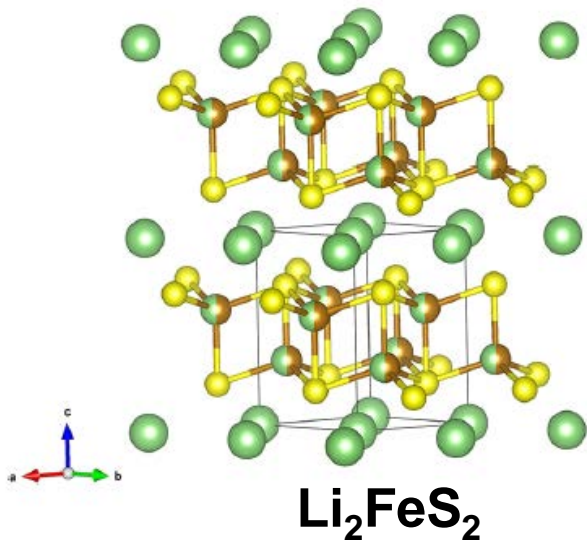
# Topochemical syntheses driven by redox-active (sulfide) anion

*Pseudo-lamellar host lattice* :  $[\text{LaO}]^+$ ,  $[\text{Sr}_2\text{MnO}_2]^{2+} \dots$

Extended array of sulfide dimers



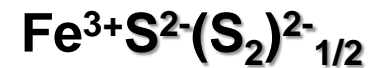
# Structure of twice-delithiated $\text{Li}_2\text{FeS}_2$ ?



-Li



*Structure unknown  
for 30 years*



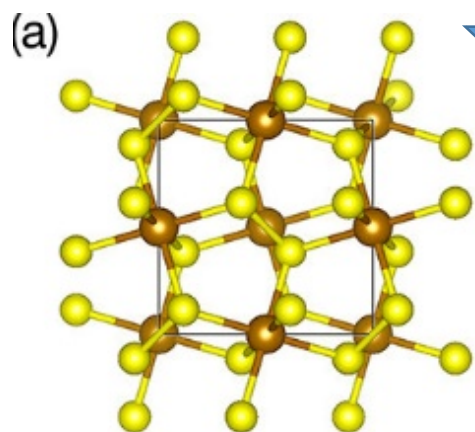
- **IR spectroscopy :**
  - $d(\text{S}-\text{S}) = 2.12 \text{ \AA}$
- **Mössbauer and EXAFS :**
  - $\text{Fe}^{3+}\text{HS}$  ion in **tetrahedral sites**
  - $d(\text{Fe}-\text{S}) = 2.28 \text{ \AA}$
  - $\text{S}^{2-}/\text{S}_2^{2-} = 2/1$
- **Lamellar structure** expected due to the stacking of  $\text{LiFeS}_2$  and Li layers.



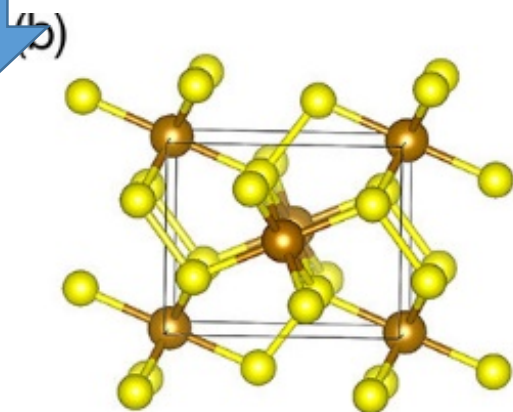
# Structure of twice-delithiated $\text{Li}_2\text{FeS}_2$ ?

3D

A collection of logos for 3D materials databases and software. From left to right: ICSD (Inorganic Crystal Structure Database), The Materials Project, AFLOW (Automatic - FLOW for Materials Discovery), NOMAD (NOVEL MATERIALS DISCOVERY), OQMD (The Open Quantum Materials Database), and MATERIALSCLOUD (with sub-logos for LEARN, WORK, DISCOVER, EXPLORE, ARCHIVE).



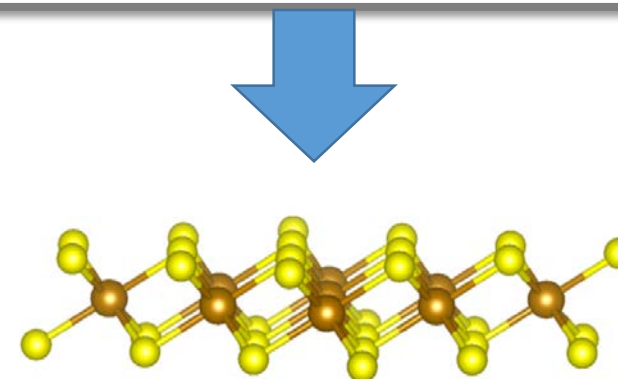
NM Pyrite



NM Marcasite

2D

A collection of logos for 2D materials databases and software. From top to bottom: Materials Cloud two-dimensional crystals database (MC2D), 2D Materials Encyclopedia, MAGICS (Materials Genome Innovation for Computational Software), UNREAL ENGINE 5, Stanford | ENGINEERING Materials Science &amp; Engineering, and MATDAT.



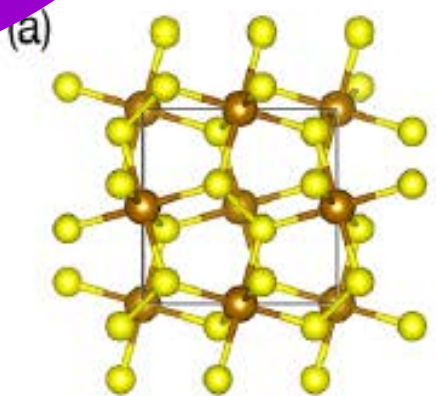
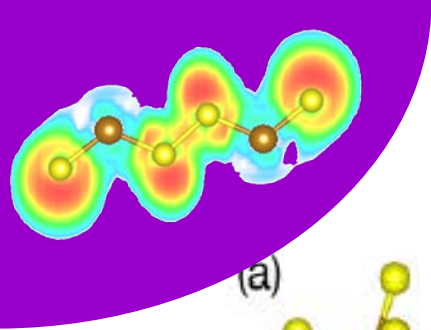
FM  $P-3m1$  1T-2D

( $\text{MoS}_2$ -like)

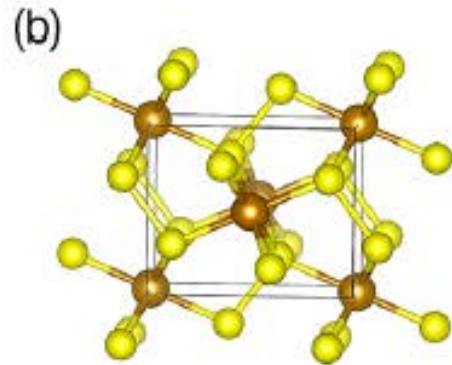




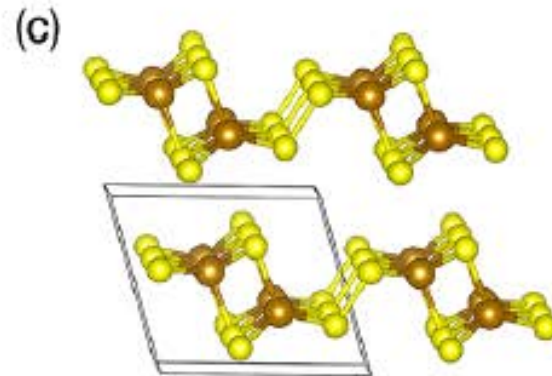
# Crystal Structure Prediction of FeS<sub>2</sub>



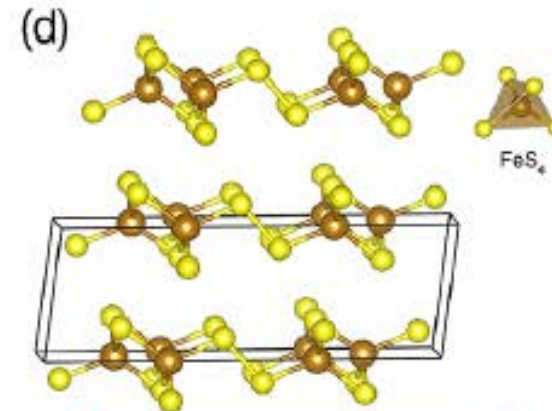
NM Pyrite  
 $\Delta E = 0.000$  eV/f.u.  
(Ground state)



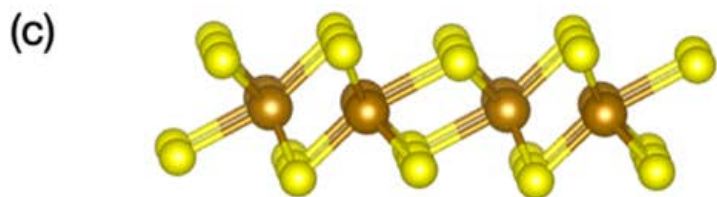
NM Marcasite  
 $\Delta E = 0.009$  eV/f.u.



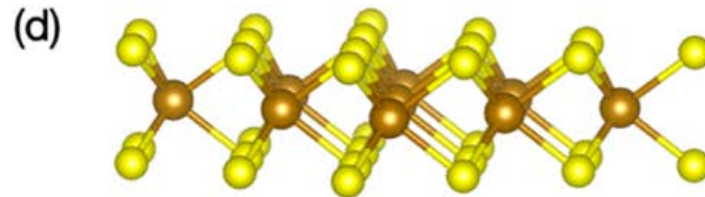
AFM *P*-1-3D  
 $\Delta E = 0.282$  eV/f.u.



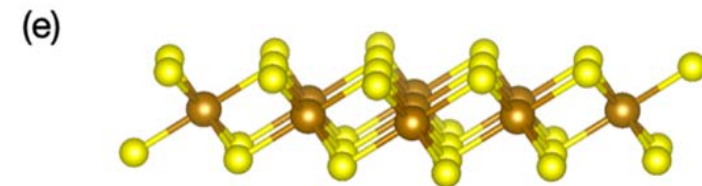
AFM *C*2/*m*-3D  
 $\Delta E = 0.425$  eV/f.u.



AFM *P*2<sub>1</sub>/*m* 1T'-2D  
 $\Delta E = 0.713$  eV/f.u.



FM *P*-6*m*2 1H-2D  
 $\Delta E = 1.254$  eV/f.u.

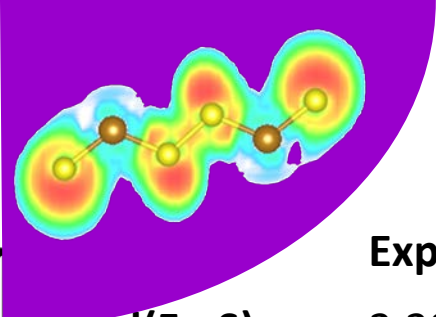


Dynamical unstable

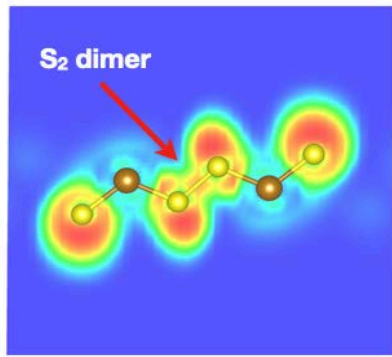


FM *P*-3*m*1 1T-2D  
 $\Delta E = 0.777$  eV/f.u.

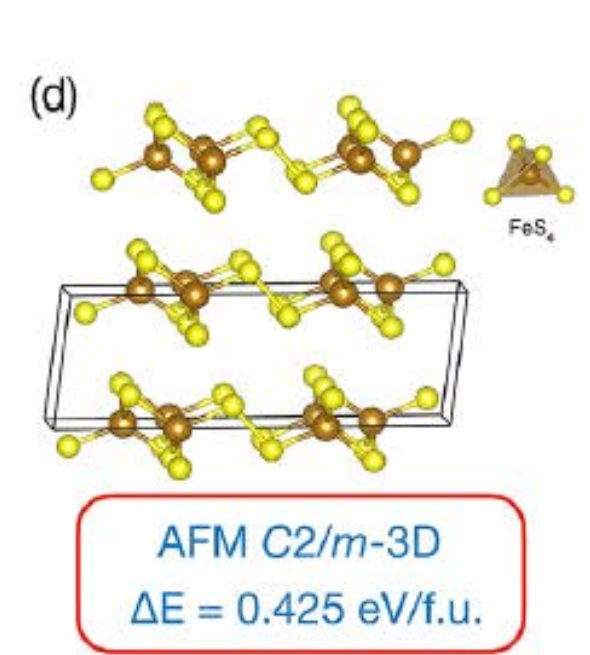
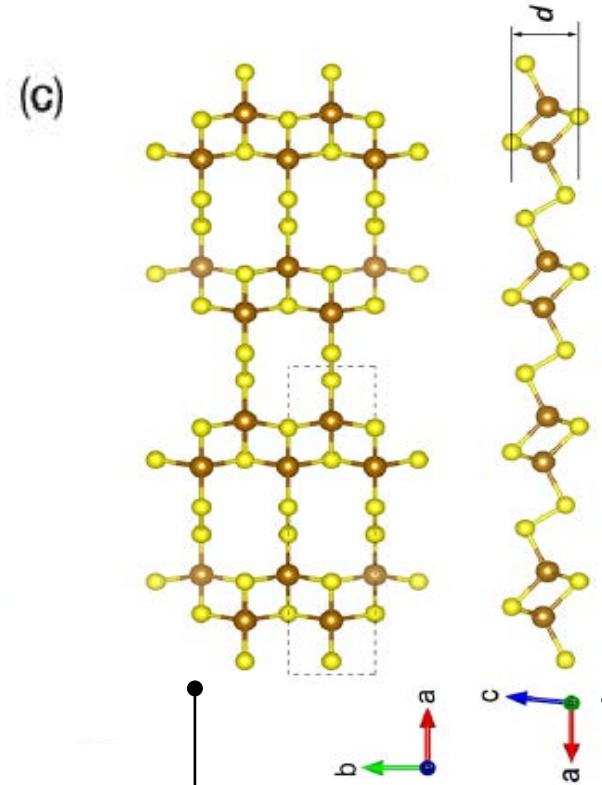
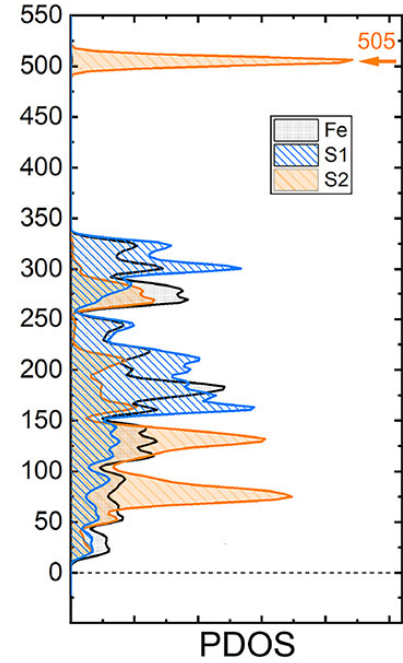
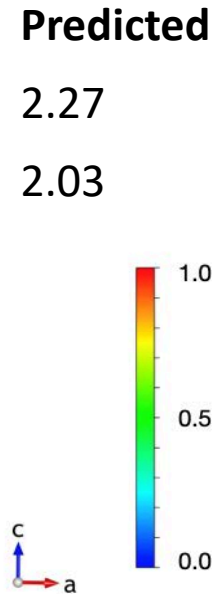
# Crystal Structure Prediction of FeS<sub>2</sub>



	Exp.	Predicted
d(Fe-S)	2.28	2.27
d(S-S)	2.12	2.03



C2/m-2D FeS<sub>2</sub>

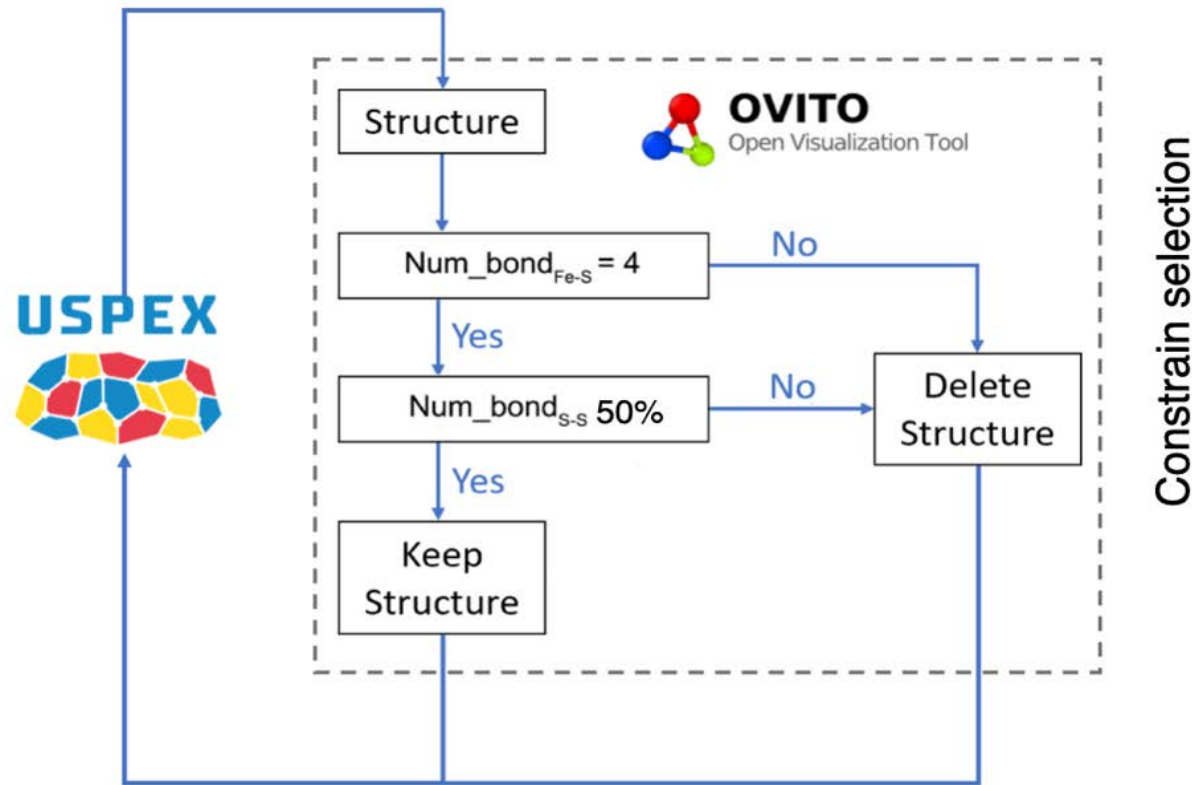
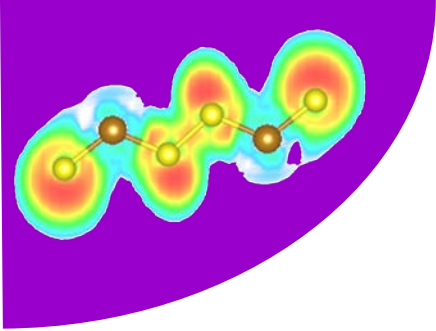


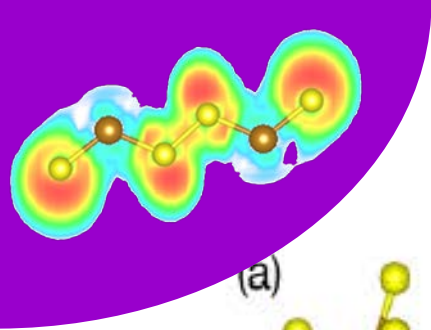
Corrugated 2/∞[FeS<sub>2</sub>] layer made of blocks linked to each other via S-S bonds

1/∞[Fe<sub>2</sub>S<sub>4</sub>] ribbons consisting of edge-sharing FeS<sub>4</sub> tetrahedra with zigzag iron rows

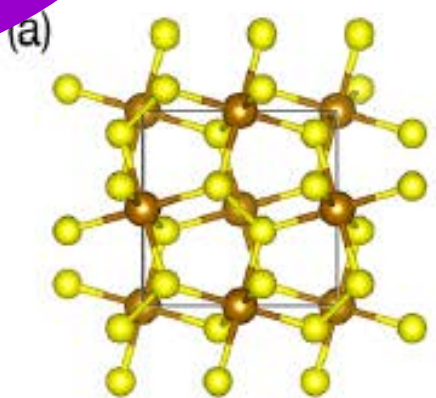
**Strongly differs from classical MS<sub>2</sub> slabs !**

# Constrained Crystal Structure Prediction of FeS<sub>2</sub>

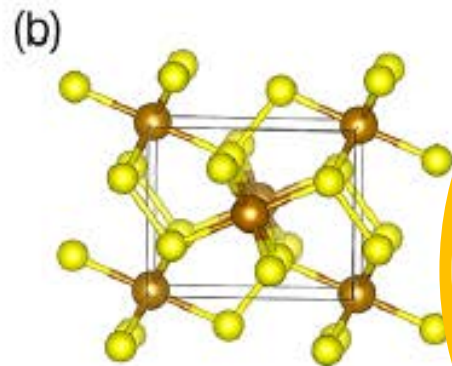




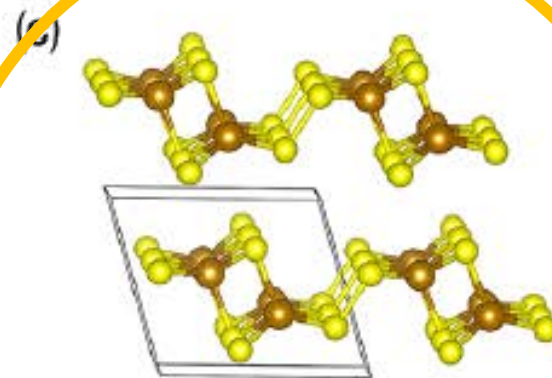
# What about the newly unraveled and more stable FeS<sub>2</sub> phase ?



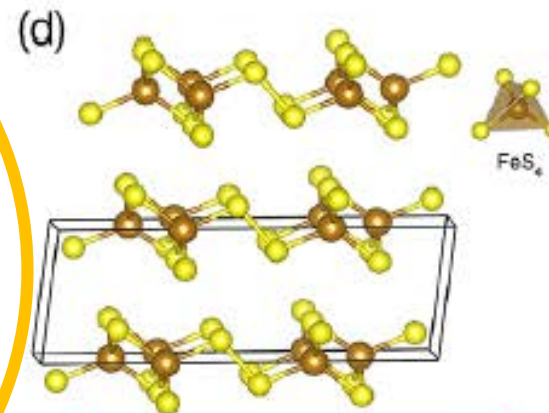
NM Pyrite  
 $\Delta E = 0.000$  eV/f.u.  
 (Ground state)



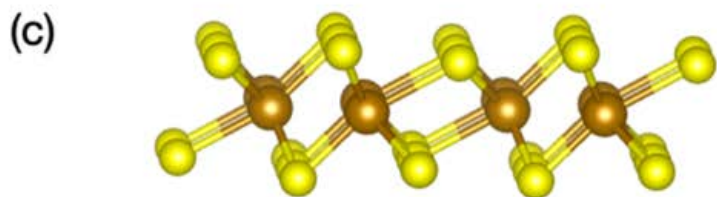
NM Marcasite  
 $\Delta E = 0.009$  eV/f.u.



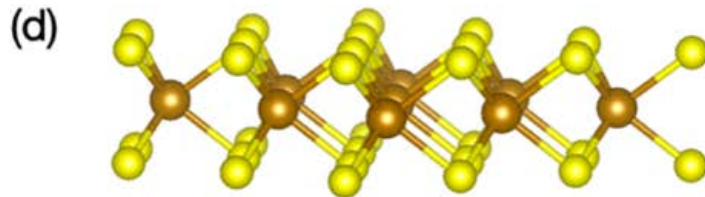
AFM *P-1-3D*  
 $\Delta E = 0.282$  eV/f.u.



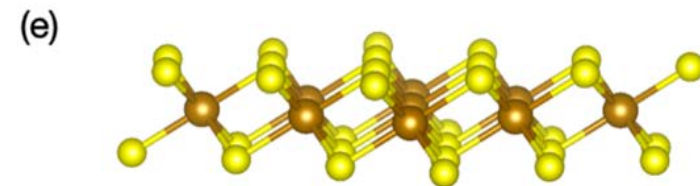
AFM *C2/m-3D*  
 $\Delta E = 0.425$  eV/f.u.



AFM *P2<sub>1</sub>/m 1T'-2D*  
 $\Delta E = 0.713$  eV/f.u.



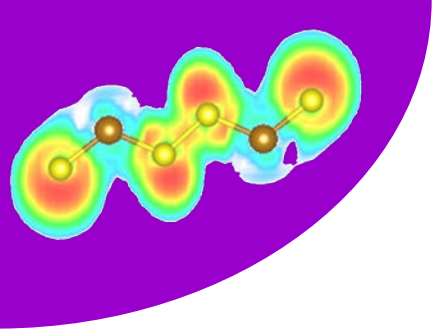
FM *P-6m2 1H-2D*  
 $\Delta E = 1.254$  eV/f.u.



Dynamical unstable



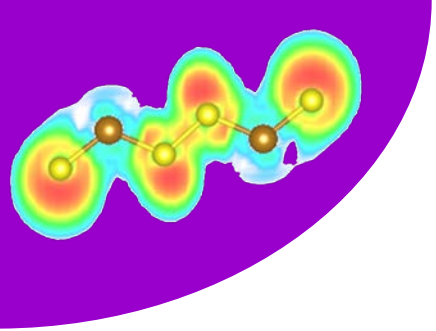
FM *P-3m1 1T-2D*  
 $\Delta E = 0.777$  eV/f.u.



# Is it really compressing time if there is no synthesability criterion ?



Our approach : *coupling* advances in *computational* material and material *synthesis* design  
by selecting one synthesis route



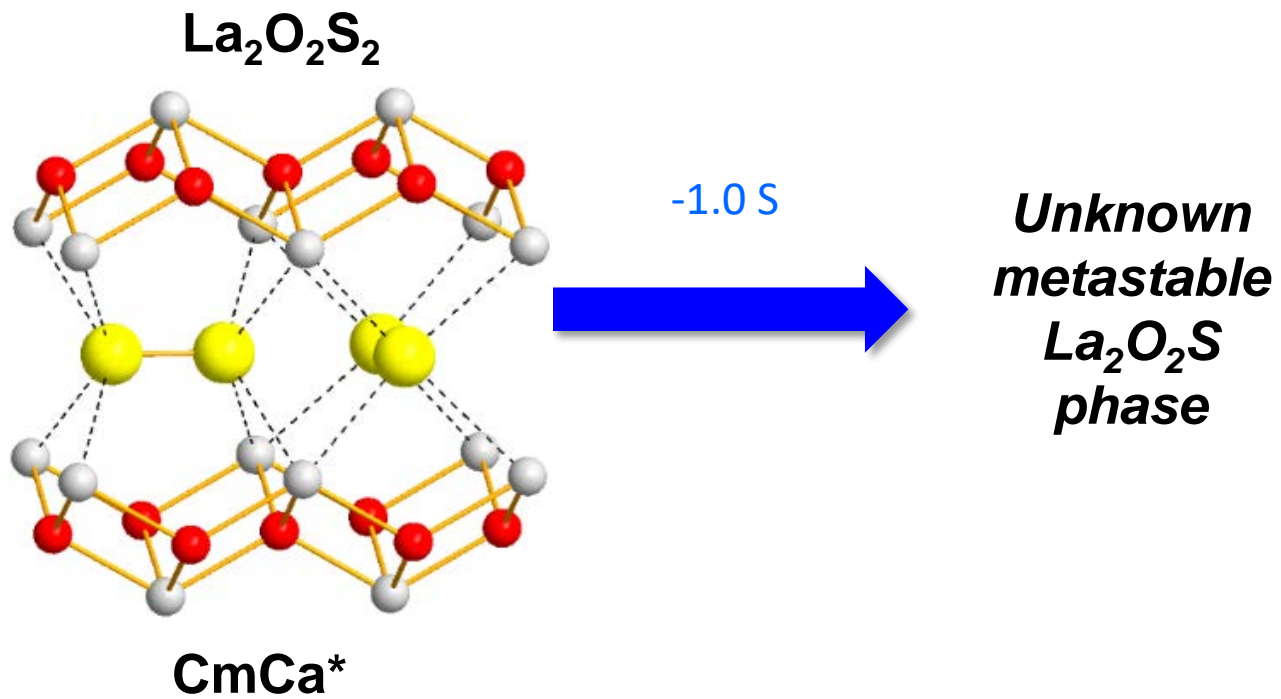
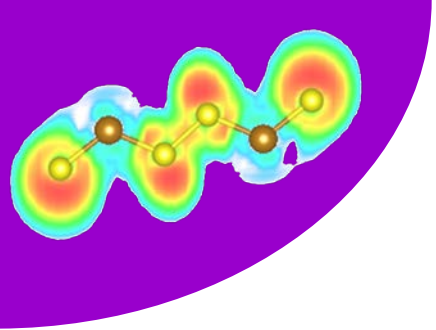
## Still a need to predict ?



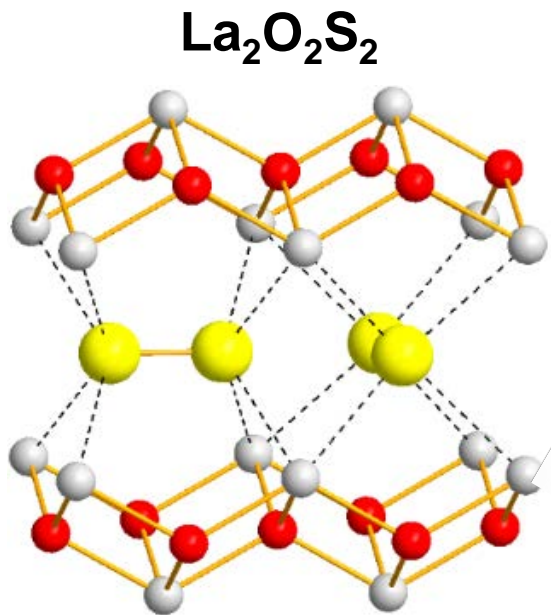
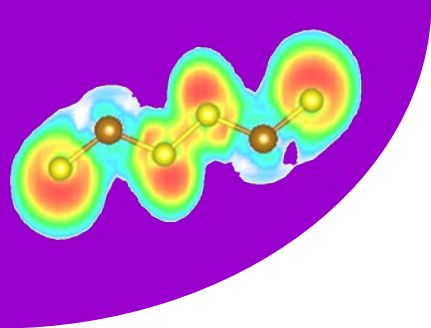
***Structure characterisation  
of metastable samples  
is not a trivial task***



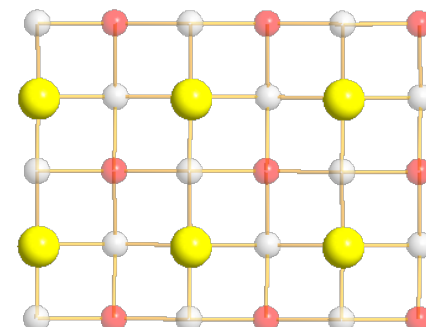
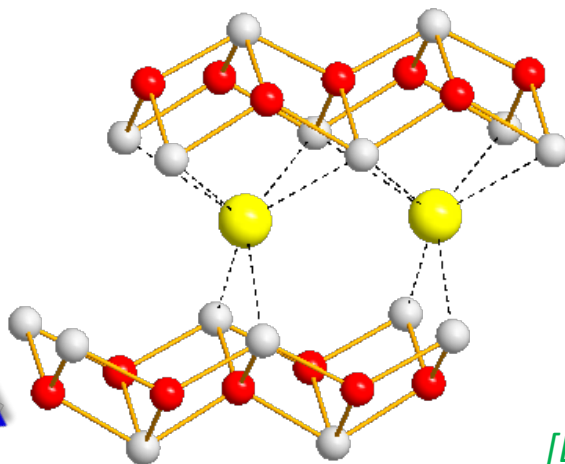
# Topochemical S (de)intercalation of $\text{La}_2\text{O}_2\text{S}_2$



# S (de)intercalation of $\text{La}_2\text{O}_2\text{S}_2$



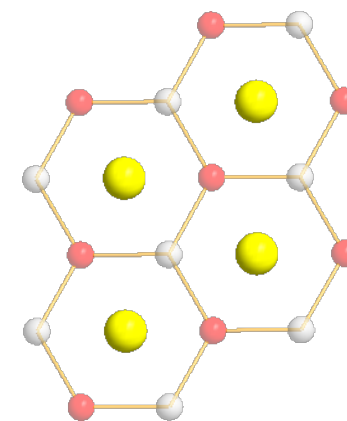
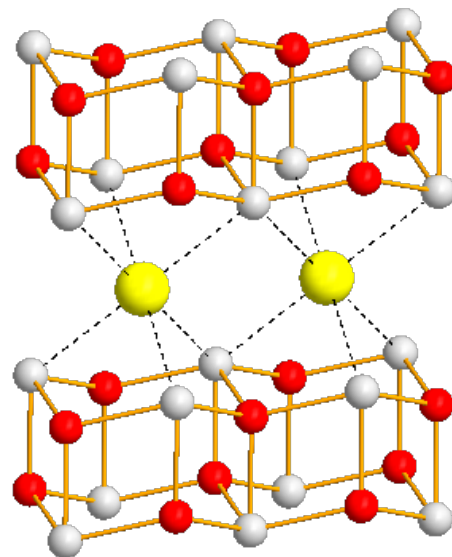
- 1.0 S



$[\text{LaO}_4\text{S}_2]$  and  $[\text{LaO}_4\text{S}_4]$  polyhedra

*Predicted Pattern*  
*oA-La<sub>2</sub>O<sub>2</sub>S*  
*Metastable*  
*(SG: Amm2)*

$\Delta E=72\text{meV}$



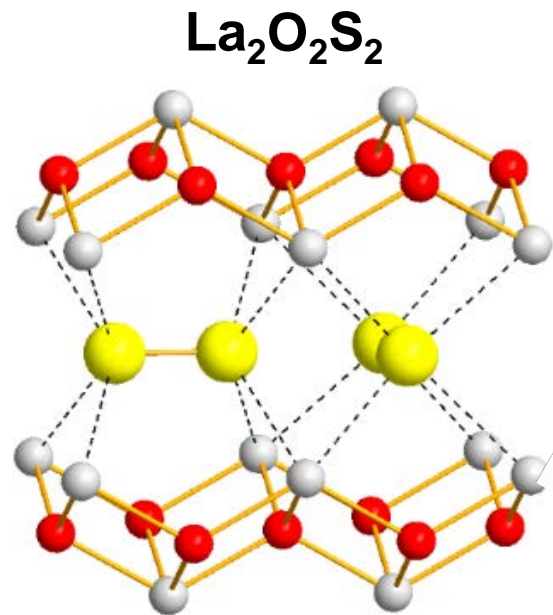
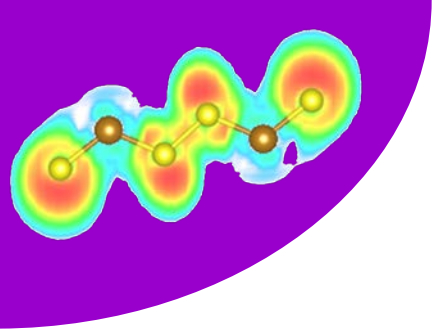
$[\text{LaO}_3\text{S}_3]$  polyhedra only

Morosin et al.  
*Acta Cryst. B*  
(1973)

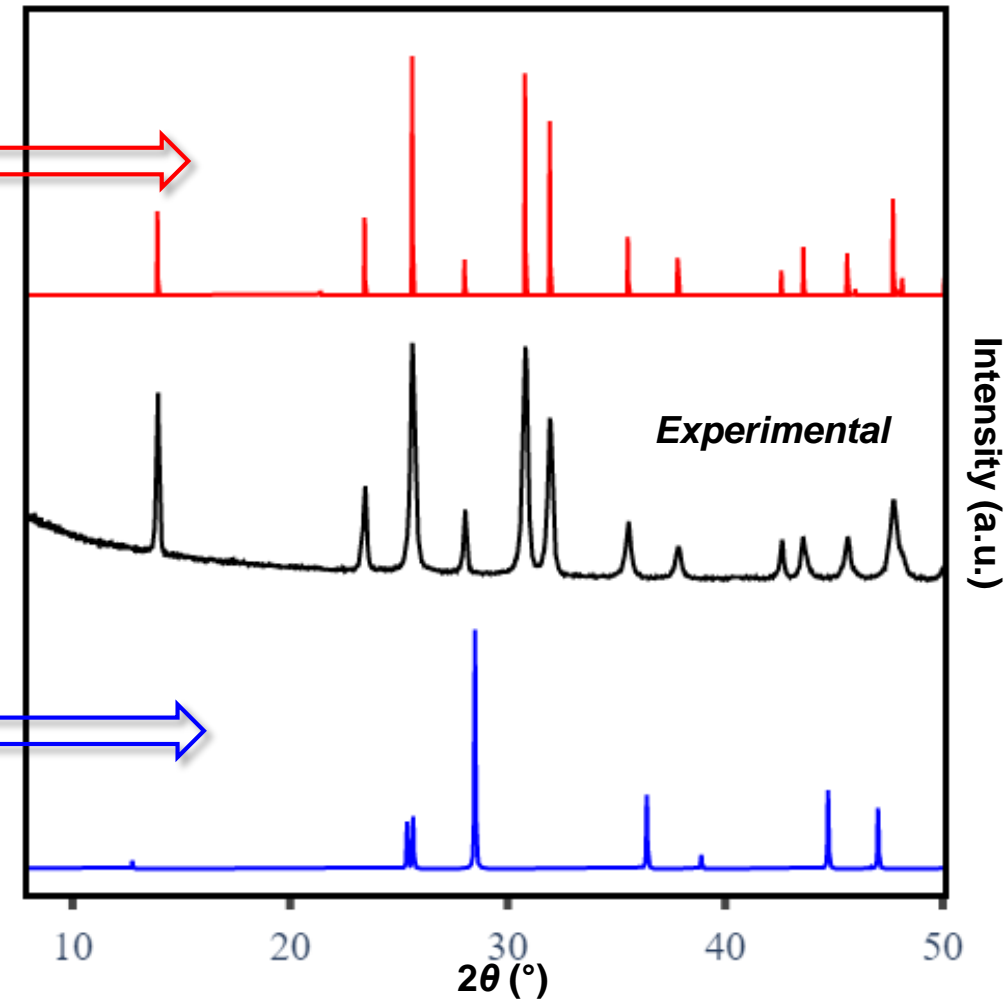
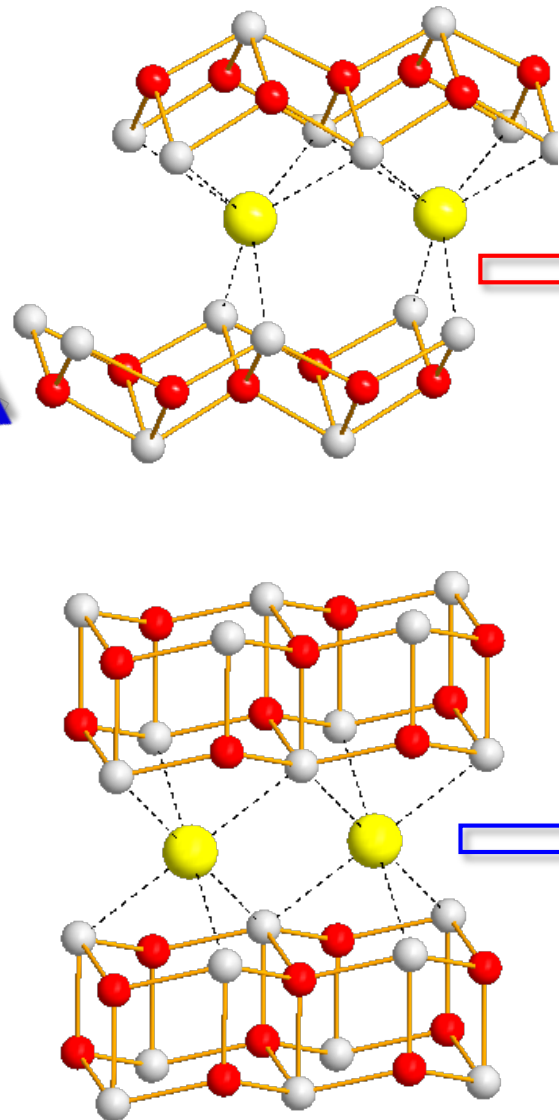
*Predicted Pattern*  
*hp-La<sub>2</sub>O<sub>2</sub>S*  
*Stable*  
*(SG: P $\bar{3}$ m1)*



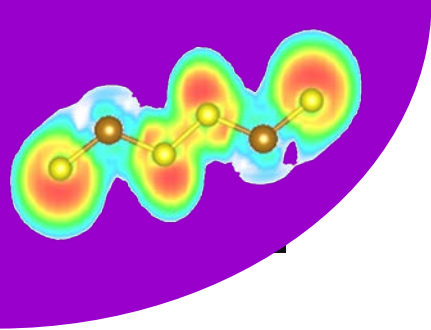
# S (de)intercalation of $\text{La}_2\text{O}_2\text{S}_2$



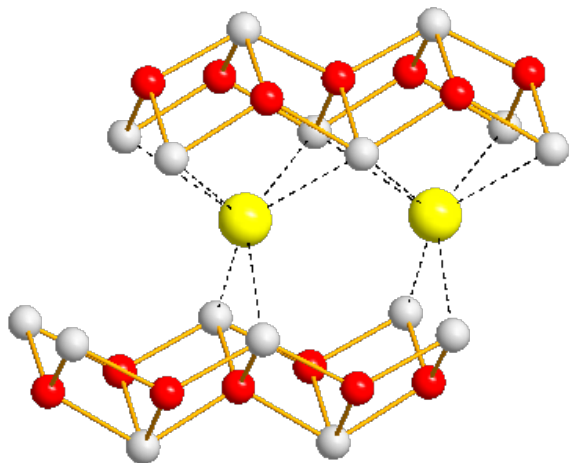
-1.0 S



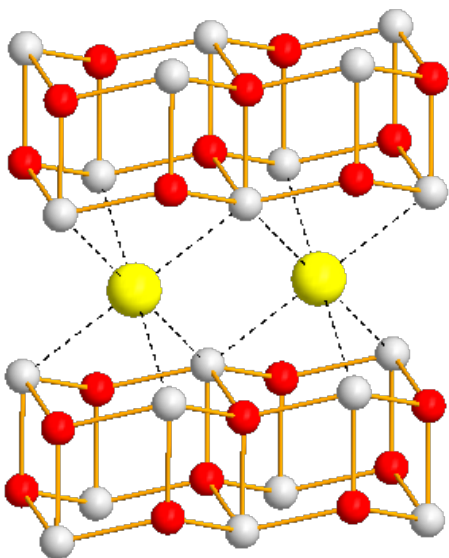
# Structure-related luminescence properties



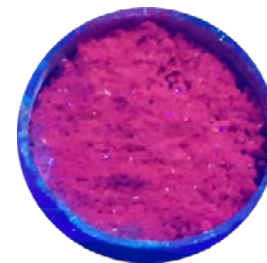
**Predicted Pattern**  
**oA-La<sub>2</sub>O<sub>2</sub>S**  
**Metastable**  
**(SG: Amm2)**



**Predicted Pattern**  
**hp-La<sub>2</sub>O<sub>2</sub>S**  
**Stable**  
**(SG: P $\bar{3}$ m1)**



Eu<sup>3+</sup> doping

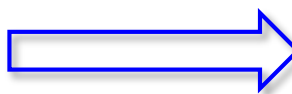


oA-La<sub>2</sub>O<sub>2</sub>S : Eu (1%)

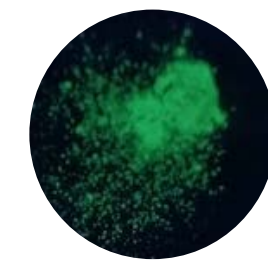
Pr<sup>3+</sup> doping



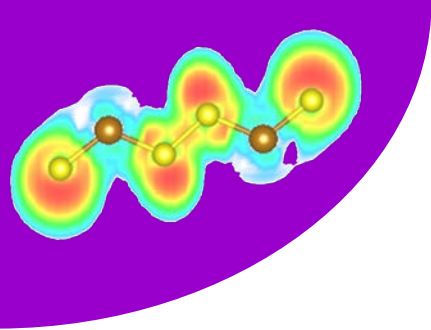
oA-La<sub>2</sub>O<sub>2</sub>S : Pr (1%)



hp-La<sub>2</sub>O<sub>2</sub>S : Eu (1%)

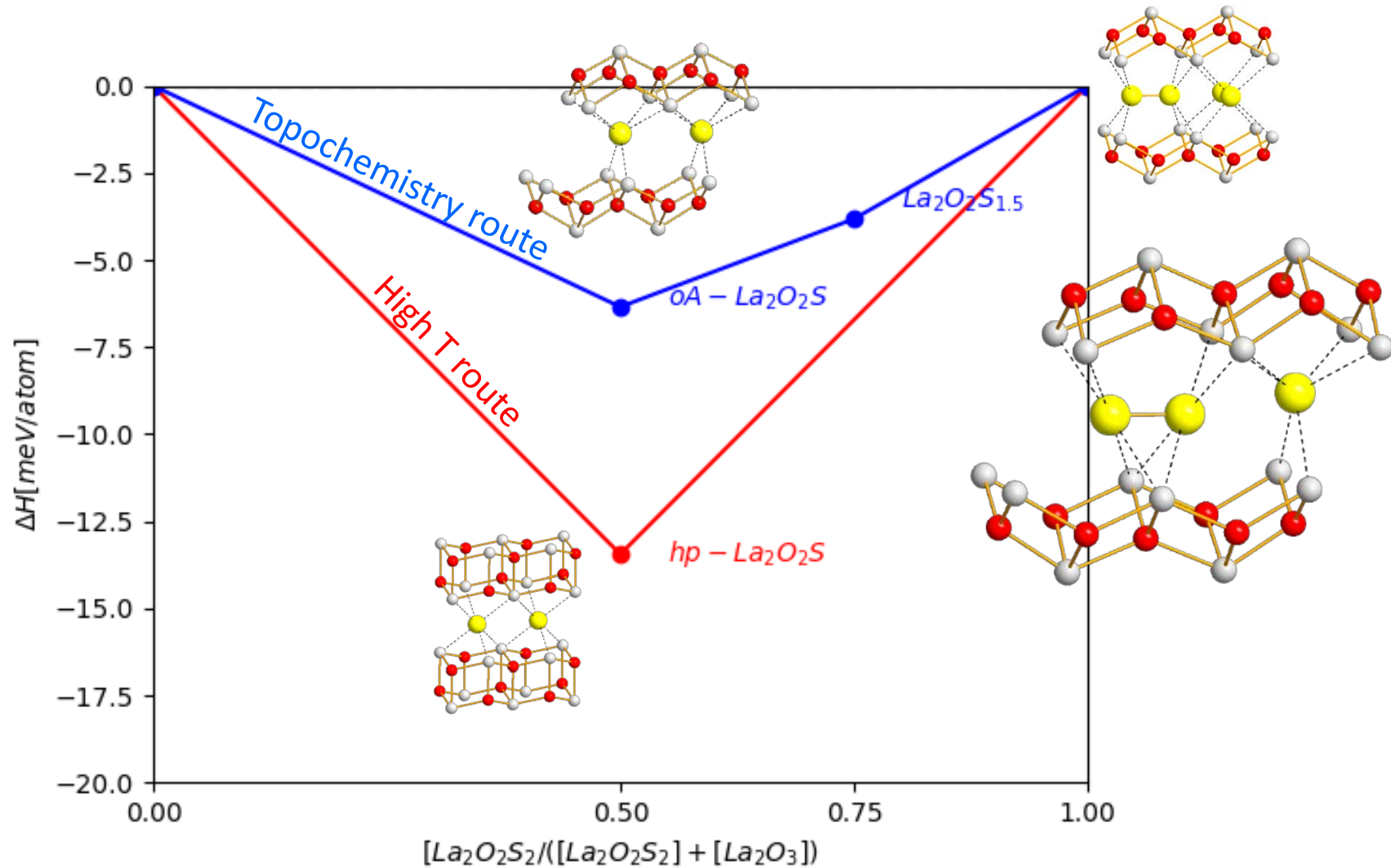


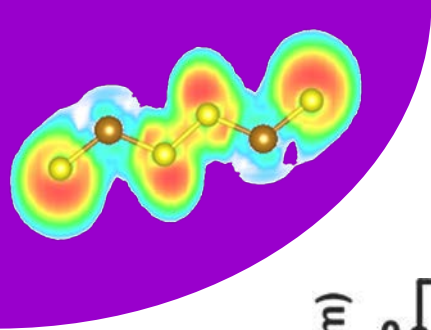
hp-La<sub>2</sub>O<sub>2</sub>S : Pr (1%)



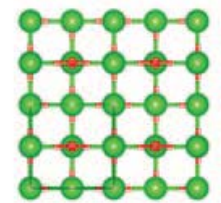
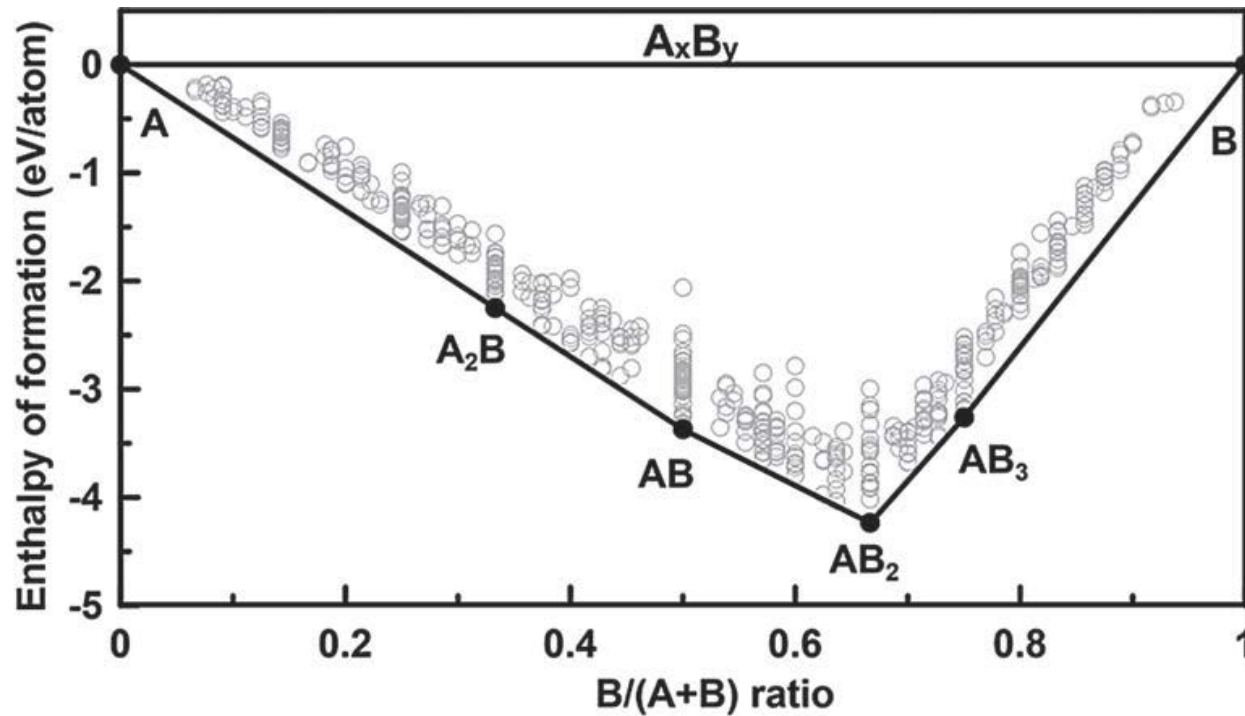
# Towards a phase diagram function of the synthesis route

Simple structural filiation with gliding of fluorine sheets

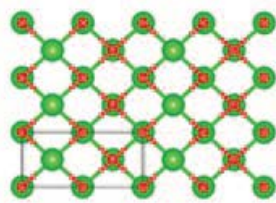




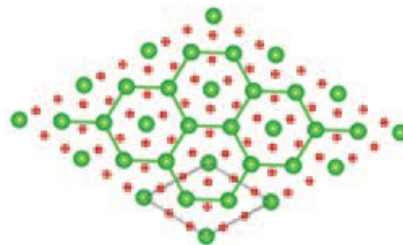
# Convex hull when composition is unknown (stable compounds)



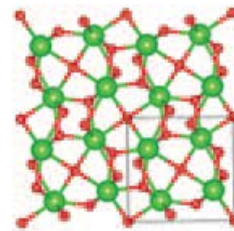
$A_2B$  ( $I4_1/amd$ )



AB ( $Pmma$ )



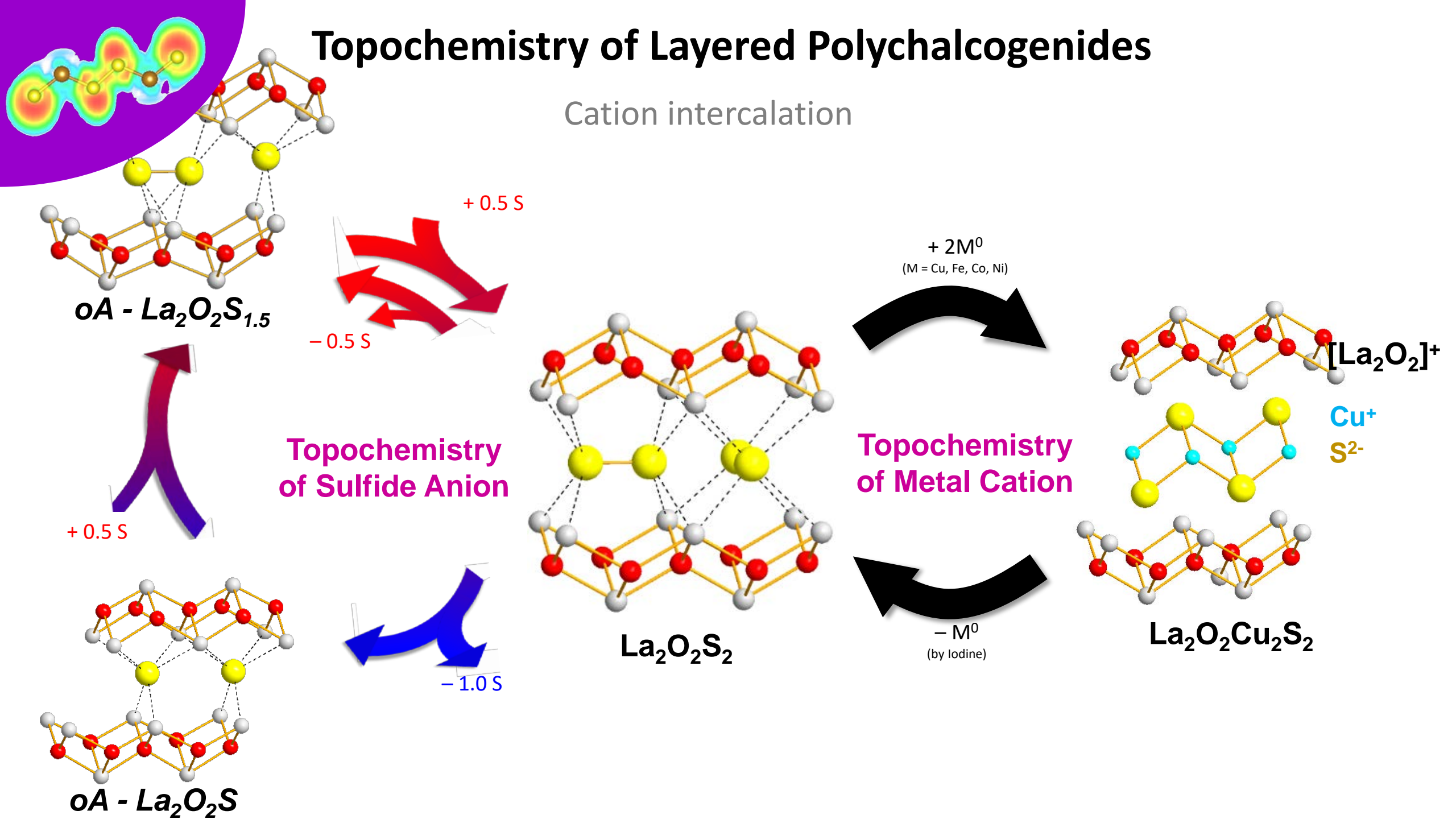
$AB_2$  ( $P\bar{6}32m$ )

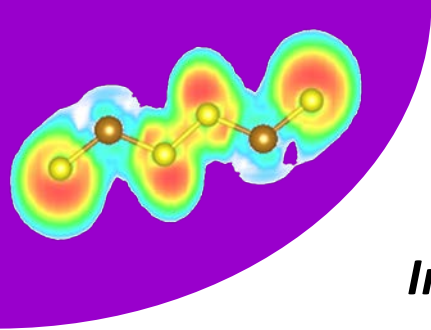


$AB_3$  ( $Pnnm$ )

# Topochemistry of Layered Polychalcogenides

Cation intercalation

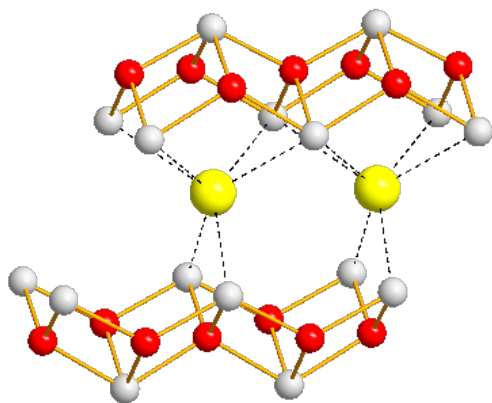




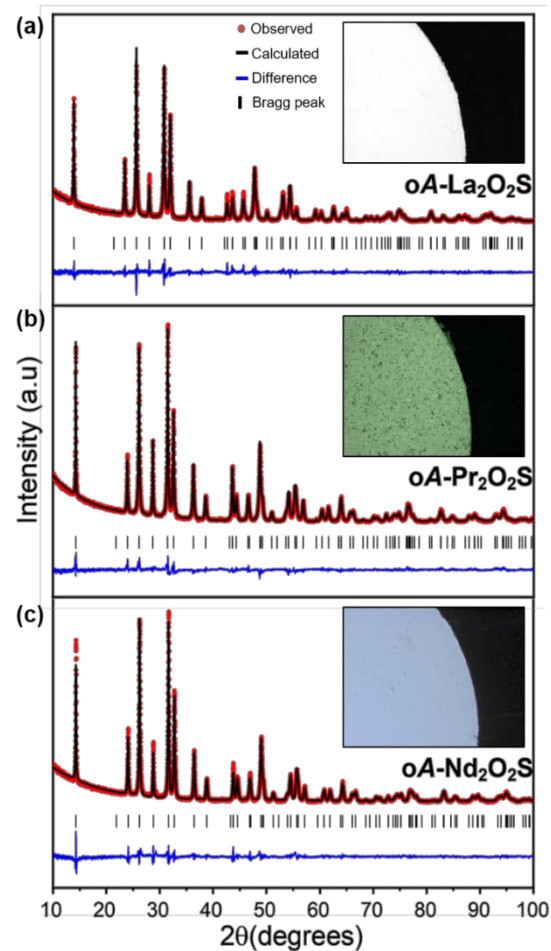
# Rationalizing cation substitution

*In the host lattice*

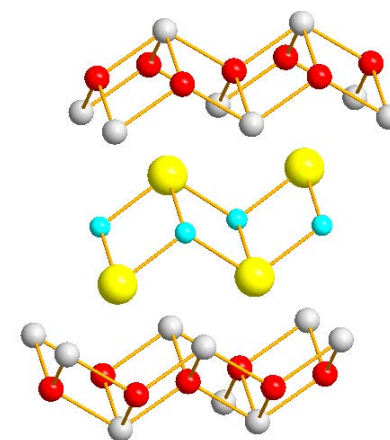
$\text{oA-Ln}_2\text{O}_2\text{S}$  (Ln=La, Pr, Nd)



L-B.Mvélé, S. Sasaki, P. Deniard,  
Y. Tsujimoto, E. Janod, C. Guillot-  
Deudon, M Caldes, IB, B.  
Corraze, S. Jobic, L. Cario, *Chem  
Mater.*, submitted



*In the confined layer*



$[\text{La}_2\text{O}_2]^+$

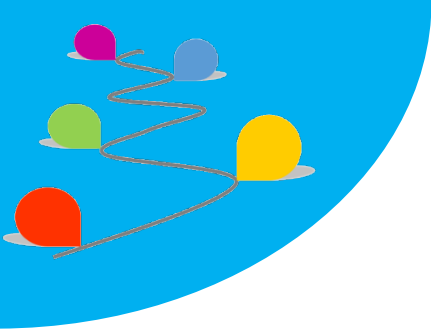
$\text{Cu}^+$

$\text{S}^{2-}$

$\text{La}_2\text{O}_2\text{Cu}_2\text{S}_2$

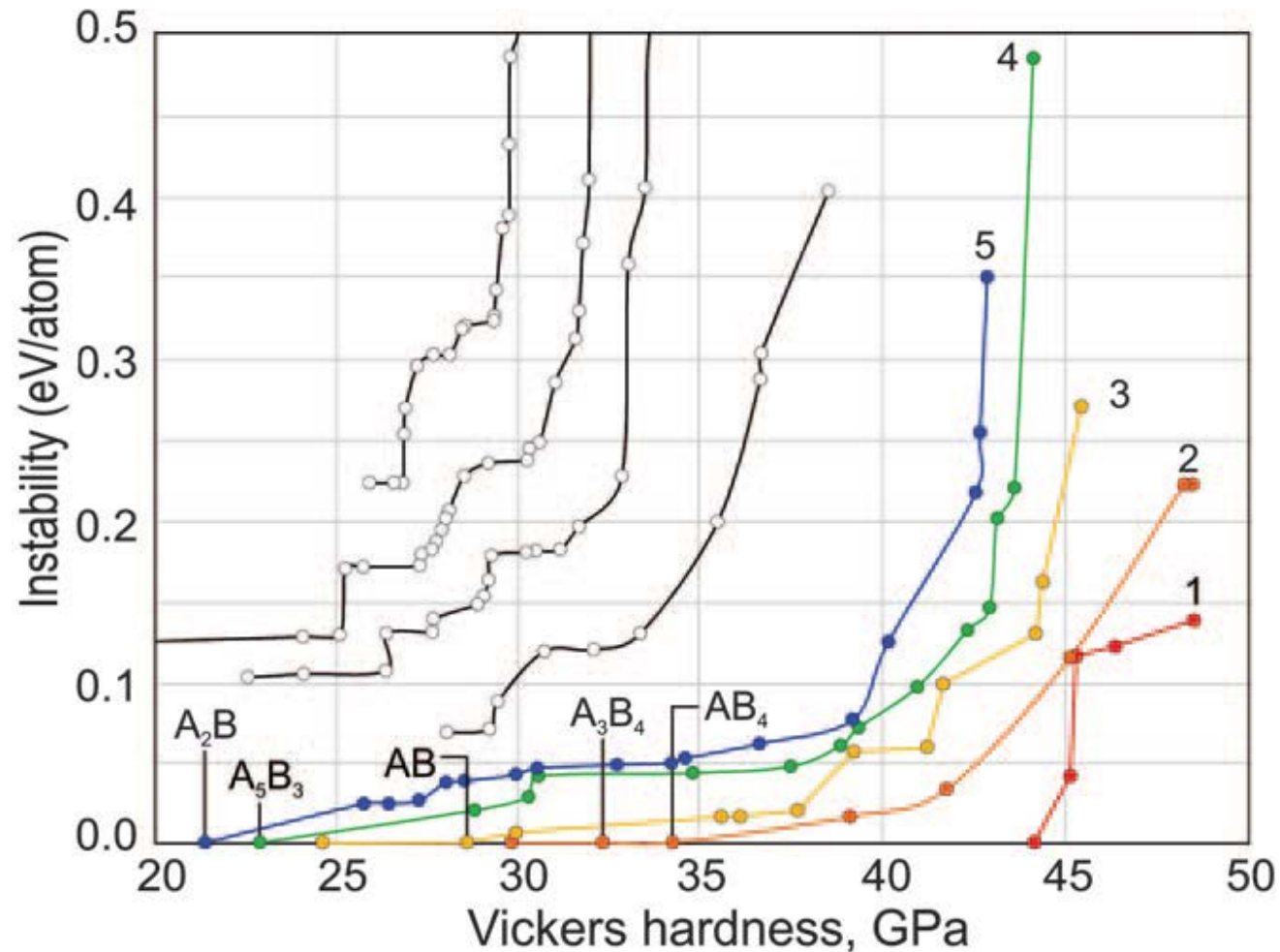
S. Sasaki, IB et al, *in preparation*

IB, S. Sasaki, E. Janod, B. Corraze, S. Jobic, L. Cario, *in  
preparation*



# Metastable with optimised desired property

## Multi-target optimization

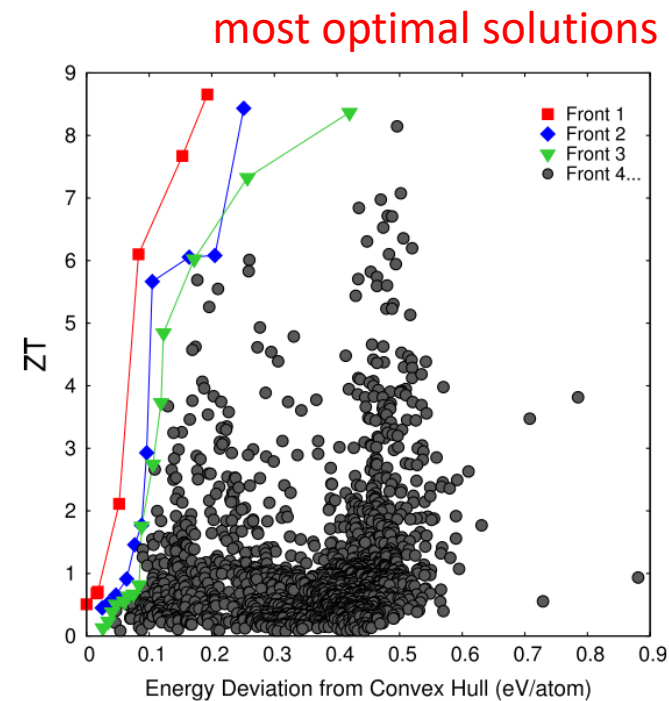
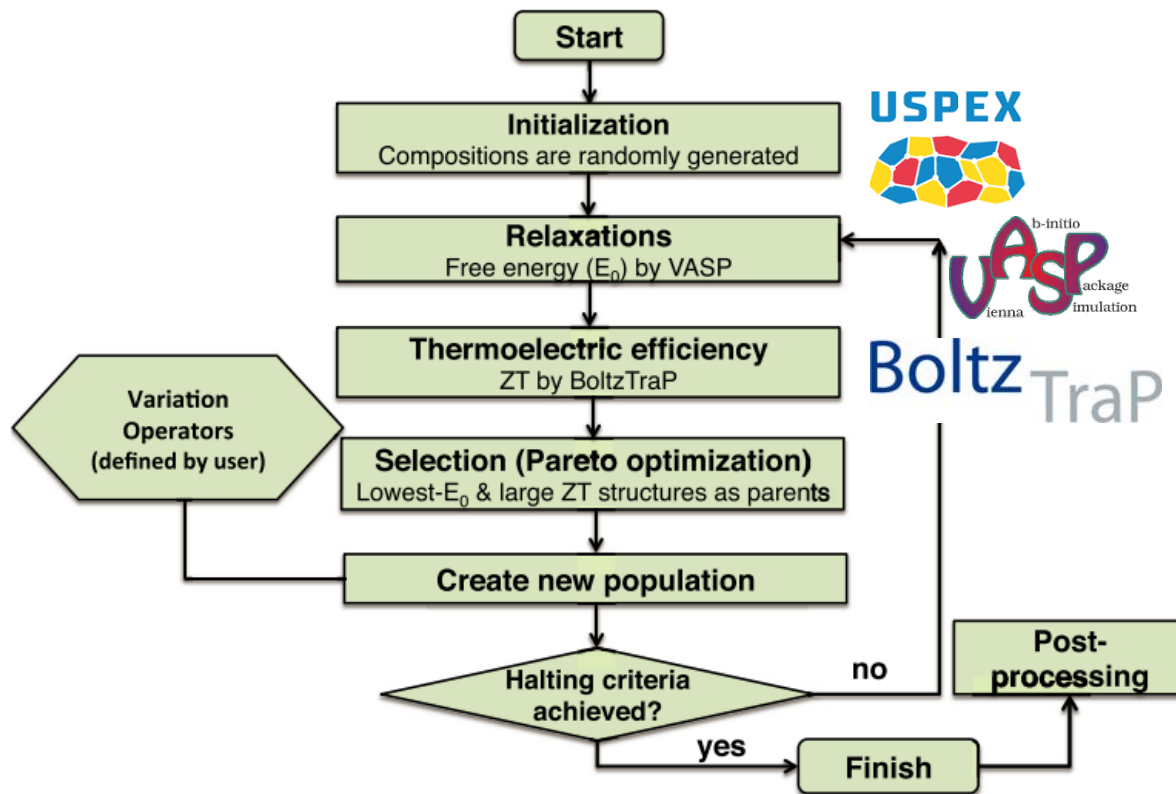


- enthalpy
- structure density
- volume of system unit cell (for 3D)
- Structure hardness
- degree of order
- structural quasientropy
- Elastic bulk modulus calculated using ML model.
- Elastic shear modulus calculated using ML model.
- Elastic Young's modulus calculated using ML model.
- Elastic Poisson's ratio calculated using ML model.
- Elastic Pugh's modulus ratio calculated using ML model.
- Elastic Vickers hardness calculated using ML model.
- Elastic fracture toughness calculated using ML model.

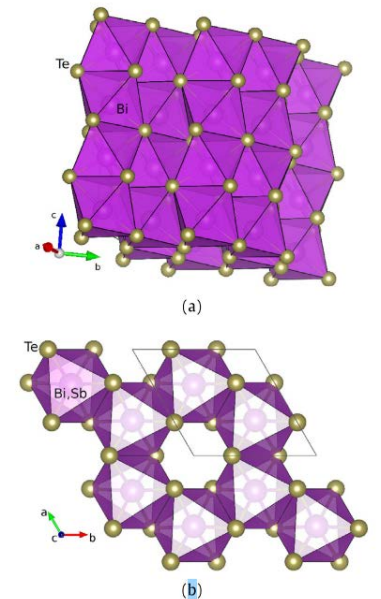


# Computational design of thermoelectric materials

seek structures that are not only the most stable, but at the same time, possess a large figure of merit ZT

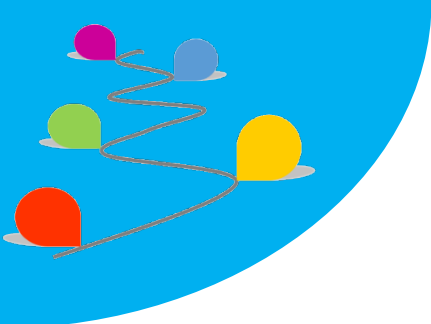


P63cm phases



$\text{Bi}_2\text{Te}_3\text{-Sb}_2\text{Te}_3$



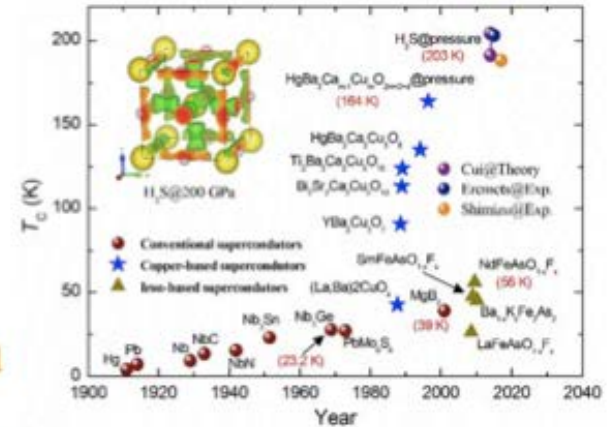
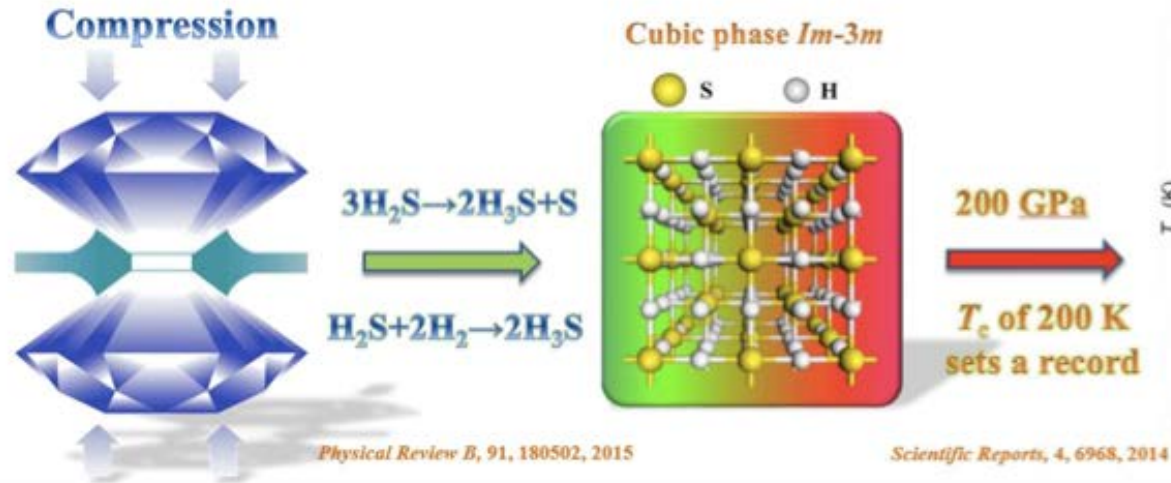
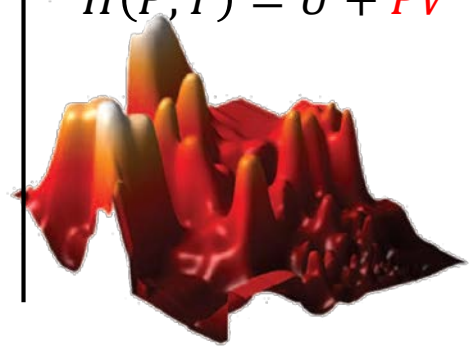


# Superconductors

Pressure diagram for bulk compounds is nowadays routine

## New sulfur hydride $H_3S$ and its exciting superconductivity under high pressures

$$H(P, T) = U + PV$$



D. Duan, Y. Liu, F. Tian, D. Li, X. Huang, Z. Zhao, H. Yu, B. Liu, W. Tian & . Cui (2014). *Nat. Sci. Rep.* **4**, 6968.

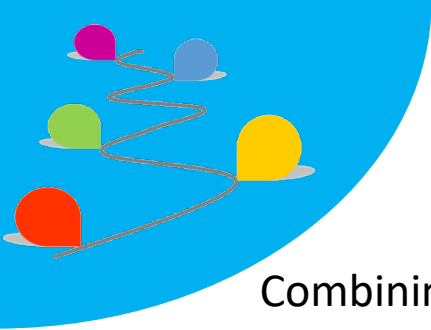
D. Duan, X. Huang, F. Tian, D. Li, H. Yu, Y. Liu, Y. Ma, B. Liu, and T. Cui (2015). *Phys. Rev. B* **91**(18), 180502.

## ADVANCED MATERIALS

Research Article | [Full Access](#)

### Effect of Magnetic Impurities on Superconductivity in LaH<sub>10</sub>

Dmitrii V. Semenov ✉ Ivan A. Troyan, Andrey V. Sadakov, Di Zhou ✉ Michele Galasso, Alexander G. Kvashnin, Anna G. Ivanova, Ivan A. Kruglov, Alexey A. Bykov ... [See all authors](#) ▾



# Take-home message

Combining structure prediction with diffraction or other experimental methods has provided a very successful approach to solving crystal structures of metastable compounds.

## Win-win :

- **information on Z** greatly simplifies the searching.
- structure prediction could yield **better models for structural refinement**

Material discovery can tackle desired target, unknown stoichiometries and even **search through the periodic table**

“For most bulk materials synthesizing and processing procedures are reasonably well established. However, theoretical approaches have limited analytical power for predicting viable synthetic routes towards making entirely new materials. State-of-the-art in materials design needs to be complemented with substantial efforts in advancing the field of synthesis design.

To increase predictive ability of material synthesizability, it is necessary to ***define both equilibrium and out-of-equilibrium descriptors that control synthetic routes and outcomes.***”

Kirstin Alberi et al 2019, The 2019 materials by design roadmap, *J. Phys. D: Appl. Phys.* 52 013001

*Thank you for your attention !*

*Open to questions and collaborations*