Materials Discovery Guided by Artificial Intelligence

Artem R. Oganov

Skolkovo Institute of Science and Technology, Russia

Crystal structure determines physical properties.
Crystal structure determination was a major breakthrough.

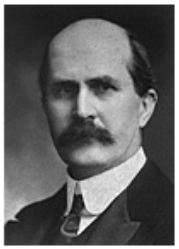
The Nobel Prize in Physics 1914

"for his discovery of the diffraction of X-rays by crystals"

Max von Laue

The Nobel Prize in Physics 1915

"for their services in the analysis of crystal structure by means of X-rays"



Sir William Henry Bragg

William Lawrence Bragg

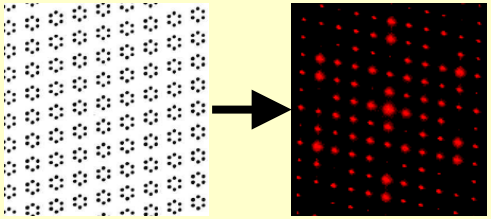
The Nobel Prize in Chemistry 1985

"for their outstanding achievements in the development of direct methods for the determination of crystal structures"

Herbert A. Hauptman

Jerome Karle

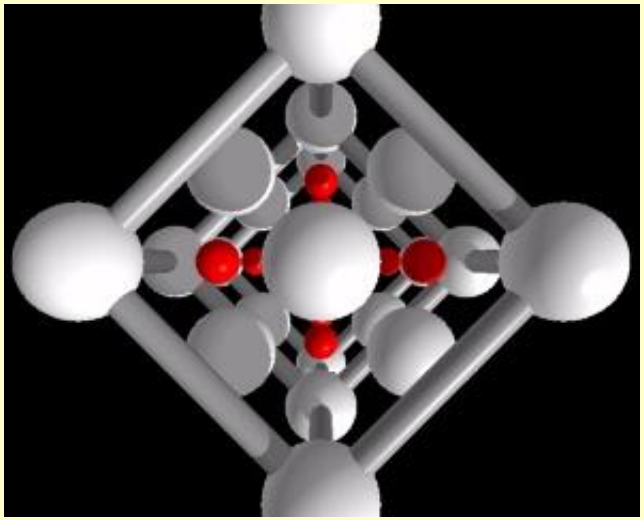
(from http://nobelprize.org)



Structure Diffraction

Zincblende ZnS.

One of the first solved structures (1912-1913)



Briefly about big data

Big data have predictive power

We have:
~300,000 experimental (& >500,000 theoretical) inorganic crystal structures. For many of them, we have some physical properties (and for many, we don't).

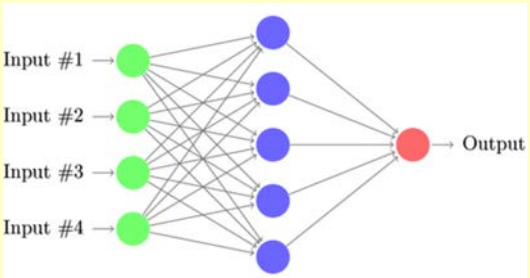
Many studied syntheses of compounds.

Many studied industrial production processes.

Big data analysis gives:
Fast predictions

Predictions where the full calculation would be too complicated.

Predictions where there's no theory.

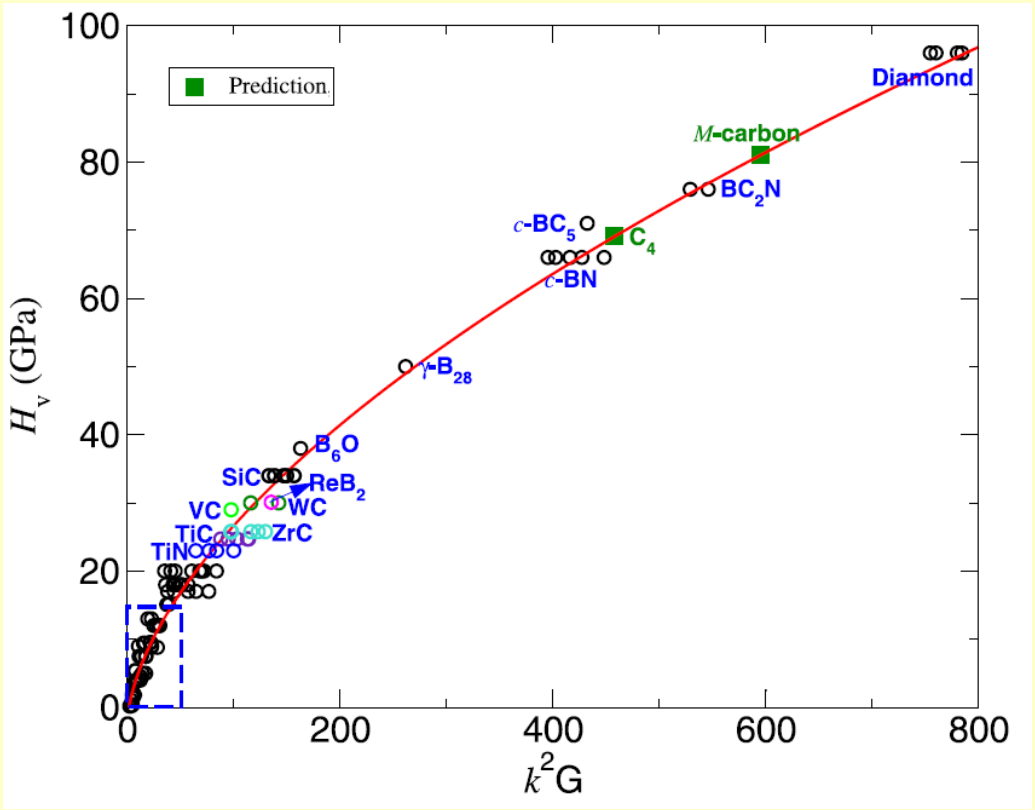


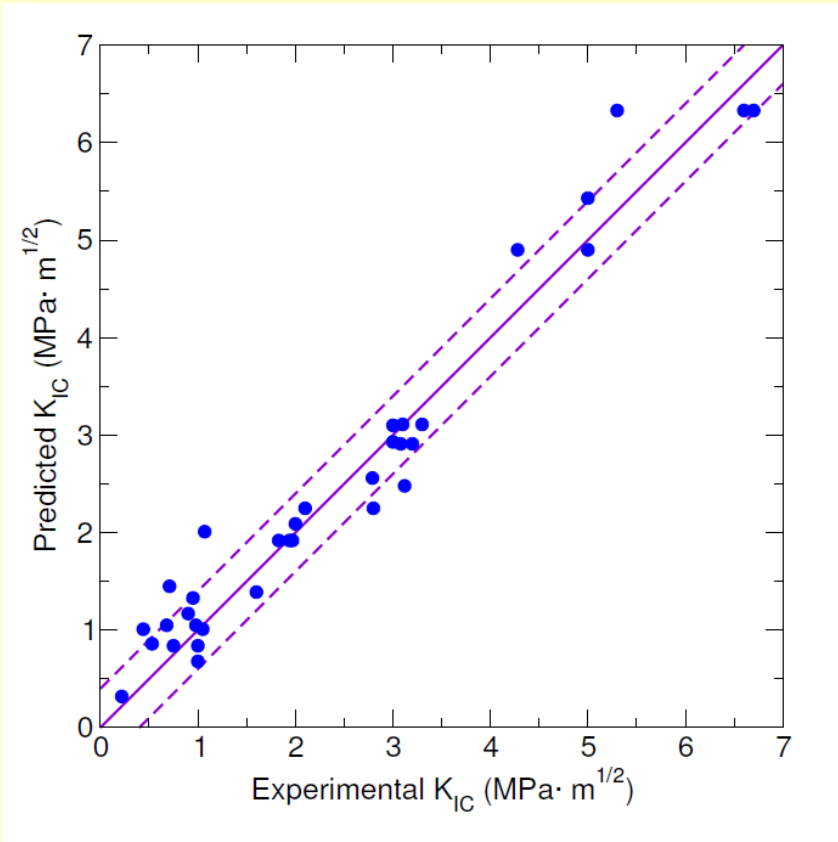
			Ti = 50	Zr = 90	? = 180	
			V = 51	Nb = 94	Ta = 182	
			Cr = 52	Mo = 96	W = 186	
			Mn = 55	Rh = 104,4	Pt = 197,4	
			Fe = 56	Ru = 104,4	Ir = 198	
		Ni =	Co = 59	Pd = 106,6	Os = 199	
H = 1			Cu = 63,4	Ag = 108	Hg = 200	
	Be = 9,4	Mg = 24	Zn = 65,2	Cd = 112		
	B = 11	Al = 27,4	? = 68	Ur = 116	Au = 197?	
	C = 12	Si = 28	? = 70	Sn = 118		
	N = 14	P = 31	As = 75	Sb = 122	Bi = 210?	
	O = 16	S = 32	Se = 79,4	Te = 128?		
	F = 19	Cl = 35,5	Br = 80	J = 127		
Li = 7	Na = 23	K = 39	Rb = 85,4	Cs = 133	Tl = 204	
		Ca = 40	Sr = 87,6	Ba = 137	Pb = 207	
		? = 45	Ce = 92			
		?Er = 56	La = 94			
		?Yt = 60	Di = 95			
		?In = 75,6	Th = 118?			

Predicting complex properties: hardness [Chen et al., 2011] and fracture toughness [Niu & Oganov, 2019]

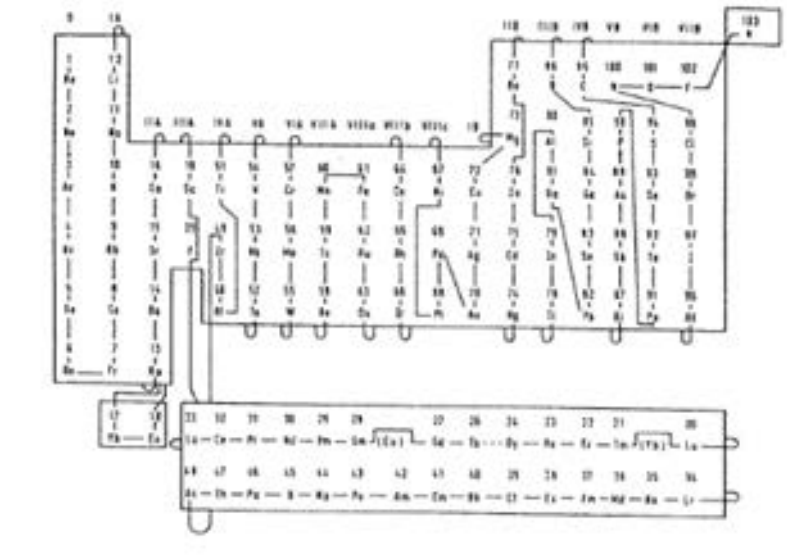
$$H_v = 2\left(k^2 G\right)^{0.585} - 3$$

$$K_{IC} = (1+\alpha)\cdot V^{1/6}\cdot G\cdot (B/G)^{1/2}$$





Mendeleev number (Pettifor, 1984). Prediction of stability, structure, and properties of materials



Mendeleev numbers of the elements

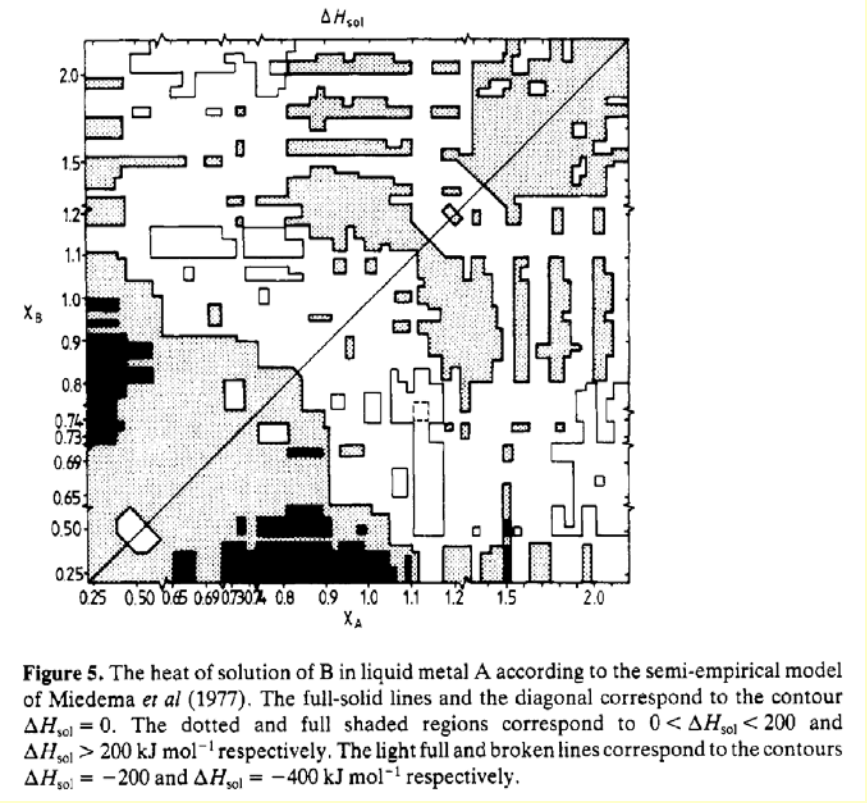
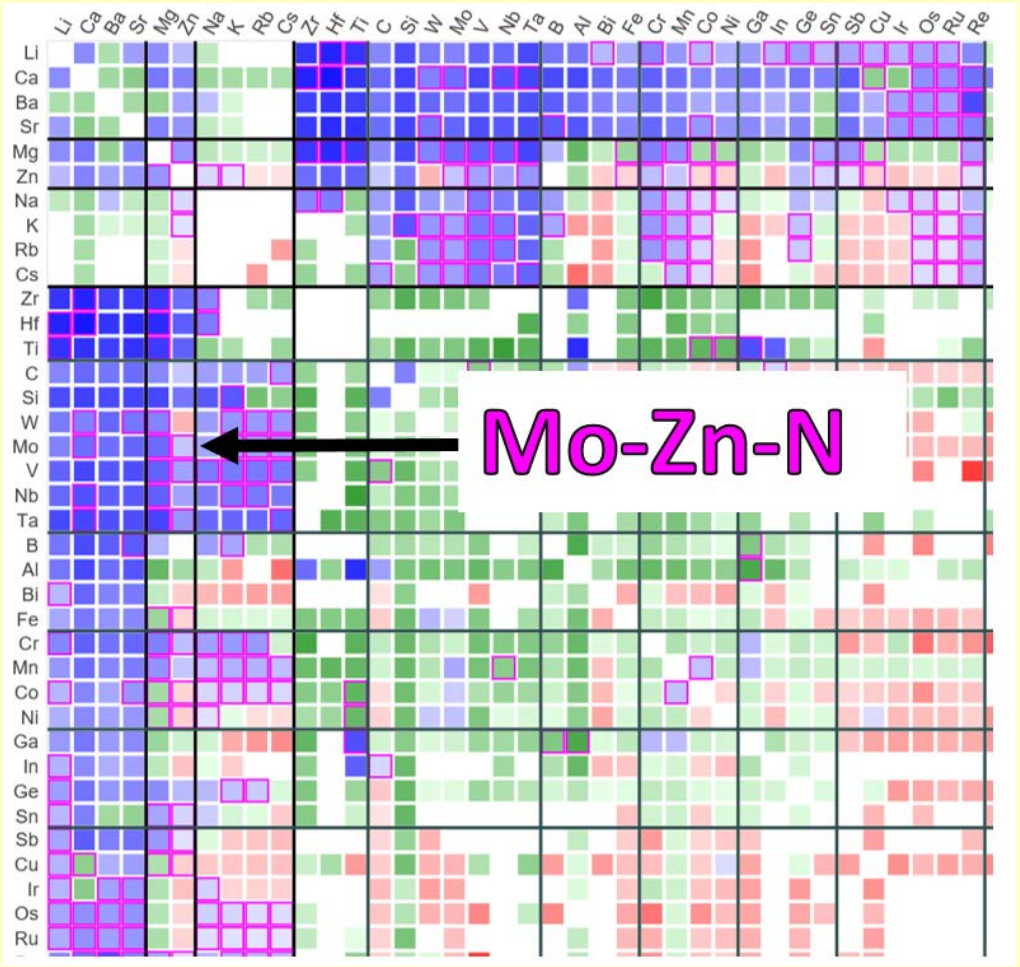


Figure 5. The heat of solution of B in liquid metal A according to the semi-empirical model of Miedema *et al* (1977). The full-solid lines and the diagonal correspond to the contour $\Delta H_{sol} = 0$. The dotted and full shaded regions correspond to $0 < \Delta H_{sol} < 200$ and $\Delta H_{sol} > 200$ kJ mol^{-1} respectively. The light full and broken lines correspond to the contours $\Delta H_{sol} = -200$ and $\Delta H_{sol} = -400$ kJ mol^{-1} respectively.

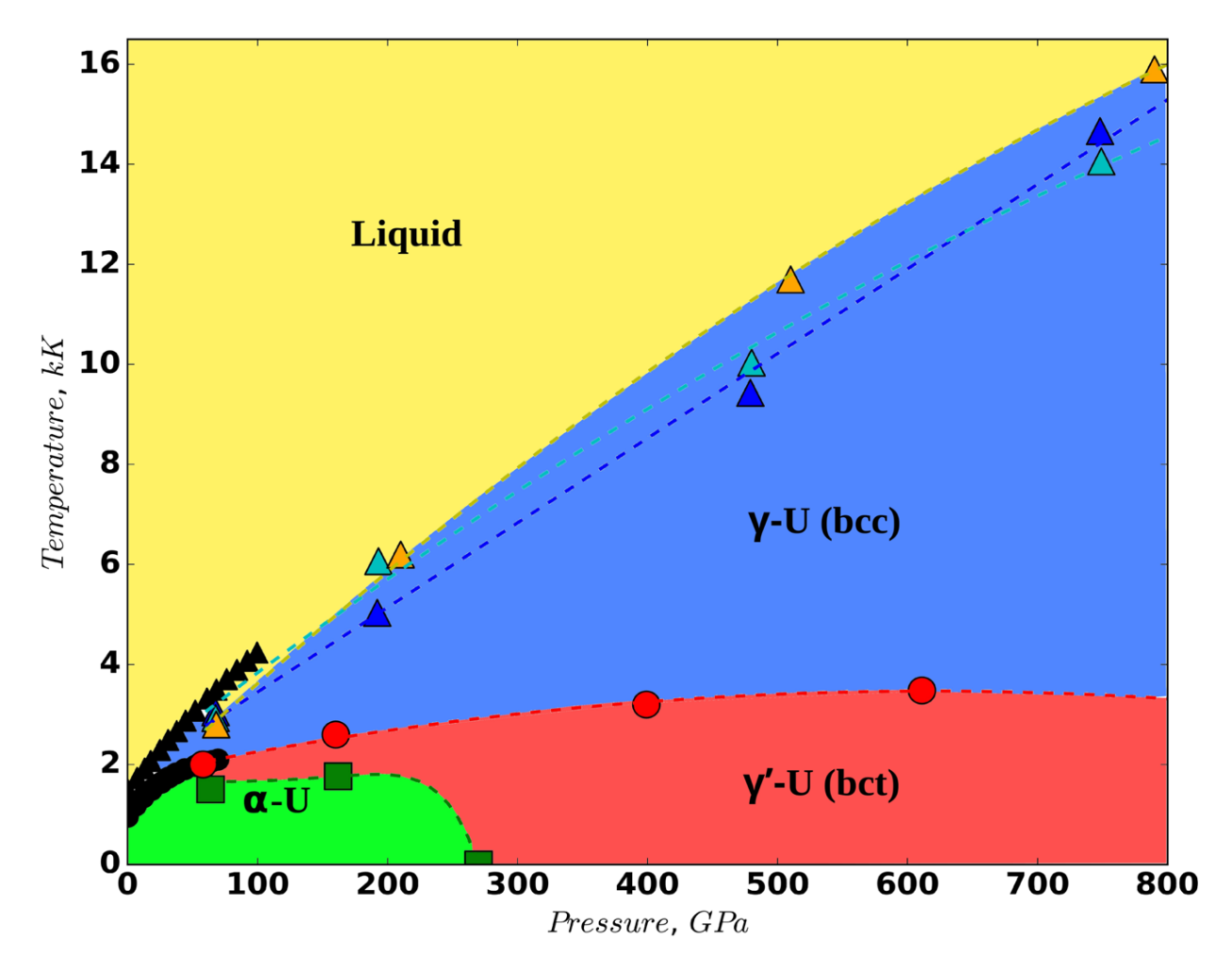
Enthalpies of formation of compounds

Example: search for new ternary nitrides (Sun, 2019)



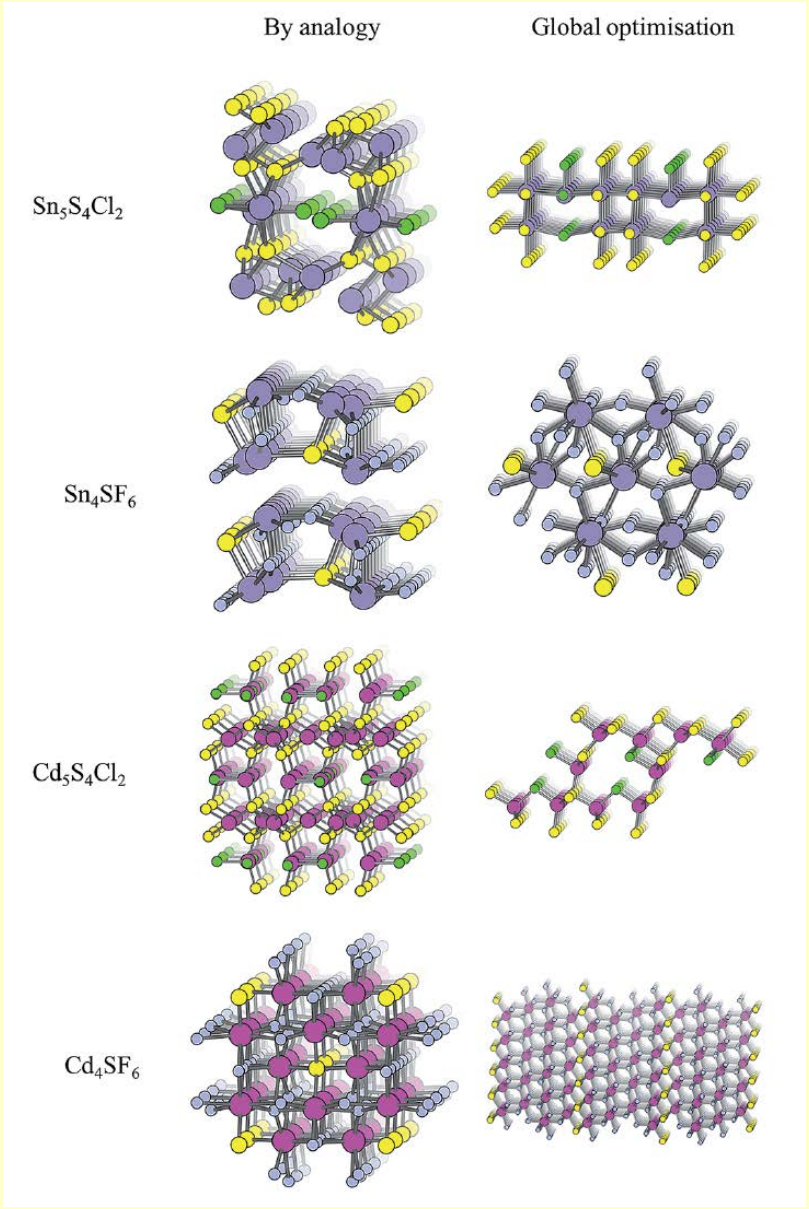
916 systems
246 contain stable nitrides
In 127 of them nitrides were not experimentally known

Machine learning can be used for accelerating simulations without loss of accuracy



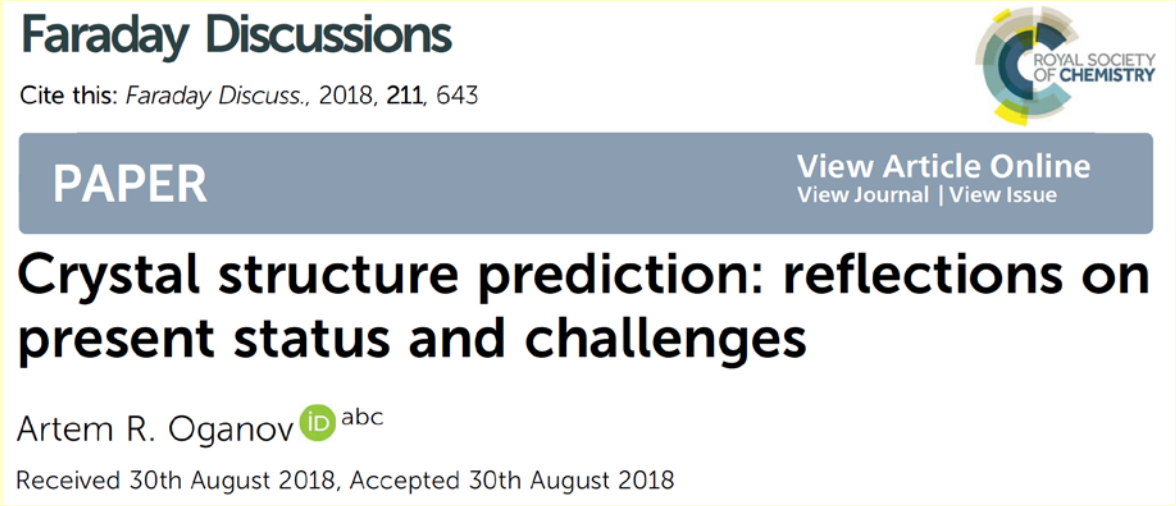
Phase diagram of uranium obtained with machine learning (Kruglov & Oganov, 2020)

Data mining gives quick-n-reasonable answers, but beware!



Davies & Oganov (2018): 4 interesting semiconductors, and for each data mining gave a wrong structure:
-for $Sn_5S_4Cl_2$ by 24.7 meV/atom,
-for Sn_4SF_6 by 5.1 meV/atom,
-for Cd_4SF_6 by 0.2 meV/atom,
-for $Cd_5S_4Cl_2$ by 33.3 meV/atom.

Briefly about crystal structure prediction



Faraday Discussions

Cite this: *Faraday Discuss.*, 2018, **211**, 643

ROYAL SOCIETY OF CHEMISTRY

PAPER

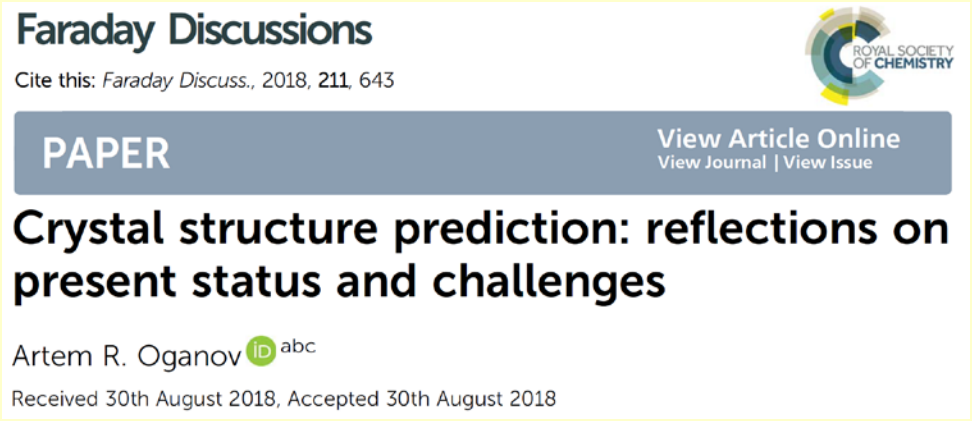
View Article Online
View Journal | View Issue

Crystal structure prediction: reflections on present status and challenges

Artem R. Oganov [abc]

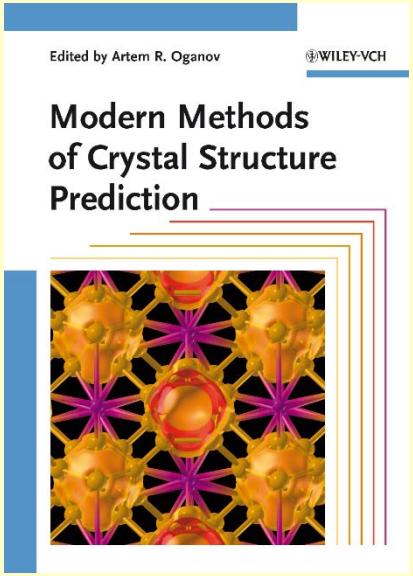
Received 30th August 2018, Accepted 30th August 2018

Useful reviews and books

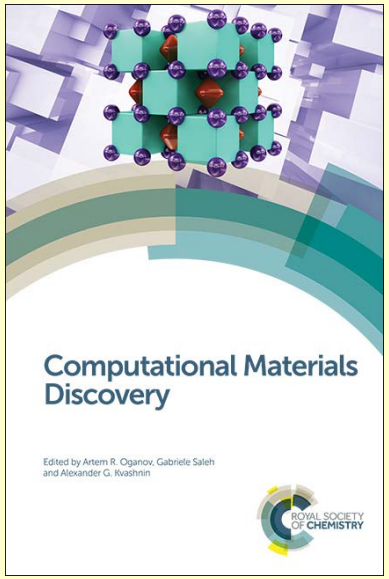


***Faraday Discussions* (2018)**

***Nature Reviews Materials* (2019)**



2011



2018

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

Are Crystal Structures Predictable?

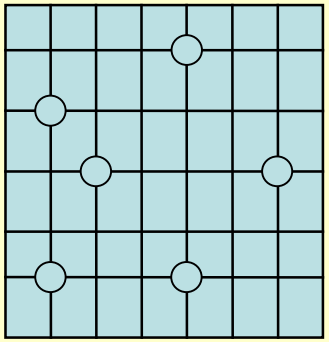
ANGELO GAVEZZOTTI*

"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 present state of affairs

Need to find GLOBAL energy minimum.

Trying all structures is impossible:

$$C = \frac{1}{(V/\delta^3)} \frac{(V/\delta^3)!}{[(V/\delta^3) - N]!N!}$$



N_{atoms}	Variants	CPU time
1	**1**	**1 sec.**
10	**10^{11}**	**10^{3} yrs.**
20	**10^{25}**	**10^{17} yrs.**
30	**10^{39}**	**10^{31} yrs.**

RESEARCH NEWS

Crystal structure prediction – evolutionary or revolutionary crystallography?

S. L. Chaplot and K. R. Rao CURRENT SCIENCE, VOL. 91, NO. 11, 10 DECEMBER 2006

Overview of USPEX (Oganov & Glass, *J.Chem.Phys.* 2006)

The USPEX project
(Universal Structure Predictor: Evolutionary Xtallography)

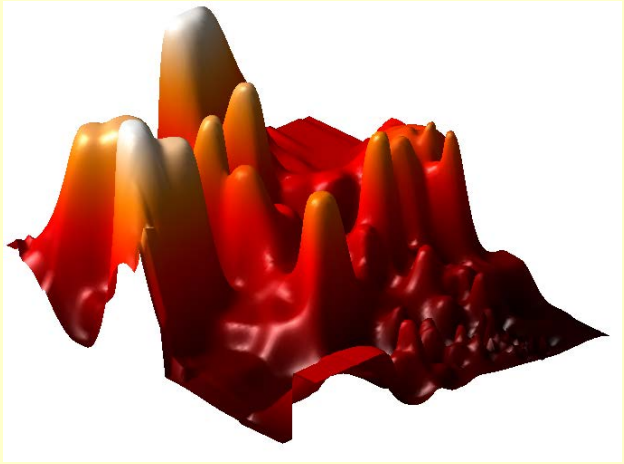
http://uspex-team.org

[Oganov A.R., Glass C.W., *J.Chem.Phys.* 124, 244704 (2006)]

- **Combination of evolutionary algorithm and quantum-mechanical calculations.**
- **>6100 users.**

- **Solves «intractable» problem of structure prediction**
  -3D, 2D, 1D, 0D –systems,
  -prediction of phase transition mechanisms.

- **Interfaced with: VASP, Quantum Espresso, CASTEP, FHI-aims, ABINIT, Siesta, Gaussian, ORCA, ATK, DFTB, MOPAC, GULP, LAMMPS, Tinker, DMACRYS**



Energy landscape of Au_8Pd

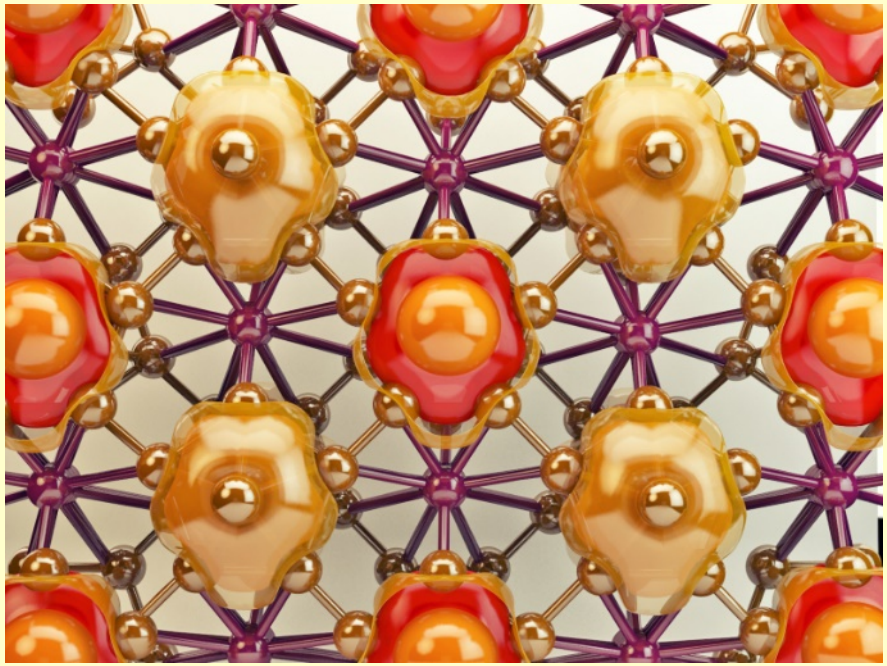
W. Kohn

J. P. Perdew

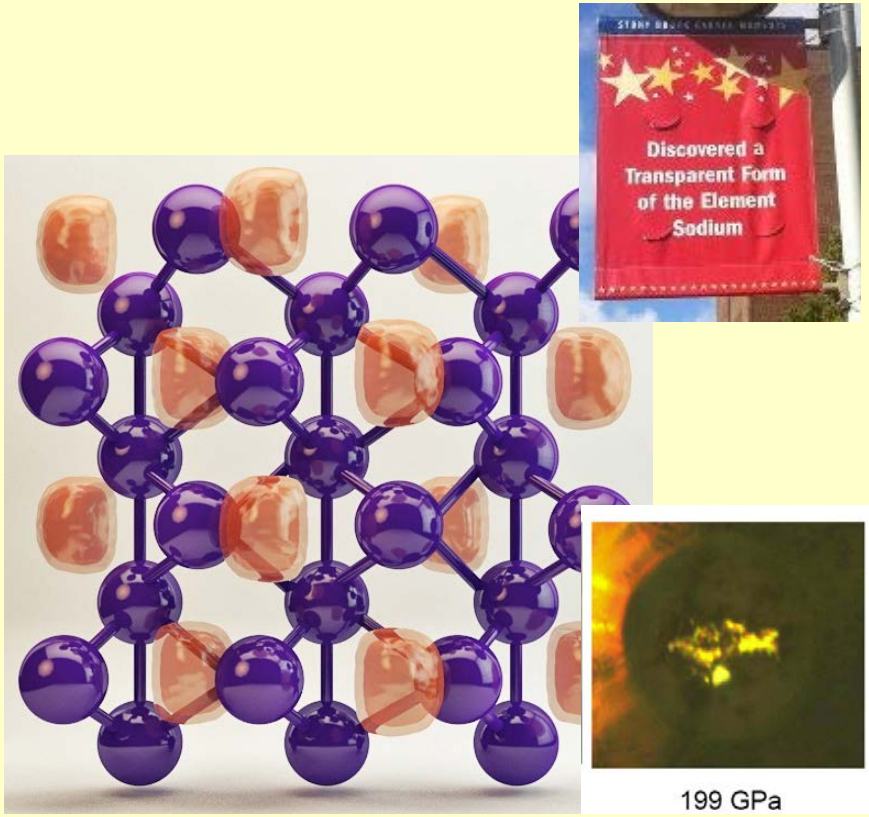
$$(-\frac{\nabla^2}{2}+v_{e-n}[\rho(\mathbf{r})]+v_H[\rho(\mathbf{r})]+v_{xc}[\rho(\mathbf{r})])\phi_i(\mathbf{r})=\varepsilon_i\phi_i(\mathbf{r})$$

$$E_{GGA,\mathrm{xc}}=\int \mathrm{d}\mathbf{r}\, F_{\mathrm{xc}}(\rho,\frac{|\nabla\rho|}{2k_{\mathrm{F}}\rho(\mathbf{r})})\rho(\mathbf{r})e_{\mathrm{x}}[\rho(\mathbf{r})]$$

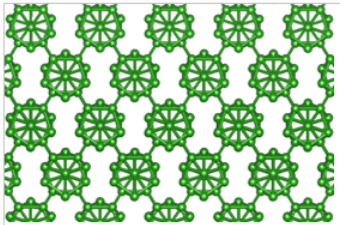
Predicting new crystal structures without empirical information



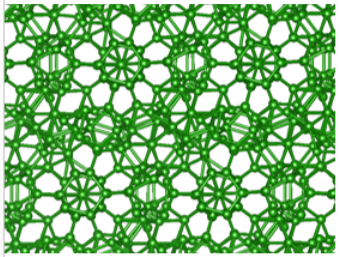
New superhard structure of boron
(Oganov et al., *Nature*, 2009)



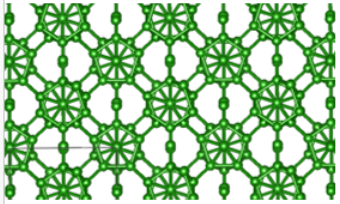
High-pressure transparent
allotrope of sodium
(Ma, Eremets, Oganov, *Nature*, 2009)



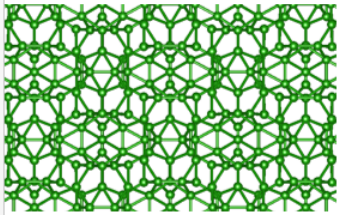
α-boron
E^{DFT} = –6.706 eV/atom
Atoms: 12,
Space group: *R*-3*m*,
Core-hours: 10^3 AL-MTP vs. $3 \cdot 10^3$ DFT
$|E^{DFT} - E^{MTP}|$ = 28.6 meV/atom



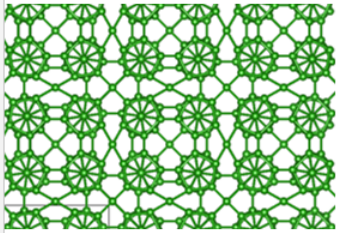
β-boron approximant
E^{DFT} = –6.704 eV/atom,
Atoms: 106,
Space group: *P*1,
Core-hours: $7 \cdot 10^3$ AL-MTP vs. $6.6 \cdot 10^7$ DFT
$|E^{DFT} - E^{MTP}|$ = 10.1 meV/atom



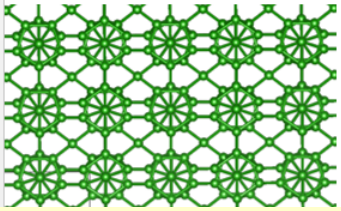
γ-boron
E^{DFT} = –6.678 eV/atom
Atoms: 28,
Space group: *Pnnm*,
Core-hours: $2 \cdot 10^3$ AL-MTP vs. $2.5 \cdot 10^4$ DFT
$|E^{DFT} - E^{MTP}|$ = 58.1 meV/atom



E^{DFT} = –6.667 eV/atom,
Atoms: 54,
Space group: *Im*-3,
Core-hours: $3 \cdot 10^3$ AL-MTP vs. $3.5 \cdot 10^5$ DFT
$|E^{DFT} - E^{MTP}|$ = 7.3 meV/atom



E^{DFT} = –6.667 eV/atom,
Atoms: 52,
Space group: *P*-42*m*,
Core-hours: $3 \cdot 10^3$ AL-MTP vs. $3.2 \cdot 10^5$ DFT
$|E^{DFT} - E^{MTP}|$ = 37.3 meV/atom



E^{DFT} = –6.665 eV/atom,
Atoms: 26,
Space group: *Cccm*,
Core-hours: $2 \cdot 10^3$ AL-MTP vs. $2.1 \cdot 10^4$ DFT
$|E^{DFT} - E^{MTP}|$ = 13.6 meV/atom

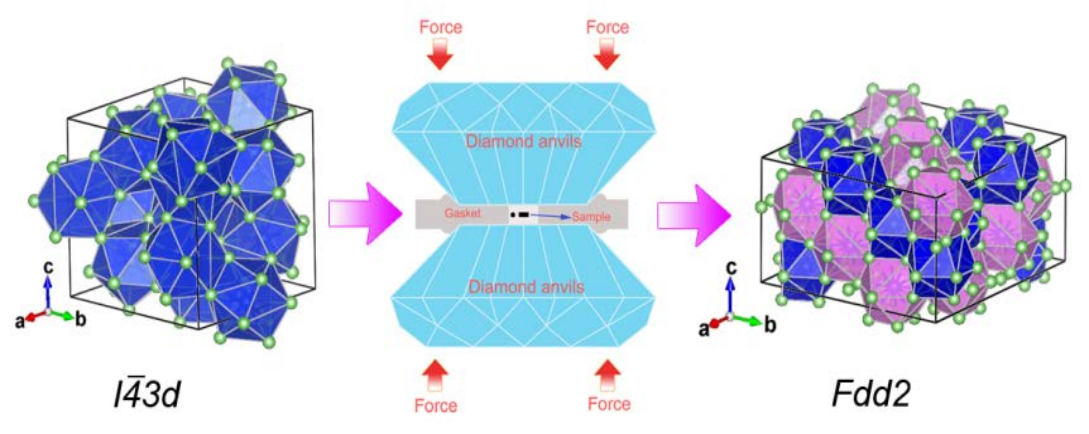
Handling complexity with machine learning: boron allotropes

(E.Podryabinkin, E. Tikhonov, A. Shapeev, A.R. Oganov, PRB, 2019)

- **ML potential with active learning (Shapeev, 2018). 800 parameters.**
- **MAE = 11 meV/atom.**
- **Reproduced α-, β-, γ-, T52 phases of boron.**
- **Predicted low-energy metastable cubic cI54 phase.**
- **Speedup by >100 times.**

Structure of $Li_{15}Si_4$ with 152 atoms/cell: record complexity, promise for Li-batteries

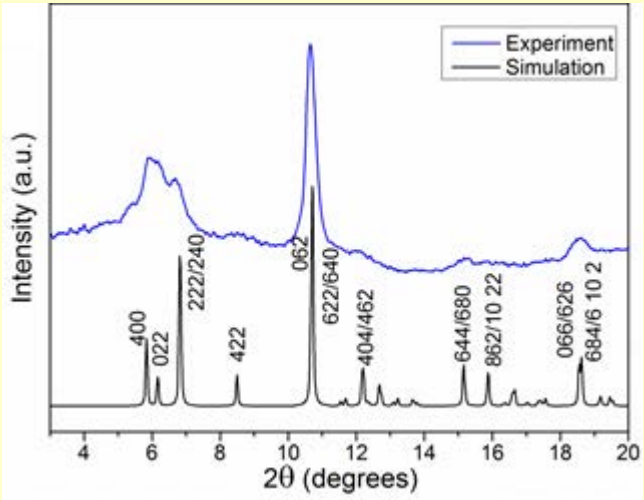
Li-Si



Structural transformation of $Li_{15}Si_4$ at 7 GPa. New phase has more attractive properties for use in Li-batteries.

Evolutionary metadynamics is a hybrid of:
-Metadynamics (Martonak, Laio, Parrinello, PRL 2003)
-Evolutionary algorithm USPEX (Oganov & Glass, JCP 2006)

It includes q-vectors and allows system size to change spontaneously



XRD of *Fdd*2-$Li_{15}Si_4$ at 18 GPa

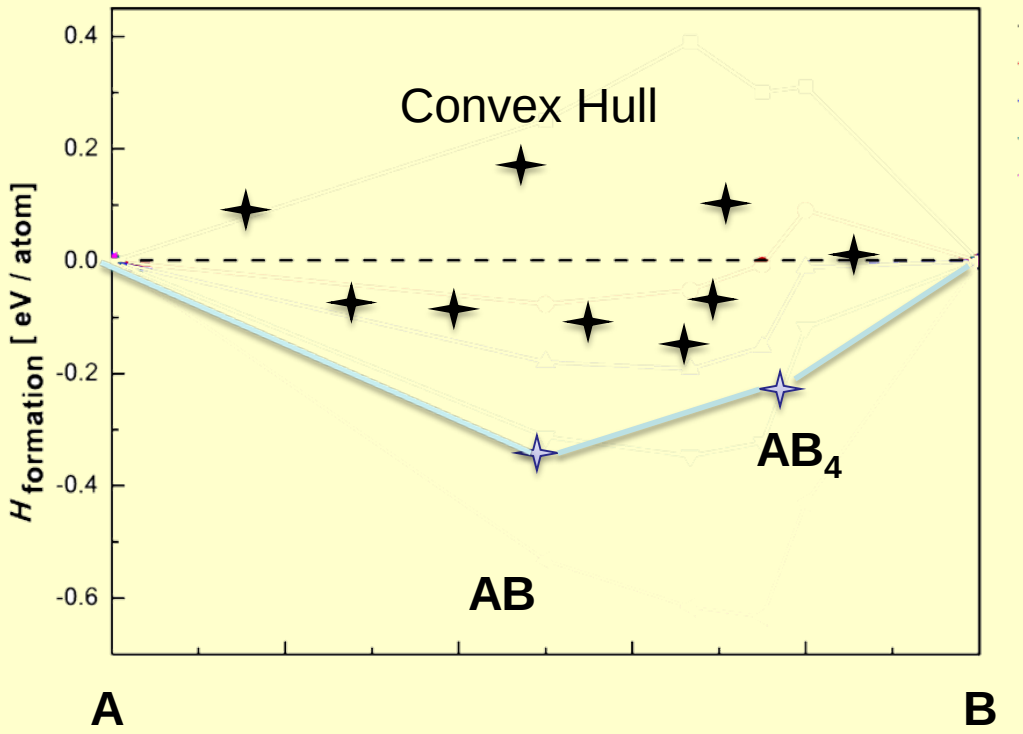
[Zeng & Oganov, *Adv. Energy Mat.*, 2015]

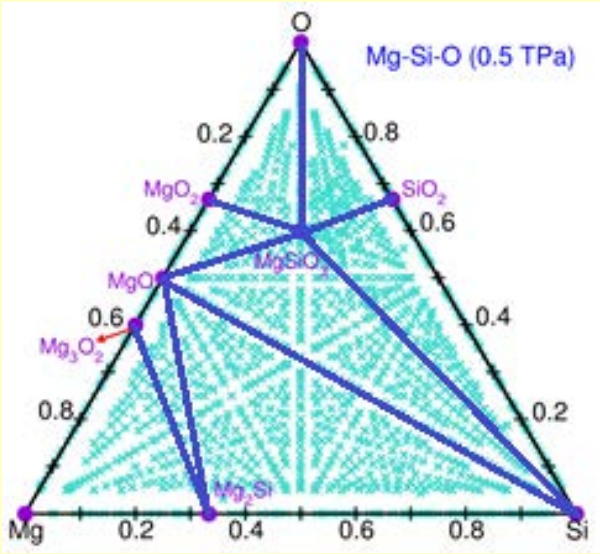
Prediction of stable structure for a given chemical composition is possible.

Now, let’s predict the chemical composition!

USPEX can automatically find all stable compounds in a multicomponent system.

Thermodynamic stability in variable-composition systems





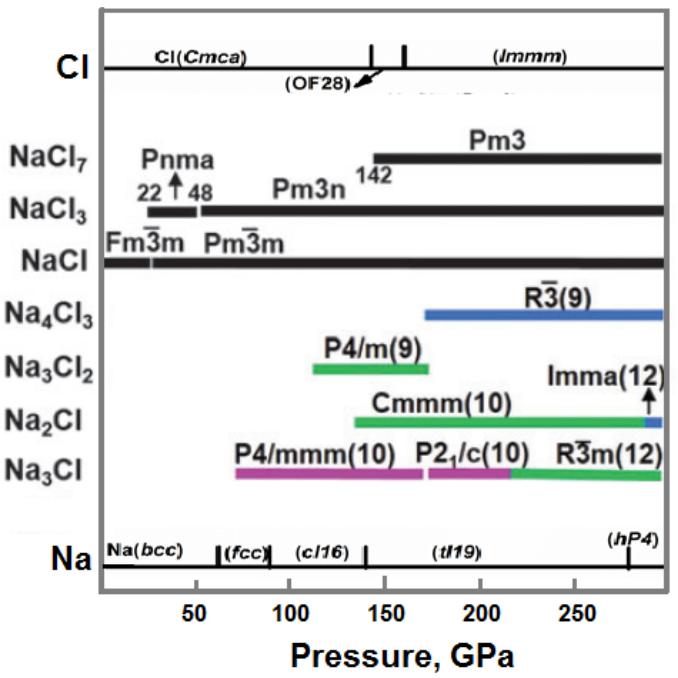
3-component convex hull: Mg-Si-O system at 500 GPa (Niu & Oganov, *Sci. Rep.* 2015)

Stable structure must be below all the possible decomposition lines !!

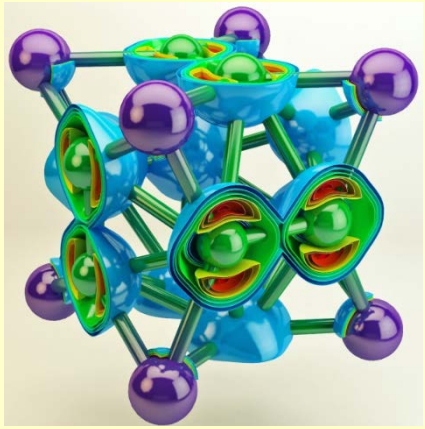
Na-Cl

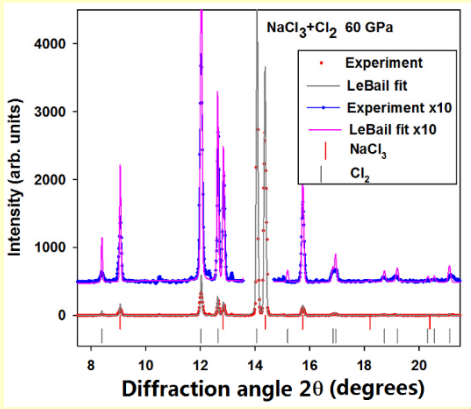
Predictive power of modern methods:

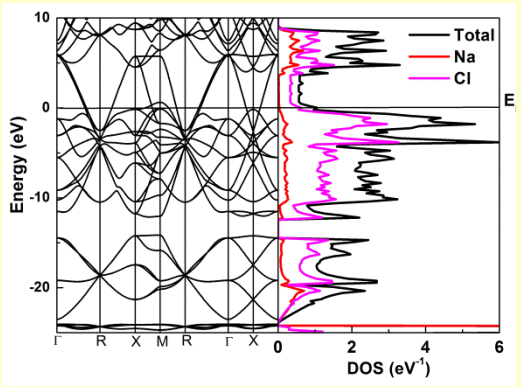
Na_3Cl, Na_2Cl, Na_3Cl_2, NaCl, $NaCl_3$, $NaCl_7$ are stable under pressure [Zhang, Oganov, et al. *Science*, 2013].

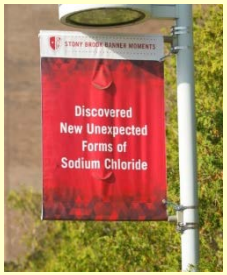


Stability fields of sodium chlorides









$NaCl_3$: atomic and electronic structure, and experimental XRD pattern

Chemical anomalies:

-Divalent Cl in Na_2Cl!
-Coexistence of metallic and ionic blocks in Na_3Cl!
-Positively charged Cl in $NaCl_7$!

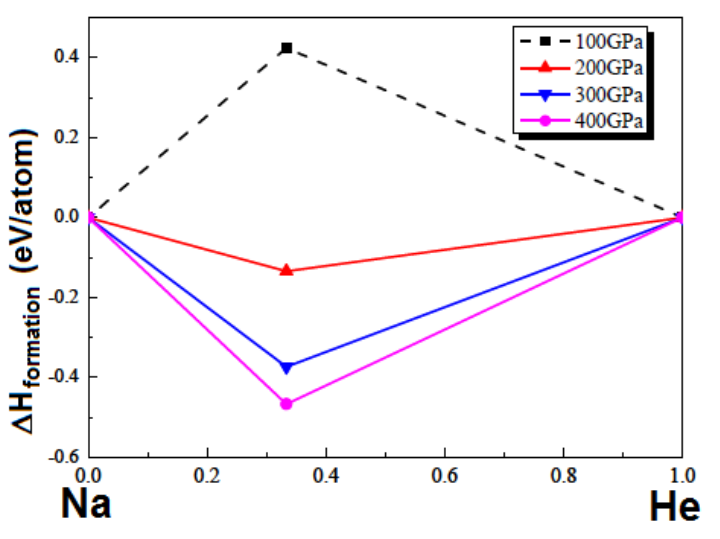
[Zhang, Oganov, et al., *Science* (2013)]
[Saleh & Oganov, *PCCP* (2015)]

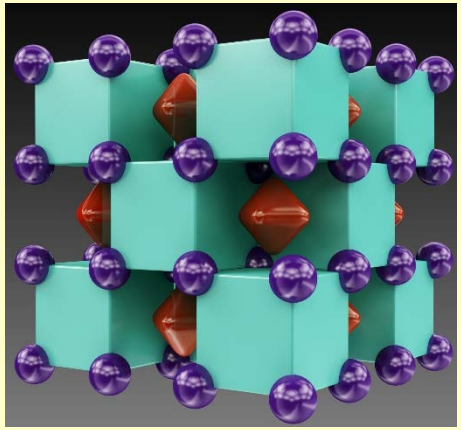
Helium chemistry? Yes!

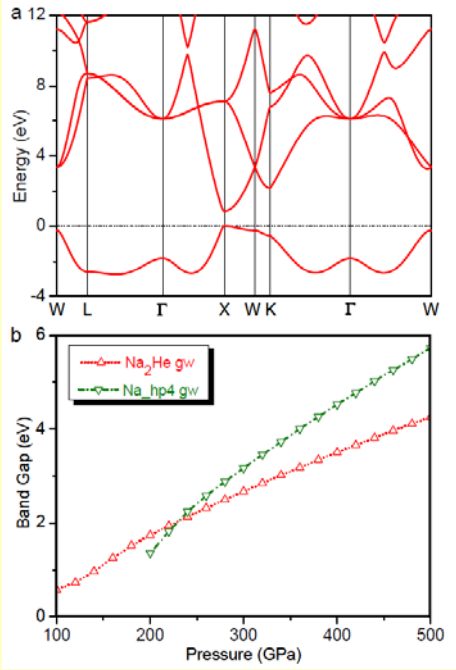
[Dong, Oganov, Goncharov, *Nature Chemistry* 2017]

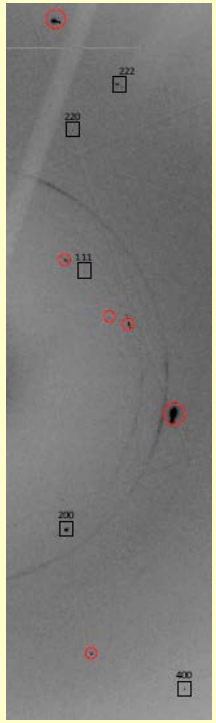
Na-He

- **Helium is the 2nd most abundant element in the Universe (24 wt.%).**
- **No stable compounds are known at normal conditions. Under pressure: van der Waals compound $NeHe_2$ (Loubeyre et al., 1993).**





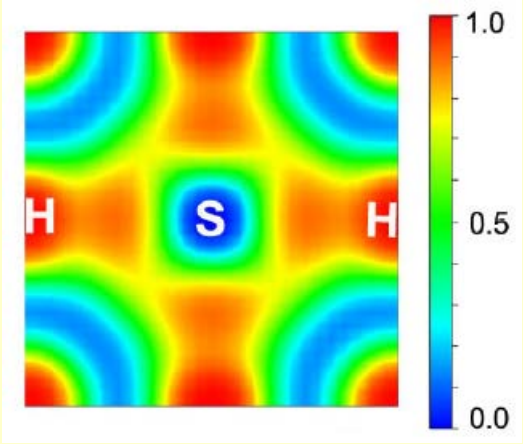




1. **Na_2He is stable at >113 GPa, at least up to 1000 GPa.**
2. **New stable helium compounds: Na_2HeO (Dong & Oganov, 2017); CaF_2He, MgF_2He (Liu, 2018).**

Highest-Tc superconductivity: new record, 203 Kelvin (Duan et al., *Sci. Rep.* 4, 6968 (2014))

H-S



OPEN

SUBJECT AREAS:
THEORY AND COMPUTATION
CONDENSED MATTER PHYSICS

Received
7 July 2014

Accepted
29 September 2014

Published
10 November 2014

Correspondence and requests for materials should be addressed to T.C. (cuitian@jlu.edu.cn)

Pressure-induced metallization of dense $(H_2S)_2H_2$ with high-T_c superconductivity

Defang Duan[1,2], Yunxian Liu[1], Fubo Tian[1], Da Li[1], Xiaoli Huang[1], Zhonglong Zhao[1], Hongyu Yu[1], Bingbing Liu[1], Wenjing Tian[2] & Tian Cui[1]

[1]State Key Laboratory of Superhard Materials, College of physics, Jilin University, Changchun, 130012, P. R. China, [2]State Key Laboratory of Supramolecular Structure and Materials, Jilin University, Changchun, 130012, P. R. China.

The high pressure structures, metallization, and superconductivity of recently synthesized H_3-containing compounds $(H_2S)_2H_2$ are elucidated by *ab initio* calculations. The ordered crystal structure with *P*1 symmetry is determined, supported by the good agreement between theoretical and experimental X-ray diffraction data, equation of states, and Raman spectra. The *Cccm* structure is favorable with partial hydrogen bond symmetrization above 37 GPa. Upon further compression, H_2 molecules disappear and two intriguing metallic structures with *R*3*m* and *Im*-3*m* symmetries are reconstructive above 111 and 180 GPa, respectively. The predicted metallization pressure is 111 GPa, which is approximately one-third of the currently suggested metallization pressure of bulk molecular hydrogen. Application of the Allen-Dynes-modified McMillan equation for the *Im*-3*m* structure yields high T_c values of 191 K to 204 K at 200 GPa, which is among the highest values reported for H_2-rich van der Waals compounds and MH_3 type hydride thus far.

Conventional superconductivity at 203 kelvin at high pressures in the sulfur hydride system

A. P. Drozdov, M. I. Eremets, I. A. Troyan, V. Ksenofontov & S. I. Shylin

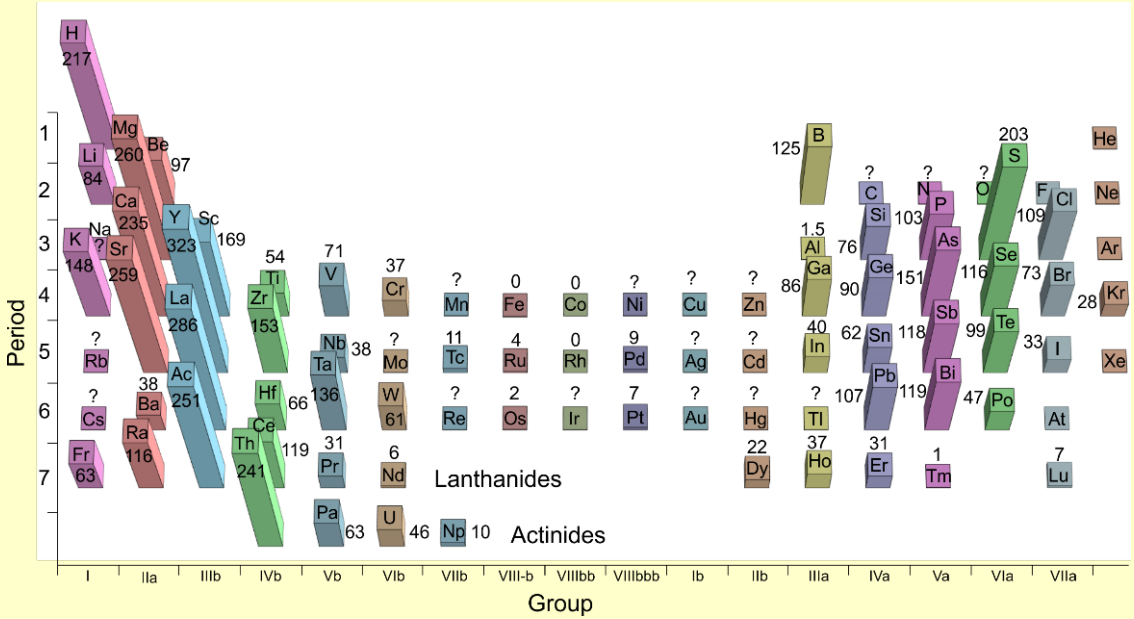
Nature (2015) | doi:10.1038/nature14964
Received 25 June 2015 | Accepted 22 July 2015 | Published online 17 August 2015

A superconductor is a material that can conduct electricity without resistance below a superconducting transition temperature, T_c. The highest T_c that has been achieved to date is in the copper oxide system[1]: 133 kelvin at ambient pressure[2] and 164 kelvin at high pressures[3]. As the nature of superconductivity in these materials is still not fully understood (they are not conventional superconductors), the prospects for achieving still higher transition temperatures by this route are not clear. In contrast, the Bardeen–Cooper–Schrieffer theory of conventional superconductivity gives a guide for achieving high T_c with no theoretical upper bound—all that is needed is a favourable combination of high-frequency phonons, strong electron–phonon coupling, and a high density of states[4]. These conditions can in principle be fulfilled for metallic hydrogen and covalent compounds dominated by hydrogen[5, 6], as hydrogen atoms provide the necessary high-frequency phonon modes as well as the strong electron–phonon coupling. Numerous calculations support this idea and have predicted transition temperatures in the range 50–235 kelvin for many hydrides[7], but only a moderate T_c of 17 kelvin has been observed experimentally[8]. Here we investigate sulfur hydride[9], where a T_c of 80 kelvin has been predicted[10]. We find that this system transforms to a metal at a pressure of approximately 90 gigapascals. On cooling, we see signatures of superconductivity: a sharp drop of the resistivity to zero and a decrease of the transition temperature with magnetic field, with magnetic susceptibility measurements confirming a T_c of 203 kelvin. Moreover, a pronounced isotope shift of T_c in sulfur deuteride is suggestive of an electron–phonon mechanism of superconductivity that is consistent with the Bardeen–Cooper–Schrieffer scenario. We argue that the phase responsible for high-T_c superconductivity in this system is likely to be H_3S, formed from H_2S by decomposition under pressure. These findings raise hope for the prospects for achieving room-temperature superconductivity in other hydrogen-based materials.

- **Old record Tc=135 K (Schilling, 1993) is broken: theorists (T. Cui, 2014) predicted new compound H_3S with Tc~200 K.**
- **Confirmed by A. Drozdov et al. (*Nature* 525, 73 (2015)).**

Superconductivity is linked with Mendeleev's Table

[Semenok & Oganov, *JPCL*, 2018; *Curr. Opinion Solid St. Mater. Sci.*, 2020]

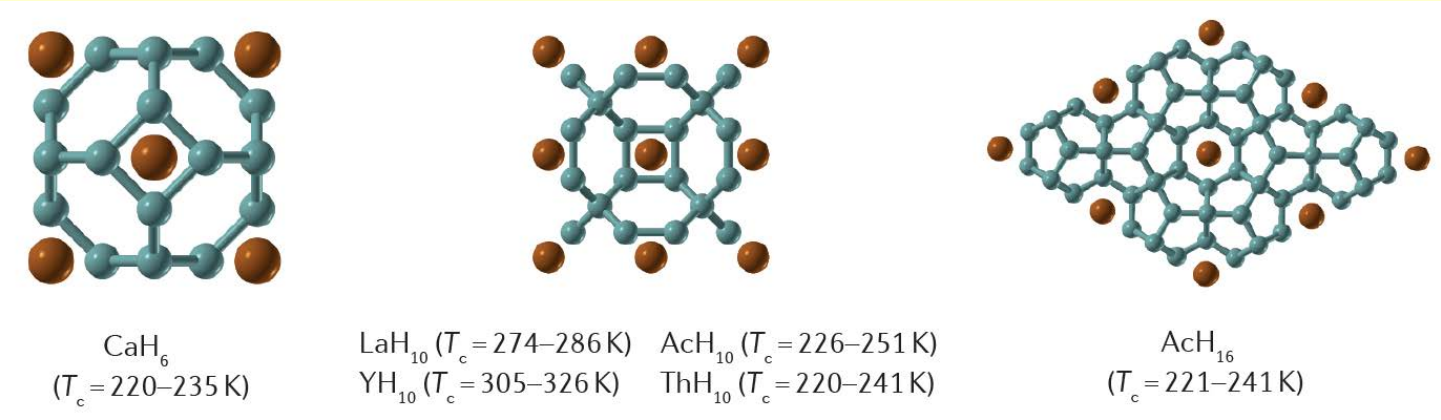


LaH_{10}: record Tc (260 K @ 190 GPa) (Somayazulu et al., 2019).

Test of idea: Th and Ac hydrides have high-Tc superconductivity.

ThH_{10}: Tc=241 K at 100 GPa (Kvashnin & Oganov, 2018).

Distribution of Tc for hydrides

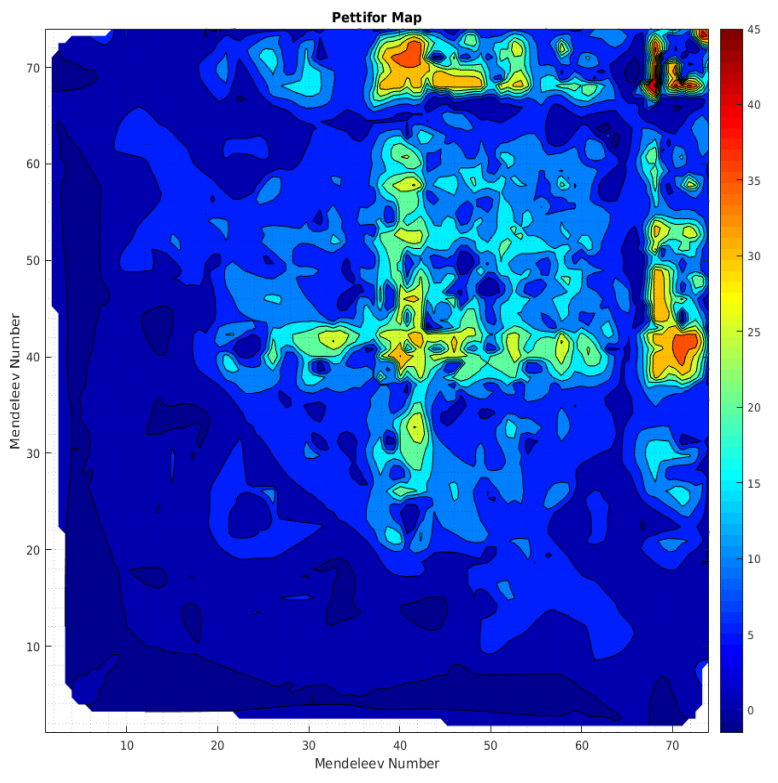


We can simultaneously optimize composition, structure, stability and other properties for a given chemical system.

Now, let’s predict the best material(s) among all possible chemical systems!

Mendelevian Search – breakthrough method for discovering best materials among all possible compounds

[Allahyari & Oganov, *NPJ Comp. Mat.*, 2020]



- **118 elements**
- **7021 binary systems**
- **273937 ternaries**
- **In each system - ∞ possible structures**

Mendeleev Number – a way to arrange elements and compounds by properties

[Pettifor, 1984; Allahyari & Oganov, *NPJ Comp. Mat.*, 2020]

Mendeleev Number	Atom	Mendeleev Number	Atom	Mendeleev Number	Atom
1	Fr	32	Tl	62	Po
2	Cs	33	U	63	Fe
3	Rb	34	Pa	64	Cu
4	K	35	Zr	65	Co
5	Ra	36	Pu	66	As
6	Ba	37	Np	67	Ni
7	Sm	38	Nb	68	Kr
8	Gd	39	Ta	69	Mo
9	Eu	40	In	70	I
10	Sr	41	Pb	71	Pd
11	Tm	42	Cd	72	Ir
12	Pm	43	Xe	73	Os
13	Ca	44	Ti	74	P
14	Na	45	Al	75	Ru
15	Ac	46	Bi	76	Pt
16	La	47	Sn	77	At
17	Yb	48	Hg	78	Rh
18	Tb	49	Zn	79	W
19	Y	50	Ga	80	Rn
20	Dy	51	V	81	Se
21	Ho	52	Mn	82	B
22	Ce	53	Sb	83	Au
23	Er	54	Te	84	S
24	Li	55	Cr	85	Br
25	Th	56	Ag	86	H
26	Lu	57	Be	87	C
27	Pr	58	Ge	88	Cl
28	Nd	59	Re	89	N
29	Mg	60	Si	90	O
30	Sc	61	Tc	91	F
31	Hf				

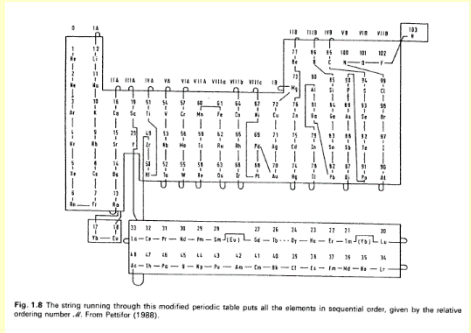
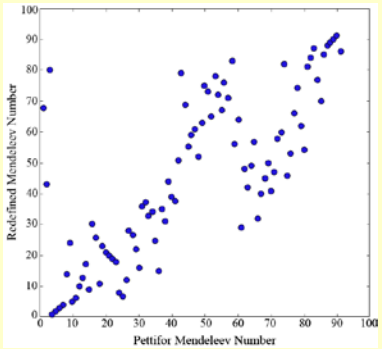
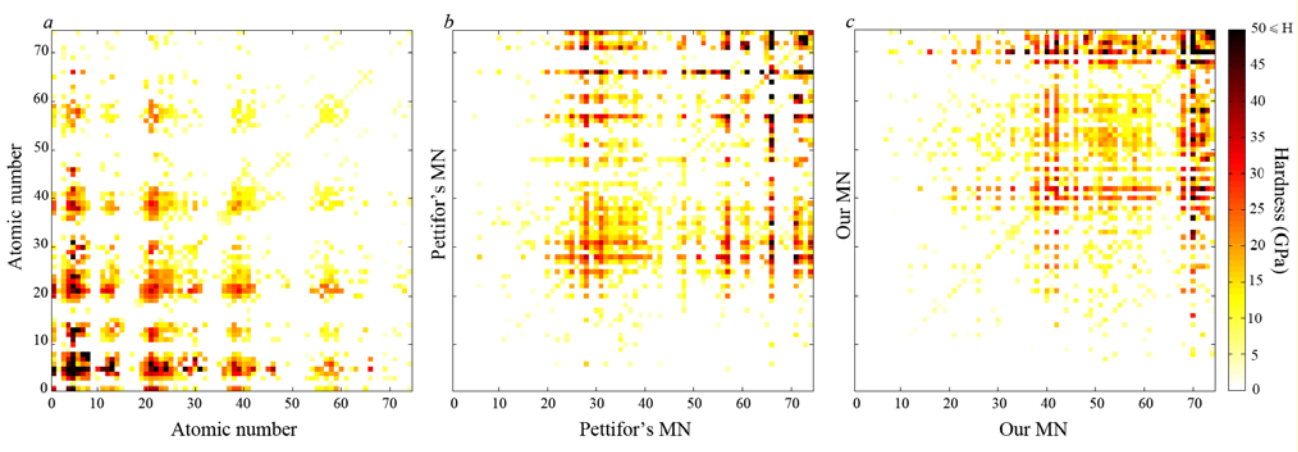


Fig. 1.8 The string running through this modified periodic table puts all the elements in sequential order, given by the relative ordering number, A'. From Pettifor (1988).

Pettifor's construction



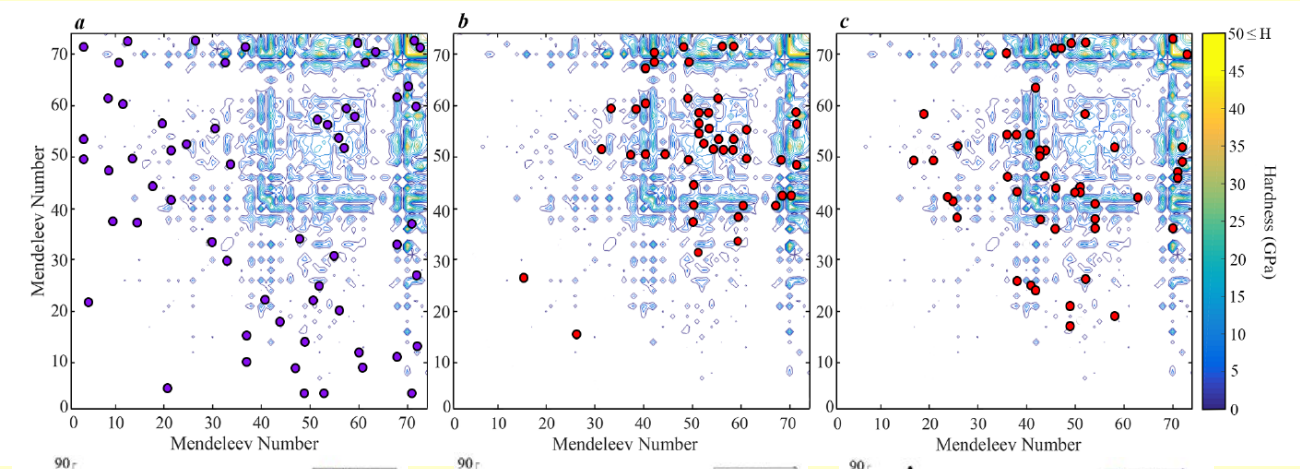
Comparison with Pettifor's numbers

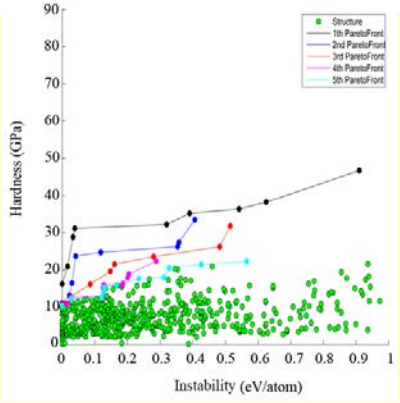


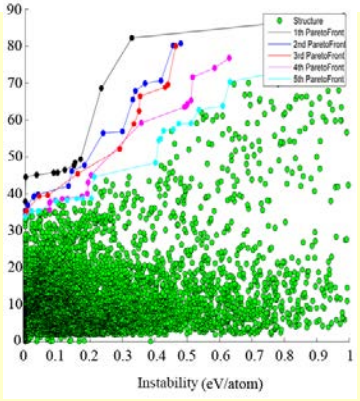
Grouping of hardness by (a) sequential number, (b) Pettifor's Mendeleev number, (c) our Mendeleev number

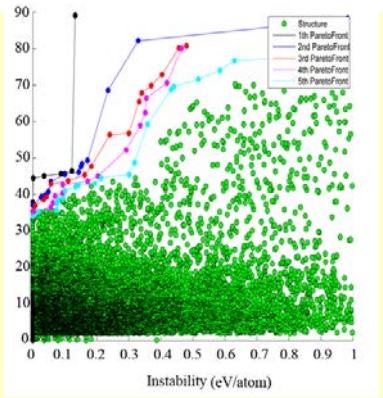
Mendelevian search for the hardest possible material: diamond and lonsdaleite are found!

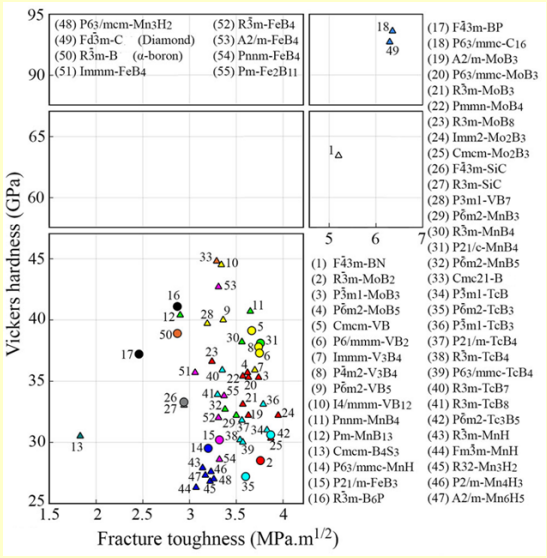
1st generation **5th generation** **10th generation**





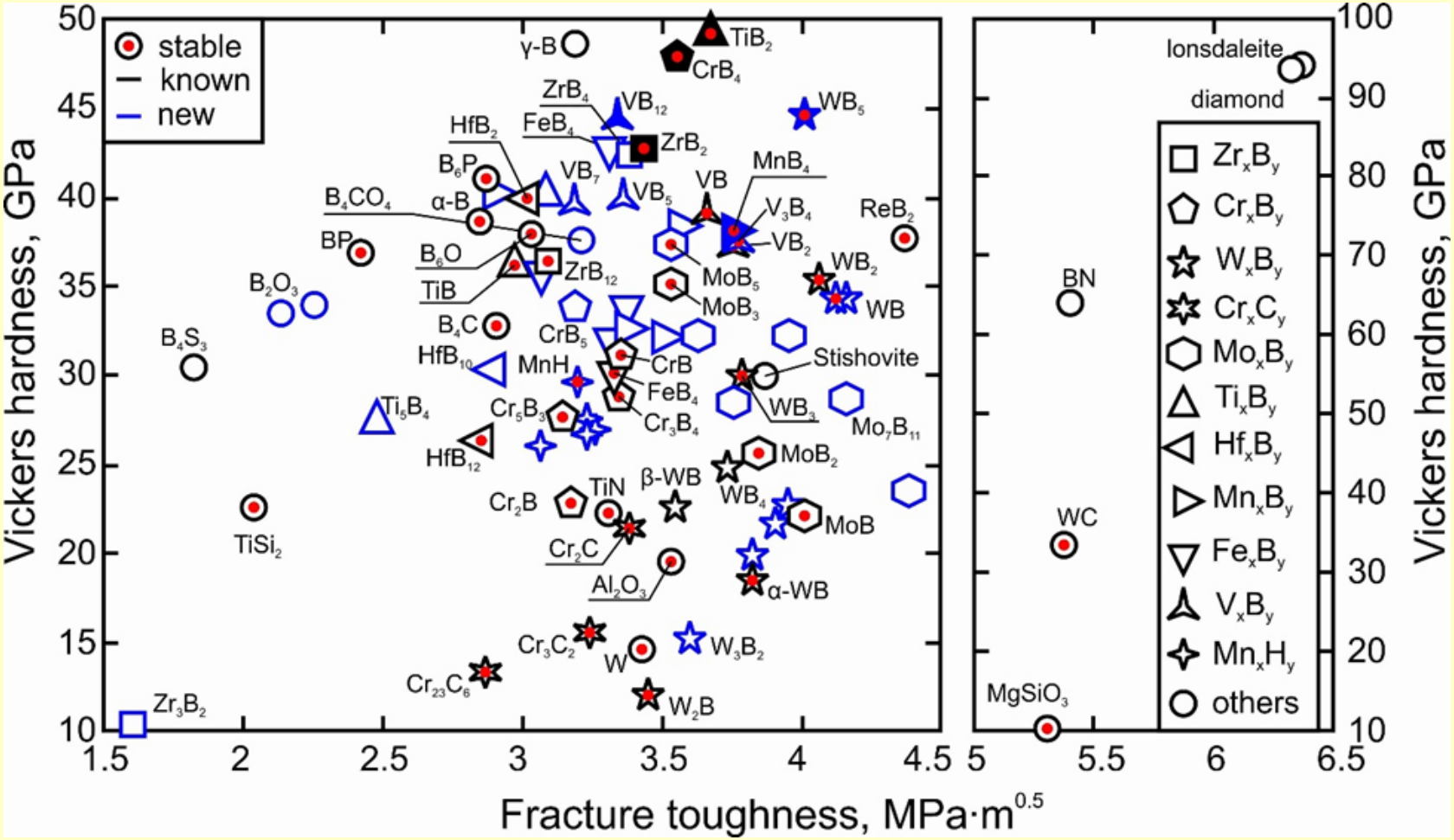






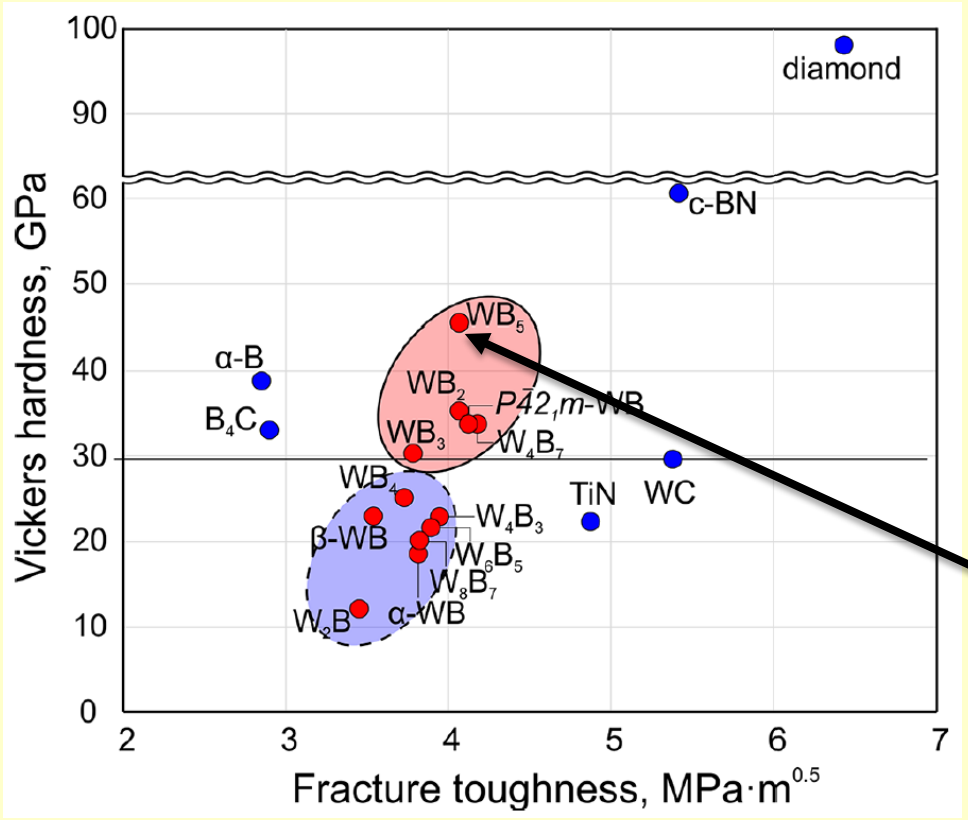
“Treasure” map of superhard materials

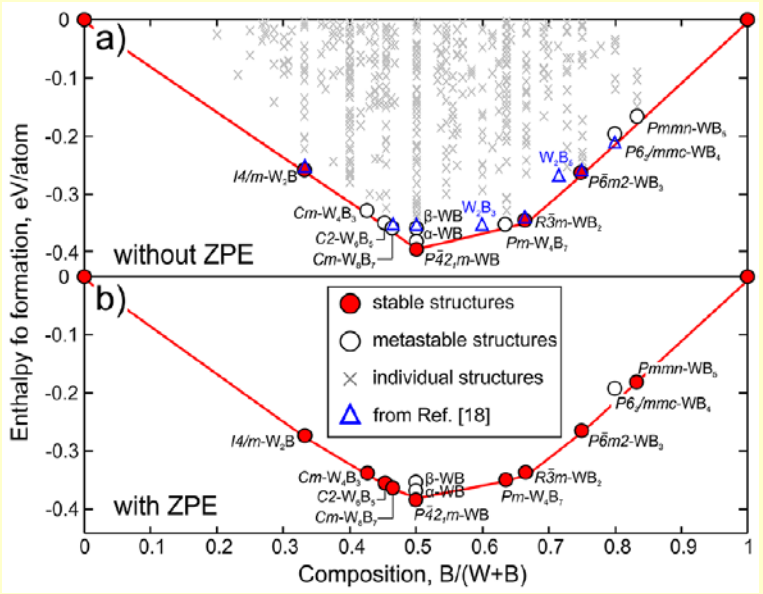
[Kvashnin, Allahyari, Oganov, *J. Appl. Phys.*, 2019]

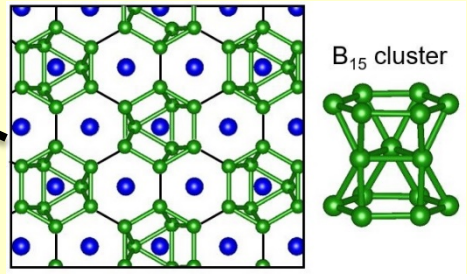


WB_{5-x}: hard & tough material

[Kvashnin & Oganov, *J. Phys. Chem. Lett.*, 2018; *Advanced Science*, 2020]





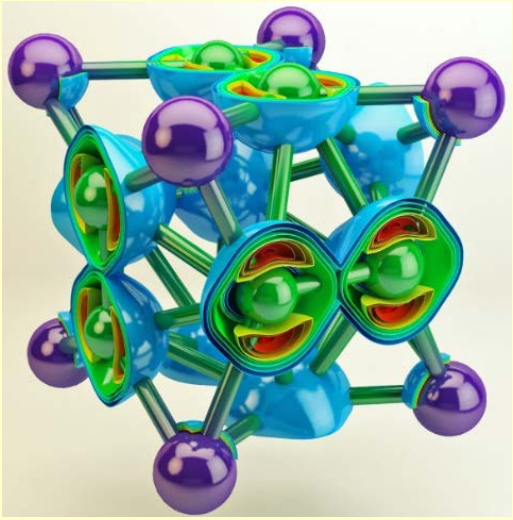


Structure of WB_5

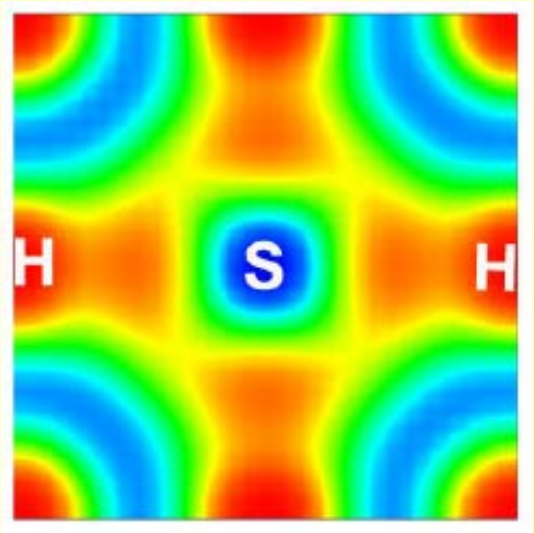
Tungsten carbide WC - standard

Synthesized by V. Filonenko

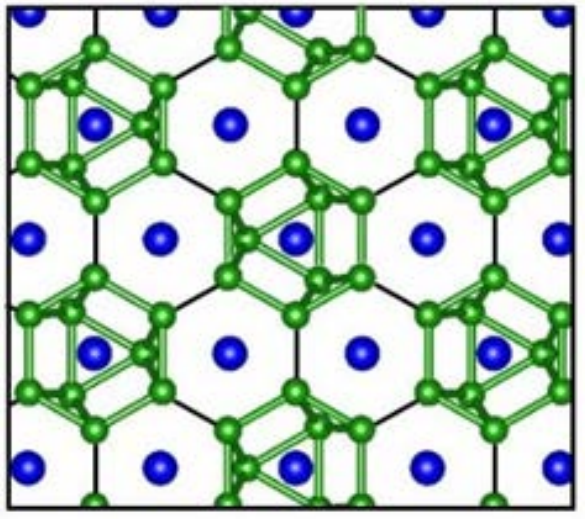
Advanced algorithms predict new supermaterials and help us understand nature



Unusual chemistry at extreme conditions



New record of high-Tc superconductivity



New superhard materials

Our team. Where great minds do NOT think alike


Q. Zhu


X. Dong


A. Goncharov


V.A. Blatov

USPEX Computational Materials Discovery