第18回情報論的学習理論ワークショップ (IBIS2015)
The 18th Information-Based Induction Sciences Workshop
2015.11.25～28 つくば国際会議場

マテリアルズインフォマティクスの現状と将来展望

[1]京都大学　工学研究科　材料工学専攻
[2]京都大学　構造材料元素戦略拠点(ESISM)
[3]ﾌｧｲﾝｾﾗﾐｯｸｽｾﾝﾀｰ(JFCC)　ナノ構造研究所
[4] NIMS 情報統合型物質・材料研究拠点(MI[2]I)

田中　功

米国のチャレンジ

2011 June

材料ゲノム：アメリカ製造業のルネサンスへ

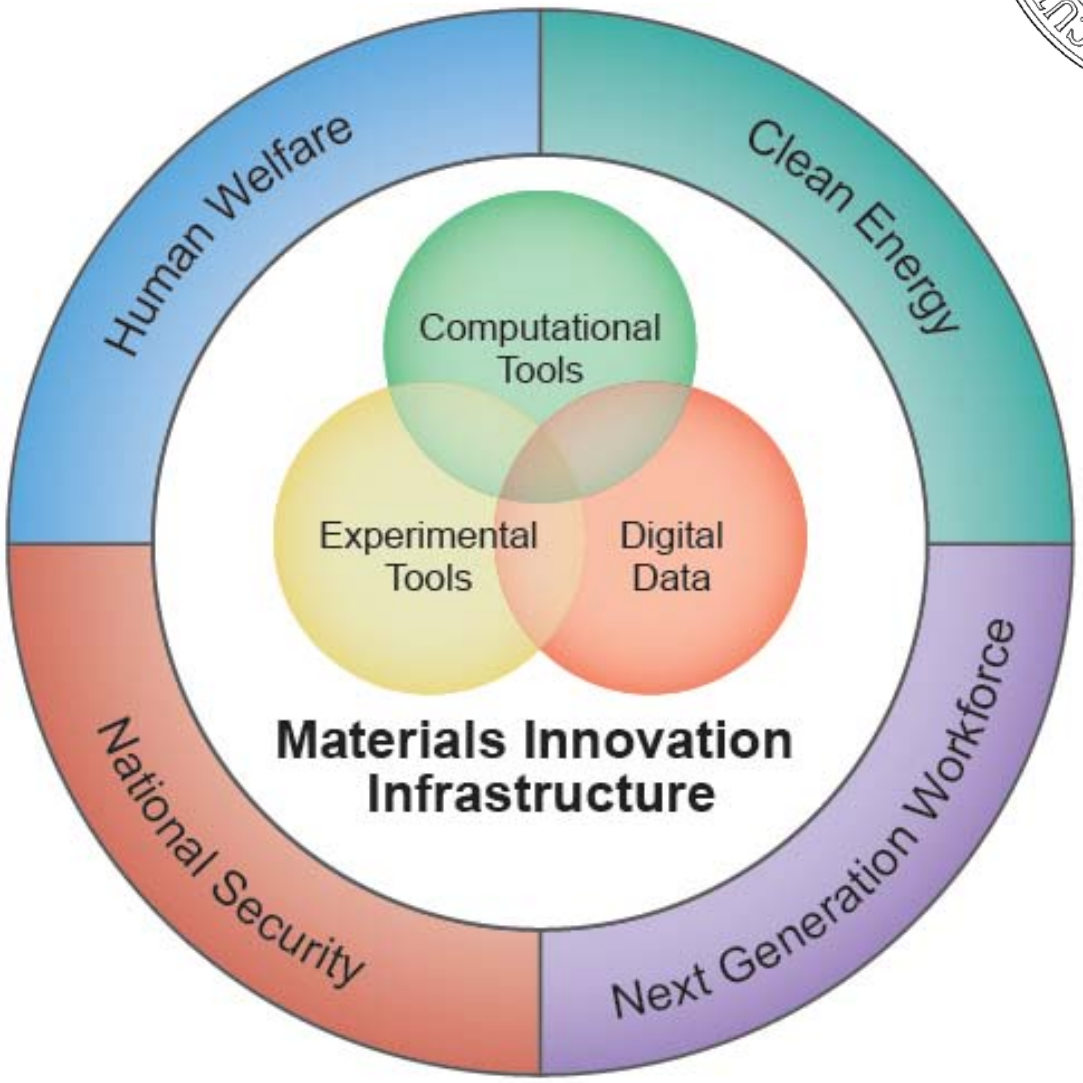
To double the speed with which we discover, develop, and manufacture new materials.　新材料の発見・開発・製品化のスピードを２倍に.

FY12 budget includes $100 million to launch the Materials Genome Initiative.

年間予算　1億USドル＝120億日本円

Materials Genome Initiative
for Global Competitiveness

June 2011



EUのチャレンジ１

NOMAD
NOVEL MATERIALS DISCOVERY

EUROPEAN CENTER OF EXCELLENCE (CoE) FOR NOVEL MATERIALS DISCOVERY (NOMAD)

Funded by the EU Commission under the Horizon2020 programs

http://nomad-lab.eu/

The NOMAD CoE develops a ***Materials Encyclopedia*** and ***Big-Data Analytics*** tools for materials science and engineering. Eight complementary research groups of highest scientific standing in computational materials science along with four high-performance computer centers form the synergetic core of this CoE.

EUのチャレンジ２

MARVEL

Materials' Revolution: Computational Design and Discovery of Novel Materials

Swiss National Science Foundation

Vertical Projects

VP1 - Novel Materials Physics

VP2 - Novel Materials Applications

Nicola Marzari
Director
EPFL, Lausanne

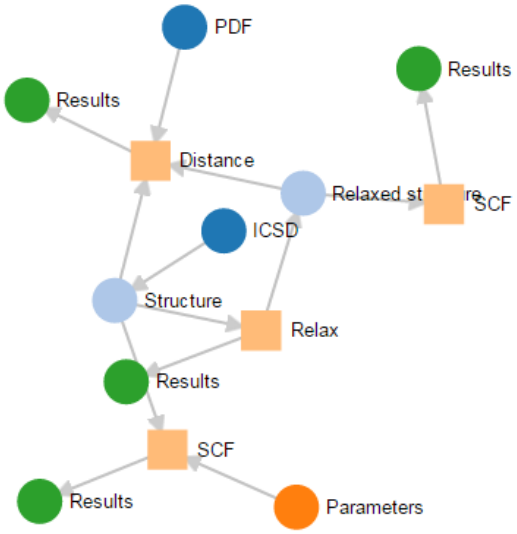
Horizontal Projects

HP3 - Advanced Quantum Simulations

HP4 - Advanced Sampling Methods

HP5 - Materials Informatics

AiiDA
Automated Interactive Infrastructure
Database for Computational Scie



Platform Projects

PP6 - Informatics

PP7 - Experiments

日本のチャレンジ

2012 科学研究費・新学術領域
『ナノ構造情報』

2013 採択

NANO INFO

2014 実施中

2015

2016

2017

『マテリアルズ・インフォマティクス』
JSTでの調査

NIMSｲﾉﾍﾞｰｼｮﾝﾊﾌﾞ開始

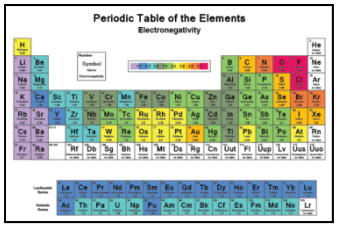
JSTさきがけ開始

マテリアルズインフォマティクスの常法

ニーズ

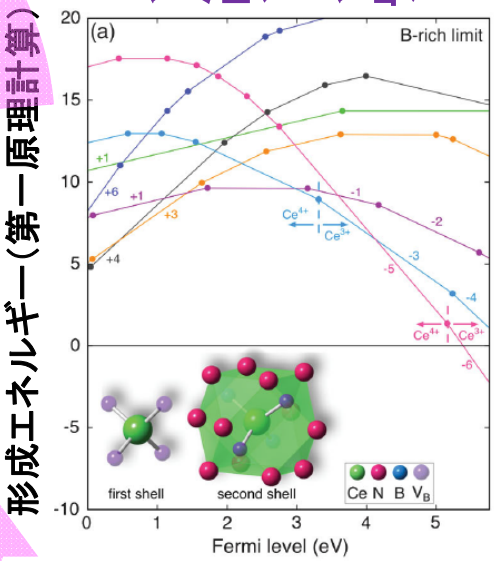
材料設計

合成実験

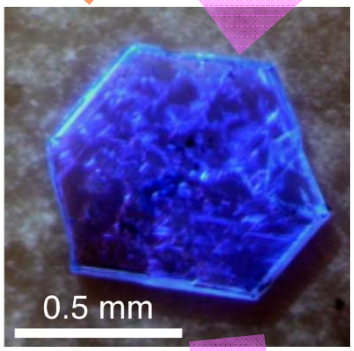


最適化

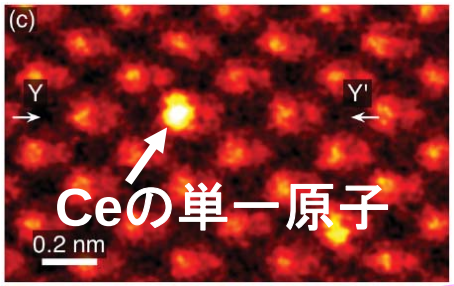
シミュレーション



特性



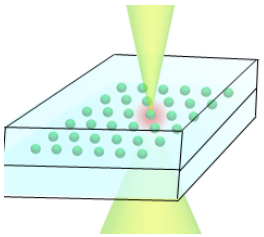
情報論的
学習理論

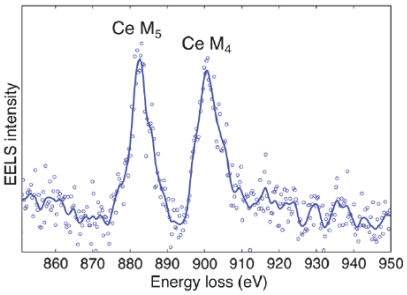


解析実験

理論計算

シーズ



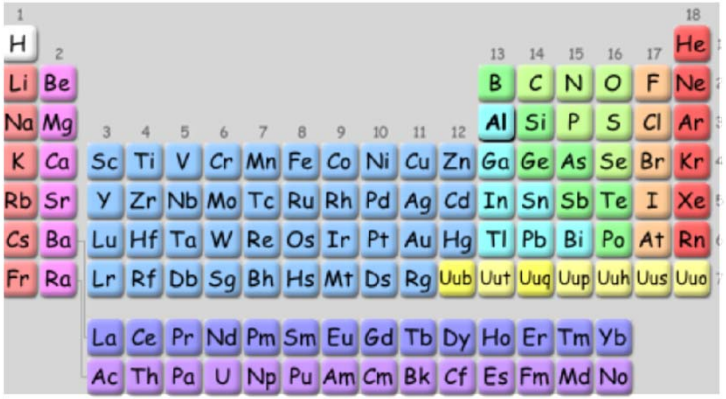


解析

物理法則：
古典力学，電磁気学，
熱・統計力学，量子力学
等に立脚

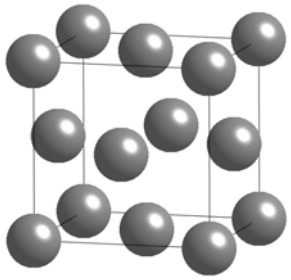
NANO INFO

（無機）材料科学のキャストの例： 結晶性物質



Periodic Table

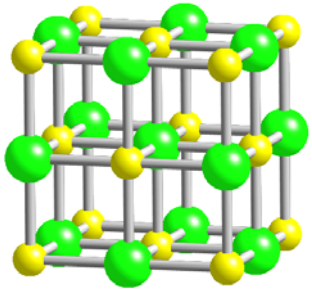
Al



face centered cubic structure

light metal tube

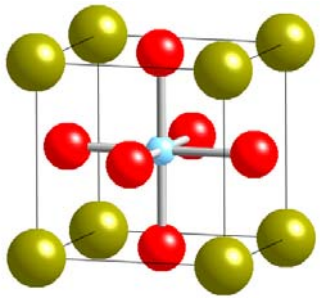
NaCl



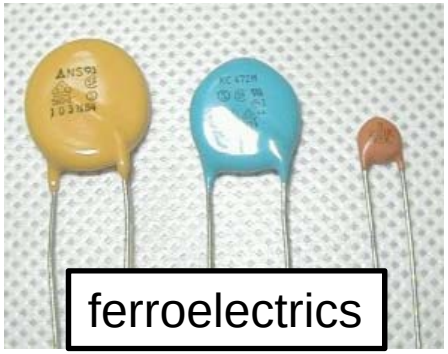
rock salt structure

table salt

$BaTiO_3$

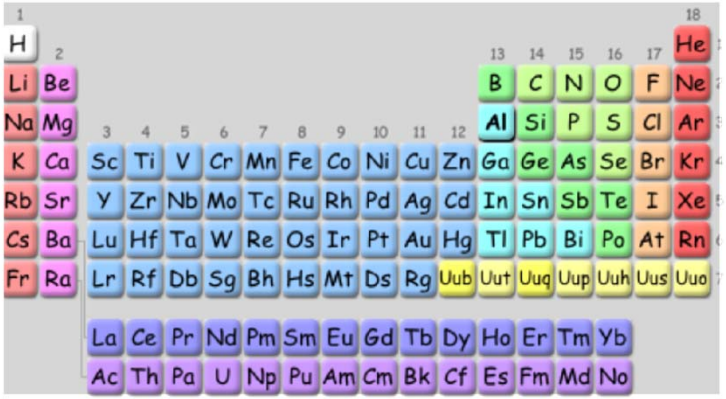


perovskite structure



ferroelectrics

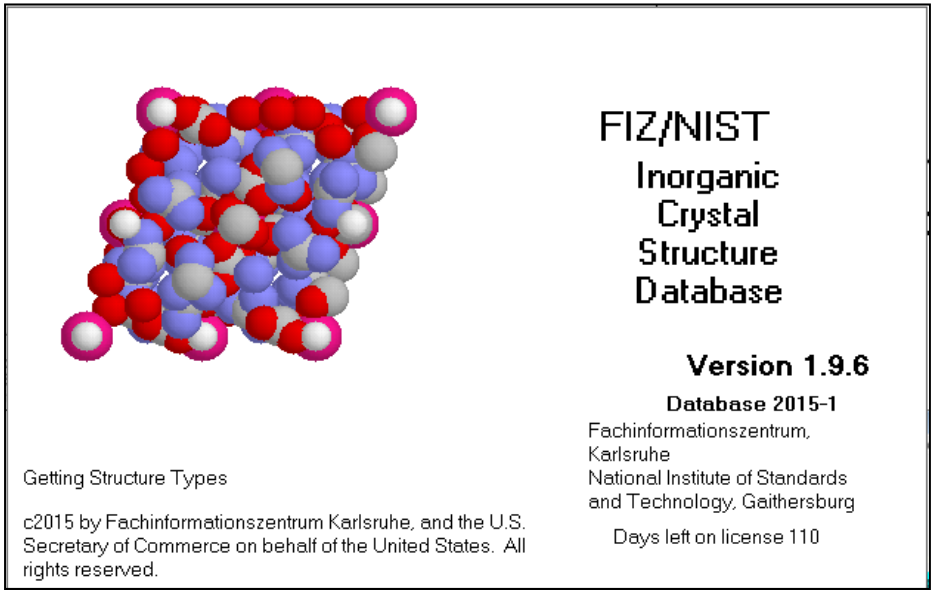
Enormous number of chemical combinations



Periodic Table

Number of chemical elements	Number of chemical combinations (only for simple composition ratio)
1	~100
2	~100,000
3	~10,000,000
4	~1,000,000,000

Inorganic Crystal Structure Database (ICSD)



World largest database
for known inorganic crystals.

177,000 crystal structures
→ 55,000 structures
excluding duplicates.

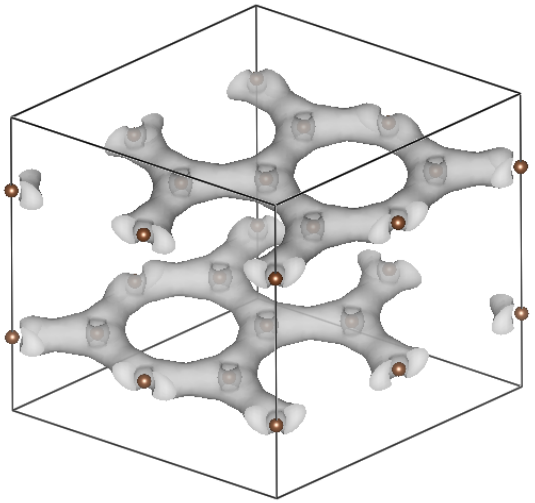
9,100 structure prototypes
(e.g. rock-salt, perovskite, ...)

Number of chemical elements	Number of structure prototypes in ICSD
1	120
2	1,700
3	4,700
4	4,300

First Principles Calculations

Useful to fill the gap between enormous number of chemical combinations and experimental database.

Electronic structure calculations entirely based on quantum theory without using empirical parameters.



Total energy

Force/Stress

Electronic/Magnetic structures

Physical properties

. . . .

First Principles Calculations

INPUT

Space Group

Unit Cell Parameters

Atoms / Atomic Number and Coordinates

Density Functional Theory Calculations

$$\left(-\frac{1}{2}\nabla^2 + V_{ion}(\vec{r}) + \int \frac{\rho(\vec{r}')}{|\vec{r}-\vec{r}'|}d\vec{r}' + v_{XC}(\vec{r})\right)\psi_i(\vec{r}) = \varepsilon_i\psi_i(\vec{r})$$

OUTPUT

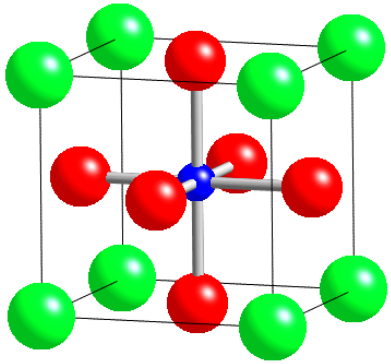
Total Energy

Stress / Force

Electronic Structures

structure optimization

Computational Cost 0 (CC0)



$SrTiO_3$

(cubic perovskite)

Space Group $Pm\bar{3}m$

1 x Sr + 1 x Ti + 3 x O

= 5 atoms/unit cell

Computational Cost

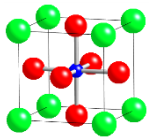
CC0 x 10 = CC1

Cost for First Principles Calculations

Computational Cost with 1 core

CC0	1min	Single shot GGA
CC1	1hour	Structure opt.
CC2		2x2x2 supercell Single shot GGA
CC3	1day	Phonon frequency
CC4	1 week	Lattice thermal conductivity
CC5	1 year	Diffusivity by FPMD simulation

1000 atoms cell Single shot GGA



PC cluster

5 nodes x 20 cores =100 cores

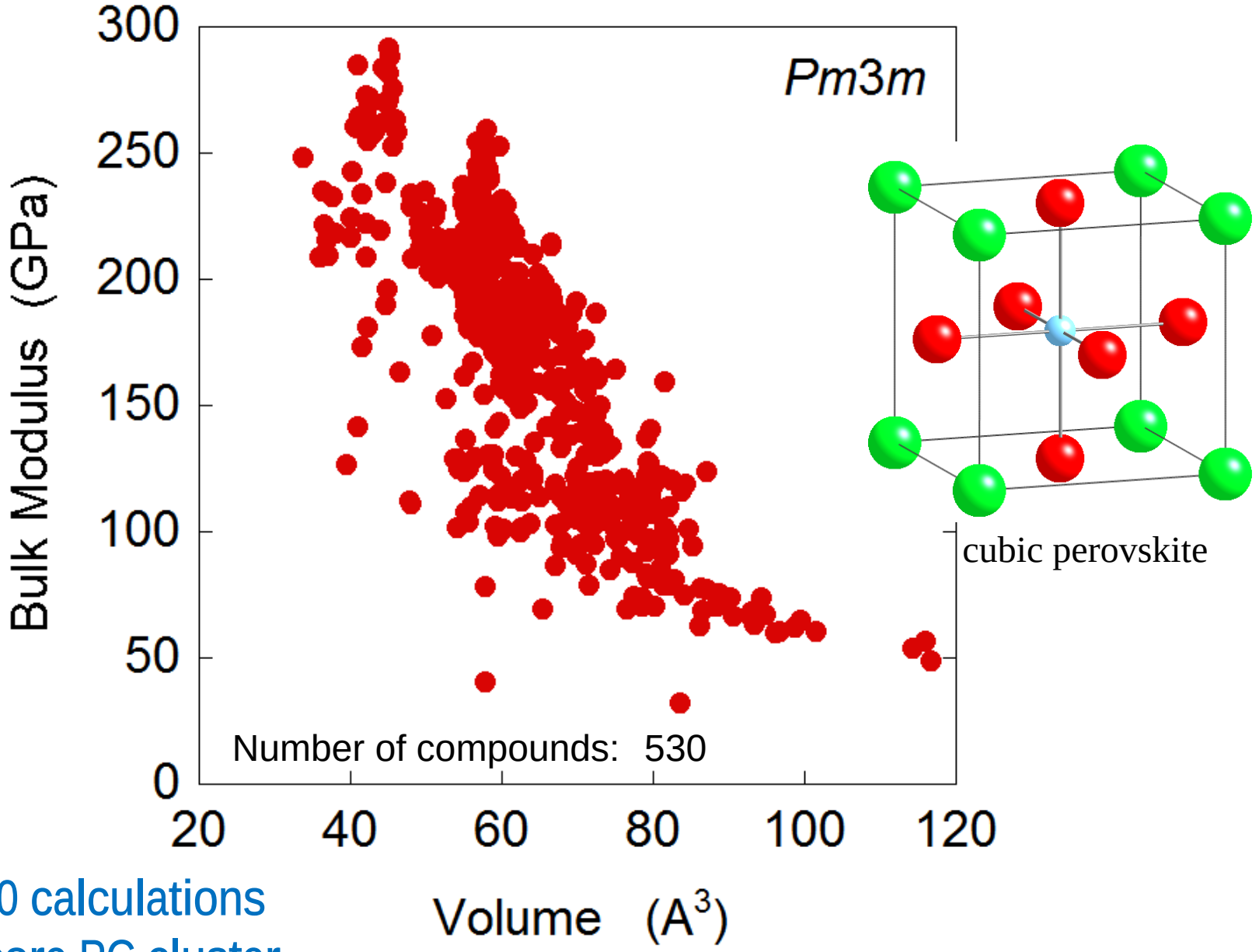
price ~ USD30k (JPY 4M)

10 nodes

Computational Cost with 100 cores

	1min
	1hour
CC5	1day
CC6	1 week
CC7	1 year

A set of DFT calculations during a lunch break

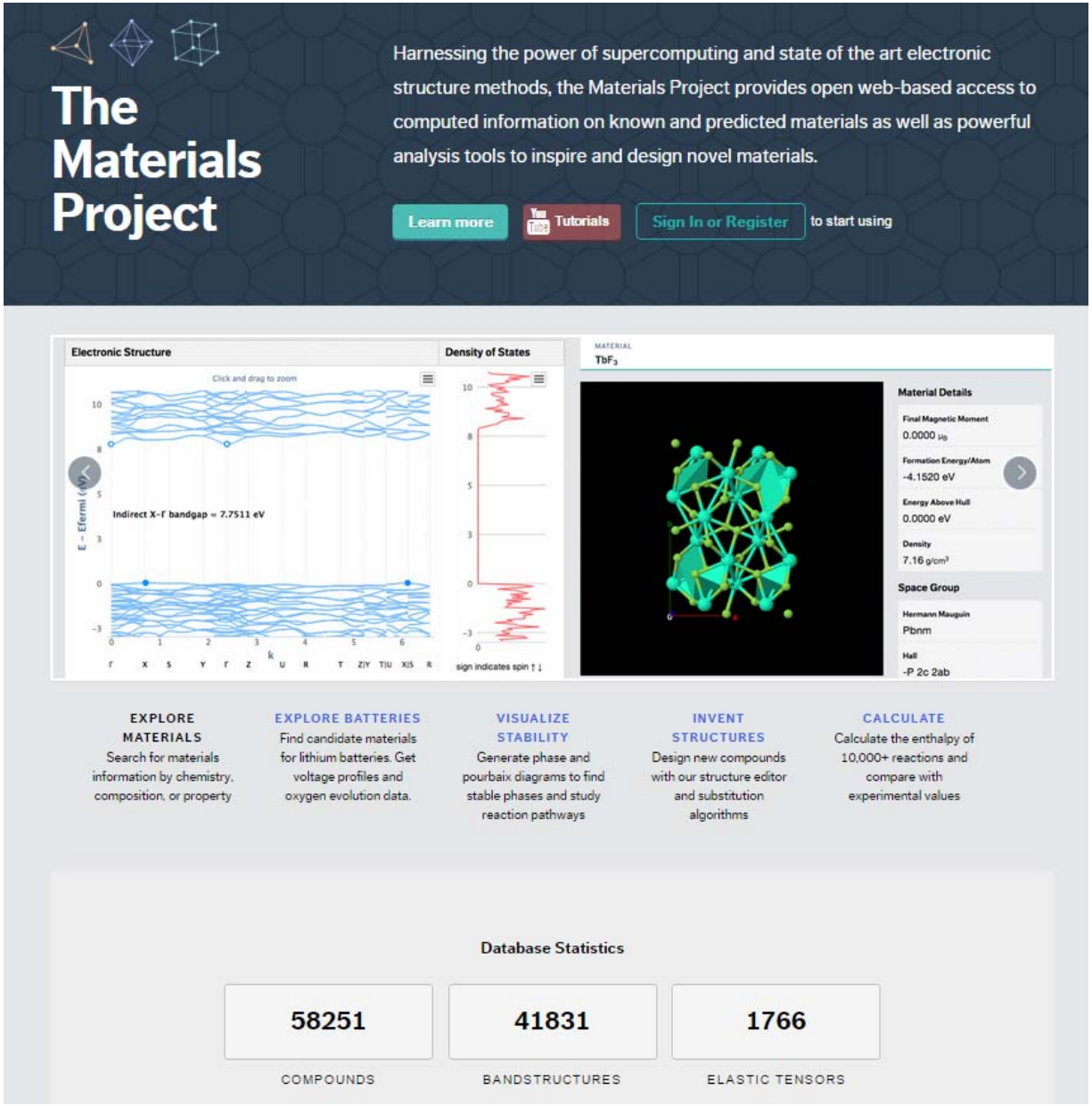


CC1 x 530 calculations
by a 100 core PC cluster

Materials Project @MIT

https://materialsproject.org/

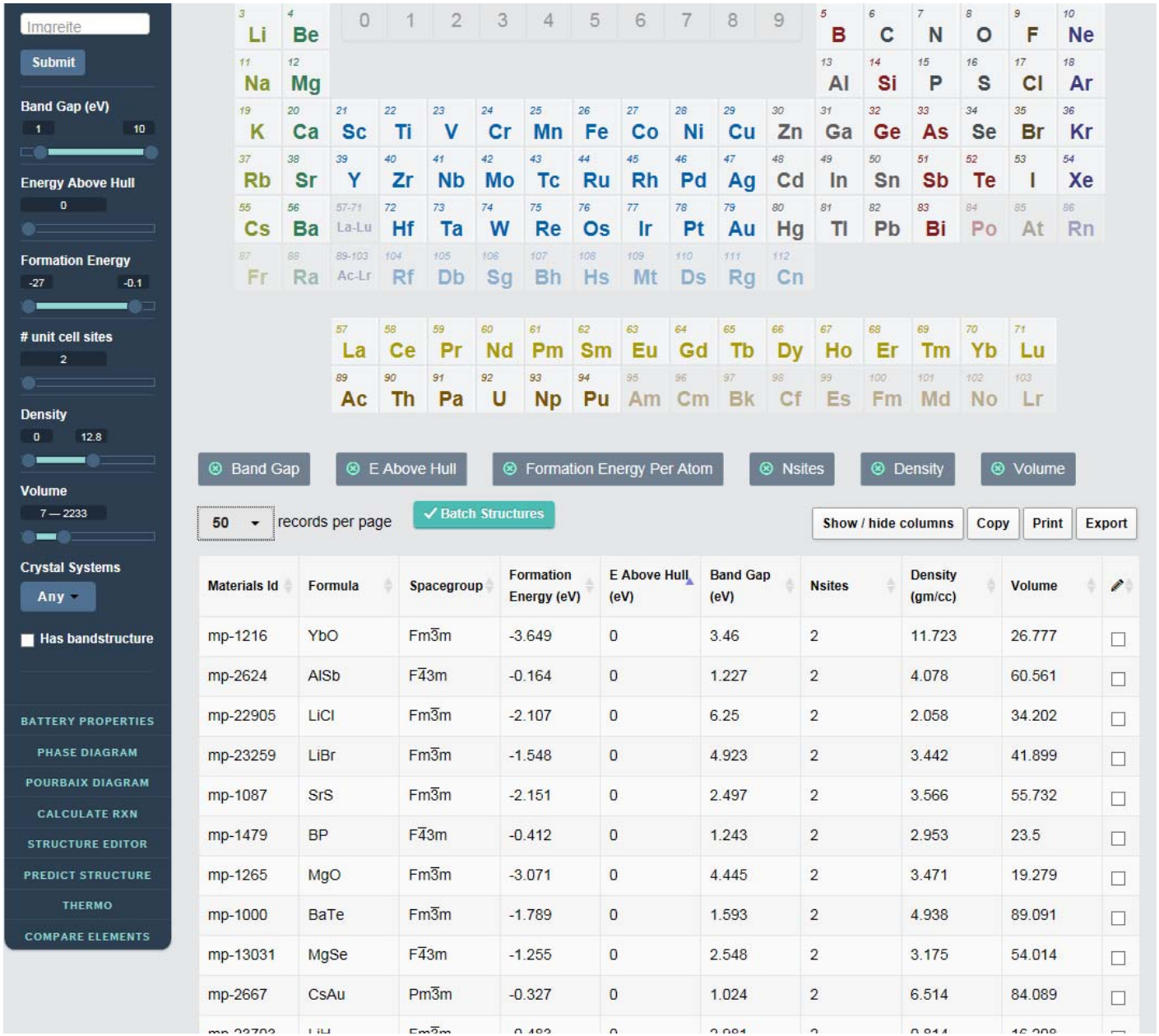
ICSD収録の構造データについて，生成エネルギーやバンドギャップなどの第一原理計算結果



Materials Project @MIT

https://materialsproject.org/

ICSD収録の構造データについて，生成エネルギーやバンドギャップなどの第一原理計算結果



Materials Id	Formula	Spacegroup	Formation Energy (eV)	E Above Hull (eV)	Band Gap (eV)	Nsites	Density (gm/cc)	Volume
mp-1216	YbO	$Fm\bar{3}m$	-3.649	0	3.46	2	11.723	26.777
mp-2624	AlSb	$F\bar{4}3m$	-0.164	0	1.227	2	4.078	60.561
mp-22905	LiCl	$Fm\bar{3}m$	-2.107	0	6.25	2	2.058	34.202
mp-23259	LiBr	$Fm\bar{3}m$	-1.548	0	4.923	2	3.442	41.899
mp-1087	SrS	$Fm\bar{3}m$	-2.151	0	2.497	2	3.566	55.732
mp-1479	BP	$F\bar{4}3m$	-0.412	0	1.243	2	2.953	23.5
mp-1265	MgO	$Fm\bar{3}m$	-3.071	0	4.445	2	3.471	19.279
mp-1000	BaTe	$Fm\bar{3}m$	-1.789	0	1.593	2	4.938	89.091
mp-13031	MgSe	$F\bar{4}3m$	-1.255	0	2.548	2	3.175	54.014
mp-2667	CsAu	$Pm\bar{3}m$	-0.327	0	1.024	2	6.514	84.089

Discovery of new low thermal conductivity materials

First-Principles Anharmonic Lattice-Dynamics Calculations and Bayesian Optimization




Atsuto Seko, Atsushi Togo, Hiroyuki Hayashi,

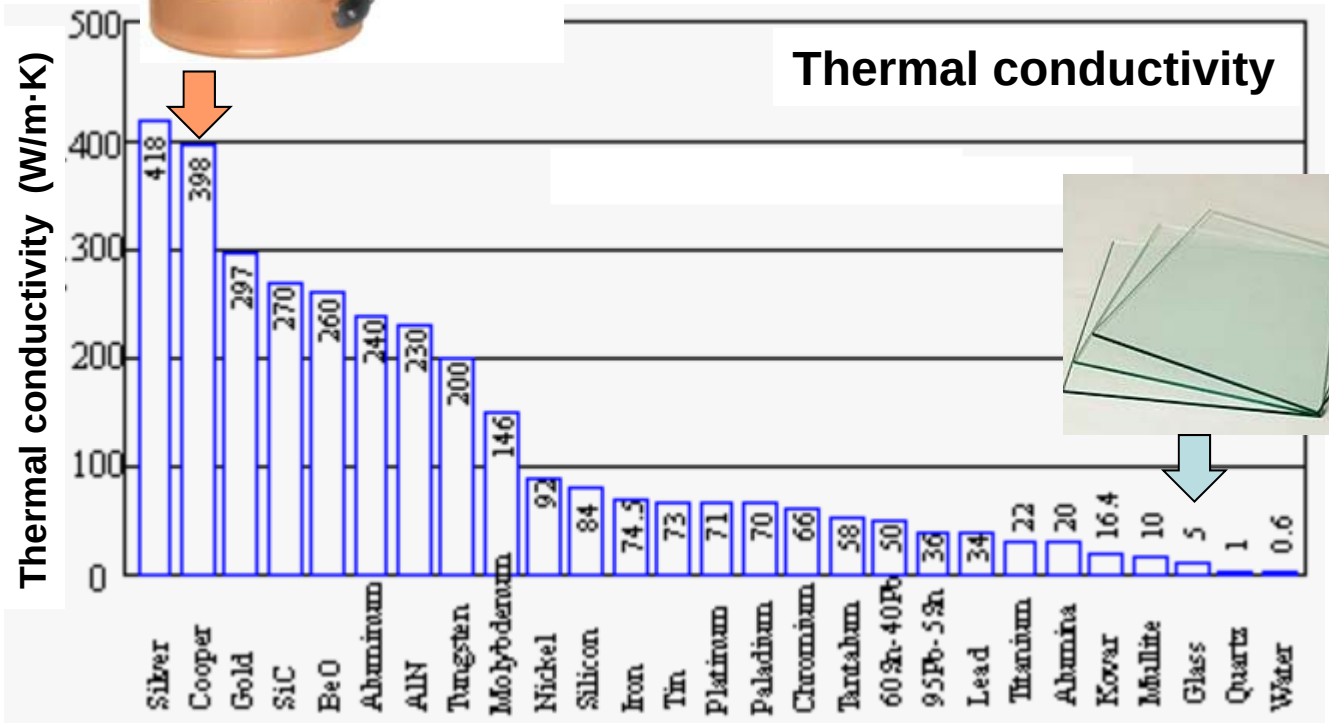
Koji Tsuda, Laurent Chaput, and IT

A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, and IT PRL 115, 205901 (2015)

Discovery of ultra-low thermal conductivity materials of 0.1W/m·K level in ICSD database (55,000 crystals)

diamond

>2000



Background

Thermal Conductivity (TC)

$$\kappa = \kappa_{electonic} + \kappa_{lattice}$$

Lattice Thermal Conductivity (LTC)

- ✓ **Reliable experimental dataset is limited (< 100 crystals).**
- ✓ **Reliable first principles calculations are very expensive.**

  (CC5 class : 1 day/100 cores for 1 simple crystal.)
- ✓ **Little knowledge to predict LTC deductively.**

  (Chemistry-structure-LTC relationships are not clear.)

Materials search has been made through "modification" of known compounds showing high/low LTC.

Thermoelectric materials

Essential for utilizing otherwise waste heat.

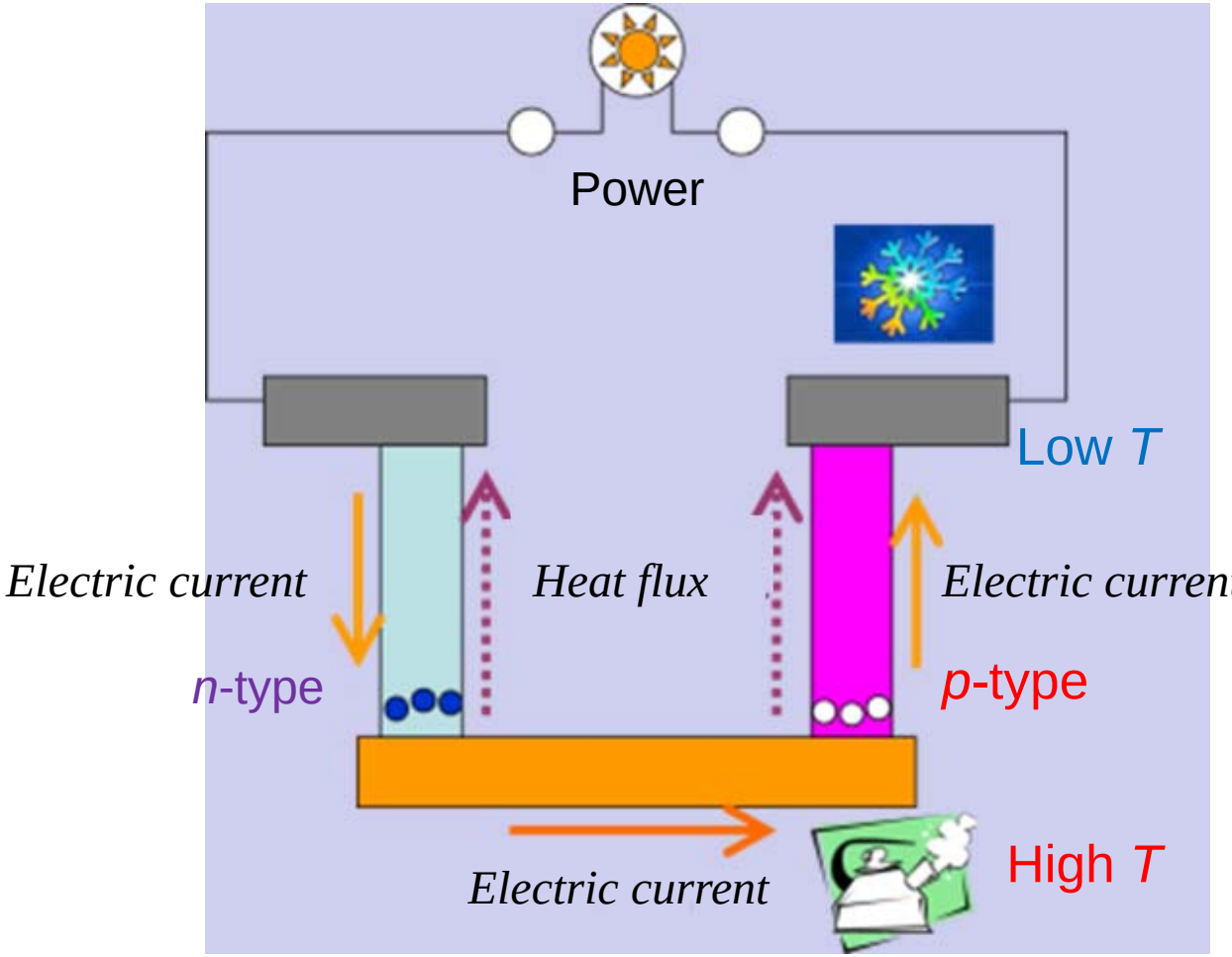


Figure of Merit

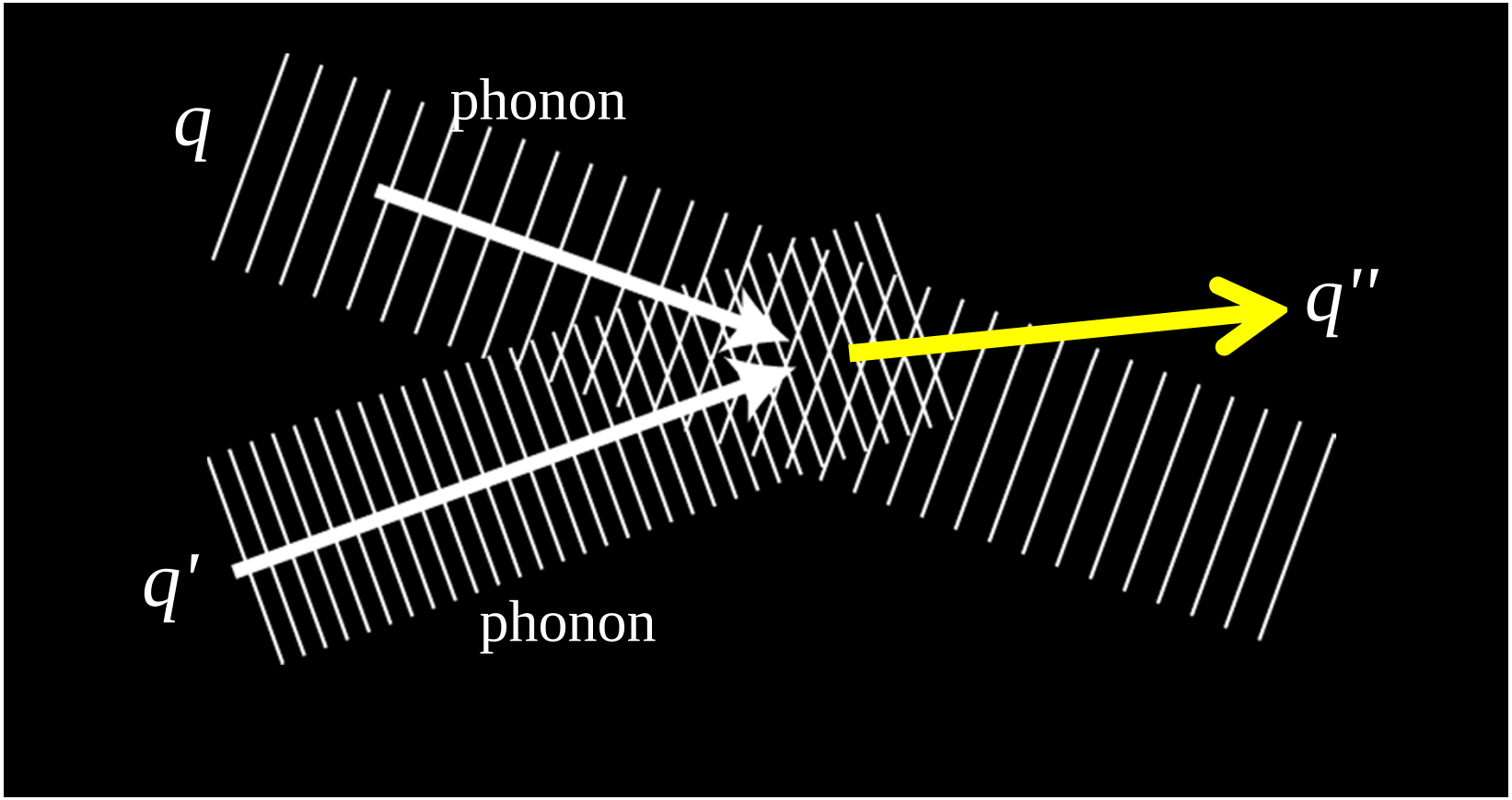
$$ZT = \frac{S^2\sigma}{\kappa} T$$

S : Seebeck coefficient
σ : electrical conductivity
κ : thermal conductivity

Physical origin of thermal resistivity

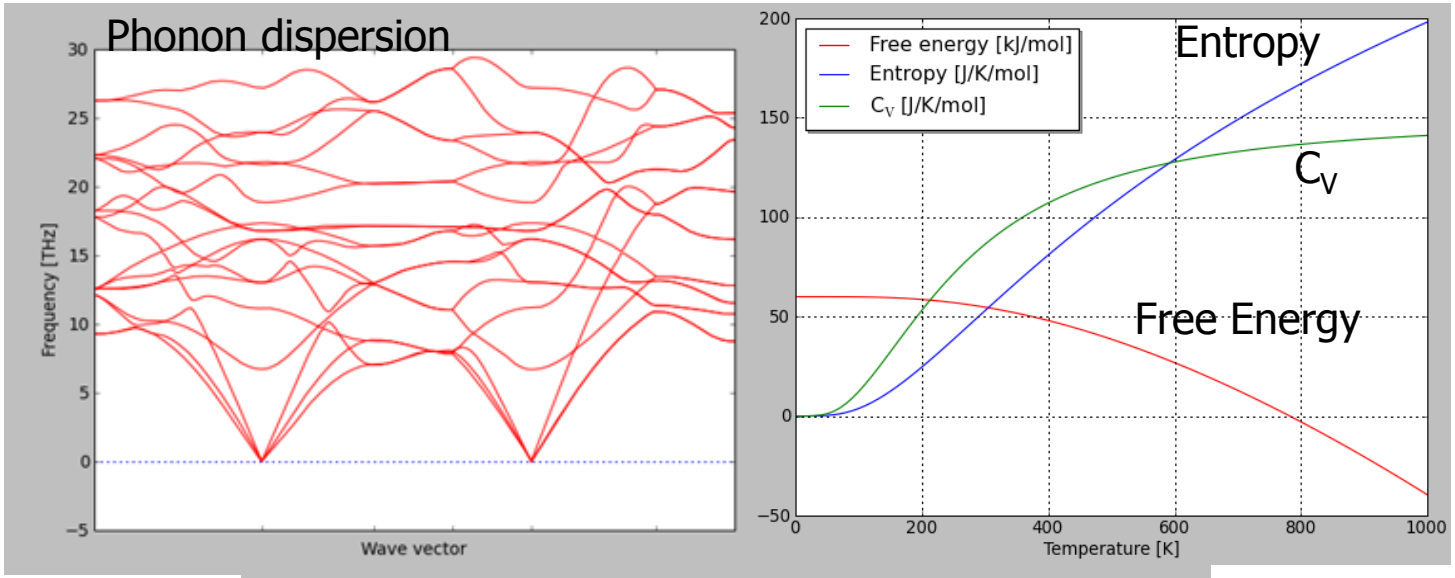
Phonon-phonon scattering (phonon anharmonicity)

Harmonic phonons do not interact.

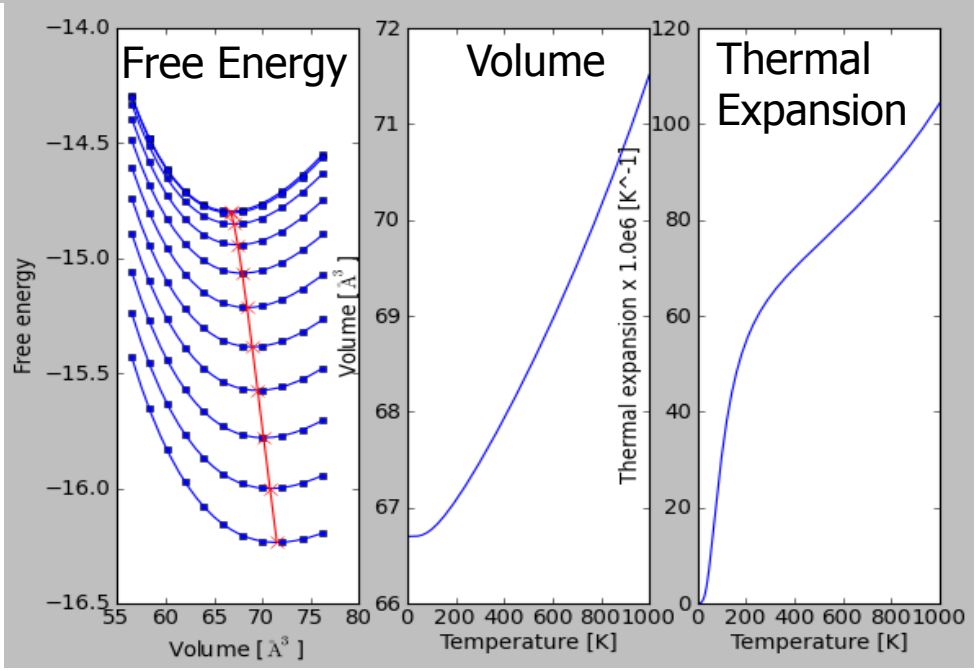


phonopy: open source for ab-initio phonon calcs

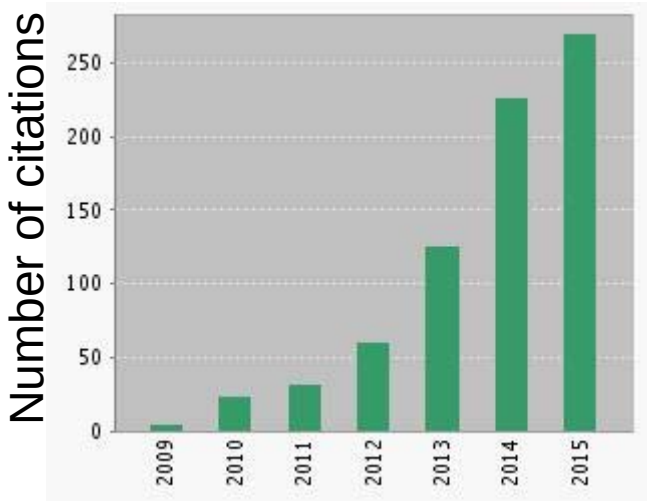
http://phonopy.sf.net/



Atsushi TOGO
Kyoto Univ.

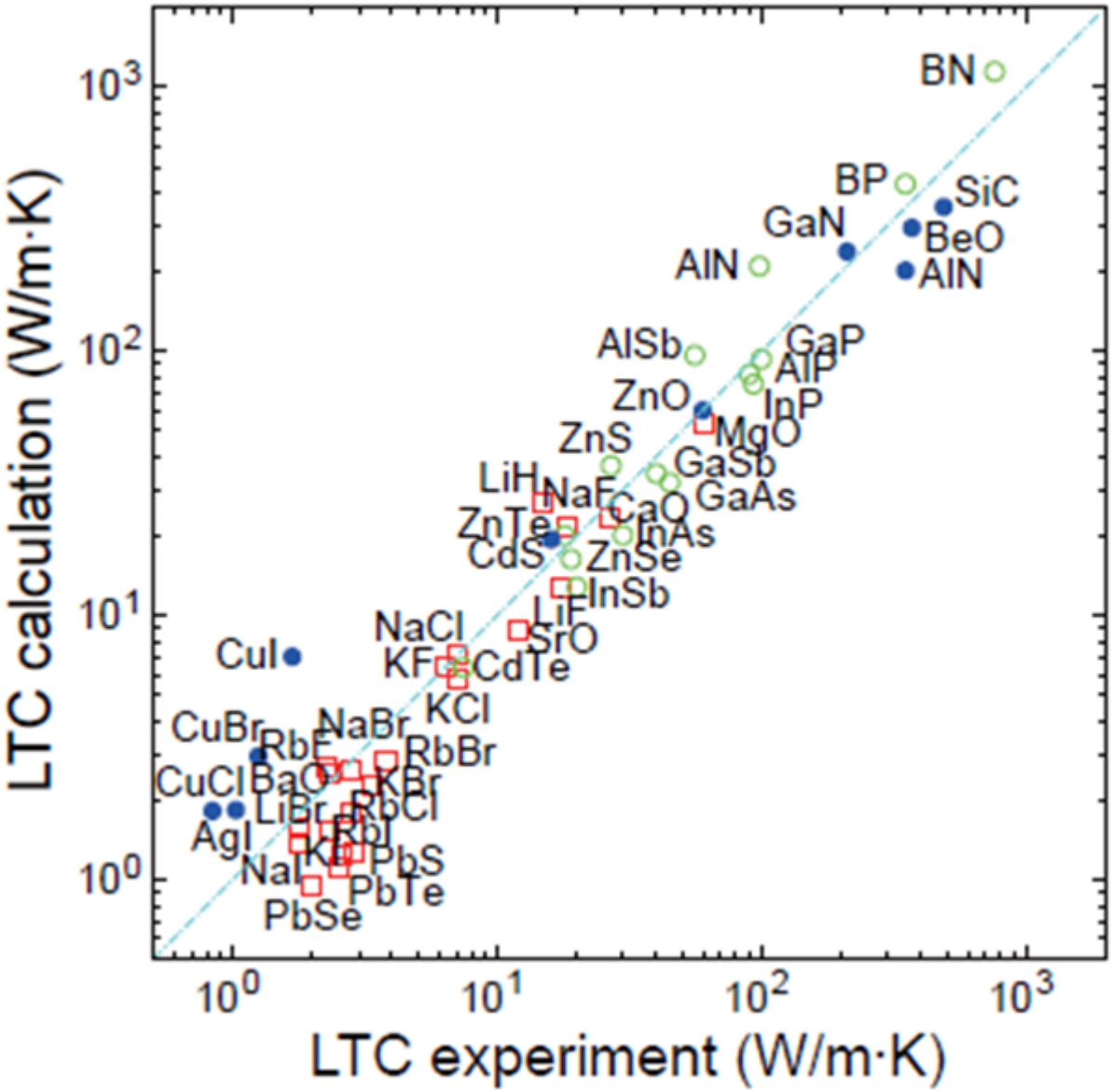


Web of Science 2015 Oct



Lattice Thermal Conductivity (LTC)@300K

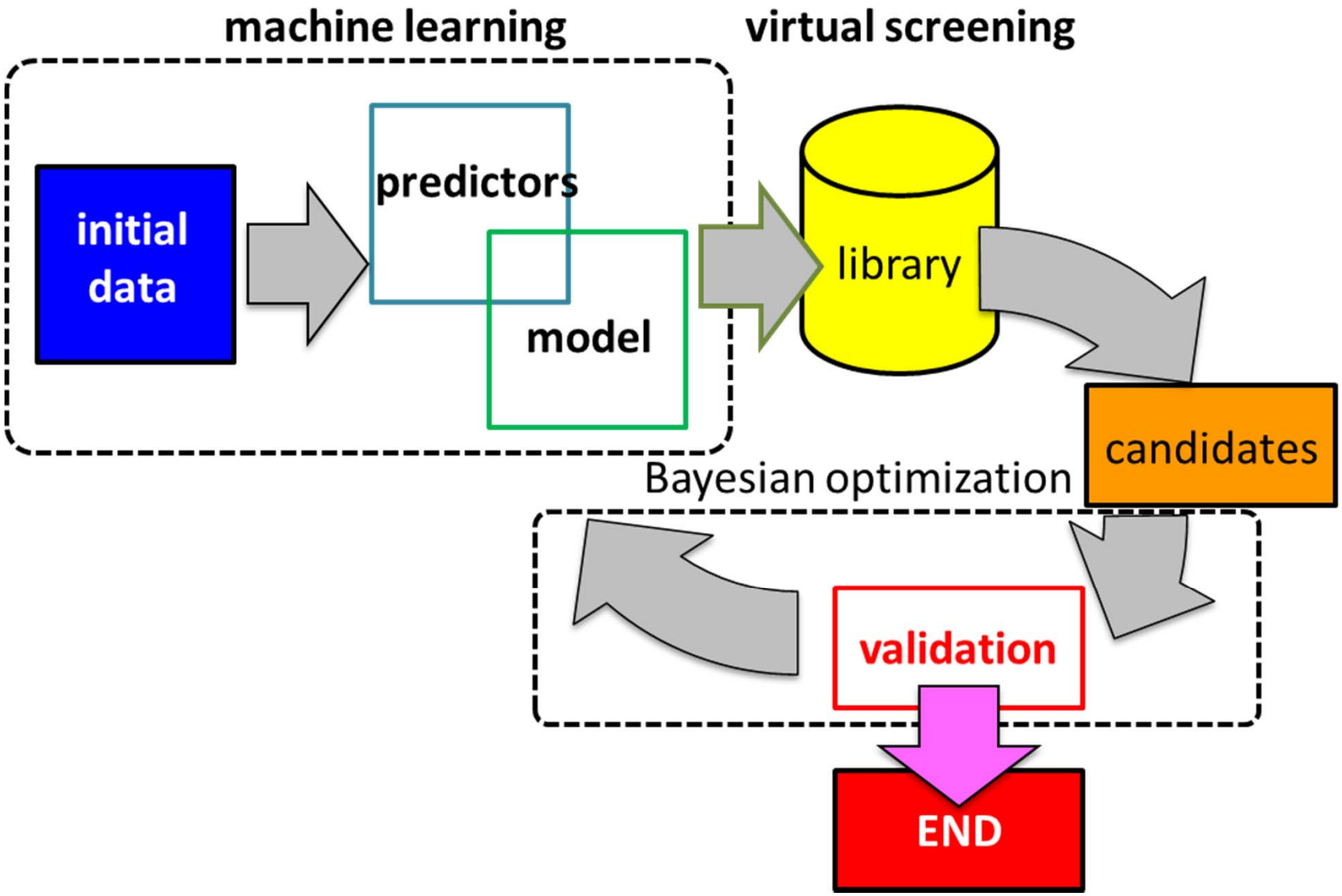
First principles calculation vs. Experimental data



CC5 class : 1 day by 100 cores for 1 simple crystal

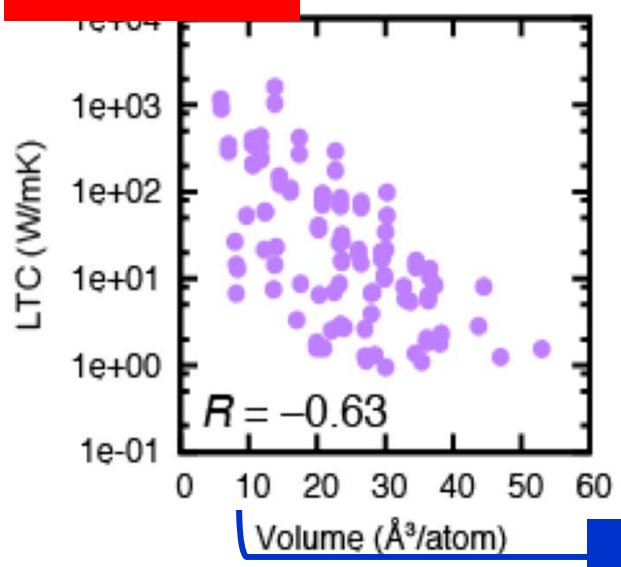
Reliable calculations whose accuracy are comparable to experiments!

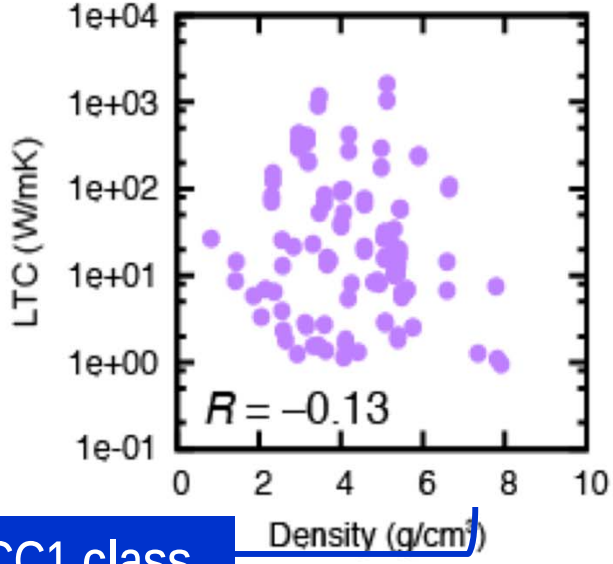
Virtual Screening

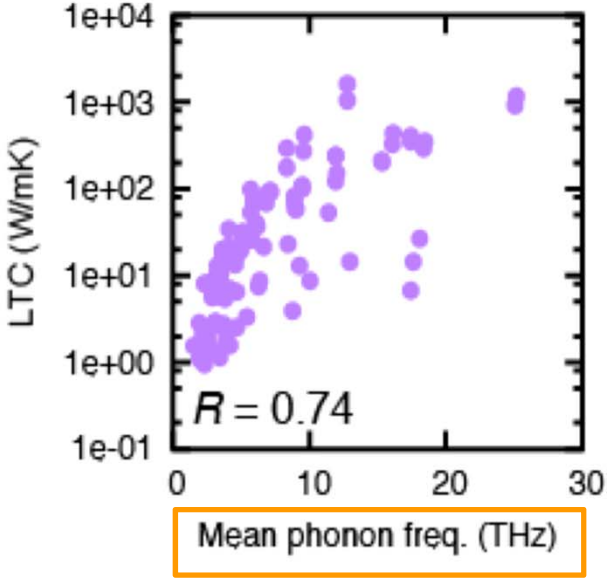


Correlations between LTC and predictors

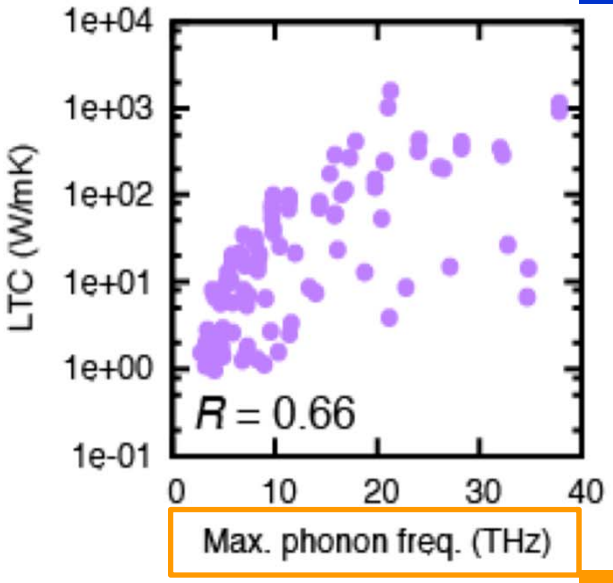
CC5 class

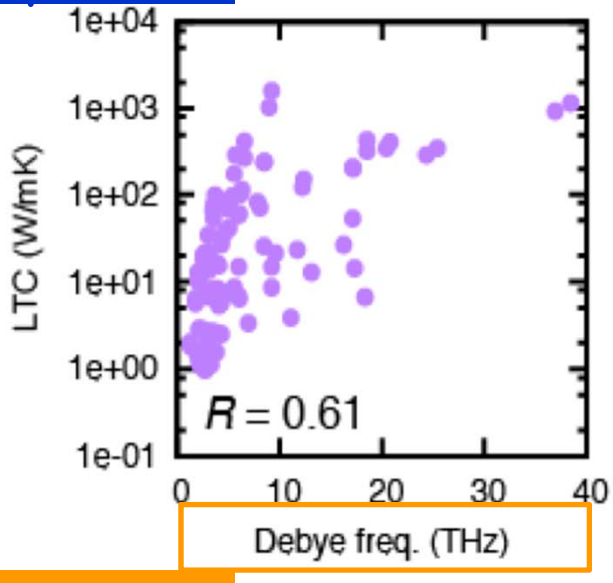


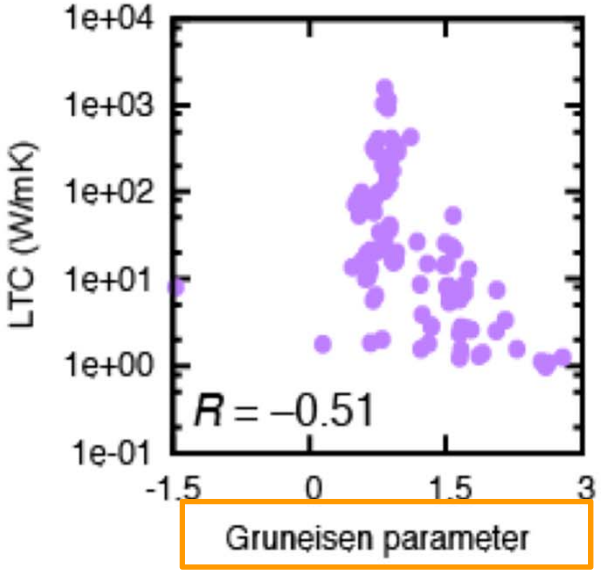




CC1 class



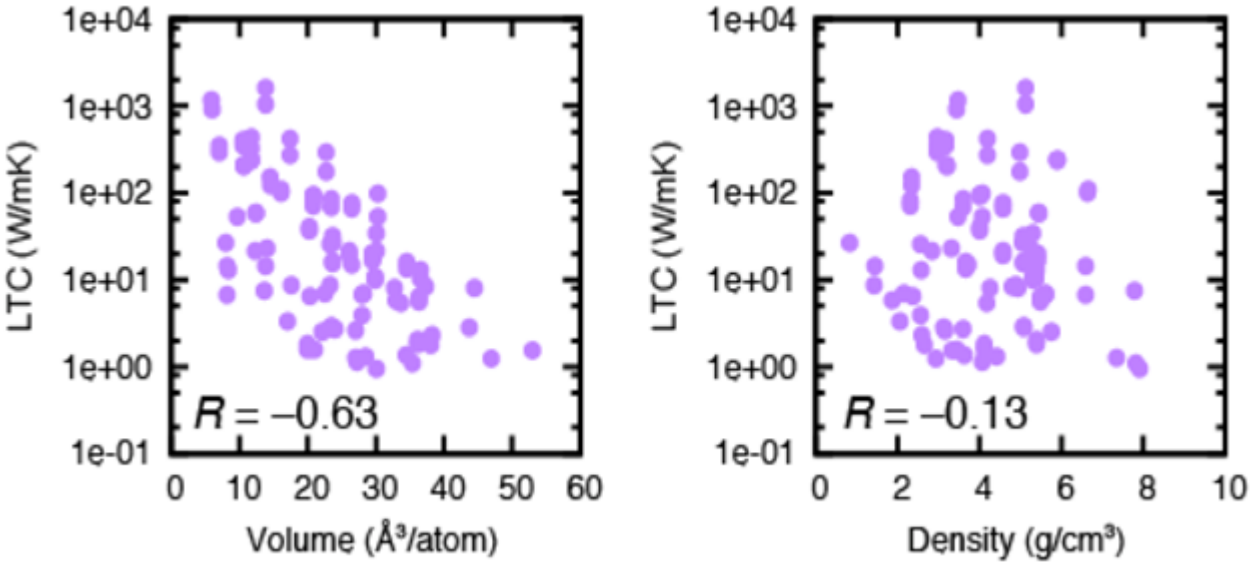




CC3 class

Candidates of descriptors for LTC

Model 1: Volume V and Density ρ



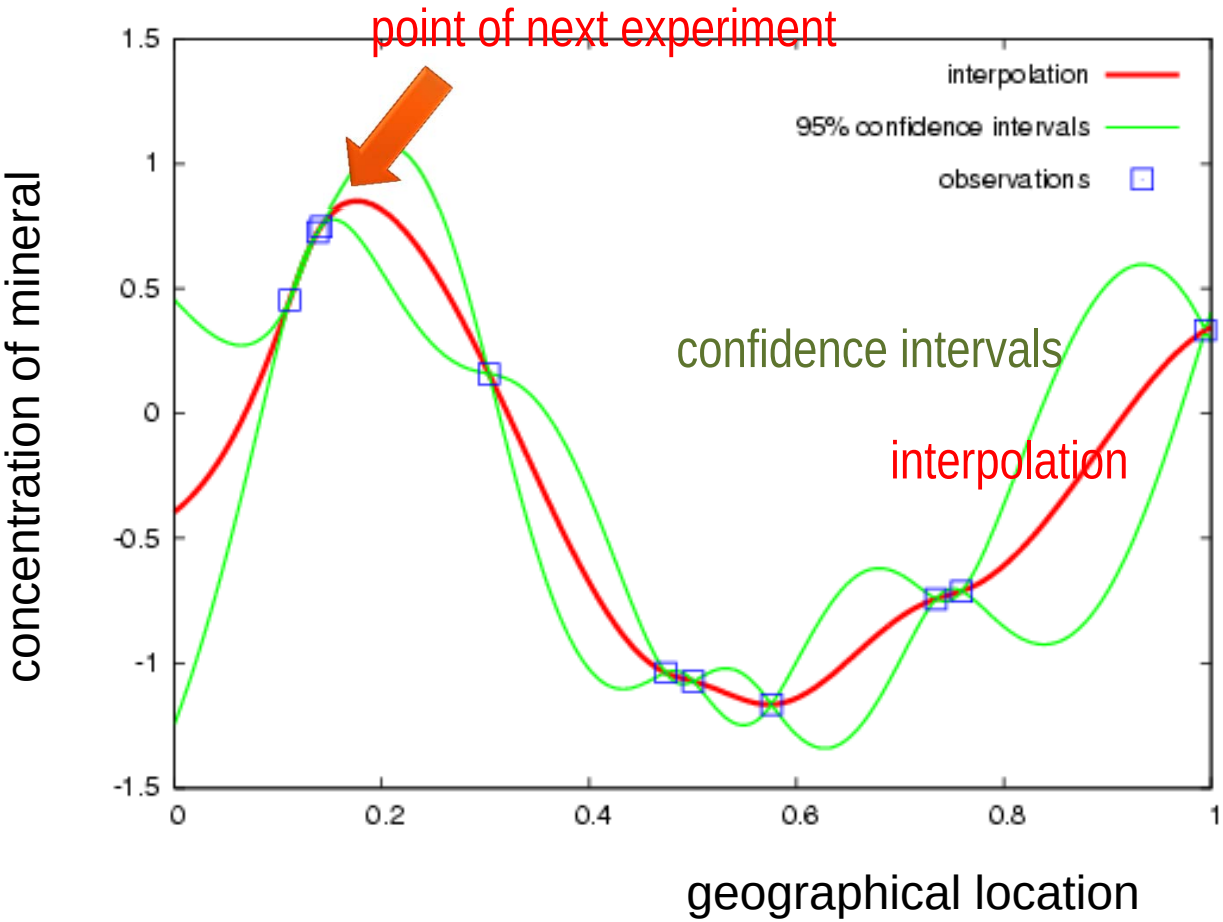
Model 2: Model 1 + primitive elemental descriptors

	H	Li	Be	B	C	N	O	F	…
LiH	1	1	0	0	0	0	0	0	
LiF	0	1	0	0	0	0	0	1	
BeO	0	0	1	0	0	0	1	0	
BN	0	0	0	1	0	1	0	0	
…									

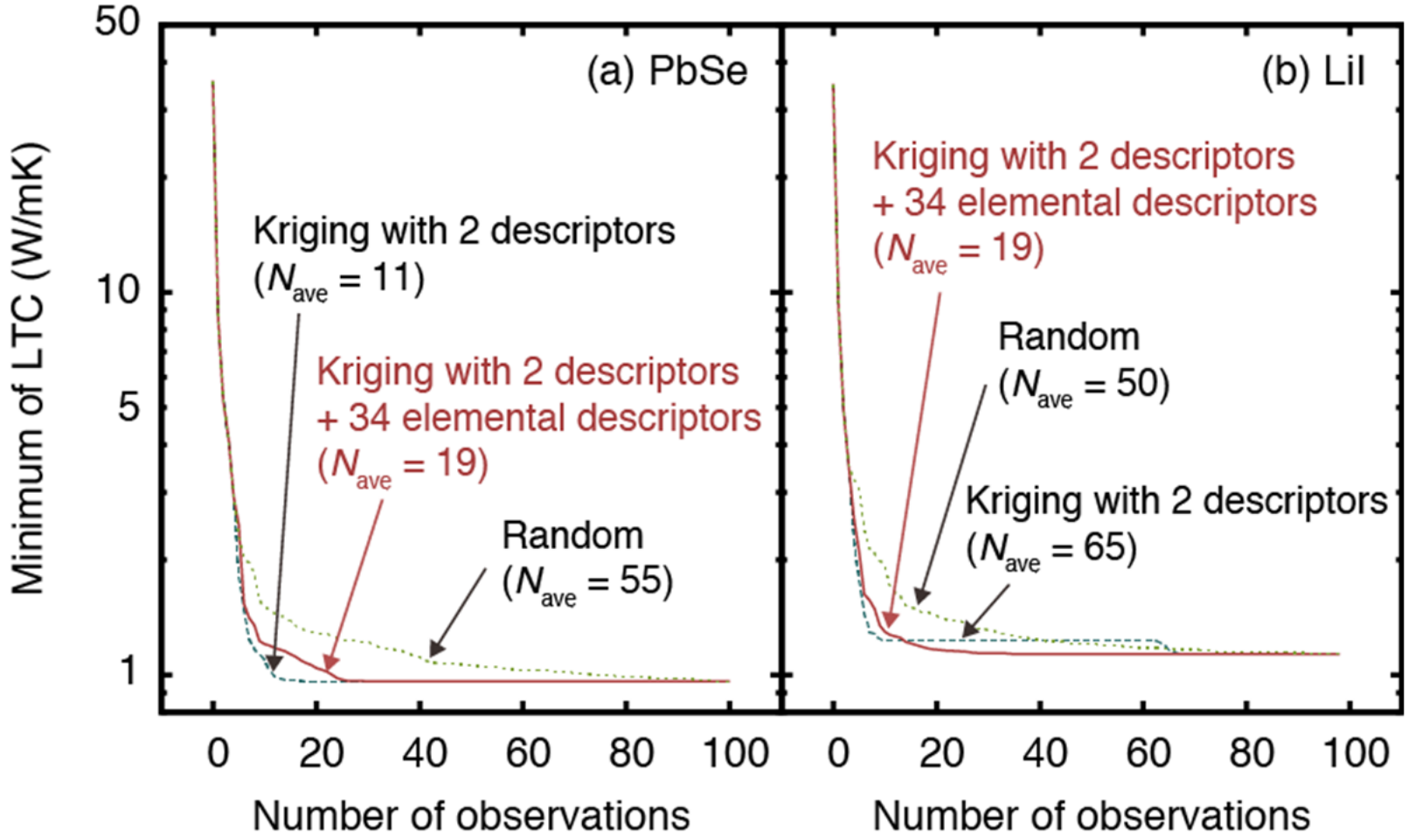
Toward discovery of new low LTC materials

Kriging

Interpolation method for which the interpolated values are modeled by a Gaussian process governed by prior covariances, originally from geo-statistics.



Kriging using 101 ab-initio LTC data



First prediction was made using N_0 (=5) LTC data and GPR.

Use of 34 elemental descriptors on top of V and ρ is found to be effective to overcome an outlier problem.

Ranking of LTC for 54,779 compounds in ICSD library

"Virtual screening"

Z-score

$$Z(\boldsymbol{x}^*) = [f(\boldsymbol{x}^*) - f_{\text{best}}]/\sqrt{v(\boldsymbol{x}^*)}$$

$$f = -\log \kappa_L$$

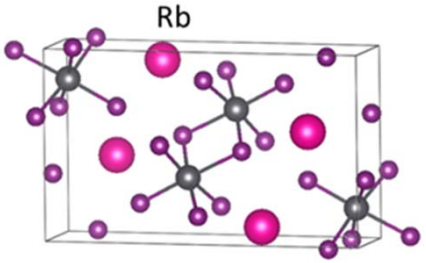
The higher Z-score, the lower predicted LTC !

	Materials Id	Formula	Z-score
1	mp-23517	RbPbI3	1.90
2	mp-571465	PbIBr	1.76
3	mp-28564	Rb4PbBr6	1.56
4	mp-23053	PbICl	1.56
5	mp-22997	PbBrCl	1.56
6	mp-23043	RbPb2Br5	1.45
7	mp-22883	PbI2	1.44
8	mp-567503	PbI2	1.43
9	mp-540839	CsPbI3	1.40
10	mp-569595	PbI2	1.34
11	mp-567246	PbI2	1.34
12	mp-23436	Cs4PbBr6	1.34
13	mp-505148	ThPbI6	1.34
14	mp-22893	PbI2	1.34
15	mp-567199	PbI2	1.32
16	mp-580202	PbI2	1.32
17	mp-600089	CsPbBr3	1.31
18	mp-640058	PbI2	1.31
19	mp-574189	PbI2	1.30
20	mp-567542	PbI2	1.30
21	mp-540789	PbI2	1.30
22	mp-567178	PbI2	1.29
23	mp-567629	CsPbBr3	1.29
24	mp-672671	PbI2	1.28
25	mp-561320	PbS	1.27
26	mp-680205	PbI2	1.25
27	mp-29883	Rb3PbCl5	1.25
28	mp-567681	CsPbBr3	1.24
29	mp-608081	Rb3Pb4Au	1.18
30	mp-27552	TlPbI3	1.17
31	mp-674972	Rb6Pb5Cl16	1.15
32	mp-31508	KPb2Br5	1.15
33	mp-771814	CaPbI6	1.14
34	mp-23037	CsPbCl3	1.14
35	mp-675524	CsPbCl3	1.14
36	mp-675022	CsPbCl3	1.14
37	mp-23475	Rb2PbCl6	1.13
38	mp-29212	Tl4PbI6	1.12
39	mp-581775	Cs5(KPb6)3	1.09
40	mp-23380	Tl3PbI5	1.08
41	mp-630851	Cs3NaPb4	1.07
42	mp-574070	Cs4Pb9	1.05
43	mp-674339	Tl6PbI10	1.04
44	mp-23425	Cs2PbCl6	1.03
45	mp-771691	CaPbI6	1.01
46	mp-756313	CaPbI4	0.99
47	mp-673703	Rb3Bi7Pb3(IO)10	0.99
48	mp-680463	Rb4Pb9	0.97
49	mp-755943	CaPbI4	0.96
50	mp-622294	Hg2Pb(SBr)2	0.96
51	mp-21525	RbPb	0.96
52	mp-571638	Rb2Cu(BrCl)2	0.90
53	mp-756136	CaPbI4	0.90
54	mp-31317	La3PbI3	0.88
55	mp-756451	CaPbI4	0.88
56	mp-753670	CaPbI4	0.87
57	mp-569879	Cs2Ta6PbCl18	0.86
58	mp-570753	Tl3PbBr5	0.86
59	mp-27451	Tl3PbBr5	0.85
60	mp-569238	Cs3LiI4	0.85
61	mp-755977	CaPbI4	0.84
62	mp-680159	CsTa6PbCl18	0.84
63	mp-557719	PbS	0.82
64	mp-755056	CaPbI4	0.82
65	mp-771827	Ca3Pb3I14	0.81
66	mp-754540	CaPbI4	0.81
67	mp-622106	Ba5Pb3	0.81
68	mp-771877	Ca3Pb3I14	0.80
69	mp-621612	AgPb2Br5	0.80
70	mp-21246	Ba2Pb	0.78
71	mp-37163	Ag2HgI4	0.78
72	mp-569465	Cs2Nb6PbCl18	0.77
73	mp-554116	BaPb2BrF5	0.76
74	mp-674993	KPb2Cl5	0.73
75	mp-607267	KPb2Cl5	0.72
76	mp-3	Cs	0.71
77	mp-27662	CsI2Br	0.71
78	mp-20136	BaPb	0.70
79	mp-573579	Cs	0.70
80	mp-672241	Cs	0.70
81	mp-11832	Cs	0.70
82	mp-1	Cs	0.69
83	mp-639727	Cs	0.69
84	mp-569225	Mo6PbI14	0.67
85	mp-541112	ZrI4	0.66
86	mp-28077	PbBr2	0.66
87	mp-569866	Cs6K7	0.64
88	mp-31288	La5Pb3I	0.64
89	mp-554245	BaPb2IF5	0.64
90	mp-613652	PbClF	0.63
91	mp-5811	CsPbF3	0.61
92	mp-570355	PbCl4	0.61
93	mp-20282	CsPbF3	0.60
94	mp-579536	CsCu2ICl2	0.60
95	mp-22969	PbIF	0.59
96	mp-559470	AgPbBrO	0.58
97	mp-674359	Tl3PbCl5	0.58
98	mp-30519	Tl3PbCl5	0.58
99	mp-23066	Pb5(SI3)2	0.56

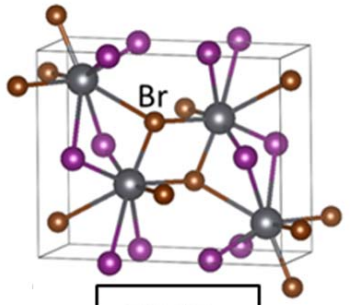
Top 10 lowest LTC compounds among 54,779

Virtual screening of 54,779 compounds in MPD library

ranking	Z-score	compound	
		formula	space group
1	1.90	$PbRbI_3$	*Pnma*
2	1.76	PbIBr	*Pnma*
3	1.56	$PbRb_4Br_6$	*R*-3*c*
4	1.56	PbICl	*Pnma*
5	1.56	PbClBr	*Pnma*
7	1.44	PbI_2	*R*-3*m*
8	1.43	PbI_2	*P*63*mc*

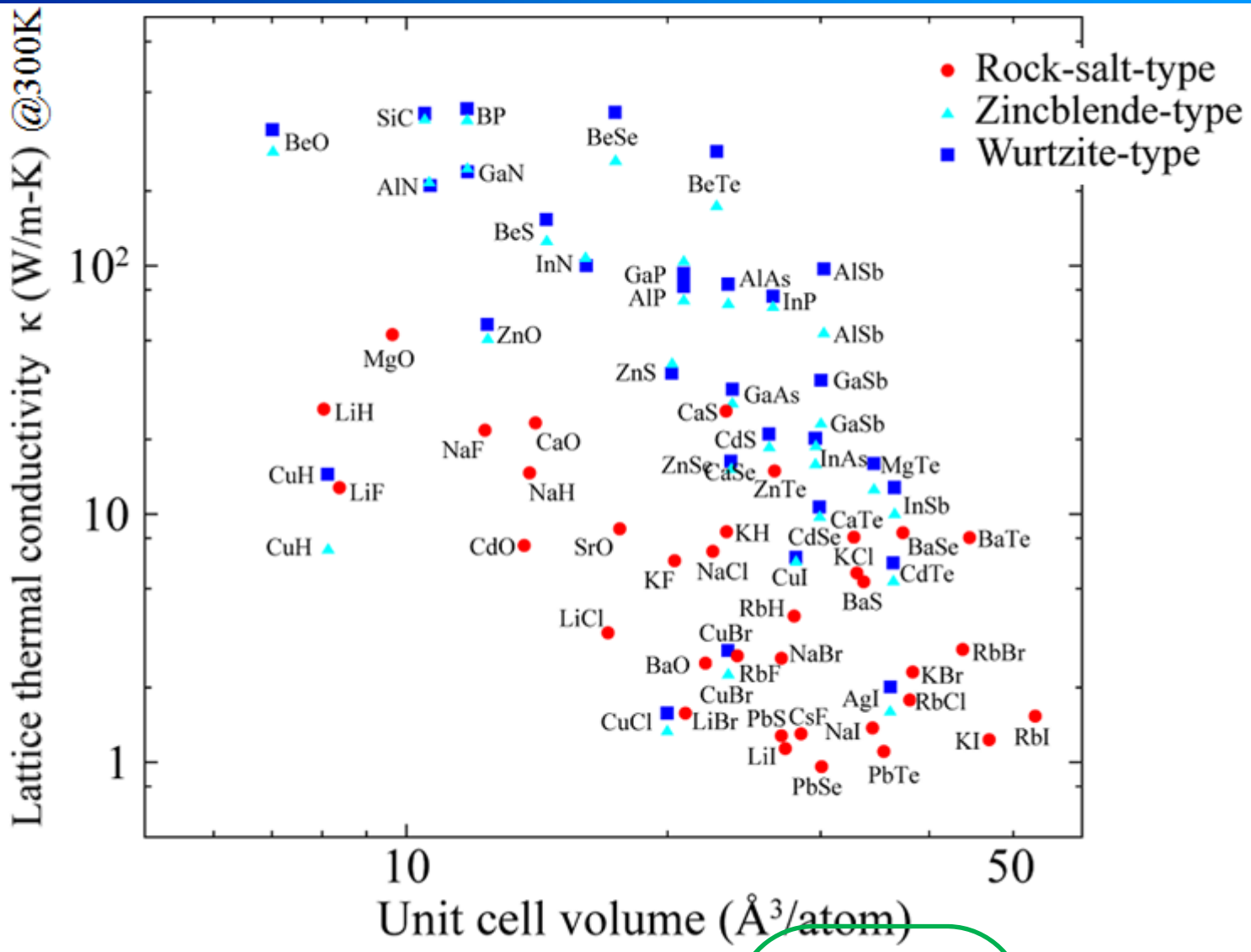


$RbPbI_3$

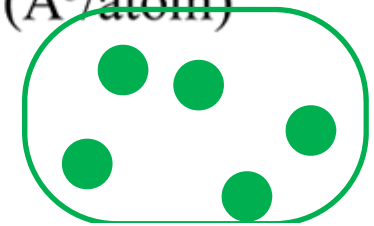


PbIBr

Lattice Thermal Conductivity (LTC)



Newly discovered !!



Thermoelectric materials

Essential for utilizing otherwise waste heat.

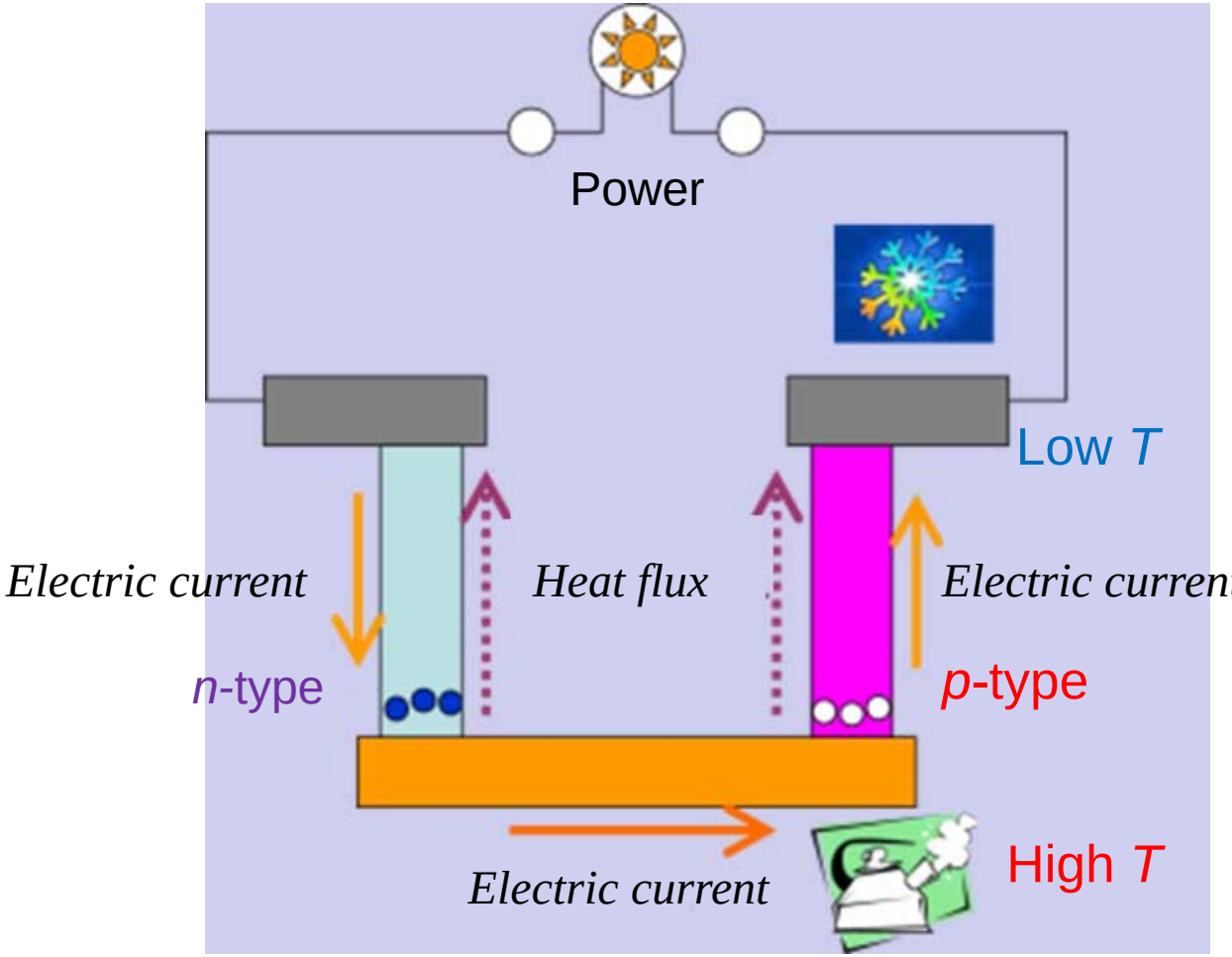


Figure of Merit

$$ZT = \frac{S^2\sigma}{\kappa} T$$

S : Seebeck coefficient
σ : electrical conductivity
κ : thermal conductivity

Top 10 lowest LTC compounds among 54,779

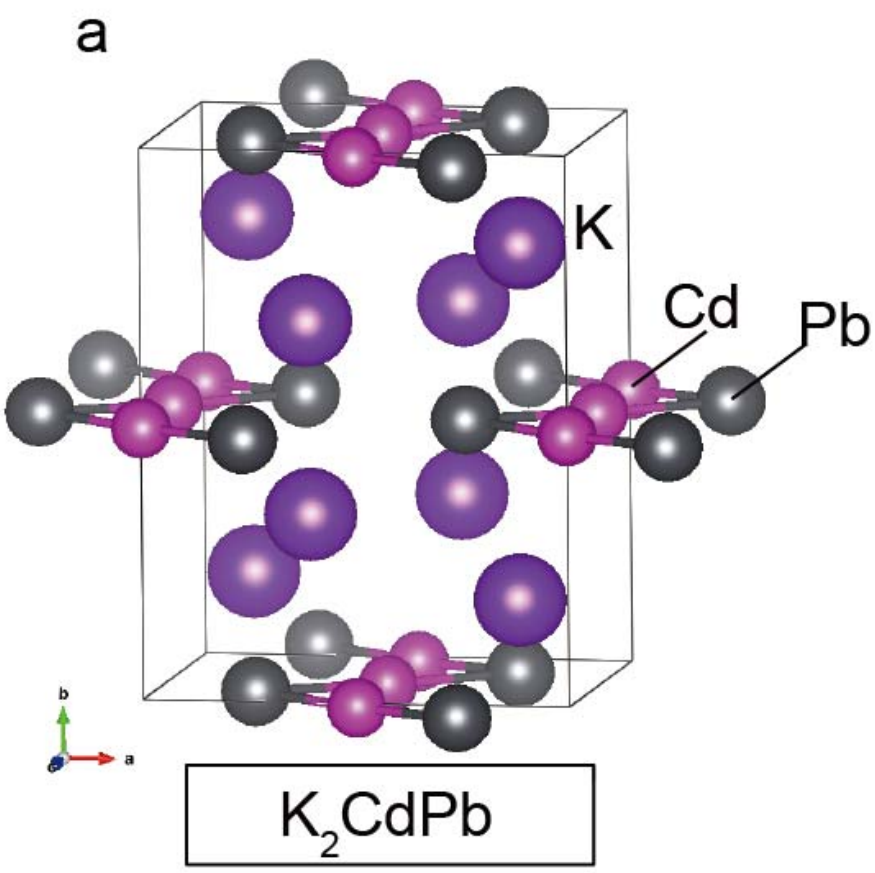
Virtual screening of 54,779 compounds in MPD library

ranking	Z-score	compound		*ab initio* LTC (W/m·K)	Band Gap (eV)
		formula	space group		
1	1.90	$PbRbI_3$	*Pnma*	0.10	2.46
2	1.76	PbIBr	*Pnma*	0.13	2.56
3	1.56	$PbRb_4Br_6$	*R*-3*c*	0.08	3.90
4	1.56	PbICl	*Pnma*	0.18	2.72
5	1.56	PbClBr	*Pnma*	0.09	3.44
7	1.44	PbI_2	*R*-3*m*	0.29	2.42
8	1.43	PbI_2	*P*63*mc*	0.29	2.45
121	0.39	K_2CdPb	*Ama*2	0.45	0.18
144	0.29	$Cs_2[PdCl_4]I_2$	*I*4/*mmm*	0.31	0.88

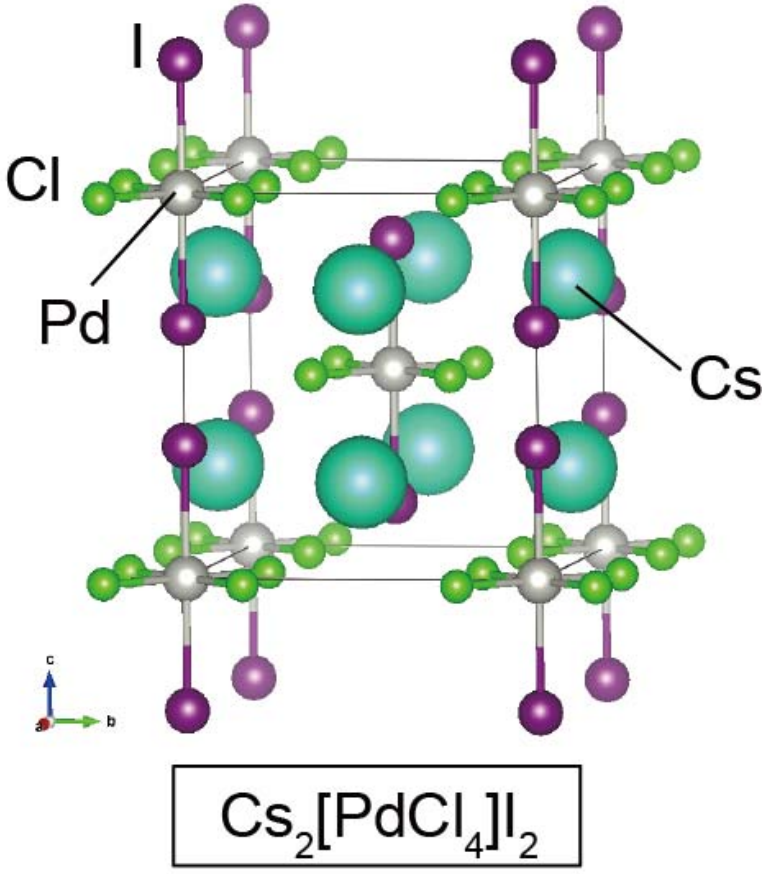
Newly discovered candidates for thermoelectrics

with low LTC of <0.5 W/mK (@300 K)

and narrow band gap of <1 eV.



K_2CdPb



$Cs_2[PdCl_4]I_2$

A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, and IT PRL 115, 205901 (2015)

Fin