

From meta-data to meta-properties with only a bit of meta-physics

Stefano Curtarolo

Materials Science - Duke University – USA

Center for Materials Genomics – USA

C. Toher, C. Oses, R. Friedrich, D. Hicks, F. Rose



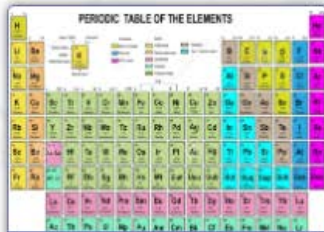
U.S. DEPARTMENT OF
ENERGY

UC San Diego

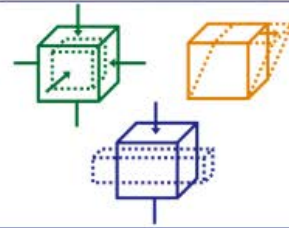
AFLOW.org: data/ metadata and applications

aflow.org

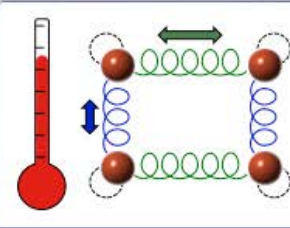
Apps and Docs



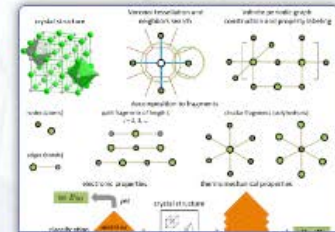
MendeLIB search



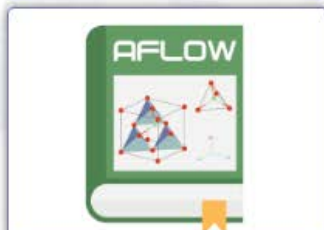
Elastic properties



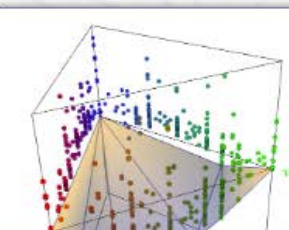
Thermal properties



AFLOW-ML



Crystal prototypes



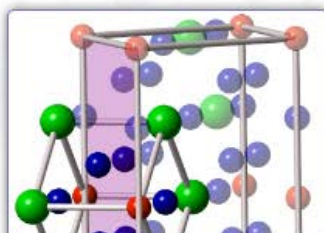
AFLOW-CHULL



AFLOW-online



Binary alloy library



Superalloys search



REST-API Docs



AFLOW π



PAOFLOW

AFLOW.org: web portal

aflow.org

~3 million entries (AUIDs)

The screenshot shows the AFLOW.org web portal. At the top left is the AFLOW logo with the tagline "Automatic - FLOW for Materials Discovery". A navigation bar contains links for HOME, CONSORTIUM, PUBLICATIONS, and FORUM. The main content area features a welcome message and a grid of property categories. A search bar is at the bottom.

Welcome to AFLOW, a globally available database of **2,744,589** material compounds with over **488,536,842** calculated properties, and growing.

323,383 band structures	80,162 Bader charges	5,791 elastic properties	5,779 thermal properties
1,620 binary systems	351,413 binary entries	30,031 ternary systems	1,882,672 ternary entries
	150,569 quaternary systems	437,755 quaternary entries	

AFLOW also offers online applications for property predictions using [machine learning](#), [crystal structure databases](#), and the generation of [convex hulls](#).

SD Number, [Aflowlib Unique Identifier](#), or advanced search string (e.g. Mg & Sn & Cu) database.

Search

~535 million properties

breakdown of groups

AFLOW.org: search page

The screenshot shows the AFLOW.org search interface. At the top, the AFLOW logo is displayed with the tagline "Automatic - FLOW for Materials Discovery". A navigation bar includes links for HOME, CONSORTIUM, PUBLICATIONS, FORUM, SRC, and SEARCH. Below this is a search bar with the text "Search Aflow" and a "Search" button indicating "(60375 Compounds)".

The main search area features a periodic table of elements. A callout box for element 'X' shows its properties: Atomic #, element name, [electrons], [density], [T₀], [lattice], [crystal], and [Debye]. To the right of the periodic table are logical operators: and, not, or, xor, (, and).

Below the periodic table, there are filters for "Results Per Page" (set to 40), "Total # of Results" (1000), and "# of Species" (input field). A grid of category buttons includes: All Metals, Alkali Metals, Alkaline Earths, Transition Metals, Lanthanides, Other Metals, Nonmetals, Group 3A, Group 4A, Group 5A, Chalcogens, and Halogens.

A "Search Displayable Column Properties" section contains a search box and a grid of property categories: Chemistry, Crystal, Electronics, Thermodynamics, Magnetics, Scintillation, Mechanical, and Calculation.

At the bottom, a welcome message states: "Welcome to AFLOW, a globally available database of 2,532,342 material compounds with over 341,866,170 calculated properties, and growing." Below this are logos for Duke University, U.S. DEPARTMENT OF ENERGY, Homeland Security, CRAY THE SUPERCOMPUTER COMPANY, and UNT.

Four orange callout boxes with arrows point to specific features: "Element group selection" points to the category buttons; "Element selection" points to the periodic table; "Property Filters" points to the "Search Displayable Column Properties" section; and "Number of species" points to the "# of Species" input field.

Element group selection

Element selection

Property Filters

Number of species

aflow.org/advanced.php

Web portal for HUMANS

[First]

searches

[4]

ENTRY

$\text{As}_2\text{Cu}_1\text{O}_8\text{Zn}_2$ [75f66067131a579e]

$\text{As}_4\text{Cu}_4\text{H}_4\text{O}_{20}\text{Zn}_4$ [521293f6fc9719f4]

$\text{Au}_1\text{Cu}_1\text{Zn}_2$ [a5c27d501796fcf2]

human readable output

Au1Cu1Zn2

LICENSE: The data included within the aflow.org repository is free for scientific, academic and

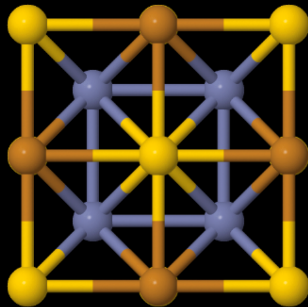
PERMANENT URL: http://afLOW.org/material.php?id=Au1Cu1Zn2_ICSD_611778

AFLOW.ORG web entry generator V3.1.224 [built=2019-05-22]

Space Group: Fm-3m (#225)

AFLOW.org consortium (AFLOW v3.1.224)
entry=Au1Cu1Zn2_ICSD_611778

Spacegroup = Fm-3m (#225)
a=6.1108Å, b=6.1108Å, c=6.1108Å
α=90°, β=90°, γ=90°



Supercell:

2x2x2

X

Be

Visualizati

Ball & St

Rotation

axis: a

Crystallogr

h: 2

Show

Symmetry

48 operator

Backgro

Identity

rotation C2

rotation C2

rotation C2

RESET

Bader Isos

Cutoff = 0.20

Au

Cu

Zn

All

ICSD entry [info]:

611778

PP - Species Version [species_pp_version]:

Cu_pv:PAW_PBE:06Sep2000

PP - Species ZVAL [species_pp_ZVAL]:

17.0000

Total CPU hours [calculation_time*calculation_cores]/3600:

5.7487 hours

Number of cores [calculation_cores]:

16 (MPI)

AFLOW Version [afLOW_version]:

afLOW30847

AFLOW Data Source [data_source]:

afLOWlib

Pseudopotentials type [pp_type]:

PAW_PBE

PP - Species Version [species_pp_version]:

Zn:PAW_PBE:06Sep2000

PP - Species ZVAL [species_pp_ZVAL]:

12.0000

Total Wall-time [calculation_time*3600]:

0.3593 hours

Processor - CPU Model [node_CPU_Model]:

AMD_Opteron(tm)_Processor_6378_

AFLOW Catalog [catalog]:

ICSD

AFLOW Isos [isos]:

bader,bands,lock,magnetic,thermodynamics,afLOWlib

AFLOW-LUID [uid]:

afLOW:a5c27d501796fcf2

Composition per cell [composition]:

Au1Cu1Zn2

PP - Species Version [species_pp_version]:

Au:PAW_PBE:06Sep2000

PP - Species ZVAL [species_pp_ZVAL]:

11.0000

LDAAU [T,(L),(U),(F)] [data_TLUF]:

2,2,2,2;4,4,7,5;0,0,0 (type:(angular-1):(eV):(eV))

Memory Used [calculation_memory]:

543.7300 MB

Processor - CPU MHz [node_CPU_MHz]:

1400MHz

AFLOW Data API [data_api]:

aapi1.2

AFLOW-LIB entry version [afLOWlib_version]:

3.1.222

AFLOW-LIB entry version [afLOWlib_version]:

afLOWlib

afLOW-LIB entry version [afLOWlib_version]:

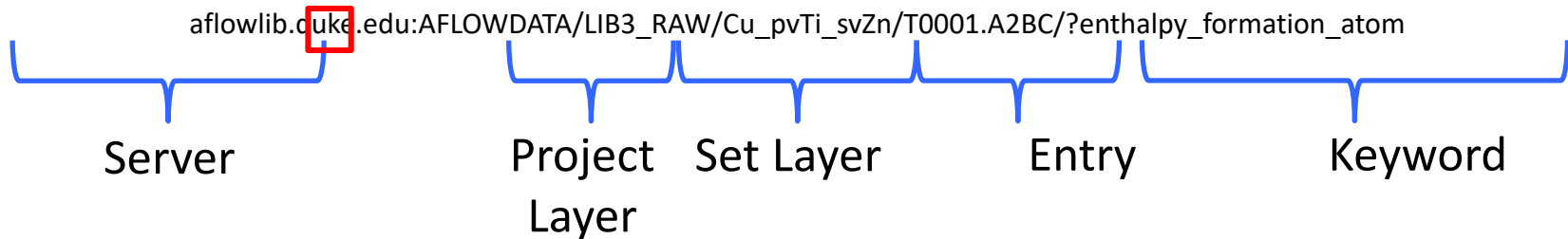
3.1.222

objects for machines: REST-API

- The AFLOW REST-API AFLOW URL (AURL) has the form:

`<server>:AFLOWDATA/<project>/<set>/<entry>/?<keyword>`

- The AURL for the formation enthalpy per atom of the ternary Heusler structure T0001.A2BC with composition Cu_2TiZn is:



- This becomes the AURL:

aflowlib.duke.edu/AFLOWDATA/LIB3_RAW/Cu_pvTi_svZn/T0001.A2BC/?enthalpy_formation_atom

- The full list of entries in the set Cu-Ti-Zn can be retrieved using:

aflowlib.duke.edu/AFLOWDATA/LIB3_RAW/Cu_pvTi_svZn/?aflowlib_entries

primitive metadata: REST-API

- Example python script to access API & download space groups (see tutorial script *restapi.py* from *aflow_api.txz*):

```
import json
```

```
from urllib.request import urlopen
```

```
SERVER = 'http://aflowlib.duke.edu/AFLOWDATA'
```

```
PROJECT = '/LIB3_RAW'
```

```
SET = '/Cu_pvTi_svZn'
```

```
ENTRIES = '/?aflowlib_entries'
```

Project: Ternary alloys

Set: CuTiZn alloy system

Keyword: List of entries in this alloy system

```
response = urlopen(SERVER+PROJECT+SET+ENTRIES).read().decode('utf-8')
```

Download list of entries with urlopen

```
entry_list = response.rstrip().split(',')
```

Loop through list of entries

```
for entry in entry_list:
```

```
    entry_sg_url = SERVER + PROJECT + SET + '/' + entry + '/?spacegroup_relax' # spacegroup relax URL
```

```
    entry_sg = urlopen(entry_sg_url).read().decode('utf-8')
```

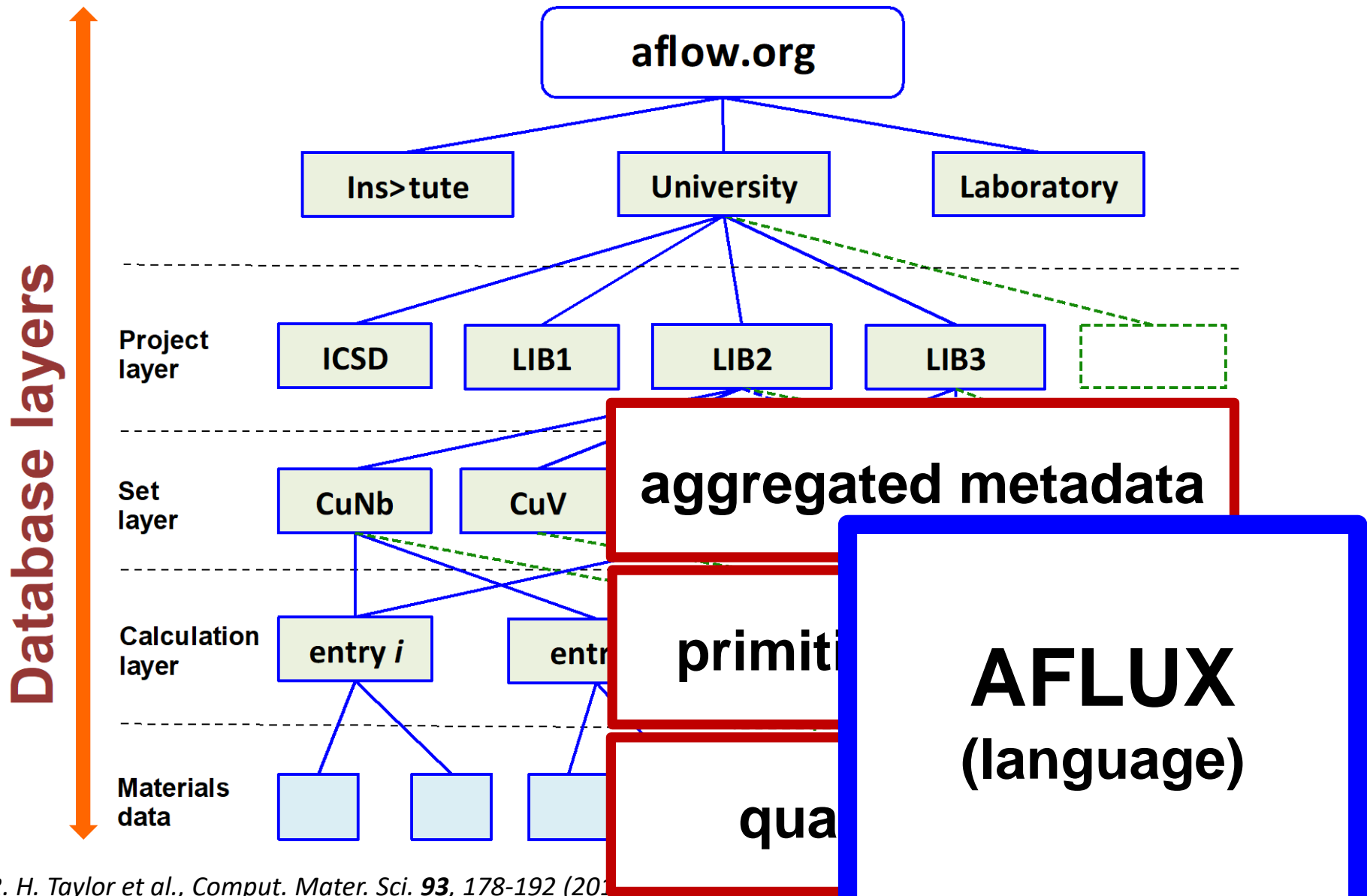
```
    print('{0}: {1}'.format(entry, entry_sg))
```

Download space group for each entry

primitive metadata for SYMMETRY: AFLOW-SYM

- Calculates symmetry descriptions for
 - *lattice* : lattice points
 - *superlattice* : *atomic sites are decorated with single species*
 - *reciprocal lattice* : *reciprocal lattice points*
 - *crystal* : lattice points + atomic sites
 - *crystal spin* : *lattice points + atomic sites + magnetic moment*
- *Provides the following operator representations:*
 - operator type
 - *Hermann-Mauguin*
 - Schoenflies
 - *transformation matrix (Cartesian)*
 - *transformation matrix (fractional)*
 - *generator matrix*
 - *so(3) coefficients ($\mathbf{L}_x, \mathbf{L}_y, \mathbf{L}_z$)*
 - *angle*
 - *axis*
 - *quaternion (vector)*
 - *quaternion (2 x 2 matrix)*
 - *quaternion (4 x 4 matrix)*
 - *su(2) coefficients (Pauli)*
 - *inversion boolean*
 - *internal translation (Cartesian)*
 - *internal translation (fractional)*
 - *atom index map*
 - *atom type map*
 - *lattice translation (Cartesian)*
 - *lattice translation (fractional)*

AFLOW.org: database organization



AFLUX: language to search for data and primitive metadata



```

url=#flowlib.duke.edu:AFLOWDATA/LIB3_RAW/AgCoMn_pv/T0002.A2BC!
auid=#flow:AgCoMn_pv/T0002.A2BC:RAW_PBE!
api=1.0!
keywords=aui,auid,aflow_api_version,code,compound,prototype,nspecies,...!
aflowlib_entry_date=20140130_20:34:00 GMT-8!
aflowlib_entry_version=30794!
aflow_version=#flow00393!
calculation_cores=1!
calculation_memory=639!
calculation_time=18347.2!
corresponding=Stefano_Sanvito_sanvito@tcd.ie!
loop=hemodynamics,bands,magnetic!
node_CPU_Cores=12!
node_CPU_MHz=2661!
node_CPU_Model=Intel(R)_Xeon(R)_CPU_X5650_@_2.67GHz!
node_RAM_GB=24!
code=vasp.4.7.38!
composition=0,1,1!
compound=Ag2Co1Mn1!
density=5.94193!
entropy=0!
entropy_atcm=0!
Egap=1!
energy=-20.4051!
energy_atcm=-5.10128!
enthalpy=-20.4051!
enthalpy_atcm=-5.10128!
enthalpy_formula=1.51248!
    
```

mandatory keywords

optional control keywords

optional materials keywords

Matchbook:

```

Matchbook ::= Unary-Not? ( Unary-Mute? Datum-string ' ( Match ' | ' ( Matchbook ' ) ) | Matchbook Bira Matchbook
    
```

referenced by:

- Matchbook
- Summons-Luifer

Match:

```

Match ::= Unary-Not? ( Unary-Loose? ( Datum-string | Datum-number ) Unary-Loose? | ' ( Match ' ) ) | Match Biral Match
    
```

referenced by:

- Match
- Matchbook

Directive:

```

Directive ::= ( Biral Unary-Mute? Datum-string ' ( ( Datum-string | Datum-number ) ( Biral ( Datum-string | Datum-number ) ) * ' ) ' ) *
    
```

Logical operator	LUX syntax
<block>	"(" and ")"
<AND>	","
<OR>	","
<NOT>	!"
<loose>	"*"
<string>	"'"
<mute>	"\$"

- Aim: Programatically expose the same functionality as the web search interface at <http://afLOW.org/advanced.php>

AFLUX: AFLOW Search-API

- AFLUX enables search functionality through query part of URI
- Supports use of several logical operators
- Operator scope can be inter-property and/or intra-property

Logical operator	AFLUX syntax	Operator scope
<block>	“(” and “)”	Intra- and inter-property
<AND>	“,”	Intra- and inter-property
<OR>	“.”	Intra- and inter-property
<NOT>	“!”	Intra-property
<loose>	“*”	Intra-property
<string>	“”	Inter-property
<mute>	“\$”	Intra-property

AFLUX: AFLOW Search-API

`http://afloplib.duke.edu/search/API/?<matchbook>,<directives>`

Search API Server

Query

- Matchbook:
 - Materials keywords with arguments
 - Example: `<server>/?species((Na:K),Cl),nspecies(2),Egap(1*,*5),energy_cell`
 - Keyword list available from schema directive: `<server>/?schema`
- Directives:
 - Formatting instructions with arguments
 - “format” takes arguments “json” and “html”
 - “catalog” takes arguments “icsd”, “lib1”, “lib2”, ...
 - “paging” controls number of entries and page displayed, sorted in ascending order of first materials keyword – order can be reversed by using a negative page number in the argument

APE: AFLOW Python Environment: search-API

- Example python script to access AFLUX search-API (see tutorial script `afluxapi.py` from `aflow_api.tgz`):

```
#!/usr/bin/env python
import json, sys, os
from urllib.request import urlopen
```

```
SERVER="http://aflowlib.duke.edu"
API="/search/API/?"
```

```
MATCHBOOK="species((Na:K),Cl),nspecies(2),Egap(2*,*5),energy_cell"
```

```
DIRECTIVES="$paging(0)"
```

```
SUMMONS=MATCHBOOK+" "+DIRECTIVES
```

```
response=json.loads(urlopen(SERVER+API+SUMMONS).read().decode("utf-8"))
```

```
for datum in response:
```

```
    bandgap=[float(x) for x in datum['Egap'].split(",")]
```

```
    energycell=[float(x) for x in datum['energy_cell'].split(",")]
```

```
    print ("{}", {}, {}".format( datum['auid'], bandgap, energycell))
```

AFLUX Server

Matchbook: materials keywords

Directive: return all entries satisfying query

JSON with query results is downloaded

What is aggregate data: ontological problem

AGGREGATED METADATA

**set of calculations giving
a particular property**

e.g. elasticity, phonons, entropy, synthesizability, etc..

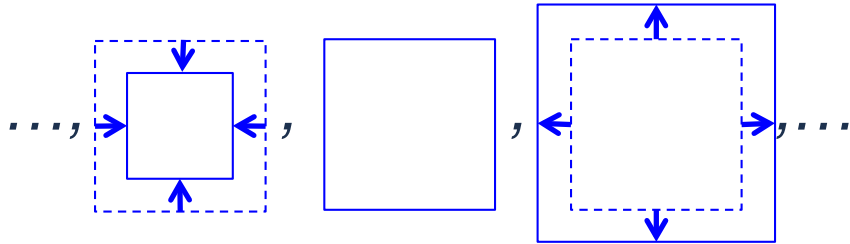
aggregated metadata requires workflow of workflows

- one workflow (e.g. band-structure) => one AUID
- workflow of workflows (phase stability) => AUID of AUIDs

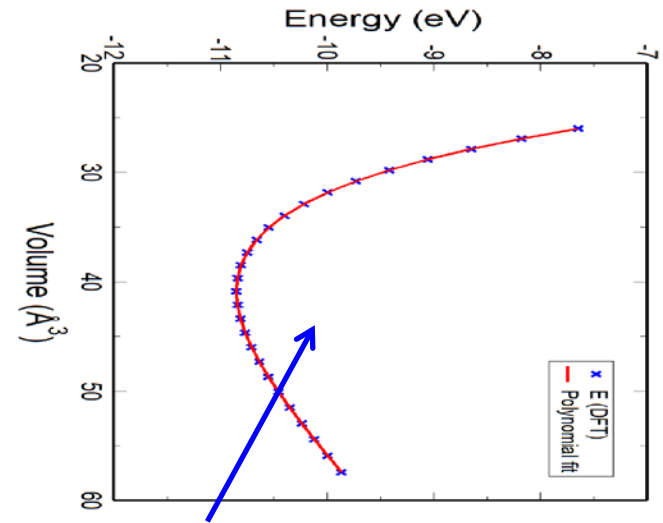
STATIC vs DYNAMIC AGGREGATED METADATA

STATIC METADATA: Debye-Grüneisen AFLOW-AGL

Different volume cells



E(V) data from DFT
(VASP, QE, etc.)



Debye temperature: $\theta_D(V)$

Grüneisen
parameter: γ

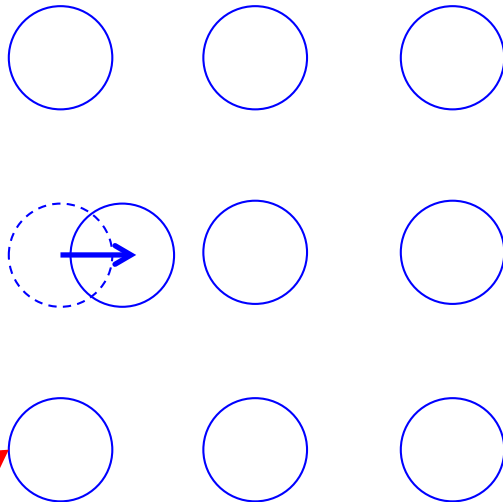
Heat capacity:
 C_V

Lattice thermal conductivity, κ_L

Fit with polynomial or eqn of
state: bulk modulus, $B(V)$

STATIC METADATA: Phonon calculations AFLOW-APL

AFLOW Phonon Library



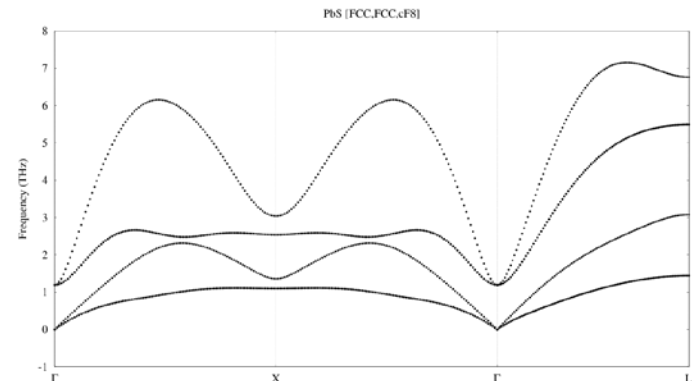
$\Delta \vec{F}$

Perturb atom, obtain change
in force on others

Interatomic force constants

$$\Phi_{ij}^{\alpha\beta} = \frac{\partial^2 V}{\partial r_i^\alpha \partial r_j^\beta}$$

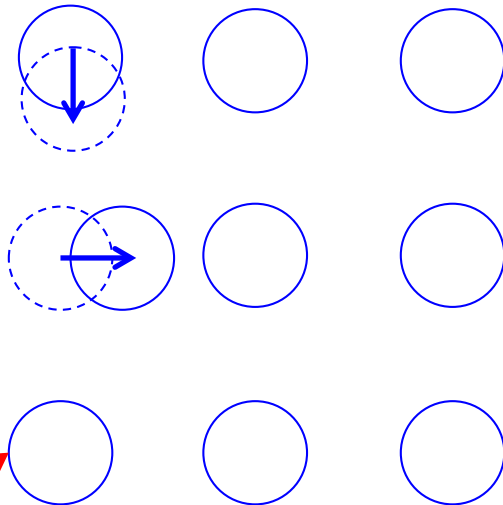
Phonon dispersion



AFLOW - www.aflowlib.org consortium

STATIC METADATA: Thermal Conductivity AFLOW-AAPL

AFLOW Anharmonic Phonon Library



$$\Delta \vec{F}$$

Perturb atoms, obtain change in force on others

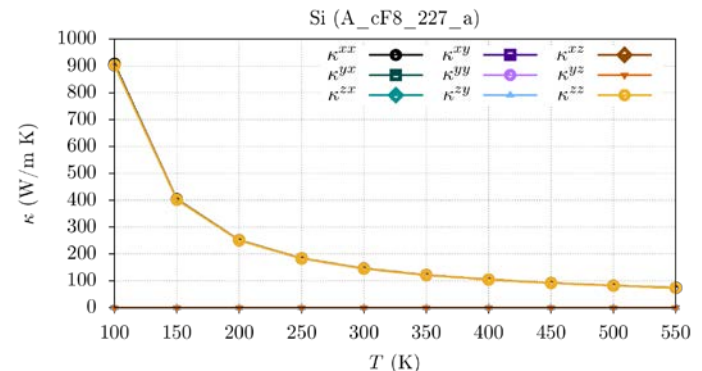
Interatomic force constants

$$\Phi_{ijk}^{\alpha\beta\gamma} = \frac{\partial^3 V}{\partial r_i^\alpha \partial r_j^\beta \partial r_k^\gamma}$$

Boltzmann Transport Equation

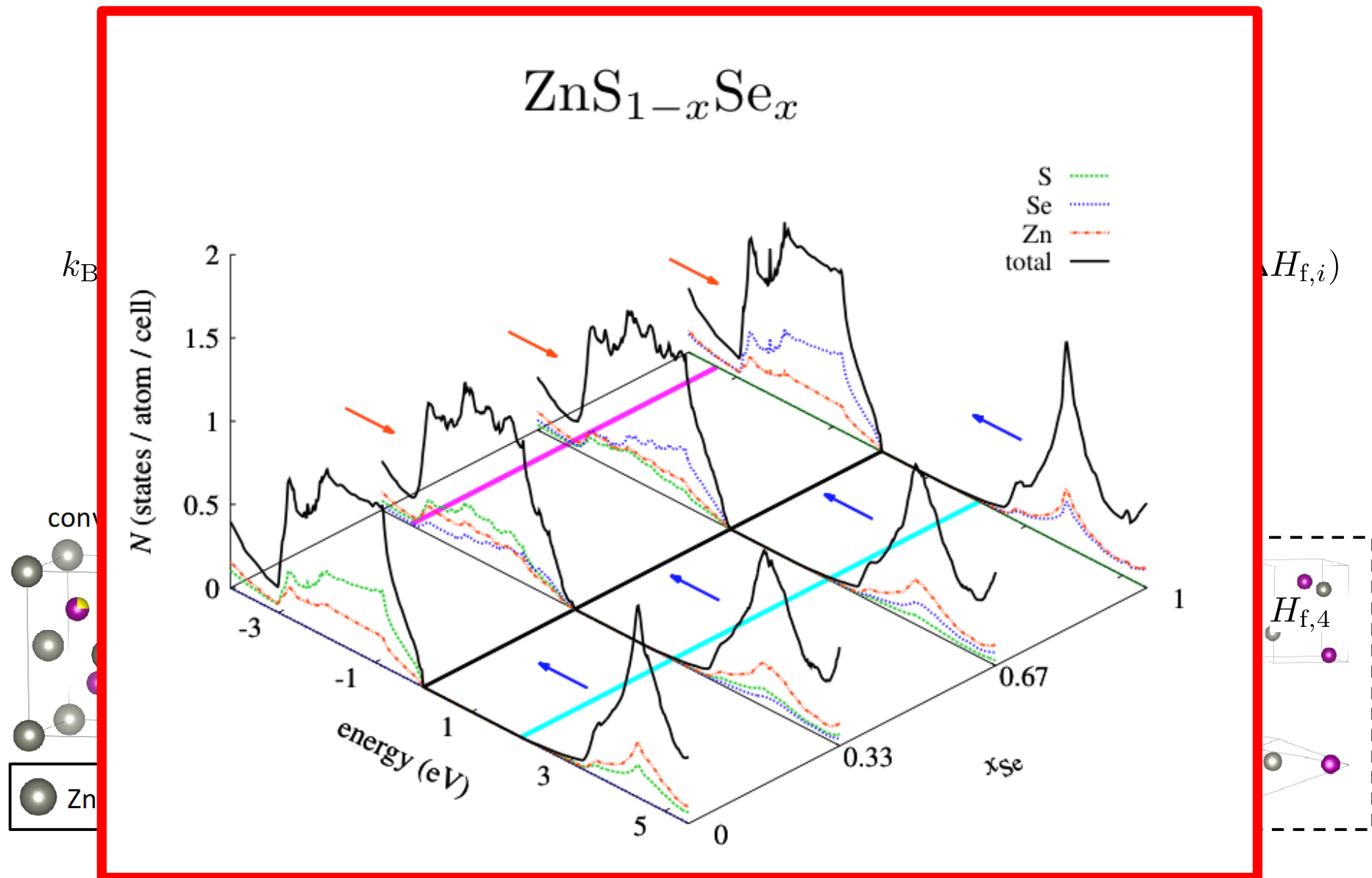
$$\kappa^{\alpha\beta} = \sum_{\lambda} C_{\lambda} v_{\lambda}^{\alpha} F_{\lambda}^{\beta}$$

Thermal conductivity



AFLOW - www.aflow.org consortium

STATIC METADATA: disorder as sum of order - AFLOW-POCC

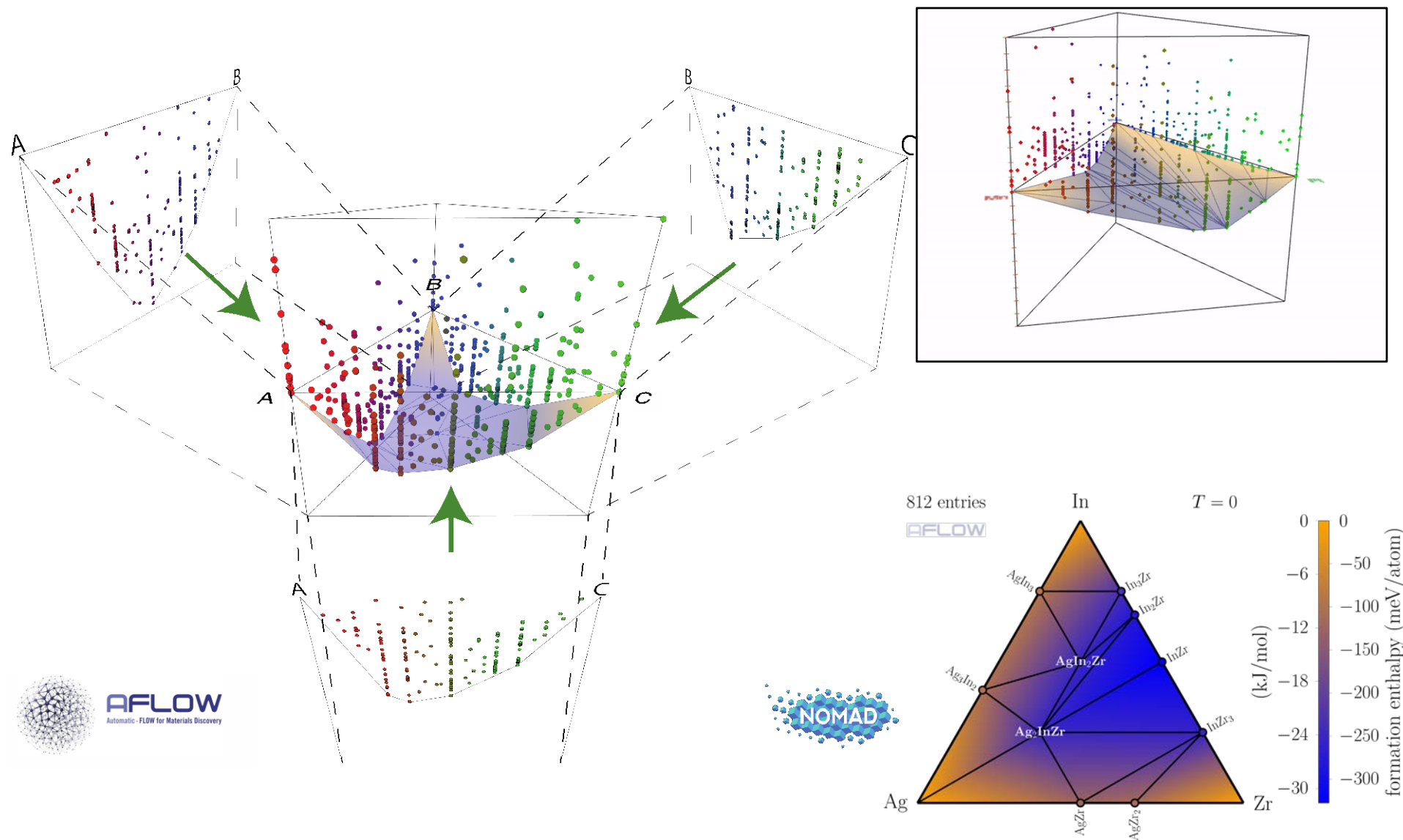


STATIC METADATA and IDENTIFIERS

aflow:e0ac186639d0a150

```
aflow_type=aggregate
method=aflow_apl
version=aflow_4
aggregate_parameters=...
aggregate_content=[
aflow:e49dd5cbb21edec5
aflow:ac91a692fbd41751
aflow:828099c6035556a5
aflow:2776de8554edceff
aflow:db62c43f06264fb3
aflow:cec73e2f39cbd4da
aflow:464f0aa2d7ebf15a
aflow:0ce04a4d6541e0e2
aflow:e6c66d8f0e02e9e4
aflow:03a40b741db4d6c4
aflow:fa78e94bd638dbef
..]
```

DYNAMIC METADATA: Convex Hull Phase AFLOW-CHULL



DYNAMIC METADATA: Convex Hull Phase AFLOW-CHULL

Go to aflow.org/aflow-chull and type/select AgAuCd

AgAuCd

APPOOL

Ag-Au-Cd summary
count=953

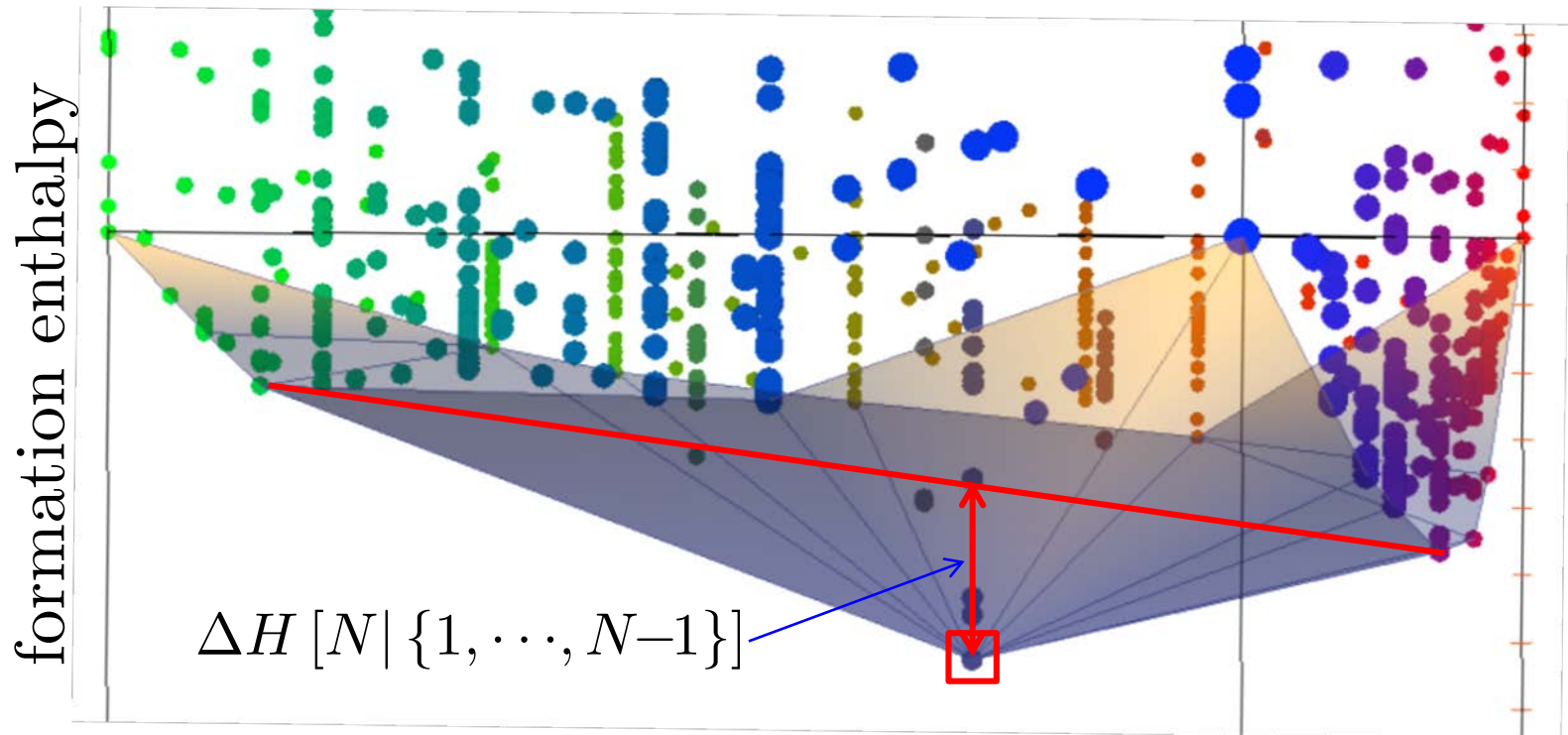
<http://aflow.org>
2018/07/27 16:42:28

Use of this data welcomes reference to the following publication:
C. Oses, E. Gossett, D. Hicks, F. Rose, M. J. Mehl, E. Perin, I. Takeuchi, S. Sanvito, M. Scheffler, V. Lederer, D. Levy, C. Toher, and S. Curtarolo,
AFLOW-CHULL: Cloud-oriented platform for autonomous phase stability analysis,
submitted arXiv:1806.06981 (2018).

POINTS DATA	REDUCED_COMPOUND	REDUCED_COMPOUND_LATEX	PROTOTYPE	PROTOTYPE_LATEX	AUID	AURL	URL_ENTRY_PAGE	SPACE_GROUP_ORIG	SPACE_GROUP_ORIG_LATEX	SPACE_GROUP_RELAT
Ag3	Ag	Ag	323	323	aflow:c2726f7969564d16	aflowlib.duke.edu:AFLOWDATA/LIB2_RAM/AgAu/323	http://aflow.org/material.php?id=aflow:c2726f7969564d16	P3_1(1)21#152	P3_1(1)21#152	Fm-3m#225
Ag3	Ag	Ag	48	48	aflow:520e54f7986a9f8ee	aflowlib.duke.edu:AFLOWDATA/LIB1_RAM/Ag/PAM_PBE:06Sep2000/AB	http://aflow.org/material.php?id=aflow:520e54f7986a9f8ee	P3_1(1)21#152	P3_1(1)21#152	Fm-3m#225
Ag1	Ag	Ag	303	303	aflow:38accf5c391aae61	aflowlib.duke.edu:AFLOWDATA/LIB2_RAM/AgAu/303	http://aflow.org/material.php?id=aflow:38accf5c391aae61	14/mmm#139	14/mmm#139	Fm-3m#225
Ag1	Ag	Ag	2	2	aflow:4852166165982ff68	aflowlib.duke.edu:AFLOWDATA/LIB2_RAM/AgAu/2	http://aflow.org/material.php?id=aflow:4852166165982ff68	Fm-3m#225	\$Fm\overline{1}n(3)m\#225\$	Fm-3m#225
Ag1	Ag	Ag	F1	F1	aflow:27c26f9621a50b08	aflowlib.duke.edu:AFLOWDATA/LIB2_RAM/AgAu/F1	http://aflow.org/material.php?id=aflow:27c26f9621a50b08	Fm-3m#225	\$Fm\overline{1}n(3)m\#225\$	Fm-3m#225
Ag1	Ag	Ag	A1	A1	aflow:b3de8b7b3385e9f	aflowlib.duke.edu:AFLOWDATA/LIB1_RAM/Ag/PAM_PBE:06Sep2000/A1	http://aflow.org/material.php?id=aflow:b3de8b7b3385e9f	Fm-3m#225	\$Fm\overline{1}n(3)m\#225\$	Fm-3m#225
Ag2	Ag	Ag	A7_A	A7_A	aflow:4345f979f9458342	aflowlib.duke.edu:AFLOWDATA/LIB2_RAM/AgAu/A7_A	http://aflow.org/material.php?id=aflow:4345f979f9458342	R-3m#166	\$R\overline{3}m(3)m\#166\$	Fm-3m#225
Ag2	Ag	Ag	A7	A7	aflow:1f7c7646d0c2b378	aflowlib.duke.edu:AFLOWDATA/LIB1_RAM/Ag/PAM_PBE:06Sep2000/A7	http://aflow.org/material.php?id=aflow:1f7c7646d0c2b378	R-3m#166	\$R\overline{3}m(3)m\#166\$	Fm-3m#225
Ag2	Ag	Ag	A3	A3	aflow:806f3a3361a004e6	aflowlib.duke.edu:AFLOWDATA/LIB1_RAM/Ag/PAM_PBE:06Sep2000/A3	http://aflow.org/material.php?id=aflow:806f3a3361a004e6	P6_3(1)/mmc#194	P6_3(1)/mmc#194	P6_3(1)/mmc#194
Ag2	Ag	Ag	117	117	aflow:68c6126a7e618e2	aflowlib.duke.edu:AFLOWDATA/LIB2_RAM/AgAu/117	http://aflow.org/material.php?id=aflow:68c6126a7e618e2	P6_3(1)/mmc#194	P6_3(1)/mmc#194	P6_3(1)/mmc#194
Ag2	Ag	Ag	n1	n1	aflow:216998e5c3852a87	aflowlib.duke.edu:AFLOWDATA/LIB2_RAM/AgAu/n1	http://aflow.org/material.php?id=aflow:216998e5c3852a87	P6_3(1)/mmc#194	P6_3(1)/mmc#194	P6_3(1)/mmc#194
Ag1	Ag	Ag	S9	S9	aflow:2b272763a4cf97b9	aflowlib.duke.edu:AFLOWDATA/LIB2_RAM/AgAu/S9	http://aflow.org/material.php?id=aflow:2b272763a4cf97b9	In-3m#229	\$In\overline{3}m(2)29\$	In-3m#229
Ac1	Au	Au	b1	b1	aflow:8baef16d9d4d4fcd	aflowlib.duke.edu:AFLOWDATA/LIB2_RAM/AgAu/b1	http://aflow.org/material.php?id=aflow:8baef16d9d4d4fcd	In-3m#229	\$In\overline{3}m(2)29\$	In-3m#229

DYNAMIC METADATA: Δ Convex Hull - AFLOW-CHULL

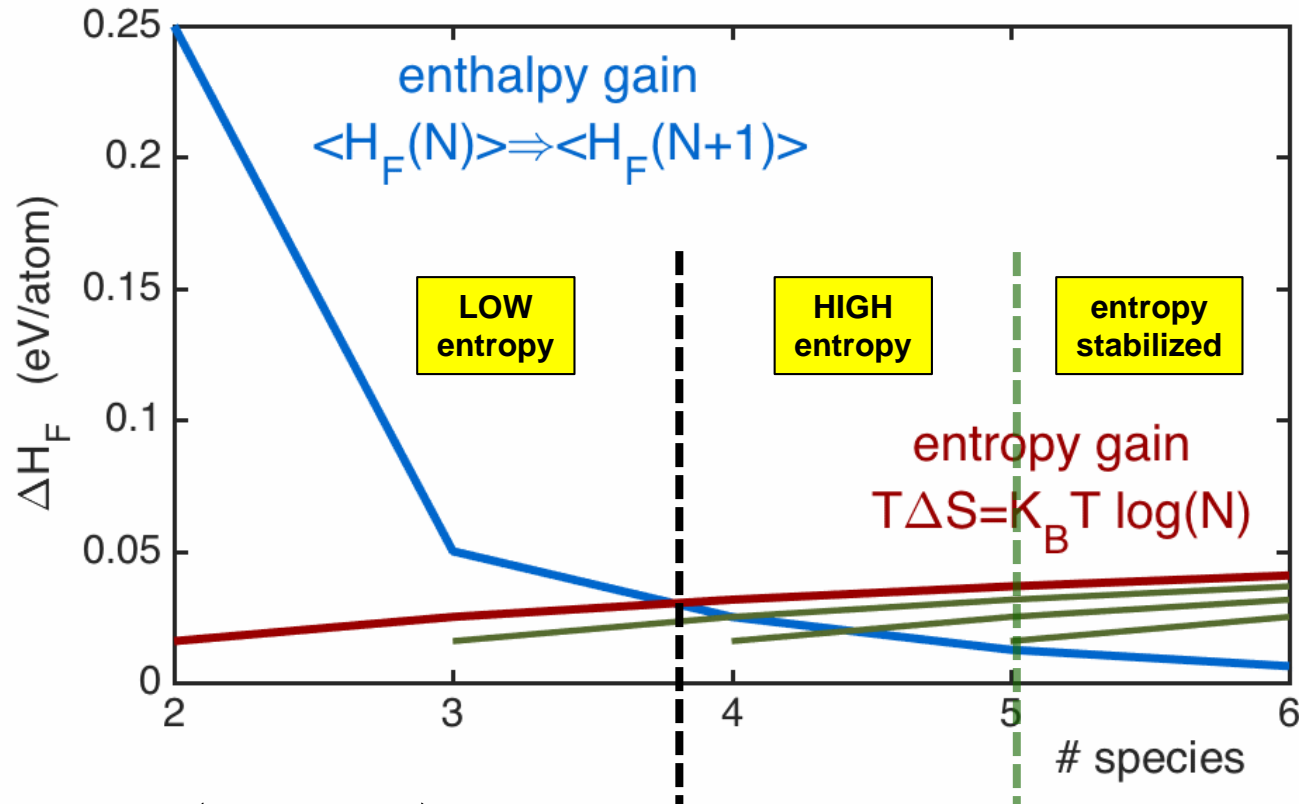
“Enthalpy gain”: energetic distance to hull formed from N-1 species components: $\Delta H [N | \{1, \dots, N-1\}]$



$$\Delta H [N | \{1, \dots, N-1\}] \equiv H_{\text{hull}} [\{1, \dots, N-1\}] - H [N]$$

if $H [N] < H_{\text{hull}} [\{1, \dots, N-1\}]$; otherwise is zero.

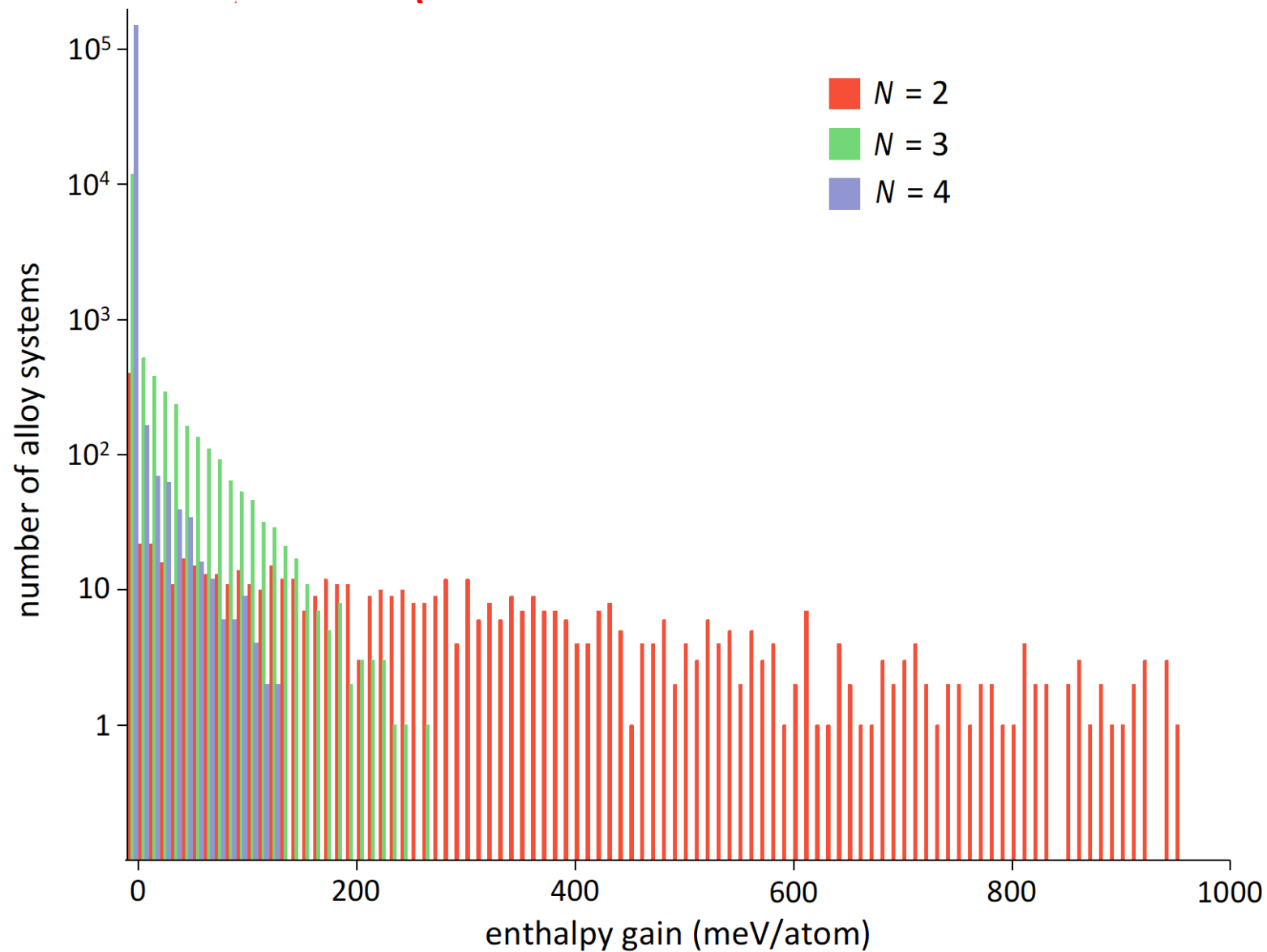
Theory from METADATA: N+1 theorem



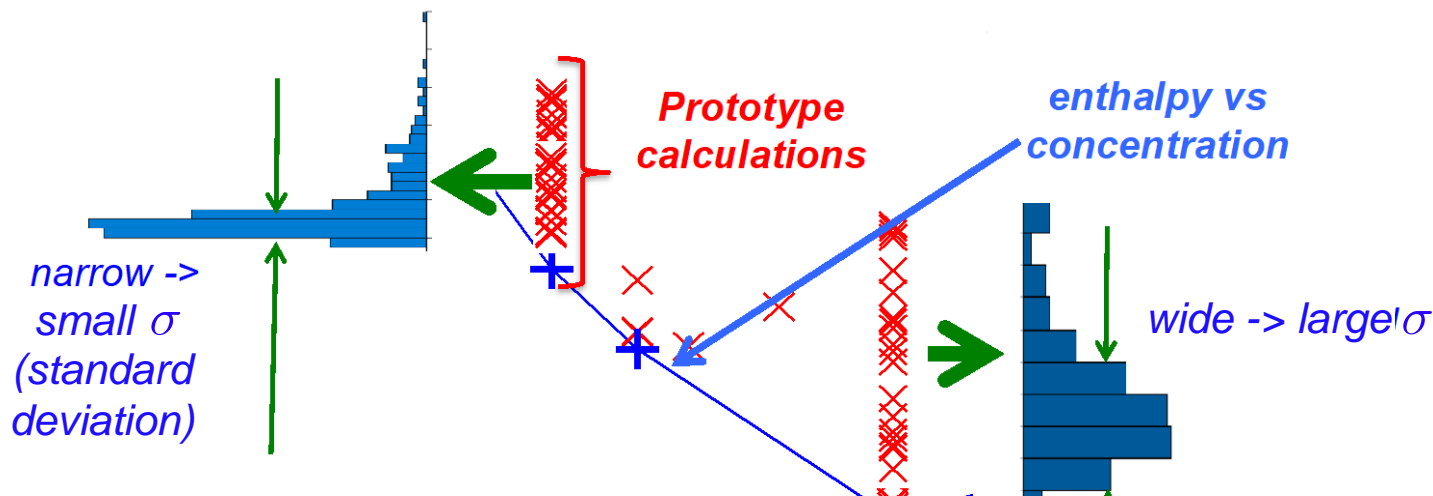
$\min(H - TS)$
 if $[\langle H(N+1) \rangle - \langle H(N) \rangle] > TS(N)$ $\leftarrow \min(H)$ $\min(-TS)$ \rightarrow
 if $[\langle H(N+1) \rangle - \langle H(N) \rangle] \lesssim TS(N)$



Materials from METADATA: The ground-state-less truth



DYNAMIC METADATA: crude ML measuring entropy

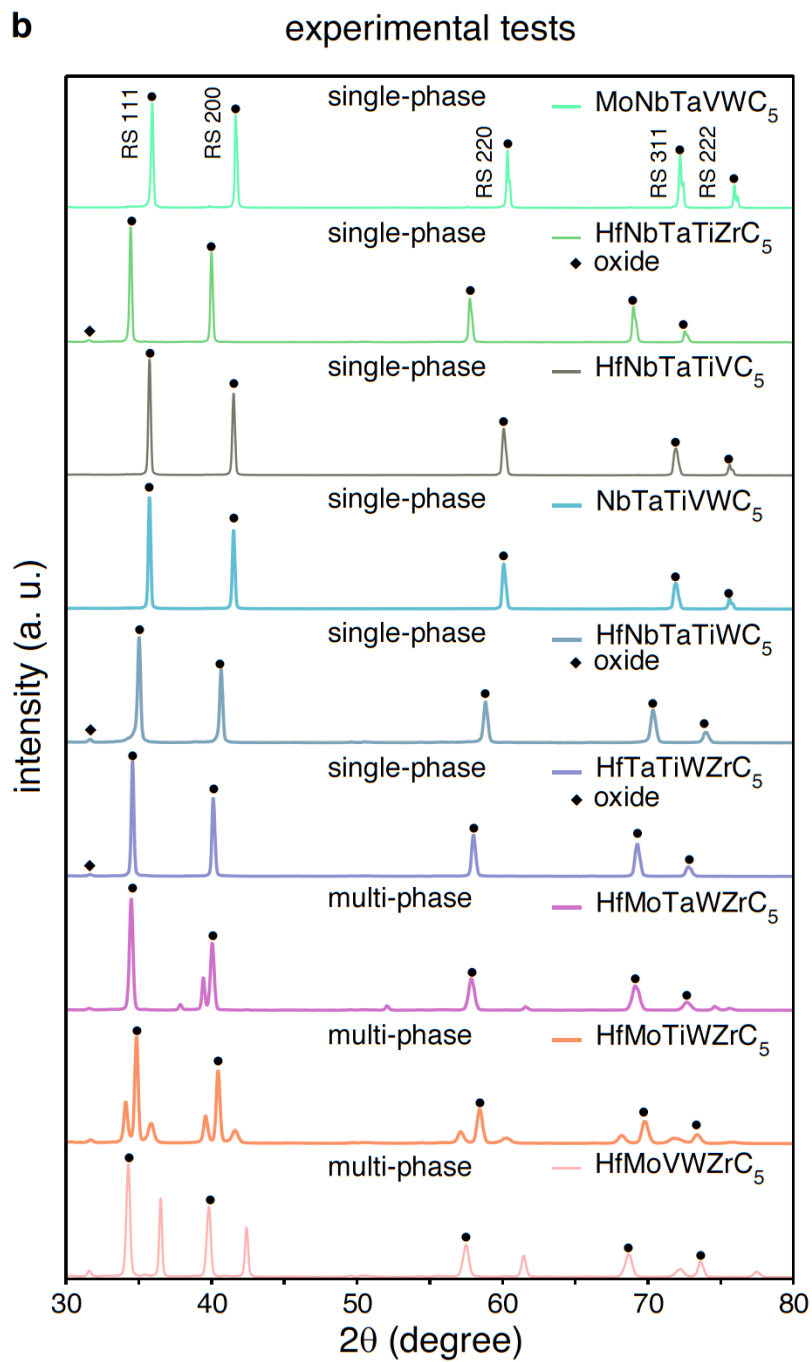
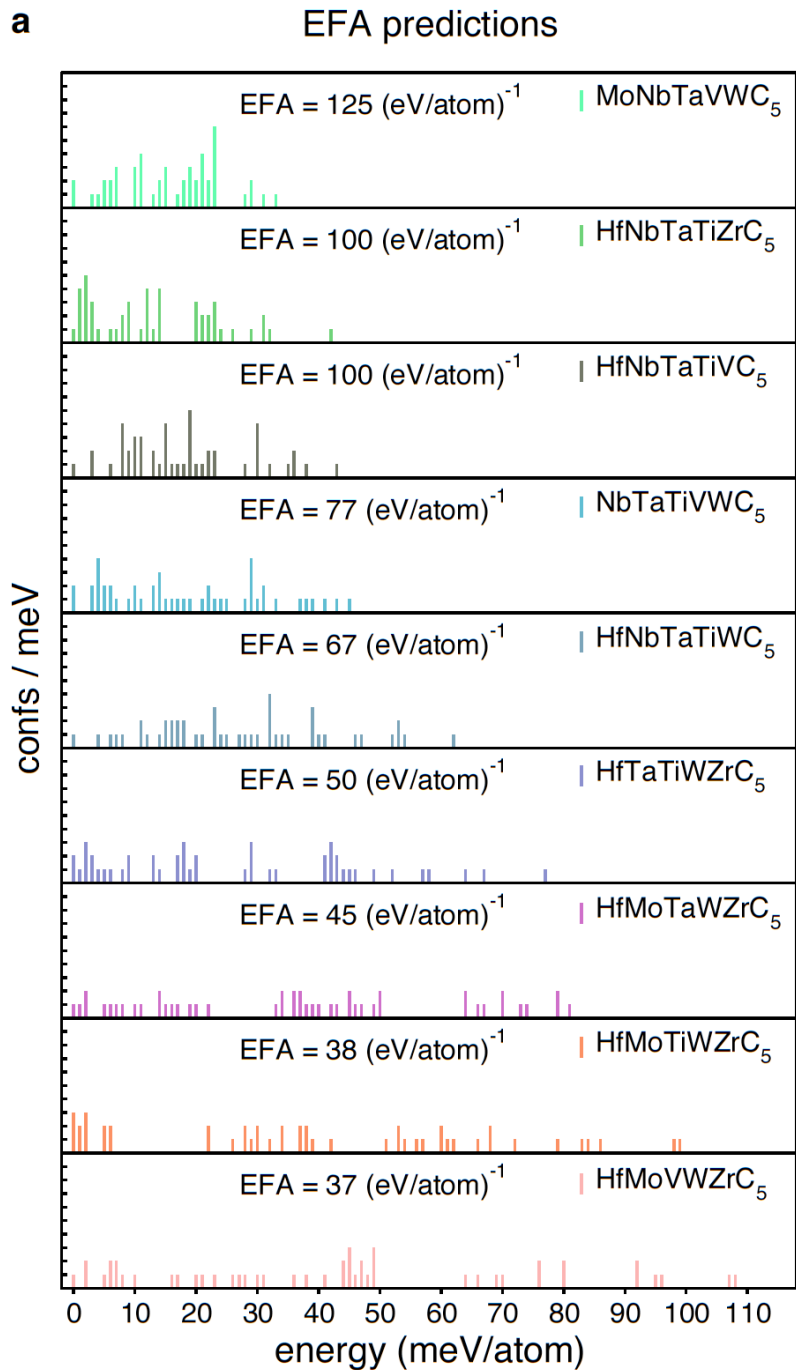


AFLOW-POCC: combinations
AFLOW-SYM: degenerations

EFA(N)

$$\sigma\{H_i(N)\} = \sqrt{\frac{\sum_{i=1}^n g_i (H_i - H_{\text{mix}})^2}{\left(\sum_{i=1}^n g_i\right) - 1}},$$

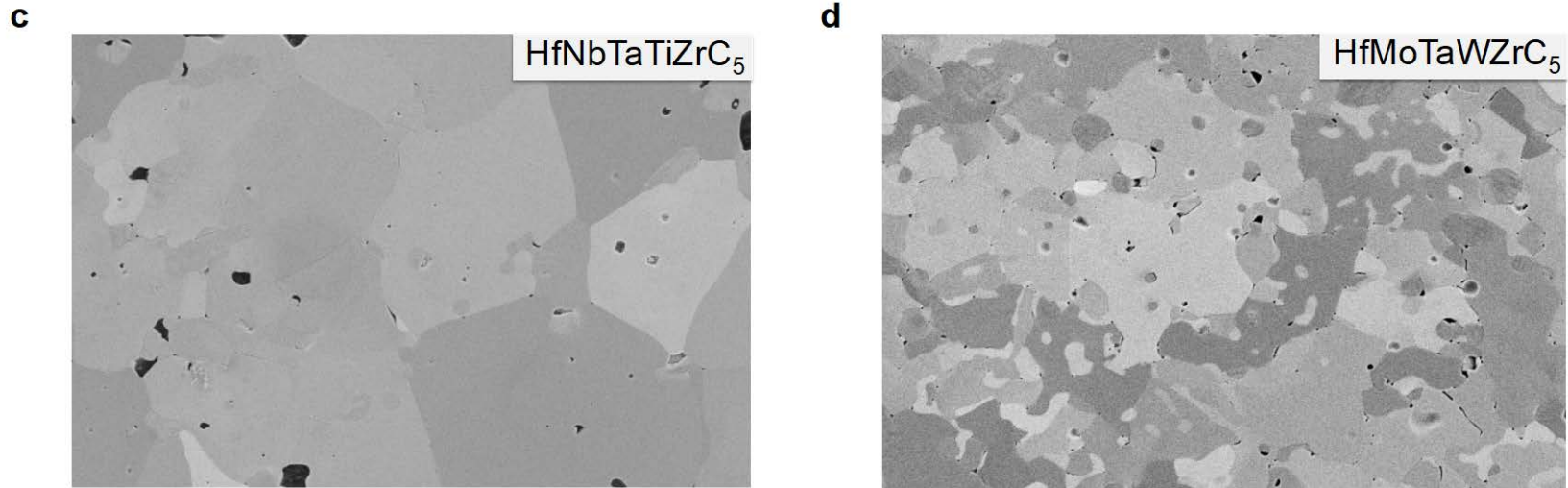
-1



OK

KO

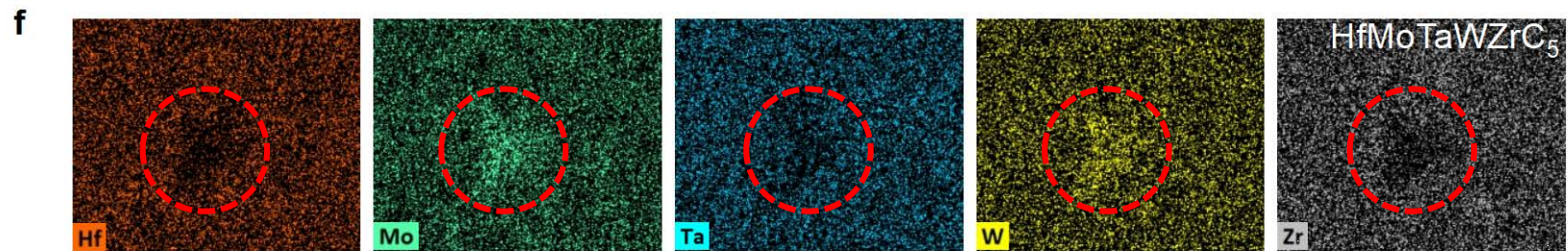
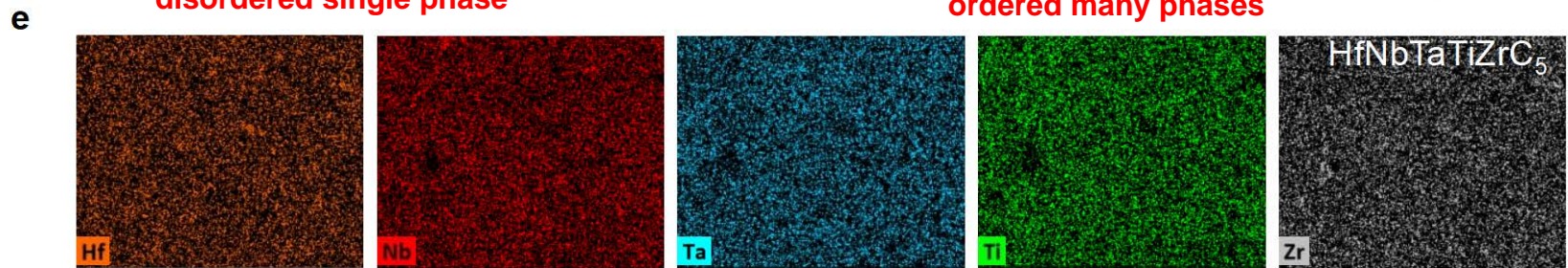
5-metal carbides: Hf Mo Nb Ta Ti V W Zr + C



**homogeneous
disordered single phase**

**non homogeneous
ordered many phases**

10 μm



10 μm

CONCLUSIONS



AFLOW tries to be as FAIR as possible.

- Findable: AUIDs, rich, searchable
- Accessible: retrievable, AFLUX-able
- Interoperable: accessible, shared and simple vocabulary
- Reusable: through python, APE, and web-apps