

Machine Learning-Assisted Design of Cesium-Based Photocathodes from First Principles

Holger-Dietrich Saßnick¹ and Caterina Cocchi^{1,2}

¹Carl von Ossietzky Universität Oldenburg, Physics Department, 26129 Oldenburg

²Humboldt-Universität zu Berlin, Physics Department and IRIS Adlershof, 12489 Berlin

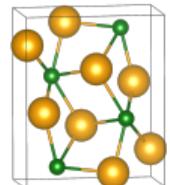
Ab initio methods like density functional theory (DFT) allow the calculation of material's properties without empirical parameters.



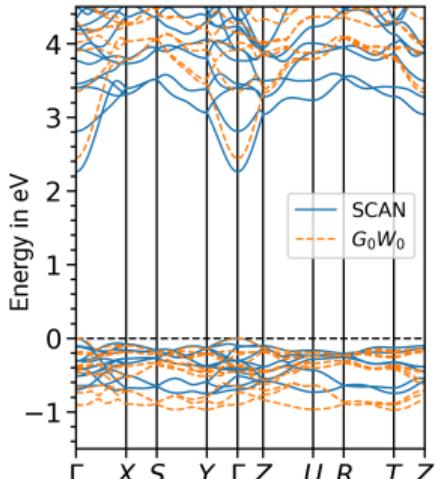
57 La Lanthanum 138.91	58 Ce Cerium 140.12	59 Pr Praseodymium 140.91	60 Nd Neodymium 144.24	61 Pm Promethium (141)	62 Sm Samarium 150.36	63 Eu Europium 151.96	64 Gd Gadolinium 157.25	65 Tb Terbium 158.93	66 Dy Dysprosium 162.50	67 Ho Holmium 164.93	68 Er Erbium 167.26	69 Tm Thulium 168.93	70 Yb Ytterbium 173.05	71 Lr Lutetium 174.97
80 Ac Actinium (227)	90 Th Thorium 232.04	91 Pa Protactinium 231.04	92 U Uranium 238.03	93 Np Neptunium (237)	94 Pu Plutonium (244)	95 Am Americium (243)	96 Cm Curium (247)	97 Bk Berkelium (247)	98 Cf Californium (251)	99 Es Einsteinium (252)	100 Fm Fermium (257)	101 Md Mendelevium (258)	102 No Nobelium (259)	103 Lr Lawrencium (266)

Ab initio methods like density functional theory (DFT) allow the calculation of material's properties without empirical parameters.

Electronic Properties



Cs_2Te
 $Pnma$ [62]

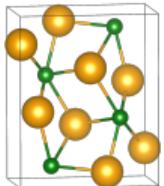


Cocchi and Saßnick, *Micromachines*
2021, 12(9), 1002;
doi: 10.3390/mi12091002

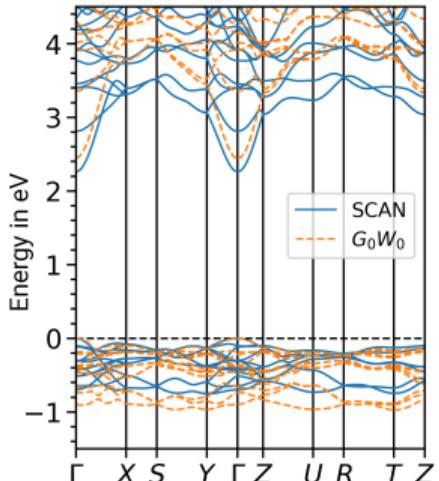
2 He Helium 4.003	10 Ne Neon 20.180
5 B Boron 10.810	6 C Carbon 12.011
7 N Nitrogen 14.007	8 O Oxygen 15.999
9 F Fluorine 18.998	10 Ar Argon 39.948
13 Al Aluminum 26.982	14 Si Silicon 28.085
15 P Phosphorus 30.974	16 S Sulfur 32.060
17 Cl Chlorine 35.450	18 Kr Krypton 83.798
31 Ga Gallium 69.723	32 Ge Germanium 72.630
33 As Arsenic 74.922	34 Se Selenium 78.971
35 Br Bromine 79.904	36 Kr Krypton 83.798
51 Sb Antimony 121.760	52 Te Tellurium 127.690
53 I Iodine 126.900	54 Xe Xenon 131.200
75 Re Rhodium 186.230	76 Os Osmium 190.230
77 Ir Iridium 192.220	78 Pt Platinum 195.080
79 Au Gold 196.970	80 Hg Mercury 200.590
81 Tl Thallium 204.38	82 Pb Lead 207.200
83 Bi Bismuth 208.980	84 Po Polonium (209)
85 At Astatine (210)	86 Rn Radon (222)
1 P Hydrogen 1.008	2 He Helium 4.003
3 Li Lithium 6.941	4 Be Beryllium 9.012
5 B Boron 10.810	6 C Carbon 12.011
7 N Nitrogen 14.007	8 O Oxygen 15.999
9 F Fluorine 18.998	10 Ar Argon 39.948
11 Na Sodium 22.988	12 Mg Magnesium 24.312
13 Al Aluminum 26.982	14 Si Silicon 28.085
15 P Phosphorus 30.974	16 S Sulfur 32.060
17 Cl Chlorine 35.450	18 Kr Krypton 83.798
19 K Potassium 39.098	20 Ca Calcium 40.078
21 Sc Scandium 44.956	22 Ti Titanium 45.988
23 V Vanadium 50.942	24 Cr Chromium 51.996
25 Mn Manganese 54.938	26 Fe Iron 55.845
27 Co Cobalt 58.933	28 Ni Nickel 58.693
29 Cu Copper 63.546	30 Zn Zinc 65.380
31 Ga Gallium 69.723	32 Ge Germanium 72.630
33 As Arsenic 74.922	34 Se Selenium 78.971
35 Br Bromine 79.904	36 Kr Krypton 83.798
37 Rb Rubidium 84.790	38 Sr Strontium 87.620
39 Y Lanthanum 88.902	40 Zr Zirconium 91.224
41 Nb Niobium 92.906	42 Ta Tantalum 101.122
43 Tc Technetium (98)	44 Ru Ruthenium 101.070
45 Rh Rhodium 102.910	46 Pd Palladium 106.420
47 Ag Silver 107.870	48 Cd Cadmium 112.410
49 In Indium 114.820	50 Sn Tin 118.710
51 Sb Antimony 121.760	52 Te Tellurium 127.690
53 I Iodine 126.900	54 Xe Xenon 131.200
55 Cs Cesium 132.915	56 Ba Barium 137.340
57 La Lanthanum 138.912	58 Ce Cerium 140.119
59 Pr Praseodymium 140.912	60 Nd Neodymium 144.242
61 Pm Promethium (141)	62 Sm Samarium 150.36
63 Eu Europium 151.96	64 Gd Gadolinium 157.25
65 Tb Terbium 158.93	66 Dy Dysprosium 162.50
67 Ho Holmium 164.93	68 Er Erbium 167.26
69 Tm Thulium 168.93	70 Yb Ytterbium 173.05
71 Lu Lutetium 174.97	
91 Tc Technetium (98)	92 Ru Ruthenium (101.070)
93 Rh Rhodium (102.910)	94 Pd Palladium (106.420)
95 Os Osmium (190.230)	96 Ir Iridium (192.220)
97 Pt Platinum (195.080)	98 Au Gold (196.970)
99 Hg Mercury (200.590)	100 Tl Thallium (204.38)
101 Pb Lead (207.200)	102 Bi Bismuth (208.980)
103 Po Polonium (209)	104 At Astatine (210)
105 Rn Radon (222)	

Ab initio methods like density functional theory (DFT) allow the calculation of material's properties without empirical parameters.

Electronic Properties

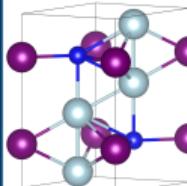


Cs₂Te
Pnma [62]

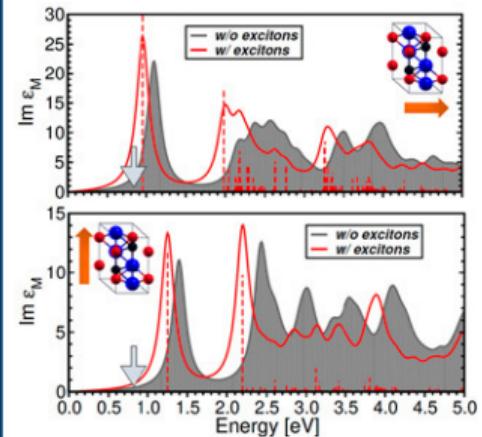


Cocchi and Saßnick, *Micromachines*
2021, 12(9), 1002;
doi: 10.3390/mi12091002

Optical Properties



NaK₂Sb
P6₃/mmc [194]

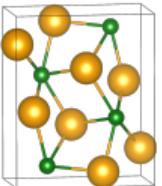


Amador, Saßnick and Cocchi, *J. Phys.: Condens. Matter* 2021, 33, 365502;
doi: 10.1088/1361-648X/ac0e70

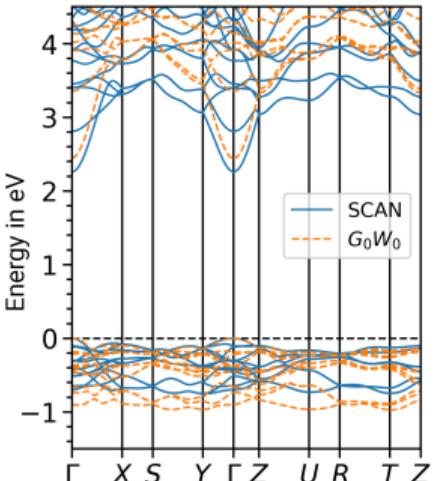
1	H	Hydrogen	4.003
2	He	Helium	4.003
3	Li	Lithium	6.941
4	Be	Boron	10.810
5	B	Carbon	12.011
6	C	Carbon	12.011
7	N	Nitrogen	14.007
8	O	Oxygen	15.999
9	F	Fluorine	18.998
10	Ne	Neon	20.180
11	Na	Sodium	22.990
12	Mg	Magnesium	24.312
13	Al	Aluminum	26.982
14	Si	Silicon	28.085
15	P	Phosphorus	30.974
16	S	Sulfur	32.060
17	Cl	Chlorine	35.450
18	Ar	Argon	39.948
19	K	Kalium	39.102
20	Ca	Calcium	40.123
21	Sc	Scandium	44.959
22	Ti	Titanium	47.867
23	V	Vanadium	50.944
24	Cr	Chromium	52.000
25	Mn	Manganese	54.938
26	Fe	Iron	55.845
27	Co	Cobalt	58.933
28	Ni	Nickel	58.693
29	Cu	Copper	63.546
30	Zn	Zinc	65.402
31	Ge	Germanium	72.630
32	As	Arsenic	74.922
33	Se	Selenium	78.971
34	Br	Bromine	79.904
35	I	Iodine	126.900
36	Xe	Xenon	131.200
37	Rb	Rubidium	140.907
38	Y	Yttrium	167.26
39	Tl	Thallium	180.902
40	Pb	Pb	207.200
41	Bi	Bismuth	208.980
42	Po	Po	(209)
43	At	At	(210)
44	Rn	Rn	(222)
45	Fr	Fr	(223)
46	Ra	Ra	(226)
47	Ac	Ac	(227)
48	Th	Th	(232)
49	Pa	Pa	(231)
50	U	U	(238)
51	Np	Np	(237)
52	Pm	Pm	(247)
53	Am	Am	(243)
54	Cm	Cm	(250)
55	Bk	Bk	(247)
56	Cf	Cf	(251)
57	Md	Md	(258)
58	No	No	(259)
59	Lr	Lr	(266)

Ab initio methods like density functional theory (DFT) allow the calculation of material's properties without empirical parameters.

Electronic Properties

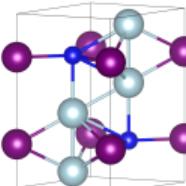


Cs_2Te
 $Pnma$ [62]

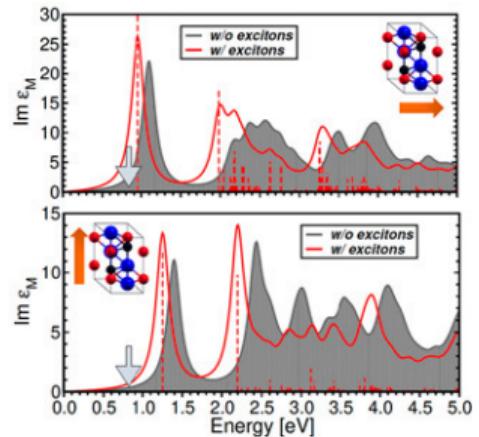


Cocchi and Saßnick, *Micromachines* 2021, 12(9), 1002;
doi: 10.3390/mi12091002

Optical Properties

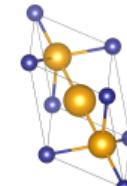


NaK_2Sb
 $P6_3/mmc$ [194]

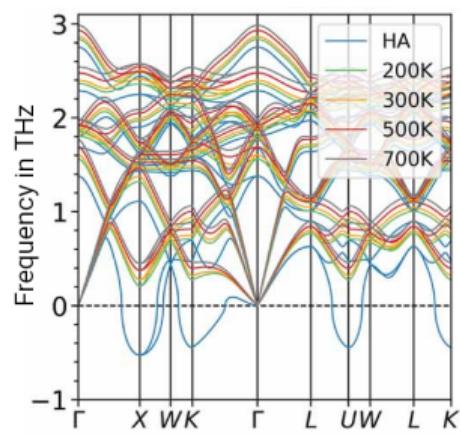


Amador, Saßnick and Cocchi, *J. Phys.: Condens. Matter* 2021, 33, 365502;
doi: 10.1088/1361-648X/ac0e70

Vibrational Properties

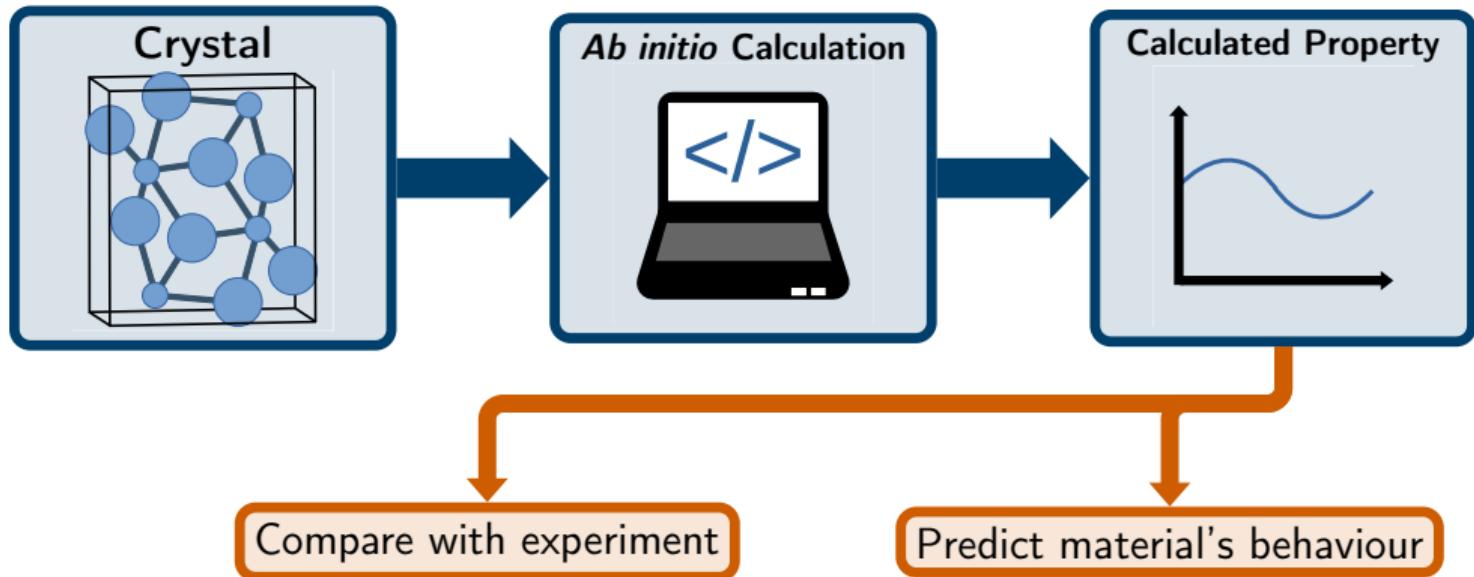


Cs_3Sb
 $Fm\bar{3}m$ [225]

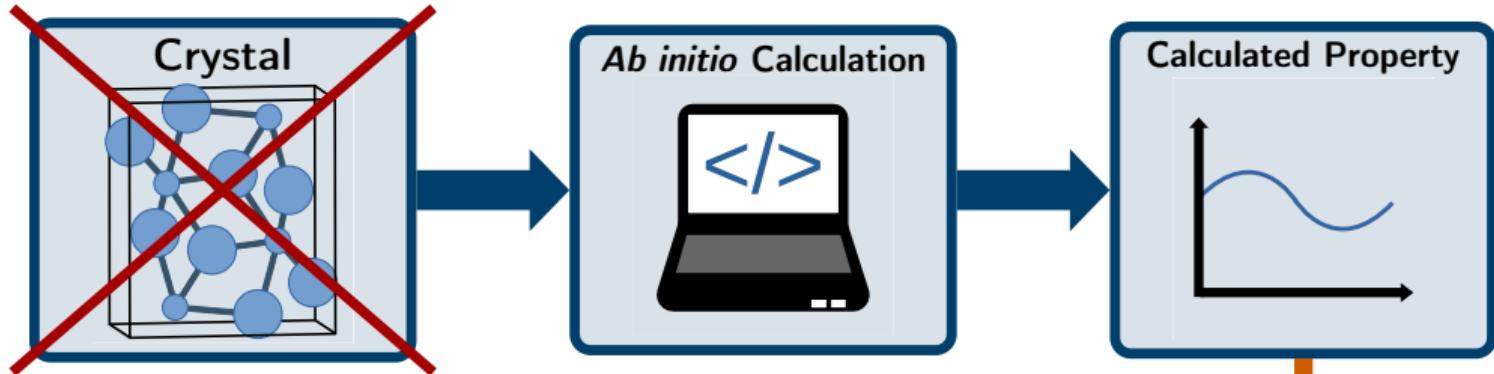


Santana-Andreo, Saßnick and Cocchi, *J. Phys. Mater.* 2024, 7, 035004;
doi: 10.1088/2515-7639/ad510b

Structural information about the material is a required input.



Structural information about the material is a required input.



- ▶ In some cases the atomistic structure of the studied material might not be known or is very complex.
- ▶ Can we find representative crystal structures without any experimental input?

Categories of Crystal Structure Prediction Methods

Methods with fixed stoichiometry:

- ▶ Minima hopping.

Goedecker, *Modern Methods of Crystal Structure Prediction* 2010, John Wiley & Sons, Ltd. Chap. 6, 131–145;
doi: 10.1002/9783527632831.ch6

- ▶ Bayesian optimization.

Yamashita et al., *Phys. Rev. Mater.* 2018, 2(1), 013803; doi: 10.1103/PhysRevMaterials.2.013803

Methods without stoichiometry constraints:

- ▶ Random structure search.

Harper et al. *Johnson Matthey Technology Review* 2020, 64.2, 103–118; doi: 10.1595/205651320X15742491027978

- ▶ Genetic algorithms.

Lyakhov et al. *Modern Methods of Crystal Structure Prediction* 2010, John Wiley & Sons, Ltd. Chap. 7, 147–180;
doi: 10.1002/9783527632831.ch7

Categories of Crystal Structure Prediction Methods

Methods with fixed stoichiometry:

- ▶ Minima hopping.

Goedecker, *Modern Methods of Crystal Structure Prediction* 2010, John Wiley & Sons, Ltd. Chap. 6, 131–145;
doi: 10.1002/9783527632831.ch6

- ▶ Bayesian optimization.

Yamashita et al., *Phys. Rev. Mater.* 2018, 2(1), 013803; doi: 10.1103/PhysRevMaterials.2.013803

Methods without stoichiometry constraints:

- ▶ Random structure search.

Harper et al. *Johnson Matthey Technology Review* 2020, 64.2, 103–118; doi: 10.1595/205651320X15742491027978

- ▶ Genetic algorithms.

Lyakhov et al. *Modern Methods of Crystal Structure Prediction* 2010, John Wiley & Sons, Ltd. Chap. 7, 147–180;
doi: 10.1002/9783527632831.ch7

All methods share large computational costs and/or are inflexible in its application.

Categories of Crystal Structure Prediction Methods

Methods with fixed stoichiometry:

- ▶ Minima hopping.

Goedecker, *Modern Methods of Crystal Structure Prediction* 2010, John Wiley & Sons, Ltd. Chap. 6, 131–145;
doi: 10.1002/9783527632831.ch6

- ▶ Bayesian optimization.

Yamashita et al., *Phys. Rev. Mater.* 2018, 2(1), 013803; doi: 10.1103/PhysRevMaterials.2.013803

Methods without stoichiometry constraints:

- ▶ Random structure search.

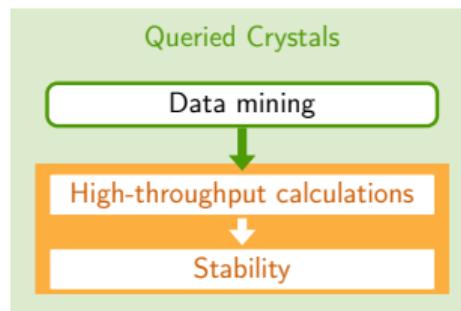
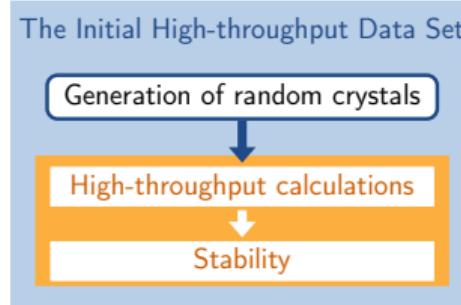
Harper et al. *Johnson Matthey Technology Review* 2020, 64.2, 103–118; doi: 10.1595/205651320X15742491027978

- ▶ Genetic algorithms.

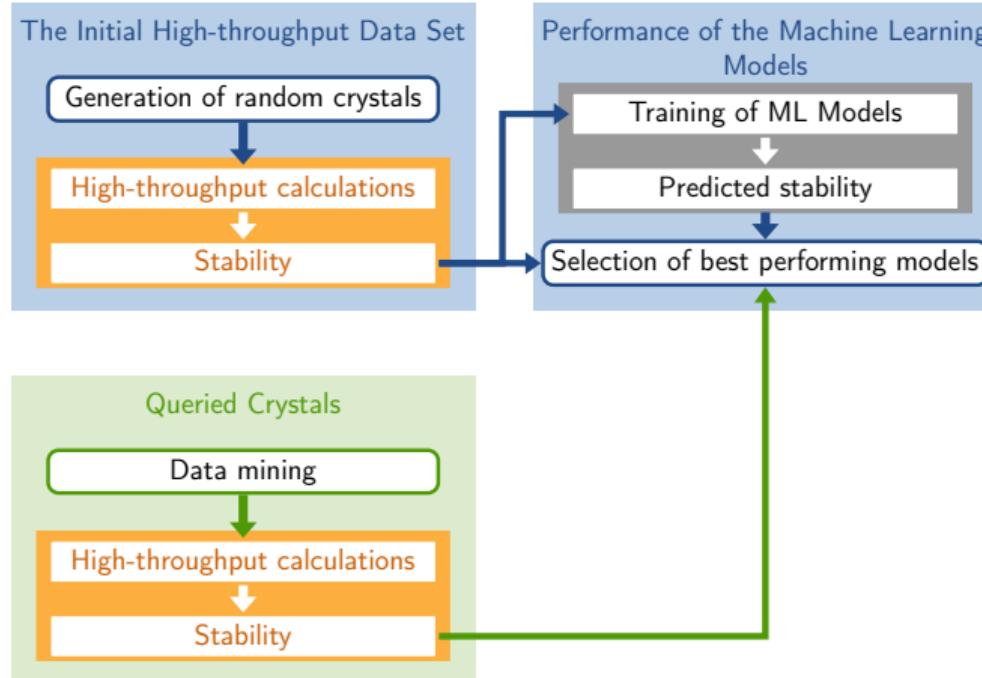
Lyakhov et al. *Modern Methods of Crystal Structure Prediction* 2010, John Wiley & Sons, Ltd. Chap. 7, 147–180;
doi: 10.1002/9783527632831.ch7

This approach combines random structure search with machine learning to alleviate these issues.

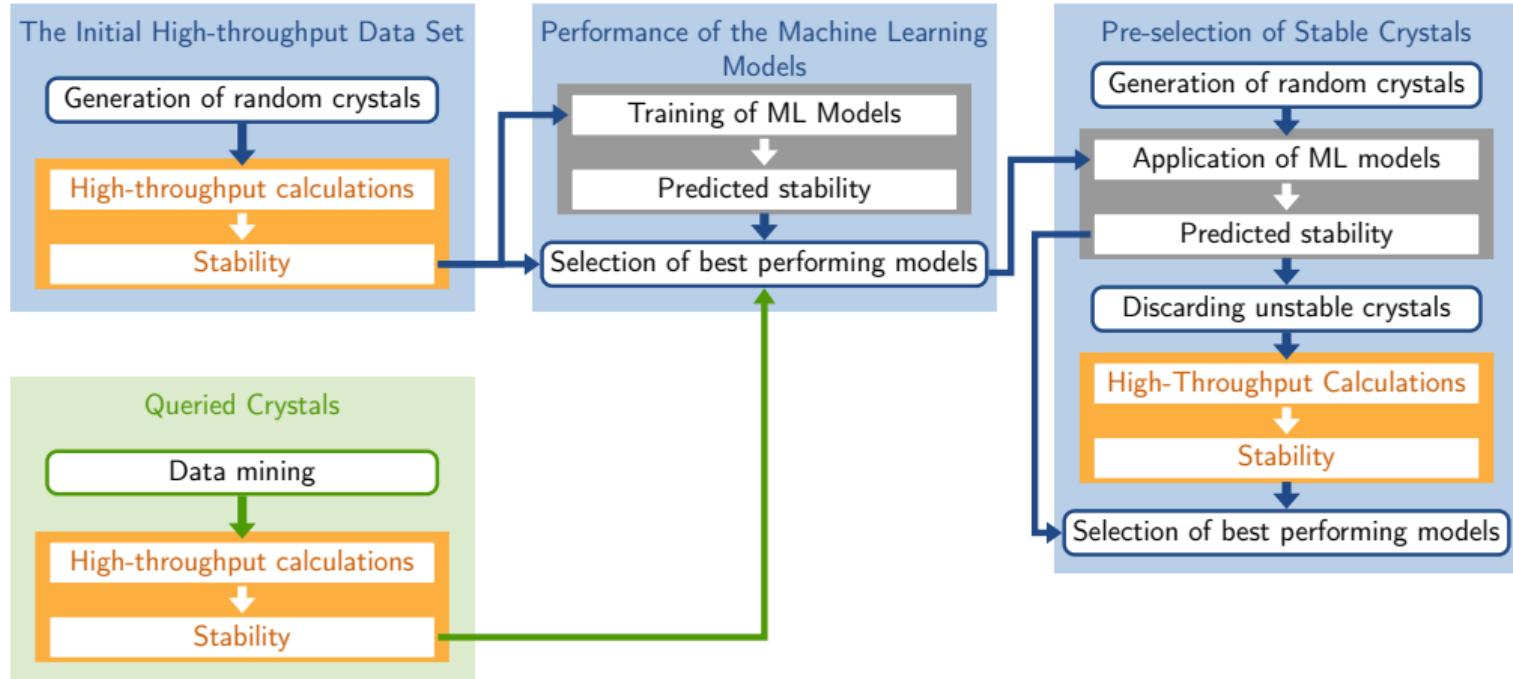
Overview of the Crystal Structure Prediction Approach



Overview of the Crystal Structure Prediction Approach



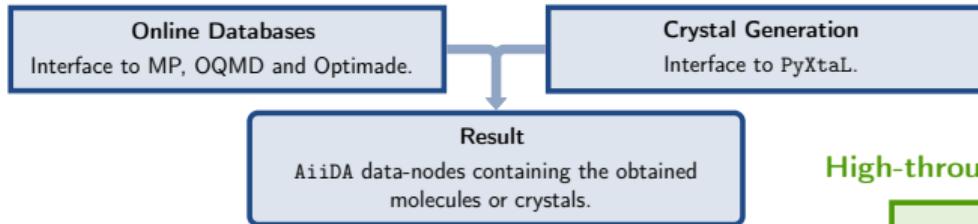
Overview of the Crystal Structure Prediction Approach



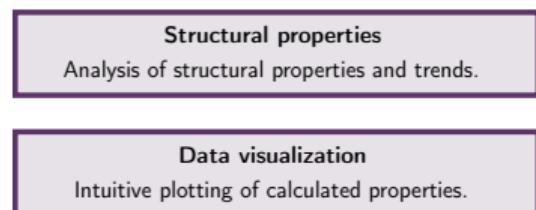
Implementation in aim²dat



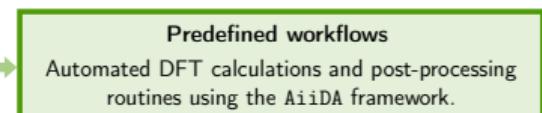
Data-mining (strct.StructureImporter).



Data analysis and visualization (strct/plots)



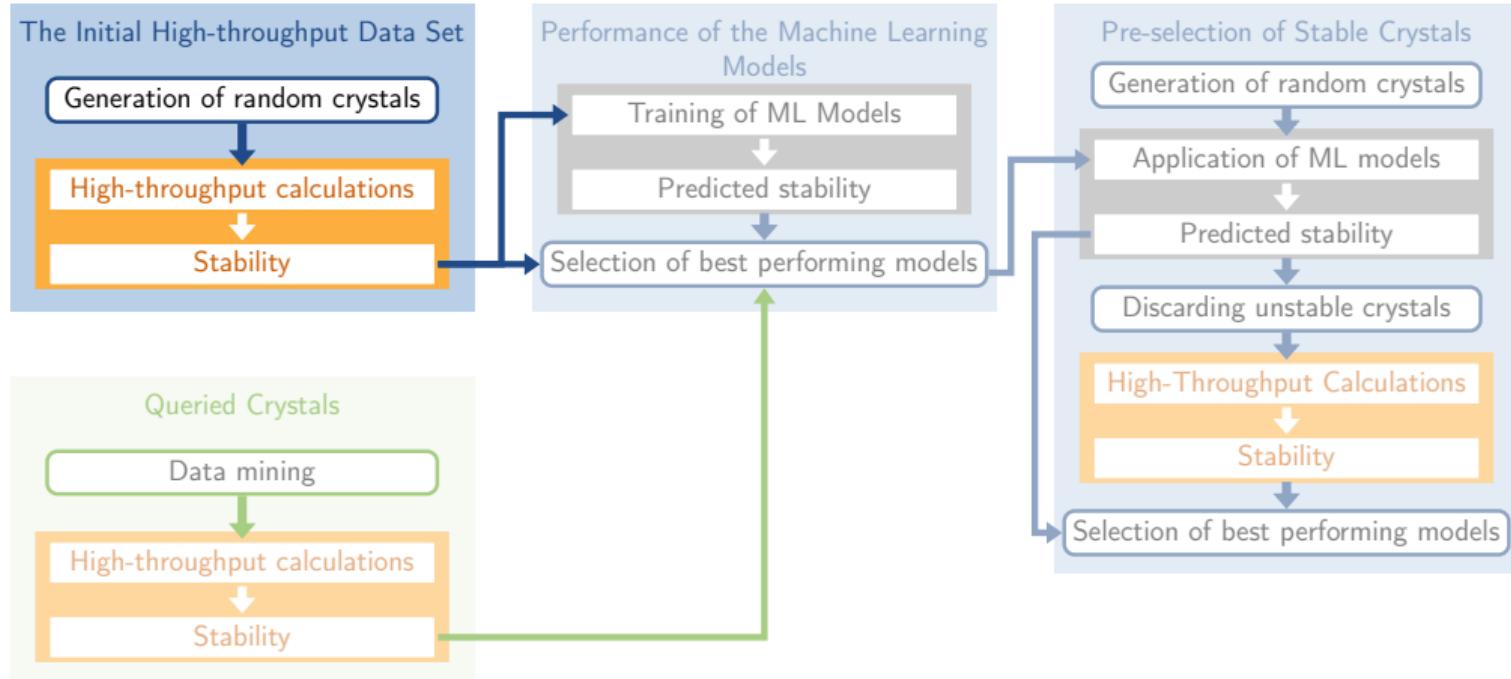
High-throughput screening (aiida_workflows).



Machine learning (ml)



Overview of the Crystal Structure Prediction Approach

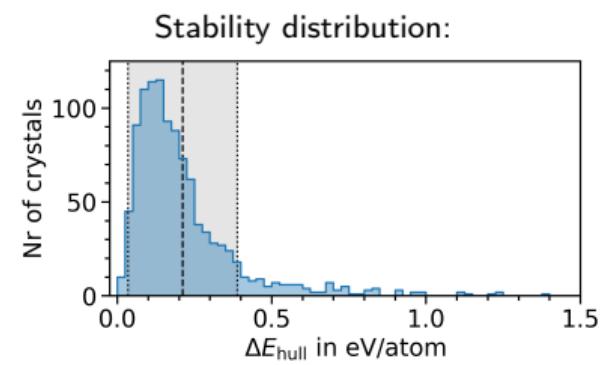
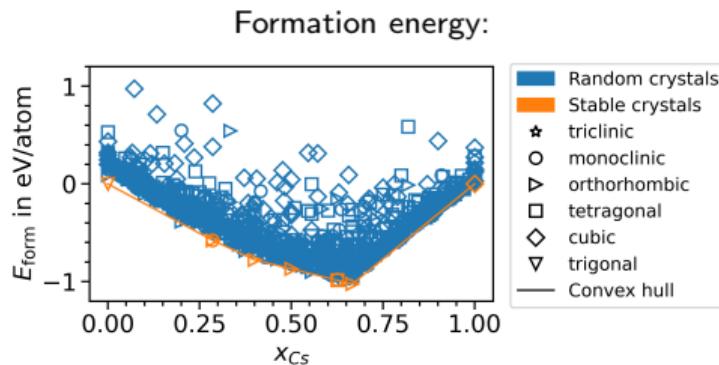


The Initial High-Throughput Data Set

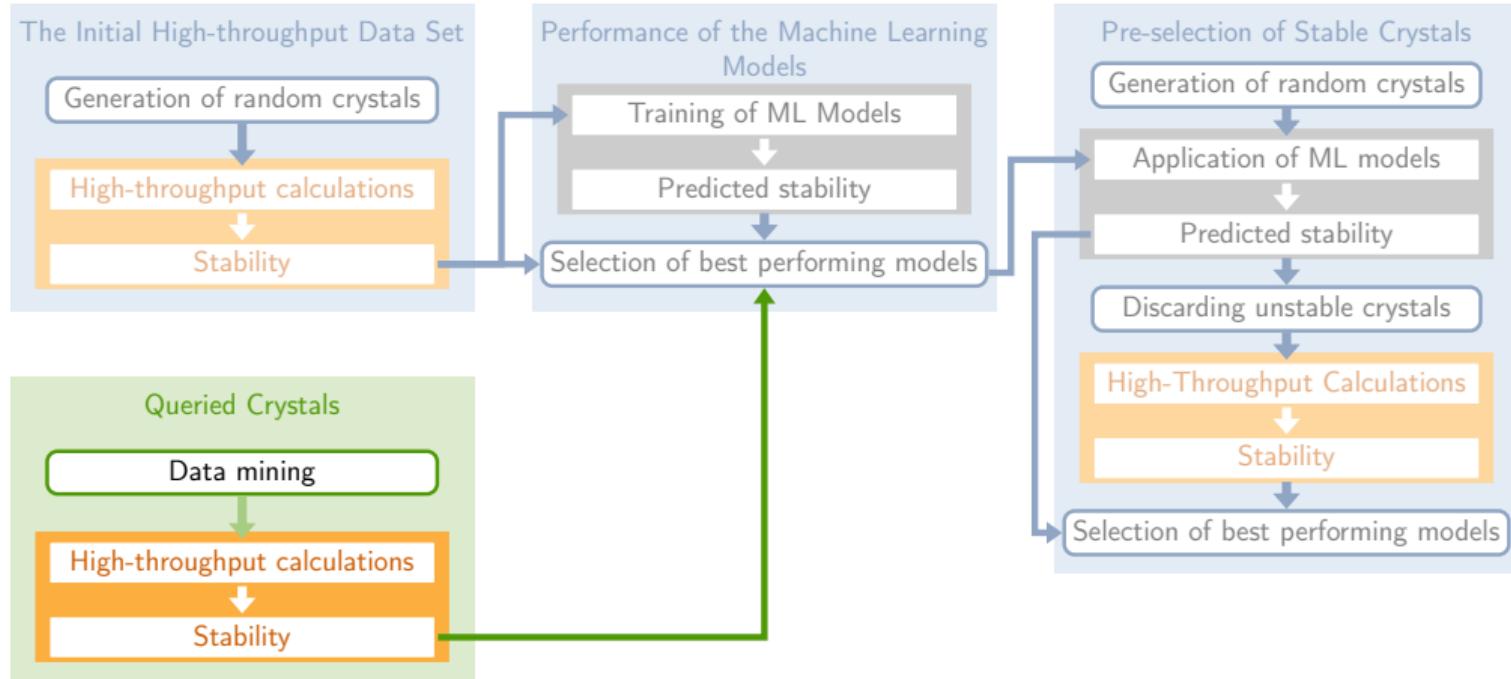
- ▶ 1073 crystal structures successfully converged.
- ▶ Formation energy:

$$E_{\text{form}}(\text{Cs}_x\text{Te}_{1-x}) = E(\text{Cs}_x\text{Te}_{1-x}) - [xE(\text{Cs}) + (1-x)E(\text{Te})]$$

- ▶ The stability (ΔE_{hull}) of each crystal is given by the distance to the convex hull.
- ▶ Average stability: 0.212 eV/atom.



Overview of the Crystal Structure Prediction Approach





Data Mining of Cs-Te Bulk Crystals



Materials project (MP)

Jain, Ong, et al., *APL Materials* 2013, 1, 011002; doi: 10.1063/1.4812323

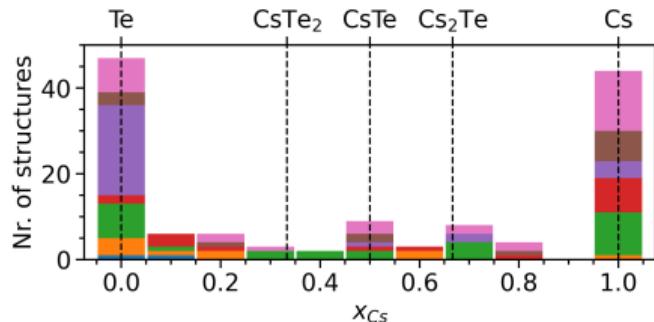


Open quantum materials database (oqmd)

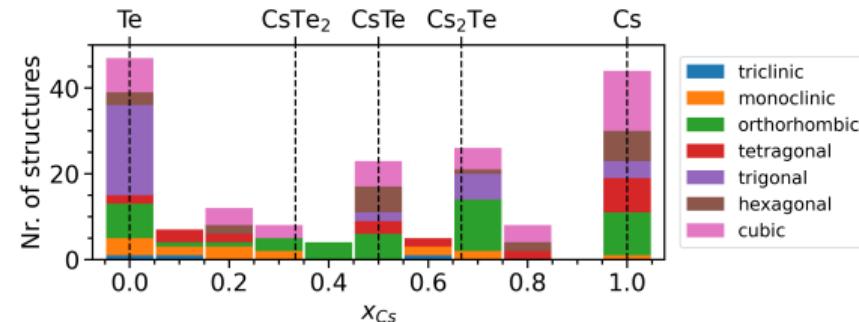
Saal, Kirklin, et al., *JOM* 2013, 65, 1501–1509; doi: 10.1007/s11837-013-0755-4

- ▶ Mined from online databases and enriched by substituting chemically similar elements.
- ▶ Distribution of the queried crystal structures:

Without ch. similar structures:



With ch. similar structures:

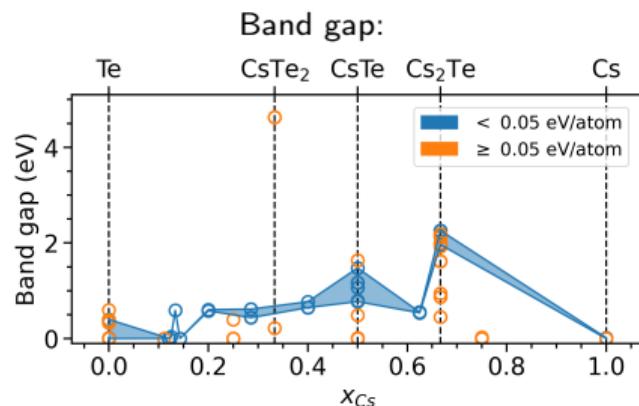
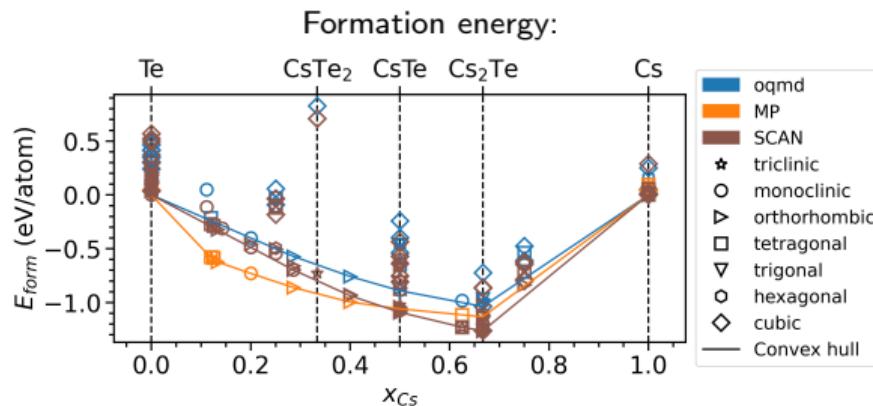


Distribution and Characteristics

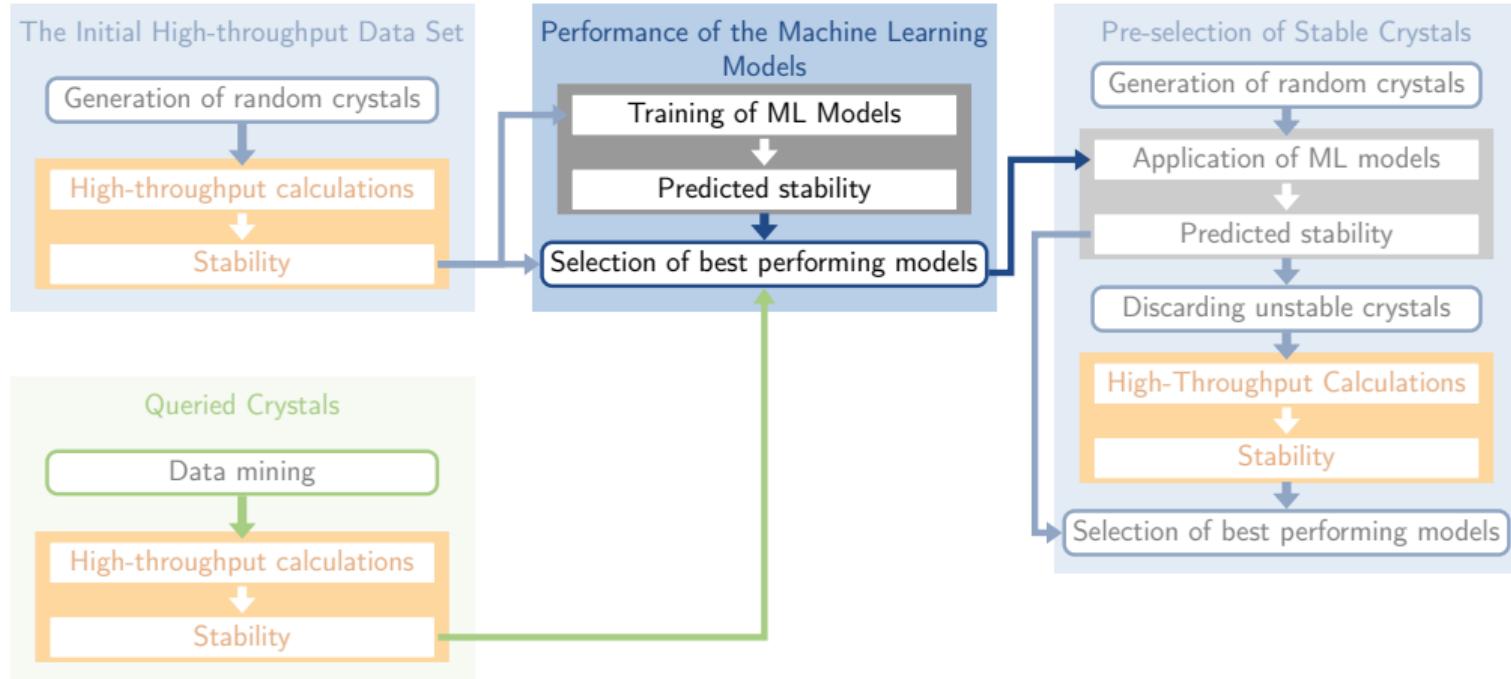
- ▶ Formation energy:

$$E_{\text{form}}(\text{Cs}_x\text{Te}_{1-x}) = E(\text{Cs}_x\text{Te}_{1-x}) - [xE(\text{Cs}) + (1-x)E(\text{Te})]$$

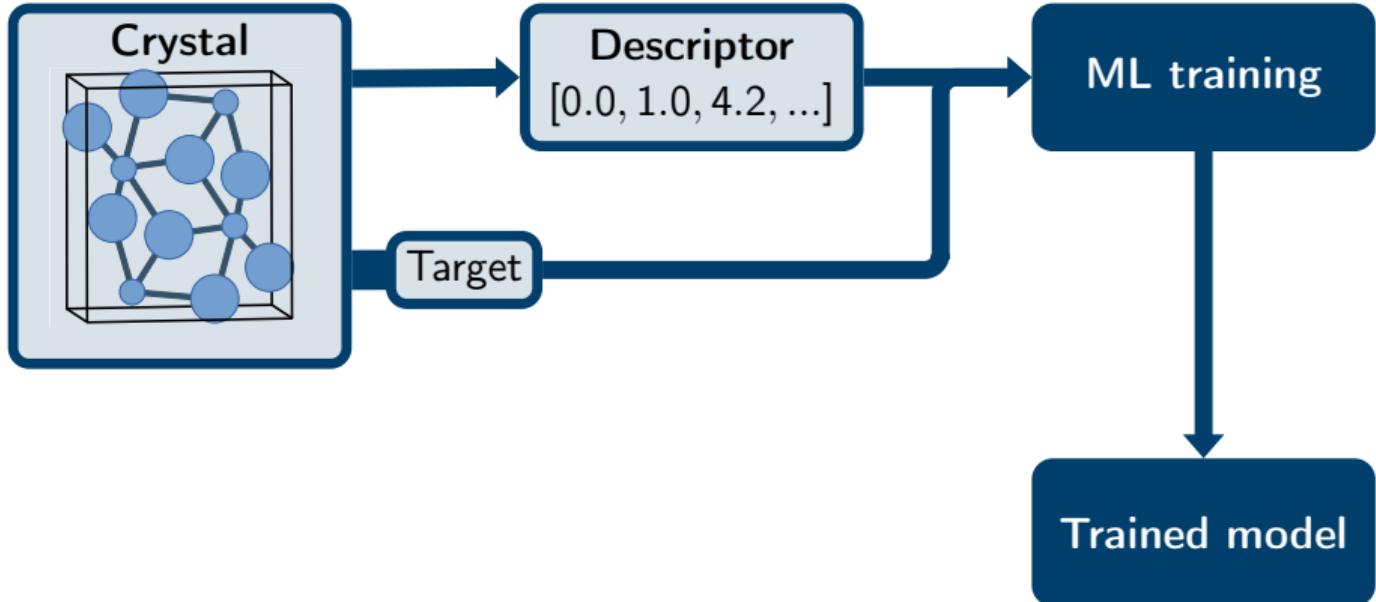
- ▶ The stability (ΔE_{hull}) of each crystal is given by the distance to the convex hull.
- ▶ Average stability: 0.136 eV/atom.
- ▶ The electronic band gap correlates with the composition for stable compounds ($\Delta E_{\text{hull}} < 0.05$ eV/atom).



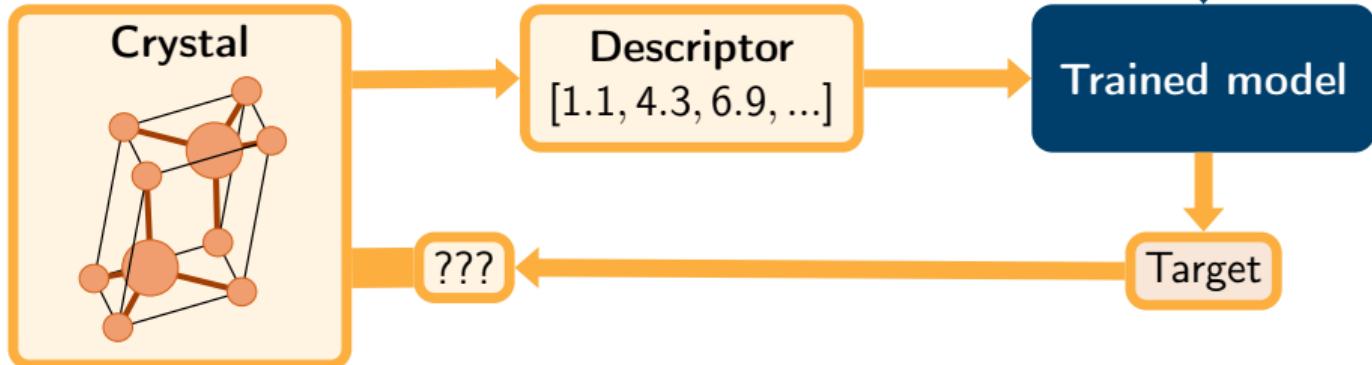
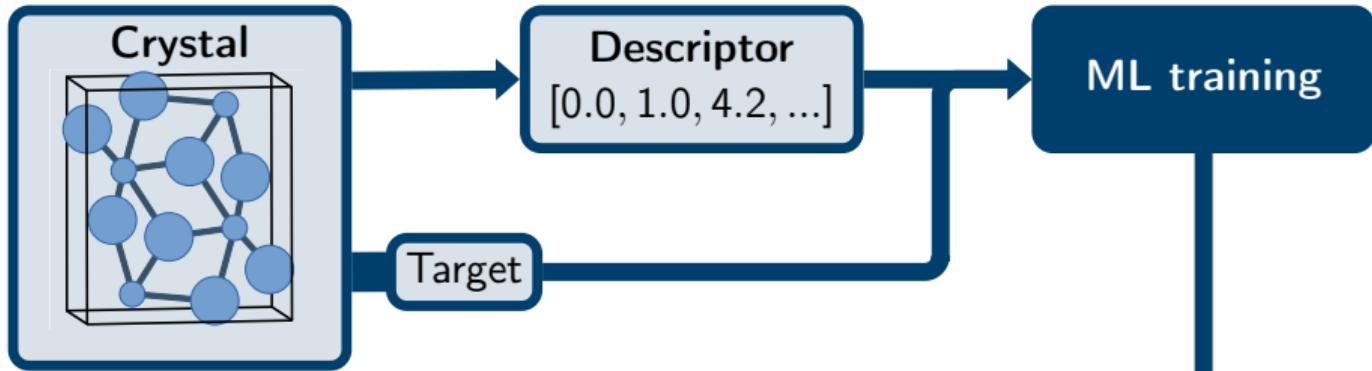
Overview of the Crystal Structure Prediction Approach



Machine Learning Structure-Property Relationships



Machine Learning Structure-Property Relationships





Used Descriptors and ML Algorithms



Used Descriptors and ML Algorithms

- ▶ *Coord.*: Statistical quantities of the coordination environment of each atomic site and composition.
 - ▶ *ET*: Extremely randomized trees
 - ▶ *KRR*: Kernel-ridge regression



Used Descriptors and ML Algorithms

- ▶ *Coord.*: Statistical quantities of the coordination environment of each atomic site and composition.
 - ▶ *ET*: Extremely randomized trees
 - ▶ *KRR*: Kernel-ridge regression
- ▶ *F-Fprint*: Averaged radial distribution function for each element-pair.
 - ▶ *KRR*: Kernel-ridge regression



Used Descriptors and ML Algorithms

- ▶ *Coord.*: Statistical quantities of the coordination environment of each atomic site and composition.
 - ▶ *ET*: Extremely randomized trees
 - ▶ *KRR*: Kernel-ridge regression
- ▶ *F-Fprint*: Averaged radial distribution function for each element-pair.
 - ▶ *KRR*: Kernel-ridge regression
- ▶ *SOAP*: Power-spectrum of the overlap matrix of a set of atom-centered Gaussian functions.
 - ▶ *KRR*: Kernel-ridge regression

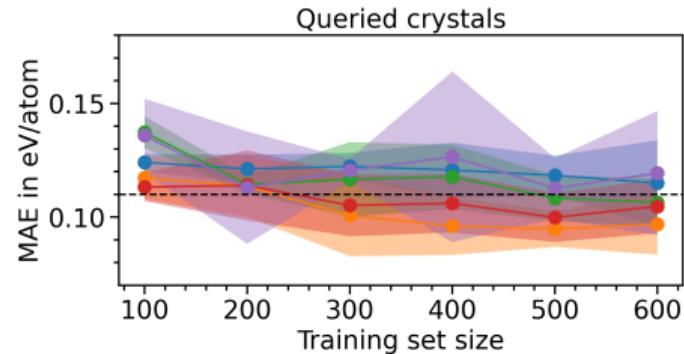
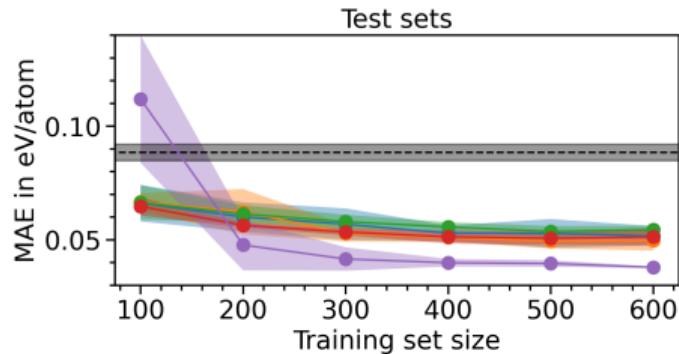
Used Descriptors and ML Algorithms

- ▶ *Coord.*: Statistical quantities of the coordination environment of each atomic site and composition.
 - ▶ *ET*: Extremely randomized trees
 - ▶ *KRR*: Kernel-ridge regression
- ▶ *F-Fprint*: Averaged radial distribution function for each element-pair.
 - ▶ *KRR*: Kernel-ridge regression
- ▶ *SOAP*: Power-spectrum of the overlap matrix of a set of atom-centered Gaussian functions.
 - ▶ *KRR*: Kernel-ridge regression
- ▶ *M3GNet*: Graph representation with atomic sites as nodes and chemical bonds as edges.
 - ▶ Graph neural network

Performance of the Machine Learning Models

Legend:

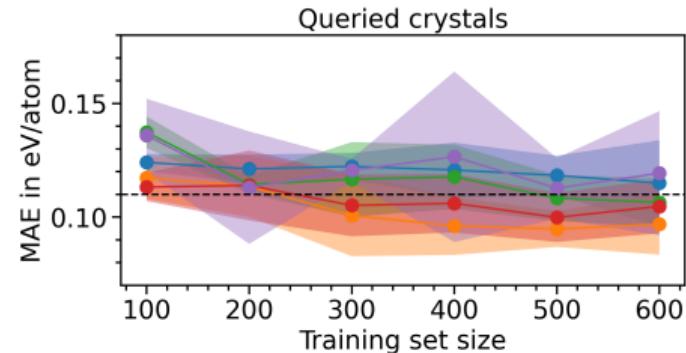
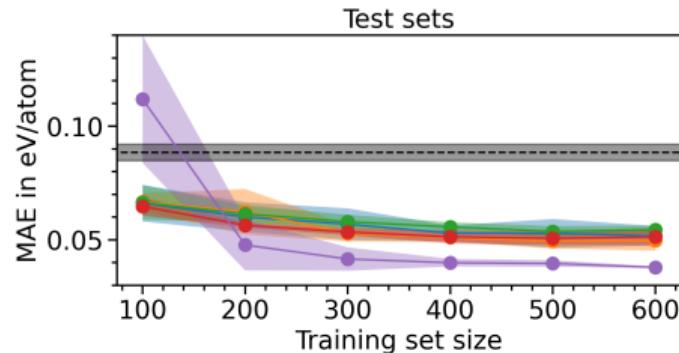
- Coord. ET
- Coord. KRR
- F-Fprint KRR
- SOAP KRR
- M3GNet TF
- M3GNet base



Performance of the Machine Learning Models

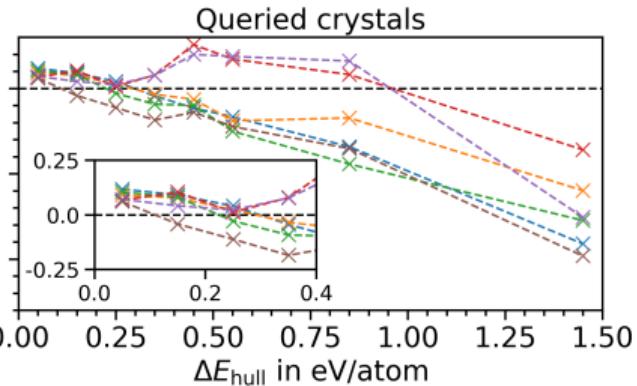
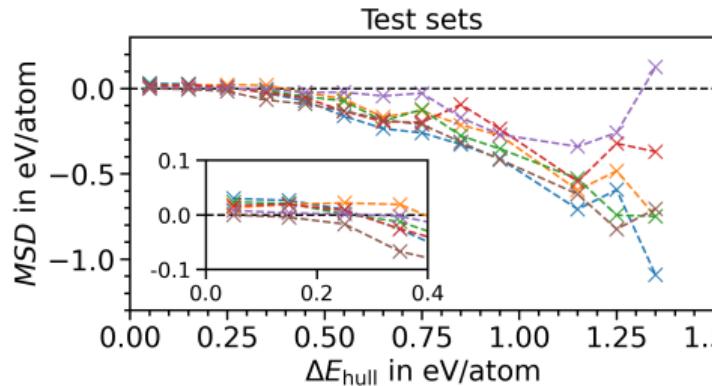
Legend:

- Coord. ET
- Coord. KRR
- F-Fprint KRR
- SOAP KRR
- M3GNet TF
- M3GNet base

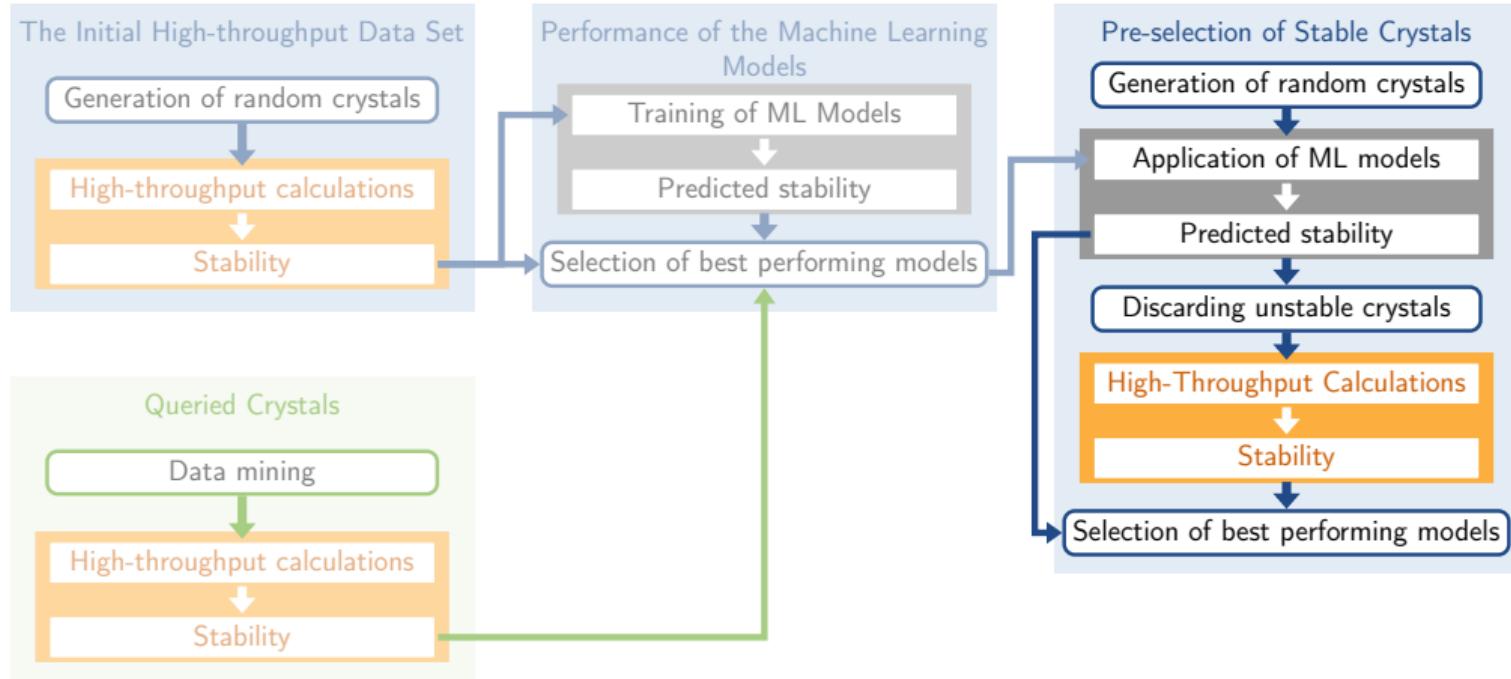


Legend:

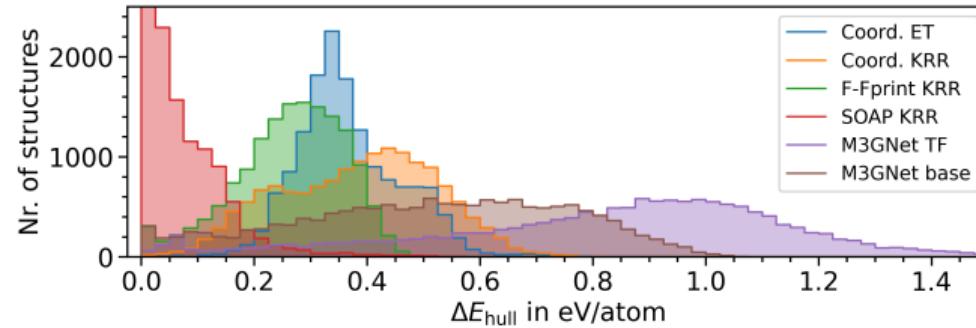
- Coord. ET
- Coord. KRR
- F-Fprint KRR
- SOAP KRR
- M3GNet TF
- M3GNet base



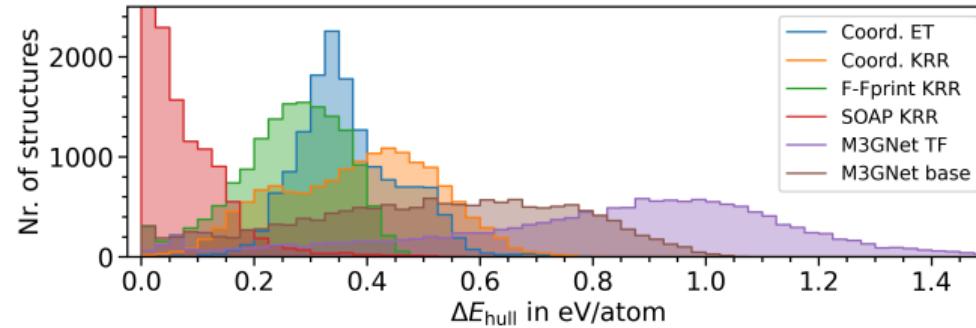
Overview of the Crystal Structure Prediction Approach



Pre-selection of Stable Crystals

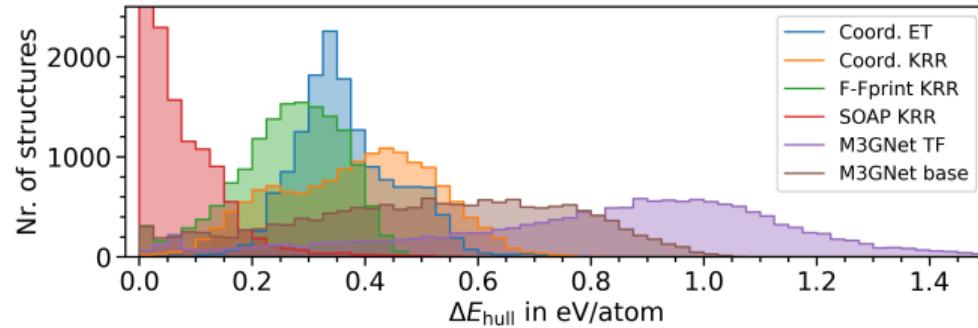


Pre-selection of Stable Crystals



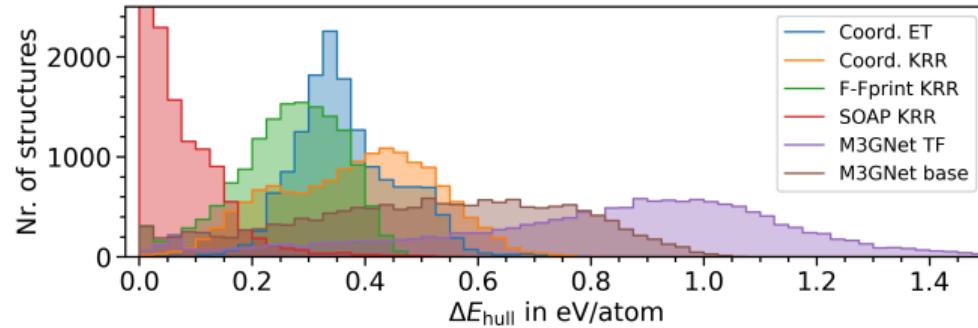
	ΔE_{Hull} in eV/at.		MAE in eV/at.	S.R. in %	Stable phases in %
	Mean	SD			
<i>Initial pool</i>	0.212	0.177	—	65.8	0.0/2.1/1.3
Coord. ET	0.178	0.109	0.081	90.0	0.0/6.0/5.0
Coord. KRR	0.164	0.098	0.101	80.0	0.0/5.0/5.0
F-Fprint KRR	0.378	0.308	0.344	34.0	0.0/0.0/0.0
SOAP KRR	0.158	0.124	0.157	63.0	0.0/0.0/3.0
M3GNet TF	0.195	0.106	0.110	90.0	1.0/5.0/1.0
M3GNet base	0.193	0.126	0.156	77.0	0.0/4.0/3.0

Pre-selection of Stable Crystals



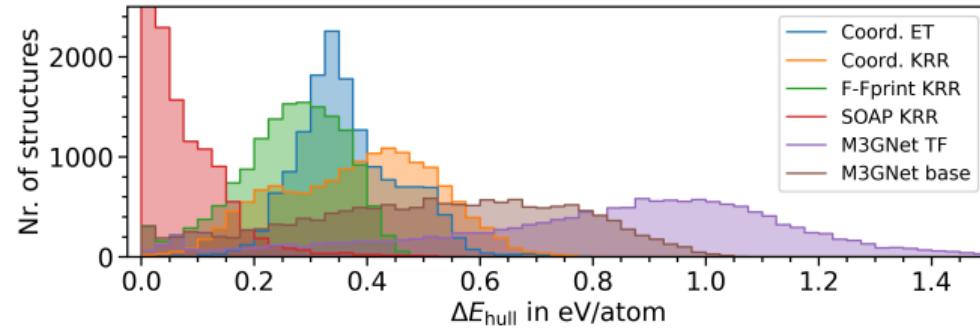
	ΔE_{Hull} in eV/at.		MAE in eV/at.	S.R. in %	Stable phases in %
	Mean	SD			
<i>Initial pool</i>	0.212	0.177	–	65.8	0.0/2.1/1.3
Coord. ET	0.178	0.109	0.081	90.0	0.0/6.0/5.0
Coord. KRR	0.164	0.098	0.101	80.0	0.0/5.0/5.0
F-Fprint KRR	0.378	0.308	0.344	34.0	0.0/0.0/0.0
SOAP KRR	0.158	0.124	0.157	63.0	0.0/0.0/3.0
M3GNet TF	0.195	0.106	0.110	90.0	1.0/5.0/1.0
M3GNet base	0.193	0.126	0.156	77.0	0.0/4.0/3.0

Pre-selection of Stable Crystals



	ΔE_{Hull} in eV/at.		MAE in eV/at.	S.R. in %	Stable phases in %
	Mean	SD			
<i>Initial pool</i>	0.212	0.177	—	65.8	0.0/2.1/1.3
Coord. ET	0.178	0.109	0.081	90.0	0.0/6.0/5.0
Coord. KRR	0.164	0.098	0.101	80.0	0.0/5.0/5.0
F-Fprint KRR	0.378	0.308	0.344	34.0	0.0/0.0/0.0
SOAP KRR	0.158	0.124	0.157	63.0	0.0/0.0/3.0
M3GNet TF	0.195	0.106	0.110	90.0	1.0/5.0/1.0
M3GNet base	0.193	0.126	0.156	77.0	0.0/4.0/3.0

Pre-selection of Stable Crystals



	ΔE_{Hull} in eV/at.		MAE in eV/at.	S.R. in %	Stable phases in %
	Mean	SD			
<i>Initial pool</i>	0.212	0.177	—	65.8	0.0/2.1/1.3
Coord. ET	0.178	0.109	0.081	90.0	0.0/6.0/5.0
Coord. KRR	0.164	0.098	0.101	80.0	0.0/5.0/5.0
F-Fprint KRR	0.378	0.308	0.344	34.0	0.0/0.0/0.0
SOAP KRR	0.158	0.124	0.157	63.0	0.0/0.0/3.0
M3GNet TF	0.195	0.106	0.110	90.0	1.0/5.0/1.0
M3GNet base	0.193	0.126	0.156	77.0	0.0/4.0/3.0

Conclusions

- ▶ The presented approach offers a straight-forward method to explore unknown materials.
 - ▶ The pre-selection step greatly improves over a random structure search.
 - ▶ The *Coord. ET* and the *M3GNet* models show the best performance.
- ▶ The implementation can be readily applied to any other material-system, e.g. ternary antimonides.
- ▶ The source code is published and freely available (<https://aim2dat.github.io/aim2dat/>).

Experimental Collaborators



Prof. Dr. Thorsten Kamps,
and team.



Bundesministerium
für Bildung
und Forschung



Niedersächsisches Ministerium
für Wissenschaft und Kultur

Sponsors



Project-id: 490940284



Project-id: nic00069,
nic00076