

機械学習 × 第一原理計算で結晶構造を予測する

劉 暢

マテリアルズインフォマティクス研究推進センター 特任助教

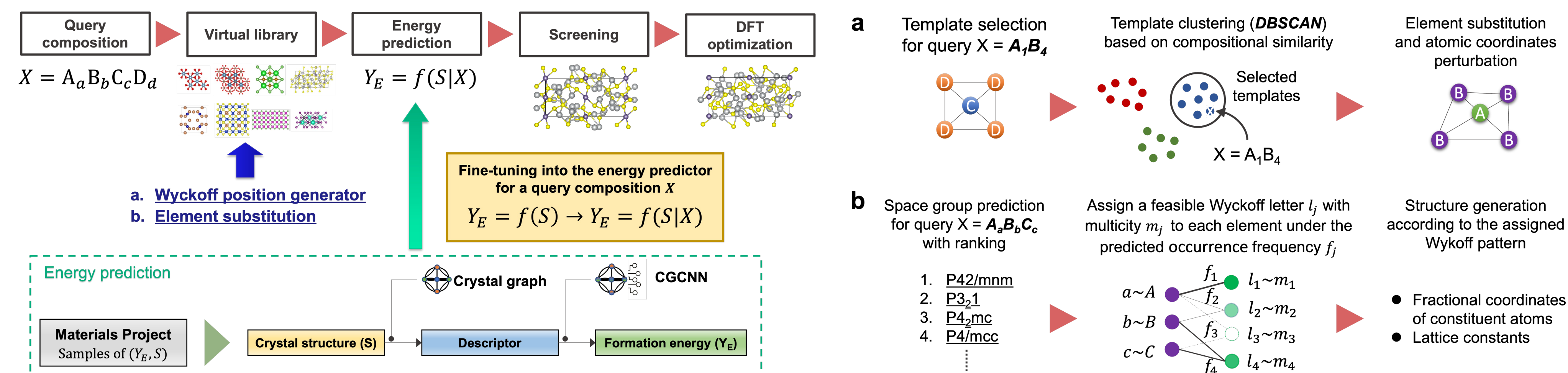
Introduction

結晶構造予測の目標は、標的化合物の組成式 X のみから、その化合物の安定に存在する結晶構造を提案することである。材料の物性の多くはその構造から由来するため、機械学習による結晶構造予測の効率・高速化は、産業界から強く期待されている。

現段階では以下の成果を得ている。

1. C. Liu *et al.*, "Shotgun crystal structure prediction using machine-learned formation energies," (submitted), DOI: [10.48550/arXiv.2305.02158](https://doi.org/10.48550/arXiv.2305.02158)
2. M. Kusaba *et al.*, "Crystal structure prediction with machine learning-based element substitution," *Comput. Mater. Sci.*, DOI: [10.1016/j.commatsci.2022.111496](https://doi.org/10.1016/j.commatsci.2022.111496)
3. Code: https://github.com/yoshida-lab/XenonPy/blob/master/samples/CSP_with_element_substitution.ipynb

Workflow



転移学習 (一点計算+CGCNN)を導入するにより反復計算は要らなくなったため、構造最適化にかかる時間が大幅に削減!

Results

Dataset I. 40 benchmark crystals selected based on the diversity of space groups, constituent elements, number of atoms, and element species; and the diversity of applications such as battery and thermoelectric materials.

Composition	Number of atoms	Space group	Wyckoff position generation	Element substitution
C	4	R3m	✓(✓/✓)	✓(✓)
Si	2	Fd3m	✓(✓/✓)	✓(✓)
GaAs	2	F43m	✓(✓/✓)	✓(✓)
ZnO	4	P6 ₃ mc	✓(✓/✓)	✓(✓)
BN	4	P6 ₃ /mmc	✓(✓/✓)	✓(✓)
LiCoO ₂	16	R3m	✓(✓/✓)	✓(✓)
Bi ₂ Te ₃	5	R3m	✓(✓/✓)	✓(✓)
Ba(FeAs) ₂	5	I4/mmm	✓(✓/✓)	✓(✓)
SiO ₂	6	I4 ₂ d	✓(✓/✓)	×(✓)
VO ₂	6	P4 ₂ /mnm	✓(✓/✓)	✓(✓)
La ₂ CuO ₄	7	I4/mmm	×(✓/✓)	✓(✓)
LiPF ₆	8	R3	✓(✓/✓)	✓(✓)
Al ₂ O ₃	10	R3c	✓(✓/✓)	✓(✓)
SrTiO ₃	10	I4/mcm	✓(✓/✓)	✓(✓)
CaCO ₃	10	R3c	✓(✓/✓)	✓(✓)
TiO ₂	12	C2/m	×(×/✓)	×(✓)
ZrO ₂	12	P2 ₁ /c	✓(✓/✓)	✓(✓)
ZrTe ₅	12	Cmcm	✓(✓/✓)	✓(✓)
V ₂ O ₅	14	Pnmm	✓(✓/✓)	×(×)
Si ₃ N ₄	14	P6 ₃ /m	✓(✓/✓)	✓(✓)
Fe ₂ O ₃	14	Fd3m	✓(✓/✓)	✓(✓)
Mn(FeO ₂) ₂	14	Fd3m	✓(✓/✓)	✓(✓)
ZnSb	16	Pbca	✓(✓/✓)	✓(✓)
CoSb ₃	16	Im3	✓(✓/✓)	✓(✓)
LiBF ₄	18	P3 ₂ 1	✓(✓/✓)	✓(✓)
Y ₂ Co ₁₇	19	R3m	✓(✓/✓)	✓(✓)
GeH ₄	20	P2 ₁ /2 ₁	✓(✓/✓)	✓(✓)
CsPb ₃	20	Pnma	×(×/✓)	✓(✓)
NaCaAlPO ₃ F ₂	24	P2 ₁ /m	×(×/✓)	✓(✓)
LiFePO ₄	28	Pnma	✓(✓/✓)	✓(✓)
Cu ₁₂ Sb ₂ S ₁₃	29	I43m	✓(✓/✓)	✓(✓)
MgB ₂	32	Imma	×(✓/✓)	×(×)
Li ₃ PS ₄	32	Pnma	✓(✓/✓)	×(×)
Cd ₃ As ₂	80	I4 ₁ /acd	✓(✓/✓)	✓(✓)
Li ₄ Ti ₃ O ₁₂	42	C2/c	✓(✓/✓)	✓(✓)
Ba ₃ CaSi ₄ (BO ₃) ₂	46	I4 ₂ m	×(×/✓)	×(×)
Ag ₃ Ge ₅	60	Pna2 ₁	×(✓/✓)	✓(✓)
Nd ₄ Fe ₄ B	68	P4 ₂ /mnm	×(×/✓)	✓(✓)
Y ₃ Al ₅ O ₁₂	80	Ia3d	✓(✓/✓)	✓(✓)
Ca ₁₁ MnSb ₁₁	104	I4 ₁ /acd	×(✓/✓)	✓(✓)
Overall			31/40 = 77.5%	33/40 = 82.5%

Element substitution (top 10): (33 + 43) / 90 = 84.4%

Wyckoff generator (top 10): (31 + 34) / 90 = 72.2%

space group prediction (top 30): (40 + 45) / 90 = 94.4%

For the space group successfully predicted cases:

➢ Fail to predict structures: 9 + 11 = 20

➢ Fail to generate the true Wyckoff labels: 3 + 3 = 6

Dataset II. 50 benchmark crystals randomly selected from the Materials Project database.

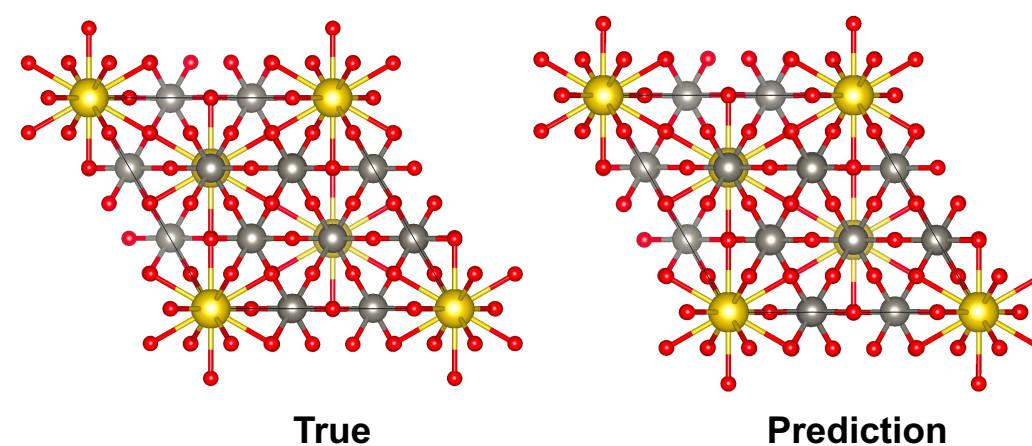
Composition	Number of atoms	Space group	Wyckoff position generation	Element substitution
CsCl	2	Fm3m	✓(✓/✓)	✓(✓)
MnAl	2	P4/mnm	✓(✓/✓)	✓(✓)
HoHfSe	3	P6m2	✓(✓/✓)	✓(✓)
ErCdRh ₂	4	Fm3m	✓(✓/✓)	✓(✓)
Eu ₂ MgTi	4	Fm3m	✓(✓/✓)	✓(✓)
Pm ₂ NiIr	4	Fm3m	✓(✓/✓)	✓(✓)
VPt ₅	4	I4/mmm	✓(✓/✓)	✓(✓)
Gd(SiO ₃) ₂	5	I4/mmm	✓(✓/✓)	✓(✓)
LaAl ₃ Au	5	I4mm	✓(✓/✓)	✓(✓)
U ₂ SbN ₂	5	I4/mmm	✓(✓/✓)	✓(✓)
MnGa(CuSe) ₂	8	I4	✓(✓/✓)	✓(✓)
SmZnPd	9	P62m	✓(✓/✓)	✓(✓)
Sn(TePd) ₂	9	I4mm	×(×/✓)	✓(✓)
V ₃ S ₂	9	I4/m	✓(✓/✓)	✓(✓)
Cs ₃ InF ₆	10	Fm3m	✓(✓/✓)	✓(✓)
Eu(CuSb) ₂	10	P4/mmm	✓(✓/✓)	✓(✓)
Rb ₂ TlAgCl ₂	10	Fm3m	✓(✓/✓)	✓(✓)
Ca ₃ Ni ₂ B ₂	12	R3m	✓(✓/✓)	✓(✓)
DyPb ₄	12	I4 ₁ /amd	✓(✓/✓)	✓(✓)
LaSiIr	12	P2 ₁ /3	✓(✓/✓)	✓(✓)
SmVO ₄	12	I4 ₁ /amd	✓(✓/✓)	✓(✓)
VCl ₃	12	P1	×(✓/✓)	✓(✓)
YbP ₃	12	P2 ₁ /m	×(✓/✓)	✓(✓)
Eu(Al ₂ Cu) ₂	13	I4/mmm	✓(✓/✓)	✓(✓)
Zr ₂ O	15	R3	×(×/✓)	×(×)
Ba ₃ Ta ₂ NiO ₈	15	P3m1	✓(✓/✓)	×(×)
K ₂ Ni ₃ S ₂	18	Fddd	✓(✓/✓)	✓(✓)
Sr(ClO ₄) ₂	18	Fddd	✓(✓/✓)	✓(✓)
LiSm ₃ IrO ₄	20	P2 ₁ /c	×(×/✓)	✓(✓)
Pr ₂ ZnPtO ₄	20	P2 ₁ /c	×(×/✓)	✓(✓)
Sc ₂ Mn ₁₂ P ₇	21	P6	✓(✓/✓)	✓(✓)
LaSi ₃ Ni ₄	24	I4 ₁ /amd	✓(✓/✓)	✓(✓)
CeCu ₂ Sn	28	Pnma	✓(✓/✓)	✓(✓)
LiP(HO ₂) ₂	32	Pna2 ₁	×(✓/✓)	✓(×)
Mg ₃ Si ₂ H ₂ O ₃	36	P6 ₃ cm	×(×/✓)	×(×)
Y ₃ Si ₃ Ir ₃	36	P6 ₃ /mmc	✓(✓/✓)	✓(✓)
Na(WO ₃) ₃	37	R3	✓(✓/✓)	✓(✓)
Sm ₂ Ni ₃₀ As ₁₃	39	P6	✓(✓/✓)	✓(✓)
BaCaGaF ₇	40	P2 ₁ /c	×(×/✓)	✓(✓)
Tm ₁₁ Sn ₁₀	42	I4/mmm	✓(✓/✓)	✓(✓)
AlH ₁₂ (ClO ₄) ₂	44	R3c	×(×/✓)	✓(✓)
K ₂ ZrSi ₂ O ₇	48	P2 ₁ /c	✓(✓/✓)	✓(×)
LiZr ₂ (PO ₄) ₃	72	P2 ₁ /c	×(×/✓)	✓(✓)
K ₅ Ag ₂ (AsSe ₃) ₃	76	Pnma	✓(✓/✓)	✓(✓)
Be ₇ Ru ₃	80	Im3	×(×/✓)	✓(✓)
Cu ₃ P ₄ (S ₂ Cl) ₃	80	Pnma	×(×/✓)	✓(✓)
Al ₃ CoO ₄	84	P3m1	×(×/✓)	✓(✓)
Li ₃ V ₃ P ₃ O ₂₀	92	P1	×(✓/✓)	✓(×)
ReBi ₃ O ₈	96	P2 ₁ /3	✓(✓/✓)	✓(✓)
Na ₄ Fe ₂ (O ₄ F) ₂	288	Pbca	×(×/✓)	✓(✓)
Overall			34/50 = 68.0%	43/50 = 86.0%

a Na(WO₃)₃ - R3 (No. 148) ※抜粋

Full formula: Na₃W₂₇O₈₁

Wyckoff letter pattern:

• Na - a W - b, c, f O - e, f, f, f



Wyckoff letters of No. 148

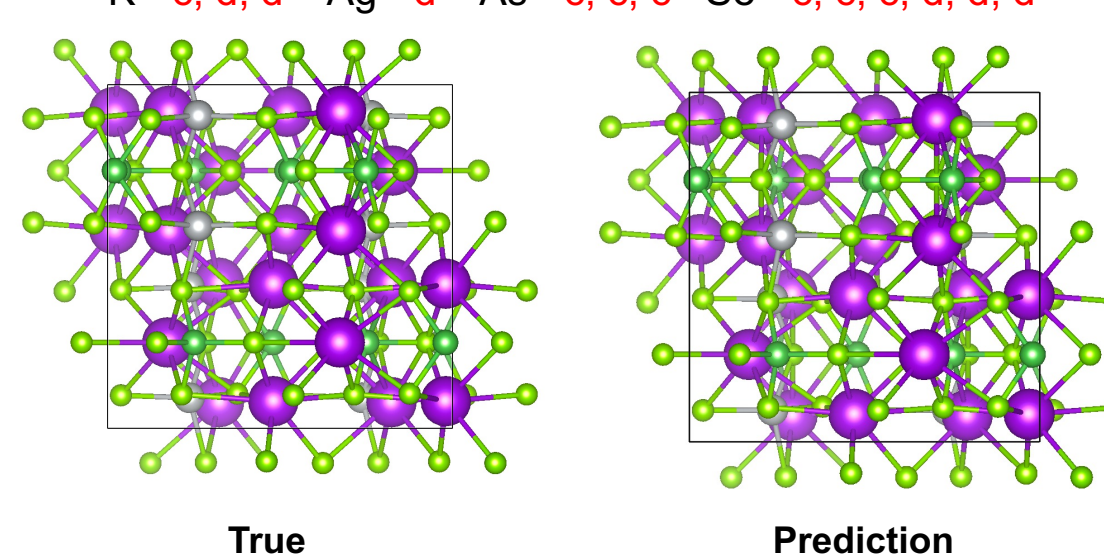
Wyckoff letter	Multiplicity	Coordinates
f	18	(x,y,z), ...
e	9	(1/2,0,0), ...
d	9	(1/2,0,1/2), ...
c	6	(0,0,z), ...
b	3	(0,0,1/2), ...
a	3	(0,0,0), ...

b K₅Ag₂(AsSe₃)₃ - Pnma (No. 62)

Full formula: K₂₀Ag₈As₁₂Se₃₆

Wyckoff letter pattern:

• K - c, d, d Ag - d As - c, c, c Se - c, c, c, d, d, d



Wyckoff letters of No. 62

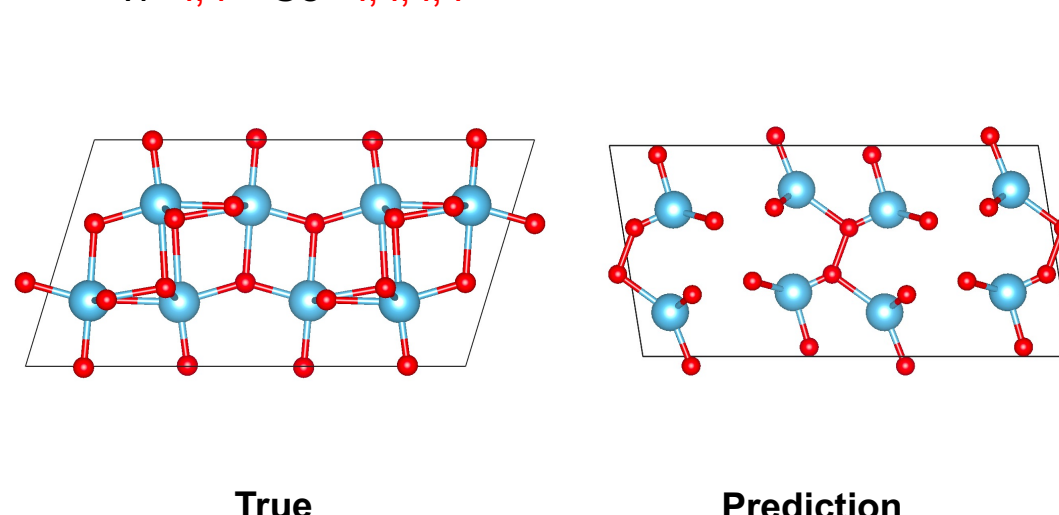
Wyckoff letter	Multiplicity	Coordinates
d	8	(x,y,z), ...
c	4	(x,1/4,z), ...
b	4	(0,0,1/2), ...
a	4	(0,0,0), ...

c TiO₂ - C2/m (No. 12)

Full formula: Ti₈O₁₆

Wyckoff letter pattern:

• Ti - i, i, i Se - i, i, i, i



Wyckoff letters of No. 12

Wyckoff letter	Multiplicity	Coordinates
j	8	(x,y,z), ...
i	4	(x,0,z), ...
h	4	(0,y,1/2), ...
g	4	(0,y,0), ...
f	4	(1/4,1/4,1/2), ...
e	4	(1/4,1/4,0), ...
d	2	(0,1/2,1/2), ...
c	2	(0,0,1/2), ...
b	2	(0,1/2,0), ...
a	2	(0,0,0), ...

高精度・高性能を同時に実現!