Supplementary Materials

CGWGAN: crystal generative framework based on Wyckoff generative adversarial network

Tianhao Su[1,#], Bin Cao[2,#], Shunbo Hu[1], Musen Li[1], Tong-Yi Zhang[1,2,*]

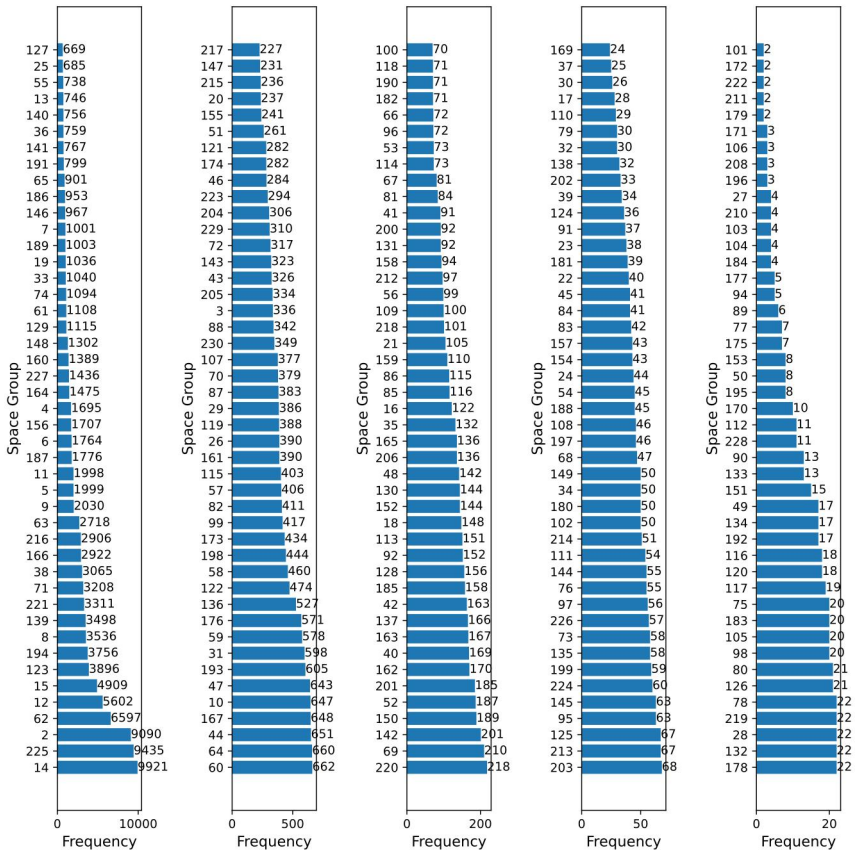
[1]Materials Genome Institute, Shanghai University, Shanghai 200444, China.
[2]Guangzhou Municipal Key Laboratory of Materials Informatics, Advanced Materials Thrust, Sustainable Energy and Environment Thrust, Hong Kong University of Science and Technology (Guangzhou), Guangzhou 511400, Guangdong, China.
[#]Authors contributed equally.

[*]**Correspondence to:** Prof. Tong-Yi Zhang, Materials Genome Institute, Shanghai University, 333 Nanchen Road, Shanghai 200444, China. E-mail: zhangty@shu.edu.cn

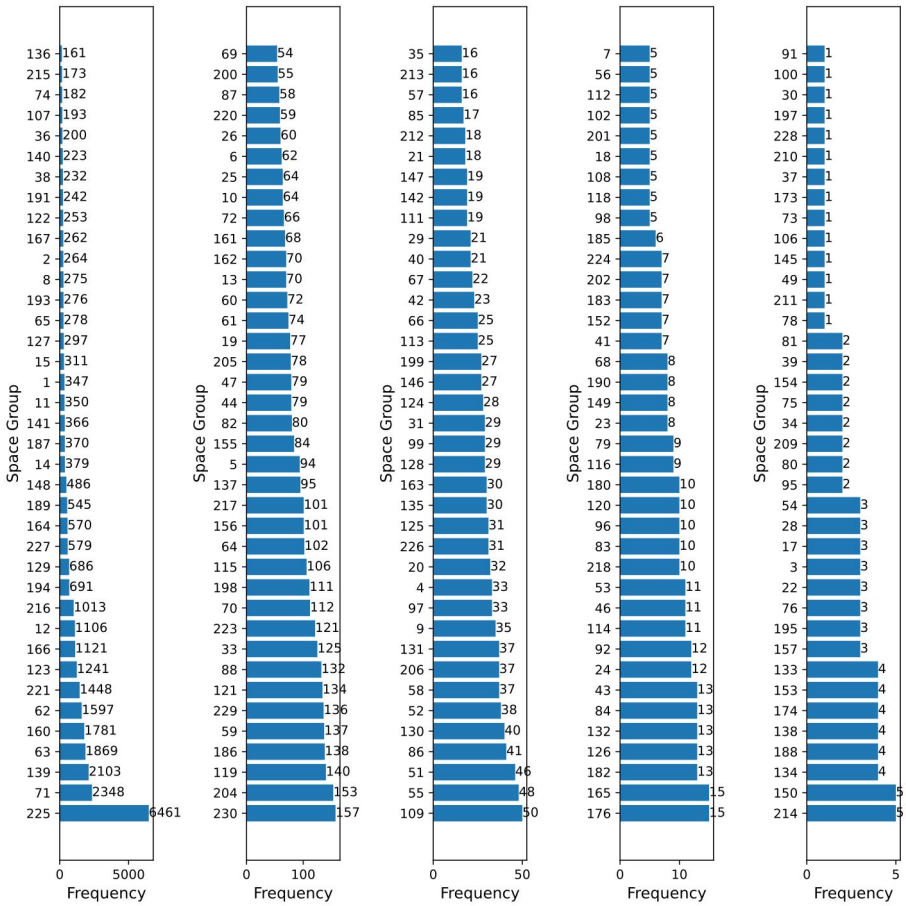
The supplementary figures

Materials Project Dataset:

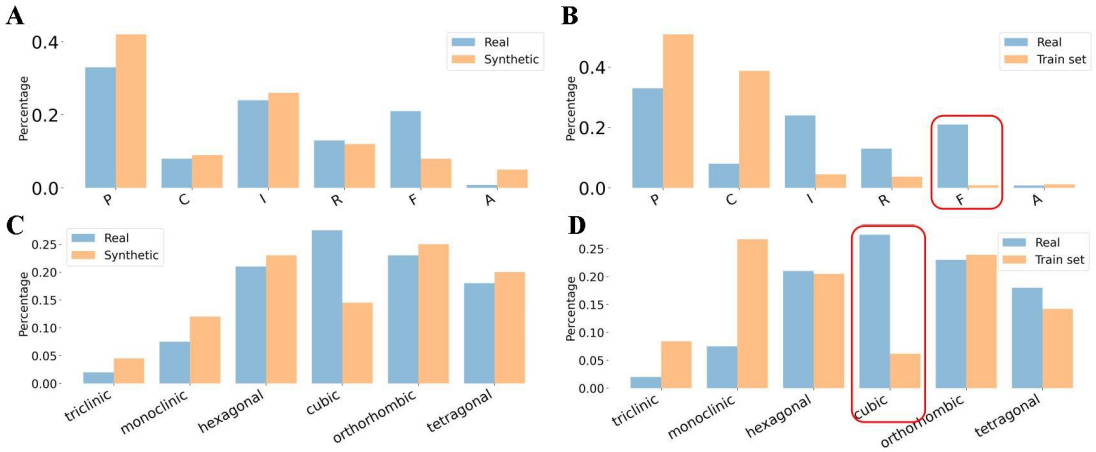
The crystal structures were retrieved from the Materials Project (MP) database, and a statistical analysis is provided here.



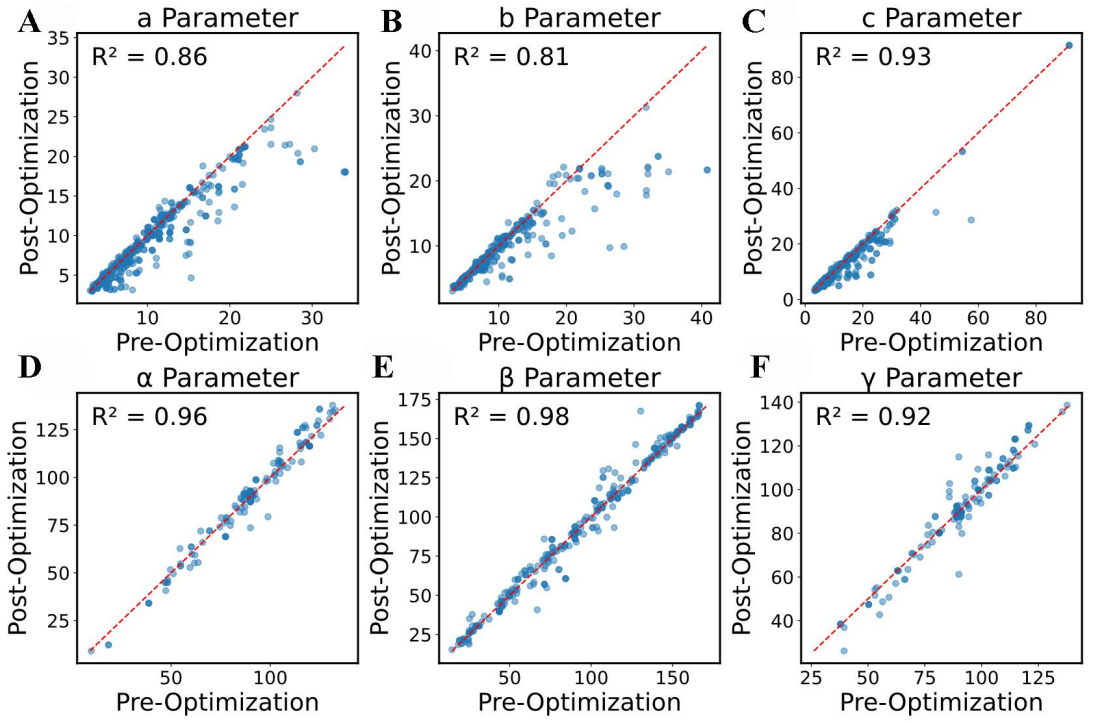
Supplementary Figure 1. The distribution of crystals in the MP database across space groups.



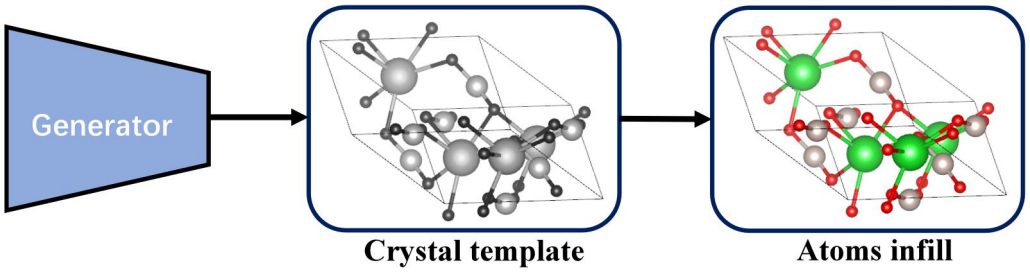
Supplementary Figure 2. The distribution of crystals in the MP database with 3-4 Wyckoff sites across space groups.



Supplementary Figure 3. (A) The six Bravais lattice styles of all MP structure (Real) for unit cell, generated samples in and (B) training set from ASU representation. (C) The crystal family, where the samples classified as Trigonal have been reassigned to Hexagonal of all in MP structure (Real) for unit cell, generated samples in and (D) training set from ASU representation.



Supplementary Figure 4. Statistics of lattice parameters for the generated structures, after filtering by M3GNet and subsequent DFT calculations of structural relaxation, (A) length parameter a, (B) length parameter b, (C) length parameter c, (D) angle parameter α, (E) angle parameter β, and (F) angle parameter γ.



Supplementary Figure 5. The workflow of crystals generation in CGWGAN.

After obtaining the template from the generator, the elements are assigned to the positions of the template in the form of permutations and combinations, and M3GNet is used to perform energy sorting and select the combination with the lowest energy to determine the atomic filling.

The phonon spectrum of the system was calculated using the widely employed first-principles-based phonon package, Phonopy. The vibrational frequencies were computed along all paths in the entire Brillouin zone. These paths were automatically supplemented by crystal symmetry through Seekpath. If the minimum phonon frequency is greater than -0.5, the system is considered to have passed the phonon stability criterion. M3GNet-calculated phonon results are available at Hugging Face and can be independently verified using the included program run_all.py. Structures subjected to rapid filtering can later undergo high-precision optimization using VASP.

The deformation ratio of lattice parameters reflects how closely the generated candidate materials approach the local minima on the potential energy surface. The proximity of these materials to saddle points or local optima helps in predicting reaction pathways and energy barriers under specific conditions, facilitating structural relaxation to converge with minimal ionic steps.

Supplementary Figure 4 shows that the structures screened and those further relaxed via DFT exhibit high correlation, with coefficients of determination R^2 of 0.86, 0.81, and 0.93 for lattice parameters a, b, and c, respectively. This indicates that the generated structures predominantly occupy low-energy sites on the potential energy surface. The interatomic distances of the candidate crystals remain reasonable after expansion through the ASU, with minimal overall expansion or compression during relaxation. Similarly, the coefficients of determination for lattice parameters α, β, and

γ are 0.96, 0.98, and 0.92, respectively, suggesting that the crystals generated by CGWGAN closely approximate physically realistic structures. This consistency ensures that the initially screened crystals do not exhibit significant structural mismatches after DFT optimization, confirming the accuracy and reliability of the material design. The high correlation coefficients reflect CGWGAN's effectiveness in capturing and retaining crucial crystal properties, which reduces the computational burden in subsequent optimization processes and indicates the inherent rationality of the generated structures.

Execution details of the convergence rate test

PyXtal requires pre-specification of elemental compositions and atomic ratios after symmetry expansion, where a priori valence combinations lend plausibility to the generated crystal structures. For benchmarking the model's convergence rate, pre-generated crystal templates were extracted, maintaining the same total number of atoms in the unit cell (ranging from 2 to 32), with identical elemental systems. The final chemical compositions of our model were entirely determined by the atomic ratios after template infill, which could lead to new compositional ratios.

The PyXtal test system included 14,400 crystals in total, with 6,941 crystals tested through template substitution. The dataset size was determined by the criterion that the overall fluctuation in the convergence rate remained below 3%, which was achieved for each PyXtal test system. The template as a whole also adhered to this standard (For convergence requirements, see the VASP parameter configuration in the SM). Structures that passed the convergence threshold were further filtered using ASE to ensure that the force on any individual atom was less than 0.01 eV/Å.

The supplementary tables

Supplementary Table 1. Algorithm: Res 1D Self-Attention Layer

Algorithm: Res 1D Self-Attention Layer
Input:
- Input tensor x with dimensions (batch_size, channels, length)
Parameters:
- in_dim: Input dimension of the tensor
- query_conv, key_conv, value_conv: Convolution layers
- gamma: Learnable scale parameter
- softmax: Softmax function for attention
1. Initialize query_conv, key_conv, value_conv with 1x1 convolution layers
2. Initialize gamma as a learnable parameter
function FORWARD(x)
proj_query <- query_conv(x) reshaped to (batch_size, -1, length)
proj_key <- key_conv(x) reshaped to (batch_size, -1, length)
proj_value <- value_conv(x) reshaped to (batch_size, -1, length)
e <- batch matrix multiplication of proj_query (permuted) and proj_key
attention <- softmax(e) along the last dimension
out <- batch matrix multiplication of proj_value and attention (permuted)
out <- reshape out to the dimensions of x
out <- gamma * out + x

Supplementary Table 2. Fractional Coordinates of the Asymmetric Units Generated for Space Group No. 99 and Space Group No. 123 Structures.

	No. 99	**No. 123**
ASU Fractional Coordinates	[0.00, 0.00, 0.50] [0.25, 0.25, 0.25] [0.25, 0.25, 0.50] [0.00, 0.25, 0.25]	[0.00, 0.00, 0.00] [0.25, 0.25, 0.25] [0.25, 0.25, 0.00] [0.00, 0.25, 0.25]
group operation set	1 "x, y, z" 2 "-x, -y, z" 3 "-y, x, z" 4 "y, -x, z" 5 "x, -y, z" 6 "-x, y, z" 7 "-y, -x, z" 8 "y, x, z"	1 "x, y, z" 2 "-x, -y, z" 3 "-y, x, z" 4 "y, -x, z" 5 "-x, y, -z" 6 "x, -y, -z" 7 "y, x, -z" 8 "-y, -x, -z" 9 "-x, -y, -z" 10 "x, y, -z" 11 "y, -x, -z" 12 "-y, x, -z" 13 "x, -y, z" 14 "-x, y, z" 15 "-y, -x, z" 16 "y, x, z"

Supplementary Table 3. training parameters employed a WGAN architecture

Parameter Name	Value / Type
Generator Learning Rate	0.0001
Discriminator Learning Rate	0.0004
Cosine annealing period T	10
Minibatch size	64
Random Initialization	Xavier Initialization
Number of Epochs	10000
Discriminator Update Steps	5
Noise Vector Dimension	100

The Xavier initialization method is determined based on the number of input neurons n_{in} and output neurons n_{out} of a neural network layer. Specifically, in this work, the weights are randomly sampled from a normal distribution with a mean of 0 and a variance of $\frac{2}{n_{in}+n_{out}}$.

Supplementary Table 4. The output of AFLOW Xtalfinder from the inputting of the 7 crystals.

Input	Output
Space group No. 1 crystal.	formatting issues
Space group No. 25 crystal.	Prototype does not match
Space group No. 35 crystal.	Prototype does not match
Space group No. 71 crystal.	Prototype does not match
Space group No. 99 crystal.	Prototype does not match
Space group No. 123 crystal.	formatting issues
Space group No. 166 crystal.	formatting issues

The efficiencies of the DFT calculations and M3GNet screening steps have been investigated and are reported in Table S5 and the linked file (https://github.com/WPEM/CGWGAN/eff.zip).

Supplementary Table 5. The comparison of VASP and M3GNet computing efficiencies.

Parameter	**M3GNet**	**VASP**
CPU Model	C86-7185	C86-7185
Architecture	X86_64	X86_64
Total Cores	1	128
Ionic Steps Time	0.26 sec	3.13 sec

Supplementary Table 6. Detailed comparison of novel crystals generated by CGWGAN and those recorded in the MP database.

Compositions	SG	PG	Total atoms	E_f	Source
$BaRuO_6$	*P*1 *(1)*	1	8	-2.01 eV	CGWGAN
$BaRuO_3$	*Pmm*2 *(25)*	$mm2$	5	-1.93 eV	CGWGAN
$BaRu_2O_3$	*Cmm*2 *(35)*	$mm2$	6	-1.94 eV	CGWGAN
$BaRuO_2$	*Immm (71)*	mmm	4	-1.87 eV	CGWGAN
$BaRuO_3$	*P*4*mm (99)*	$4mm$	5	-1.97 eV	CGWGAN
$BaRuO_3$	*P*4/*mmm (123)*	$4/mmm$	5	-1.97 eV	CGWGAN
$Ba_3Ru_3O_9$	$R\bar{3}m$ *(166)*	$\bar{3}m$	15	-2.04 eV	CGWGAN
$Ba_2Ru_7O_{18}$	$P\bar{1}$ *(2)*	$\bar{1}$	27	-1.75 eV*	MP
$Ba_4Ru_3O_{10}$	*P*12$_1$/*c*1 *(14)*	$2/m$	34	-2.29 eV*	MP
$Ba_5Ru_2O_{11}$	*C*2/*c (15)*	$2/m$	72	-2.25 eV*	MP
$Ba_4Ru_3O_{10}$	*Cmce (64)*	mmm	34	-2.29 eV*	MP
$Ba(RuO_2)_6$	*P*4/*n (85)*	$4/m$	76	-1.59 eV*	MP
Ba_2RuO_4	*I*4/*mmm (139)*	$4/mmm$	14	-2.39 eV*	MP
$BaRuO_3$	$R\bar{3}m$ *(166)*	$2/m$	30	-2.19 eV*	MP
$BaRuO_5$	$R\bar{3}c$ *(167)*	$\bar{3}m$	126	-1.78 eV*	MP
$Ba_5(RuO_5)_2$	*P*6$_3$/*mmc (194)*	$6/mmm$	34	-2.44 eV*	MP
$BaRuO_3$	*P*6$_3$/*mmc (194)*	$6/mmm$	20	-2.19 eV*	MP
$BaRuO_3$	*P*6$_3$/*mmc (194)*	$6/mmm$	30	-2.16 eV*	MP
$BaRuO_3$	$Pm\bar{3}m$ *(221)*	$m\bar{3}m$	5	-2.12 eV*	MP

*The formation energy values are retrieved from the MP database.

To make a comparison of CGWGAN with existing crystal structure prediction methods, one must have criterions and Jiao et al. (2024)[1] proposed three criterions to compare generated crystals, which are Validity, Coverage and Property Statistics. The structural validity evaluates whether the structures of generated samples are valid, and the structural validity rate is defined as the ratio of samples with a minimal pairwise distance greater than 0.5 Å over the number of total generated samples. The minimal pairwise distance in CGWGAN is defined as 1 Å, double the criterion of 0.5 Å, which means that these generated samples pass the 0.5 Å criterion validity may fail the 1 Å criterion validity. The composition validity criterion is based the electron valence and

charge neutrality solved by SMACT (Davies et al., 2019)[2], where no fractional charges are allowed in the crystals. This classical picture is not consistent with ab-initial calculation results, in which the charge partition gives fractional charges to each ion in a crystal cell. Therefore, the composition validity is not considered here. The coverage assesses the generated crystals having how much coverage with available crystals. For this purpose, a testing data subset must be randomly taken from the used dataset and these testing data cannot be seen by the crystal generation model. The coverage recall (COV-R) and precision (COV-P) metrics represent the percentage of crystals in the testing set that match those in the generated samples within a specified fingerprint distance threshold. The property statistics are represented by three Wasserstein distances between the generated and testing structures, specifically focusing on density (d_ρ), formation energy (d_E), and the number of elements. As mentioned in main text, CGWGAN generates crystal templates first, then fills-in atoms of the same or different elements by designer, and then uses M3GNet to estimate the formation energy and phonon spectrum. Clearly, the accuracies of estimated formation energy and phonon spectrum rely completely on M3GNet. Only these generated crystals passing M3GNet will be further examined by the ab-initial calculations.

The comparison employs the Perov-5 dataset (Castelli et al., 2012)[3], which includes 18,928 perovskite crystals with similar structures but distinct compositions and each structure contains exactly five atoms per unit crystal cell. The structural validity and coverage metrics are calculated on all 10,000 generated samples and the about 20% of the Perov-5 dataset is randomly selected as the testing set, which are all consistent with Jiao's experimental settings. Table S7 lists the comparison of CGWGAN with existing crystal structure prediction methods. The comparison indicates that CGWGAN is comparable to these recently developed methods.

Supplementary Table 7. The structural validity and coverage metrics of various crystal structure prediction methods, where other baseline results are from Xie et al. (2021)[4] and Jiao et al.

Method	0.5 Å Validity	1.0 Å Validity	COV-R	COV-P	d_ρ	d_E
FTCP	0.24	-	0.00	0.00	10.27	156.0
Cond-DFC-VAE	73.60	-	73.92	10.13	2.268	4.111
G-SchNet	99.92	-	0.18	0.23	1.625	4.746
P-G-SchNet	79.63	-	0.37	0.25	0.2755	1.388
CDVAE	**100.0**	-	99.45	98.46	0.1258	0.0264
DiffCSP	**100.0**	-	**99.74**	98.27	0.1110	**0.0263**
DiffCSP++	**100.0**	-	99.60	**98.80**	0.0661	0.0405
CGWGAN	**100.0**	**100.0**	99.50	93.90	0.0821	0.0271

A Case study of crystal embedding:

The embedding vector consists of three main components: the fractional coordinates of the ASU, the space group number, and the lattice parameters a, b, c, α, β, and γ. The initial embedding vector for a crystal with three ASU units is as follows:

$$\overline{\boldsymbol{C}}^i = [0.5,0.5,0.5,0.75,0.75,0.75,0.75,0.75,0.25,216,6.53,6.53,6.53,90,90,90]$$

This vector includes the 3x3 fractional coordinates, one space group number, and six lattice constants.

To enrich the crystal embedding, our model embeds three or four ASU structures together. Therefore, the practical embedding contains 3x4 fractional coordinates. In cases where only three ASUs are present, a placeholder is used for the fourth set of coordinates, with the values (−2,−2,−2). The resulting embedding vector for three ASU crystals is:

$$\overline{\boldsymbol{C}}^i = [0.5,0.5,0.5,0.75,0.75,0.75,0.75,0.75,0.25,-2,-2,\\ -2,216,6.53,6.53,6.53,90,90,90]$$

The INCAR in the standard version of VASP.6.4.3.

The pseudopotential is the official GGA and the K-point density has been specified in INCAR via KSPACING:

SYSTEM = CGWGAN

Algo = All

LWAVE = .FALSE.

LCHARG = .FALSE.

ISPIN = 1

ISMEAR = 1

SIGMA = 0.2

KSPACING=0.2

NSW = 300

IBRION = 2

ISIF = 3

POTIM = 0.1

EDIFF = 1e-7

Reference

[1] Jiao R, Huang W, Liu Y, et al. Space group constrained crystal generation[J]. arXiv preprint arXiv:2402.03992, 2024.

[2] Davies D W, Butler K T, Jackson A J, et al. Smact: Semiconducting materials by analogy and chemical theory[J]. Journal of Open Source Software, 2019, 4(38): 1361.

[3] Castelli I E, Landis D D, Thygesen K S, et al. New cubic perovskites for one-and two-photon water splitting using the computational materials repository[J]. Energy & Environmental Science, 2012, 5(10): 9034-9043.

[4] Xie T, Fu X, Ganea O E, et al. Crystal diffusion variational autoencoder for periodic material generation[J]. arXiv preprint arXiv:2110.06197, 2021.