

## Supplementary Information

*Uncertainty-Aware Active Learning Reveals Reliability Limits in Lead-Free Halide Perovskite Screening*

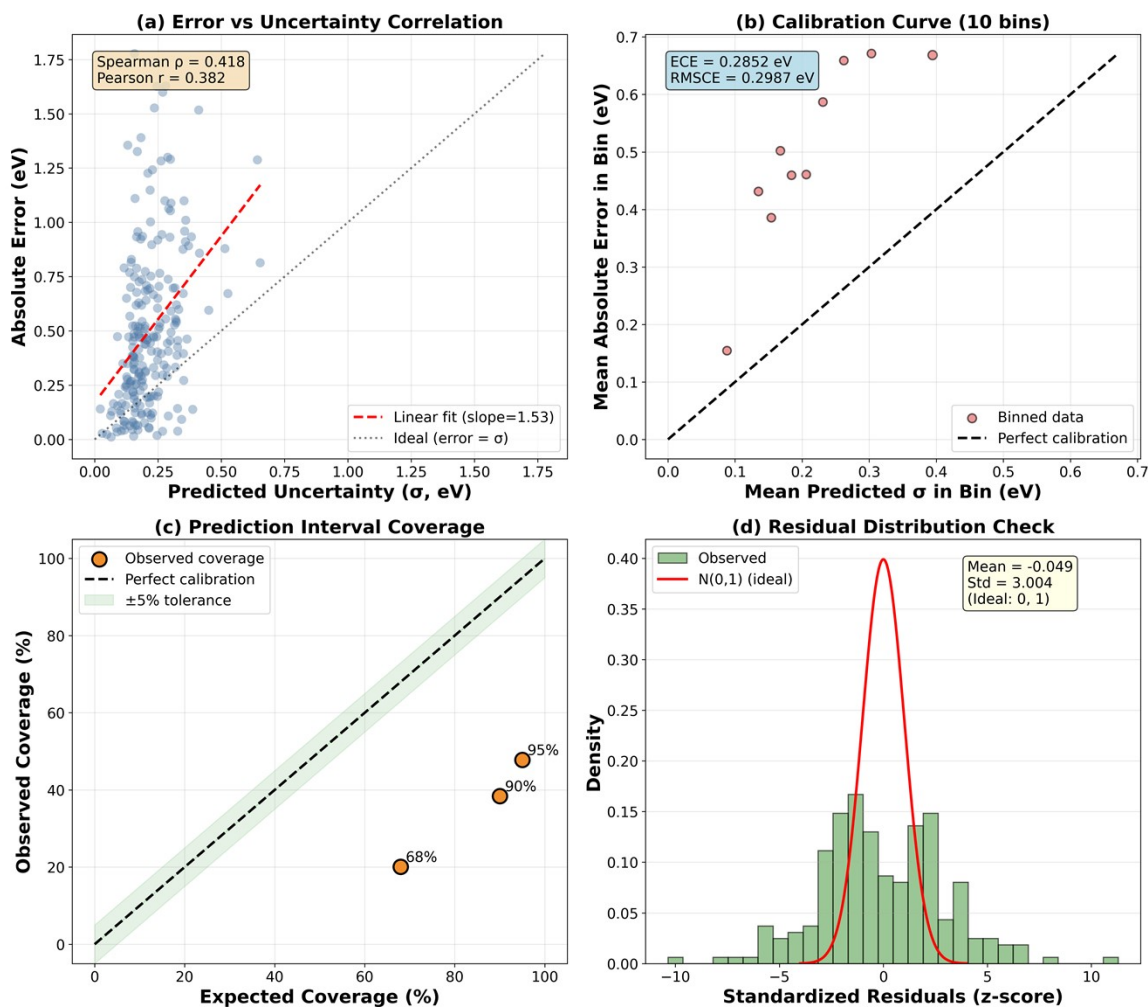


Fig. S1. Uncertainty calibration diagnostics on the held-out test set: (a) predicted  $\sigma$  vs. absolute error with Pearson and Spearman correlations; (b) expected vs. observed coverage for prediction intervals from 10% to 90%; (c) residual-normalised Z-score histogram compared to a standard normal.

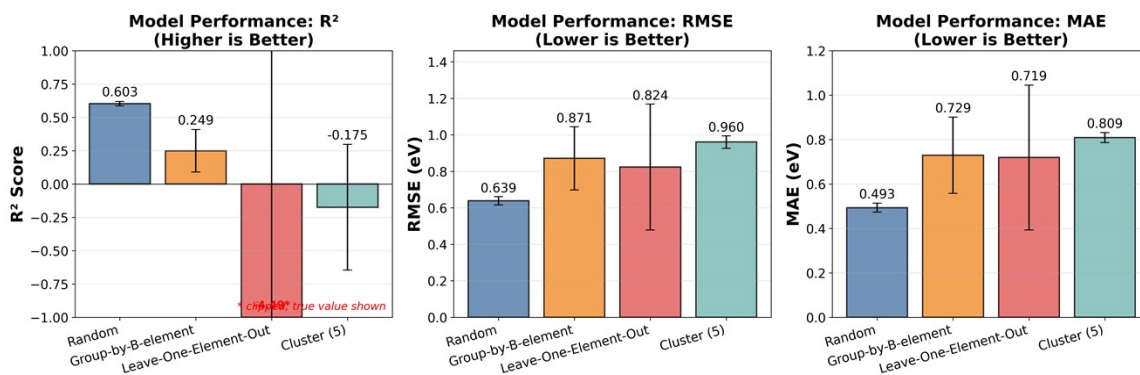


Fig. S2. Generalisation comparison under four data-splitting strategies: Random (80/20), GroupKFold by B-site cation, Leave-One-Element-Out (LOEO), and feature-space KMeans clustering. Panels show (a) R<sup>2</sup>, (b) RMSE, and (c) MAE with error bars from 5-fold cross-validation. Numerical values are annotated above each bar.

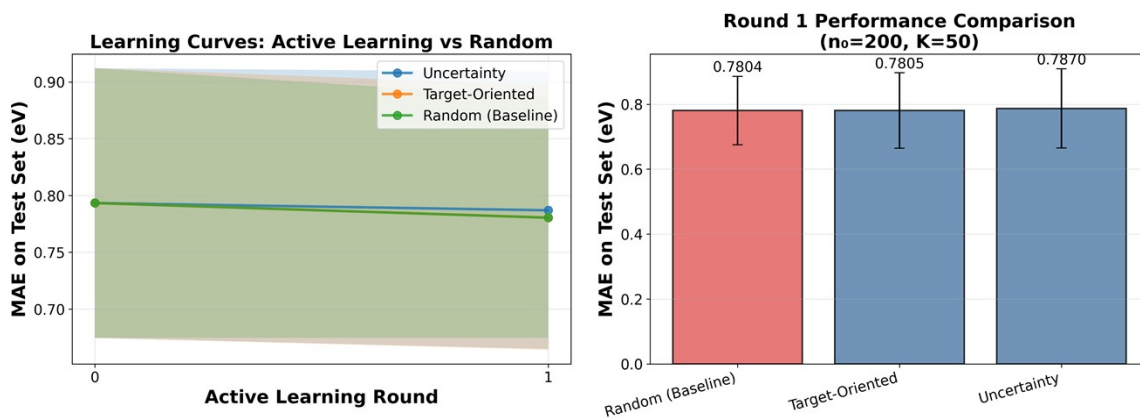


Fig. S3. Retrospective pool-based active learning simulation over six acquisition rounds (n<sub>o</sub> = 200, K = 50 per round). Comparison of Random, Uncertainty, Greedy, and Hybrid acquisition strategies; y-axis shows relative MAE change (%) from the initial model.

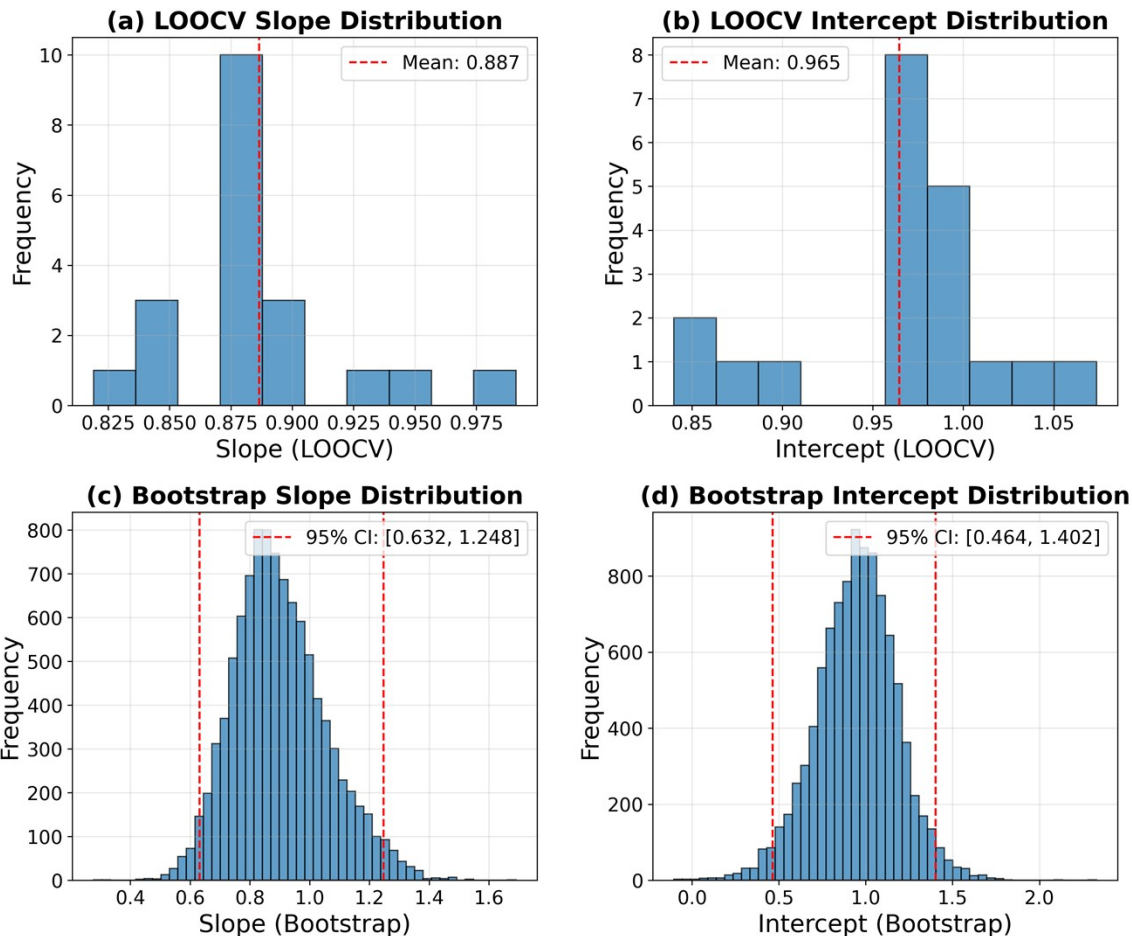


Fig. S4. Calibration robustness diagnostics: (a) leave-one-out cross-validation (LOOCV) of the 20-compound calibration set, showing predicted vs. actual experimental band gap for each held-out compound; (b) bootstrap distribution of slope and intercept (10,000 resamples) with 95% confidence intervals.

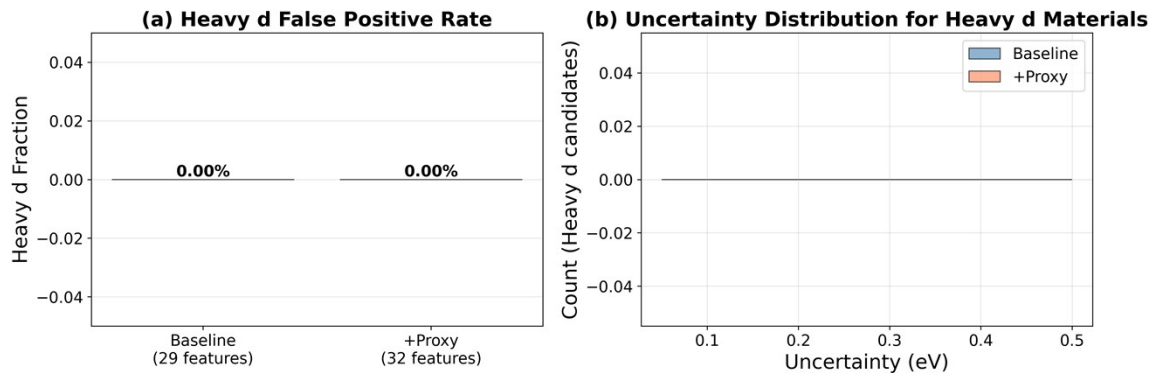


Fig. S5. Proxy feature ablation: fraction of heavy d/p-element compounds in the ML-priority set under the baseline 29-feature model vs. the proxy-augmented model (35+ descriptors including SOC proxy, d-electron count, and Madelung-like terms).

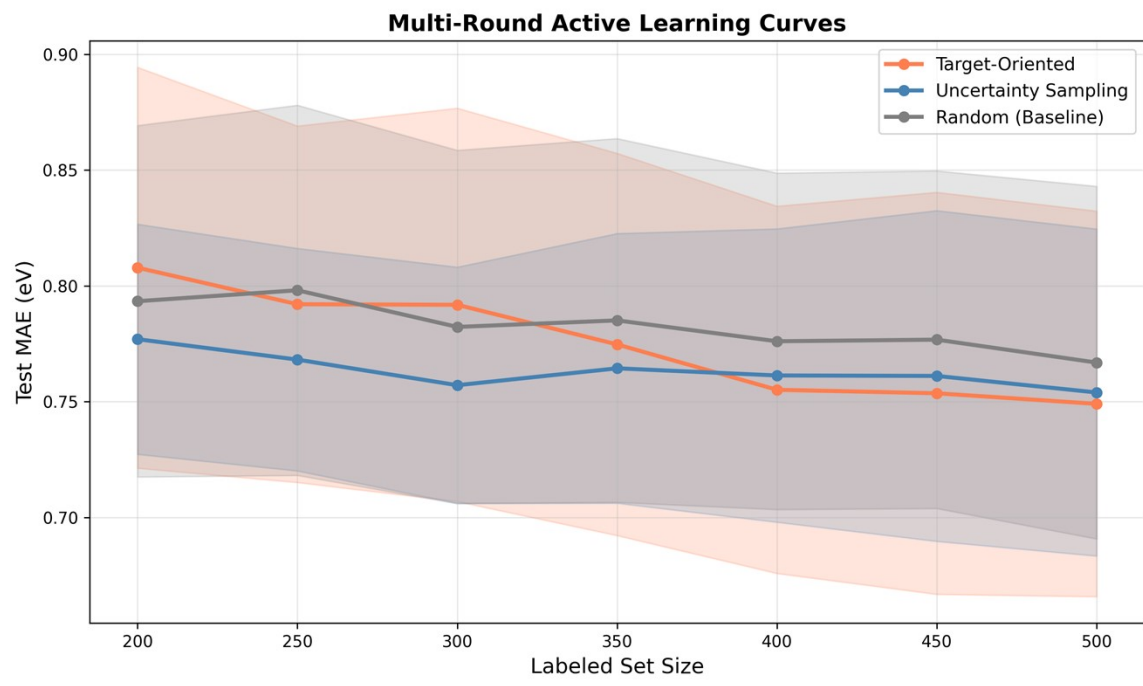


Fig. S6. Multi-round active learning curves showing MAE trajectory over 6 acquisition rounds for each strategy (Random, Uncertainty, Greedy, Hybrid), averaged over 10 seeds with  $\pm 1$  standard-error bands.

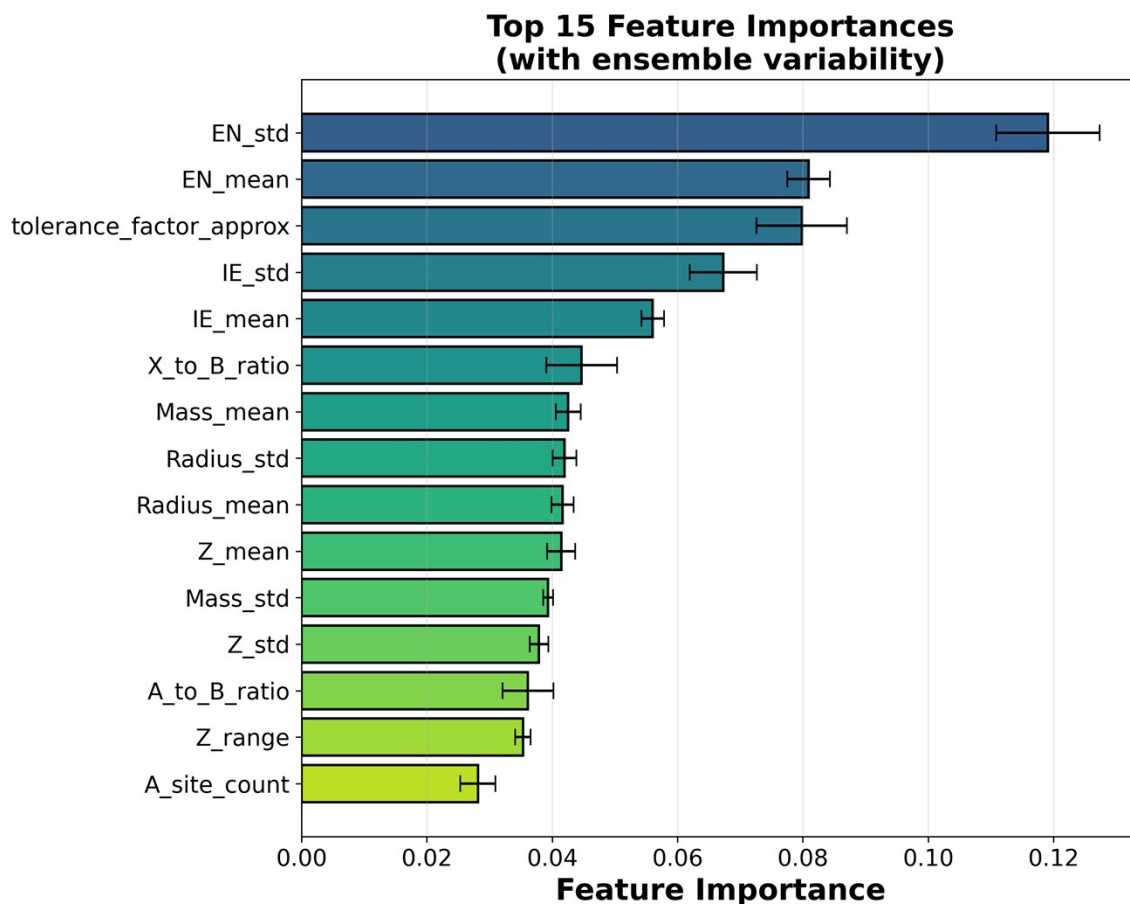


Fig. S7. Feature importance ranked by mean ensemble variability (standard deviation of feature importances across 15 ensemble members). Top-10 features are highlighted.

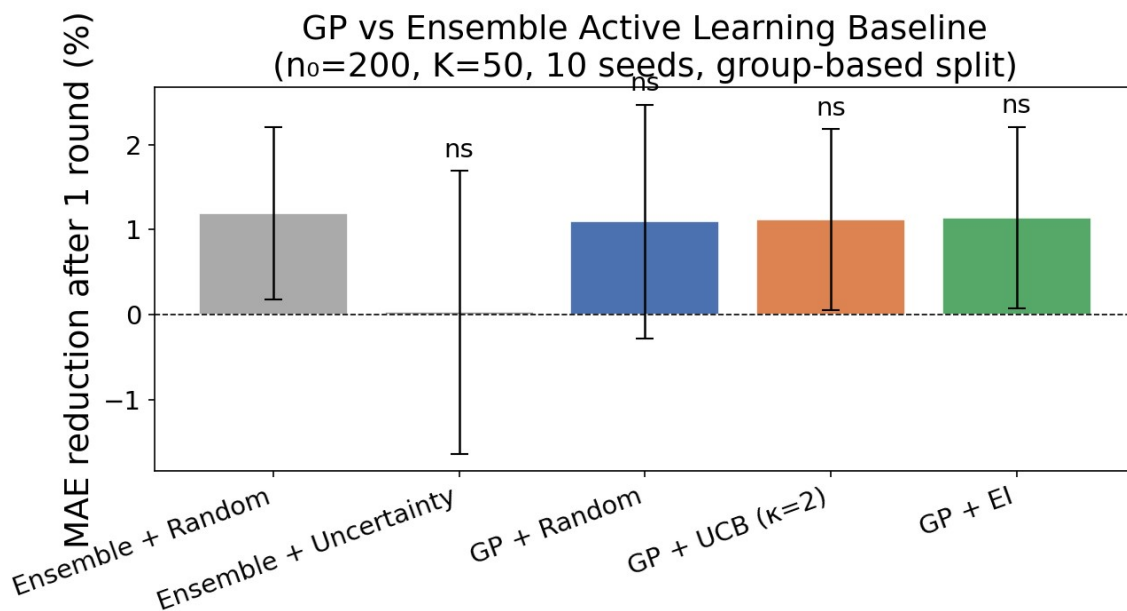


Fig. S8. GP vs. ensemble active learning baseline comparison. Mean MAE reduction (%) after one acquisition round ( $n_0 = 200$ ,  $K = 50$  additional points) across 10 random seeds, for five acquisition strategies: Ensemble + Random (grey), Ensemble + Uncertainty (dark grey), GP + Random (blue), GP + UCB  $\kappa = 2$  (orange), and GP + EI (green). Error bars show  $\pm 1$  standard error across seeds. 'ns' indicates no significant difference from the Ensemble + Random baseline (paired t-test, all  $p > 0.9$ ).

**Table S1. Complete list of 29 compositional descriptors used as model input features.**

#	Feature Name	Description
1	mean_AtomicRadius	Weighted mean atomic radius (Å)
2	std_AtomicRadius	Weighted std of atomic radius
3	mean_Electronegativity	Weighted mean Pauling electronegativity
4	std_Electronegativity	Weighted std of electronegativity
5	mean_AtomicWeight	Weighted mean atomic mass (u)
6	std_AtomicWeight	Weighted std of atomic mass
7	mean_MeltingPoint	Weighted mean melting point (K)
8	std_MeltingPoint	Weighted std of melting point
9	mean_AtomicNumber	Weighted mean atomic number Z
10	std_AtomicNumber	Weighted std of Z
11	mean_FirstIonizationEnergy	Weighted mean first IE (eV)
12	std_FirstIonizationEnergy	Weighted std of first IE
13	num_elements	Number of distinct elements
14	num_sites	Total number of formula units
15	A_site_count	Number of A-site atoms per formula unit
16	B_site_count	Number of B-site atoms per formula unit
17	X_site_count	Number of X-site (halide) atoms per formula unit
18	A_mean_radius	Mean atomic radius of A-site elements
19	B_mean_radius	Mean atomic radius of B-site elements
20	X_mean_radius	Mean atomic radius of X-site (halide) elements
21	A_mean_EN	Mean electronegativity of A-site elements
22	B_mean_EN	Mean electronegativity of B-site elements
23	X_mean_EN	Mean electronegativity of X-site elements
24	tolerance_factor_approx	Goldschmidt tolerance factor $t = (r_A + r_X) / [\sqrt{2} (r_B + r_X)]$
25	is_double_perovskite	Binary flag: 1 if two distinct B-site elements, else 0
26	halide_mix_I	Fractional I content on X-site
27	halide_mix_Br	Fractional Br content on X-site
28	halide_mix_Cl	Fractional Cl content on X-site
29	B_site_Z_range	Range of atomic numbers across B-site elements

**Table S2. The 20 inorganic benchmark compounds used for PBE-to-experimental band gap calibration.**

Formula	Type	PBE gap (eV)	Exp. gap (eV)	MP-ID	Reference
CsSnI <sub>3</sub>	Sn	0.258	1.31	mp-27214	Chung et al. Nature 485, 486 (2012)
CsSnBr <sub>3</sub>	Sn	0.438	1.75	mp-27537	Stoumpos et al. Inorg. Chem. 52, 9019 (2013)
CsSnCl <sub>3</sub>	Sn	1.021	2.90	mp-29241	Stoumpos et al. Inorg. Chem. 52, 9019 (2013)
CsGeI <sub>3</sub>	Ge	0.553	1.63	mp-27800	Stoumpos et al. Inorg. Chem. 54,

					2757 (2015)
CsGeBr <sub>3</sub>	Ge	0.952	2.32	mp-28279	Stoumpos et al. Inorg. Chem. 54, 2757 (2015)
CsGeCl <sub>3</sub>	Ge	1.636	3.67	mp-28786	Stoumpos et al. Inorg. Chem. 54, 2757 (2015)
Cs <sub>3</sub> Bi <sub>2</sub> I <sub>9</sub>	Bi	1.451	2.03	mp-27531	Lehner et al. Angew. Chem. 54, 8546 (2015)
Cs <sub>3</sub> Bi <sub>2</sub> Br <sub>9</sub>	Bi	2.015	2.60	mp-29957	Hoye et al. Chem. Eur. J. 22, 2605 (2016)
Cs <sub>3</sub> Bi <sub>2</sub> Cl <sub>9</sub>	Bi	2.629	3.30	mp-31131	Hoye et al. Chem. Eur. J. 22, 2605 (2016)
Cs <sub>3</sub> Sb <sub>2</sub> I <sub>9</sub>	Sb	1.305	2.05	mp-27466	Saparov et al. Chem. Mater. 27, 5622 (2015)
Cs <sub>3</sub> Sb <sub>2</sub> Br <sub>9</sub>	Sb	1.793	2.67	mp-29568	Vargas et al. J. Phys. Chem. Lett. 8, 1412 (2017)
Cs <sub>3</sub> Sb <sub>2</sub> Cl <sub>9</sub>	Sb	2.422	3.56	mp-31130	Vargas et al. J. Phys. Chem. Lett. 8, 1412 (2017)
Cs <sub>2</sub> AgBiBr <sub>6</sub>	double	1.143	1.95	mp-985241	Slavney et al. J. Am. Chem. Soc. 138, 2138 (2016)
Cs <sub>2</sub> AgBiCl <sub>6</sub>	double	1.936	2.77	mp-27163	Slavney et al. J. Am. Chem. Soc. 138, 2138 (2016)
Cs <sub>2</sub> AgInCl <sub>6</sub>	double	2.266	3.23	mp-862977	Volonakis et al. J. Phys. Chem. Lett. 8, 772 (2017)
Cs <sub>2</sub> AgSbBr <sub>6</sub>	double	0.627	1.64	mp-1102078	Deng et al. Adv. Funct. Mater. 30, 2002131 (2020)
Cs <sub>2</sub> SnI <sub>6</sub>	vacancy	0.300	1.26	mp-676481	Saparov et al. Chem. Mater. 28, 2315 (2016)
Cs <sub>2</sub> SnBr <sub>6</sub>	vacancy	0.852	1.82	mp-675760	Kaltzoglou et al. J. Phys. Chem. C 120, 11777 (2016)
Cs <sub>2</sub> SnCl <sub>6</sub>	vacancy	1.795	3.90	mp-30878	Kaltzoglou et al. J. Phys. Chem. C 120, 11777 (2016)
Cs <sub>2</sub> PdBr <sub>6</sub>	vacancy	0.593	1.60	mp-1079158	Babu et al. J. Phys. Chem. C 124, 10580 (2020)