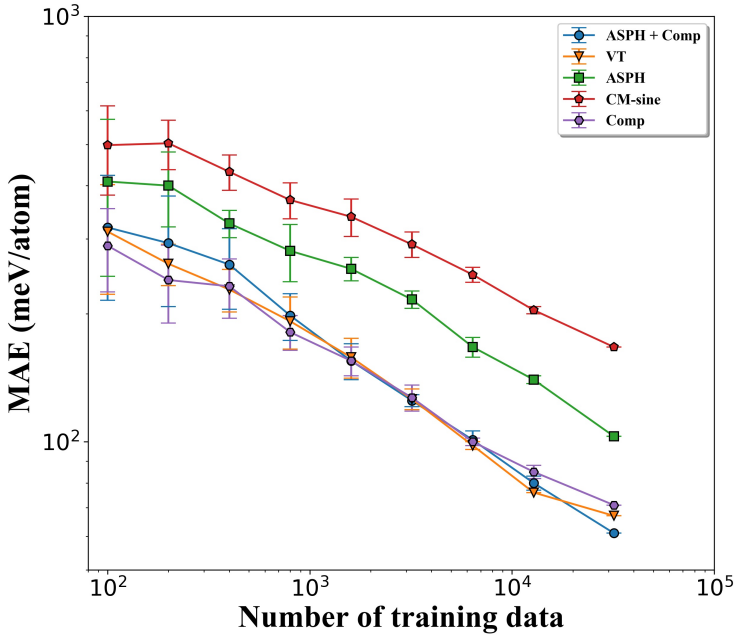
Supplementary Information for Topological representations of crystalline compounds for the machine-learning prediction of materials properties

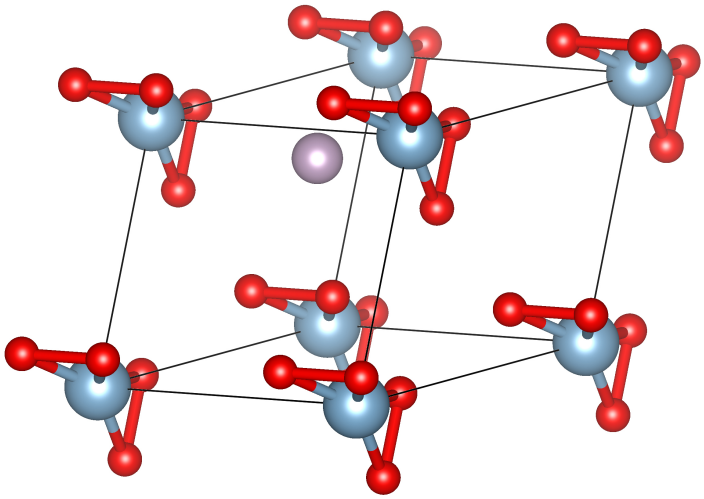
Yi Jiang[1], Dong Chen[1,2], Xin Chen[1], Tangyi Li[1], Guo-Wei Wei[*2] and Feng Pan[*1]

[1] *School of Advanced Materials, Peking University, Shenzhen Graduate School, Shenzhen 518055, People's Republic of China*
[2] *Department of Mathematics, Michigan State University, MI 48824, USA*

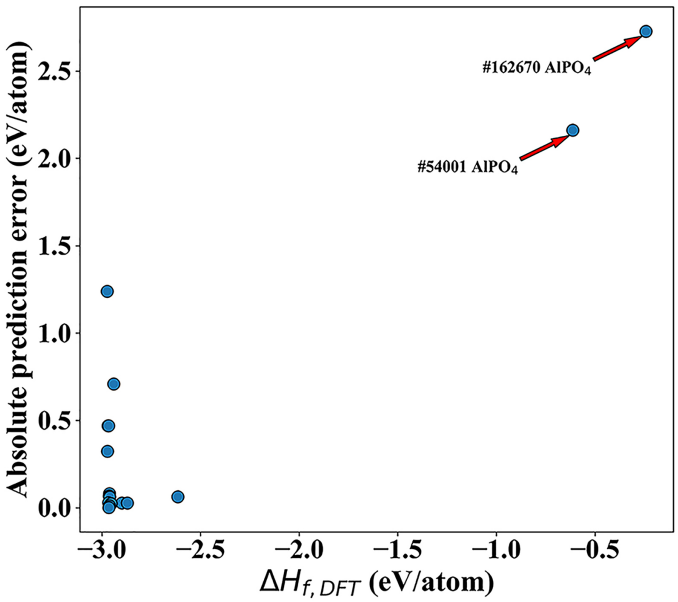
[*] Corresponding author: weig@msu.edu, panfeng@pkusz.edu.cn



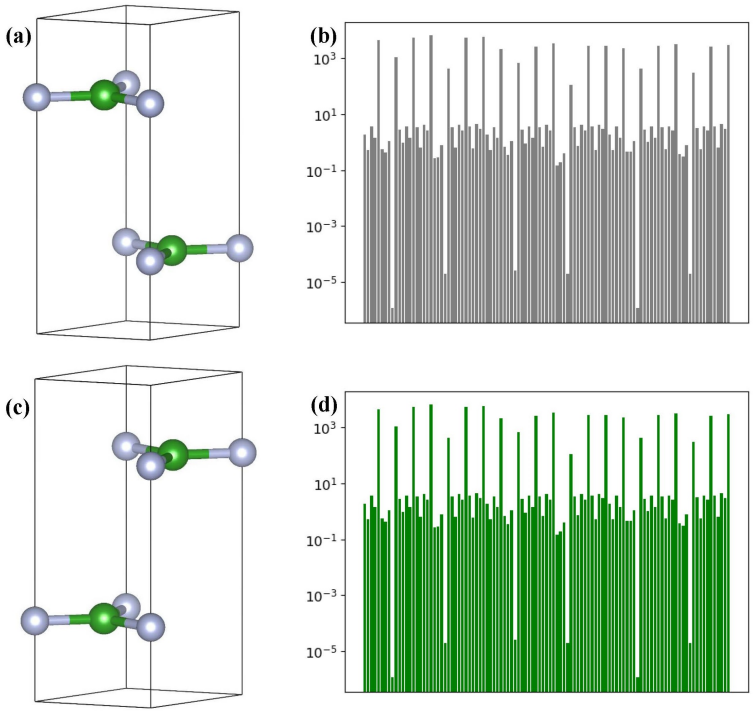
Supplementary Figure 1. The prediction performance of multiple representations with respect to the training samples. *Comp* means the compositional feature described by Ward [1].



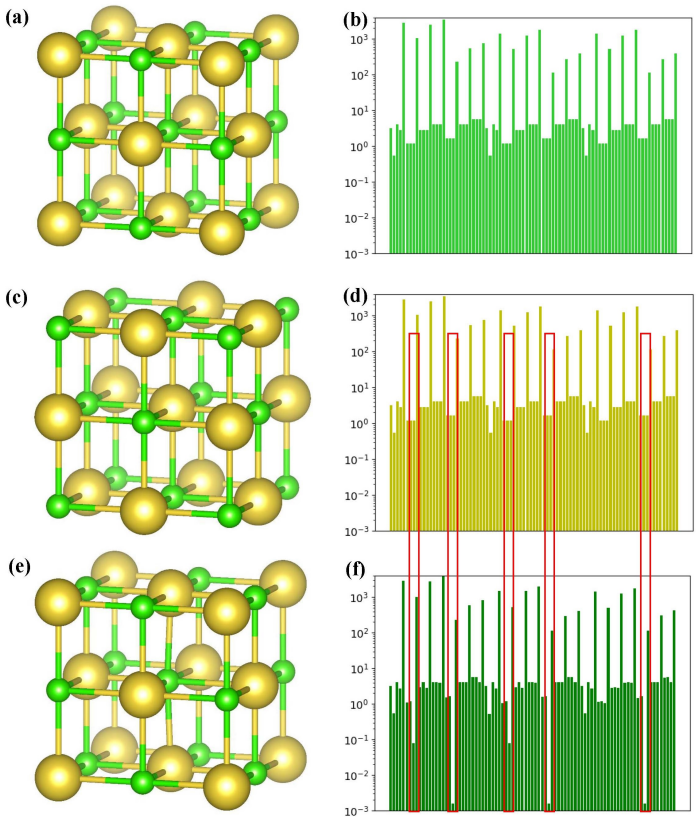
Supplementary Figure 2. The crystal structure of $AlPO_4$ (ICSD #162670) associated with the worst prediction. It behaves abnormally with a relatively short O-O bond of 1.52 Å and the coordination number of the P atom is zero.



Supplementary Figure 3. Illustration of DFT-calculated formation enthalpies and the associated absolute errors in our predictions for some forstereoisomerisms of $AlPO_4$.



Supplementary Figure 4. The unit cell of BN and its rotated one, and their respective feature. (a) The structure of original BN and (b) its feature vector. (c) The structure of rotated BN and (d) its feature vector.



Supplementary Figure 5. The unit cell of original NaCl, translated NaCl and ones with tiny atomic position change, and their features, respectively. (a) The structure of original NaCl and (b) its feature vector. (c) The structure of translated NaCl and (d) its feature vector. (e) The structure of NaCl with tiny changes on atomic positions and (f) its feature vector. The red frames indicates the changed features.

References

[1] Logan Ward, Ankit Agrawal, Alok Choudhary, and Christopher Wolverton. A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Computational Materials*, 2:16028, 2016.