

# Automation of machine learning driven interatomic potential generation for predicting vibrational properties

Christina Ertural<sup>1\*</sup>, Volker L. Deringer<sup>2</sup>, Janine George<sup>1,3</sup> (email: christina.ertural@bam.de)

<sup>1</sup>Federal Institute for Materials, Research and Testing, Department Materials Chemistry, Unter den Eichen 87, 12205 Berlin, Germany.

<sup>2</sup>Department of Chemistry, Inorganic Chemistry Laboratory, University of Oxford, Oxford OX1 3QR, United Kingdom

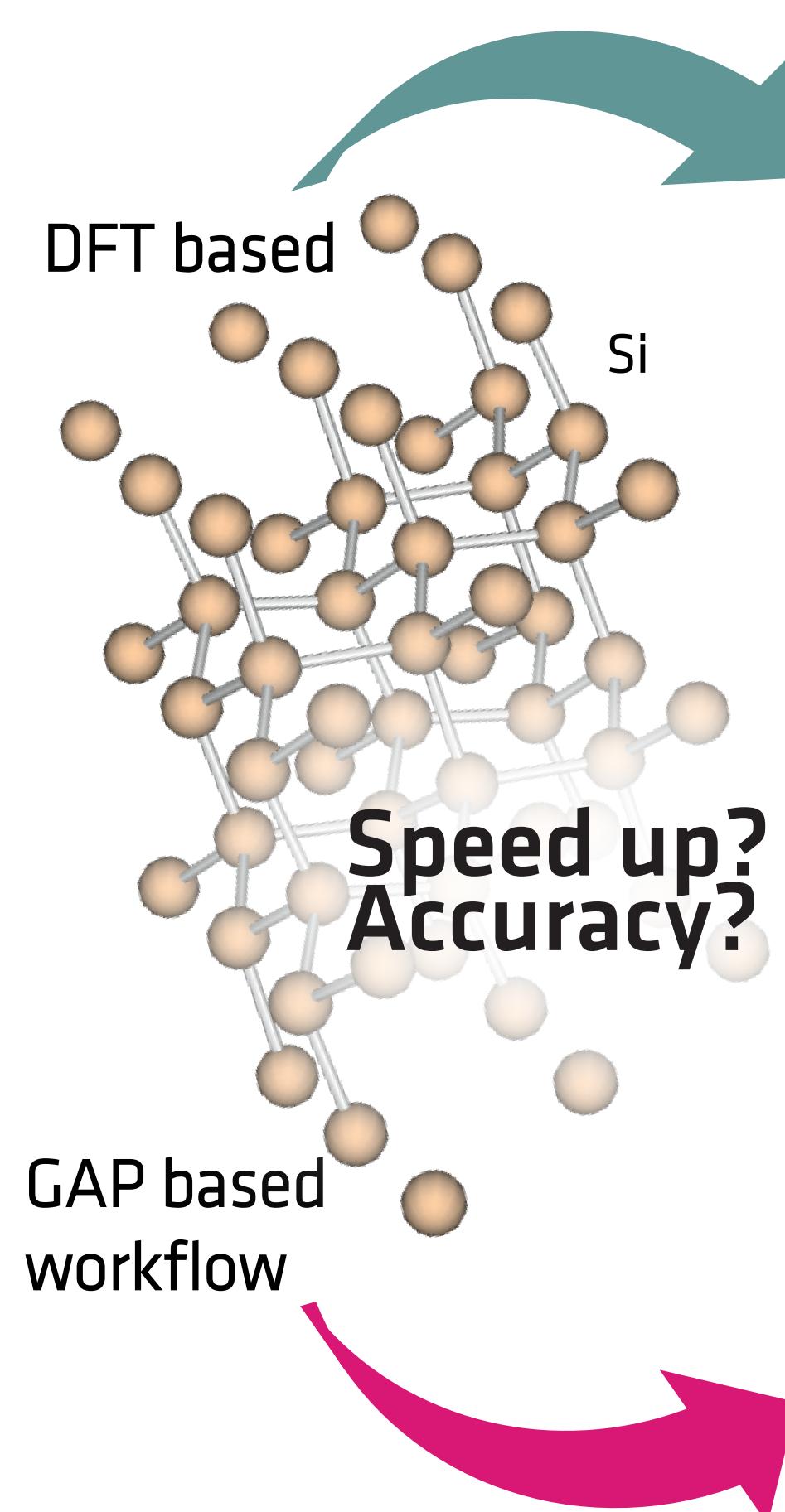
<sup>3</sup>Friedrich Schiller University Jena, Institute of Condensed Matter Theory and Solid-State Optics, Max-Wien-Platz 1, 07743 Jena, Germany.

## Introduction

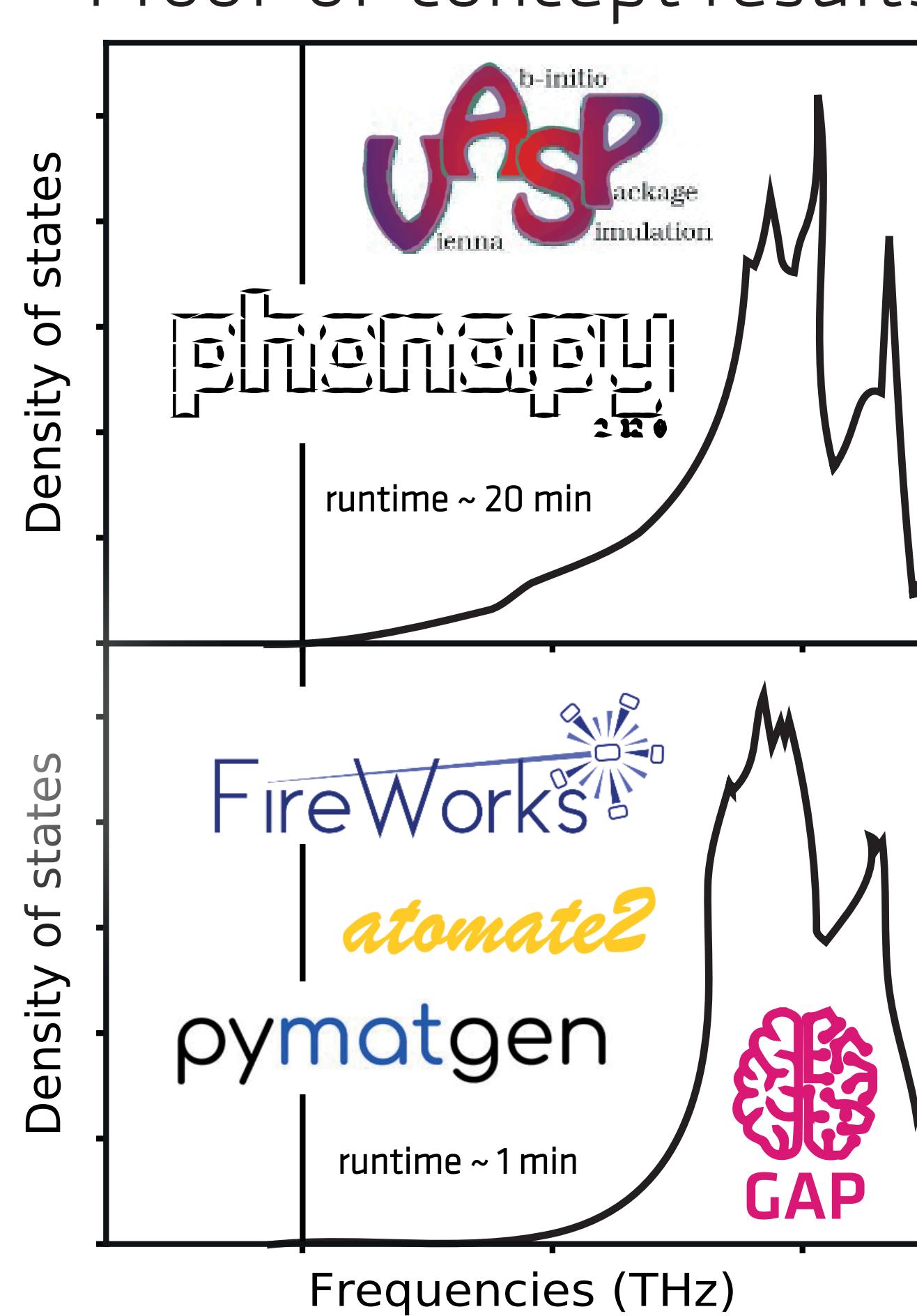
Investigating the phononic properties is beneficial for predicting low thermal conductivity thermoelectric materials.<sup>1-3</sup> Employing density functional theory<sup>4</sup> takes many calculation steps and consumes a lot of computational resources.<sup>5,6</sup> Using machine learning driven interatomic potentials (MLIP, e.g., Gaussian approximation potential,<sup>8</sup> GAP) opens up a faster route to phonons<sup>7</sup> but in most cases, the potentials are specifically tailored for a certain compound. In this work, we automate the generation of such MLIPs in a Python code-based workflow, based on the automation tools atomate<sup>9</sup> and pymatgen<sup>10</sup> which combines the automatic DFT computations with the automated fitting of GAPs. Automation enables easier testing, benchmarking and validation.<sup>11</sup> We aim to provide the workflow-generated potentials for storage in databases.

## Methodology

### The idea

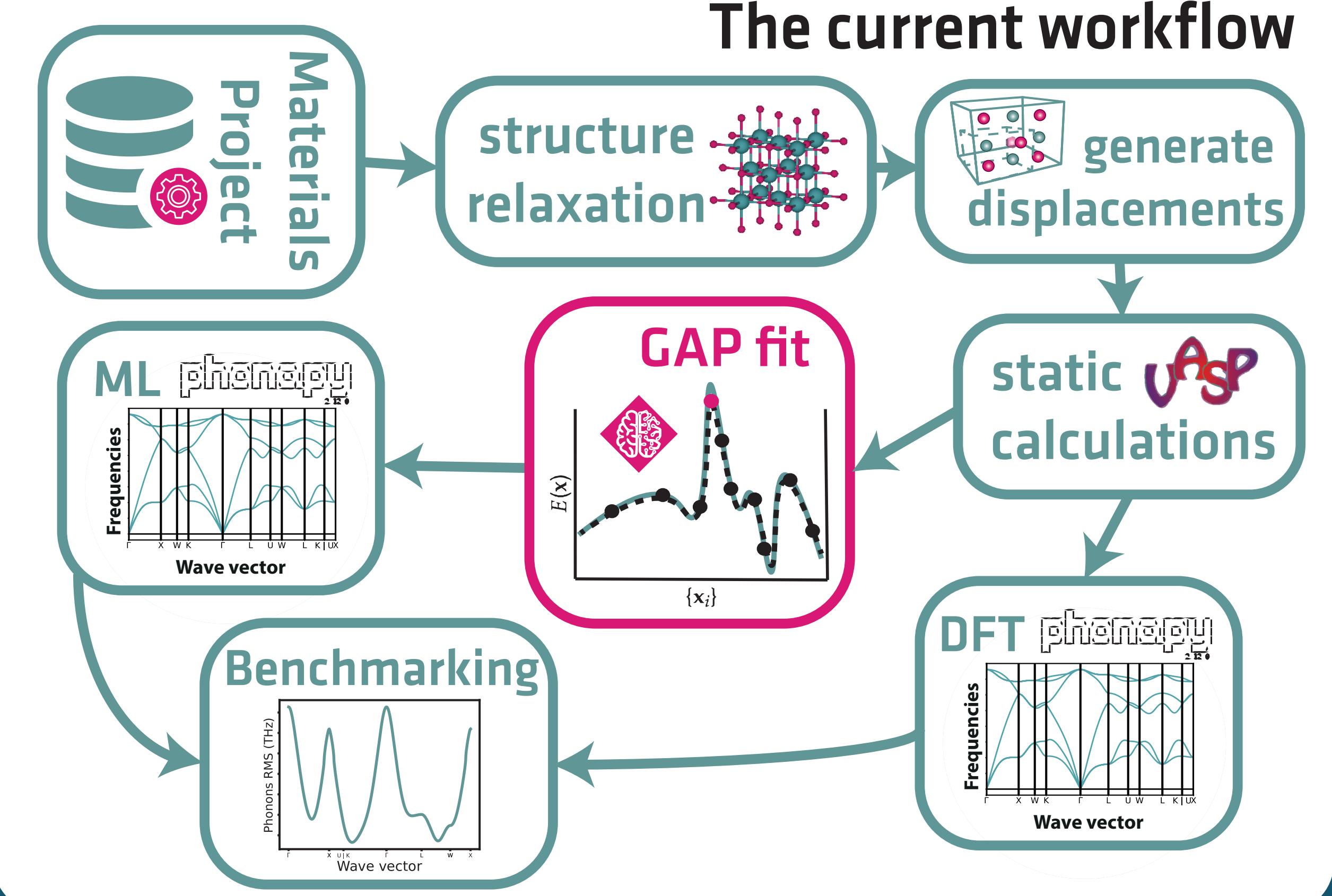


### Proof-of-concept results



Progress 80%

### The current workflow

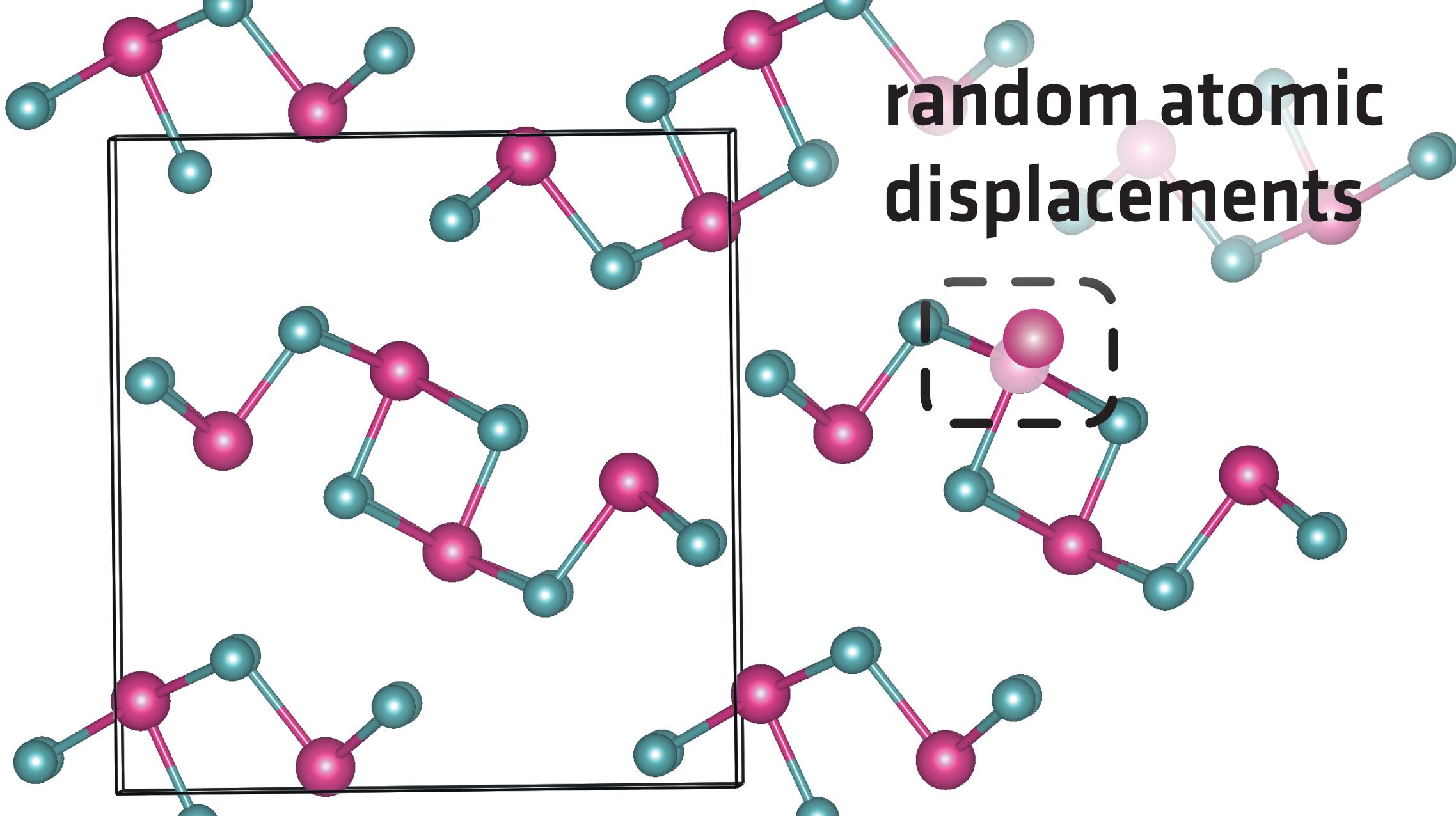


### Outlook

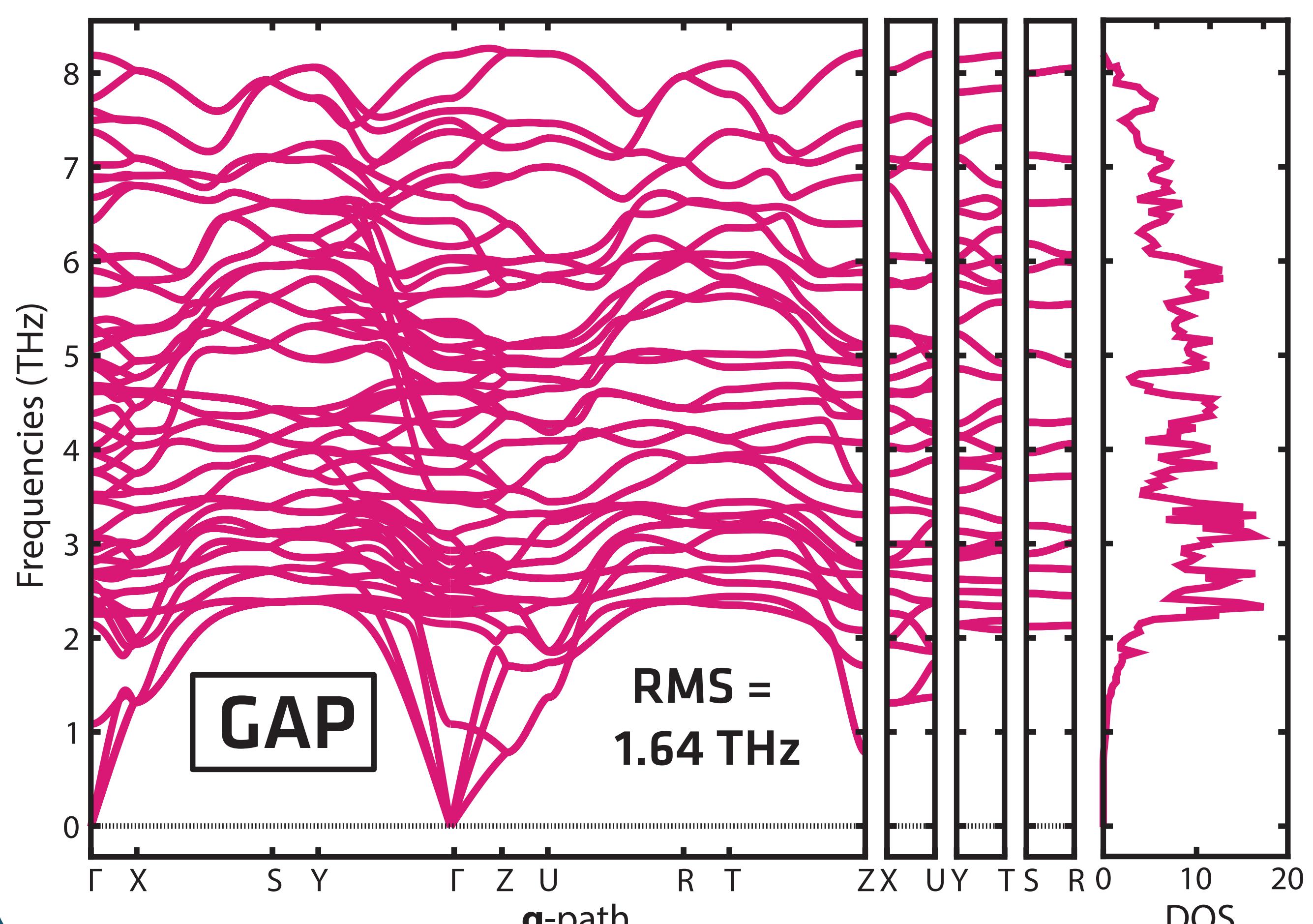
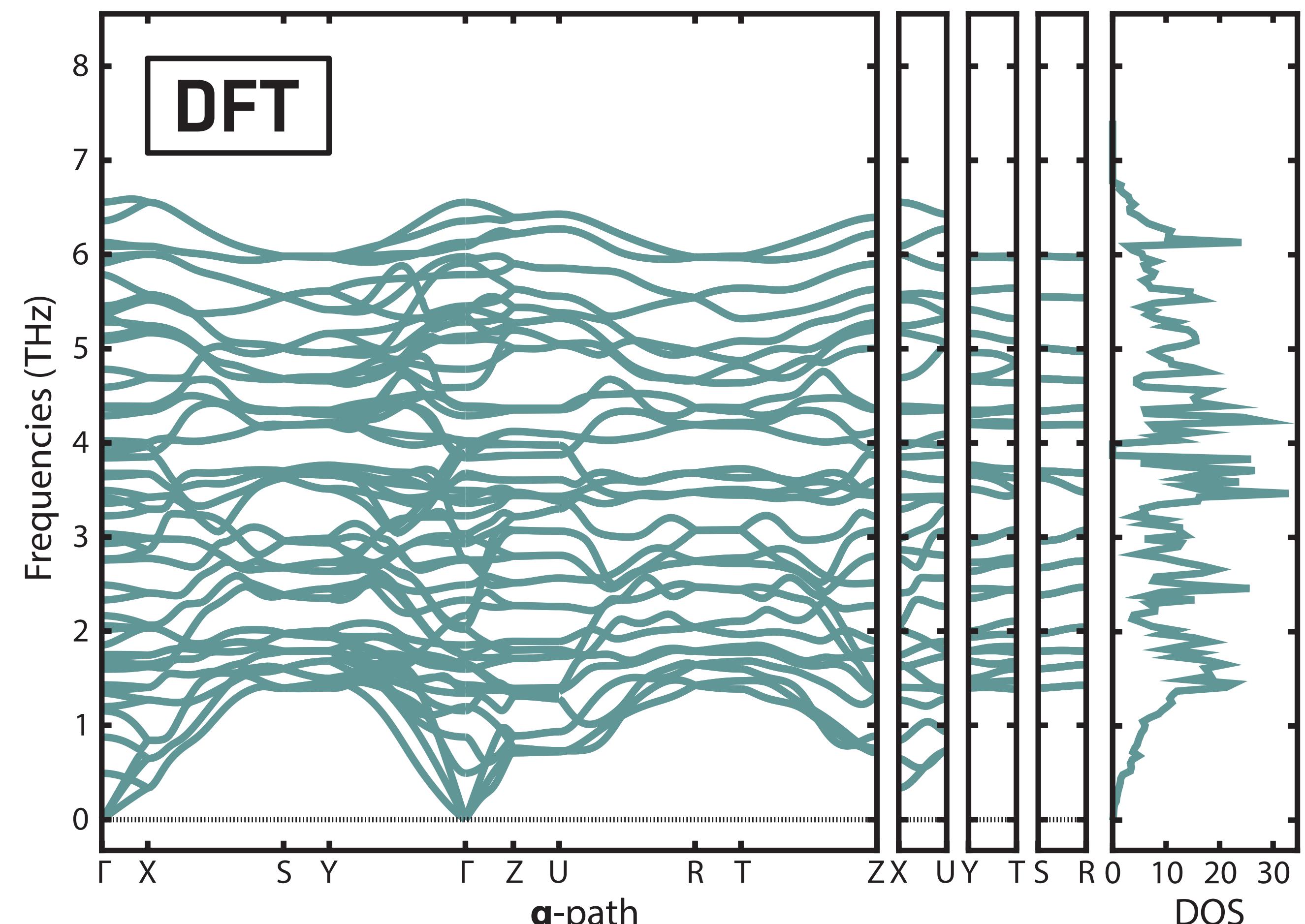
- Include a variety of MLIPs (SNAP, MTP, ACE, etc.) as well as defective, paracrystalline and amorphous structures
- Extend workflow features (like database sources & dataset generation)
- Optimize the MLIP fitting process with active learning, pre-training/fine-tuning and other methods
- Improve and extend validation/benchmarking beyond RMS using a variation of error metrics and physically guided validation<sup>11</sup>

## Preliminary results

Starting with a small test ML data set of 120 structures for  $\text{Sb}_2\text{Se}_3$  (mp-2160) to train a GAP



### Initial results – work in progress!



(1) Snyder, G. J.; Toberer, E. S. Nat. Mater. 2008, 7 (2), 105–114. (2) He, J.; Xia, Y.; Lin, W.; Pal, K.; Zhu, Y.; Kanatzidis, M. G.; Wolverton, C. Adv. Funct. Mater. 2022, 32 (14) 2108532.. (3) Wehmeyer, G.; Yabuki, T.; Monachon, C.; Wu, J.; Dames, C. Appl. Phys. Rev. 2017, 4 (4), 041304. (4) Togo, A.; Tanaka, I. Scr. Mater. 2015, 108, 1–5. (5) Hellman, O.; Steneteg, P.; Abrikosov, I. A.; Srimak, S. I. Phys. Rev. B 2013, 87 (10), 104111. (6) Eriksson, F.; Fransson, E.; Erhart, P. Adv. Theory Simul. 2019, 2 (5). (7) George, J.; Hautier, G.; Bartók, A. P.; Csányi, G.; Deringer, V. L. J. Chem. Phys. 2020, 153 (4), 044104. (8) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Phys. Rev. Lett. 2010, 104 (13), 136403. (9) Mathew, K.; Montoya, J. H.; Faghanninia, A.; Dwarkanath, S.; Aykol, M.; Tang, H.; Chu, I.; Smidt, T.; Bocklund, B.; Horton, M.; Dagdelen, J.; Wood, B.; Liu, Z.-K.; Neaton, J.; Ong, S. P.; Persson, K.; Jain, A. Comput. Mater. Sci. 2017, 139, 140–152. (10) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Comput. Mater. Sci. 2013, 68, 314–319. (11) Morrow, J. D.; Gardner, J. L. A.; Deringer, V. L. J. Chem. Phys. 2023, 158 (12), 121501.