AQFORGE: BRIDGING GENERATIVE MODELS AND PROPERTY PREDICTION FOR MATERIALS DISCOVERY

Shivang Agarwal, Kevin Ryczko, Rodrigo Wang
SandboxAQ
Palo Alto, CA 94301, USA
rodrigo.wang@sandboxaq.com

ABSTRACT

Designing and characterizing new materials with tailored properties is critical in fields such as catalysis, energy storage, and solid-state materials. Despite significant technological advances, including the development of generative models and universal machine learning force fields, these tools often operate in isolation rather than as integrated components of a comprehensive workflow. Additionally, while the computation of energies and forces remains highly valuable and the focus of many studies, it often falls short for accurately predicting certain task-specific macroscopic properties. To address these limitations, we propose an end-to-end workflow that extends the capabilities of current state-of-the-art works and fully automates the design and discovery of materials, with a particular emphasis on calculating downstream properties. By integrating and validating existing approaches, we ensure the robustness of the workflow and demonstrate its utility with a few illustrative use cases.

1 INTRODUCTION

The discovery and design of novel materials with desired properties has long been an essential focus for advancing fields like catalysis, energy storage, electronics, and polymer engineering (Alberi et al., 2018). Applications in these areas often depend on finding materials that meet specific performance criteria (Osman et al., 2021; Zhao et al., 2019; 2020). Traditional approaches, relying on experimental trials and computational modeling, can be time-consuming, resource-intensive, and limited in scope (de Pablo et al., 2019; Sinnott, 2013). Recent advancements in machine learning (ML) and computational methodologies have accelerated the process of material discovery by enabling a more efficient exploration of the materials design space.

ML force fields (MLFFs) (Liao et al., 2023; Batatia et al., 2022; Chen & Ong, 2022; Deng et al., 2023) and generative models (Jiao et al., 2023; Xie et al., 2021; Zeni et al., 2023; Gómez-Bombarelli et al., 2018; Jin & Merz Jr, 2024; Kadan et al., 2024) are among the most promising tools in this domain. MLFFs provide accurate atomic-level predictions, reducing computational costs compared to traditional *ab initio* methods. Similarly, generative models have shown remarkable success in proposing novel candidate materials with desired properties. However, these methodologies are often developed and applied in isolation, limiting their integration into unified workflows and their ability to systematically link atomic-scale predictions with macroscopic, application-relevant characteristics.

Most existing technologies focus on predicting fundamental properties like energies and forces, which, while important, don't always assess a model's practical value for real-world applications. Extending these models to predict downstream properties provides a better understanding of their utility and relevance to experimental outcomes (Wen et al., 2024; Kadan et al., 2024). Integrating these models into workflows that complement generative approaches enhances their utility, allowing a smoother transition from material generation to application-specific property evaluation.

To address these challenges, we propose an end-to-end workflow – AQForge – that integrates state-of-the-art generative models, MLFFs, and property prediction tools into a cohesive framework. This workflow automates material design and discovery while bridging atomic-level predictions with

macroscopic property evaluations. Currently, it uses MatterGen (Zeni et al., 2023) for structure generation, fairchem's EquiformerV2 (Liao et al., 2023) for structural relaxation, MatTen (Wen et al., 2024) for elastic tensor computation, and Phonopy (Togo, 2023) for phonon calculations. While these choices are preliminary, they can be easily expanded to more advanced or tailored models. Additionally, these models can be extended to small molecule design by integrating technologies such as cG-SchNet Gebauer et al. (2022) and IDOLpro Kadan et al. (2024). By combining and building upon existing methods, we ensure the validity of results and expand materials informatics to include downstream property predictions.

In this work, we detail the development of this workflow and demonstrate its utility through benchmarking and illustrative use cases. These examples highlight the potential of our approach to streamline materials discovery and enable the prediction of properties across multiple scales, ultimately paving the way for more targeted and efficient materials design. As part of ongoing and future work, we will also explore how this workflow can leverage downstream computed properties to guide the generative model toward a more targeted search space through active learning.

2 AN END-TO-END WORKFLOW

A schematic overview of how AQForge streamlines materials design using different state-of-the-art (SOTA) technologies is shown in Fig. 1. The process is fully automated and requires only one input file specifying which computations to run and which models to use. Currently, MatterGen is the only implemented generative model, but other models can easily be added. The structural relaxation block supports both openly available and internally developed MLFF checkpoints, as well as *ab initio* support via Quantum ESPRESSO (Giannozzi et al., 2009) and VASP Hafner (2008). Downstream properties, including elasticity, thermal conductivity, and electronic structure, can be computed with any compatible MLFF (MatTen is used for elastic tensor predictions). Phonon density of states (DOS) are computed using Phonopy. *Ab initio* methods are used to validate results and may also be used to provide feedback to the generative model.

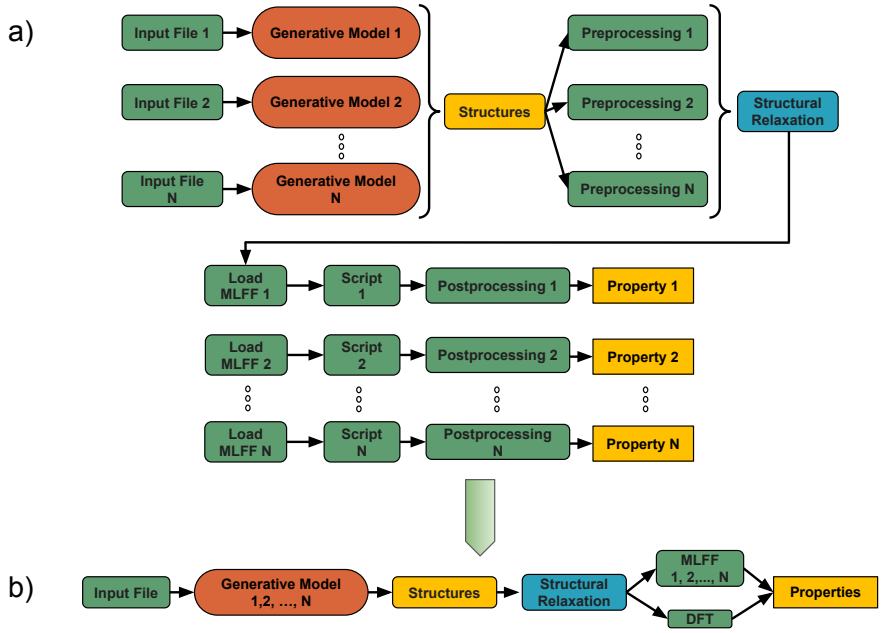


Figure 1: Schematic representation of (b) AQForge and how it reduces the I/O clutter in (a) by streamlining all tasks into a single input file.

Here, it is important to realize that downstream properties can further be used to train models for predicting macroscopic properties. For example, with sufficient data, the elastic tensor could predict yield strength of materials, extending the workflow's ability to compute macroproperties. Additionally, some generative models Zeni et al. (2023); Kadan et al. (2024) can be conditioned on

application-specific properties, enhancing the workflow's versatility. Finally, it's important to note that the goal of AQForge is not to replace existing technologies, but to enhance their capabilities by integrating them into a unified framework.

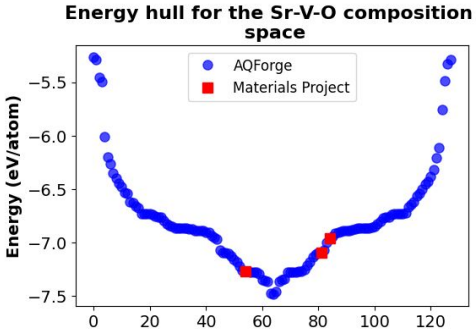
3 Example use cases

We demonstrate the utility and validity of AQForge through representative use cases. In the first, we examine two chemical spaces and compare the structures and energies against the Materials Project (MP) database (Jain et al., 2020). The structural similarity in all cases is calculated as the mean root mean squared displacement (RMSD) normalized by $[V/n]^{(1/3)}$ – where V is the volume of and n is the number of sites in the unit cell. For the second example, we follow the above procedure and additionally compute the elastic tensor and the phonon DOS, and compare it to the same benchmark.

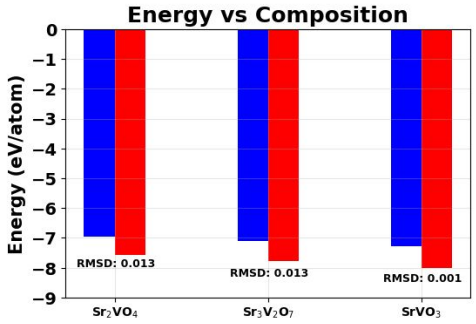
3.1 Structures and energies

We explore the chemical spaces of Sr-V-O and Li-Te, and compare the structures and energies to those reported in the MP database. The structures were generated by the base MatterGen model fine-tuned on chemical spaces, and the structural relaxation was carried out by EquiformerV2.

In fig. 2b we see that the average RMSD value for structural similarity is less than 0.01 Å and that the average error in energies is 0.421 eV/atom, demonstrating that not only does the workflow generate structures matching existing benchmark datasets but also that these structures are accurate. The errors in energies can be largely attributed to the errors reported within the MLFF, i.e., 0.316 eV/atom in this case. From the energy hull diagrams in Fig. 2, it is evident that the workflow can uncover previously unknown stable and metastable compositions. Notably, the most stable structures (Fig.2c) exhibit a 2:1 Li:Te ratio and not just the singular known Li_2Te composition. Validating and studying them using *ab initio* methods constitutes ongoing and future work.



(a) Energy hull diagram for the Sr-V-O composition space explored in this study. Each structure has less than 20 atoms.



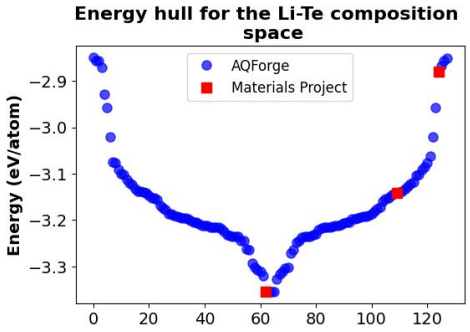
(b) Energy per atom comparison between AQForge (blue) and MP (red) structures, with normalized RMSD shown.

3.2 Elastic properties and Phonon DOS

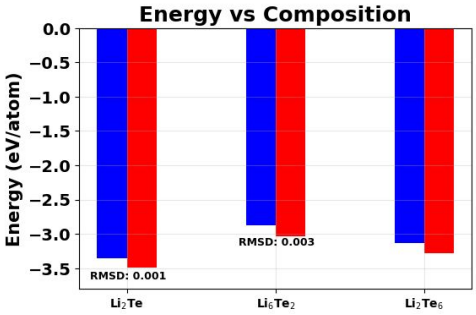
We calculate elastic properties of the above chemical systems and Al-Ti-C, and compare against MP and DFT benchmarks. Furthermore, we also compare the phonon DOS to demonstrate the versatility of properties available in AQForge. Excellent agreement is observed in the evaluation of derived elastic properties in Table 1. For a more detailed discussion on these, see Appendix A. There is also good qualitative agreement in the phonon DOS in Fig. 3, which may be improved further upon the use of finer meshes.

4 Conclusion and Future Work

In this article, we have presented an end-to-end workflow for materials property prediction and discovery. By integrating state-of-the-art machine learning models into a unified framework, our



(c) Energy hull diagram for the Li-Te composition space explored in this study. Each structure had less than 20 atoms.



(d) Energy per atom comparison between AQForge (blue) and MP (red) structures, with normalized RMSD shown.

Figure 2: An overview of the results pertaining to the chemical spaces of Sr-V-O and Li-Te. We observe in (a) and (c) that the workflow indeed re-discovers structures reported in the MP. The relative energies in (b) and (d) are consistent across the compositions with reference data.

Table 1: Comparison between the elastic properties of multiple compositions found in this study and reference values in parentheses — MP (*) and DFT (†).

Composition	Bulk Modulus (GPa)	Shear Modulus (GPa)	Poisson Ratio	RMSD (Å)
Li_2Te	25.440 (26.076*)	16.125 (18.389*)	0.238 (0.215*)	0.001*
Li_6Te_2	24.544 (25.943†)	15.747 (17.916†)	0.236 (0.219†)	0.003*
$Ti_4Al_2C_2$	136.352 (137.240*)	116.168 (115.936*)	0.168 (0.170*)	0.001*
$Ti_6Al_2C_4$	156.979 (158.241*)	126.873 (128.029*)	0.188 (0.181*)	0.002*
Sr_2VO_4	122.514 (143.501†)	68.238 (80.279†)	0.265 (0.264†)	0.015*

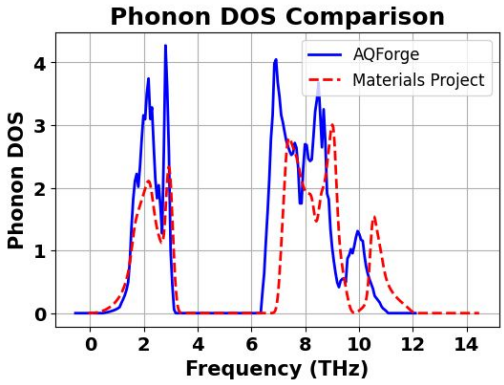


Figure 3: A comparison of the phonon DOS for Li_2Te with good agreement with baseline MP data.

approach eliminates the need for laborious independent calculations and predictions, streamlining the materials design process. Through two illustrative use cases, we demonstrate the workflow's ability to both rediscover known materials and identify novel compositions while accurately computing their properties. The workflow's capabilities can be further expanded by incorporating more advanced and diverse structure-generation models (including those for small molecules), genetic algorithms, a broader range of property prediction tools, and synthesizability metrics Antoniuk et al. (2023) to improve characterization across different application domains (see Appendix C for details). By leveraging cutting-edge machine learning techniques, we believe this workflow can extend the capabilities of existing tools and serve as a powerful tool for accelerating the discovery and design of materials and molecules with targeted properties.

REFERENCES

Kirstin Alberi, Marco Buongiorno Nardelli, Andriy Zakutayev, Lubos Mitas, Stefano Curtarolo, Anubhav Jain, Marco Fornari, Nicola Marzari, Ichiro Takeuchi, Martin L Green, et al. The 2019 materials by design roadmap. *Journal of Physics D: Applied Physics*, 52(1):013001, 2018.

Evan R Antoniuk, Gowoon Cheon, George Wang, Daniel Bernstein, William Cai, and Evan J Reed. Predicting the synthesizability of crystalline inorganic materials from the data of known material compositions. *npj Computational Materials*, 9(1):155, 2023.

Ilyes Batatia, David P Kovacs, Gregor Simm, Christoph Ortner, and Gábor Csányi. Mace: Higher order equivariant message passing neural networks for fast and accurate force fields. *Advances in Neural Information Processing Systems*, 35:11423–11436, 2022.

Chi Chen and Shyue Ping Ong. A universal graph deep learning interatomic potential for the periodic table. *Nature Computational Science*, 2(11):718–728, 2022.

Ali Davariashtiyani, Zahra Kadkhodaie, and Sara Kadkhodaei. Predicting synthesizability of crystalline materials via deep learning. *Communications Materials*, 2(1):115, 2021.

Juan J de Pablo, Nicholas E Jackson, Michael A Webb, Long-Qing Chen, Joel E Moore, Dane Morgan, Ryan Jacobs, Tresa Pollock, Darrell G Schlom, Eric S Toberer, et al. New frontiers for the materials genome initiative. *npj Computational Materials*, 5(1):41, 2019.

Bowen Deng, Peichen Zhong, KyuJung Jun, Janosh Riebesell, Kevin Han, Christopher J Bartel, and Gerbrand Ceder. Chgnet as a pretrained universal neural network potential for charge-informed atomistic modelling. *Nature Machine Intelligence*, 5(9):1031–1041, 2023.

Niklas WA Gebauer, Michael Gastegger, Stefaan SP Hessmann, Klaus-Robert Müller, and Kristof T Schütt. Inverse design of 3d molecular structures with conditional generative neural networks. *Nature communications*, 13(1):973, 2022.

Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Guido L Chiarotti, Matteo Cococcioni, Ismaila Dabo, et al. Quantum espresso: a modular and open-source software project for quantum simulations of materials. *Journal of physics: Condensed matter*, 21(39):395502, 2009.

Rafael Gómez-Bombarelli, Jennifer N Wei, David Duvenaud, José Miguel Hernández-Lobato, Benjamín Sánchez-Lengeling, Dennis Sheberla, Jorge Aguilera-Iparraguirre, Timothy D Hirzel, Ryan P Adams, and Alán Aspuru-Guzik. Automatic chemical design using a data-driven continuous representation of molecules. *ACS central science*, 4(2):268–276, 2018.

Jürgen Hafner. Ab-initio simulations of materials using vasp: Density-functional theory and beyond. *Journal of computational chemistry*, 29(13):2044–2078, 2008.

Anubhav Jain, Joseph Montoya, Shyam Dwaraknath, Nils ER Zimmermann, John Dagdelen, Matthew Horton, Patrick Huck, Donny Winston, Shreyas Cholia, Shyue Ping Ong, et al. The materials project: Accelerating materials design through theory-driven data and tools. *Handbook of Materials Modeling: Methods: Theory and Modeling*, pp. 1751–1784, 2020.

Jidon Jang, Geun Ho Gu, Juhwan Noh, Juhwan Kim, and Yousung Jung. Structure-based synthesizability prediction of crystals using partially supervised learning. *Journal of the American Chemical Society*, 142(44):18836–18843, 2020.

Rui Jiao, Wenbing Huang, Peijia Lin, Jiaqi Han, Pin Chen, Yutong Lu, and Yang Liu. Crystal structure prediction by joint equivariant diffusion. *Advances in Neural Information Processing Systems*, 36:17464–17497, 2023.

Hongni Jin and Kenneth M Merz Jr. Liganddiff: de novo ligand design for 3d transition metal complexes with diffusion models. *Journal of Chemical Theory and Computation*, 2024.

Amit Kadan, Kevin Ryczko, Erika Lloyd, Adrian Roitberg, and Takeshi Yamazaki. Guided multi-objective generative ai to enhance structure-based drug design. *arXiv preprint arXiv:2405.11785*, 2024.

Andrew Lee, Suchismita Sarker, James E Saal, Logan Ward, Christopher Borg, Apurva Mehta, and Christopher Wolverton. Machine learned synthesizability predictions aided by density functional theory. *Communications Materials*, 3(1):73, 2022.

Yi-Lun Liao, Brandon Wood, Abhishek Das, and Tess Smidt. Equiformerv2: Improved equivariant transformer for scaling to higher-degree representations. *arXiv preprint arXiv:2306.12059*, 2023.

Ahmed I Osman, Mahmoud Hefny, MIA Abdel Maksoud, Ahmed M Elgarahy, and David W Rooney. Recent advances in carbon capture storage and utilisation technologies: a review. *Environmental Chemistry Letters*, 19(2):797–849, 2021.

Susan B Sinnott. Material design and discovery with computational materials science. *Journal of Vacuum Science & Technology A*, 31(5), 2013.

Zhilong Song, Shuaihua Lu, Minggang Ju, Qionghua Zhou, and Jinlan Wang. Is large language model all you need to predict the synthesizability and precursors of crystal structures? *arXiv preprint arXiv:2407.07016*, 2024.

Atsushi Togo. First-principles phonon calculations with phonopy and phono3py. *Journal of the Physical Society of Japan*, 92(1):012001, 2023.

Mingjian Wen, Matthew K Horton, Jason M Munro, Patrick Huck, and Kristin A Persson. An equivariant graph neural network for the elasticity tensors of all seven crystal systems. *Digital Discovery*, 3(5):869–882, 2024.

Tian Xie, Xiang Fu, Octavian-Eugen Ganea, Regina Barzilay, and Tommi Jaakkola. Crystal diffusion variational autoencoder for periodic material generation. *arXiv preprint arXiv:2110.06197*, 2021.

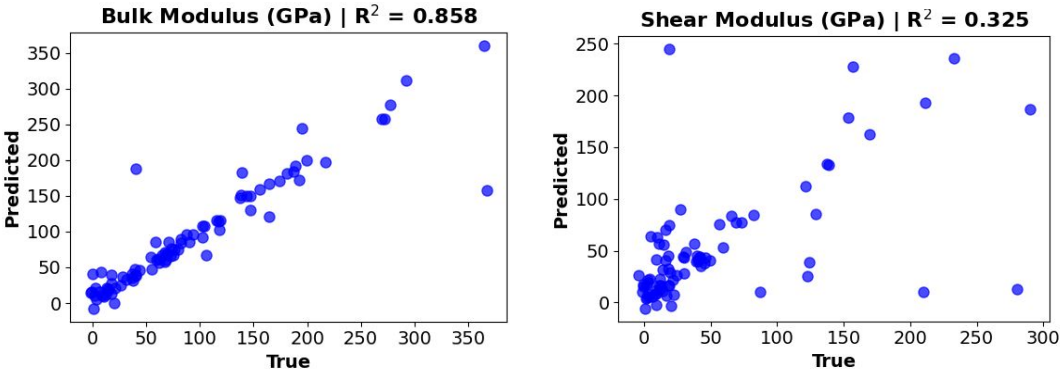
Claudio Zeni, Robert Pinsler, Daniel Zügner, Andrew Fowler, Matthew Horton, Xiang Fu, Sasha Shysheya, Jonathan Crabbé, Lixin Sun, Jake Smith, et al. Mattergen: a generative model for inorganic materials design. *arXiv preprint arXiv:2312.03687*, 2023.

Qing Zhao, Sanjuna Stalin, Chen-Zi Zhao, and Lynden A Archer. Designing solid-state electrolytes for safe, energy-dense batteries. *Nature Reviews Materials*, 5(3):229–252, 2020.

Zhi-Jian Zhao, Sihang Liu, Shenjun Zha, Dongfang Cheng, Felix Studt, Graeme Henkelman, and Jinlong Gong. Theory-guided design of catalytic materials using scaling relationships and reactivity descriptors. *Nature Reviews Materials*, 4(12):792–804, 2019.

A Benchmarking Elastic Properties

Here, we provide an overview of the benchmarking results of using AQForge to predict the bulk moduli and shear moduli of material compositions. A total of 187 compositions were considered, with the ground truth coming from MP data. For every composition, 25 structures were generated using DiffCSP Jiao et al. (2023). We used DiffCSP instead of MatterGen because the data contained known formulae which is more suitable for a model geared towards crystal structure prediction (CSP) rather than generic materials discovery. However, retraining MatterGen for a CSP task and redoing this benchmarking constitutes ongoing and future work. Each structure was then relaxed using FAIR-Chem's EquirformerV2 (note that this model has been trained on the test set), and their elastic properties were computed using MatTen. The results can be seen in Fig. 4.



(a) Predictions for the bulk moduli using AQForge

(b) Predictions for the shear moduli using AQForge

Figure 4: An overview of the results of predicting the bulk and shear moduli of a set of material compositions

The mean absolute error (MAE) for bulk moduli and shear moduli predictions were 13.37 GPa and 27.08 GPa, respectively. While the bulk modulus MAE is close to the values reported in Wen et al. (2024) and Zeni et al. (2023), the errors in the shear moduli are, on average, worse by a factor of 2. However, on closer inspection, one can also observe that the errors for the shear moduli predictions are high only because of a handful data points. This demonstrates both a utility and a deficiency in the current state of AQForge, which can also be seen in the results for Sr_2VO_4 in Table 1. To address this, one can use generative models and MLFFs fine-tuned on elastic tensor data. Another avenue of improvement being explored is the computation of elastic tensors through *ab initio* methods, or a mix of *ab initio* and ML-based methods. We expect to see noticeable improvements as these methodologies are integrated within AQForge.

B Run-time Comparisons

Here, we provide a brief comparison in the computational wall-time needed to compute the elastic properties of a given structure. Note that while this comparison has been done only for elastic properties, we expect to see similar improvements for the calculation of other properties as well. Additionally, the results below do not include the resources used to generate structures, and assume that relevant structures are already saved. The results have been consolidated in Fig. 5.

While it is no surprise that the use of MLFFs considerably reduces the wall-time over traditional density functional theory (DFT) calculations, they key advantage of AQForge comes in removing any postprocessing requirements and alleviating the need to convert files from one format to another, or to prepare different input scripts or files for each MLFF. Moreover, AQForge allows the user to specify a different MLFF for each step, if needed, or to execute the same step with different MLFFs, if desired. As we continue to develop this framework further by integrating more models and a wider range of property prediction tasks, the reduction in total wall-time from end-to-end will only become more pronounced.

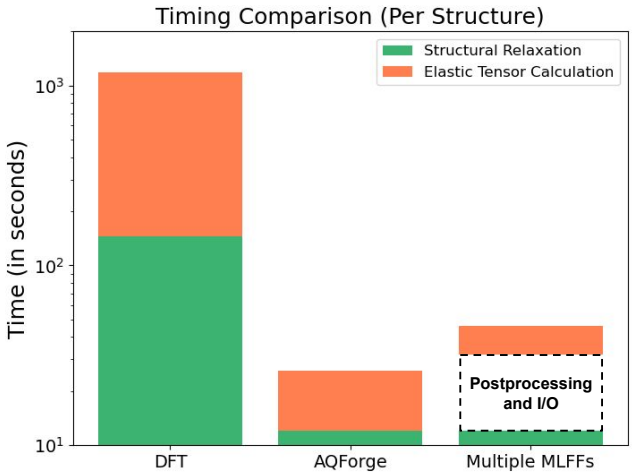


Figure 5: Average wall-time comparison for structural relaxation and the calculation of the elastic tensor for over 100 structures.

C Ongoing and Future Work

As part of ongoing development work for AQForge, we summarize here briefly the current plan to expand its capabilities in Fig. 6. We expect to actively work on improving this framework, and the development plan shown below reflects only the current state.

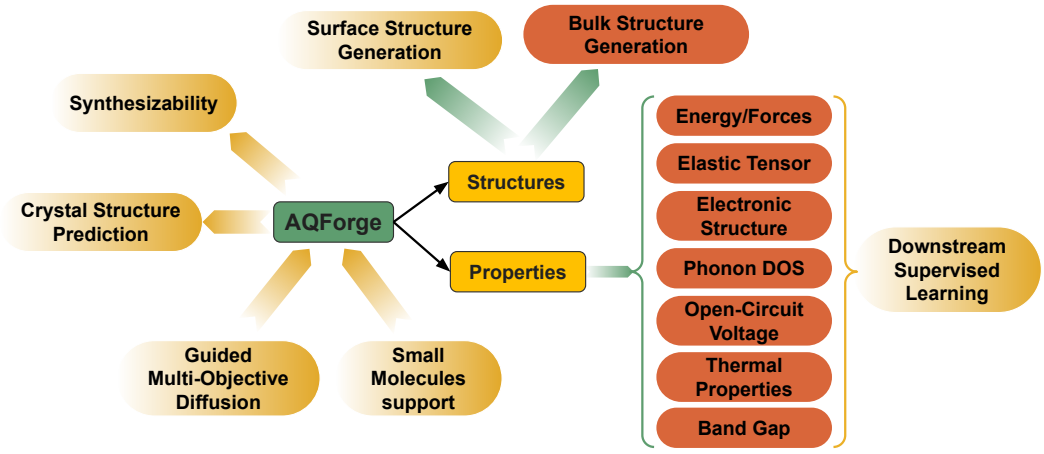


Figure 6: Ongoing and planned development work for AQForge. The orange bubbles have been fully implemented, whereas the shaded yellow bubbles constitute ongoing and future work. The road map may be expanded further.

In summary, one of the main improvements would be to integrate molecular frameworks into AQForge in addition to bulk crystalline materials. Generative models such as IDOLpro Kadan et al. (2024) and proprietary molecular property prediction models developed at SandboxAQ will form the bulk of this branch of improvement. Another area of active development is material synthesizability, which has become a rich avenue of research in recent years Lee et al. (2022); Jang et al. (2020); Davariashtiyani et al. (2021); Song et al. (2024). CSP along with surface generation and characterization form integral parts of upcoming AQForge capabilities and are expected to have significant impacts in fields such as catalysis, surface reactions, and corrosion modeling. Finally, implementing guided multi-objective diffusion within the generative models used in AQForge will

allow users to discover new materials with desired properties much more cheaply and without the need to fine-tune or retrain traditional models.

While the work above points to long-term improvements, something that may be more short-term in nature is adding additional supervised learning models that take in AQForge-predicted properties as features to predict macroscopic properties of interest, as briefly discussed in Section 2. With these, we strongly feel that AQForge has the capacity to streamline end-to-end materials generation and characterization workflows, while giving flexibility to the users to modify it as desired.