Large Scale Benchmark of Materials Design Methods

Kamal Choudhary[1,a] Daniel Wines[1] Kangming Li[2] Kevin F. Garrity[1] Vishu Gupta[3] Aldo H. Romero[4] Jaron T. Krogel[5] Kayahan Saritas[5] Addis Fuhr[6] Panchapakesan Ganesh[6] Paul R. C. Kent[7] Keqiang Yan[8] Yuchao Lin[8] Shuiwang Ji[8] Ben Blaiszik[9] Patrick Reiser[10] Pascal Friederich[11,10] Ankit Agrawal[3] Pratyush Tiwary[12] Eric Beyerle[12] Peter Minch[13] Trevor David Rhone[13] Ichiro Takeuchi[14] Robert B. Wexler[15] Arun Mannodi-Kanakkithodi[16] Elif Ertekin[17,18] Avanish Mishra[19] Nithin Mathew[19] Sterling G. Baird[20] Mitchell Wood[21] Andrew Dale Rohskopf[21] Jason Hattrick-Simpers[2] Shih-Han Wang[22] Luke E. K. Achenie[22] Hongliang Xin[22] Maureen Williams[1] Adam J. Biacchi[23] and Francesca Tavazza[1]

[1)]*Material Measurement Laboratory, National Institute of Standards and Technology, Maryland, 20899, USA.*

[2)]*Department of Materials Science and Engineering, University of Toronto, 27 King's College Cir, Toronto, ON, Canada.*

[3)]*Department of Electrical and Computer Engineering, Northwestern University, Evanston, Illinois, 60208, USA*

[4)]*Department of Physics and Astronomy, West Virginia University, Morgantown, WV 26506, USA*

[5)]*Materials Science and Technology Division, Oak Ridge National Laboratory, Oak Ridge,Tennessee 37831, USA*

[6)]*Center for Nanophase Materials Science, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, United States*

[7)]*Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA*

[8)]*Department of Computer Science and Engineering, Texas A&M University, College Station, TX 77843, USA*

[9)]*Globus, University of Chicago, Illinois, 60637, USA. Data Science and Learning Division, Argonne National Lab, Illinois, 60439, USA.*

[10)]*Institute of Nanotechnology, Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany*

[11)]*Institute of Theoretical Informatics, Karlsruhe Institute of Technology, Engler-Bunte-Ring 8, 76131 Karlsruhe, Germany*

[12)]*Department of Chemistry and Biochemistry and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, United States*

[13)]*Department of Physics, Applied Physics and Astronomy, Rensselaer Polytechnic Institute, Troy, NY 12180, USA*

[14)]*Department of Materials Science and Engineering, University of Maryland, College Park, MD 20742, USA*

[15)]*Department of Chemistry and Institute of Materials Science and Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA*

[16)]*School of Materials Engineering, Purdue University, West Lafayette, IN, 47907, USA*

[17)]*Department of Mechanical Science and Engineering, University of Illinois Urbana-Champaign, Urbana, Illinois 61801, USA*

[18)]*Materials Research Laboratory, University of Illinois Urbana-Champaign, Urbana, Illinois 61801, USA*

[19)]*Theoretical Division (T-1), Los Alamos National Laboratory, Los Alamos, NM, 87545, USA*

[20)]*Department of Materials Science and Engineering, University of Utah, Salt Lake City, UT 84108, USA*

[21)]*Center for Computing Research, Sandia National Laboratories, Albuquerque, New Mexico 87185, USA*

[22)]*Department of Chemical Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA*

[23)]*Physical Measurement Laboratory, National Institute of Standards and Technology, Maryland, 20899, USA.*

Lack of rigorous reproducibility and validation are major hurdles for scientific development across many fields. Materials science in particular encompasses a variety of experimental and theoretical approaches that require careful benchmarking. Leaderboard efforts have been developed previously to mitigate these issues. However, a comprehensive comparison and benchmarking on an integrated platform with multiple data modalities with both perfect and defect materials data is still lacking. This work introduces JARVIS-Leaderboard, an open-source and community-driven platform that facilitates benchmarking and enhances reproducibility. The platform allows users to set up benchmarks with custom tasks and enables contributions in the form of dataset, code, and meta-data submissions. We cover the following

materials design categories: Artificial Intelligence (AI), Electronic Structure (ES), Force-fields (FF), Quantum Computation (QC) and Experiments (EXP). For AI, we cover several types of input data, including atomic structures, atomistic images, spectra, and text. For ES, we consider multiple ES approaches, software packages, pseudopotentials, materials, and properties, comparing results to experiment. For FF, we compare multiple approaches for material property predictions. For QC, we benchmark Hamiltonian simulations using various quantum algorithms and circuits. Finally, for experiments, we use the inter-laboratory approach to establish benchmarks. There are 1281 contributions to 274 benchmarks using 152 methods with more than 8 million data-points, and the leaderboard is continuously expanding. The JARVIS-Leaderboard is available at the website: https://pages.nist.gov/jarvis_leaderboard/

I. INTRODUCTION

The accelerated design and characterization of materials of technological interest has been a rapidly evolving area of research in the last few decades[1]. Materials design requires approaches spanning a variety of length and time scales[2], processes, and operating conditions. For atomistic design, the methods employed may include computational approaches such as density functional theory (DFT), tight-binding (TB), force-field methods (FF), and highly accurate approaches such as quantum Monte Carlo (QMC) or quantum computation (QC). A wide range of approaches are employed above the purely atomistic level, i.e., phase-field.[3] Experimental characterization approaches include X-ray diffraction (XRD), vibroscopy, manometry, scanning probe and electron microscopy, and magnetic susceptibility measurements. Moreover, data produced from materials design techniques can be of various types: atomic/micro-structures, images, spectra, and text-documents[4–6]. The data analysis and curation methods add further complexity to benchmarking efforts.

Currently, many materials design approaches are used for scientific and industrial applications. Although there have been significant advances in each of these approaches, there is an urgent need to establish a large-scale benchmark for systematic, reproducible, transparent, and unbiased scientific development. Developing such metrology is a highly challenging task, even for one of these methods, let alone the entire galaxy of available methods. For example, only 1.5 % of CO_2 adsorption for the metal-organic framework (MOF) experiments were found to be reproducible[7]. Projects and approaches such as the materials genome and FAIR initiatives[1,8], have resulted in several well-curated datasets and benchmarks. These, in turn, have led to several materials informatics applications[9–11]. Although electronic structure approaches such as DFT tend to be more reproducible than other categories[12], a systematic effort must be made to validate these methods and estimate the error in predictions. Hence, it is highly desirable to have a large-scale benchmarking platform in the materials science field for reproducibility and method validation.

Massive progress in fields such as image recognition/image classification (ImageNet[13]), protein structure prediction (AlphaFold[14]), large language modeling (Generative pretrained transformers (GPT))[15]) has been possible primarily because of well-defined benchmarks in respective fields. Such benchmarks allow a wide community to solve problems collectively and systematically. We believe that such a universal and large-scale set of benchmarks for materials science will significantly benefit the scientific community.

To this date, several benchmarks of individual methods have already been developed. For AI methods, there have been several benchmarks and leaderboards such as MatBench[16], MoleculeNet[17], sGDML[18,19], mLEARN[20], MatScholar[21], and AtomAI[22]. For electronic structure methods, some of the notable benchmarks include the work by Lejaeghere et al.[12], Borlido et al.[23], Huber et al.[24], Zhang et al.[25], Tran et al.[26] and several other projects[27–30]. Other method benchmarks include phase-field benchmarks by Wheeler et al.[31], Lindsay et al.[32], and microscopy benchmarks such as by Wei et al.[33]. More details on some of these benchmarking efforts are provided in later sections.

In this work, we present a user-friendly, comprehensive approach to integrate the benchmarking of both computational and data-analytics methods. The JARVIS-Leaderboard framework (https://pages.nist.gov/jarvis_leaderboard/) covers a variety of categories: Artificial Intelligence (AI), Electronic Structure (ES), Force-field (FF), Quantum Computation (QC), and Experiments (EXP). It also covers various data types, including atomic structures, spectra, images, and text. This project can be used to: (1) check the state-of-the-art methods in respective fields, (2) add a contribution model on an existing benchmark, (3) add a new benchmark, (4) compare new ideas and approaches to well-known approaches. To enhance reproducibility, we encourage each contribution to (1) be from peer-reviewed articles with an associated DOI for all contributions, models, and tools, (2) include a run script to exactly reproduce the results (especially for computational tools), (3) include a metadata file with details such as team name, contact information, computational timing and software (with software version)/hardware used in order to enhance transparency.

Some of the key distinguishing factors between a usual large data-repository (such as JARVIS-DFT) and the present leaderboard effort are: 1) the leaderboard contains well-characterized/well-known samples/tasks (i.e., with digital object identifier/peer-reviewed article links) with all the scripts/metadata easily available to reproduce the results rather than just being a look-up table to find data, 2) large data repositories usually contain more variation in materials chemistry/structure and less variation of methods while the leaderboard focuses on more number method comparisons. For example, the JARVIS-DFT contains DFT data for more

[a)]Electronic mail: kamal.choudhary@nist.gov

than 80,000 materials and millions of material properties with a few specific ES methods and hence there are only a few entries for, say, the electronic bandgap of Silicon, while the leaderboard contains electronic bandgaps for Silicon using more than 17 ES methods from various contributors. Similarly, JARVIS-ALIGNN project contains AI models for more than 80 properties/tasks of materials, i.e., just one model for a well-known property such as formation energy, while there are more than 12 methods for formation energy task in the leaderboard (as discussed later).

The leaderboard can also be used to identify and focus on some of the challenges in different fields, such as (1) why is it difficult to develop a reasonably good AI model with similar accuracy to electronic structure methods?, (2) how can we reduce the computational cost of higher accuracy electronic structure predictions (i.e. bandgap predictions)?, (3) how do we identify examples of materials that require high-fidelity methods (beyond DFT accuracy) to understand their properties?, (4) how can we estimate errors across methods in property predictions and in turn identify material space where methodological improvements need to be targeted?, (5) how can we establish figures of merits for mesoscale models such as phase field?, (6) how can we make atomistic image analysis quantitative rather than qualitative?, (7) and how do we develop and benchmark multi-modal models[34]?

The JARVIS-Leaderboard is seamlessly integrated into the existing and well-established NIST-JARVIS infrastructure[35], which hosts several datasets, tools, applications, and tutorials for materials design, motivated by the materials genome initiative[1]. The framework is open access to the entire materials science community for progressing the field collectively and systematically. JARVIS (Joint Automated Repository for Various Integrated Simulations)[35] is a repository designed to automate materials discovery and optimization using classical force-field, density functional theory, machine learning calculations, and experiments. Since its creation in 2017, JARVIS has had over 10,000 users worldwide, over 40 JARVIS-associated articles have been published, and over 80,000 materials currently reside in the database. As these numbers continue to multiply, significant effort on external outreach to the materials science community has been an additional goal of JARVIS, with several events (https://jarvis.nist.gov/events/) such as the Artificial Intelligence for Materials Science (AIMS) and Quantum Matters in Materials Science (QMMS) workshops and hands-on JARVIS-Schools, which have had hundreds of participants throughout the last few years. Based on the level of success and support from the community with regard to the existing JARVIS infrastructure, we believe that the integration of the JARVIS-Leaderboard will have a similar level of engagement and success, with a growing number of contributors from all over the world (in government, academia and industry) and in different sub-fields of materials science.

II. LEADERBOARD OVERVIEW

Fig. 1 gives an example of the leaderboard interface. At the homepage, information regarding the number of methods, benchmarks, contributions, and datapoints are provided. Clicking on one of the entries (or searching in the 'Search' box) such as "formation_energy_peratom" opens a new tab with available contributions. For each contribution, links are provided to the submitted data (in .csv.zip format), reference benchmark data (in JSON file), a shell script to reproduce the contribution (run.sh file) and metadata file (metadata.json). The metadata file contains details about the team name, DOI number, software (with software version), hardware, instrument, computational timing and other relevant details of a benchmark. There are several categories for the benchmarks including AI, ES, QC, FF and EXP and their combinations (Fig. 1c). Some example contributions and a summary table are also provided on the webpage to help a user navigate through the project. The summary table breaks down the available information into categories and sub-categories of different methodologies.

JARVIS-Leaderboard is an evolving project, so additions to the project are anticipated, welcome, and easy to make. We show a general flowchart for adding a new benchmark to the leaderboard in Fig. 2. The user can populate the reference dataset (with well-defined data splits) used for a specific benchmark (e.g. for 2D exfoliation energies in JARVIS-DFT dataset using an AI method: "AI-SinglePropertyPrediction-exfoliation_energy-dft_3d-test"). AI benchmarks have predefined training/validation/test identifiers and target data in a corresponding json.zip file, while other methods have only reference test set for evaluation because they do not require model training like an AI method does. For most benchmarks in the leaderboard, experimental data is used as the reference data.

There is a helper script jarvis_populate_data.py to generate a benchmark dataset. A user can apply their method, train models, or run experiments on that dataset and prepare a csv.zip file, a metadata.json file, and also if possible, a Dockerfile and run.sh file. This step helps to reproduce the benchmark. These files are kept in a folder with the name of the folder as the team name and can be uploaded to a user's GitHub account by the automated jarvis_upload.py script. This script automatically forks the parent usnistgov/jarvis_leaderboard repo for the user, adds the team-name folder with its files in that forked repo, runs a few minimal sanity-checks on the new contribution, and then makes a pull request to the parent repo. The contribution addition and automated testings are carried out using GitHub actions. The administrators of the JARVIS-Leaderboard at NIST will verify the contributions and then finally, it will become part of the leaderboard website.

This project is available on GitHub at: https://github.com/usnistgov/jarvis_leaderboard. The administrators of the JARVIS-Leaderboard at NIST will fully oversee the upload of contributions and benchmarks. A tree structure of the repo is shown in Fig. 3. There are two main directories in the repo: (1) benchmarks (reference) and (2) leaderboard

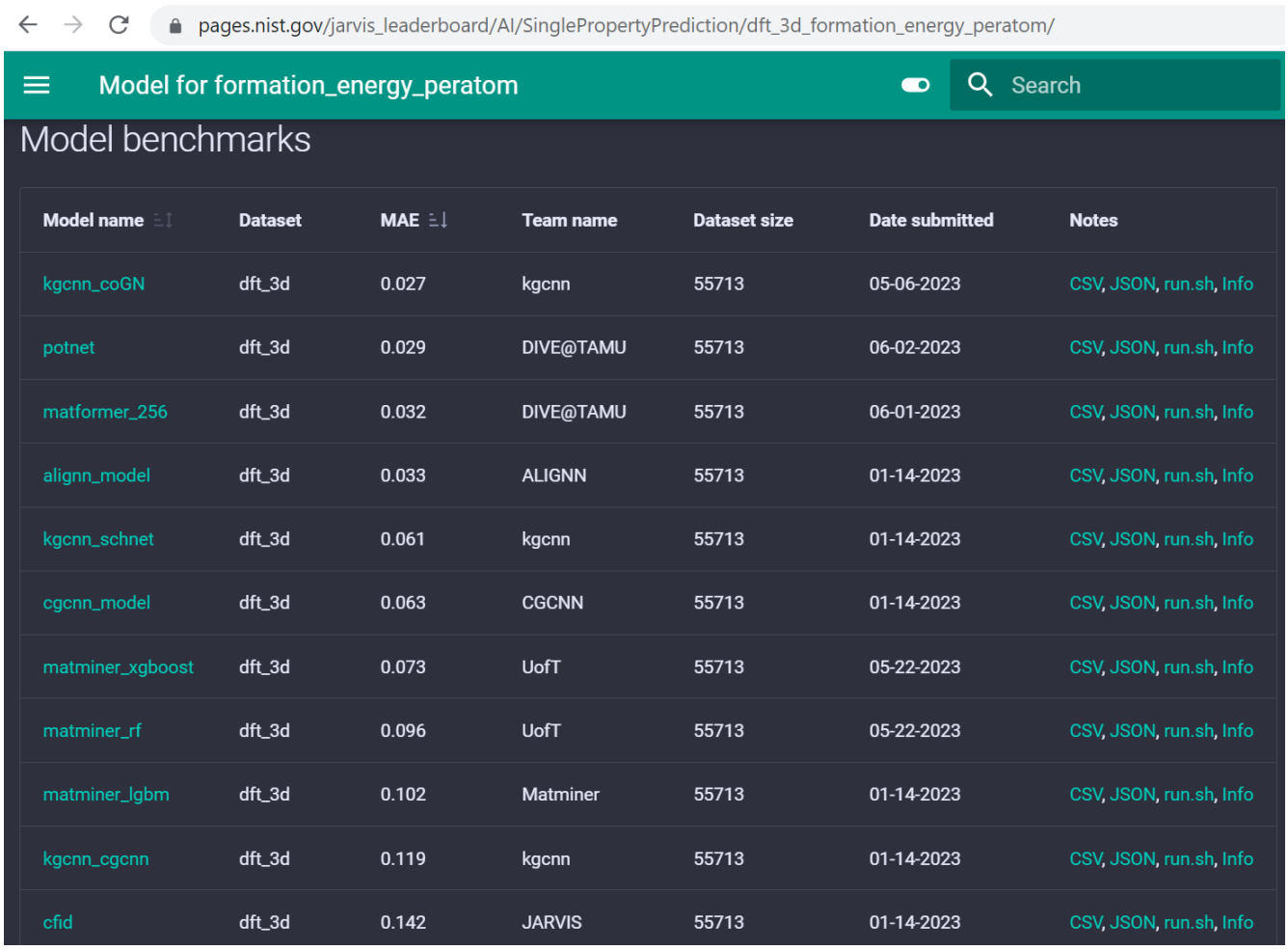


FIG. 1. Leaderboard snapshot showing an example output for AI based formation energy per atom model on the JARVIS-DFT (dft_3d) dataset. The benchmark has seven contributions so far and they are sorted based on the mean absolute error (MAE) values. Lower MAE values indicate higher accuracy. Links to individual csv.zip (AI-SinglePropertyPrediction-formation_energy_peratom-dft_3d-test-mae.csv.zip), json.zip (dft_3d_formation_energy_peratom.json.zip), shell script (run.sh) and detailed info (metadata.json) files are provided to help enhance reproducibility. Such results tables are available for each benchmark in the leaderbord.

contributions (for various leaderboard entries), as shown by the green highlighted boxes in Fig. 3.

The "benchmarks" directory has folders for the AI, ES, QC, FF, and EXP categories. Within them, there are sub-folders for specific sub-categories such as (1) SinglePropertyPrediction (where the output of a model/experiment is one single number for an entry), (2) SinglePropertyClass (where the output is class-ids, i.e., 0,1,.. instead of floating values), (3) ImageClass (for multi-class image classification), (4) textClass (for multi-label text classification), (5) MLFF (machine learning force-field), (6) Spectra (for multi-value data) and (7) EigenSolver (for Hamiltonian simulation). In each of these sub-folders, there are .json.zip files with well-defined reference datasets and available properties as also available in the JARVIS-Tools package https://jarvis-tools.readthedocs.io/en/master/databases.html. To avoid storage of large files in the GitHub repo, the actual datasets are part of JARVIS-Tools and are stored in the Figshare repository with specific DOIs and version numbers.

Next, in the "contributions" directory, there is a collection of folders that consist of .csv.zip, metadata.json files, and optionally a Dockerfile and run.sh file. The csv.zip file contains identifier (id) entries and corresponding prediction values obtained by the corresponding model/method. These test identifiers (such as JVASP-1408 in Fig. (3)) must match the test set IDs in the json.zip file in the benchmarks folder for the metric measurements to work. Each of the csv.zip files must contain six components in the filename to place the contribution in the appropriate webpage. The components are the categories (such as AI), sub-categories (such as ImageClass), property (such as bravais_lattice), dataset-name (such as stem_2d_image as available in the JARVIS-Tools database page), and data-split. For entries in the AI category, the data is in train-validation-test splits (using a fixed random number generator). For the current leaderboard format, we report the performance accuracy in the test set only. These files can

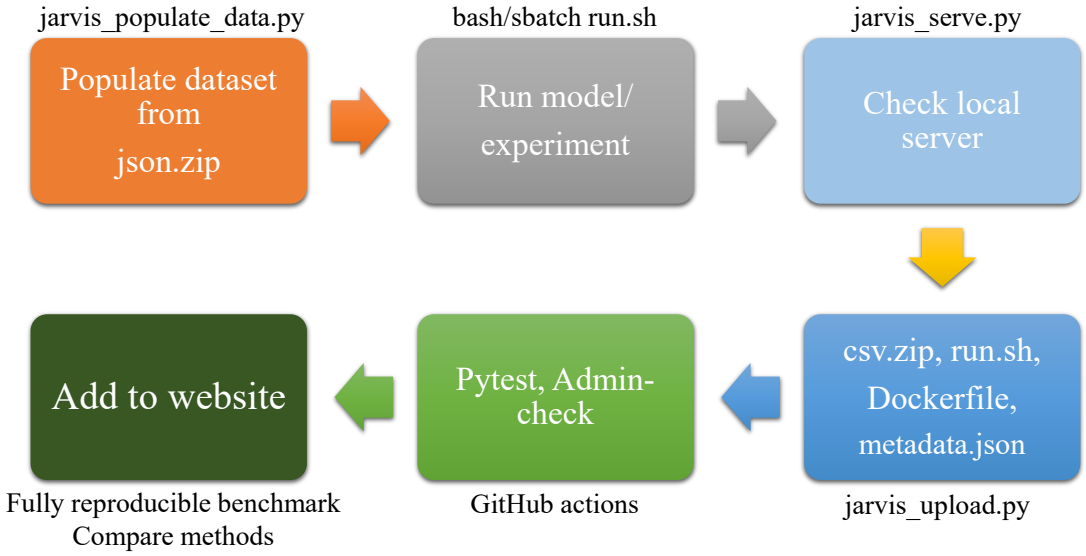


FIG. 2. A flow-chart showing the processes involved in uploading a new contribution to the leaderbaord. The jarvis_populate_data.py scripts generate a benchmark dataset. A user can apply their method, train models, or run experiments on that dataset and prepare a csv.zip, a metadata.json file, and other files in a new folder in the contributions directory. The contributions can be locally checked by the user using jarvis_server.py script. Then the folder can be uploaded to a user's GitHub account by the automated jarvis_upload.py script involving several GitHub uploading steps. The administrators of the JARVIS-Leaderboard at NIST will verify the contributions and then finally, it will become part of the leaderboard website.

be easily edited with common text editors. Each contribution folder (e.g. alignn-model) consists of one or several csv.zip files corresponding to each benchmark (such as for formation energies, bandgap, etc.).

Model-specific details are kept in the metadata.json file with *required* keys such as model_name, project_url, team_name and an email address. Users can keep other data such as the uncertainty, time taken, and instrument/software/hardware used in the metadata file as well. For computational models, the run.sh script can be used to reproduce the contributions completely as a single command line script or job submission script. If a method requires additional steps or details beyond a simple command line script, a user can upload a README file containing the additional details. For enhanced reproducibility, we also optionally allow users to include a Dockerfile and an ipython/Google-colab notebook for each benchmark. These notebooks can be used to run the contributions in the Google-cloud without downloading anything locally.

In addition, there is a "docs" directory in the JARVIS-leaderboard. The docs folder consists of a directory structure that is similar to the benchmarks folder with categories names (AI, ES, etc.), and sub-categories (such as SinglePropertyPrediction, ImageClass etc.) with markdown (.md) files that will be converted automatically into corresponding html pages for the website. For each benchmark (i.e., json.zip file), a corresponding docs entry (i.e., md file) should be present. A new benchmark must be associated with a peer-reviewed article and a DOI, in order to have trust in the reference benchmark data. A new benchmark must also be verified by the JARVIS-Leaderboard administrators.

As mentioned above, there already exist several other materials science-specific benchmarks. We compare some of these benchmarks in Table I based on the categories that are included. We find that there is no single, large-scale benchmark encompassing the various fields as in the JARVIS-Leaderboard. Also, the data format, metadata, and website for these different leaderboards vary significantly. Hence, having a uniform way to compare different methods would greatly help the materials community.

III. METHODOLOGY

The JARVIS-Leaderboard aims to provide a comprehensive framework covering a variety of length and time-scale approaches[2] to enable realistic materials design. In this section, we provide a brief overview of the methods that are currently available in the leaderboard. In this work we use the terms categories, sub-categories, methods, benchmarks, and

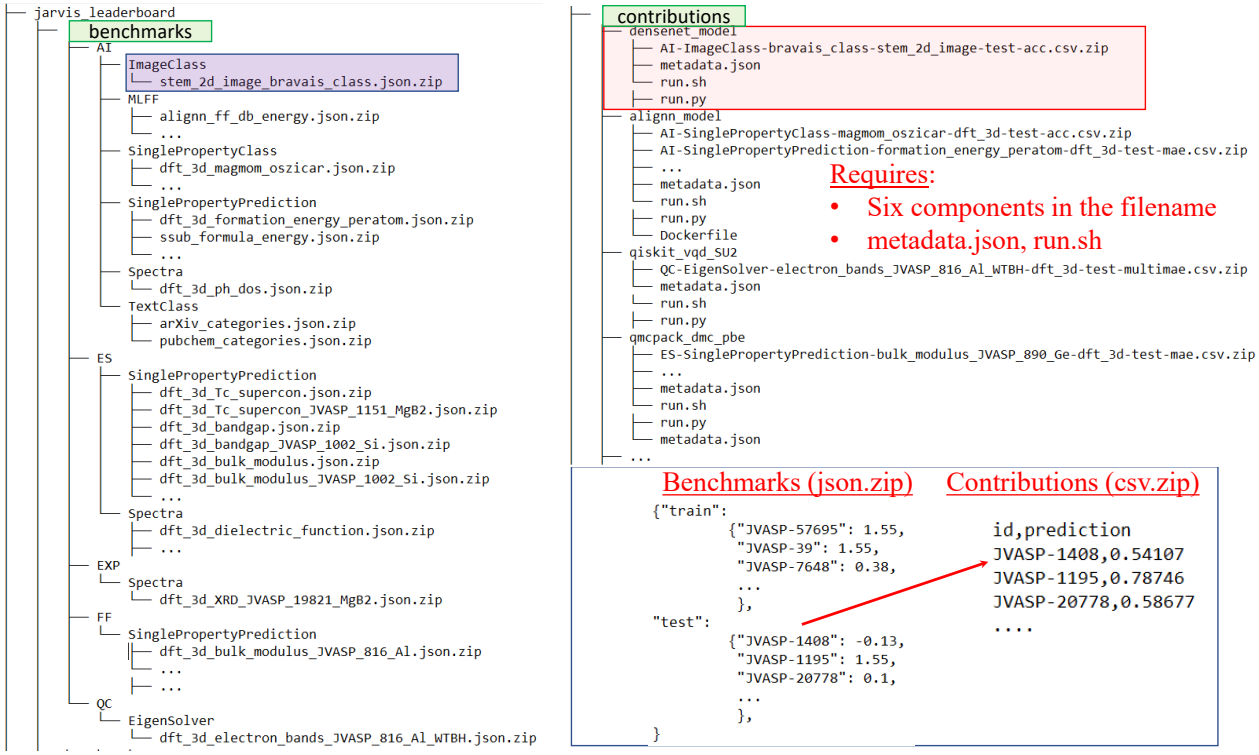


FIG. 3. A tree diagram for directory and file-structure in the leaderboard. There are two main directories in the repo: (1) benchmarks (reference) and (2) leaderboard contributions (for various leaderboard entries). In the "benchmarks" directory, there are folders for the AI, ES, QC, FF, and EXP categories. Within them, there are sub-folders for specific sub-categories. In the "contributions" directory there is a collection of folders that consists of .csv.zip, metadata.json files, and optionally a Dockerfile and run.sh file for available contributions from each method. The csv.zip file contains entries of identifier (id) and corresponding prediction values as obtained by the corresponding model/method. These test identifiers (such as JVASP-1408) must match the test set ids in the json.zip file in the benchmarks folder for the metric measurements to work.

contributions often, so we define them as follows.

Currently, there are five main "categories" in the leaderboard: Artificial Intelligence (AI), Electronic Structure (ES), Force-field (FF), Quantum Computation (QC), and Experiments (EXP). Each category is divided into "sub-categories", a list of which is provided on the website. These sub-categories include single-property-prediction, single-property-classification, atomic force prediction, text classification, text-token classification, text generation, image classification, image segmentation, image generation, spectra-prediction, and eigensolver. These sub-categories are highly flexible and new categories can be easily added. "Benchmarks" are the reference data (in the form of json.zip file, discussed later) used to calculate performance metrics for each specific contribution. "Methods" are a set of precise specifications for evaluation against a benchmark. For example, within the ES category, density functional theory (DFT) performed with the specifications of the Vienna Ab initio Simulation Package (VASP)[52,53], Perdew-Burke-Ernzerhof (PBE)[54] functional and PAW[52,53] pseudopotentials (VASP-PBE-PAW) is a method. Similarly, within the AI category, descriptor/feature-based models with specifications of MatMiner[55] chemical features and the LightGBM[56]software is a method. "Contributions" are individual data (in the form of csv.zip files) for each benchmark computed with a specific method. Each contribution files consist of six components: category (e.g. AI), sub-category (e.g. SinglePropertyPrediction), property (e.g. formation energy), dataset (e.g. dft_3d), data-split (e.g. test), metric (e.g. mae).

A. Electronic structure

Electronic structure approaches cover short length scales and short time scales with high-fidelity. There are a variety of ES methodologies such as such as tight-binding[57–59], density functional theory (DFT)[60], quantum Monte Carlo[61], dynamical mean field theory[62] and many-body perturbation theory (Green's function with screened Coulomb potential, GW methods)[63]. For each of the methodologies, there are a number of specifications to completely describe a method including the exact software, exchange-correlation functional, pseudopotential, and other relevant parameters. Example methods used in this work are given in Table II.

Each method in the ES category can have a variety of contributions. For example, using a specific method, one can calculate various properties such as bandgaps, formation energies, bulk moduli, solar cell efficiencies, and

TABLE I. Comparison of benchmark infrastructure available for materials design methods for several categories.

Projects	AI	ES	FF	QC	EXP
MoleculeNet[17]	✓	-	-	-	-
MatBench[16]	✓	-	-	-	-
OpenCatalystProject[36]	✓	-	-	-	-
SciML[37]	✓	-	-	-	-
SGDML[19]	✓	-	-	-	-
GuacaMol[38]	✓	-	-	-	-
Alchemy[39]	✓	-	-	-	-
ML4Chem[40]	✓	-	-	-	-
DGL-LifeSci[41]	✓	-	-	-	-
CCCBDB[42]	-	✓	-	-	✓
Delta-DFT[12]	-	✓	-	-	-
SSSP[43]	-	✓	-	-	-
OpenKIM[44]	-	-	✓	-	-
IPR[45]	-	-	✓	-	-
JARVIS-FF[46]	-	-	✓	-	-
Mlearn[20]	-	-	✓	-	-
QuantumVolume[47]	-	-	-	✓	-
SupermarQ[48]	-	-	-	✓	-
Olympus[49]	-	-	-	-	✓
Golem[50]	-	-	-	-	✓
HTE-MC[51]	-	-	-	-	✓
JARVIS-LB	✓	✓	✓	✓	✓

superconducting transition temperatures as well as spectral quantities such as dielectric functions. While there are more than 400 approximate exchange-correlation functionals proposed in DFT literature[122], currently, we have OptB88vdW[65], Opt86BvdW[66], LDA[64], PBE[54], PBEsol[73], GLLB-sc[78], TBmBJ[67,68], SCAN[69], r2SCAN[70], HSE06[71], in the leaderboard. We use converged k-points and cutoffs as available in the JARVIS-DFT database[123]. We have used the Vienna Ab initio Simulation Package (VASP)[52,53], ABINIT[74–76], GPAW[77] and Quantum Espresso (QE)[72] as DFT software packages, but other packages can be easily added as well. In addition, we use VASP[52,53] to perform GW calculations including "single-shot" G_0W_0 and self-consistent GW_0 methods[63]. Other ES approaches include tight-binding (TB)[57] and quantum Monte Carlo (QMC)[61]. For TB, we use the recently developed ThreeBodyTB.jl code[59] along with the Wannier90[80] code, while the QMCPACK[79] code is used for diffusion Monte Carlo (DMC)[61] calculations.

B. Force-field

Force fields can be used in molecular dynamics and Monte Carlo simulations for studying larger time and length scales compared to electronic structure methods. Traditional force fields are developed for specific chemical systems and applications and may not be transferable to other uses. It is important to check the validity of a FF before using it in a particular application. Moreover, the development of FFs is a cumbersome task. Examples of typical FFs include embedded-atom method (EAM) potentials[110] (i.e. Al099.eam.alloy for aluminum system[124]), Lennard Jones (LJ)[108] for 2D liquids, reactive empirical bond order (REBO)[111] for Si, and classical, atomistic force fields for biomolecular systems[125,126]. Recently, machine learning force fields (MLFF)[127–132] have become popular because of their higher accuracy and ease of development (such as SNAP[93] FFs). Nevertheless, early generations of MLFFs were also developed for specific types of chemistry and applications. Very recently, several MLFFs have been developed that can be used to simulate any combination of periodic table elements. Some of these FFs include M3GNET[102], ALIGNN-FF[133], and CHGNet[101]. In the leaderboard, we include benchmarks for energies, forces, and stress tensors for both specific systems and universal datasets.

Traditional FFs are available in LAMMPS[109], while MLFFs are integrated into the Atomic Simulation Environment (ASE)[134] package. Some of these MLFFs are now available in LAMMPS and other large-scale MD codes. In addition to static quantities, FFs can be used for Monte Carlo simulations, such as CO_2 adsorption in metal-organic frameworks (MOFs)[135] using the RASPA[136] code. In addition to energy, force, and stress, we also have FF benchmarks for classical properties such as the bulk modulus. For biomolecular systems, GROMACS[137] is commonly used, and we present here free energy differences and conformational state population benchmarks for three model peptides[138–140].

C. Artificial intelligence

Recently artificial intelligence methods have become popular for materials prediction across all lengths and time scales. We currently have benchmarks for four types of data used as input for the AI models: (1) atomic structure, (2) spectra, (3) images, and 4) text. AI techniques can be used for both forward prediction and inverse design. For atomic structure datasets, we use DFT datasets such as JARVIS-DFT[35], Materials Project (MP)[141], Tight binding three-body dataset (TB3)[59], Quantum-Machine 9 (QM9)[142,143]. For spectral data, we use either DFT-based spectra of, for example, electron or phonon density of states (DOS), Eliashberg functions, or numerical XRD spectra. For images, we have simulated and experimental scanning transmission electron microscope (STEM) and scanning tunneling microscopy (STM) images for 2D materials. For text data, we have used the publicly available arXiv dataset.

Currently, we have models for feature-based/tabular models (such as RandomForest[144], Gradient boosting[144], Linear regression[144]), graph based models (such as ALIGNN[94], SchNet[96], CGCNN[95], M3GNET[102], AtomVision[97], ChemNLP[98]) as well as transformers (such as OPT[106], GPT[15], and T5[107]). These models use popular AI code bases including PyTorch[145], scikit-learn[144], TensorFlow[146], LightGBM[56], JAX[147], and HuggingFace[148]. These models are used for a variety of properties such as formation energies, electron bandgaps, phonon spectra, forces, text data etc.

TABLE II. Summary of current benchmark categories and methods available in the JARVIS-Leaderboard at the time of writing. More details can be found in the individual metadata.json file. Note that the number of methods is continuously growing.

Category	General name	Method Specification
ES	DFT[64]	VASP[52,53] (PBE[54], LDA[64], OptB88vdW[65], Opt86BvdW[66], TBmBJ[67,68], SCAN[69], r2SCAN[70], HSE06[71])
		QE[72] (PBE[54], PBEsol[73])
		ABINIT[74–76] (PBE[54])
		GPAW[77] (PBE[54], LDA[64], GLLB-sc[78])
	QMC[61]	QMCPACK[79] (DMC[61])
	GW[63]	VASP[52,53] (G_0W_0[63], GW_0[63])
	TB[57]	ThreeBodyTB.jl[59] (Wannier90[80])
AI	Descriptor	CFID[81], MagPie[82], MatMiner[55,83], crystal feature model[84], ElemNet[85–87], IRNet[88–90], BRNet[91,92], SNAP[93]
,	Graph-based	ALIGNN[94], CGCNN[95], SchNet[96], AtomVision[97], ChemNLP[98], DimeNet+[99,100], CHGNet[101], M3GNET[102] kgcnn_coGN[103], Potnet[104], Matformer[105]
	Transformers	OPT[106], GPT[15], T5[107]
FF	LJ[108]	LAMMPS[109] (2D-Liquid)
	EAM[110]	LAMMPS[109] (FCC-Al)
	REBO[111]	LAMMPS[109] (Diamond-Si)
	AMBER99sb-ildn[112]	GROMACS[113] (Alanine dipeptide)
	CHARMM36m[114]	GROMACS[113] (α-aminoisobutyric acid)
QC	Algorithms	Qiskit[115] (VQE[116], VQD[117])
		PennyLane[118,119] (VQE[116], VQD[117])
	Circuits	Qiskit[115] (PauliTwo Design[115], SU(2)[115])
EXP	Diffraction	XRD (Bruker D8)
	Manometry	CO_2 adsorption FACT lab[120]
	Vibroscopy	Kevlar FAVIMAT[121]
	Magnetometry	Susceptibility (PPMS)[121]

D. Quantum computation

Quantum chemistry is one of the most promising applications of quantum computations[149]. Quantum computers with relatively few logical qubits can potentially exceed the performance of much larger classical computers because the size of Hilbert space increases exponentially with the number of electrons in the system. Predicting the energy levels of a Hamiltonian is a typical and fundamentally important problem in quantum chemistry. We use Hamiltonian simulations with quantum algorithms and compare it with classical solvers. Determination of appropriate quantum circuit for a specific QC problem is a challenging task. For example, we use the tight-binding Hamiltonians for electrons and phonons in JARVIS-DFT and evaluate the electron bandstructures using quantum algorithms (such as variational quantum eigen solver (VQE)[116] and variational quantum deflation (VQD)[117]) and with different quantum circuits (such as PauliTwo design[115] and SU(2)[115] circuits). We primarily use the Qiskit[115] software in this work through the JARVIS-Tools/AtomQC[150] interface, but other packages such as Tequila[151], Circq[152], and Pennylane[118,119] can also be easily integrated. In addition to studying algorithm and circuit architecture dependence, the leaderboard can be used for studying the noise-levels in quantum circuits across different quantum computers, which is a key issue hindering quantum computer commercialization. Currently, we are only using statevector simulators for the quantum algorithms available in the Qiskit[115] library.

E. Experiments

Some of the experiments used for benchmarking purposes are XRD, magnetometry, vibroscopy, and scanning electron microscopy (SEM) and transition electron microscopy (TEM). We purchase the samples from industrial vendors with available identifiers such as CAS-number and carry out systematic experiments in a round-robin (inter-laboratory test performed independently several times[153]) fashion (such as manometric measurements for CO_2 adsorption[120]. We carried out XRD for MgB_2 (a superconducting material) to verify its crystal structure before carrying out magnetometry measurements to determine the transition temperature. This measurement was compared with numerical XRD data. Magnetometry measurements for 2D superconductors were also conducted to compare their superconducting transition temperatures with respect to predicted or experimentally available values[121]. Strain-stress measurements were done for Kevlar for failure analysis[121]. We have several instruments such as Bruker D8, Titan, Quantum design PPMS and FAVIMAT in the leaderboard currently.

F. Metrics used

We use several metrics in the leaderboard depending on the "sub-categories" mentioned above. We use mean absolute error (MAE), accuracy (acc), multi-mae (Euclidean distance), recall-oriented understudy for gisting evaluation (ROUGE) for the singlepropertyprediction, singlepropertyclassification, spectra/eigensolver/atomic forces and textGen/textsummary subcategories respectively. As the user contributes their data to compare against the reference data (benchmarks), other complementary metrics (such as those available in the sklearn.metrics library) can be easily calculated as the raw contribution data is also made available through the website. For the sake of readability and ease of use, we primarily employ the metrics mentioned above. For single property prediction, there is only scalar values per column in the csv.zip file with id and prediction separate by comma., For spectra, forceprediction and other multi-value quantities (i.e.with multiple prediction values per id) we concatenate the array and separate by semicolon (to avoid comma convention in csv files). The benchmark data is also stored in a similar format. We provide tools to convert these csv.zip files into json or other file formats if needed. We also provided notebooks to visualize the data through Jupyter/Colab notebooks. In addition, we plan to eventually add metrics for timing, uncertainty, development cost and other details.

IV. BENCHMARKS

The benchmarks consists of experimental, density functional theory or numerical solution-based data which are well-known and are already published in peer-reviewed articles/books. Currently, we have more than 270 benchmarks in the leaderboard. Each entry in the dataset consists of a unique identifier. Most of these datasets are integrated into JARVIS-Tools already, with an associated JARVIS ID number (JID) and are backed up in Figshare, Google Drive and NIST-internal storage systems. The number of entries can vary from a few (which is especially applicable for experimental and high-accuracy computational methods, where generating a very large dataset is not feasible in terms of time and resources) to hundreds of thousands of entries in a dataset.

An overview of the dataset can be found in Fig. 4. Considering all possible entries in the dataset, we have close to 7 million datapoints. For example, an atomic structure can have multiple properties calculated, such as bandgaps and formation energies, among other properties. We find the JARVIS-DFT-3D dataset to have the largest number of entries. Considering unique systems, we can find the distribution in Fig. 4b). In this case, qe-tb (fitting dataset for ThreeBodyTB.jl[59]) is one of the largest datasets available in the leaderboard. Note that these datasets contain all varieties of data modalities such as atomic structure, images, spectra and text.

In Fig. 5, we show the fractional distribution of periodic table elements in the entire dataset. We find that the most common elements are C, N, O, Cu which is similar to the natural abundance of these elements.

Experimental results are uploaded as benchmarks (i.e. what is regarded as the reference). In the absence of experimental data, high-fidelity computational methods can be used as a reference. If there are multiple experimental measurements available in the literature, each can be individually added as separate benchmarks (i.e., different json.zip files to distinguish one benchmark from another) and users can submit contributions for each of them. As time and the materials science field progresses, certain experimental data may need to be revisited (i.e. more accurate measurements in the future or results are reported that contradict previous experimental data). As a response to this, separate reference (experimental) benchmarks can be added, and users will be able to plot and compare the evolution of these benchmarks over time.

In addition, Leaderboard users can raise an issue on GitHub pertaining to reference benchmarks. The administrators will also upload a README file which contains additional information about the experiments conducted, including associated DOI, experimental conditions and provide details if additional experiments conducted on the same material/property exist in the literature. The experimental conditions described in the README file can be important when comparing the reference benchmark to calculated results, which may be in different conditions than the experiment (i.e. the bandgap of a material is never measured at 0 K, as DFT predicts).

Contributions to the leaderboard in the form of user-submitted experimental data can be compared with previous experiments, electronic structure methods or other numerical results. ES-based contributions are benchmarked against experimental results and can be compared with other ES methods. QC data can be compared with classical computation data or exact analytical results. For FF, contributions can be compared to DFT (or other ES data) or high-level interatomic potential benchmark suites (specifically for MLFFs)[154]. For AI, a test dataset is used. Unlike other methods, AI methods can have both "train" and "test" datasets, while others have only "test" sets in the corresponding dataset. For AI methods, if the "train" dataset is not provided and only "test" is given, the benchmark can be used for checking extrapolation behavior such as vacancy formation energy benchmarks.

V. RESULTS AND DISCUSSION

Presently, the leaderboard has 5 categories, 10 subcategories, 152 methods, 274 benchmarks, 1281 contributions and 8714228 datapoints. In this section, we show a few of the hundreds of example analyses that can be carried out using the available benchmarks and contributions. In Fig. 6, we show the MAE of the AI computed formation energy and ES computed bandgap for Si for a variety of contributions in the leaderboard. In Fig. 6(a) we see the comparison of 12 AI models (each AI model had a well-defined 80:10:10 split for training, validation and testing respectively from the JARVIS-3D database) and find the kgcnn_coGN[103] has the highest accuracy/lowest error, followed by Potnet[104], Matformer[105] and ALIGNN[94] models. This can be attributed to the fact that as we include more structural information and

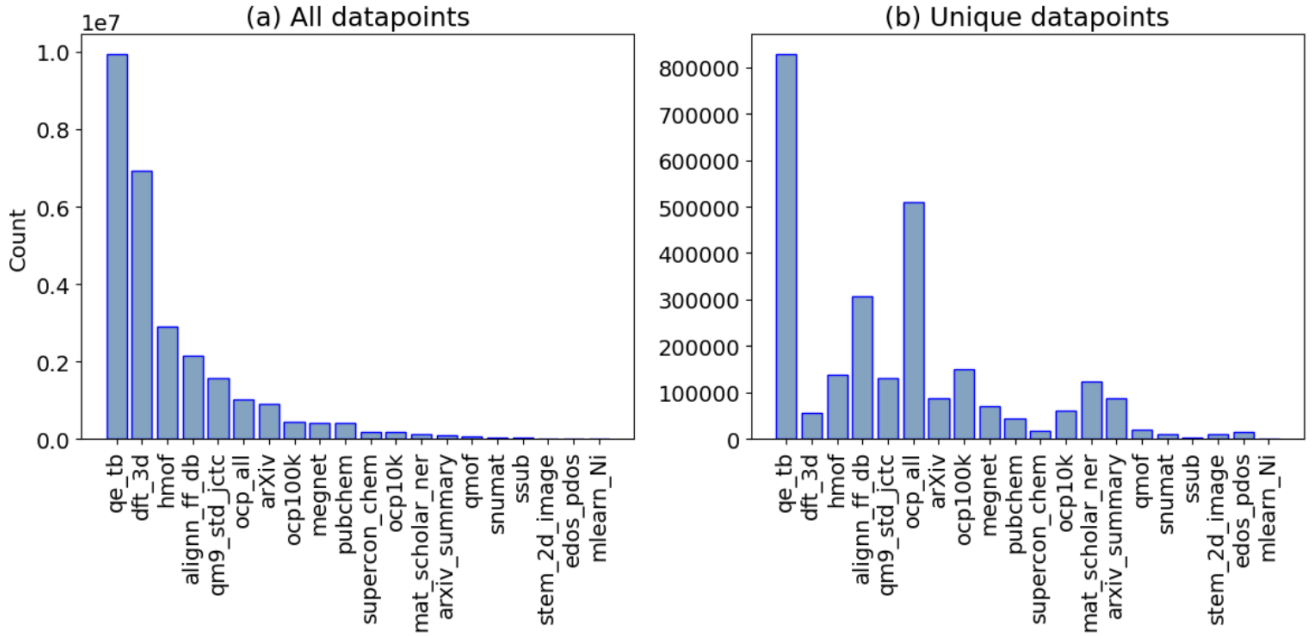


FIG. 4. Distribution of data in each dataset. (a) all entries in leaderboard, (b) entries with unique identifiers. Note that one identifier (such as JVASP-1002 for silicon) can have multiple properties (such as bandgap, bulk modulus etc.). A script to generate this figure is also provided on the leaderboard website as the leaderboard is continuously evolving.

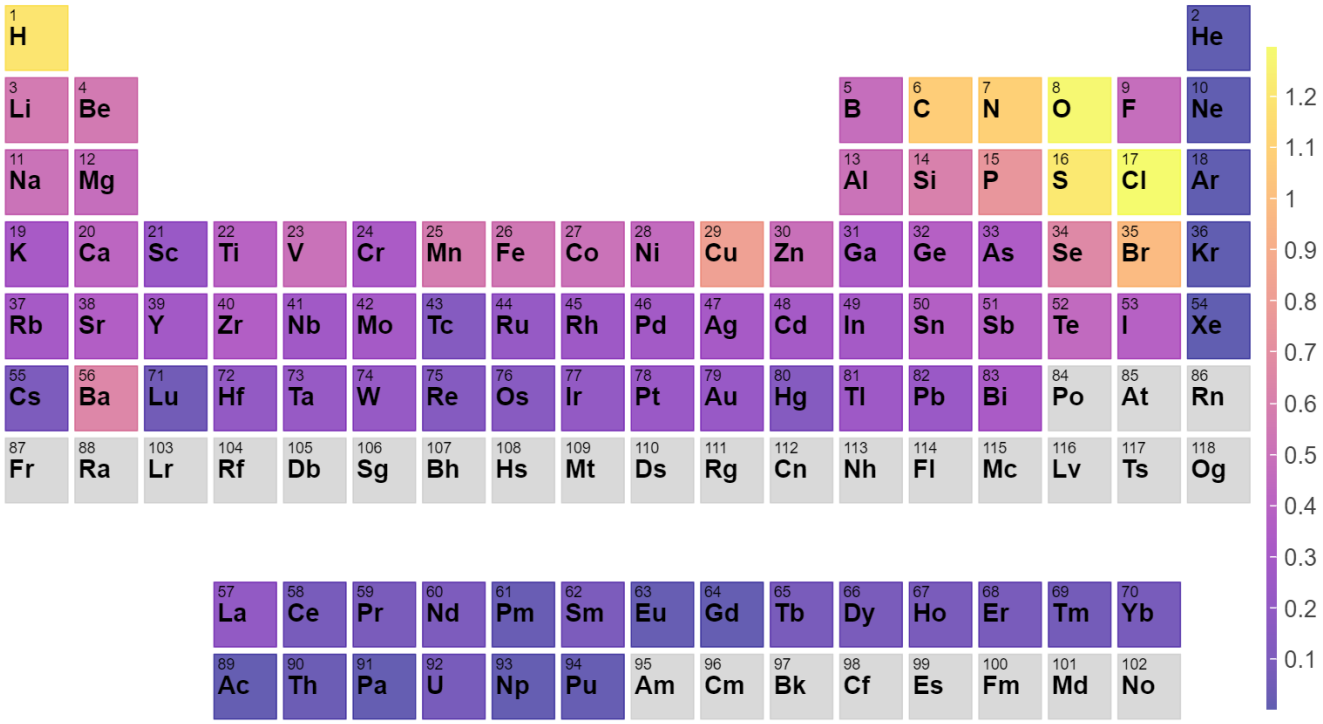


FIG. 5. Periodic table element distribution for entries in all the datasets. This is calculated by taking into account all the element specific entries normalized by total entries i.e. these are percentage probabilities.

use deep-learning methods rather than descriptor methods, we get improvement in the accuracy. Similarly, in Fig. 6(b) we compare the bandgap of Si using several methods and find GLLB-sc[78] calculated with GPAW[77] to yield the lowest error, while G_0W_0[63] (VASP[52,53]), GW_0[63] (VASP[52,53]), TBmBJ[67,68] (VASP), and DMC[61] (QMCPACK[79]) methods follow. This

can be attributed to the inclusion of the discontinuity potential (GLLB-sc[78]) or kinetic energy density (TBmBJ[67,68]) in the density functional or incorporating many-body physics (G_0W_0[63], GW_0[63], DMC[61]) into the methodology, which can lead to improved accuracy for bandgap prediction. Also, similar methods such as PBE[54] data from Open Quantum Materials Database (OQMD)[155,156], AFLOW[157] and Materials Project[141] compare well with each other. The comparisons are presented in tabular form for all the benchmarks on the leaderboard website. We have provided tools and notebooks in the leaderboard GitHub repository that can be used for making such plots for all the available benchmarks and contributions. A collections of such figures for method comparison is available in the supplementary information. We plan to eventually add interactive plots for such comparisons on the website. These tools can aid in identifying examples of materials that require high-fidelity methods beyond the accuracy of DFT in order to understand their underlying properties. In addition, these tools can be used to validate electronic structure methods and provide insight for error estimation.

Our leaderboard with a large number of benchmarks can enable a more comprehensive comparison of different methods for better revealing their respective advantages and limitations. For instance, neural networks outperform descriptor-based models by a large degree in all of the 10 regression tasks in the latest Matbench[16] leaderboard. To check if this is also the case for 44 regression benchmarks in the current JARVIS leaderboard, we compare the performance of the best descriptor-based model to that of the best neural network. As shown in Fig. 7, the best neural network outperform the best descriptor-based model in 34 tasks, but only 14 out of 44 (32 %) tasks see a performance difference by more than 20 %. This indicates that descriptor-based models are still competitive with respect to neural networks, especially considering their better interpretability and orders of magnitude lower training cost[83,158]. Notably, the best descriptor-based model is found to outperform the best neural network in 10 tasks including those with 10^4-10^5 training data, opening up interesting questions and potential direction to further model improvement. For instance, the inferior performance of neural networks in the regression tasks for the heat capacity and hMOF data may be related to the recently revealed incapability of graph neural networks in capturing periodicity[159].

Although a metric such as the MAE can be useful to compare methods for a specific benchmark, it is difficult to compare across different methods, since MAE values can differ substantially. Hence, we use the mean absolute deviation (MAD, computed with respect to the average value of the training data as a baseline/random-guess model) to MAE ratio for both AI and ES single-property-prediction categories. Mean absolute deviation values act as a baseline/random-guessing model for the benchmark and contributed models should have MAE performance better than MAD values. We show the MAD/MAE ratios for AI and ES benchmarks in Fig. 8. We find that the MAD/MAE values range from 2 to 50. MAD/MAE values close to 1 suggest low predictive power. We observe that quantum properties such as the bandgap have lower MAD/MAE than classical quantities (quantities that do not require quantum mechanical simulations) such as total energy or bulk modulus. Interestingly, such trends for classical vs. quantum quantities are observed for both the AI and ES approaches.

In addition to making bar plots as shown in Fig. 6 and Fig. 8, the raw data available in benchmarks and contributions can be presented in various other forms such as scatter plots, band-structures, adsorption spectra, and diffraction spectra. In Fig. 9, we show example comparisons of different methods for AI, ES, QC and EXP categories including (a) formation-energy-per atom model using AI, (b) bulk modulus predictions using ES, (c) electronic bandstructure of Al using QE with different quantum circuits[150], (d) CO_2 capture for zeolite at several labs in round-robin fashion[120]. In Fig. 9a), we find that formation energy is one of the easiest quantities to train AI models and even simple chemistry only-based models can perform reasonably well (i.e., cfid_chem). Including more structural features (such as bond angles and dihedral angles) and using deep learning models (such as graph neural network vs descriptor based models) further helps improve accuracy. Similarly, for ES example for predicting bulk modulus, we find irrespective of DFT based method used, they are in relatively close agreement with experimental bulk modulus data as shown in Fig. 9b). In Fig 9c), we find that the selection of a quantum circuit is critically important for predicting electronic band structures well. Here, we used 6 different quantum[150] circuits and found the SU(2)[115] circuit to compare well with classical computer-based electronic bandstructures. This can be attributed to various entanglements captured in the SU(2)[115] circuits that may be missing in other circuits. Finally, for experimental inter-laboratory/round-robin type measurements of the zeolite CO_2 isotherm, we find excellent agreement across different labs[120].

VI. SUMMARY

We have developed a large-scale and expandable benchmark platform for materials design methods, which is interfaced with the existing JARVIS infrastructure at NIST. This platform is an open-access and community-driven infrastructure to enhance reproducibility, method validation, and transparency. We welcome researchers around the globe to join efforts in solving critical materials design problems, enhance scientific understanding and accelerate industrial development. There are already more than 1200 contributions to the JARVIS-Leaderboard and we expect these contributions to grow to several thousand in the coming few years. The leaderboard will provide clarity on the strengths of different approaches, which is increasingly needed as the number of methods and implementations continues to multiply. We believe this community effort will improve transferability, validation, error quantification, and reproducibility throughout materials science.

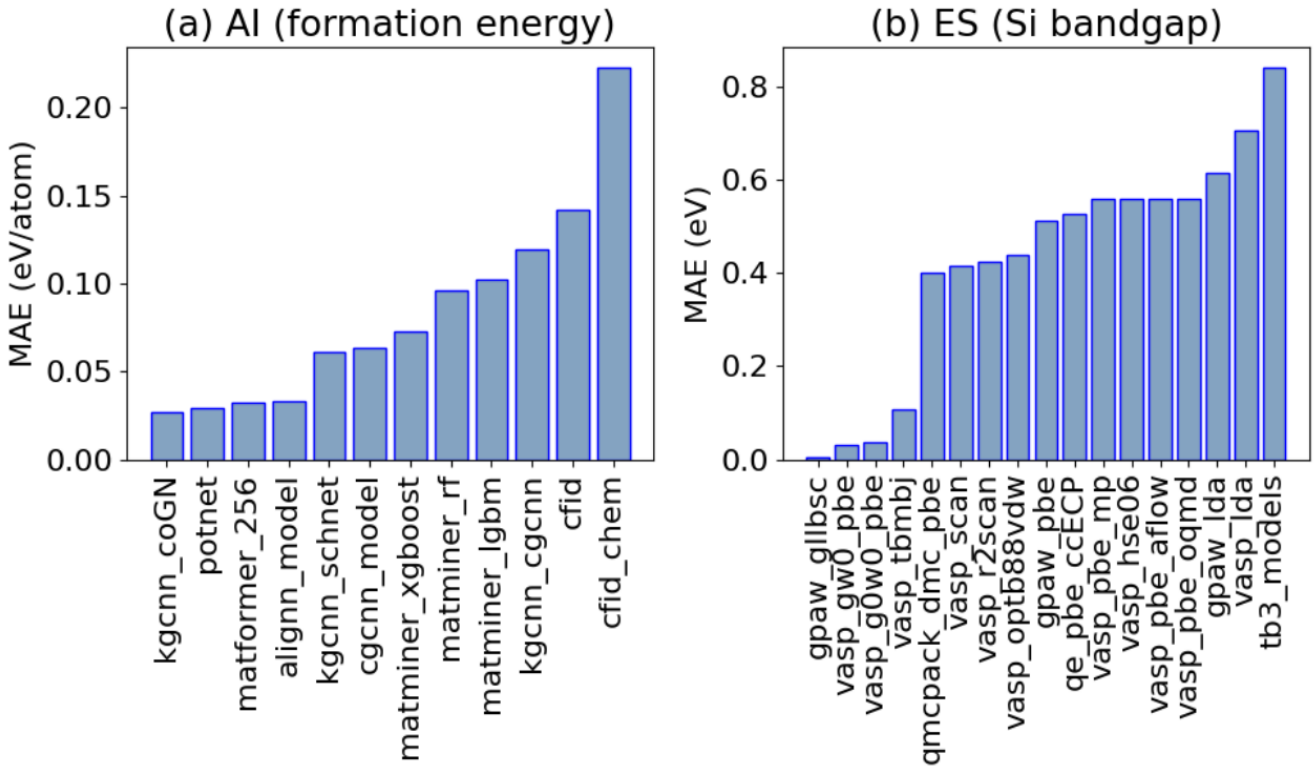


FIG. 6. Example mean absolute errors for (a) AI formation energy for test set with 5572 materials in JARVIS-DFT 3D dataset and (b) ES Si (JARVIS-DFT ID: JVASP-1002) bandgap benchmarks. We provide Jupyter/Google colab notebooks to plot such comparisons for all available benchmarks. Also, similar analysis figures for all the available benchmarks are available in the supplementary information.

VII. NOTES

Please note certain equipment, instruments, software, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply the recommendation or endorsement of any product or service by NIST, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The publisher acknowledges the US government license to provide public access under the DOE Public Access Plan (https://energy.gov/downloads/doe-public-access-plan). The Los Alamos National Laboratory is operated by the Triad National Security, LLC, for the National Nuclear Security Administration of U.S. Department of Energy (Contract No. 89233218CNA000001).

ACKNOWLEDGEMENTS

K.C., D.W., K.F.G., A.F., A.J.B., M.W., and F.T. thank the National Institute of Standards and Technology for funding, computational, and data-management resources. K.C. thanks the computational support from XSEDE (Extreme Science and Engineering Discovery Environment) computational resources under allocation number TG-DMR 190095. Contributions from K.C. were supported by the financial assistance award 70NANB19H117 from the U.S. Department of Commerce, National Institute of Standards and Technology. J.T.K., K.S. and P.R.C.K. were supported by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division, as part of the Computational Materials Sciences Program and Center for Predictive Simulation of Functional Materials. AHR also recognizes the support of West Virginia Higher Education Policy Commission under the call Research Challenge Grant (RCG) program. N.M. and A.M. acknowledge support from the U.S. Department of Energy through the LANL LDRD Programs under grant no. 20210036DR and 20220814PRD4, respectively. V.G. and A.A. were supported by NIST award 70NANB19H005 and NSF award CMMI-2053929. S.H.W. especially thanks to the NSF Non-Academic Research Internships for Graduate Students (INTERN) program (CBET-1845531) for supporting part of the work in NIST under the guidance of K.C.

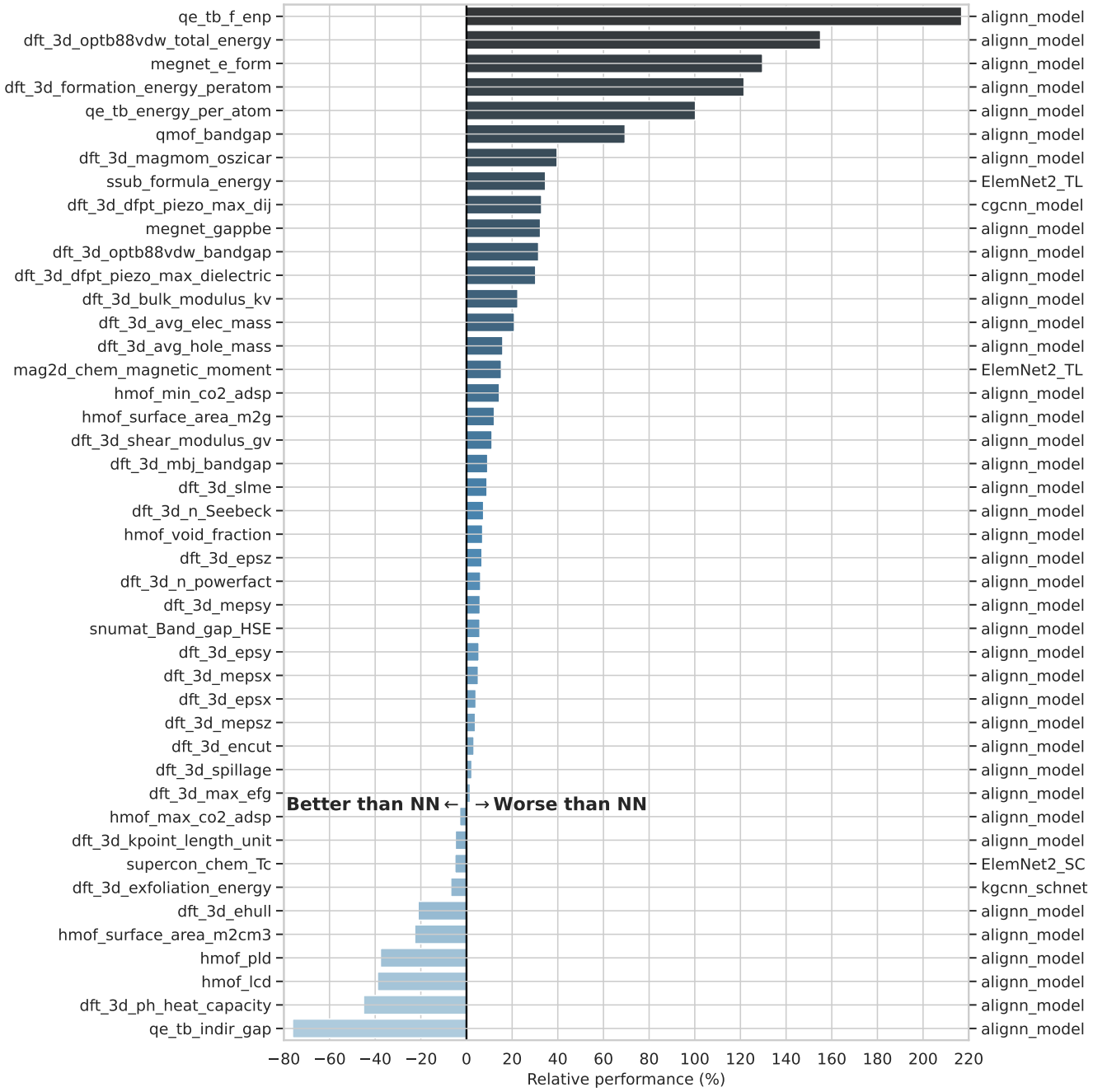


FIG. 7. Relative performance computed as the ratio of the MAE of the best descriptor-based model to that of the best neural networks in the AI regression benchmarks. The benchmark name and the corresponding best performing neural network are indicated in the left and right y axis, respectively. For all the considered AI benchmarks, the best descriptor-based model is the tree-based model using Magpie[82] and Voronoi-tessellation[160] features.

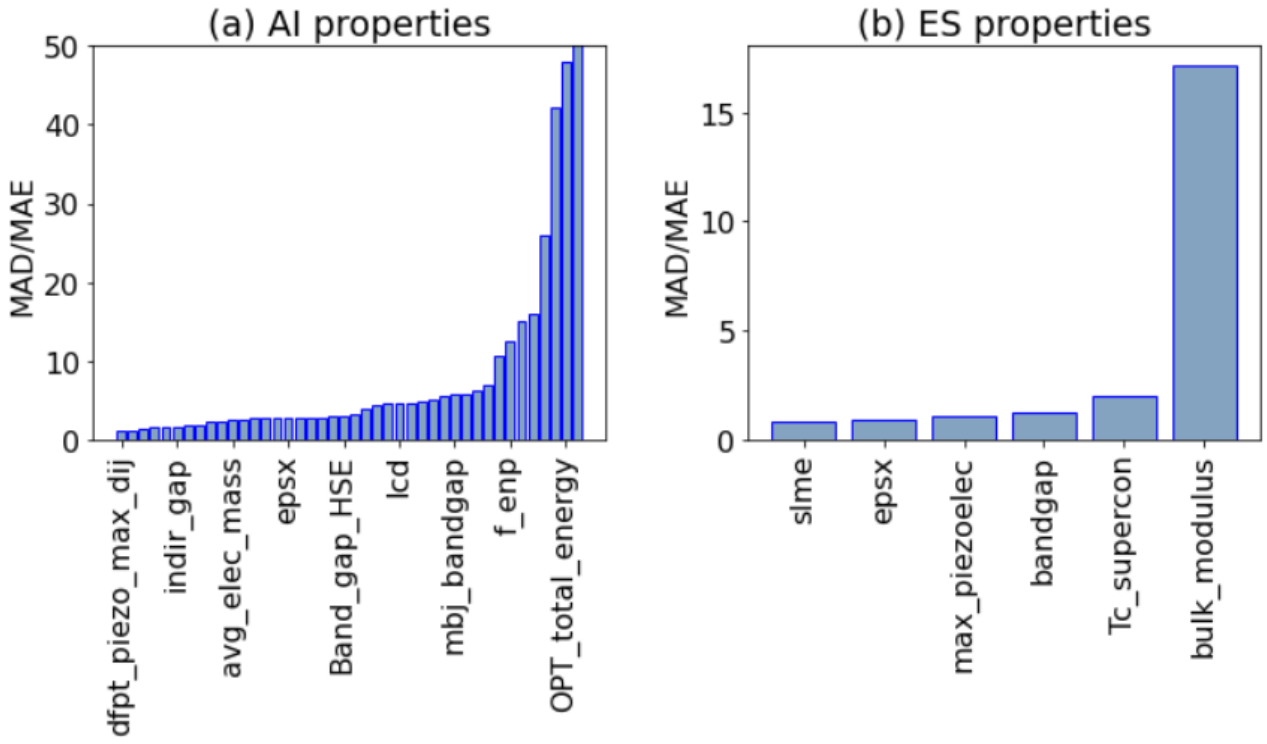


FIG. 8. Mean absolute deviation (MAD) to mean absolute error (MAE) ratio for (a) AI and (b) electronic structure methods. MAD:MAE serves as uniform criteria for comparing performances of models.

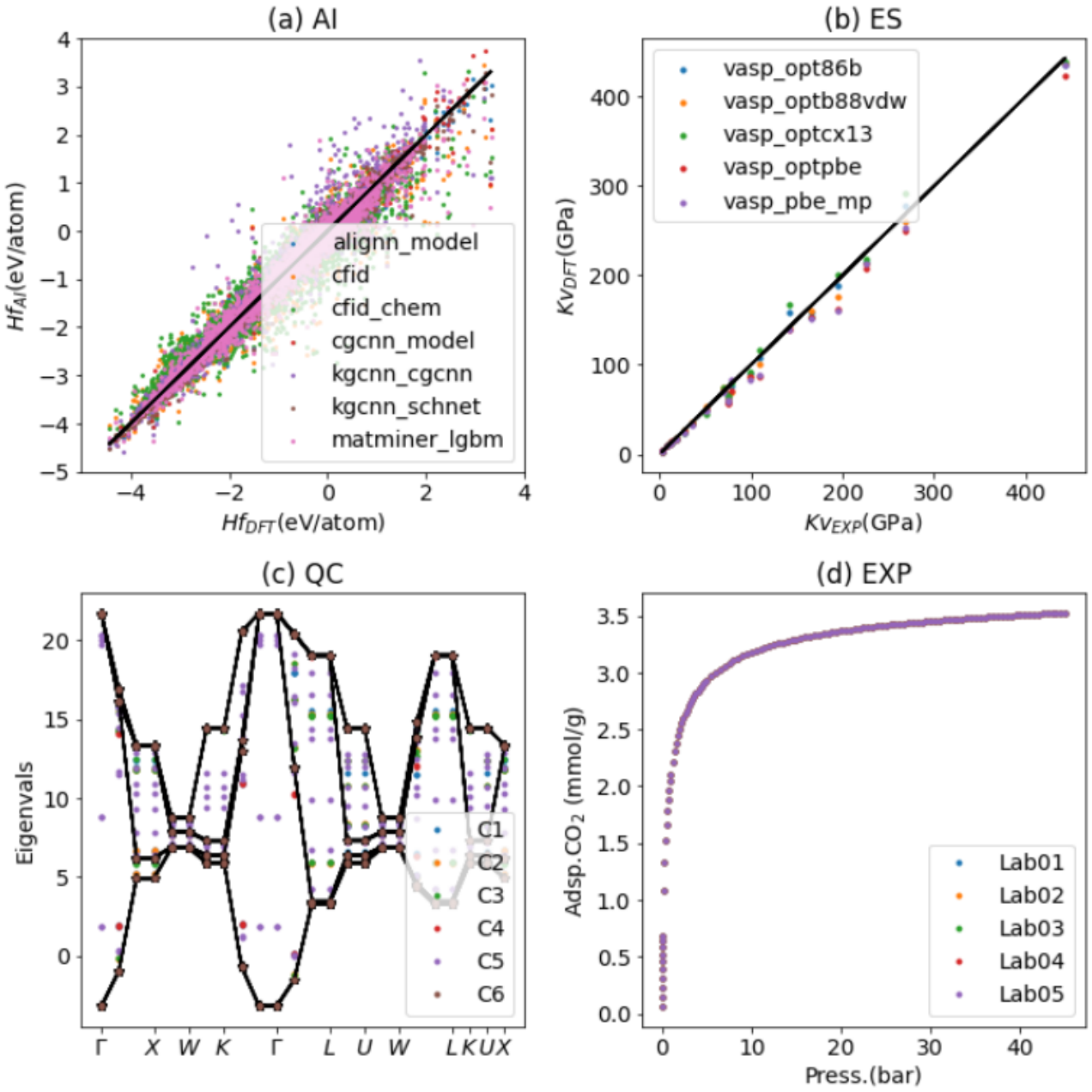


FIG. 9. Example results for AI, ES, QC and EXP results. (a) formation-energy-per atom model using AI for JARVIS-DFT 3D dataset with 5572 materials in the test set, (b) bulk modulus predictions using ES methods for 21 materials, (c) electronic bandstructure of Aluminum using QC methods with different quantum circuits on a coarse k-point mesh, (d) CO_2 capture for zeolite (ZSM-5) at several labs in inter-laboratory/round-robin fashion.

VIII. REFERENCES

[1] C. H. Ward, C. H. Ward, and J. A. Warren, *Materials genome initiative: materials data* (US Department of Commerce, National Institute of Standards and Technology, 2015).

[2] W. D. Callister *et al.*, *Fundamentals of materials science and engineering*, Vol. 471660817 (Wiley London, 2000).

[3] L.-Q. Chen, "Phase-field models for microstructure evolution," Annual Review of Materials Research **32**, 113–140 (2002), https://doi.org/10.1146/annurev.matsci.32.112001.132041.

[4] A. Agrawal, K. Gopalakrishnan, and A. Choudhary, "Materials image informatics using deep learning," in *Handbook on Big Data and Machine Learning in the Physical Sciences: Volume 1. Big Data Methods in Experimental Materials Discovery*, World Scientific Series on Emerging Technologies, edited by "" ("WorldScientific, 2020) pp. 205–230.

[5] K. Choudhary, B. DeCost, C. Chen, A. Jain, F. Tavazza, R. Cohn, C. W. Park, A. Choudhary, A. Agrawal, S. J. Billinge, *et al.*, "Recent advances and applications of deep learning methods in materials science," npj Computational Materials **8**, 59 (2022).

[6] D. J. Audus, K. Choudhary, B. L. DeCost, A. G. Kusne, F. Tavazza, and J. A. Warren, "Artificial intelligence for materials," in *Artificial Intelligence for Science*, Chap. Chapter 23, pp. 413–430.

[7] J. Park, J. D. Howe, and D. S. Sholl, "How reproducible are isotherm measurements in metal–organic frameworks?" Chemistry of Materials **29**, 10487–10495 (2017).

[8] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, *et al.*, "The fair guiding principles for scientific data management and stewardship," Scientific data **3**, 1–9 (2016).

[9] A. Agrawal and A. Choudhary, "Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science," APL Mater. **4**, 053208 (2016).

[10] J. Rickman, T. Lookman, and S. Kalinin, "Materials informatics: From the atomic-level to the continuum," Acta Materialia **168**, 473–510 (2019).

[11] A. Agrawal and A. Choudhary, "Deep materials informatics: Applications of deep learning in materials science," MRS Communications **9**, 779–792 (2019).

[12] K. Lejaeghere, G. Bihlmayer, T. Björkman, P. Blaha, S. Blügel, V. Blum, D. Caliste, I. E. Castelli, S. J. Clark, A. Dal Corso, *et al.*, "Reproducibility in density functional theory calculations of solids," Science **351**, aad3000 (2016).

[13] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, *et al.*, "Imagenet large scale visual recognition challenge," International journal of computer vision **115**, 211–252 (2015).

[14] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, *et al.*, "Highly accurate protein structure prediction with alphafold," Nature **596**, 583–589 (2021).

[15] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, *et al.*, "Language models are few-shot learners," Advances in neural information processing systems **33**, 1877–1901 (2020).

[16] A. Dunn, Q. Wang, A. Ganose, D. Dopp, and A. Jain, "Benchmarking materials property prediction methods: the matbench test set and automatminer reference algorithm," npj Computational Materials **6**, 138 (2020).

[17] Z. Wu, B. Ramsundar, E. N. Feinberg, J. Gomes, C. Geniesse, A. S. Pappu, K. Leswing, and V. Pande, "Moleculenet: a benchmark for molecular machine learning," Chemical science **9**, 513–530 (2018).

[18] S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, and K.-R. Müller, "Machine learning of accurate energy-conserving molecular force fields," Science advances **3**, e1603015 (2017).

[19] S. Chmiela, H. E. Sauceda, I. Poltavsky, K.-R. Müller, and A. Tkatchenko, "sgdml: Constructing accurate and data efficient molecular force fields using machine learning," Computer Physics Communications **240**, 38–45 (2019).

[20] Y. Zuo, C. Chen, X. Li, Z. Deng, Y. Chen, J. Behler, G. Csányi, A. V. Shapeev, A. P. Thompson, M. A. Wood, *et al.*, "Performance and cost assessment of machine learning interatomic potentials," The Journal of Physical Chemistry A **124**, 731–745 (2020).

[21] L. Weston, V. Tshitoyan, J. Dagdelen, O. Kononova, A. Trewartha, K. A. Persson, G. Ceder, and A. Jain, "Named entity recognition and normalization applied to large-scale information extraction from the materials science literature," J. Chem. Inf. Model. **59**, 3692–3702 (2019).

[22] M. Ziatdinov, A. Ghosh, T. Wong, and S. V. Kalinin, "Atomai: A deep learning framework for analysis of image and spectroscopy data in (scanning) transmission electron microscopy and beyond," arXiv preprint arXiv:2105.07485 (2021).

[23] P. Borlido, T. Aull, A. W. Huran, F. Tran, M. A. Marques, and S. Botti, "Large-scale benchmark of exchange–correlation functionals for the determination of electronic band gaps of solids," Journal of chemical theory and computation **15**, 5069–5079 (2019).

[24] S. P. Huber, E. Bosoni, M. Bercx, J. Bröder, A. Degomme, V. Dikan, K. Eimre, E. Flage-Larsen, A. Garcia, L. Genovese, *et al.*, "Common workflows for computing material properties using different quantum engines," npj Computational Materials **7**, 136 (2021).

[25] G.-X. Zhang, A. M. Reilly, A. Tkatchenko, and M. Scheffler, "Performance of various density-functional approximations for cohesive properties of 64 bulk solids," New Journal of Physics **20**, 063020 (2018).

[26] R. Tran, J. Lan, M. Shuaibi, S. Goyal, B. M. Wood, A. Das, J. Heras-Domingo, A. Kolluru, A. Rizvi, N. Shoghi, *et al.*, "The open catalyst 2022 (oc22) dataset and challenges for oxide electrocatalysis," arXiv preprint arXiv:2206.08917 (2022).

[27] P. Jurečka, J. Šponer, J. Černỳ, and P. Hobza, "Benchmark database of accurate (mp2 and ccsd (t) complete basis set limit) interaction energies of small model complexes, dna base pairs, and amino acid pairs," Physical Chemistry Chemical Physics **8**, 1985–1993 (2006).

[28] B. Brauer, M. K. Kesharwani, S. Kozuch, and J. M. Martin, "The s66x8 benchmark for noncovalent interactions revisited: Explicitly correlated ab initio methods and density functional theory," Physical Chemistry Chemical Physics **18**, 20905–20925 (2016).

[29] R. A. Mata and M. A. Suhm, "Benchmarking quantum chemical methods: Are we heading in the right direction?" Angewandte Chemie International Edition **56**, 11011–11018 (2017).

[30] D. E. Taylor, J. G. Ángyán, G. Galli, C. Zhang, F. Gygi, K. Hirao, J. W. Song, K. Rahul, O. Anatole von Lilienfeld, R. Podeszwa, *et al.*, "Blind test of density-functional-based methods on intermolecular interaction energies," The Journal of chemical physics **145**, 124105 (2016).

[31] D. Wheeler, T. Keller, S. J. DeWitt, A. M. Jokisaari, D. Schwen, J. E. Guyer, L. K. Aagesen, O. G. Heinonen, M. R. Tonks, P. W. Voorhees, *et al.*, "Pfhub: the phase-field community hub," Journal of Open Research Software **7** (2019).

[32] A. D. Lindsay, D. R. Gaston, C. J. Permann, J. M. Miller, D. Andrš, A. E. Slaughter, F. Kong, J. Hansel, R. W. Carlsen, C. Icenhour, L. Harbour, G. L. Giudicelli, R. H. Stogner, P. German, J. Badger, S. Biswas, L. Chapuis, C. Green, J. Hales, T. Hu, W. Jiang, Y. S. Jung, C. Matthews, Y. Miao, A. Novak, J. W. Peterson, Z. M. Prince, A. Rovinelli, S. Schunert, D. Schwen, B. W. Spencer, S. Veeraraghavan, A. Recuero, D. Yushu, Y. Wang, A. Wilkins, and C. Wong, "2.0 - MOOSE: Enabling massively parallel multiphysics simulation," SoftwareX **20**, 101202 (2022).

[33] J. Wei, B. Blaiszik, A. Scourtas, D. Morgan, and P. M. Voyles, "Benchmark tests of atom segmentation deep learning models with a consistent dataset," arXiv preprint arXiv:2207.10173 (2022).

[34] R. Vuorio, S.-H. Sun, H. Hu, and J. J. Lim, "Multimodal model-agnostic meta-learning via task-aware modulation," in *Advances in Neural Information Processing Systems*, Vol. 32, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., 2019).

[35] K. Choudhary, K. F. Garrity, A. C. Reid, B. DeCost, A. J. Biacchi, A. R. Hight Walker, Z. Trautt, J. Hattrick-Simpers, A. G. Kusne, A. Centrone, *et al.*, "The joint automated repository for various integrated simulations (jarvis) for data-driven materials design," npj computational materials **6**, 173 (2020).

[36] L. Chanussot, A. Das, S. Goyal, T. Lavril, M. Shuaibi, M. Riviere, K. Tran, J. Heras-Domingo, C. Ho, W. Hu, *et al.*, "Open catalyst 2020 (oc20) dataset and community challenges," Acs Catalysis **11**, 6059–6072 (2021).

[37] J. T. et al, "Scimlbench: A benchmarking suite for ai for science," (2021).

[38] N. Brown, M. Fiscato, M. H. Segler, and A. C. Vaucher, "Guacamol: benchmarking models for de novo molecular design," Journal of chemical information and modeling **59**, 1096–1108 (2019).

[39] G. Chen, P. Chen, C.-Y. Hsieh, C.-K. Lee, B. Liao, R. Liao, W. Liu, J. Qiu, Q. Sun, J. Tang, *et al.*, "Alchemy: A quantum chemistry dataset for benchmarking ai models," arXiv preprint arXiv:1906.09427 (2019).

[40] M. E. Khatib and W. A. de Jong, "Ml4chem: A machine learning package for chemistry and materials science," arXiv preprint arXiv:2003.13388 (2020).

[41] F. Broccatelli, R. Trager, M. Reutlinger, G. Karypis, and M. Li, "Benchmarking accuracy and generalizability of four graph neural networks using large in vitro adme datasets from different chemical spaces," Molecular Informatics **41**, 2100321 (2022).

[42] R. D. Johnson *et al.*, "Nist computational chemistry comparison and benchmark database," http://srdata. nist. gov/cccbdb (2006).

[43] G. Prandini, A. Marrazzo, I. E. Castelli, N. Mounet, and N. Marzari, "Precision and efficiency in solid-state pseudopotential calculations," npj Computational Materials **4**, 72 (2018).

[44] D. S. Karls, M. Bierbaum, A. A. Alemi, R. S. Elliott, J. P. Sethna, and E. B. Tadmor, "The openkim processing pipeline: A cloud-based automatic material property computation engine," The Journal of Chemical Physics **153**, 064104 (2020).

[45] L. M. Hale, Z. T. Trautt, and C. A. Becker, "Evaluating variability with atomistic simulations: the effect of potential and calculation methodology on the modeling of lattice and elastic constants," Modelling and Simulation in Materials Science and Engineering **26**, 055003 (2018).

[46] K. Choudhary, A. J. Biacchi, S. Ghosh, L. Hale, A. R. H. Walker, and F. Tavazza, "High-throughput assessment of vacancy formation and surface energies of materials using classical force-fields," Journal of Physics: Condensed Matter **30**, 395901 (2018).

[47] A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation, and J. M. Gambetta, "Validating quantum computers using randomized model circuits," Physical Review A **100**, 032328 (2019).

[48] T. Tomesh, P. Gokhale, V. Omole, G. S. Ravi, K. N. Smith, J. Viszlai, X.-C. Wu, N. Hardavellas, M. R. Martonosi, and F. T. Chong, "Supermarq: A scalable quantum benchmark suite," in *2022 IEEE International Symposium on High-Performance Computer Architecture (HPCA)* (IEEE, 2022) pp. 587–603.

[49] F. Häse, M. Aldeghi, R. J. Hickman, L. M. Roch, M. Christensen, E. Liles, J. E. Hein, and A. Aspuru-Guzik, "Olympus: a benchmarking framework for noisy optimization and experiment planning," (2020), arXiv:2010.04153 [stat.ML].

[50] M. Aldeghi, F. Häse, R. J. Hickman, I. Tamblyn, and A. Aspuru-Guzik, "Golem: an algorithm for robust experiment and process optimization," Chemical Science **12**, 14792–14807 (2021).

[51] J. R. Hattrick-Simpers, A. Zakutayev, S. C. Barron, Z. T. Trautt, N. Nguyen, K. Choudhary, B. DeCost, C. Phillips, A. G. Kusne, F. Yi, *et al.*, "An inter-laboratory study of zn–sn–ti–o thin films using high-throughput experimental methods," ACS combinatorial science **21**, 350–361 (2019).

[52] G. Kresse and J. Furthmüller, "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set," Physical review B **54**, 11169 (1996).

[53] G. Kresse and J. Furthmüller, "Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set," Computational materials science **6**, 15–50 (1996).

[54] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," Phys. Rev. Lett. **77**, 3865–3868 (1996).

[55] L. Ward, A. Dunn, A. Faghaninia, N. E. Zimmermann, S. Bajaj, Q. Wang, J. Montoya, J. Chen, K. Bystrom, M. Dylla, K. Chard, M. Asta, K. A. Persson, G. J. Snyder, I. Foster, and A. Jain, "Matminer: An open source toolkit for materials data mining," Computational Materials Science **152**, 60–69 (2018).

[56] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "Lightgbm: A highly efficient gradient boosting decision tree," Advances in neural information processing systems **30**, 3146–3154 (2017).

[57] W. A. Harrison, *Electronic structure and the properties of solids: the physics of the chemical bond* (Courier Corporation, 2012).

[58] K. F. Garrity and K. Choudhary, "Database of wannier tight-binding hamiltonians using high-throughput density functional theory," Scientific data **8**, 106 (2021).

[59] K. F. Garrity and K. Choudhary, "Fast and accurate prediction of material properties with three-body tight-binding model for the periodic table," Phys. Rev. Mater. **7**, 044603 (2023).

[60] R. M. Martin, *Electronic structure: basic theory and practical methods* (Cambridge university press, 2020).

[61] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, "Quantum Monte Carlo simulations of solids," Rev. Mod. Phys. **73**, 33–83 (2001).

[62] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. Marianetti, "Electronic structure calculations with dynamical mean-field theory," Reviews of Modern Physics **78**, 865 (2006).

[63] G. Onida, L. Reining, and A. Rubio, "Electronic excitations: density-functional versus many-body green's-function approaches," Rev. Mod. Phys. **74**, 601–659 (2002).

[64] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," Phys. Rev. **136**, B864–B871 (1964).

[65] J. Klimeš, D. R. Bowler, and A. Michaelides, "Chemical accuracy for the van der waals density functional," Journal of Physics: Condensed Matter **22**, 022201 (2009).

[66] J. c. v. Klimeš, D. R. Bowler, and A. Michaelides, "Van der waals density functionals applied to solids," Phys. Rev. B **83**, 195131 (2011).

[67] F. Tran and P. Blaha, "Importance of the kinetic energy density for band gap calculations in solids with density functional theory," The Journal of Physical Chemistry A **121**, 3318–3325 (2017).

[68] D. P. Rai, M. P. Ghimire, and R. K. Thapa, "A dft study of bex (x = s, se, te) semiconductor: Modified becke johnson (mbj) potential," Semiconductors **48**, 1411–1422 (2014).

[69] J. Sun, A. Ruzsinszky, and J. P. Perdew, "Strongly constrained and appropriately normed semilocal density functional," Phys. Rev. Lett. **115**, 036402 (2015).

[70] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun, "Accurate and numerically efficient r2scan meta-generalized gradient approximation," *The Journal of Physical Chemistry Letters*, The Journal of Physical Chemistry Letters **11**, 8208–8215 (2020).

[71] J. Heyd, G. E. Scuseria, and M. Ernzerhof, "Hybrid functionals based on a screened coulomb potential," The Journal of Chemical Physics **118**, 8207–8215 (2003), https://doi.org/10.1063/1.1564060.

[72] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, *et al.*, "Quantum espresso: a modular and open-source software project for quantum simulations of materials," Journal of physics: Condensed matter **21**, 395502 (2009).

[73] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke, "Restoring the density-gradient expansion for exchange in solids and surfaces," Phys. Rev. Lett. **100**, 136406 (2008).

[74] X. Gonze, F. Jollet, F. A. Araujo, D. Adams, B. Amadon, T. Applencourt, C. Audouze, J.-M. Beuken, J. Bieder, A. Bokhanchuk, *et al.*, "Recent developments in the abinit software package," Computer Physics Communications **205**, 106–131 (2016).

[75] A. H. Romero, D. C. Allan, B. Amadon, G. Antonius, T. Applencourt, L. Baguet, J. Bieder, F. Bottin, J. Bouchet, E. Bousquet, F. Bruneval, G. Brunin, D. Caliste, M. Côté, J. Denier, C. Dreyer, P. Ghosez, M. Giantomassi, Y. Gillet, O. Gingras, D. R. Hamann, G. Hautier, F. Jollet, G. Jomard, A. Martin, H. P. C. Miranda, F. Naccarato, G. Petretto, N. A. Pike, V. Planes, S. Prokhorenko, T. Rangel, F. Ricci, G.-M. Rignanese, M. Royo, M. Stengel, M. Torrent, M. J. van Setten, B. V. Troeye, M. J. Verstraete, J. Wiktor, J. W. Zwanziger, and X. Gonze, "Abinit: Overview, and focus on selected capabilities," J. Chem. Phys. **152**, 124102 (2020).

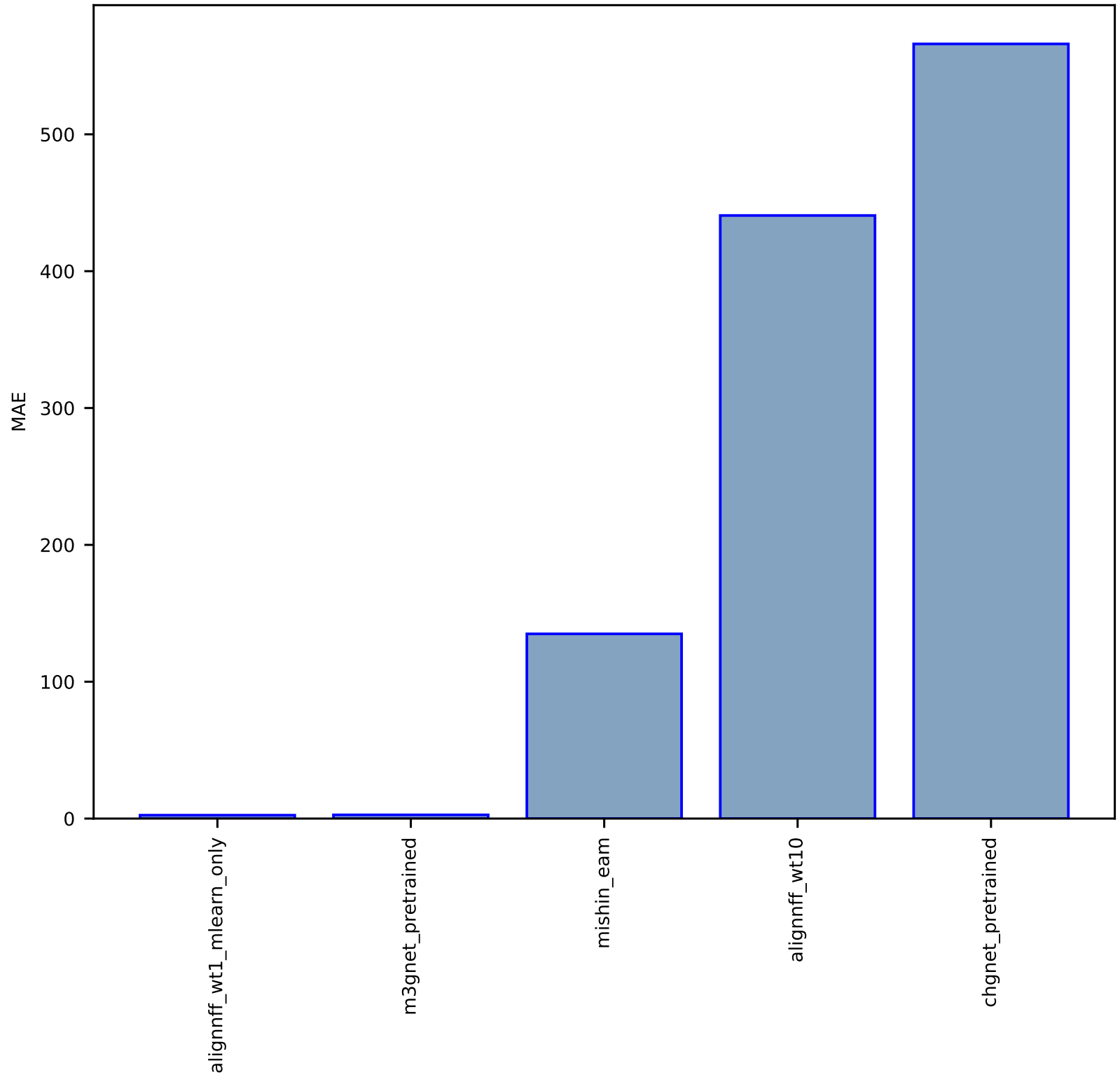
[76] X. Gonze, B. Amadon, G. Antonius, F. Arnardi, L. Baguet, J.-M. Beuken, J. Bieder, F. Bottin, J. Bouchet, E. Bousquet, N. Brouwer, F. Bruneval, G. Brunin, T. Cavignac, J.-B. Charraud, W. Chen, M. Côté, S. Cottenier, J. Denier, G. Geneste, P. Ghosez, M. Giantomassi, Y. Gillet, O. Gingras, D. R. Hamann, G. Hautier, X. He, N. Helbig, N. Holzwarth, Y. Jia, F. Jollet, W. Lafargue-Dit-Hauret, K. Lejaeghere, M. A. L. Marques, A. Martin, C. Martins, H. P. C. Miranda, F. Naccarato, K. Persson, G. Petretto, V. Planes, Y. Pouillon, S. Prokhorenko, F. Ricci, G.-M. Rignanese, A. H. Romero, M. M. Schmitt, M. Torrent, M. J. van Setten, B. V. Troeye, M. J. Verstraete, G. Zérah, and J. W. Zwanziger, "The abinit project: Impact, environment and recent developments," Comput. Phys. Commun.

248, 107042 (2020).
[77] J. Enkovaara, C. Rostgaard, J. J. Mortensen, J. Chen, M. Dułak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H. A. Hansen, H. H. Kristoffersen, M. Kuisma, A. H. Larsen, L. Lehtovaara, M. Ljungberg, O. Lopez-Acevedo, P. G. Moses, J. Ojanen, T. Olsen, V. Petzold, N. A. Romero, J. Stausholm-Møller, M. Strange, G. A. Tritsaris, M. Vanin, M. Walter, B. Hammer, H. Häkkinen, G. K. H. Madsen, R. M. Nieminen, J. K. Nørskov, M. Puska, T. T. Rantala, J. Schiøtz, K. S. Thygesen, and K. W. Jacobsen, "Electronic structure calculations with gpaw: a real-space implementation of the projector augmented-wave method," Journal of Physics: Condensed Matter **22**, 253202 (2010).
[78] M. Kuisma, J. Ojanen, J. Enkovaara, and T. T. Rantala, "Kohn-sham potential with discontinuity for band gap materials," Phys. Rev. B **82**, 115106 (2010).
[79] J. Kim, A. D. Baczewski, T. D. Beaudet, A. Benali, M. C. Bennett, M. A. Berrill, N. S. Blunt, E. J. L. Borda, M. Casula, D. M. Ceperley, *et al.*, "Qmcpack: an open source ab initio quantum monte carlo package for the electronic structure of atoms, molecules and solids," Journal of Physics: Condensed Matter **30**, 195901 (2018).
[80] A. A. Mostofi, J. R. Yates, G. Pizzi, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, "An updated version of wannier90: A tool for obtaining maximally-localised wannier functions," Computer Physics Communications **185**, 2309–2310 (2014).
[81] K. Choudhary, B. DeCost, and F. Tavazza, "Machine learning with force-field-inspired descriptors for materials: Fast screening and mapping energy landscape," Physical review materials **2**, 083801 (2018).
[82] L. Ward, A. Agrawal, A. Choudhary, and C. Wolverton, "A general-purpose machine learning framework for predicting properties of inorganic materials," npj Computational Materials **2**, 16028 (2016).
[83] K. Li, D. Persaud, K. Choudhary, B. DeCost, M. Greenwood, and J. Hattrick-Simpers, "On the redundancy in large material datasets: efficient and robust learning with less data," arXiv preprint arXiv:2304.13076 (2023).
[84] R. B. Wexler, G. S. Gautam, E. B. Stechel, and E. A. Carter, "Factors governing oxygen vacancy formation in oxide perovskites," J. Am. Chem. Soc. **143**, 13212–13227 (2021).
[85] D. Jha, L. Ward, A. Paul, W.-k. Liao, A. Choudhary, C. Wolverton, and A. Agrawal, "Elemnet: Deep learning the chemistry of materials from only elemental composition," Scientific Reports **8**, 17593 (2018).
[86] D. Jha, K. Choudhary, F. Tavazza, W.-k. Liao, A. Choudhary, C. Campbell, and A. Agrawal, "Enhancing materials property prediction by leveraging computational and experimental data using deep transfer learning," Nat. Commun **10**, 1–12 (2019).
[87] V. Gupta, K. Choudhary, F. Tavazza, C. Campbell, W.-k. Liao, A. Choudhary, and A. Agrawal, "Cross-property deep transfer learning framework for enhanced predictive analytics on small materials data," Nature communications **12**, 6595 (2021).
[88] D. Jha, L. Ward, Z. Yang, C. Wolverton, I. Foster, W.-k. Liao, A. Choudhary, and A. Agrawal, "Irnet: A general purpose deep residual regression framework for materials discovery," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, edited by "" (2019) pp. 2385–2393.
[89] D. Jha, V. Gupta, L. Ward, Z. Yang, C. Wolverton, I. Foster, W.-k. Liao, A. Choudhary, and A. Agrawal, "Enabling deeper learning on big data for materials informatics applications," Scientific reports **11**, 4244 (2021).
[90] D. Jha, V. Gupta, W.-k. Liao, A. Choudhary, and A. Agrawal, "Moving closer to experimental level materials property prediction using ai," Scientific reports **12**, 1–9 (2022).
[91] V. Gupta, W.-k. Liao, A. Choudhary, and A. Agrawal, "Brnet: Branched residual network for fast and accurate predictive modeling of materials properties," in *Proceedings of the 2022 SIAM international conference on data mining (SDM)* (SIAM, 2022) pp. 343–351.
[92] V. Gupta, A. Peltekian, W.-k. Liao, A. Choudhary, and A. Agrawal, "Improving deep learning model performance under parametric constraints for materials informatics applications," Scientific Reports **13**, 9128 (2023).
[93] C. Chen, Z. Deng, R. Tran, H. Tang, I.-H. Chu, and S. P. Ong, "Accurate force field for molybdenum by machine learning large materials data," Phys. Rev. Mater. **1**, 043603 (2017).
[94] K. Choudhary and B. DeCost, "Atomistic line graph neural network for improved materials property predictions," npj Computational Materials **7**, 185 (2021).
[95] T. Xie and J. C. Grossman, "Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties," Phys. Rev. Lett. **120**, 145301 (2018).
[96] K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller, "Schnet – a deep learning architecture for molecules and materials," The Journal of Chemical Physics **148**, 241722 (2018), https://doi.org/10.1063/1.5019779.
[97] K. Choudhary, R. Gurunathan, B. DeCost, and A. Biacchi, "Atomvision: A machine vision library for atomistic images," Journal of Chemical Information and Modeling (2023).
[98] K. Choudhary and M. L. Kelley, "Chemnlp: A natural language processing based library for materials chemistry text data," (2022), 10.48550/ARXIV.2209.08203.
[99] J. Gasteiger, J. Groß, and S. Günnemann, "Directional message passing for molecular graphs," in *International Conference on Learning Representations (ICLR)* (2020).
[100] J. Gasteiger, S. Giri, J. T. Margraf, and S. Günnemann, "Fast and uncertainty-aware directional message passing for non-equilibrium molecules," in *Machine Learning for Molecules Workshop, NeurIPS* (2020).
[101] B. Deng, P. Zhong, K. Jun, K. Han, C. J. Bartel, and G. Ceder, "CHGNet: Pretrained universal neural network potential for charge-informed atomistic modeling," arXiv preprint arXiv:2302.14231 (2023).
[102] C. Chen and S. P. Ong, "A universal graph deep learning interatomic potential for the periodic table," Nature Computational Science **2**, 718–728 (2022).
[103] P. Reiser, A. Eberhard, and P. Friederich, "Graph neural networks in tensorflow-keras with raggedtensor representation (kgcnn)," Software Impacts , 100095 (2021).
[104] Y. Lin, K. Yan, Y. Luo, Y. Liu, X. Qian, and S. Ji, "Efficient approximations of complete interatomic potentials for crystal property prediction," in *Proceedings of the 40th International Conference on Machine Learning* (2023).
[105] K. Yan, Y. Liu, Y. Lin, and S. Ji, "Periodic graph transformers for crystal material property prediction," in *The 36th Annual Conference on Neural Information Processing Systems* (2022) pp. 15066–15080.
[106] S. Zhang, S. Roller, N. Goyal, M. Artetxe, M. Chen, S. Chen, C. Dewan, M. Diab, X. Li, X. V. Lin, T. Mihaylov, M. Ott, S. Shleifer, K. Shuster, D. Simig, P. S. Koura, A. Sridhar, T. Wang, and L. Zettlemoyer, "Opt: Open pre-trained transformer language models," (2022), arXiv:2205.01068 [cs.CL].
[107] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer," (2020), arXiv:1910.10683 [cs.LG].
[108] J. E. Jones and S. Chapman, "On the determination of molecular fields.—i. from the variation of the viscosity of a gas with temperature," Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **106**, 441–462 (1924), https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1924.0081.
[109] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, "LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales," Comp. Phys. Comm. **271**, 108171 (2022).
[110] M. S. Daw and M. I. Baskes, "Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals," Phys. Rev. B **29**, 6443–6453 (1984).
[111] J. Tersoff, "New empirical approach for the structure and energy of covalent systems," Phys. Rev. B **37**, 6991–7000 (1988).
[112] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, "Improved side-chain torsion potentials for the amber ff99sb protein force field," Proteins: Structure, Function, and Bioinformatics **78**, 1950–1958 (2010).
[113] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. Berendsen, "Gromacs: fast, flexible, and free," Journal of computational chemistry **26**, 1701–1718 (2005).
[114] S. Mehdi, D. Wang, S. Pant, and P. Tiwary, "Accelerating all-atom simulations and gaining mechanistic understanding of biophysical systems

through state predictive information bottleneck," Journal of chemical theory and computation **18**, 3231–3238 (2022).
[115]"IBM Quantum," https://quantum-computing.ibm.com (2021).
[116]A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, "A variational eigenvalue solver on a photonic quantum processor," Nature Communications **5**, 4213 (2014).
[117]O. Higgott, D. Wang, and S. Brierley, "Variational Quantum Computation of Excited States," Quantum **3**, 156 (2019).
[118]V. Bergholm, J. Izaac, M. Schuld, C. Gogolin, S. Ahmed, V. Ajith, M. S. Alam, G. Alonso-Linaje, B. AkashNarayanan, A. Asadi, J. M. Arrazola, U. Azad, S. Banning, C. Blank, T. R. Bromley, B. A. Cordier, J. Ceroni, A. Delgado, O. D. Matteo, A. Dusko, T. Garg, D. Guala, A. Hayes, R. Hill, A. Ijaz, T. Isacsson, D. Ittah, S. Jahangiri, P. Jain, E. Jiang, A. Khandelwal, K. Kottmann, R. A. Lang, C. Lee, T. Loke, A. Lowe, K. McKiernan, J. J. Meyer, J. A. Montañez-Barrera, R. Moyard, Z. Niu, L. J. O'Riordan, S. Oud, A. Panigrahi, C.-Y. Park, D. Polatajko, N. Quesada, C. Roberts, N. Sá, I. Schoch, B. Shi, S. Shu, S. Sim, A. Singh, I. Strandberg, J. Soni, A. Száva, S. Thabet, R. A. Vargas-Hernández, T. Vincent, N. Vitucci, M. Weber, D. Wierichs, R. Wiersema, M. Willmann, V. Wong, S. Zhang, and N. Killoran, "Pennylane: Automatic differentiation of hybrid quantum-classical computations," (2022), arXiv:1811.04968 [quant-ph].
[119]J. M. Arrazola, S. Jahangiri, A. Delgado, J. Ceroni, J. Izaac, A. Száva, U. Azad, R. A. Lang, Z. Niu, O. D. Matteo, R. Moyard, J. Soni, M. Schuld, R. A. Vargas-Hernández, T. Tamayo-Mendoza, C. Y.-Y. Lin, A. Aspuru-Guzik, and N. Killoran, "Differentiable quantum computational chemistry with pennylane," (2023), arXiv:2111.09967 [quant-ph].
[120]H. G. T. Nguyen, L. Espinal, R. D. van Zee, M. Thommes, B. Toman, M. S. L. Hudson, E. Mangano, S. Brandani, D. P. Broom, M. J. Benham, K. Cychosz, P. Bertier, F. Yang, B. M. Krooss, R. L. Siegelman, M. Hakuman, K. Nakai, A. D. Ebner, L. Erden, J. A. Ritter, A. Moran, O. Talu, Y. Huang, K. S. Walton, P. Billemont, and G. De Weireld, "A reference high-pressure co2 adsorption isotherm for ammonium zsm-5 zeolite: results of an interlaboratory study," Adsorption **24**, 531–539 (2018).
[121]A. Engelbrecht-Wiggans, F. Burni, E. Guigues, S. Jiang, T. Huynh, Z. Tsinas, D. Jacobs, and A. Forster, "Effects of temperature and humidity on high-strength p-aramid fibers used in body armor," Textile Research Journal **90**, 2428–2440 (2020), https://doi.org/10.1177/0040517520918232.
[122]S. Lehtola, C. Steigemann, M. J. Oliveira, and M. A. Marques, "Recent developments in libxc—a comprehensive library of functionals for density functional theory," SoftwareX **7**, 1–5 (2018).
[123]K. Choudhary and F. Tavazza, "Convergence and machine learning predictions of monkhorst-pack k-points and plane-wave cut-off in high-throughput dft calculations," Computational materials science **161**, 300–308 (2019).
[124]K. Choudhary, F. Y. P. Congo, T. Liang, C. Becker, R. G. Hennig, and F. Tavazza, "Evaluation and comparison of classical interatomic potentials through a user-friendly interactive web-interface," Scientific Data **4**, 160125 (2017).
[125]D. A. Case, T. E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K. M. Merz Jr., A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, "The amber biomolecular simulation programs," J. Comput. Chem. **26**, 1668–1688 (2005).
[126]J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. De Groot, H. Grubmüller, and A. D. MacKerell Jr, "Charmm36m: an improved force field for folded and intrinsically disordered proteins," Nature methods **14**, 71–73 (2017).
[127]I. Novoselov, A. Yanilkin, A. Shapeev, and E. Podryabinkin, "Moment tensor potentials as a promising tool to study diffusion processes," Computational Materials Science **164**, 46–56 (2019).
[128]R. Drautz, "Atomic cluster expansion for accurate and transferable interatomic potentials," Phys. Rev. B **99**, 014104 (2019).
[129]A. P. Bartók, R. Kondor, and G. Csányi, "On representing chemical environments," Phys. Rev. B **87**, 184115 (2013).
[130]L. Zhang, J. Han, H. Wang, R. Car, and W. E, "Deep potential molecular dynamics: A scalable model with the accuracy of quantum mechanics," Phys. Rev. Lett. **120**, 143001 (2018).
[131]V. Botu and R. Ramprasad, "Adaptive machine learning framework to accelerate ab initio molecular dynamics," International Journal of Quantum Chemistry **115**, 1074–1083 (2015), https://onlinelibrary.wiley.com/doi/pdf/10.1002/qua.24836.
[132]J. S. Smith, B. Nebgen, N. Mathew, J. Chen, N. Lubbers, L. Burakovsky, S. Tretiak, H. A. Nam, T. Germann, S. Fensin, and K. Barros, "Automated discovery of a robust interatomic potential for aluminum," Nature Communications **12** (2021), 10.1038/s41467-021-21376-0.
[133]K. Choudhary, B. DeCost, L. Major, K. Butler, J. Thiyagalingam, and F. Tavazza, "Unified graph neural network force-field for the periodic table: solid state applications," Digital Discovery (2023).
[134]A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. B. Jensen, J. Kermode, J. R. Kitchin, E. L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, and K. W. Jacobsen, "The atomic simulation environment—a python library for working with atoms," Journal of Physics: Condensed Matter **29**, 273002 (2017).
[135]K. Choudhary, T. Yildirim, D. W. Siderius, A. G. Kusne, A. McDannald, and D. L. Ortiz-Montalvo, "Graph neural network predictions of metal organic framework co2 adsorption properties," Computational Materials Science **210**, 111388 (2022).
[136]D. Dubbeldam, S. Calero, D. E. Ellis, and R. Q. Snurr, "Raspa: molecular simulation software for adsorption and diffusion in flexible nanoporous materials," Molecular Simulation **42**, 81–101 (2016), https://doi.org/10.1080/08927022.2015.1010082.
[137]S. Páll, A. Zhmurov, P. Bauer, M. Abraham, M. Lundborg, A. Gray, B. Hess, and E. Lindahl, "Heterogeneous parallelization and acceleration of molecular dynamics simulations in GROMACS," The Journal of Chemical Physics **153** (2020), 10.1063/5.0018516, 134110, https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0018516/16736127/134110_1_online.pdf.
[138]S.-T. Tsai, Z. Smith, and P. Tiwary, "Sgoop-d: Estimating kinetic distances and reaction coordinate dimensionality for rare event systems from biased/unbiased simulations," Journal of Chemical Theory and Computation **17**, 6757–6765 (2021).
[139]S. Mehdi, D. Wang, S. Pant, and P. Tiwary, "Accelerating all-atom simulations and gaining mechanistic understanding of biophysical systems through state predictive information bottleneck," Journal of Chemical Theory and Computation **18**, 3231–3238 (2022).
[140]D. Wang and P. Tiwary, "State predictive information bottleneck," The Journal of Chemical Physics **154**, 134111 (2021).
[141]A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, "Commentary: The materials project: A materials genome approach to accelerating materials innovation," APL Materials **1**, 011002 (2013), https://doi.org/10.1063/1.4812323.
[142]L. Ruddigkeit, R. van Deursen, L. C. Blum, and J.-L. Reymond, "Enumeration of 166 billion organic small molecules in the chemical universe database gdb-17," Journal of Chemical Information and Modeling **52**, 2864–2875 (2012).
[143]R. Ramakrishnan, P. O. Dral, M. Rupp, and O. A. von Lilienfeld, "Quantum chemistry structures and properties of 134 kilo molecules," Scientific Data **1** (2014).
[144]F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research **12**, 2825–2830 (2011).
[145]A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," (2019), arXiv:1912.01703 [cs.LG].
[146]M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on het-

erogeneous systems," (2015), software available from tensorflow.org.
[147] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, "JAX: composable transformations of Python+NumPy programs," (2018).
[148] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, J. Davison, S. Shleifer, P. von Platen, C. Ma, Y. Jernite, J. Plu, C. Xu, T. L. Scao, S. Gugger, M. Drame, Q. Lhoest, and A. M. Rush, "Huggingface's transformers: State-of-the-art natural language processing," (2020), arXiv:1910.03771 [cs.CL].
[149] M. A. Nielsen and I. L. Chuang, "Quantum computation and quantum information," Phys. Today **54**, 60 (2001).
[150] K. Choudhary, "Quantum computation for predicting electron and phonon properties of solids," Journal of Physics: Condensed Matter **33**, 385501 (2021).
[151] J. S. Kottmann, S. Alperin-Lea, T. Tamayo-Mendoza, A. Cervera-Lierta, C. Lavigne, T.-C. Yen, V. Verteletskyi, P. Schleich, A. Anand, M. Degroote, S. Chaney, M. Kesibi, N. G. Curnow, B. Solo, G. Tsilimigkounakis, C. Zendejas-Morales, A. F. Izmaylov, and A. Aspuru-Guzik, "Tequila: a platform for rapid development of quantum algorithms," Quantum Science and Technology **6**, 024009 (2021).
[152] C. Developers, "Cirq," (2022), See full list of authors on Github: https://github .com/quantumlib/Cirq/graphs/contributors.
[153] R. H. PIERSON and E. A. FAY, "Guidelines for interlaboratory testing programs," Analytical Chemistry **31**, 25A–49A (1959).
[154] Y. Zuo, C. Chen, X. Li, Z. Deng, Y. Chen, J. Behler, G. Csányi, A. V. Shapeev, A. P. Thompson, M. A. Wood, and S. P. Ong, "Performance and cost assessment of machine learning interatomic potentials," The Journal of Physical Chemistry A **124**, 731–745 (2020).
[155] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, "Materials design and discovery with high-throughput density functional theory: The open quantum materials database (oqmd)," JOM **65**, 1501–1509 (2013).
[156] S. Kirklin, J. E. Saal, B. Meredig, A. Thompson, J. W. Doak, M. Aykol, S. Rühl, and C. Wolverton, "The open quantum materials database (oqmd): assessing the accuracy of dft formation energies," npj Computational Materials **1**, 15010 (2015).
[157] S. Curtarolo, W. Setyawan, G. L. Hart, M. Jahnatek, R. V. Chepulskii, R. H. Taylor, S. Wang, J. Xue, K. Yang, O. Levy, M. J. Mehl, H. T. Stokes, D. O. Demchenko, and D. Morgan, "Aflow: An automatic framework for high-throughput materials discovery," Computational Materials Science **58**, 218–226 (2012).
[158] K. Li, B. DeCost, K. Choudhary, M. Greenwood, and J. Hattrick-Simpers, "A critical examination of robustness and generalizability of machine learning prediction of materials properties," npj Computational Materials **9**, 55 (2023).
[159] S. Gong, T. Xie, Y. Shao-Horn, R. Gomez-Bombarelli, and J. C. Grossman, "Examining graph neural networks for crystal structures: limitations and opportunities for capturing periodicity," arXiv preprint arXiv:2208.05039 (2022).
[160] L. Ward, R. Liu, A. Krishna, V. I. Hegde, A. Agrawal, A. Choudhary, and C. Wolverton, "Including crystal structure attributes in machine learning models of formation energies via voronoi tessellations," Physical Review B **96**, 024104 (2017).

Fig S1:AI-MLFF-energy-mlearn_Ni-test-mae.csv.zip



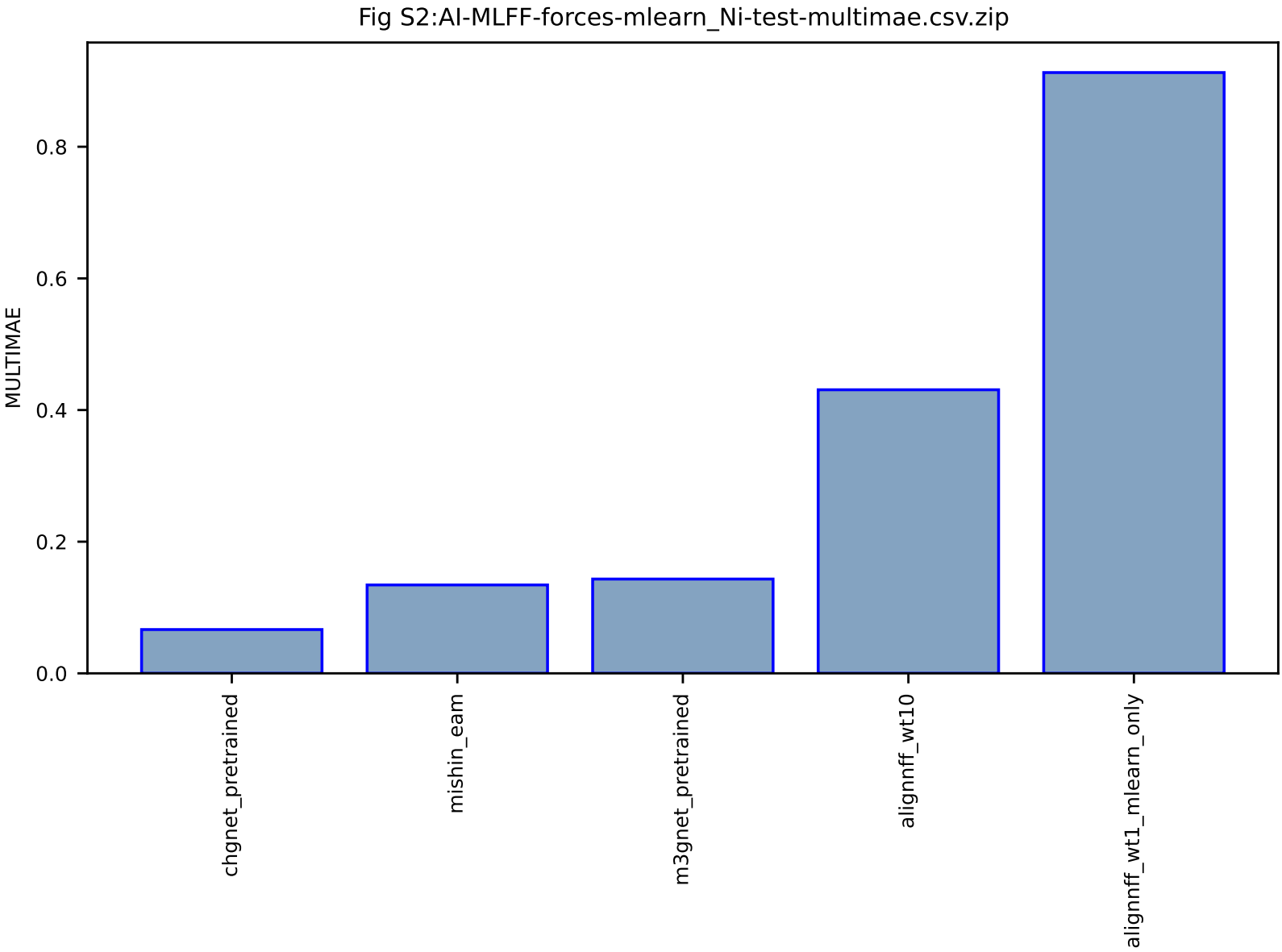


Fig S2:AI-MLFF-forces-mlearn_Ni-test-multimae.csv.zip
MULTIMAE
0.0
0.2
0.4
0.6
0.8
chgnet_pretrained
mishin_eam
m3gnet_pretrained
alignnff_wt10
alignnff_wt1_mlearn_only

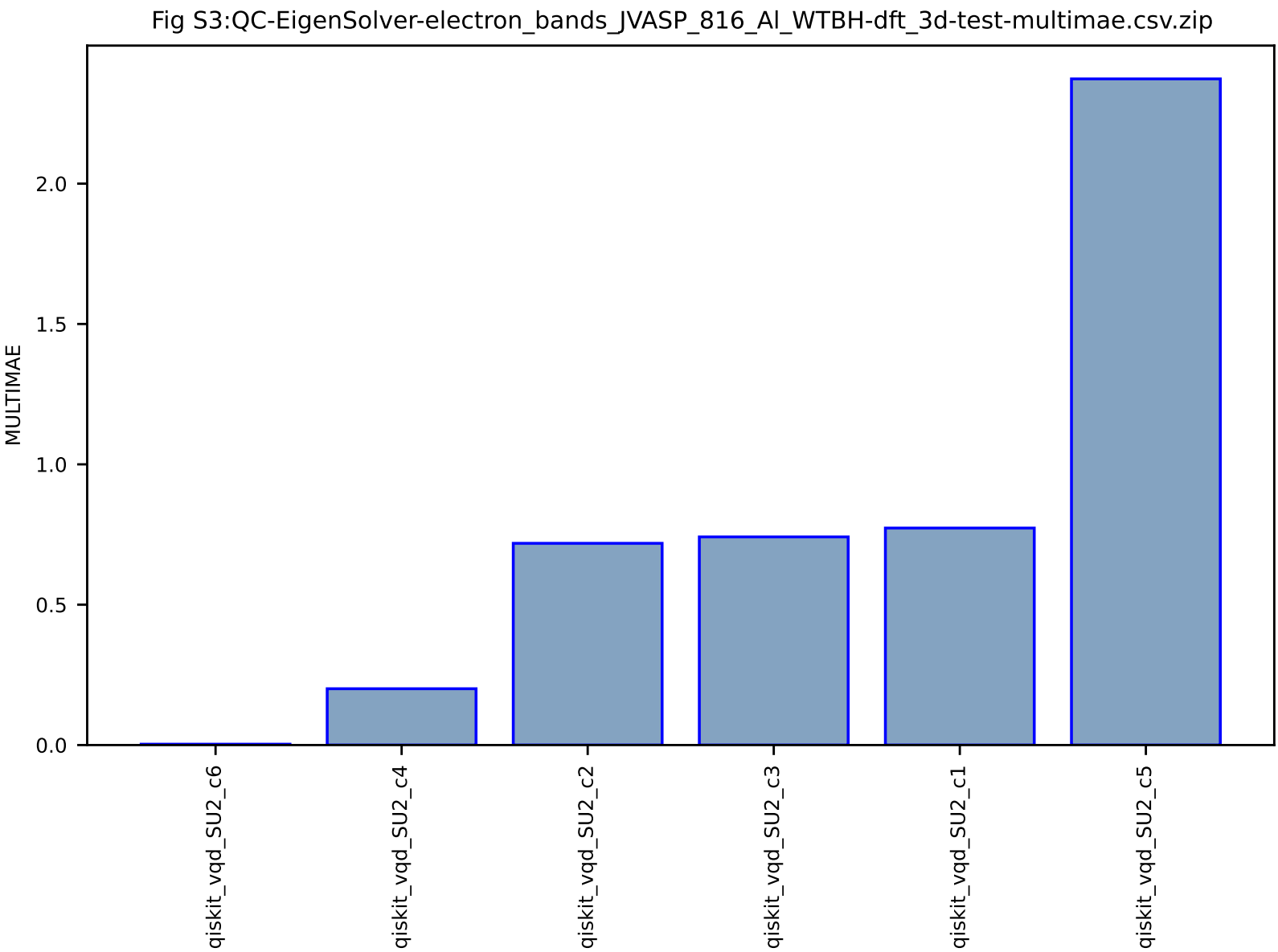


Fig S3:QC-EigenSolver-electron_bands_JVASP_816_Al_WTBH-dft_3d-test-multimae.csv.zip
MULTIMAE
0.0
0.5
1.0
1.5
2.0
qiskit_vqd_SU2_c6
qiskit_vqd_SU2_c4
qiskit_vqd_SU2_c2
qiskit_vqd_SU2_c3
qiskit_vqd_SU2_c1
qiskit_vqd_SU2_c5

Fig S4:FF-SinglePropertyPrediction-bulk_modulus_JVASP_867_Cu-dft_3d-test-mae.csv.zip

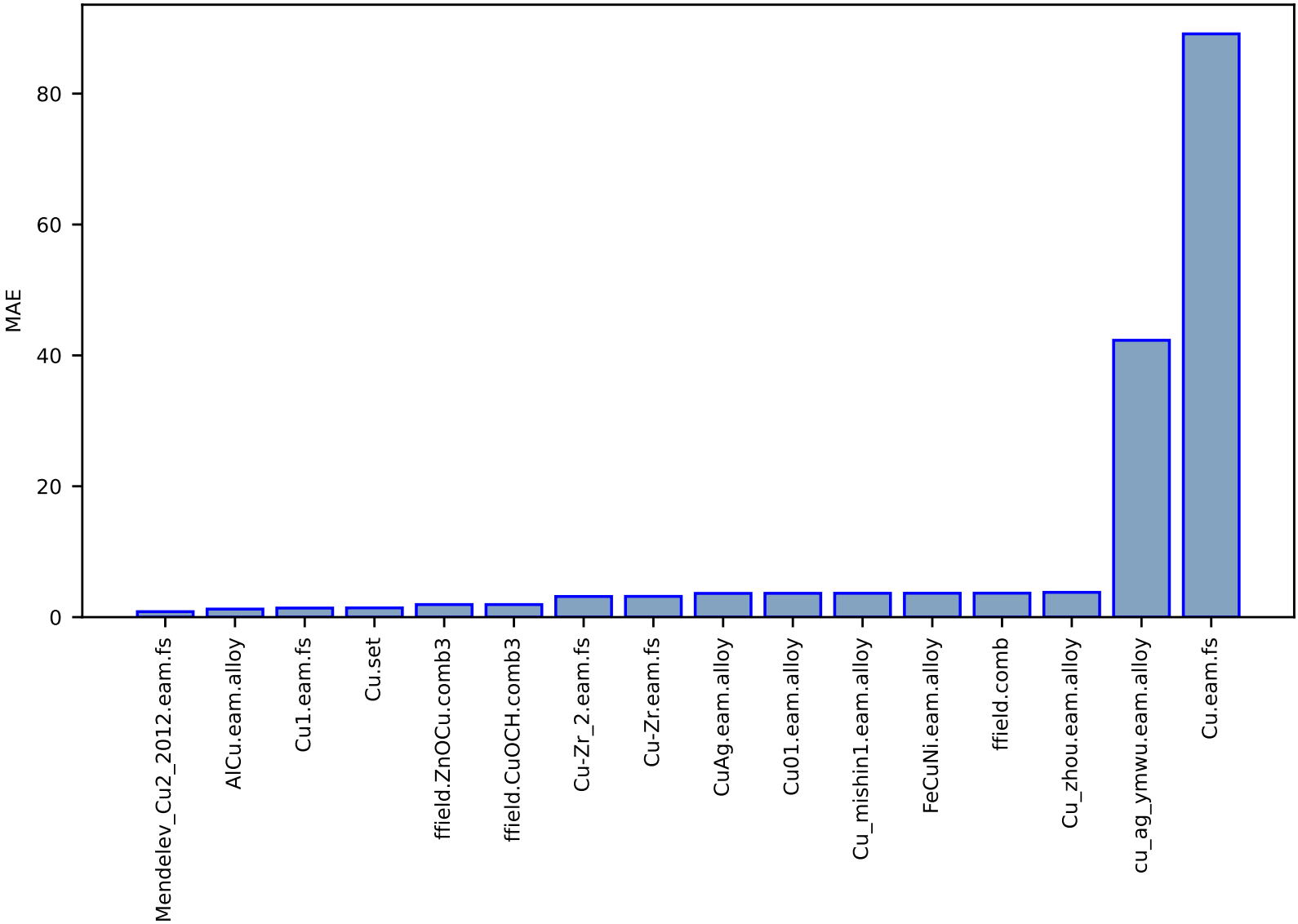


Fig S5:FF-SinglePropertyPrediction-bulk_modulus_JVASP_816_Al-dft_3d-test-mae.csv.zip

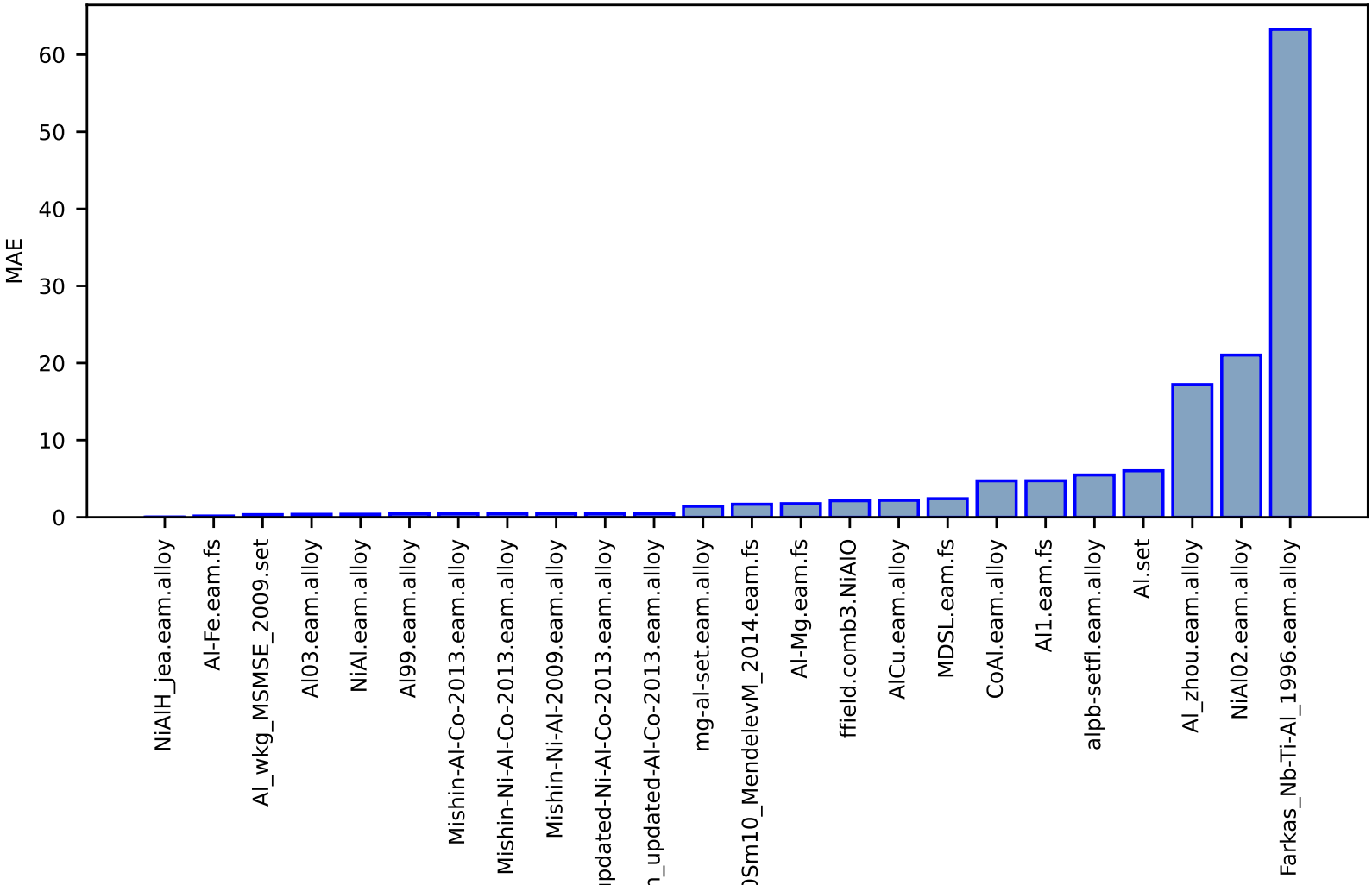


Fig S6:EXP-Spectra-co2_RM_8852-nist_isodb-test-multimae.csv.zip

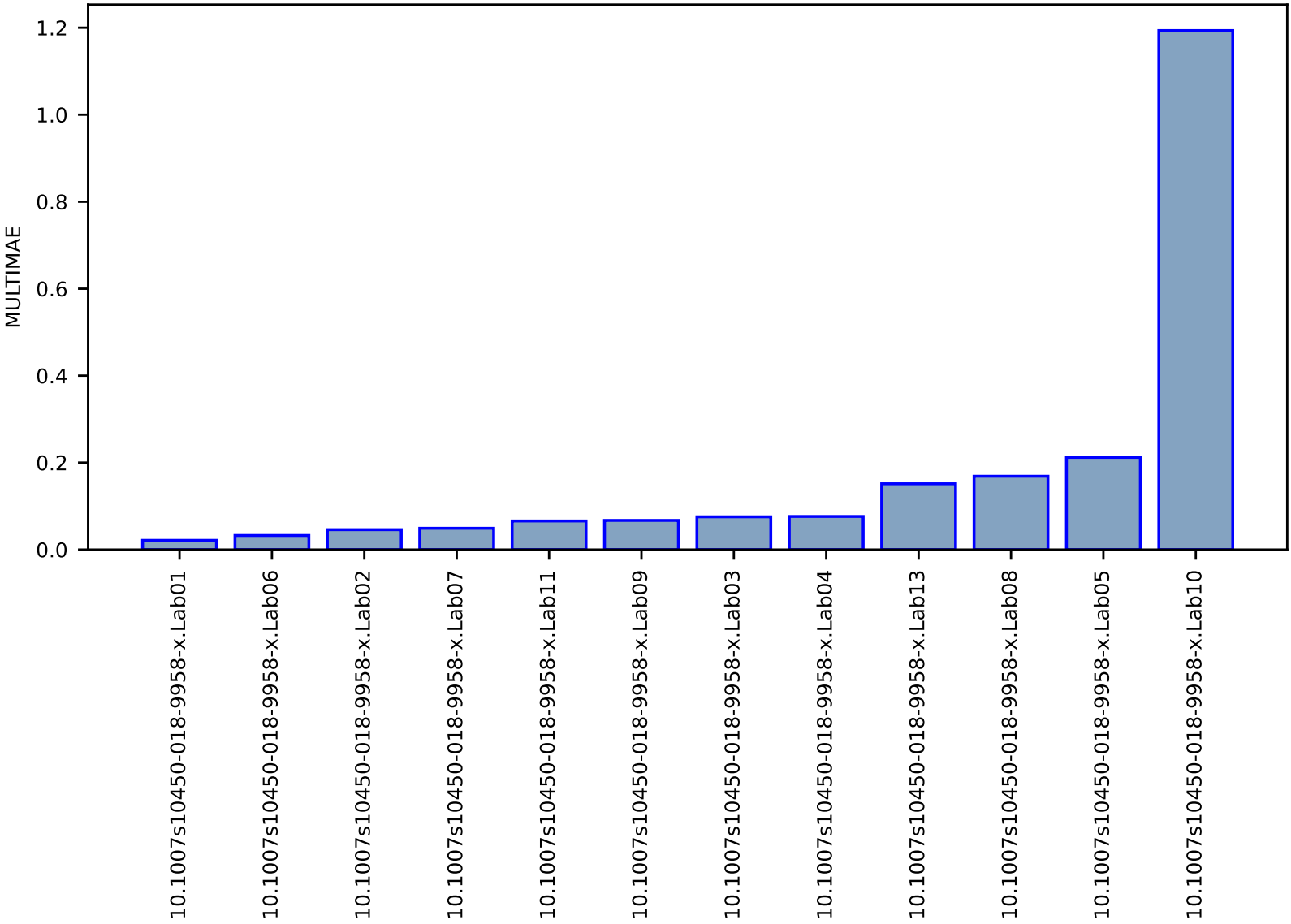


Fig S7:AI-TextClass-categories-pubchem-test-acc.csv.zip

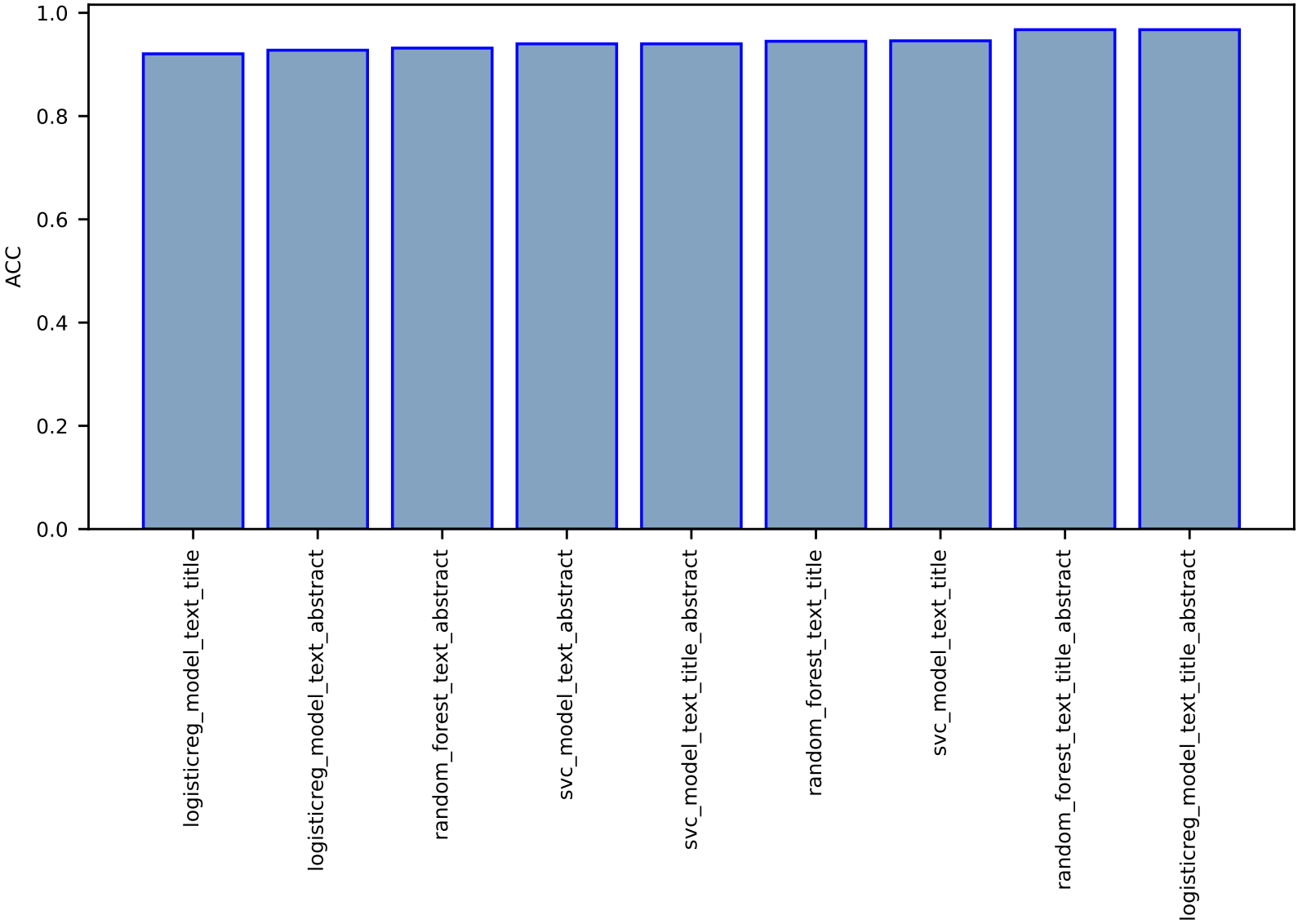


Fig S8:AI-TextClass-categories-arXiv-test-acc.csv.zip

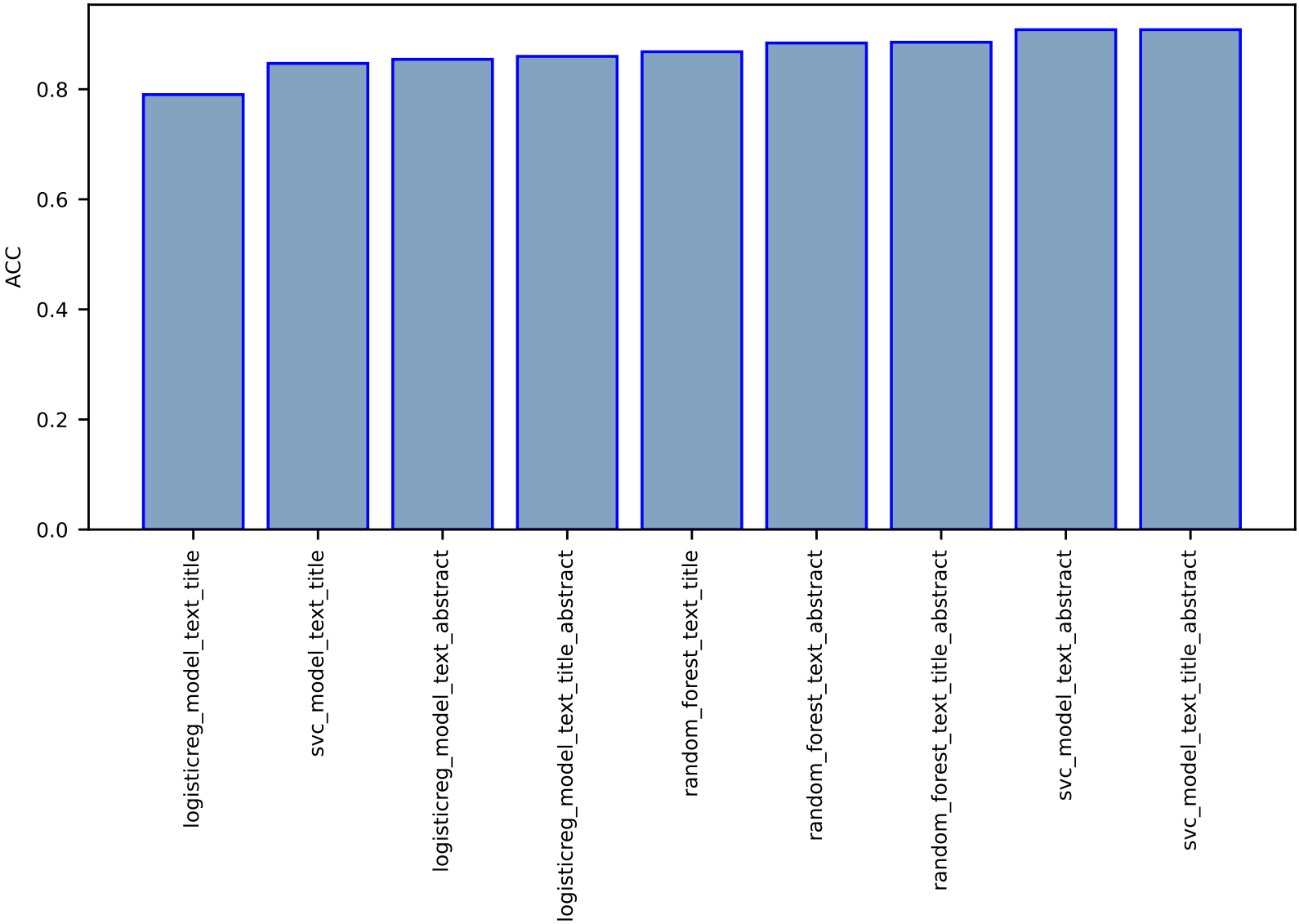


Fig S9:AI-ImageClass-bravais_class-stem_2d_image-test-acc.csv.zip

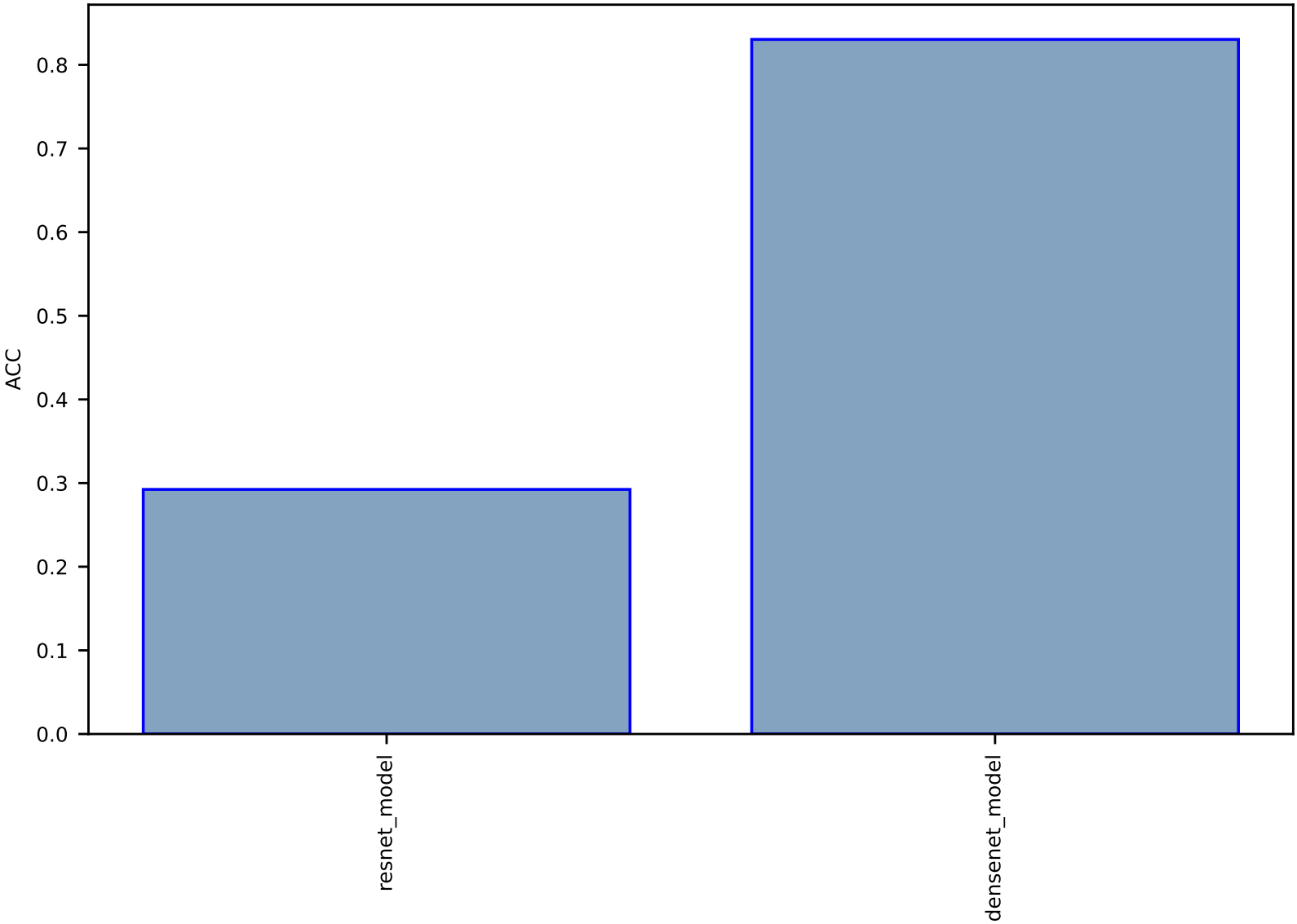
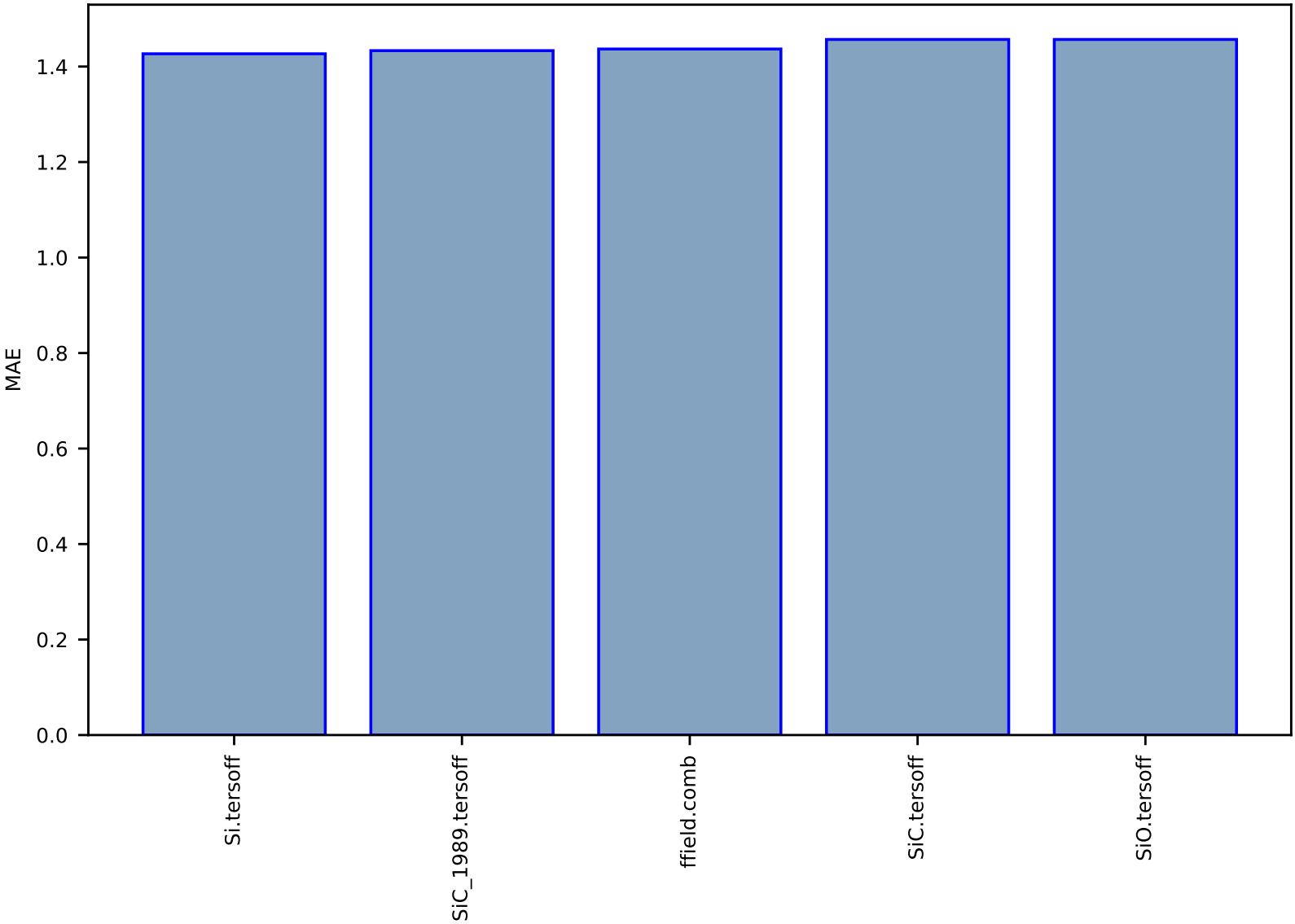


Fig S10:FF-SinglePropertyPrediction-bulk_modulus_JVASP_1002_Si-dft_3d-test-mae.csv.zip



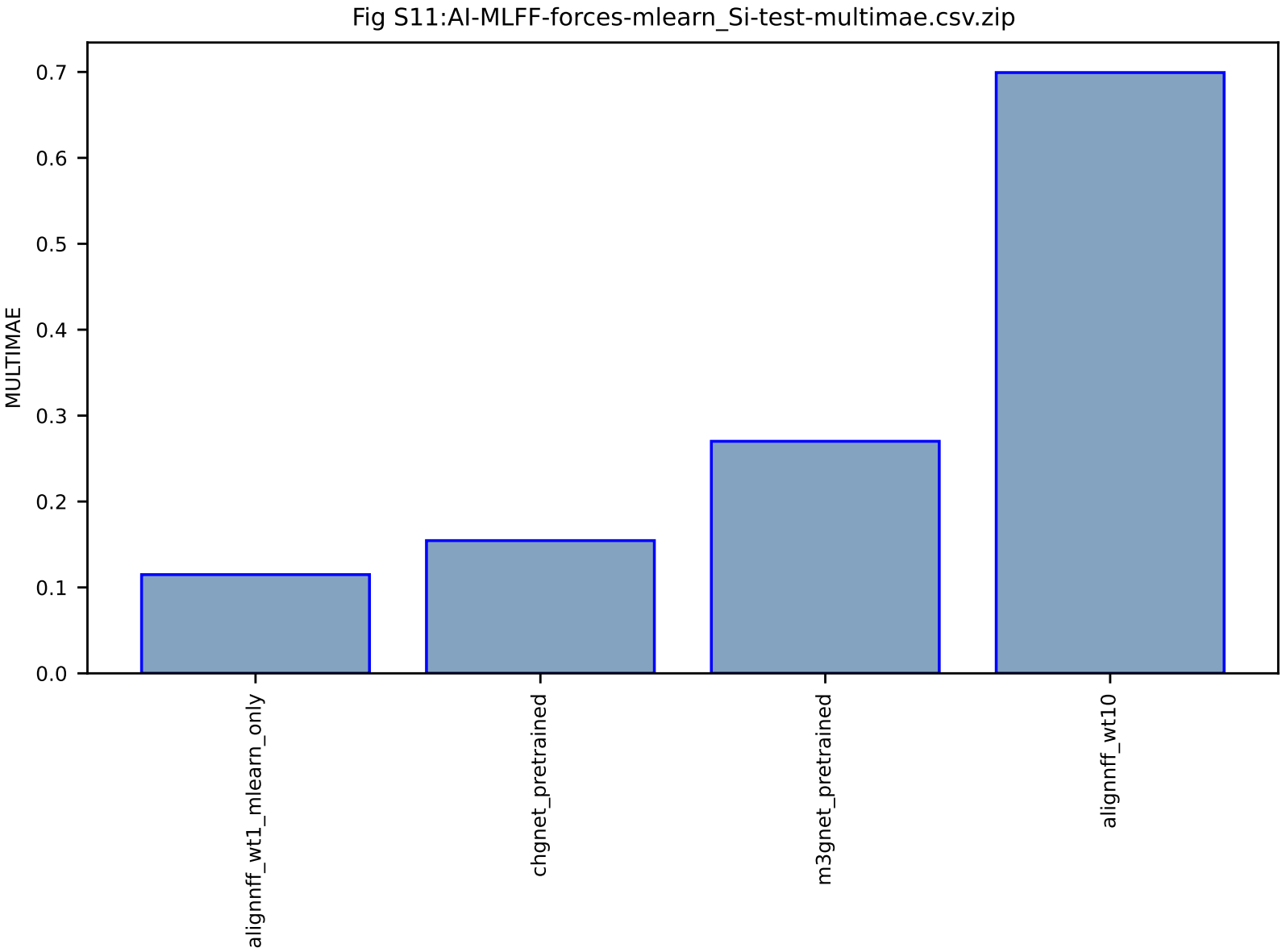


Fig S11:AI-MLFF-forces-mlearn_Si-test-multimae.csv.zip
MULTIMAE
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
alignnff_wt1_mlearn_only
chgnet_pretrained
m3gnet_pretrained
alignnff_wt10

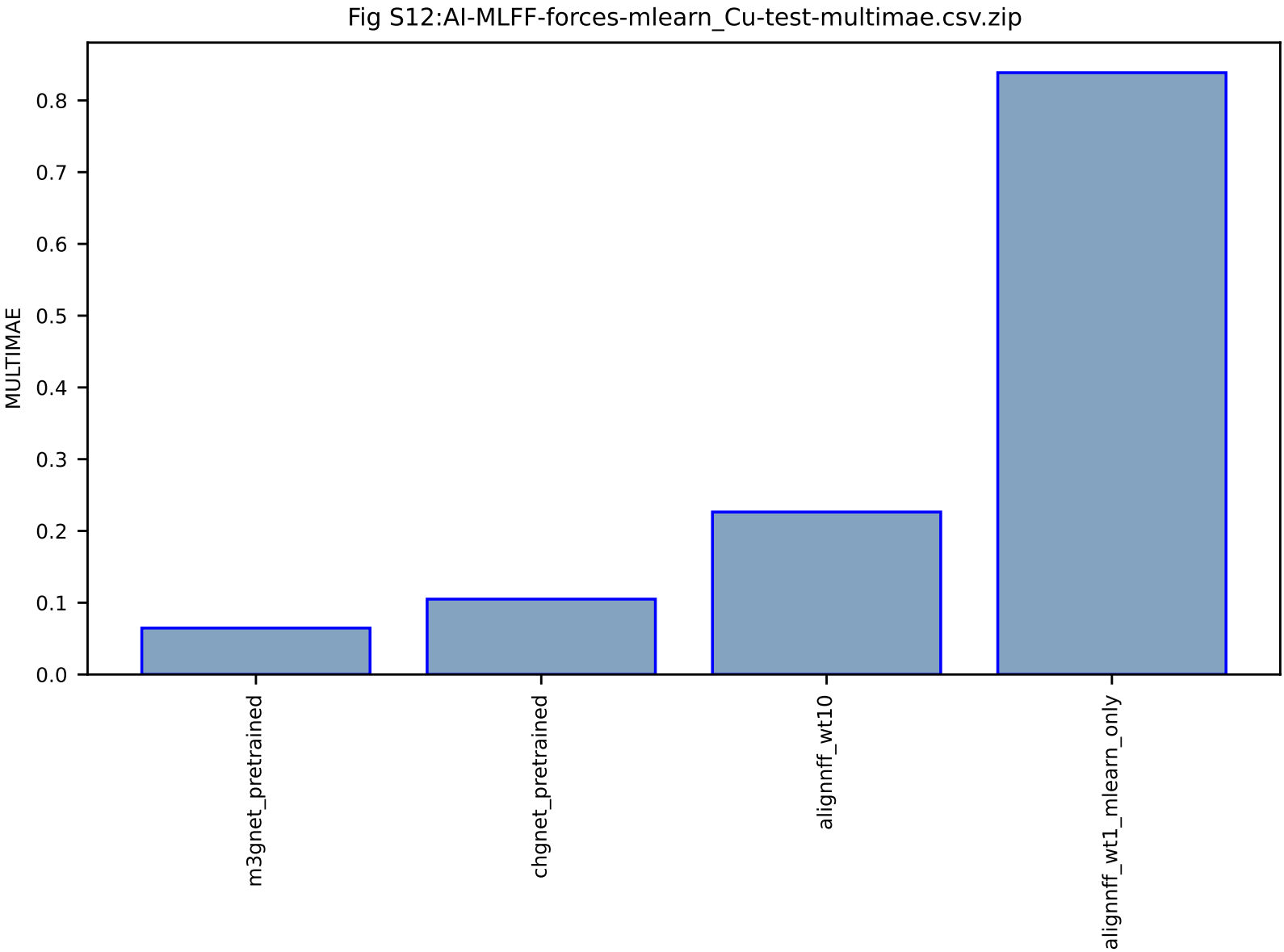


Fig S12:AI-MLFF-forces-mlearn_Cu-test-multimae.csv.zip
MULTIMAE
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
m3gnet_pretrained
chgnet_pretrained
alignnff_wt10
alignnff_wt1_mlearn_only

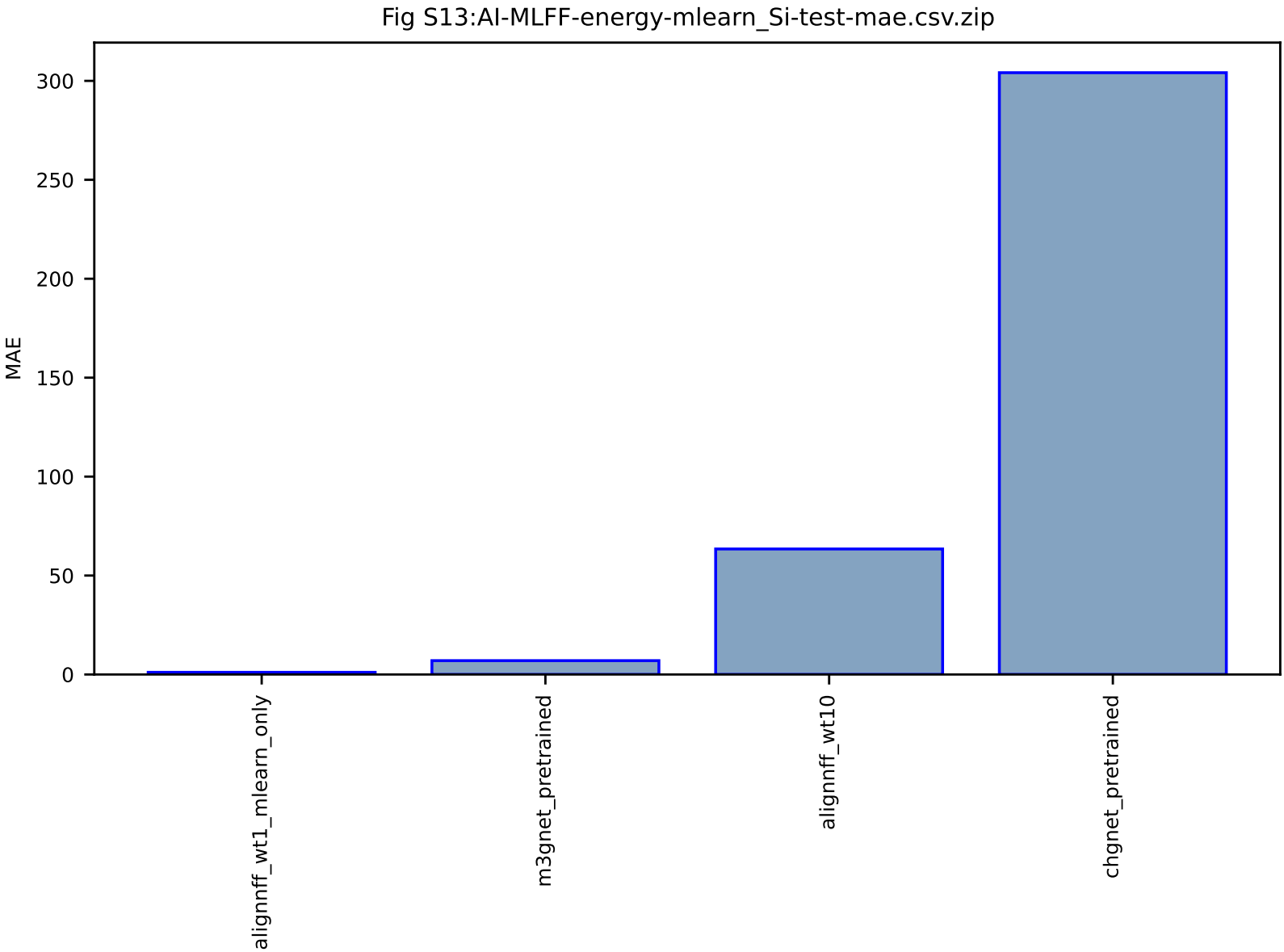


Fig S13:AI-MLFF-energy-mlearn_Si-test-mae.csv.zip
MAE
300
250
200
150
100
50
0
alignnff_wt1_mlearn_only
m3gnet_pretrained
alignnff_wt10
chgnet_pretrained

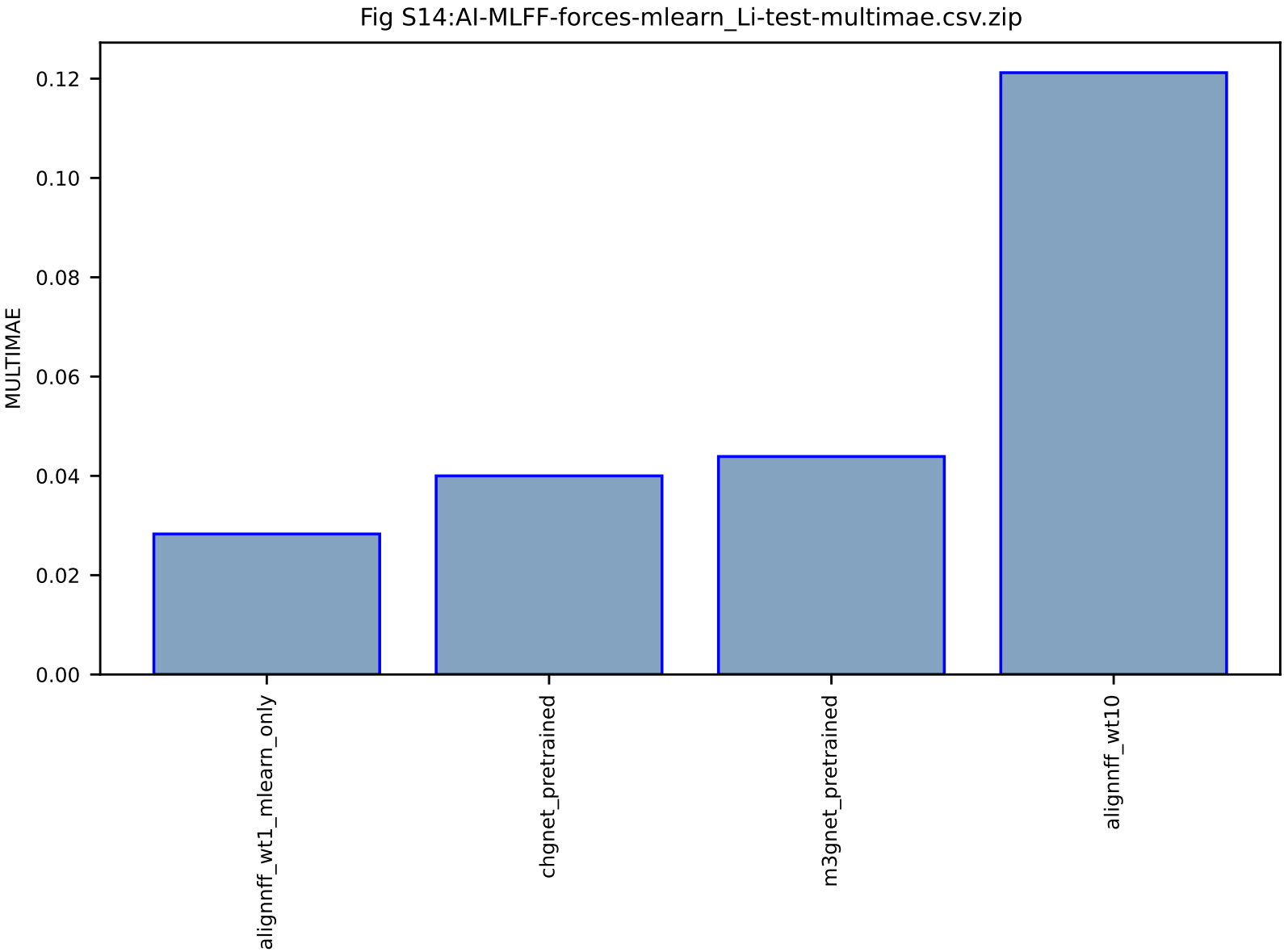


Fig S14:AI-MLFF-forces-mlearn_Li-test-multimae.csv.zip
MULTIMAE
0.00
0.02
0.04
0.06
0.08
0.10
0.12
alignnff_wt1_mlearn_only
chgnet_pretrained
m3gnet_pretrained
alignnff_wt10

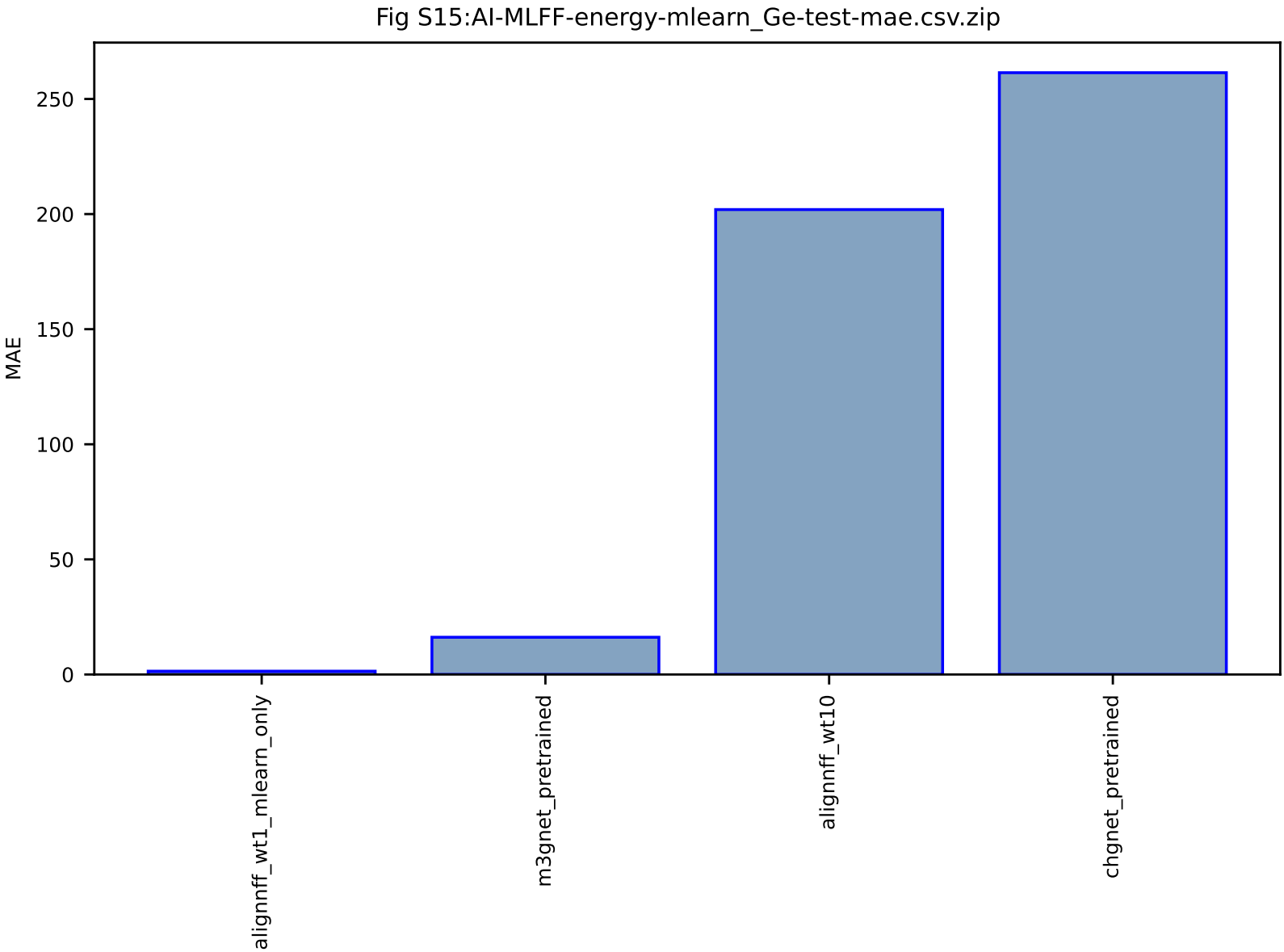


Fig S15:AI-MLFF-energy-mlearn_Ge-test-mae.csv.zip
MAE
0
50
100
150
200
250
alignnff_wt1_mlearn_only
m3gnet_pretrained
alignnff_wt10
chgnet_pretrained

Fig S16:AI-MLFF-forces-mlearn_Mo-test-multimae.csv.zip

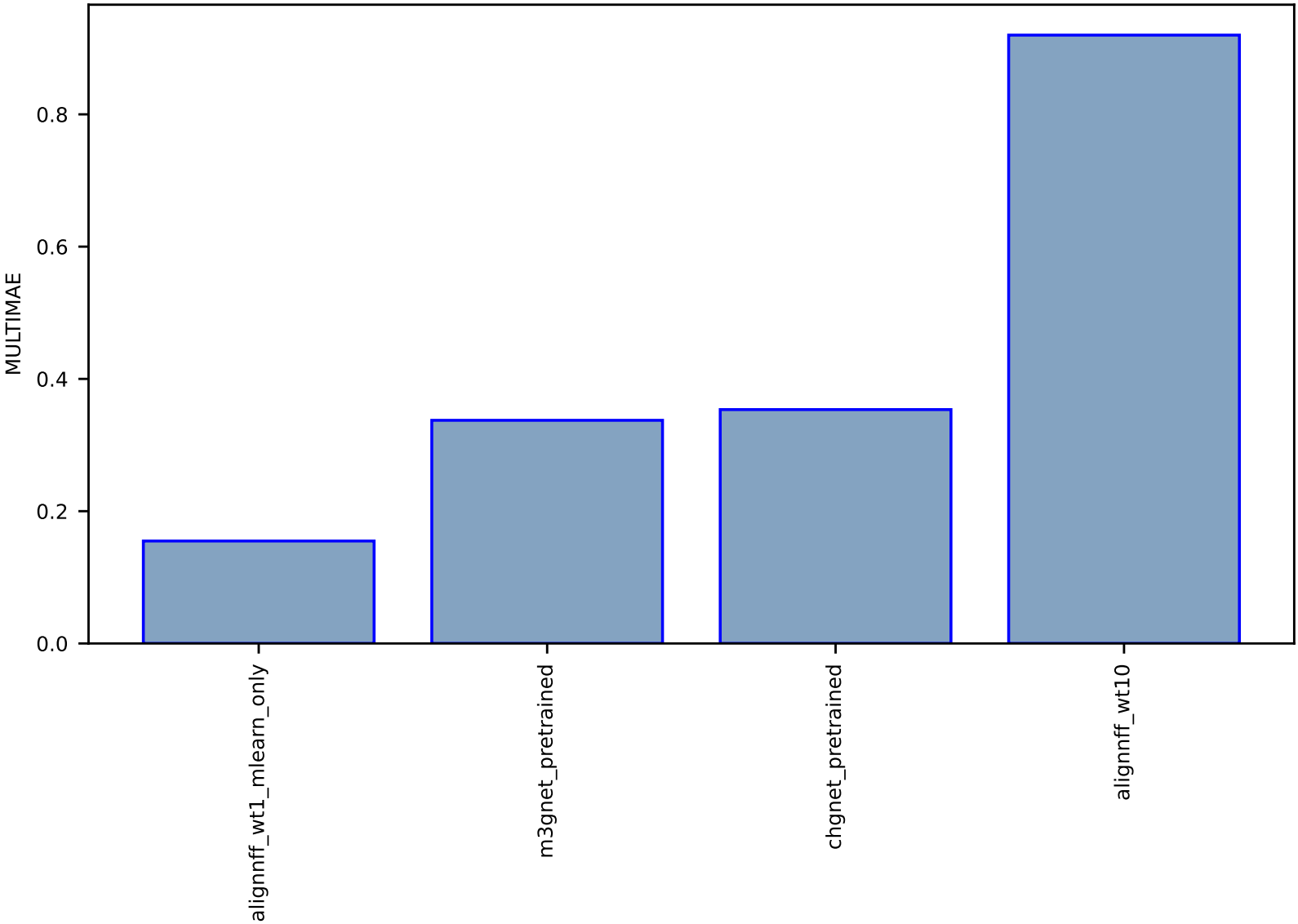
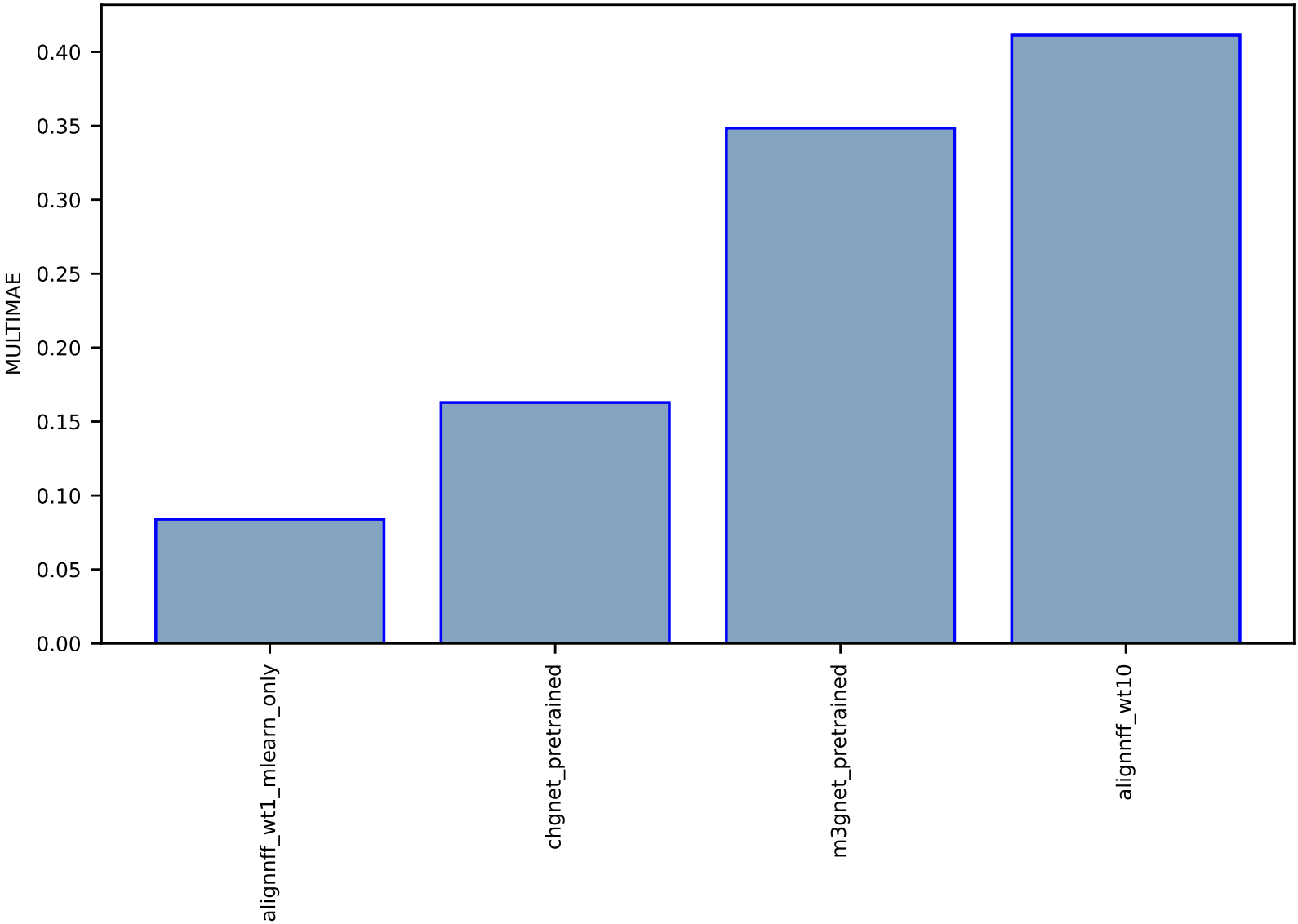


Fig S17:AI-MLFF-forces-mlearn_Ge-test-multimae.csv.zip



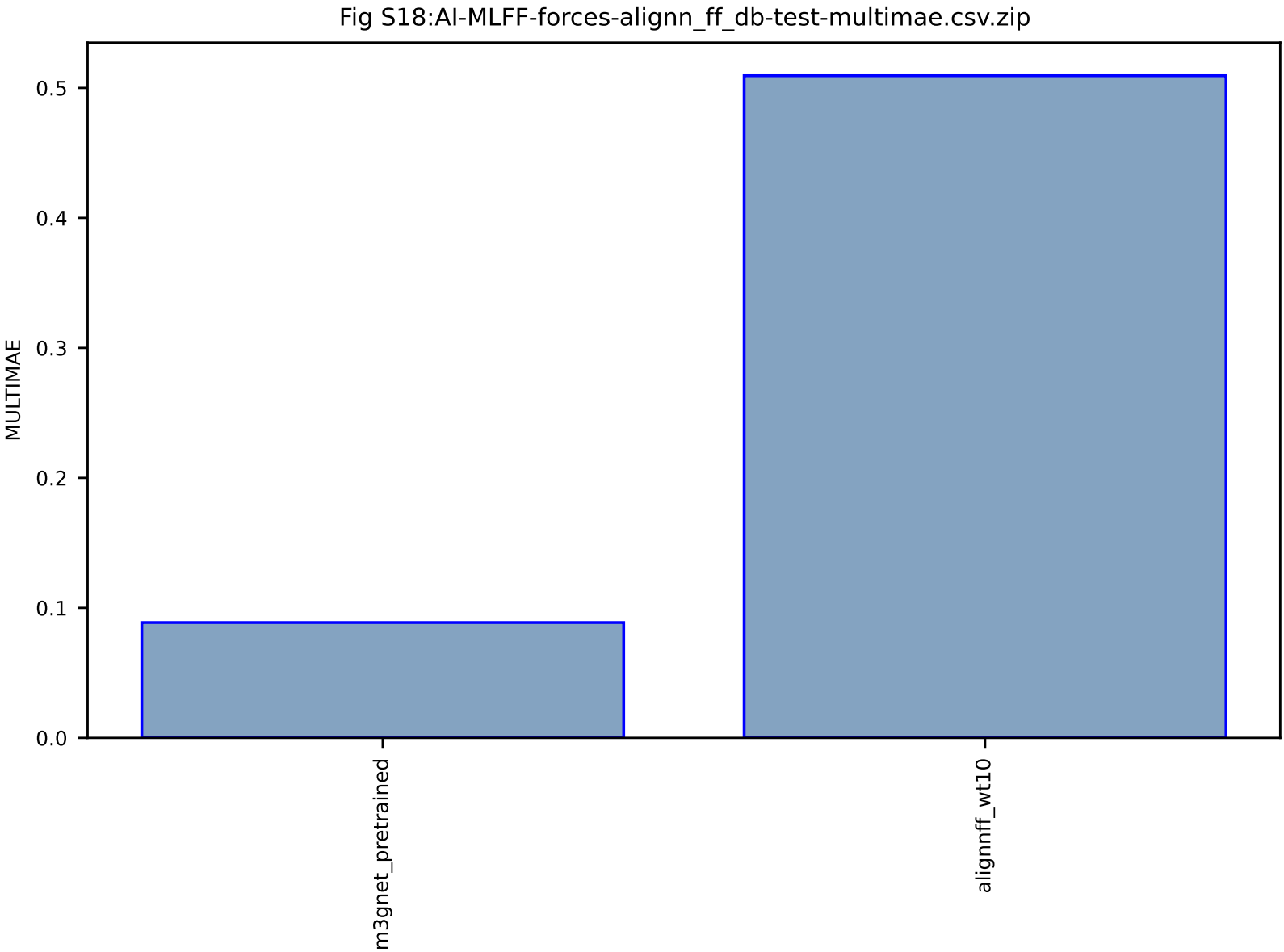


Fig S18:AI-MLFF-forces-alignn_ff_db-test-multimae.csv.zip
MULTIMAE
0.0
0.1
0.2
0.3
0.4
0.5
m3gnet_pretrained
alignnff_wt10

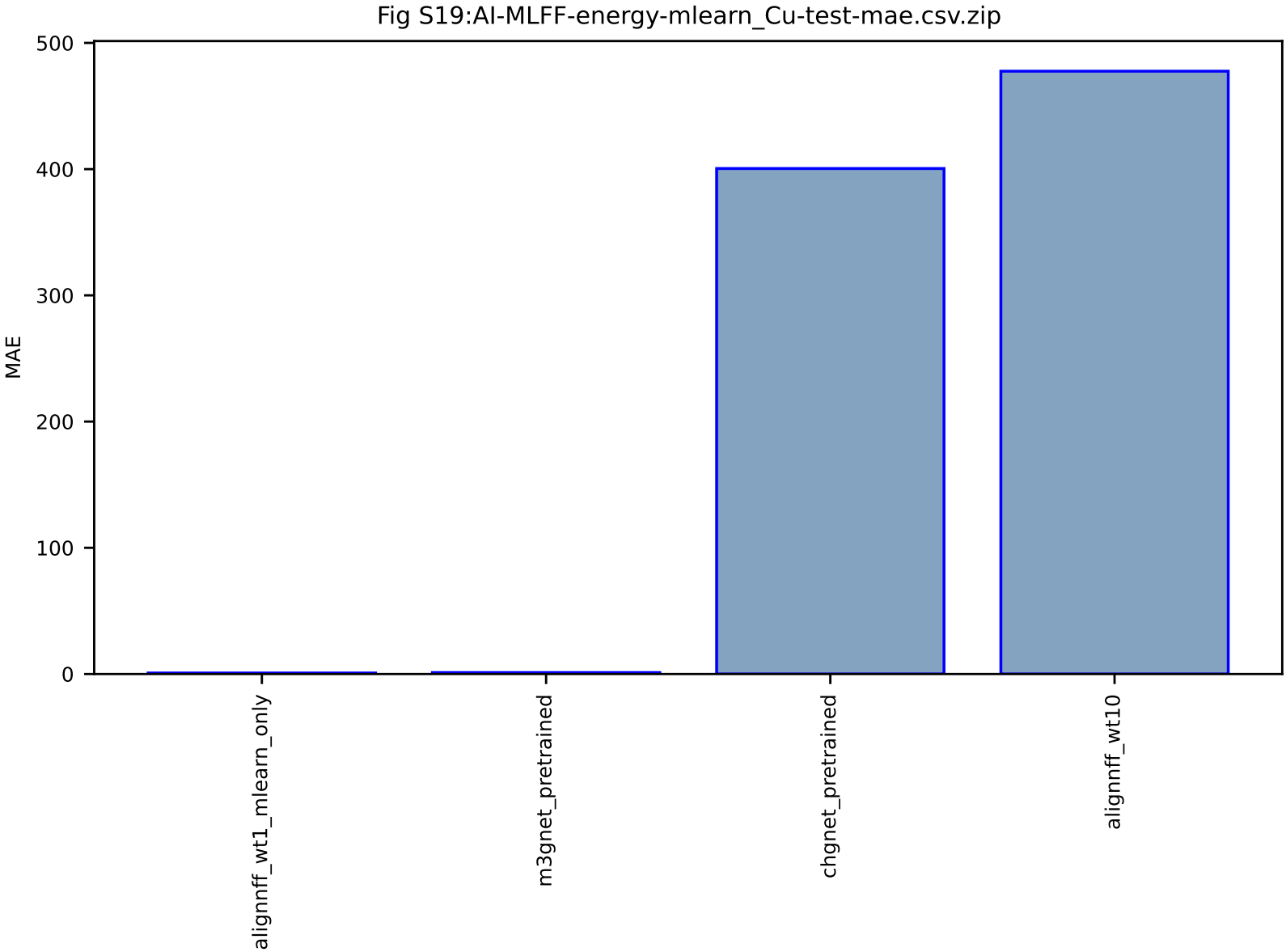


Fig S19:AI-MLFF-energy-mlearn_Cu-test-mae.csv.zip
MAE
500
400
300
200
100
0
alignnff_wt1_mlearn_only
m3gnet_pretrained
chgnet_pretrained
alignnff_wt10

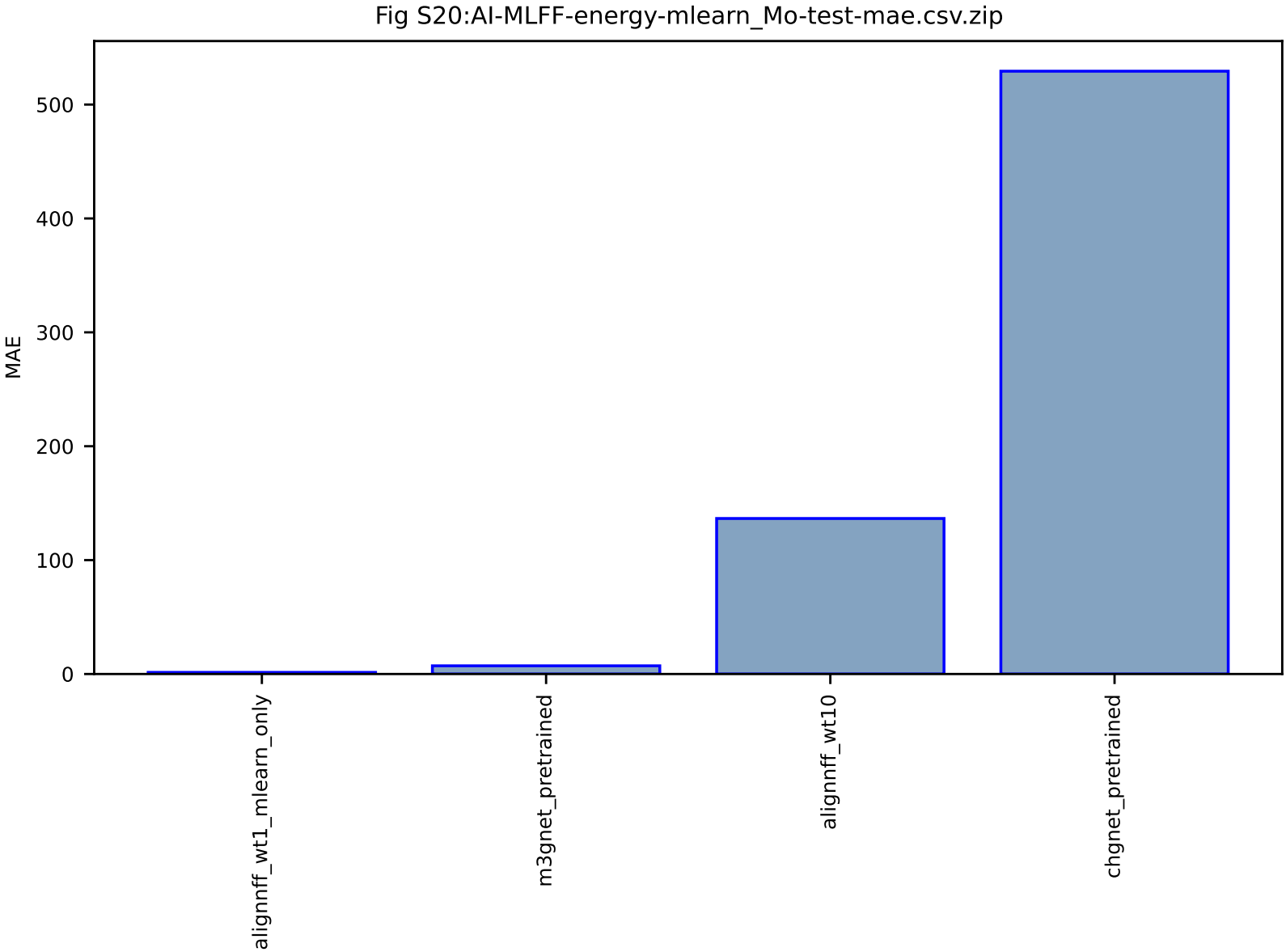


Fig S20:AI-MLFF-energy-mlearn_Mo-test-mae.csv.zip
MAE
0
100
200
300
400
500
alignnff_wt1_mlearn_only
m3gnet_pretrained
alignnff_wt10
chgnet_pretrained

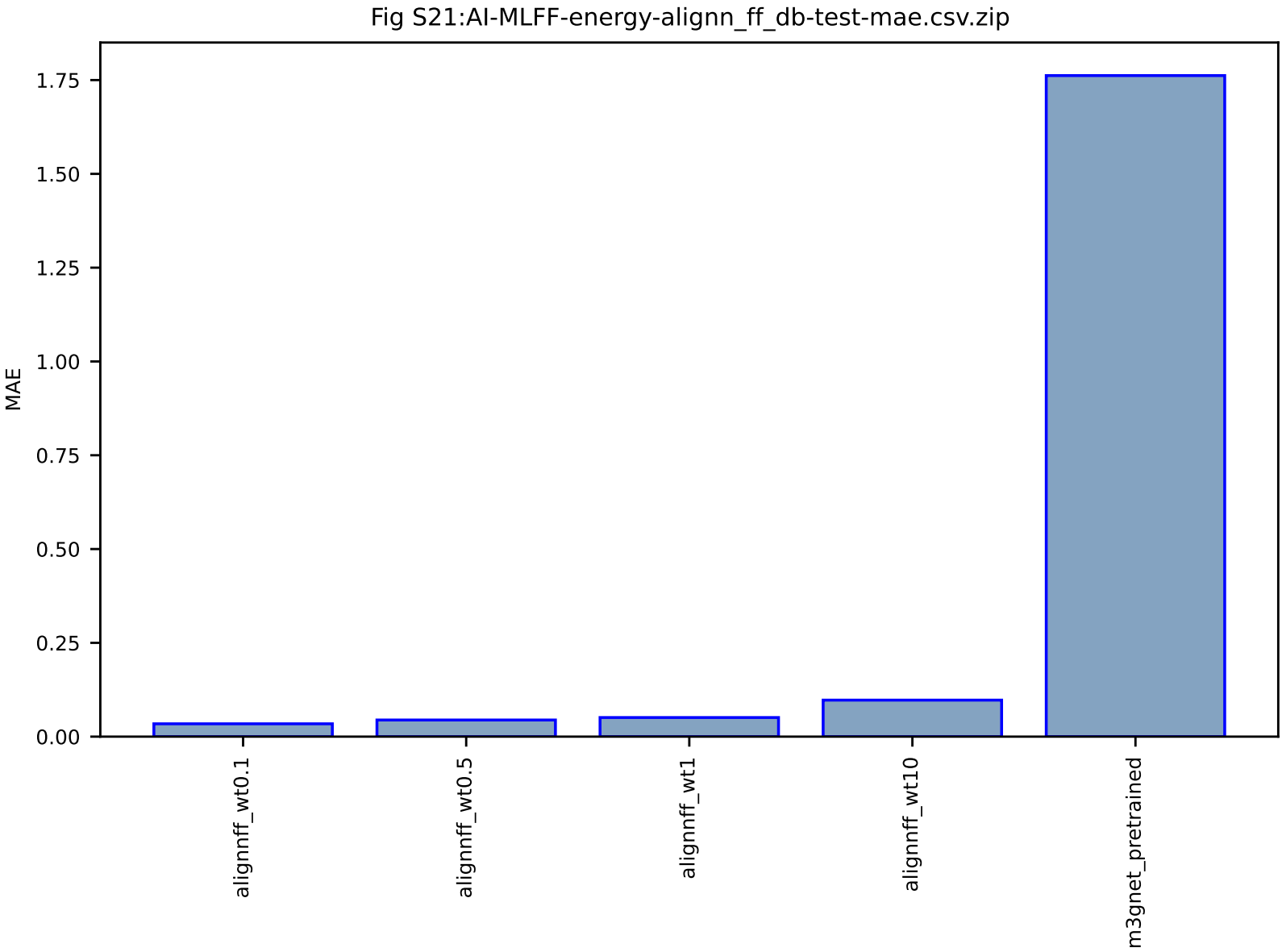


Fig S21:AI-MLFF-energy-alignn_ff_db-test-mae.csv.zip
MAE
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
alignnff_wt0.1
alignnff_wt0.5
alignnff_wt1
alignnff_wt10
m3gnet_pretrained

Fig S22:AI-MLFF-energy-mlearn_Li-test-mae.csv.zip

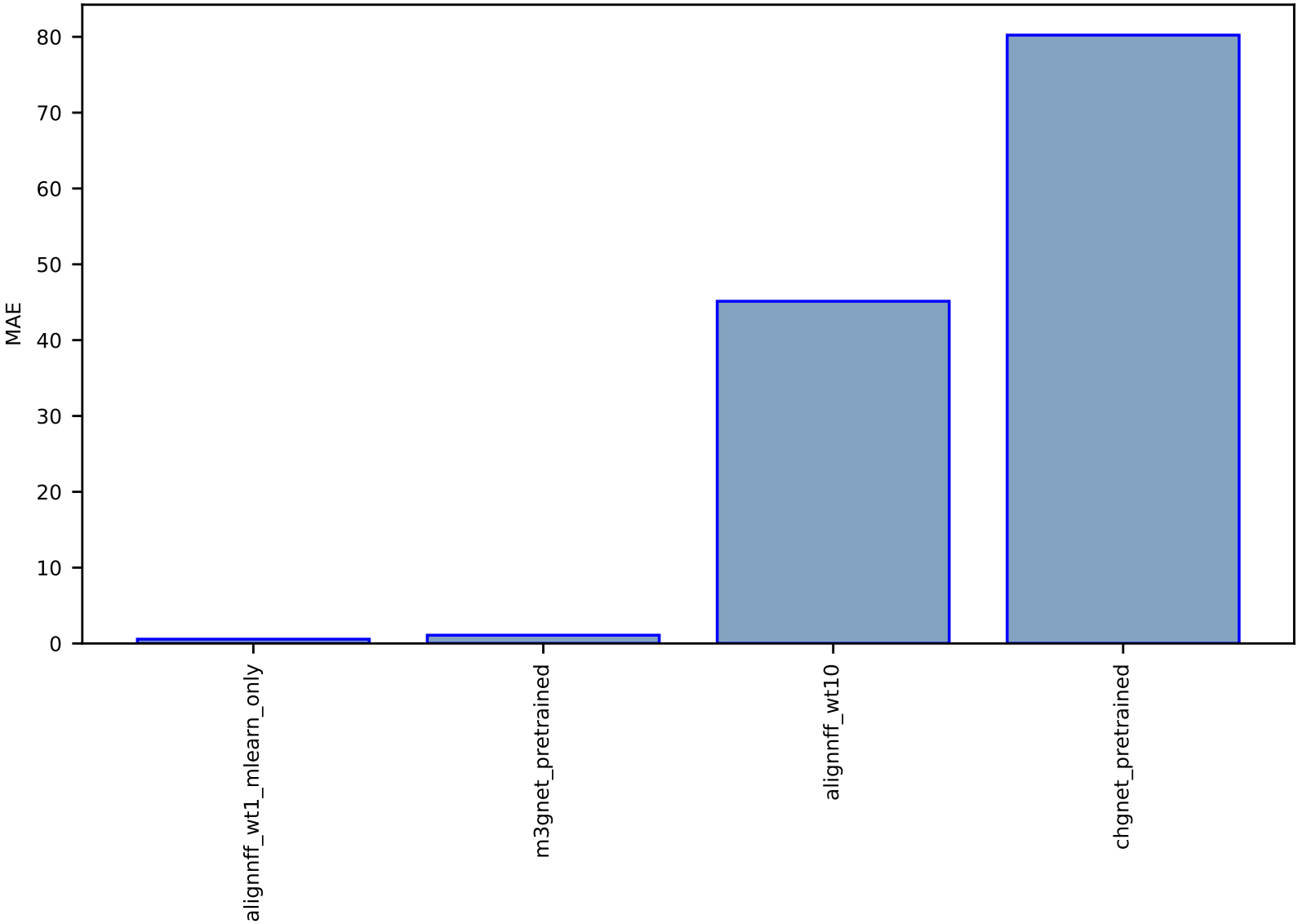


Fig S23:AI-SinglePropertyPrediction-ef-vacancydb_oxides_train_test-test-mae.csv.zip

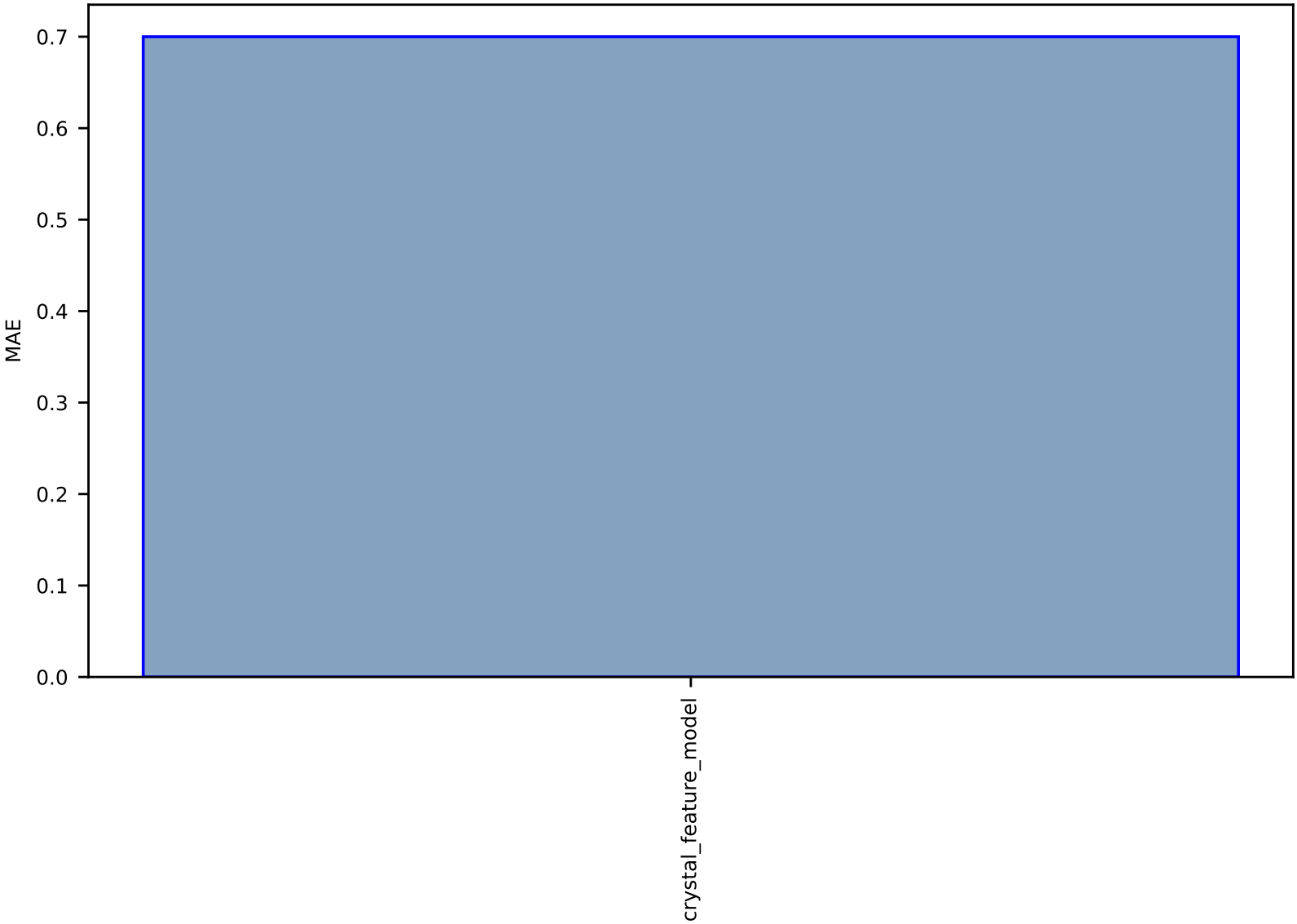


Fig S24:EXP-Spectra-stress_strain_vibroscopy_kevlar129-midas-test-multimae.csv.zip

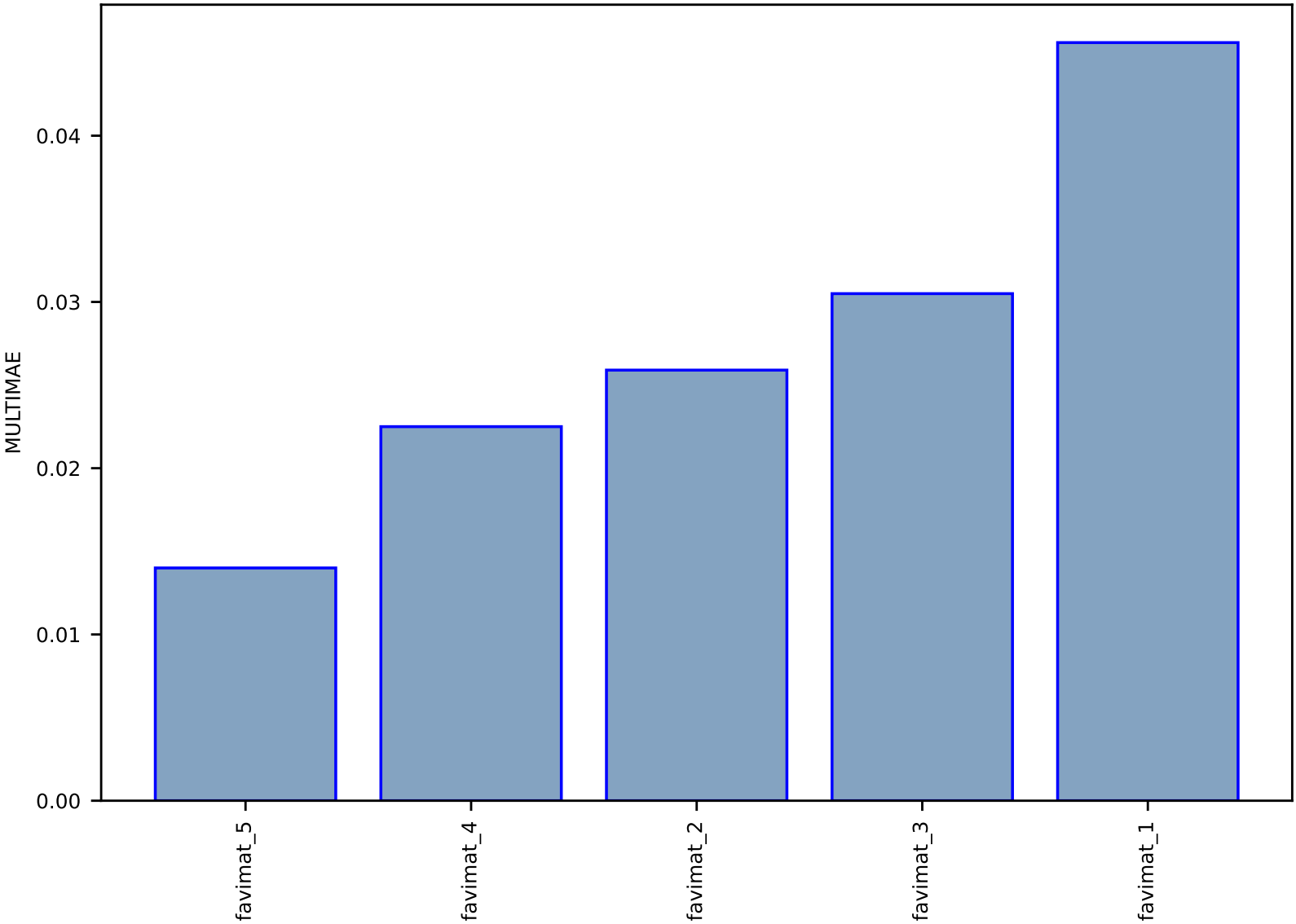


Fig S25:AI-SinglePropertyPrediction-formula_energy-ssub-test-mae.csv.zip

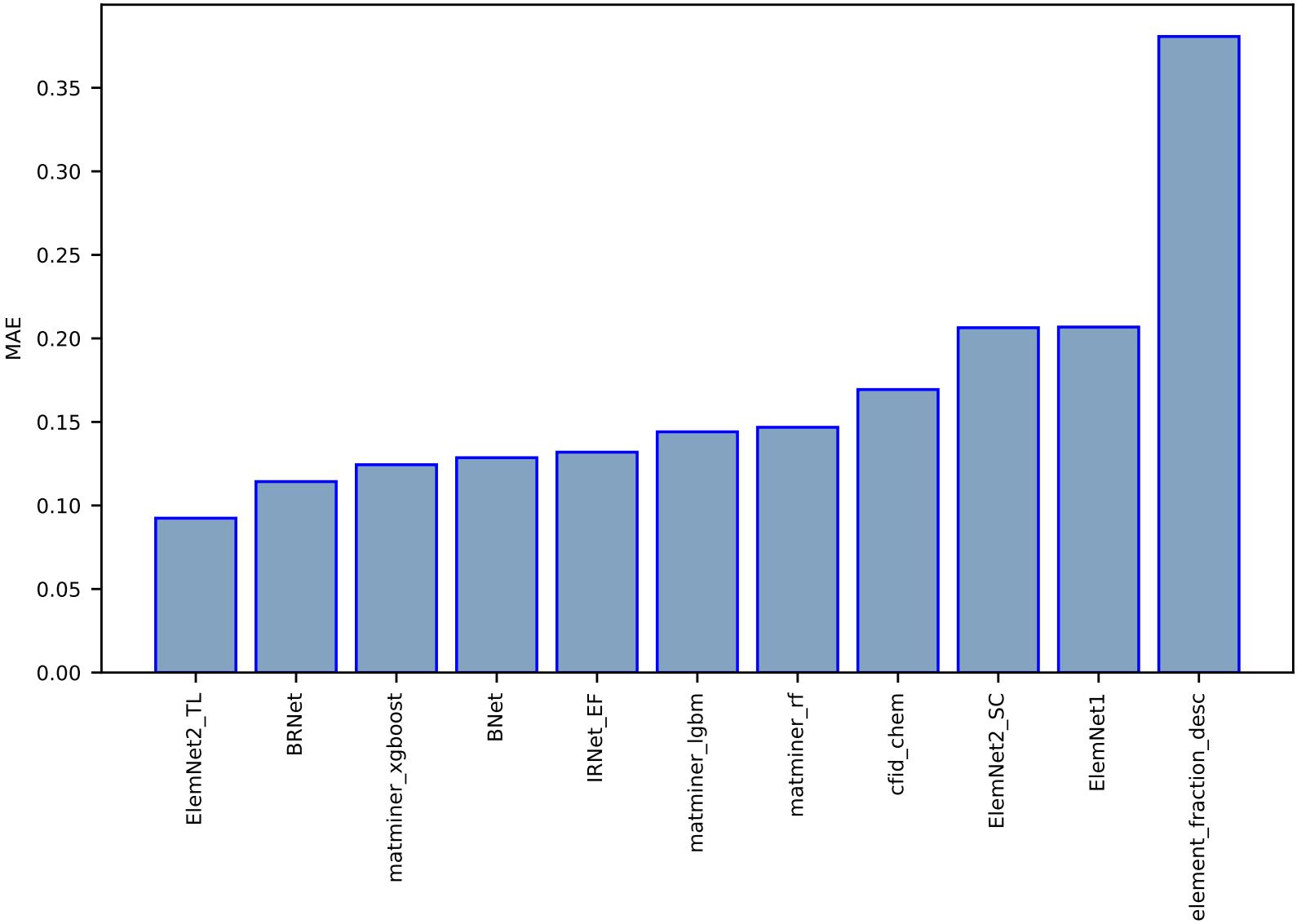


Fig S26:AI-SinglePropertyPrediction-Tc-supercon_chem-test-mae.csv.zip

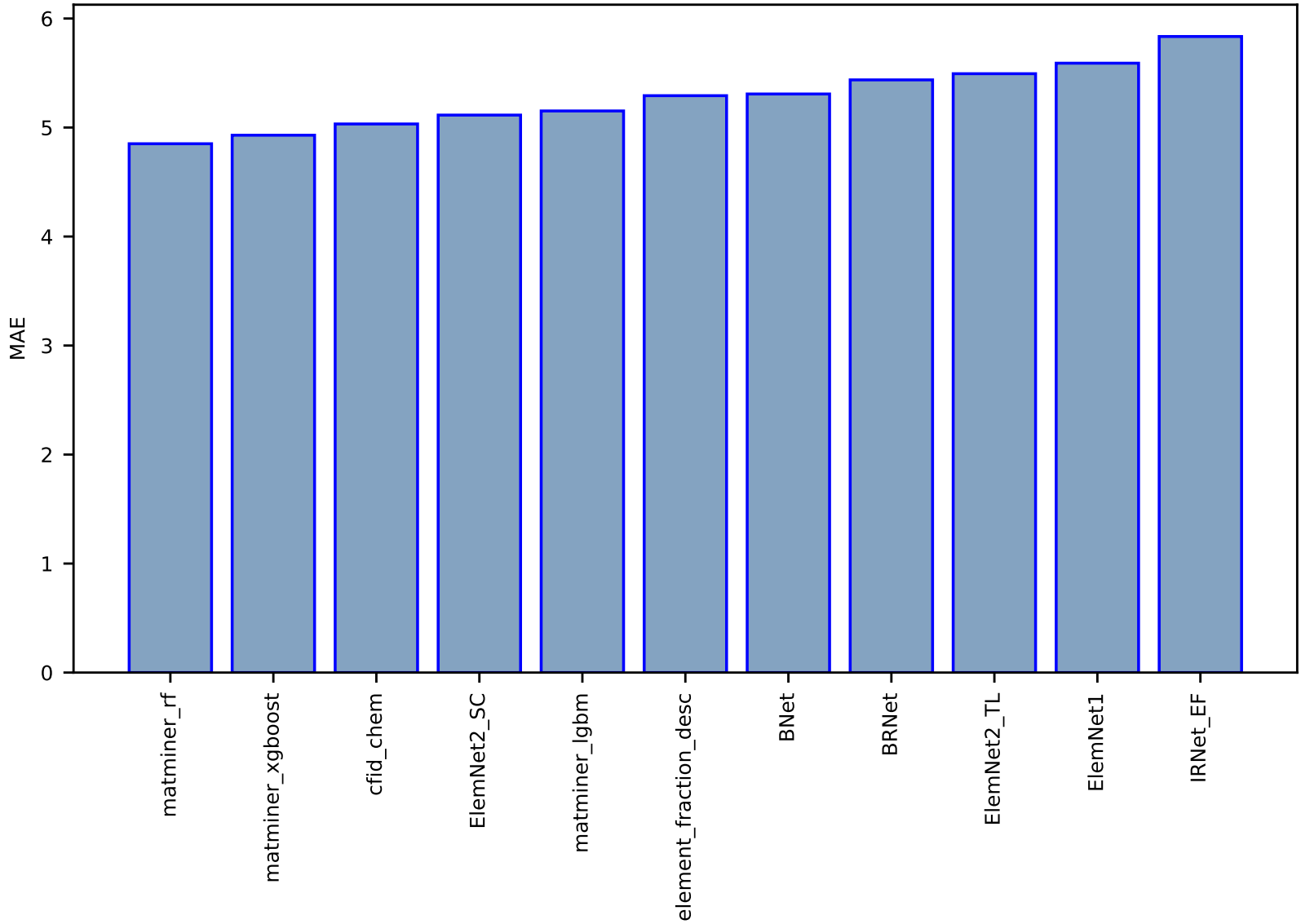


Fig S27:AI-SinglePropertyPrediction-magnetic_moment-mag2d_chem-test-mae.csv.zip

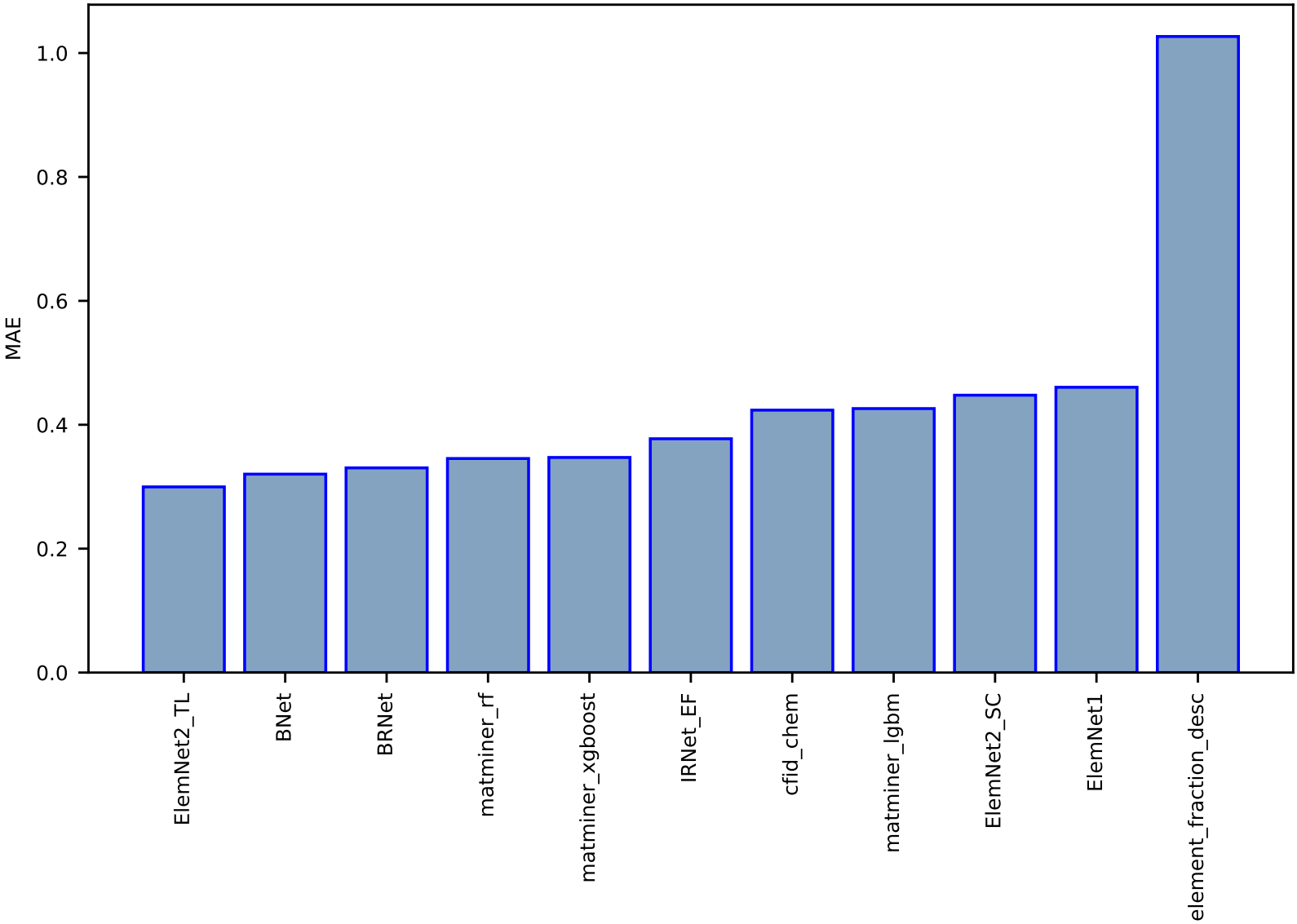


Fig S28:AI-SinglePropertyPrediction-dfpt_piezo_max_dij-dft_3d-test-mae.csv.zip

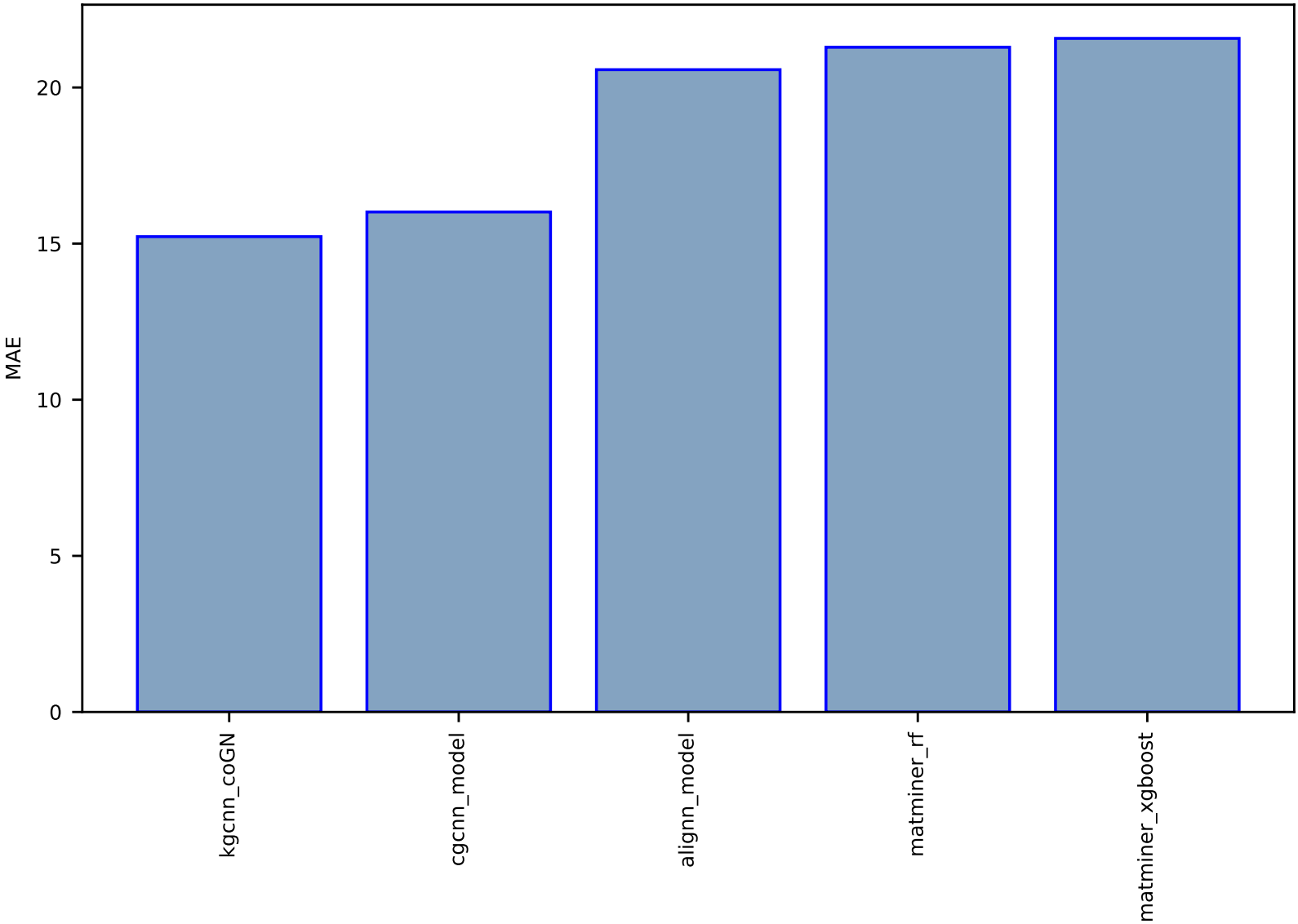


Fig S29:AI-SinglePropertyPrediction-mepsz-dft_3d-test-mae.csv.zip

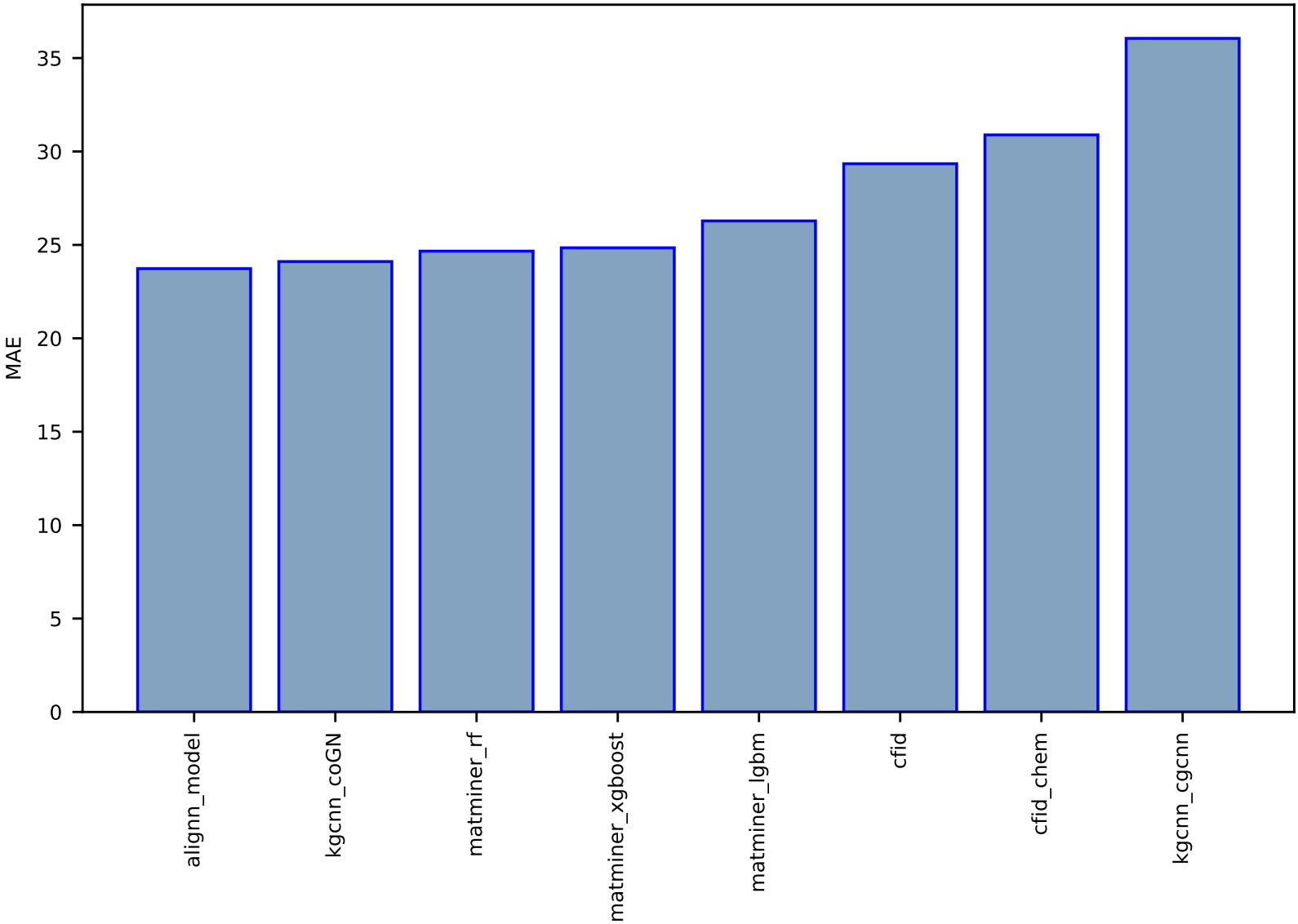


Fig S30:AI-SinglePropertyPrediction-epsx-dft_3d-test-mae.csv.zip

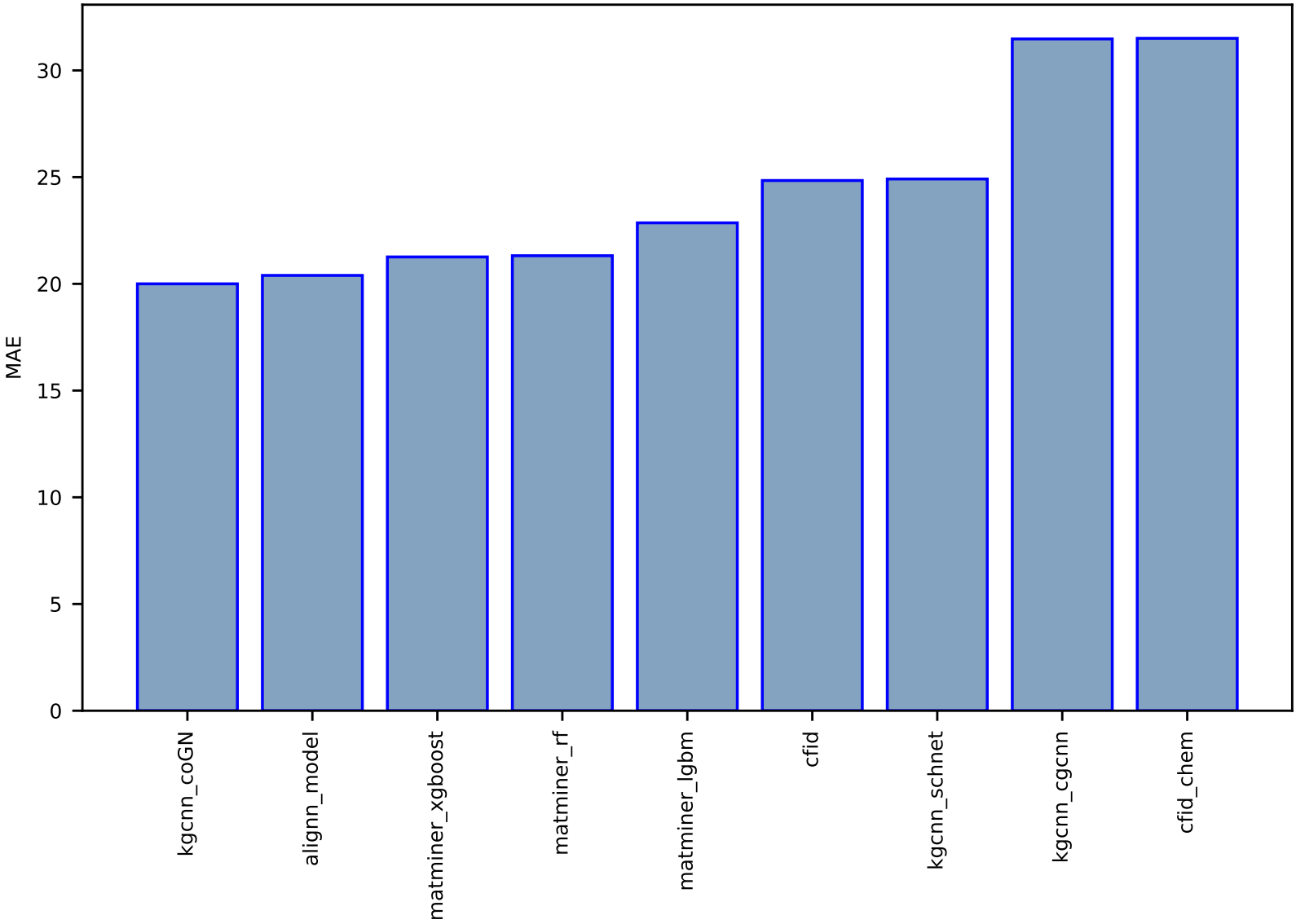


Fig S31:AI-SinglePropertyPrediction-exfoliation_energy-dft_3d-test-mae.csv.zip

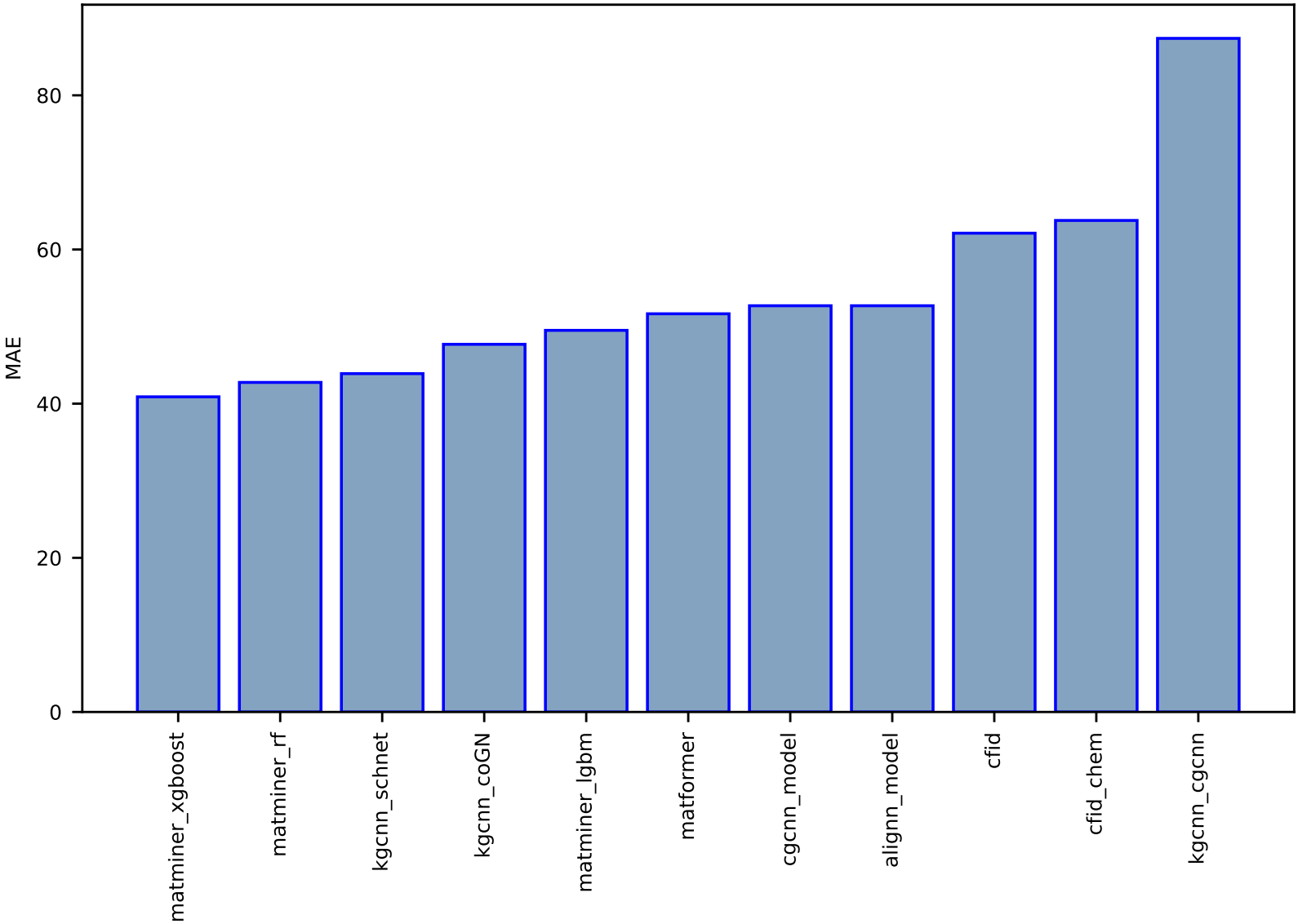


Fig S32:AI-SinglePropertyPrediction-n_powerfact-dft_3d-test-mae.csv.zip

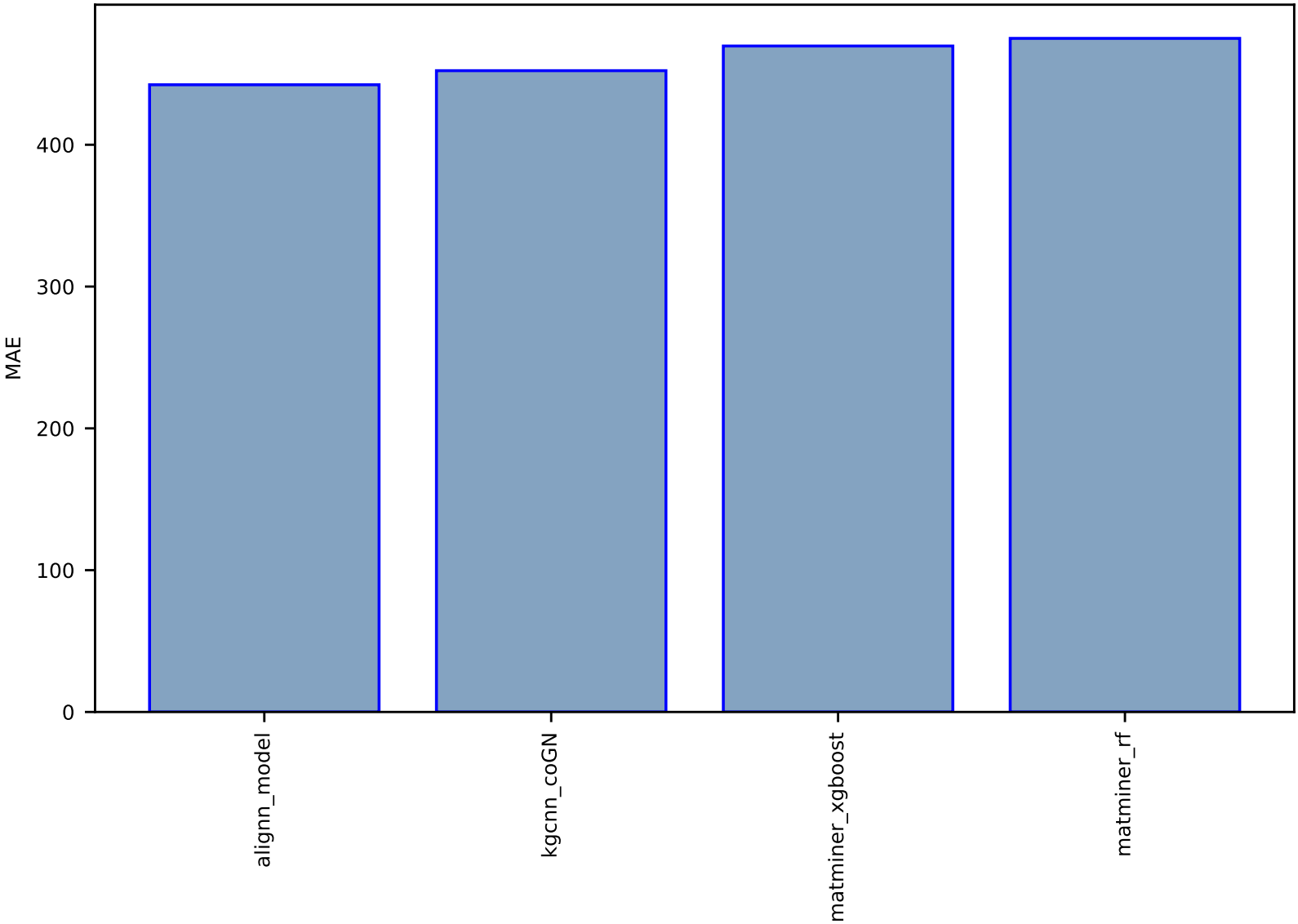


Fig S33:AI-SinglePropertyPrediction-shear_modulus_gv-dft_3d-test-mae.csv.zip

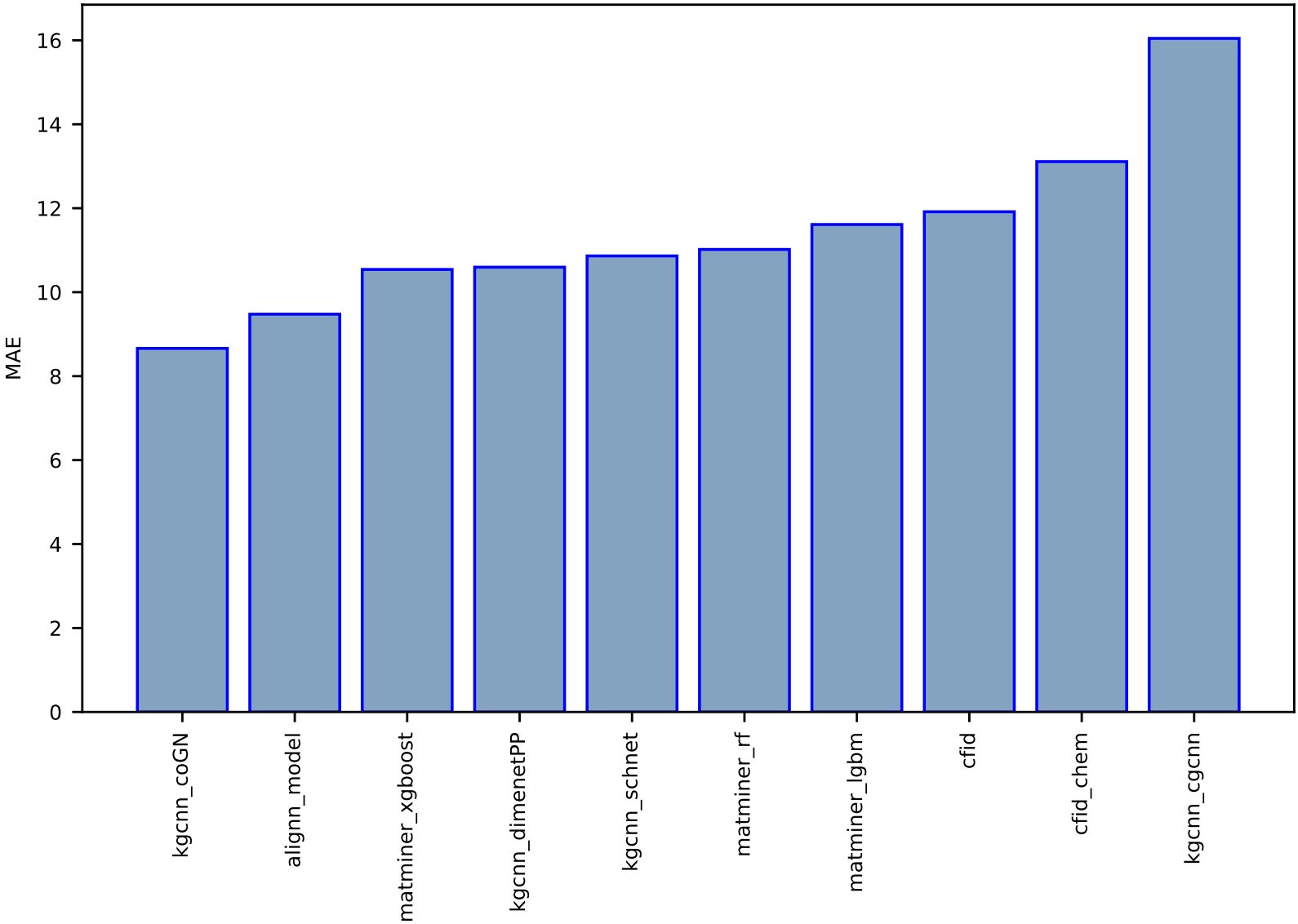


Fig S34:AI-SinglePropertyPrediction-kpoint_length_unit-dft_3d-test-mae.csv.zip

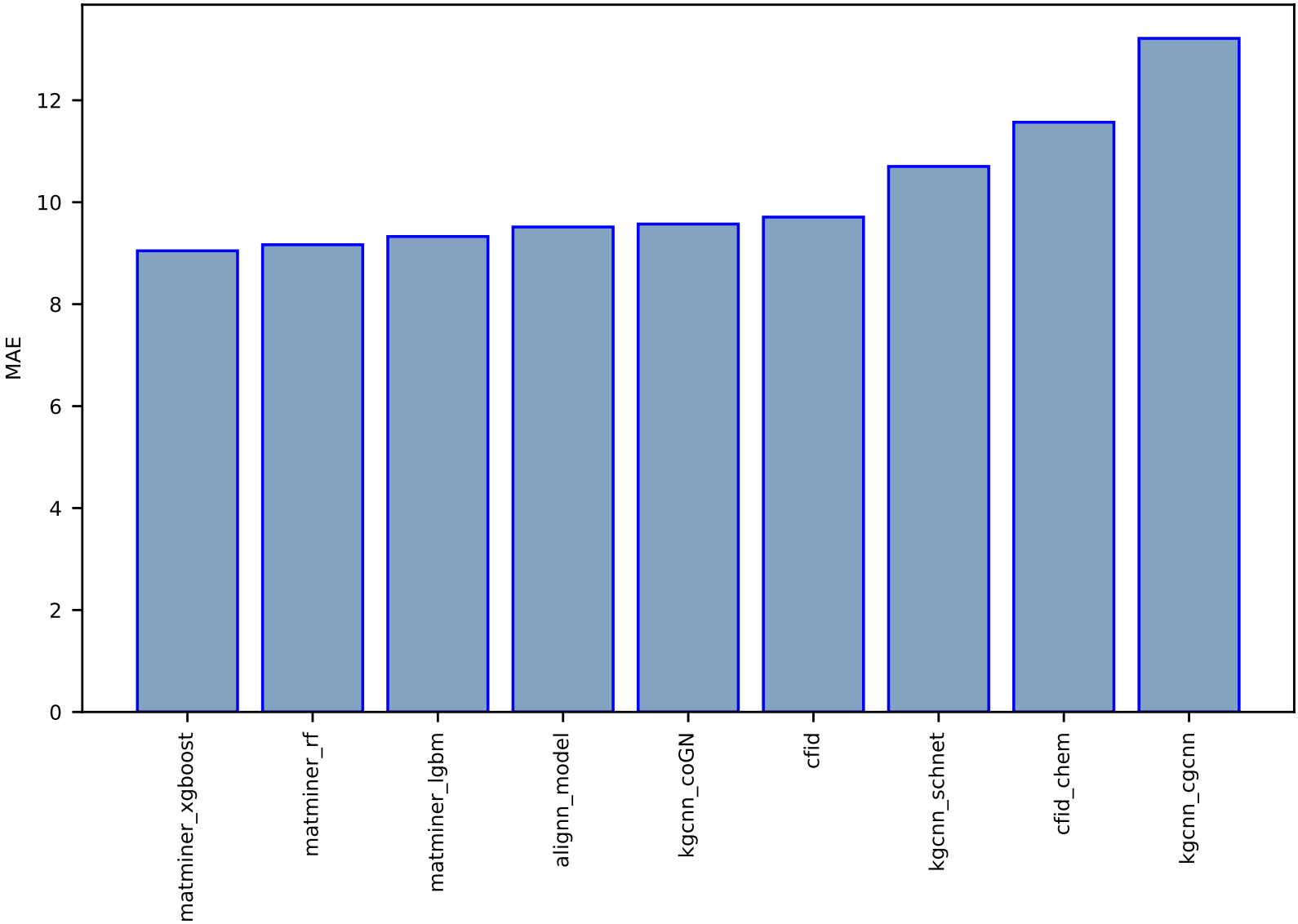


Fig S35:AI-SinglePropertyPrediction-optb88vdw_bandgap-dft_3d-test-mae.csv.zip

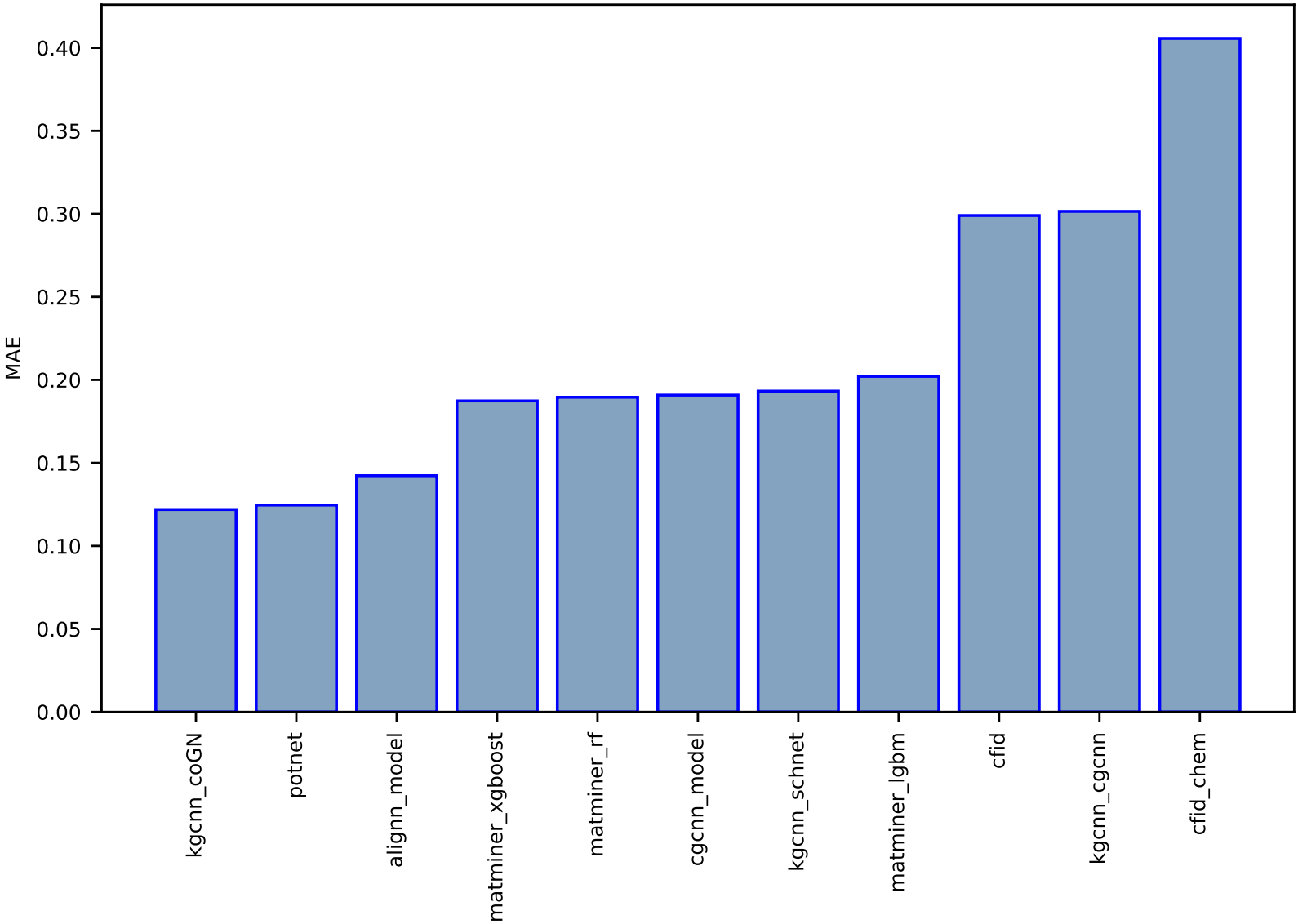


Fig S36:AI-SinglePropertyPrediction-epsz-dft_3d-test-mae.csv.zip

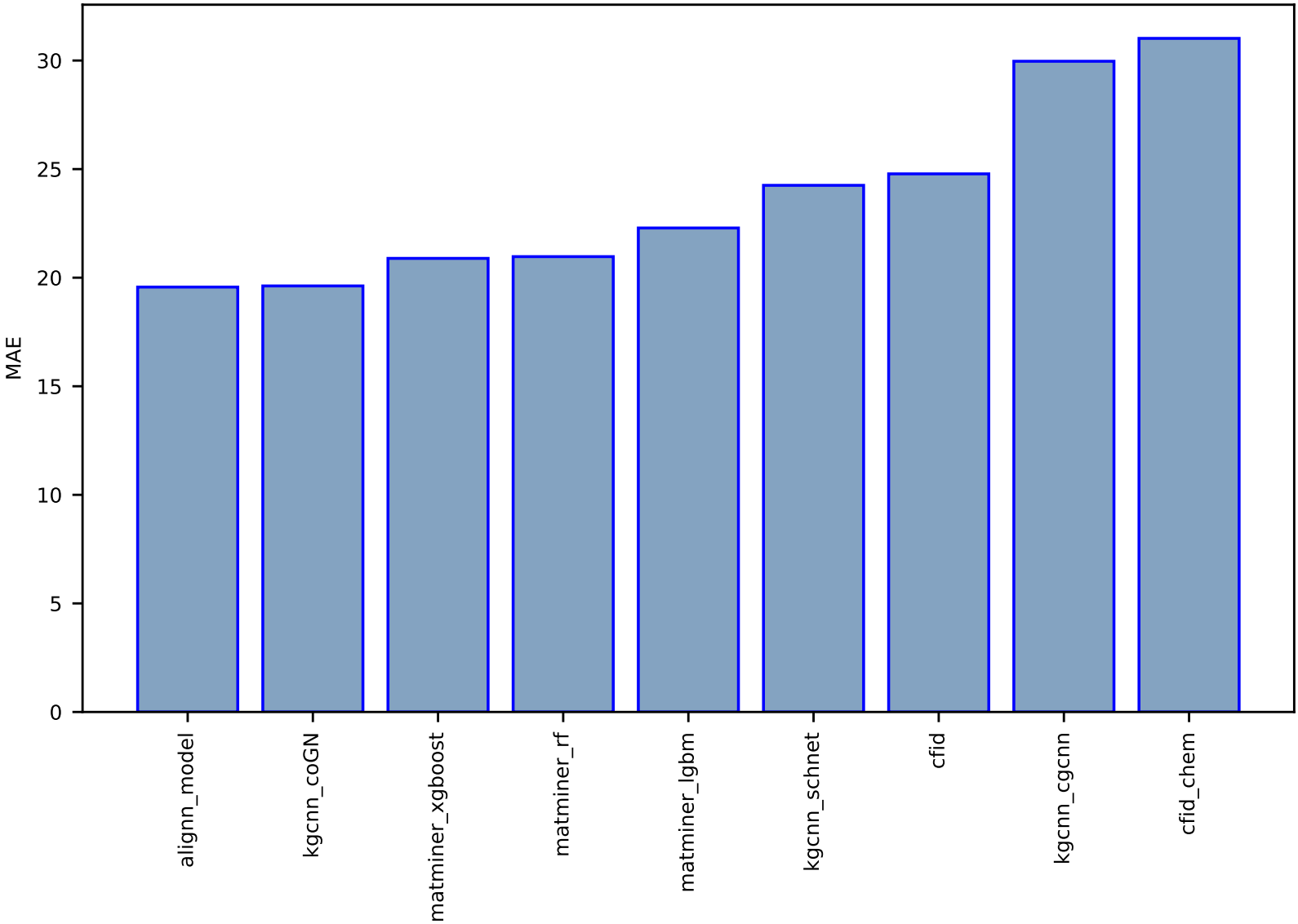


Fig S37:AI-SinglePropertyPrediction-spillage-dft_3d-test-mae.csv.zip

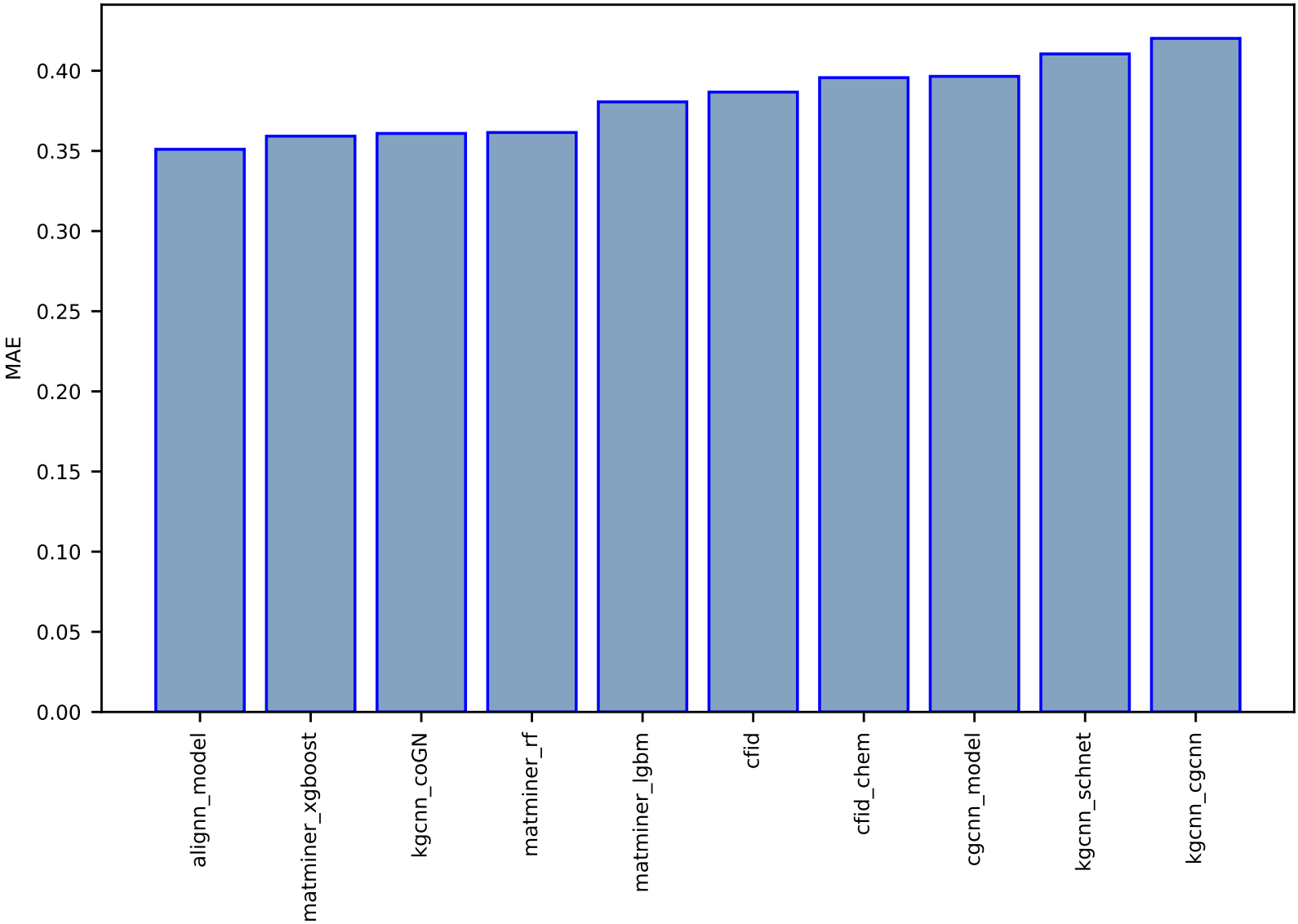


Fig S38:AI-SinglePropertyPrediction-optb88vdw_total_energy-dft_3d-test-mae.csv.zip

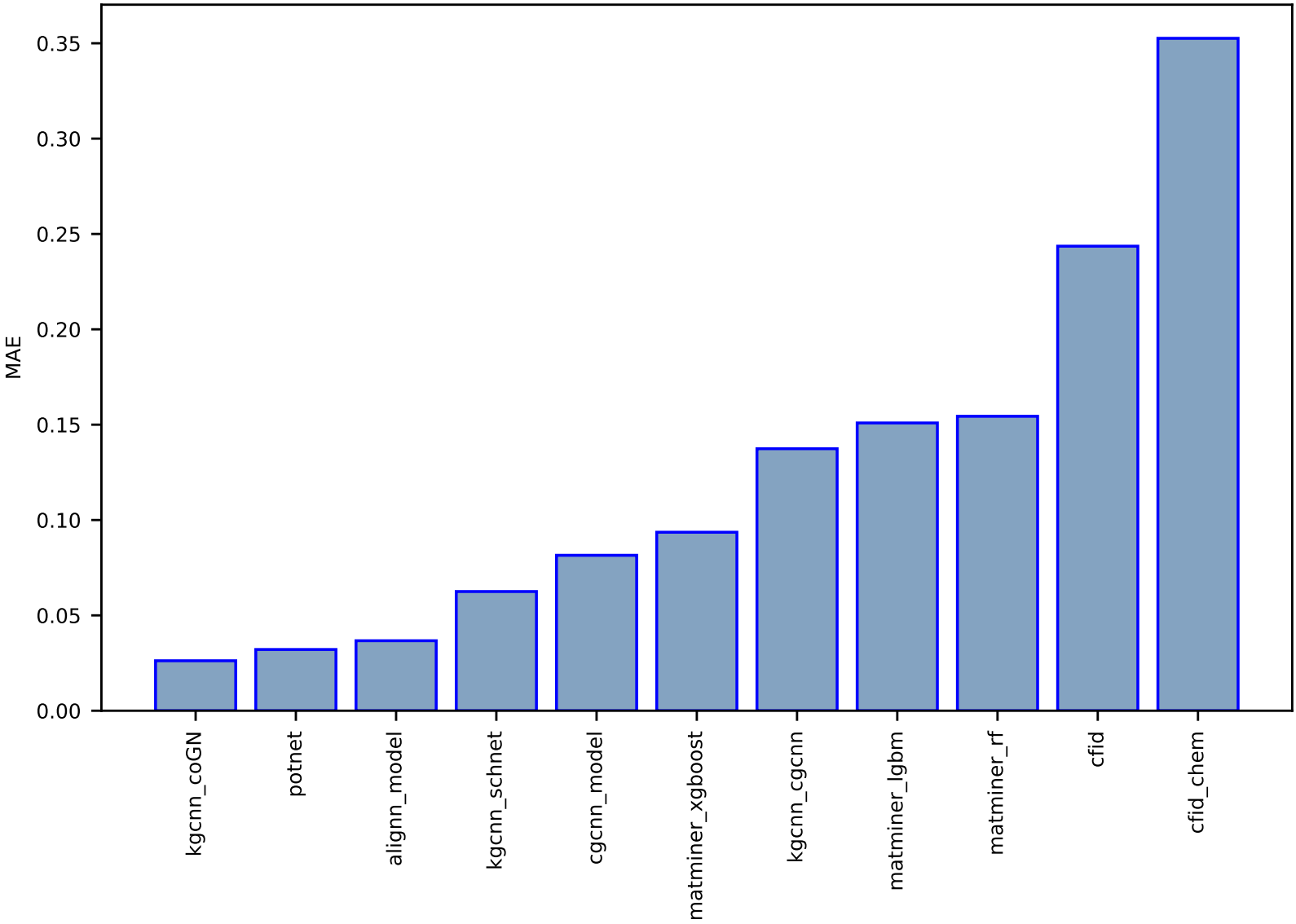


Fig S39:AI-SinglePropertyPrediction-mepsy-dft_3d-test-mae.csv.zip

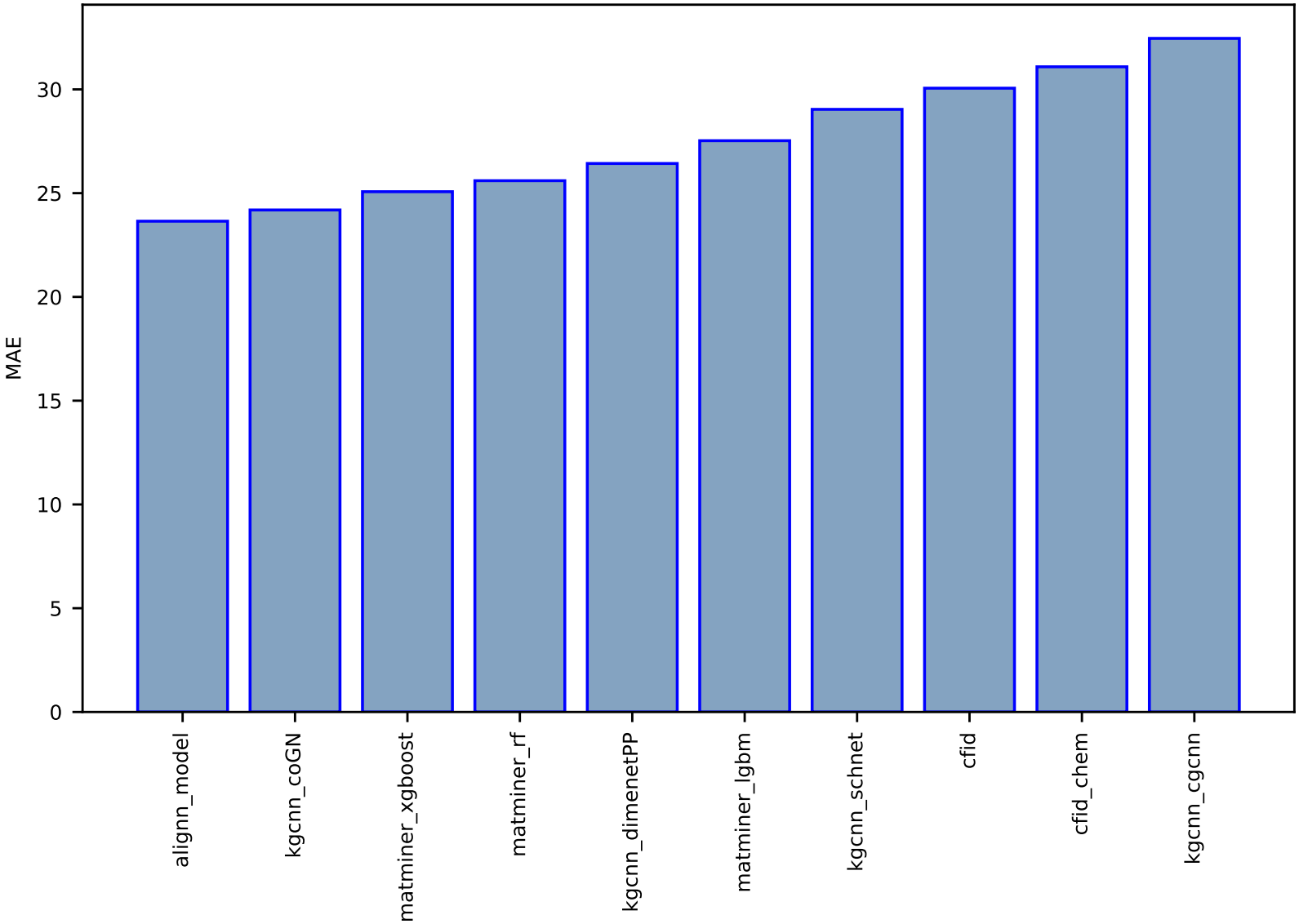


Fig S40:AI-SinglePropertyPrediction-mepsx-dft_3d-test-mae.csv.zip

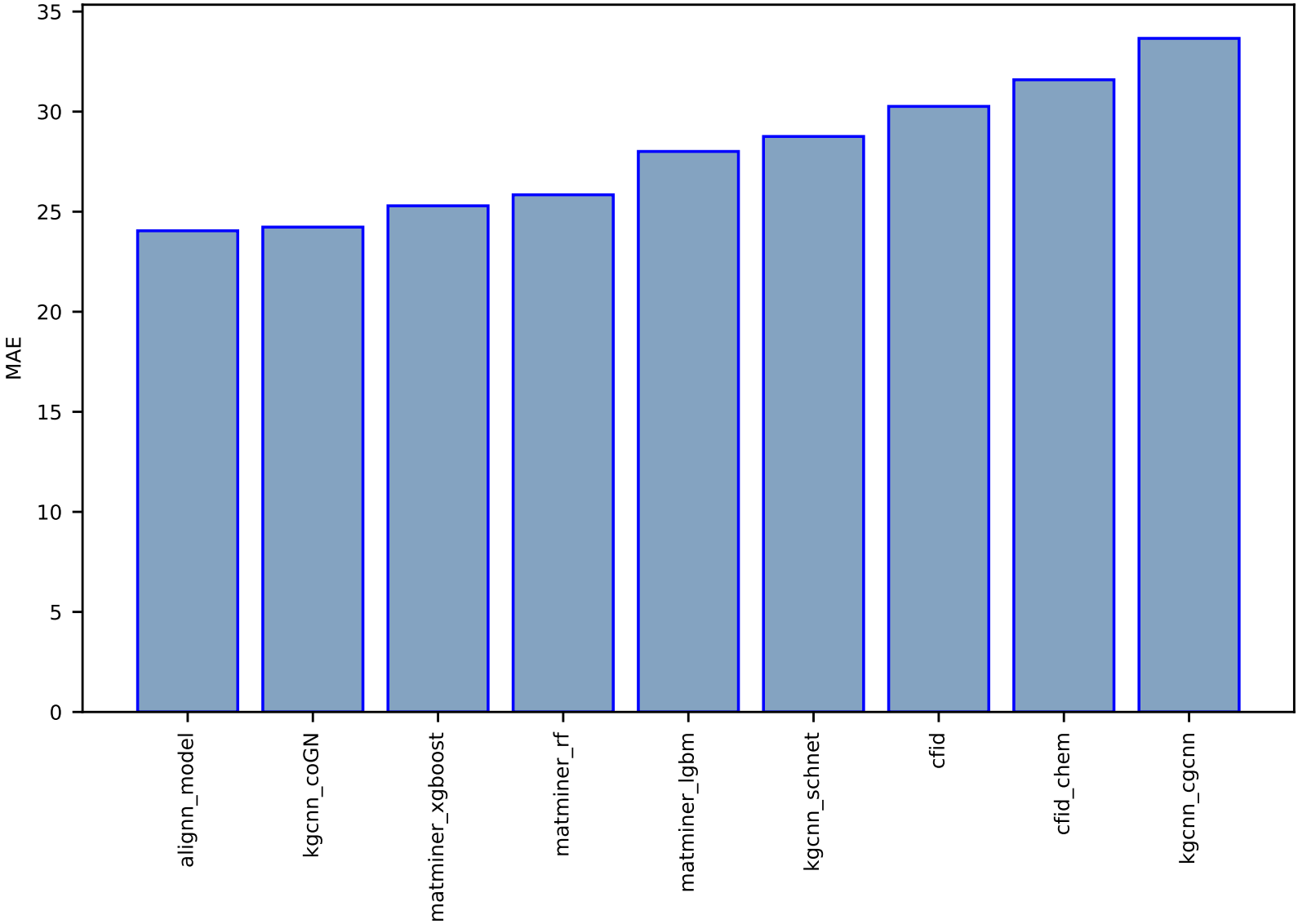


Fig S41:AI-SinglePropertyPrediction-ph_heat_capacity-dft_3d-test-mae.csv.zip

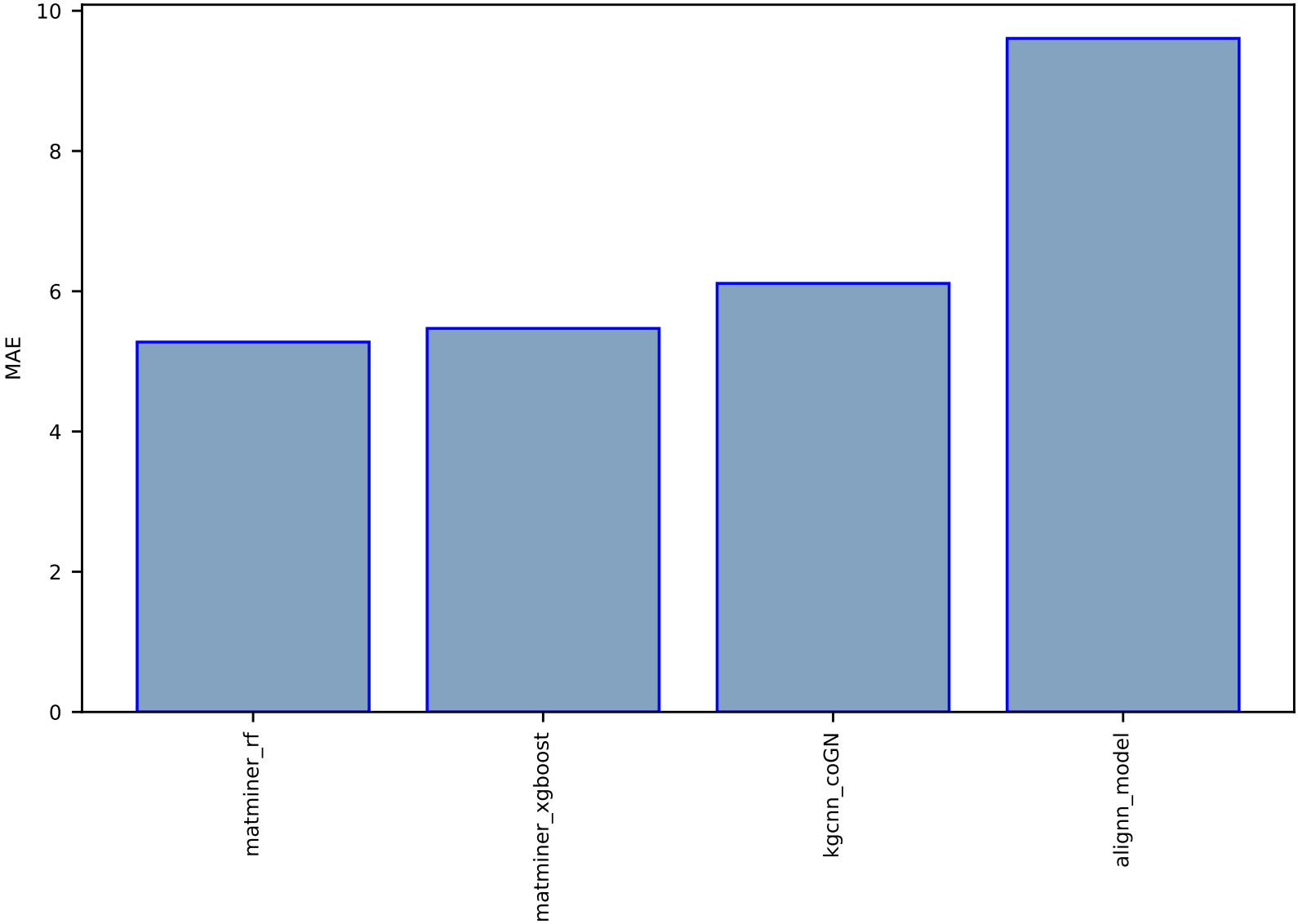


Fig S42:AI-SinglePropertyPrediction-indir_gap-qe_tb-test-mae.csv.zip

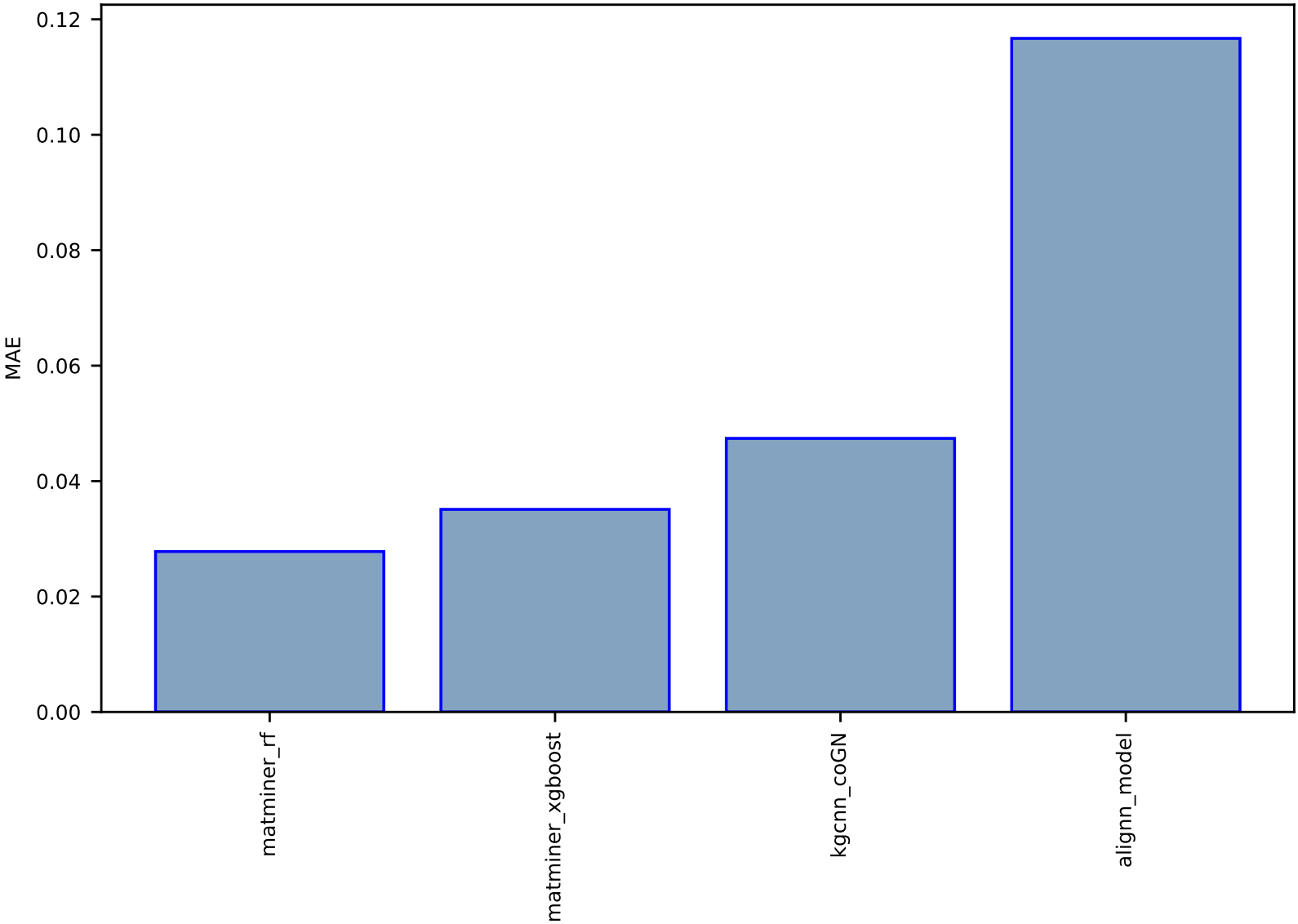


Fig S43:AI-SinglePropertyPrediction-formation_energy_peratom-dft_3d-test-mae.csv.zip

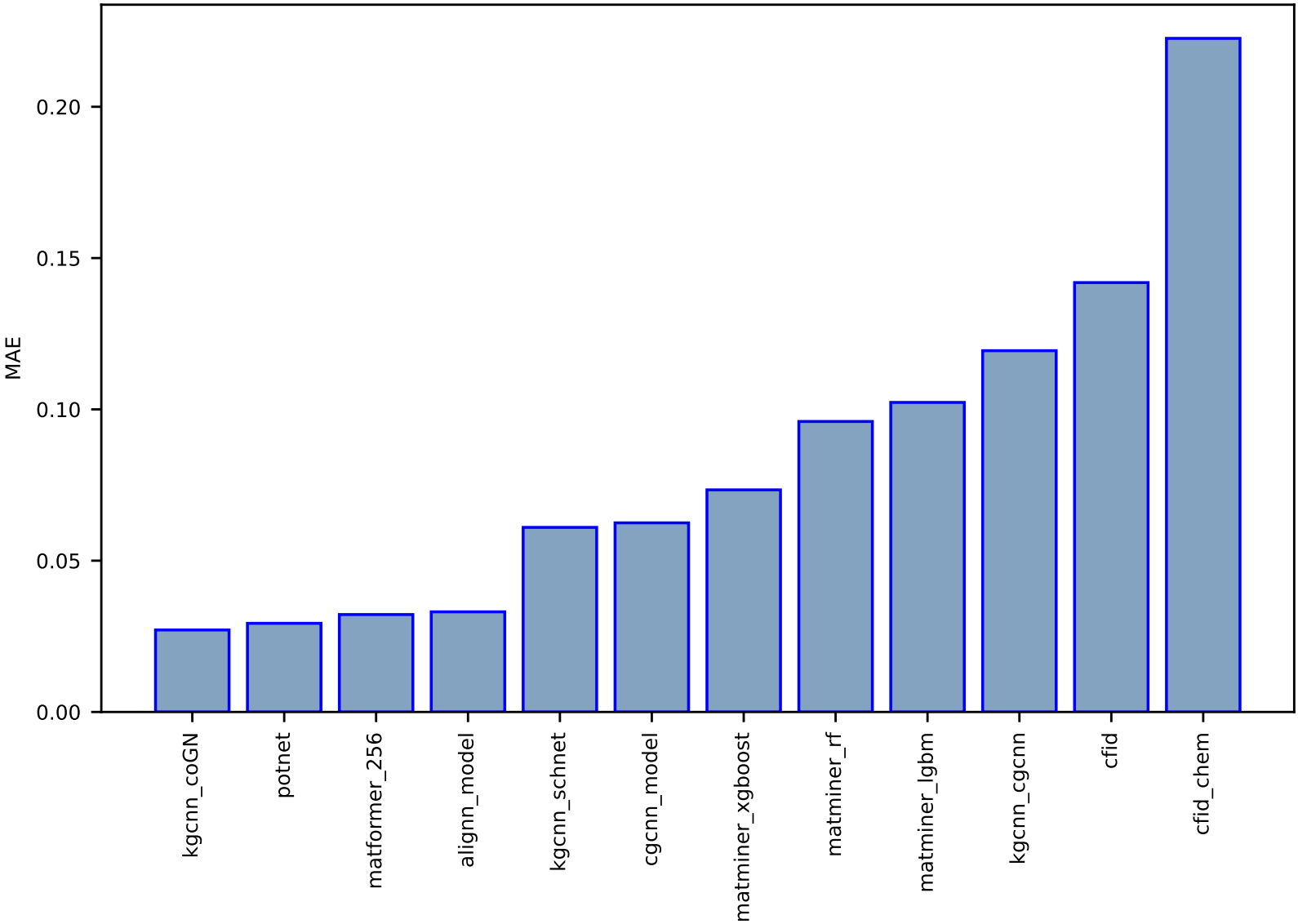


Fig S44:ES-SinglePropertyPrediction-bandgap_JVASP_39_AlN-dft_3d-test-mae.csv.zip

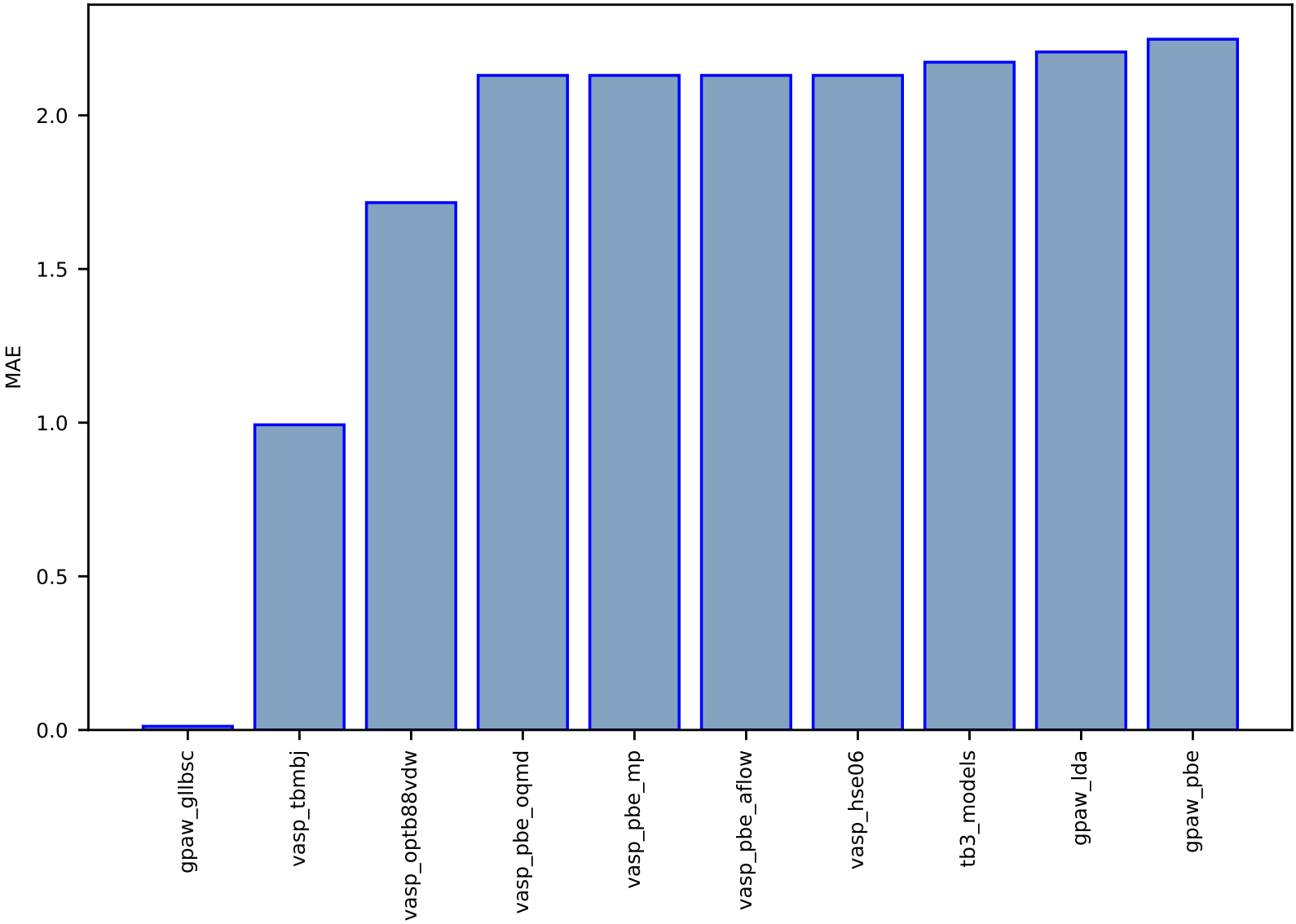


Fig S45:ES-SinglePropertyPrediction-bandgap_JVASP_1216_Cu2O-dft_3d-test-mae.csv.zip

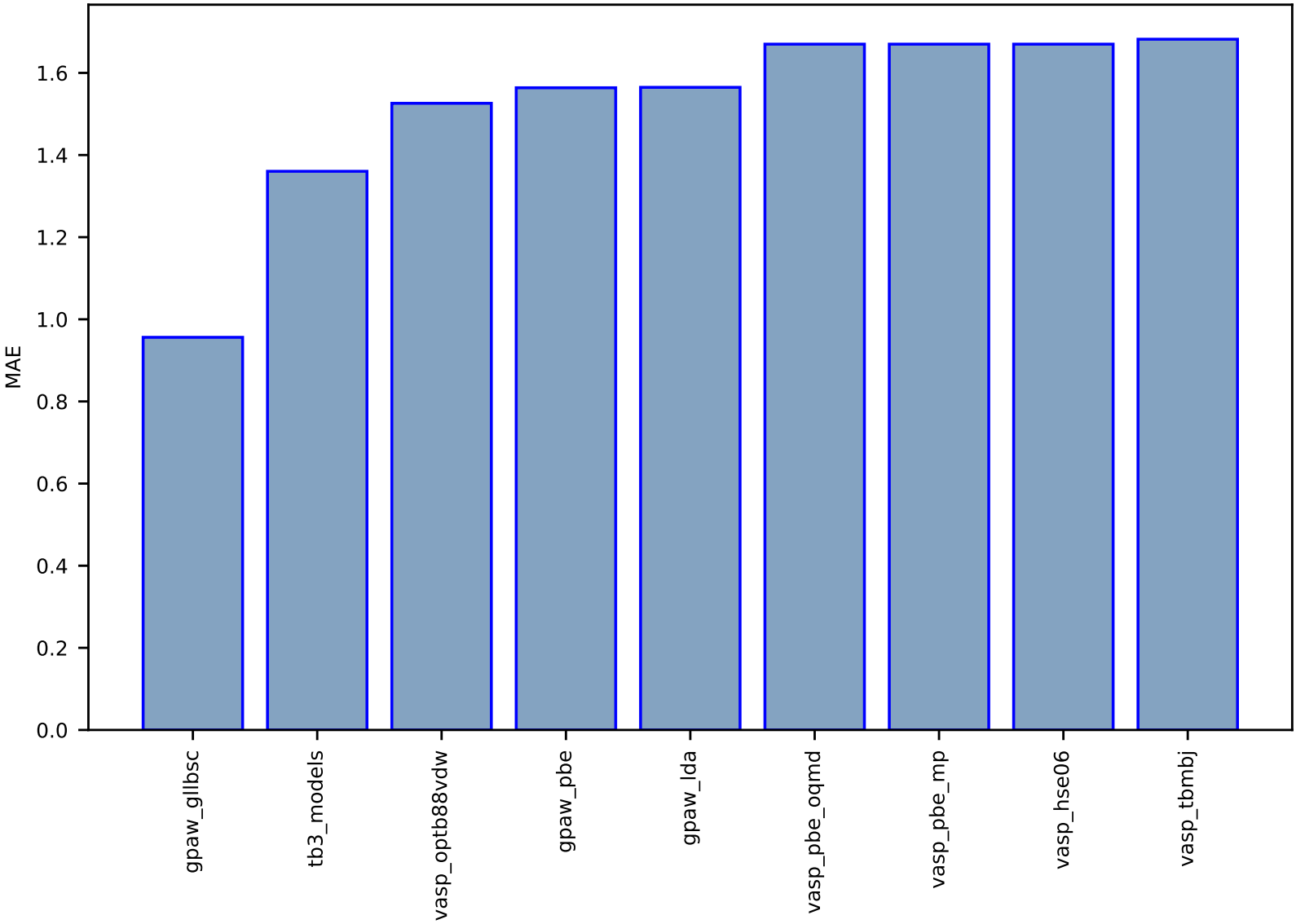


Fig S46:ES-SinglePropertyPrediction-bandgap_JVASP_1393_GaP-dft_3d-test-mae.csv.zip

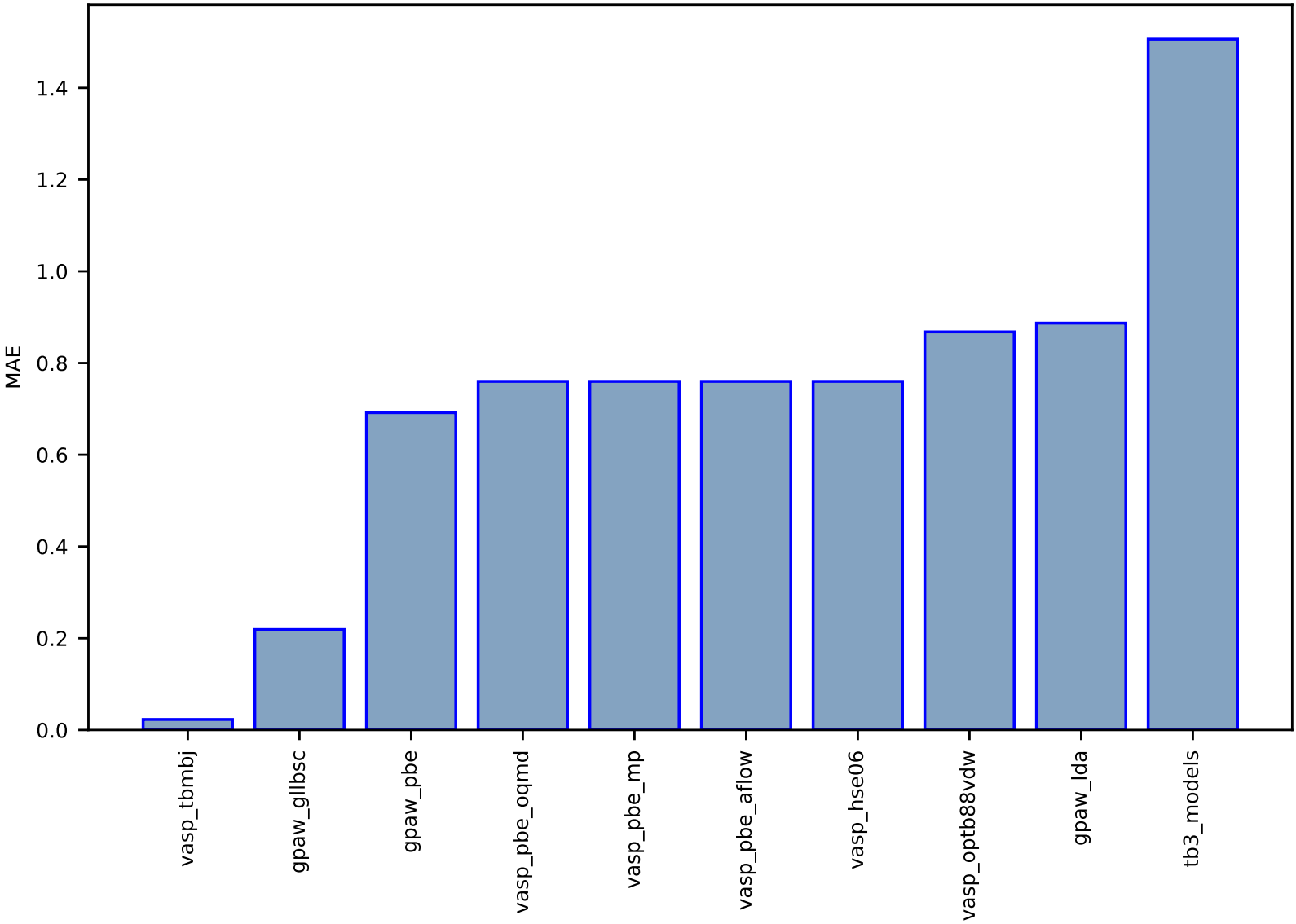


Fig S47:ES-SinglePropertyPrediction-bandgap_JVASP_1198_ZnTe-dft_3d-test-mae.csv.zip

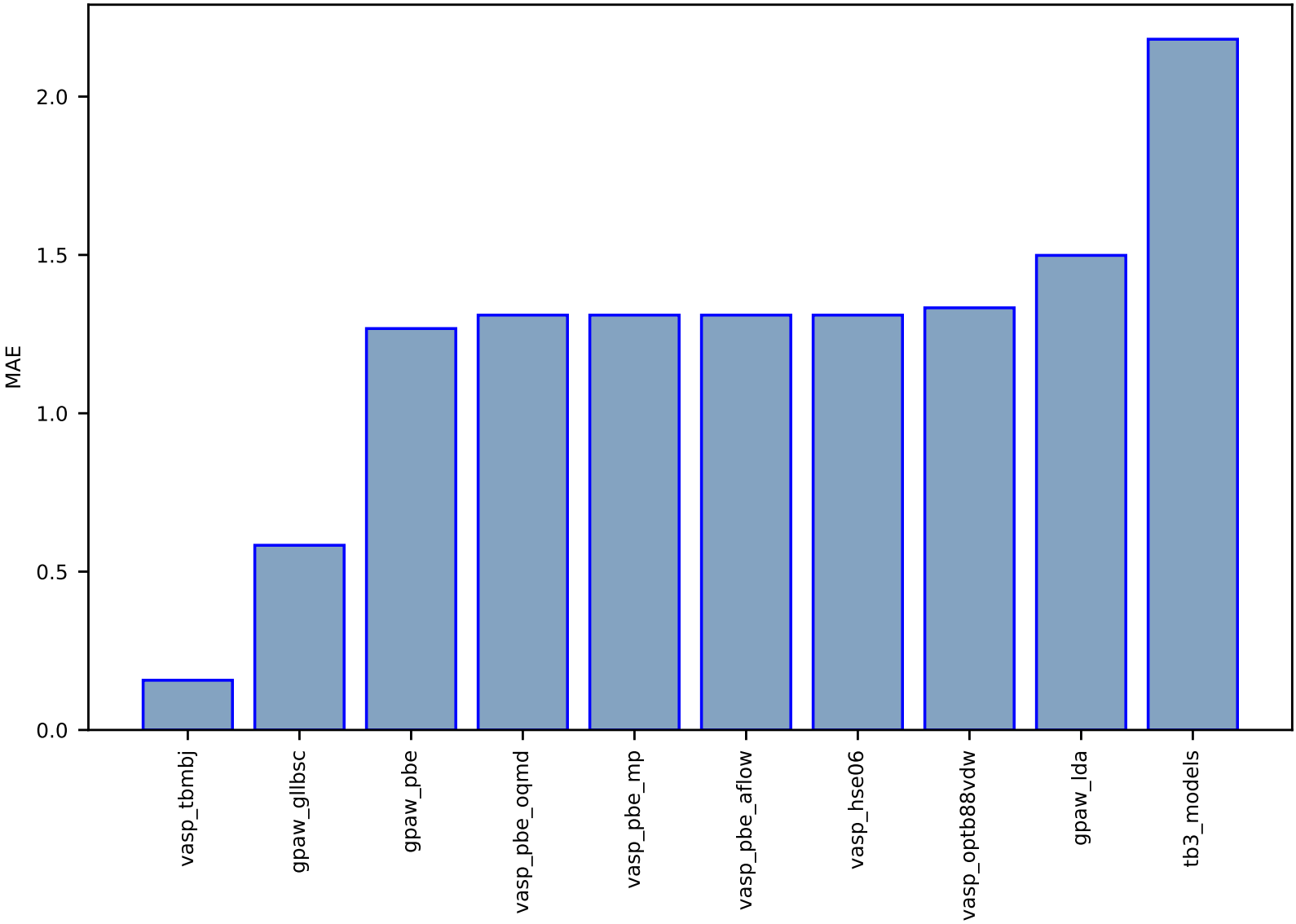


Fig S48:ES-SinglePropertyPrediction-slme_JVASP_8554_InCuSe2-dft_3d-test-mae.csv.zip

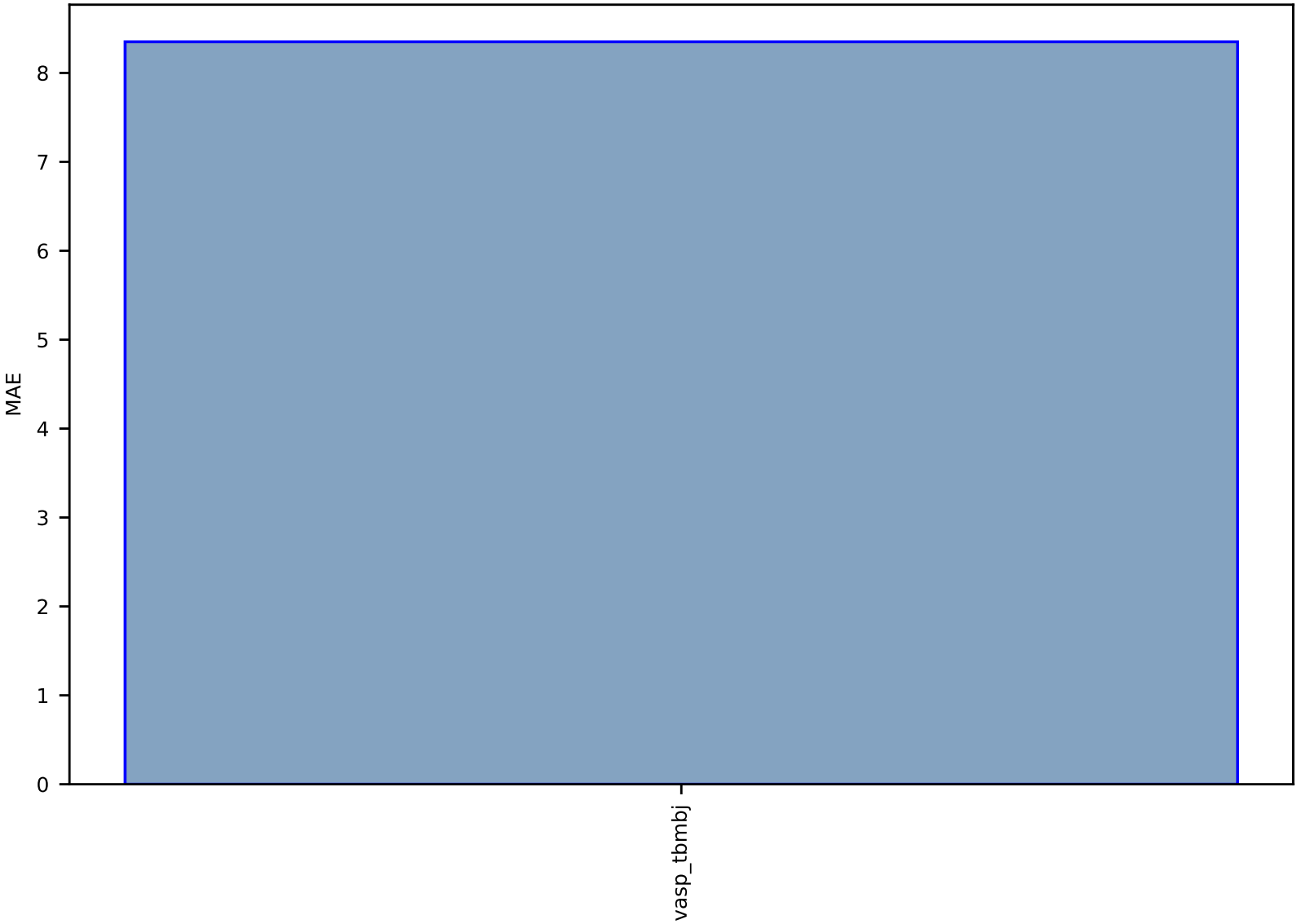


Fig S49:ES-SinglePropertyPrediction-bandgap_JVASP_8169_GaN-dft_3d-test-mae.csv.zip

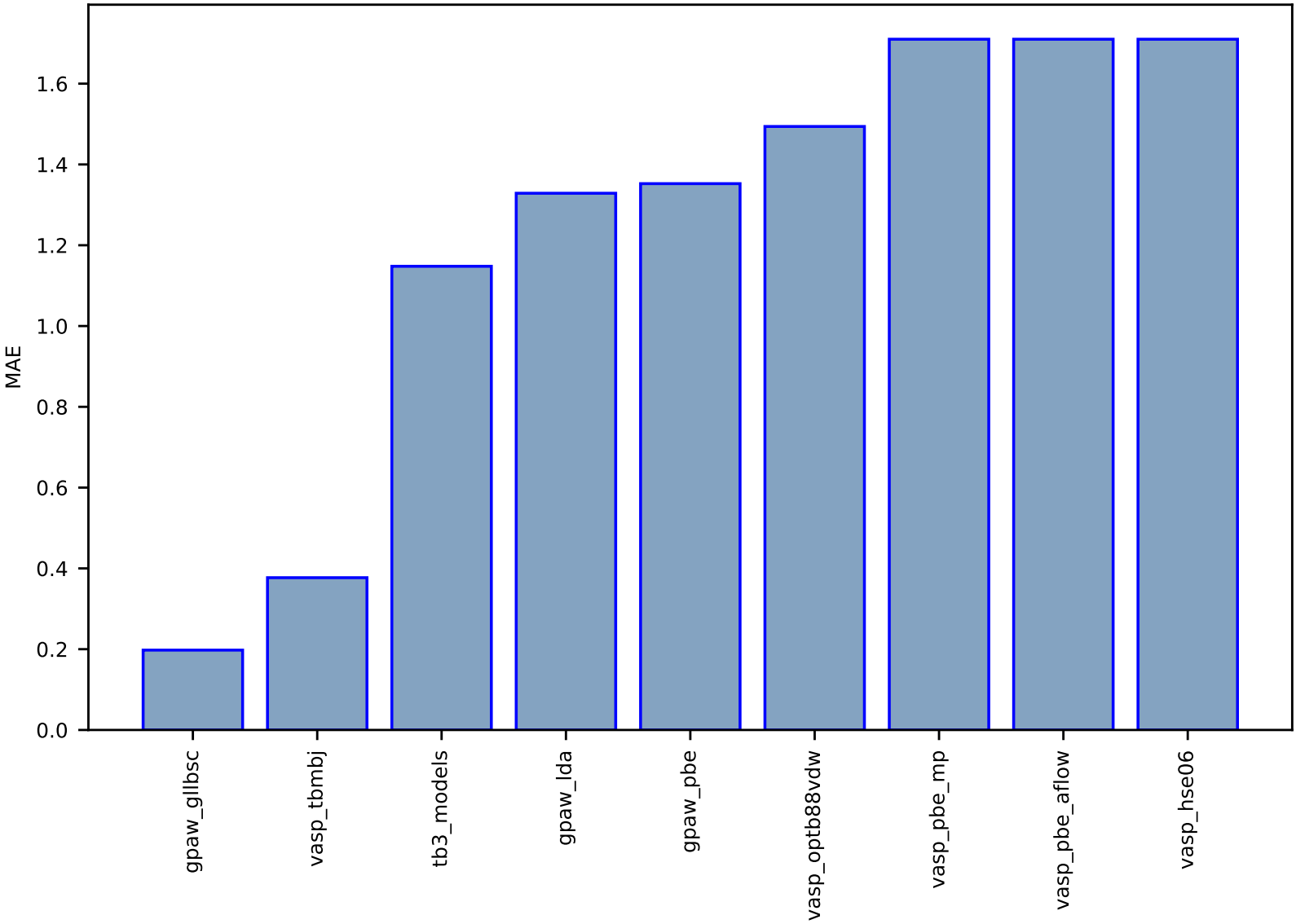


Fig S50:ES-SinglePropertyPrediction-bandgap_JVASP_1002_Si-dft_3d-test-mae.csv.zip

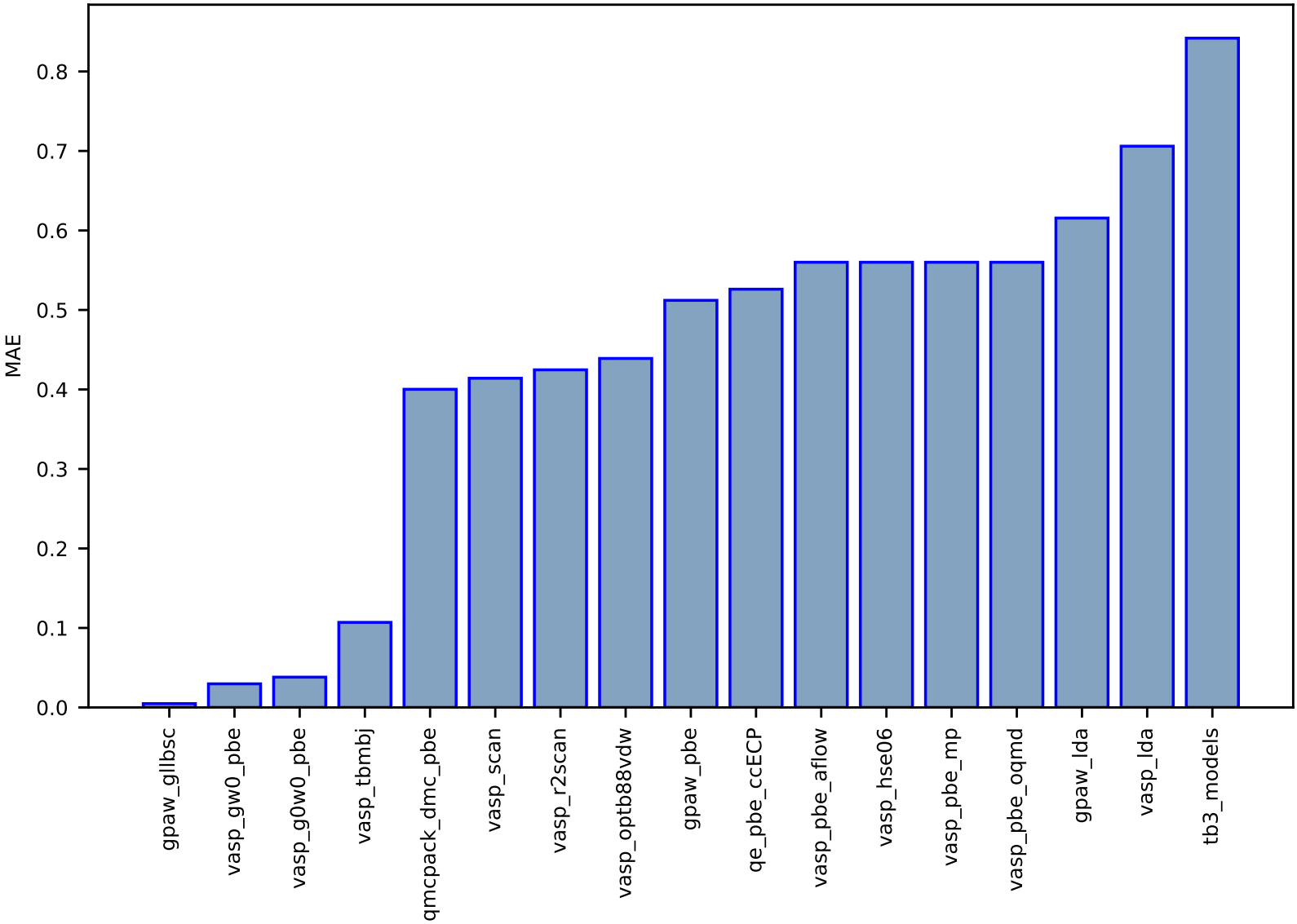


Fig S51:ES-SinglePropertyPrediction-bandgap_JVASP_96_ZnSe-dft_3d-test-mae.csv.zip

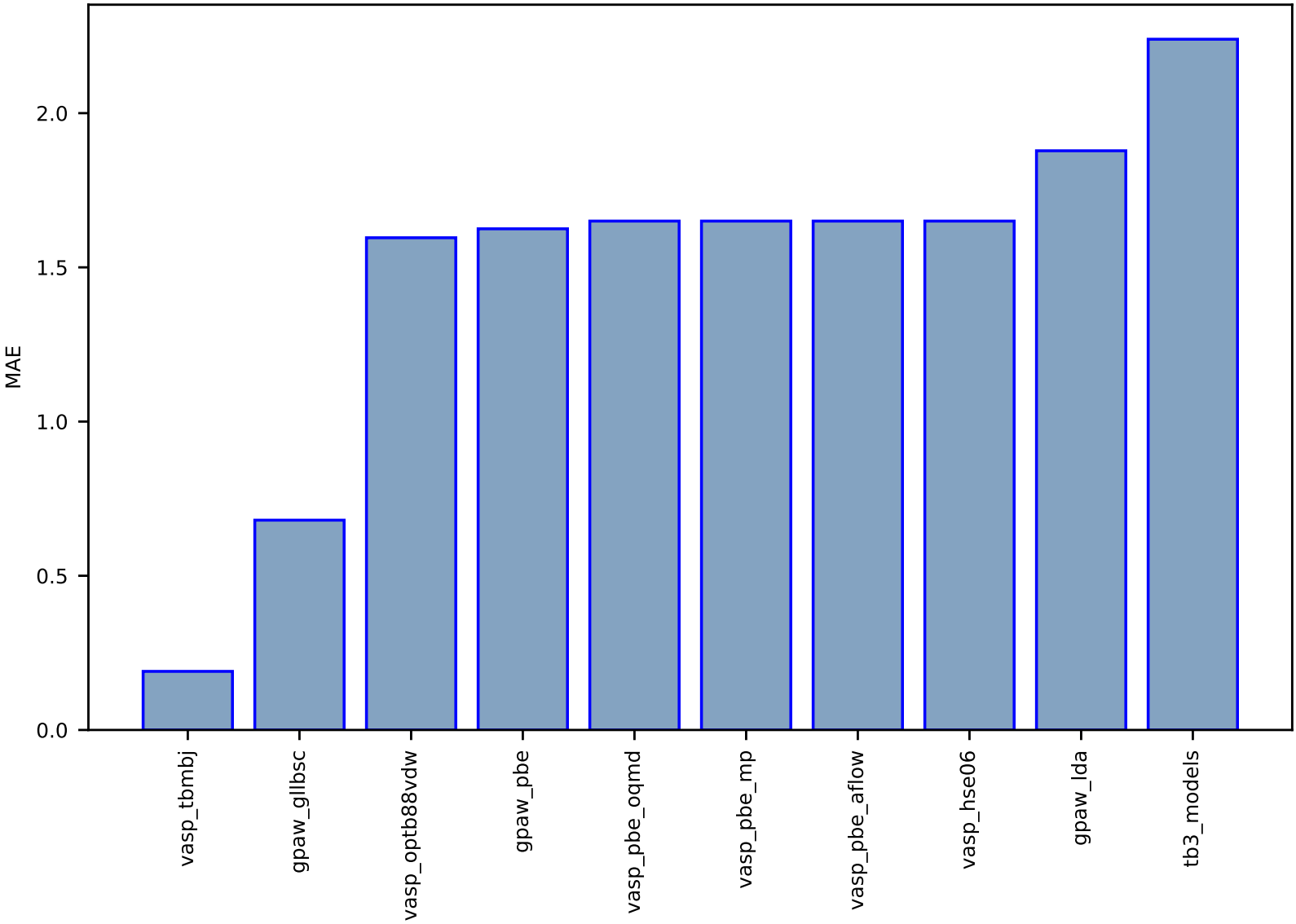


Fig S52:ES-SinglePropertyPrediction-bandgap_JVASP_54_MoS2-dft_3d-test-mae.csv.zip

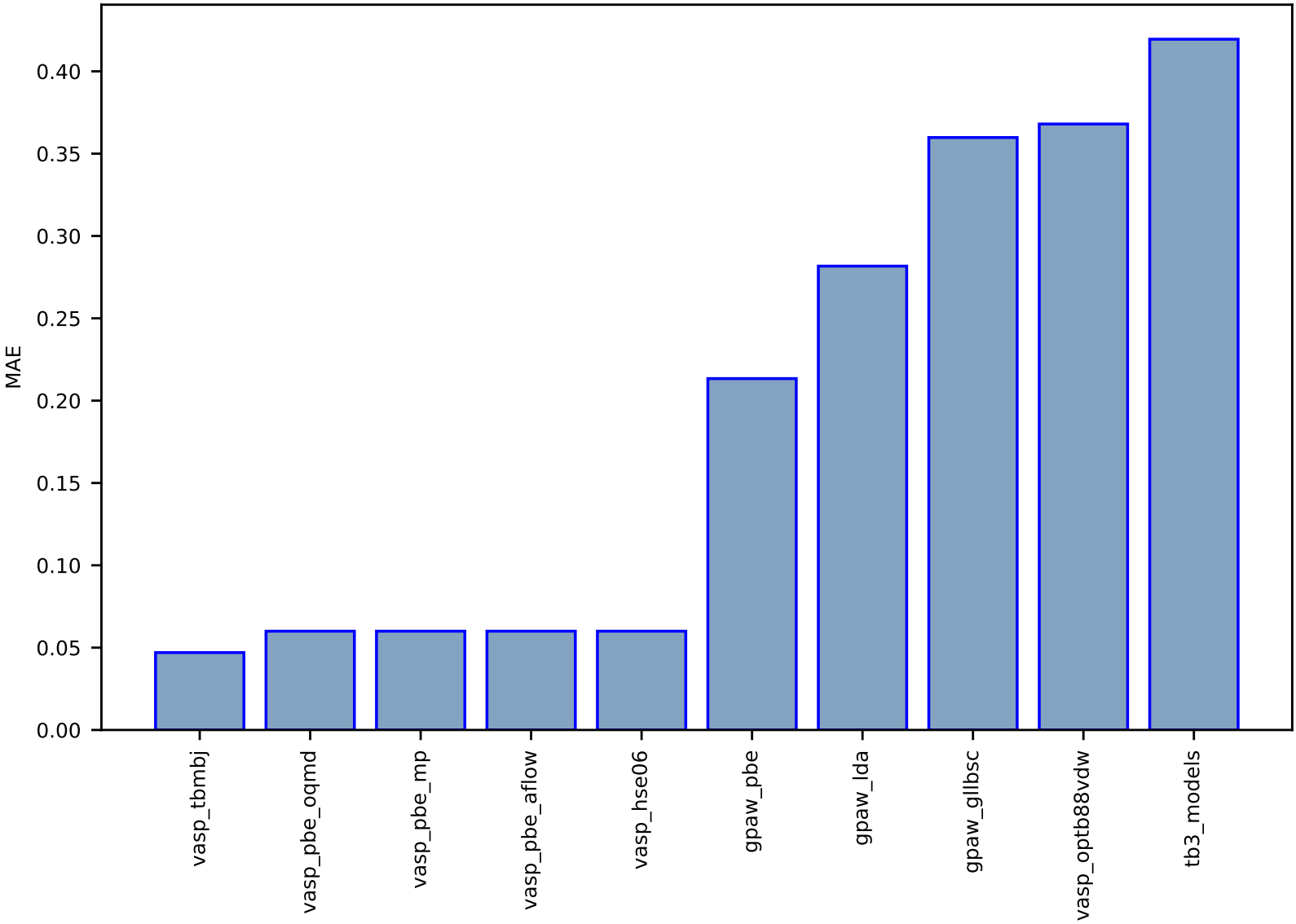


Fig S53:ES-SinglePropertyPrediction-bandgap_JVASP_7844_AlN-dft_3d-test-mae.csv.zip

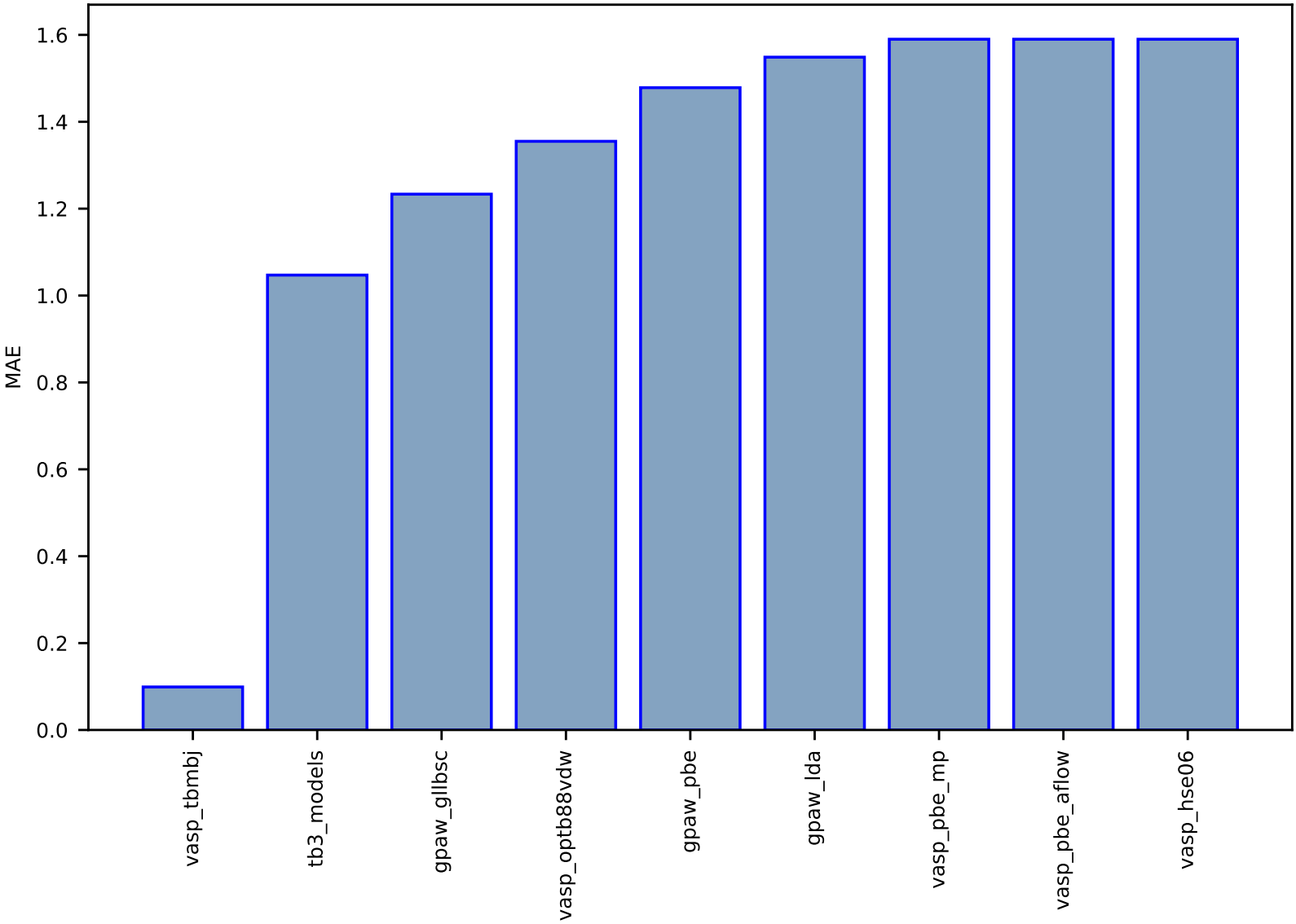


Fig S54:ES-SinglePropertyPrediction-bandgap_JVASP_1300_MgS-dft_3d-test-mae.csv.zip

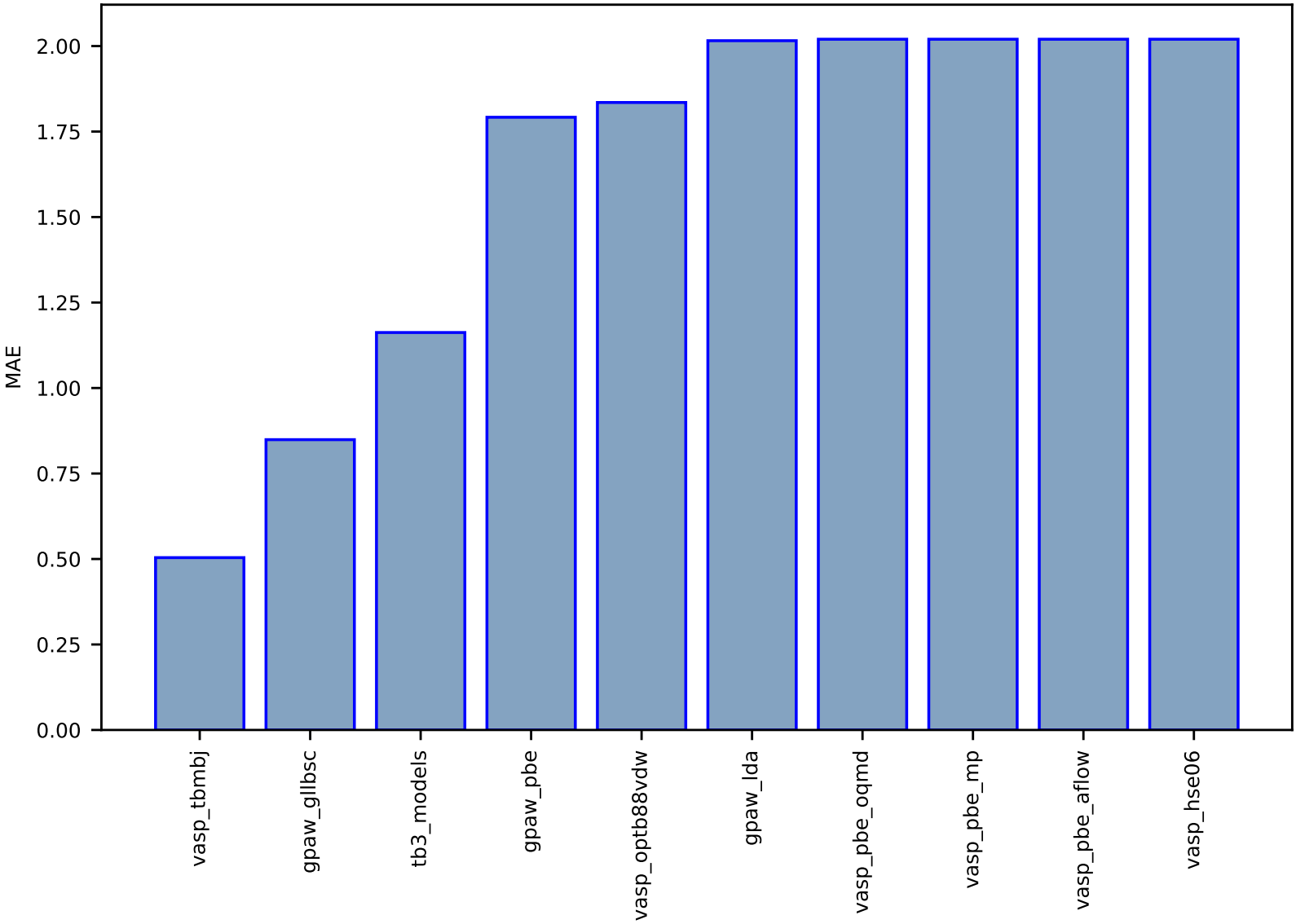


Fig S55:ES-SinglePropertyPrediction-bandgap_JVASP_7762_MgTe-dft_3d-test-mae.csv.zip

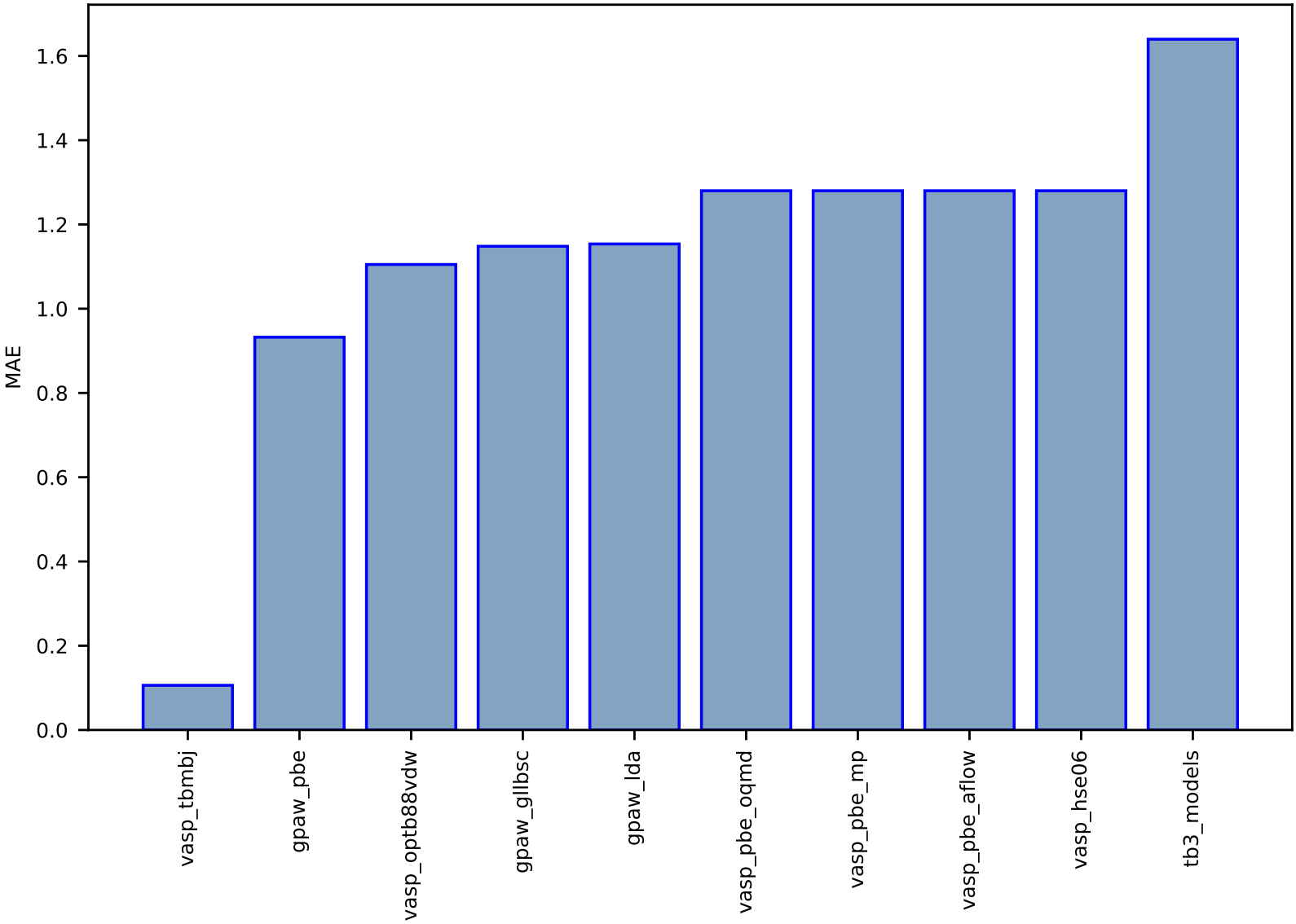


Fig S56:ES-SinglePropertyPrediction-bandgap_JVASP_75_WSe2-dft_3d-test-mae.csv.zip

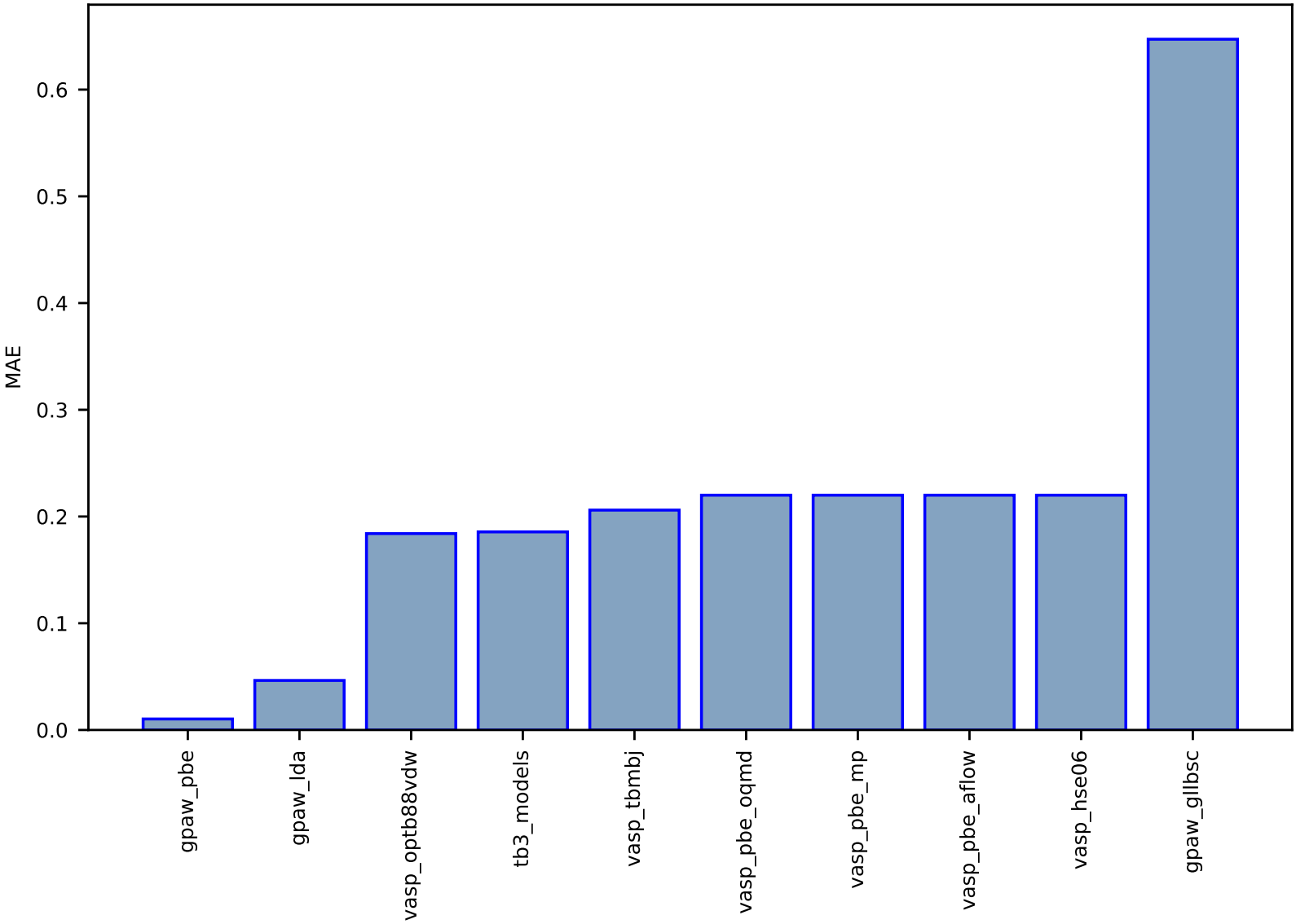


Fig S57:ES-SinglePropertyPrediction-slme_JVASP_7112_H6PbCI3N-dft_3d-test-mae.csv.zip

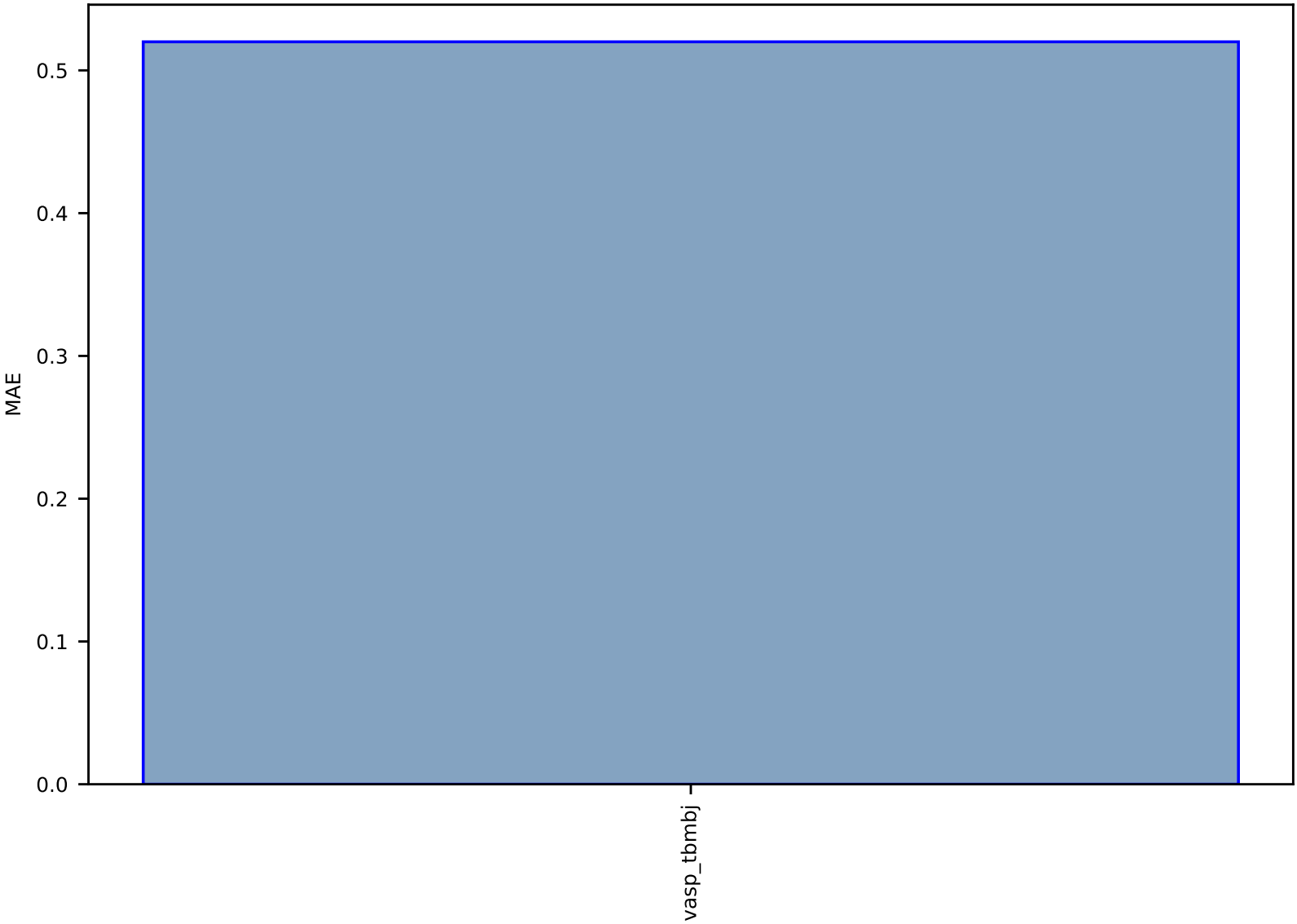


Fig S58:ES-SinglePropertyPrediction-bandgap_JVASP_1130_LiF-dft_3d-test-mae.csv.zip

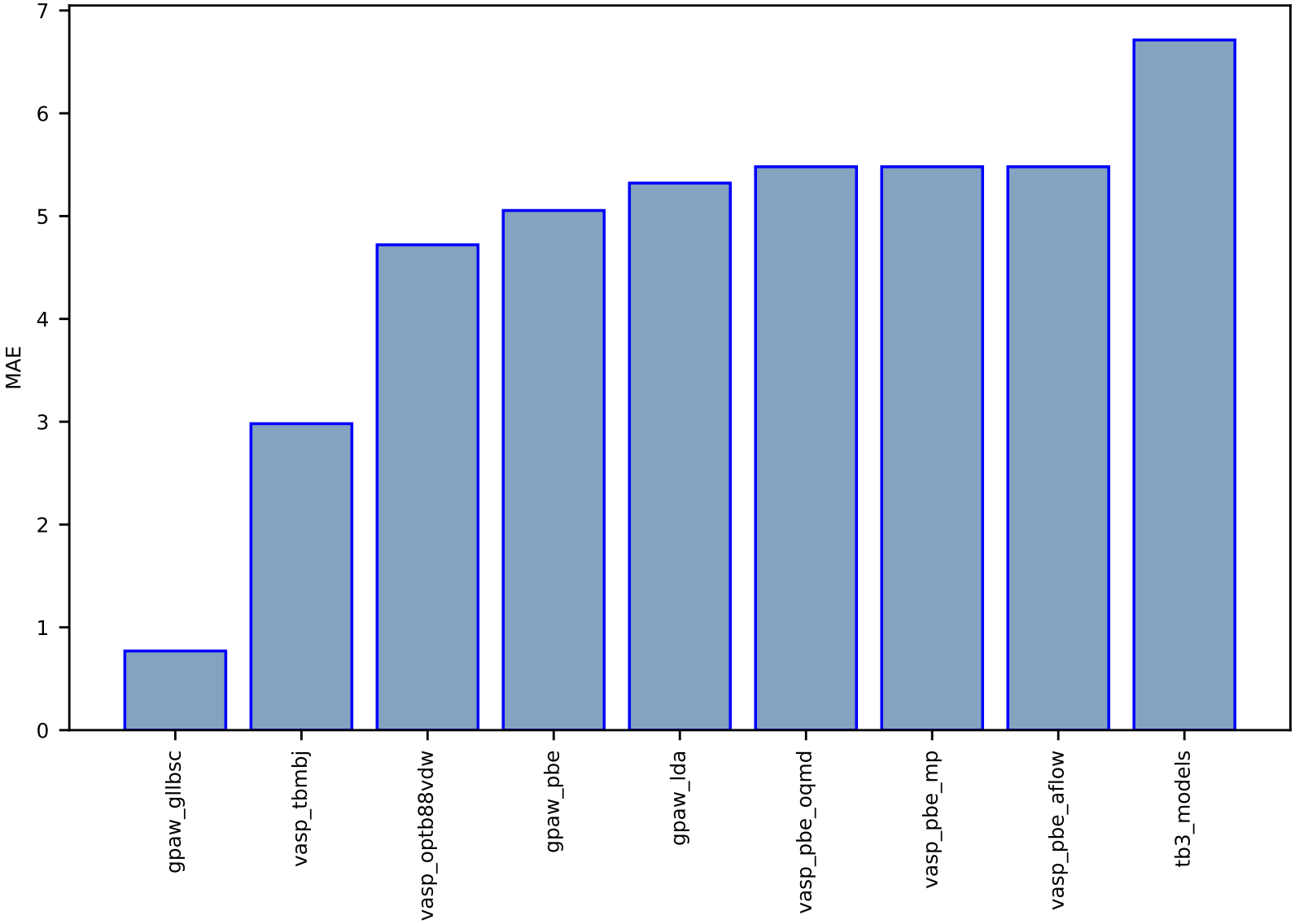


Fig S59:ES-SinglePropertyPrediction-bandgap_JVASP_57_MoSe2-dft_3d-test-mae.csv.zip

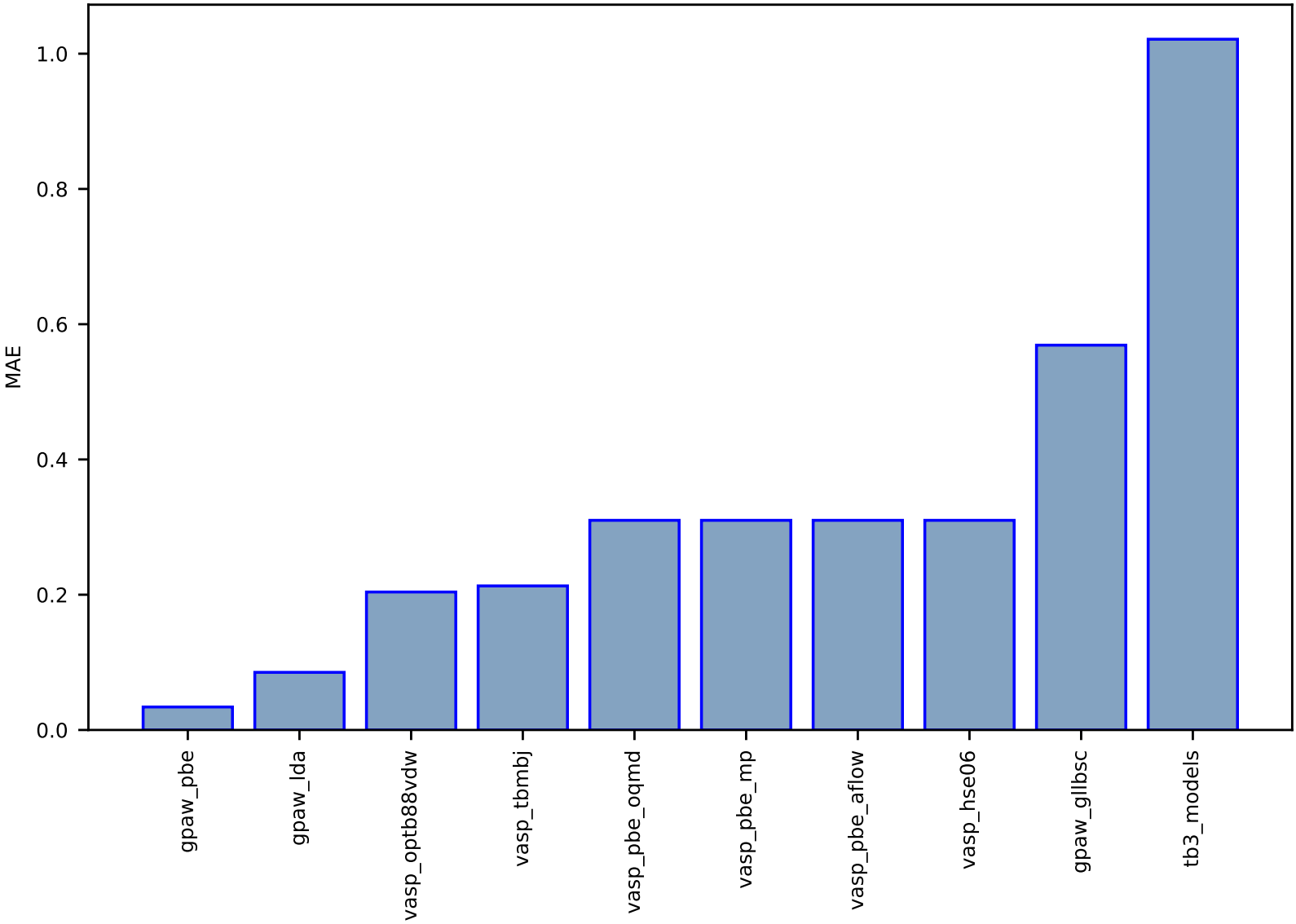


Fig S60:ES-SinglePropertyPrediction-slme_JVASP_266_InP-dft_3d-test-mae.csv.zip

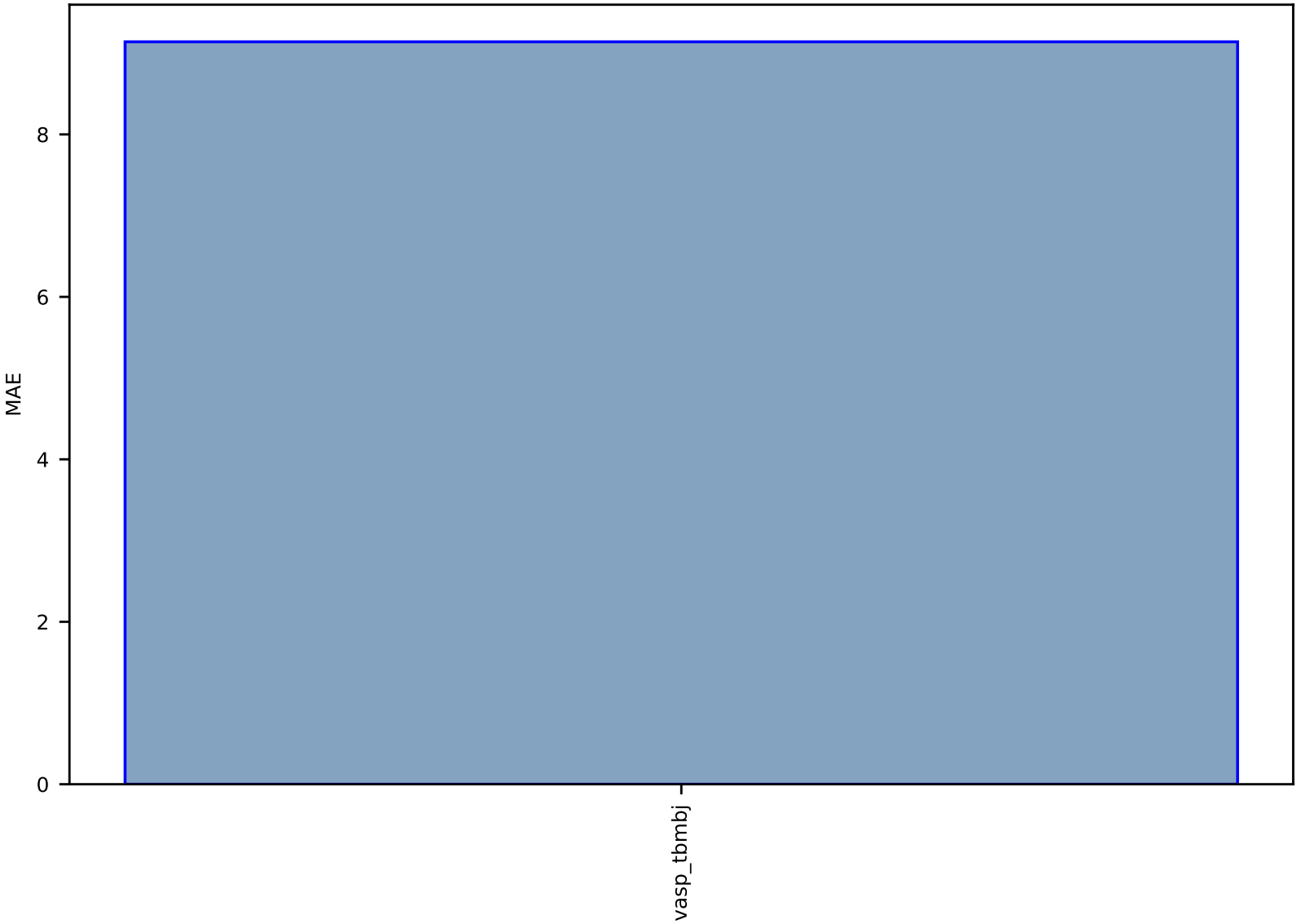


Fig S61:ES-SinglePropertyPrediction-slme-dft_3d-test-mae.csv.zip

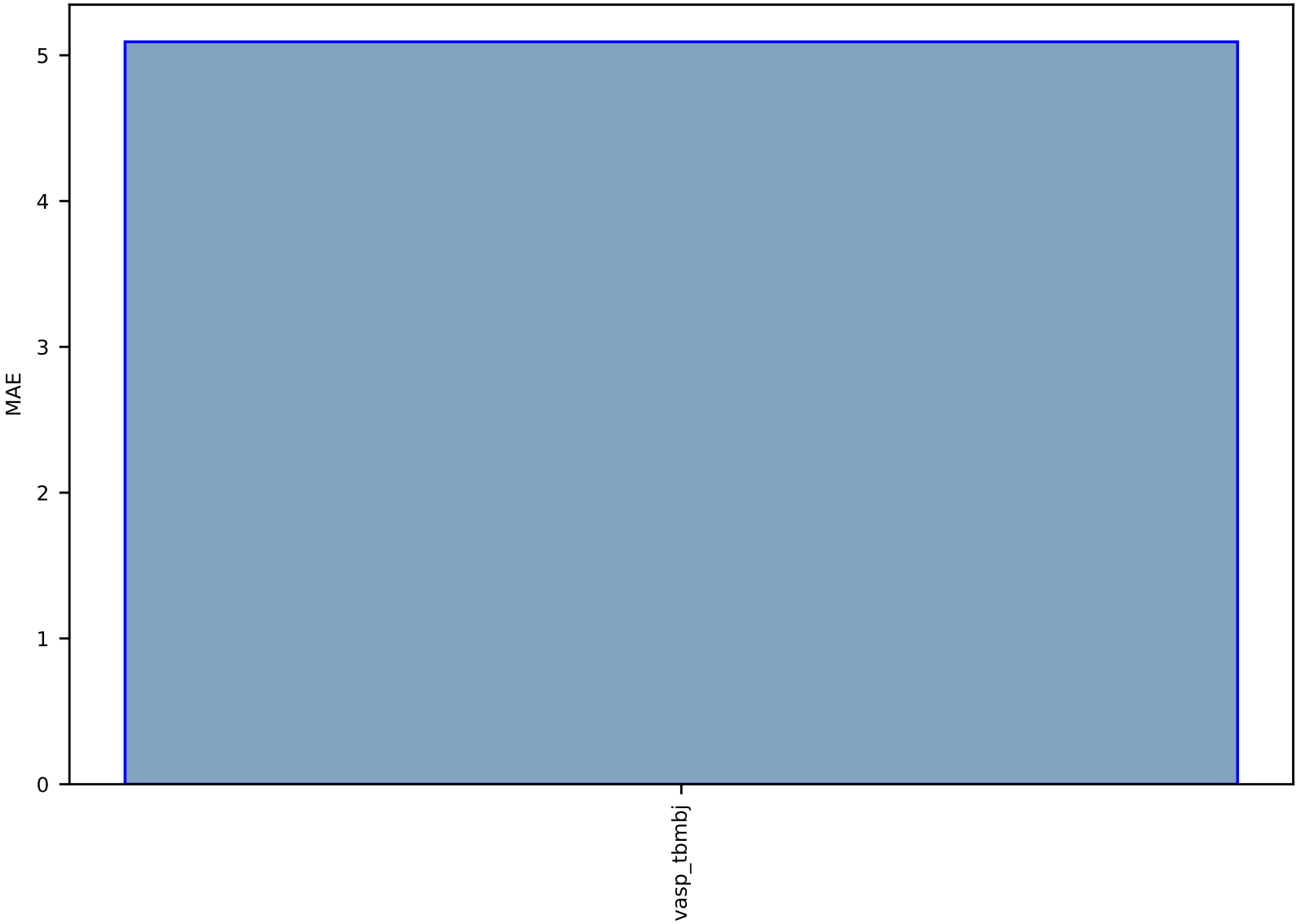


Fig S62:ES-SinglePropertyPrediction-bandgap_JVASP_1405_CaO-dft_3d-test-mae.csv.zip

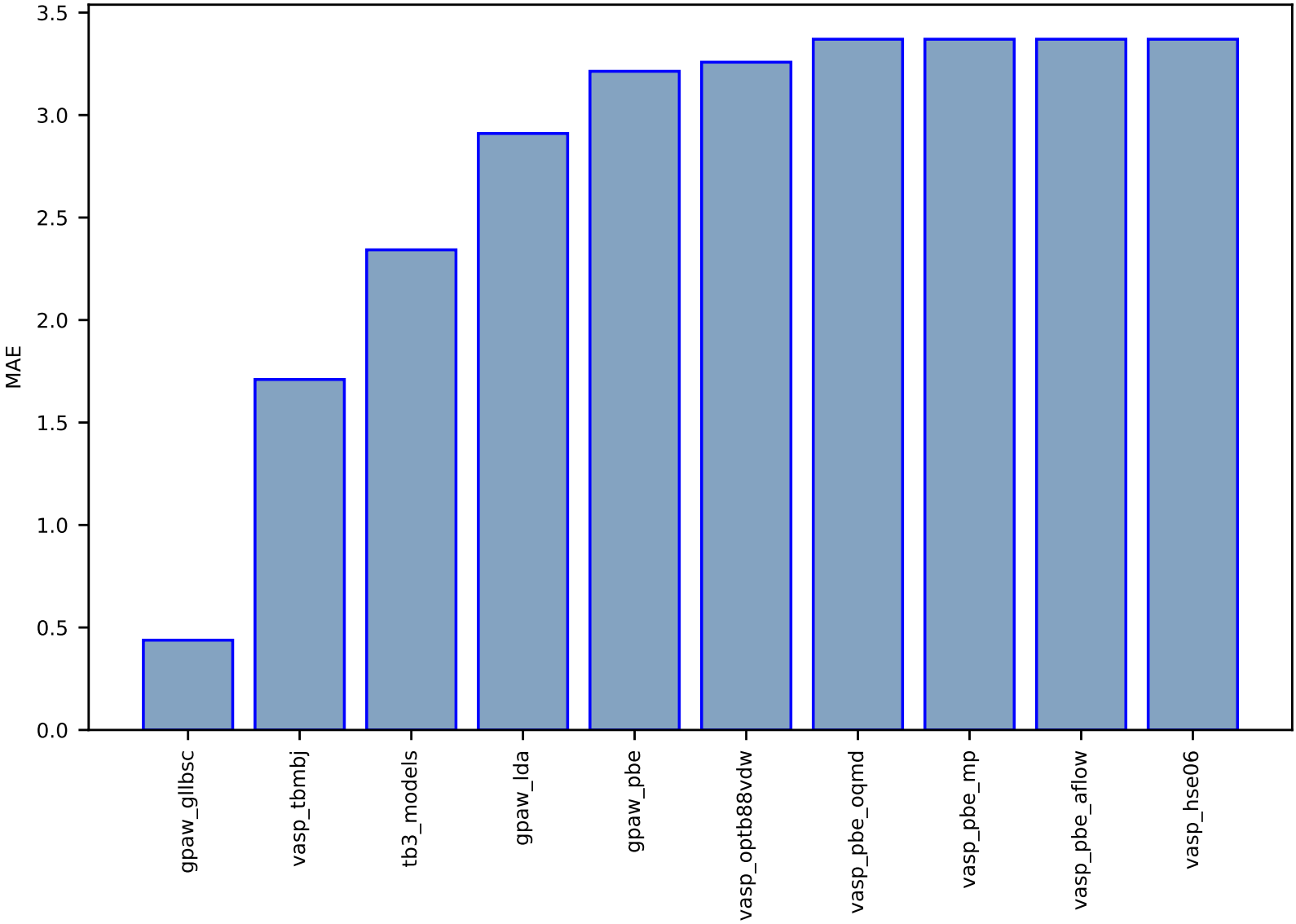


Fig S63:ES-SinglePropertyPrediction-bandgap_JVASP_8158_SiC-dft_3d-test-mae.csv.zip

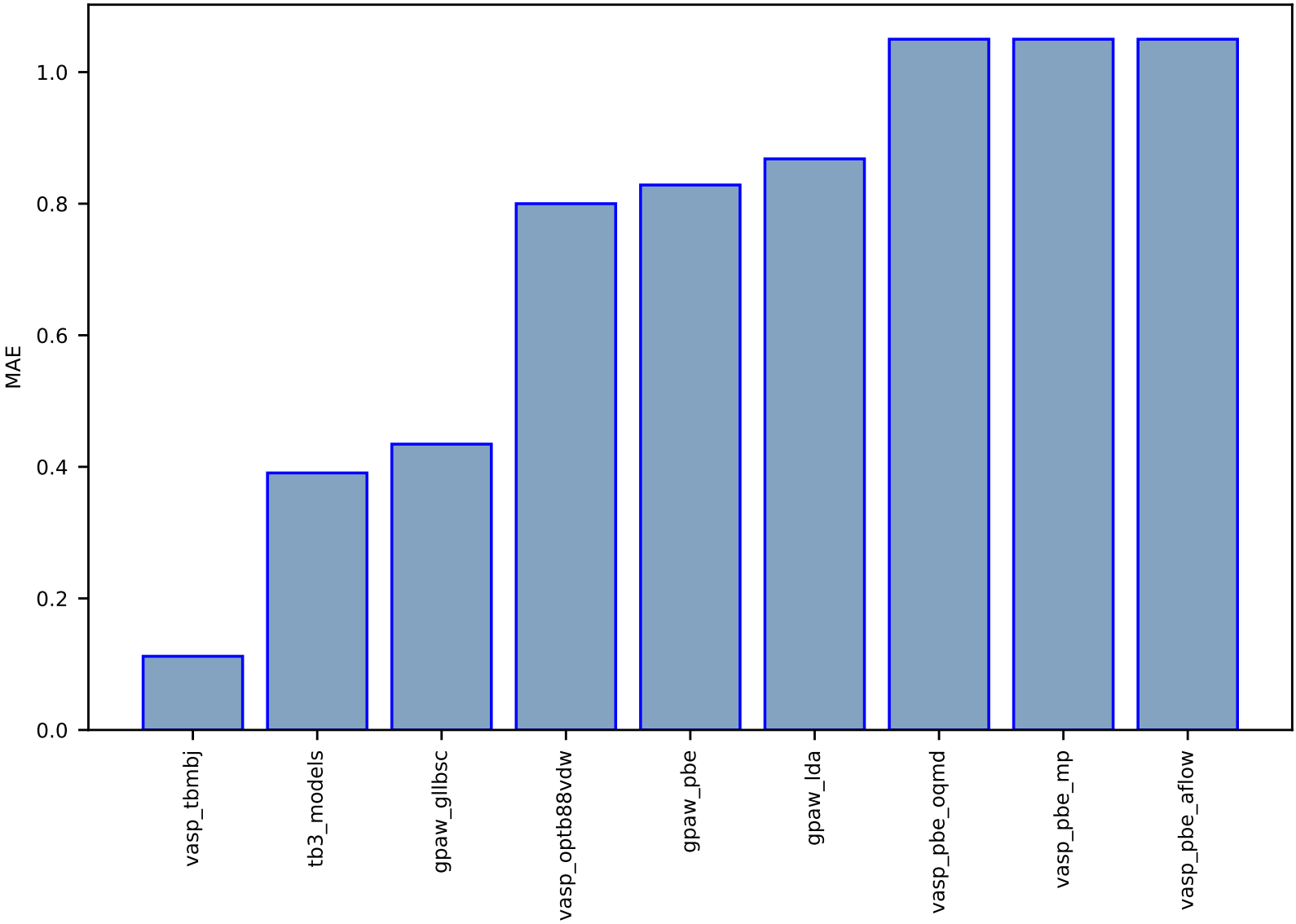


Fig S64:ES-SinglePropertyPrediction-bandgap_JVASP_1189_InSb-dft_3d-test-mae.csv.zip

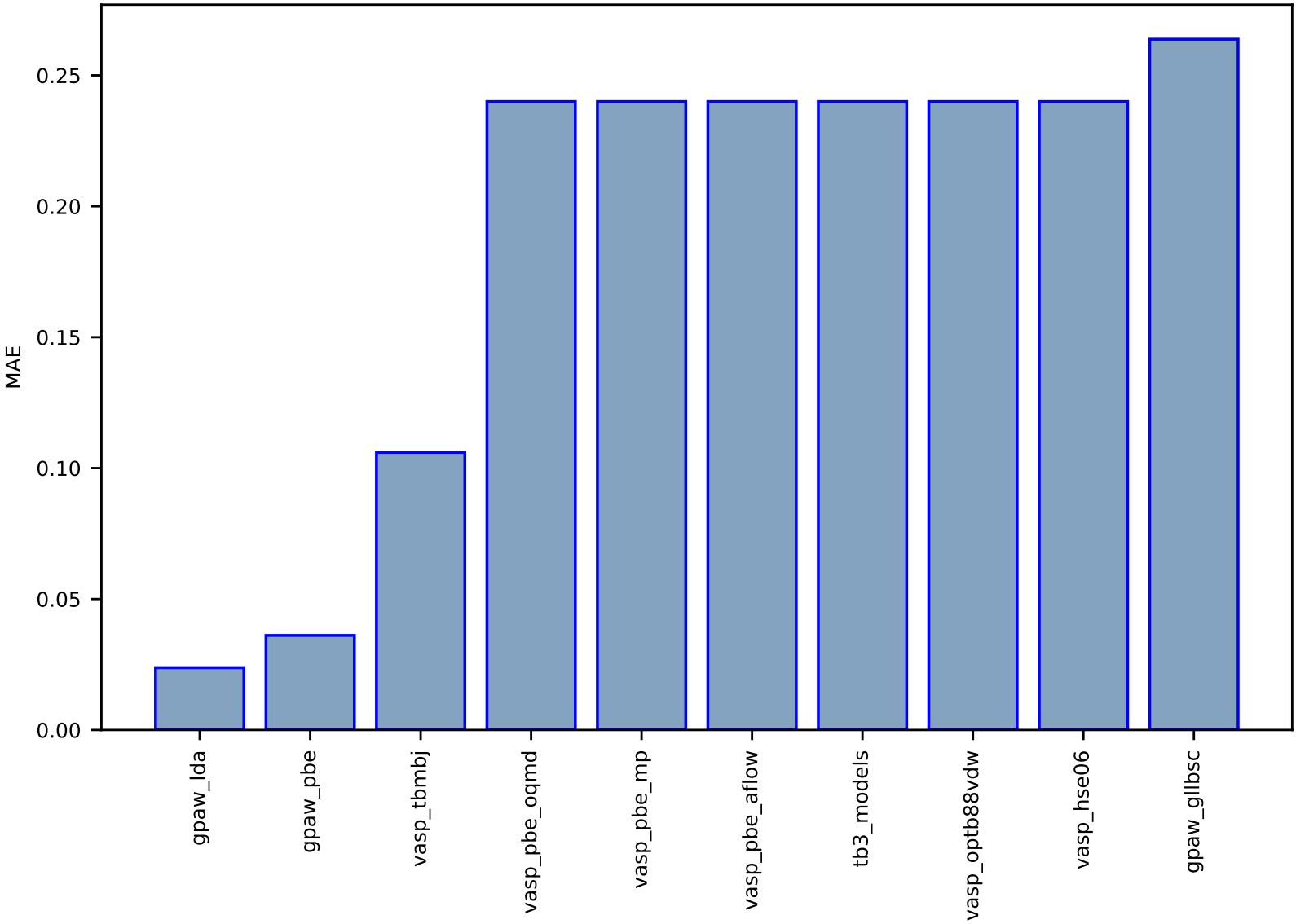


Fig S65:ES-SinglePropertyPrediction-bandgap_JVASP_299_SnSe-dft_3d-test-mae.csv.zip

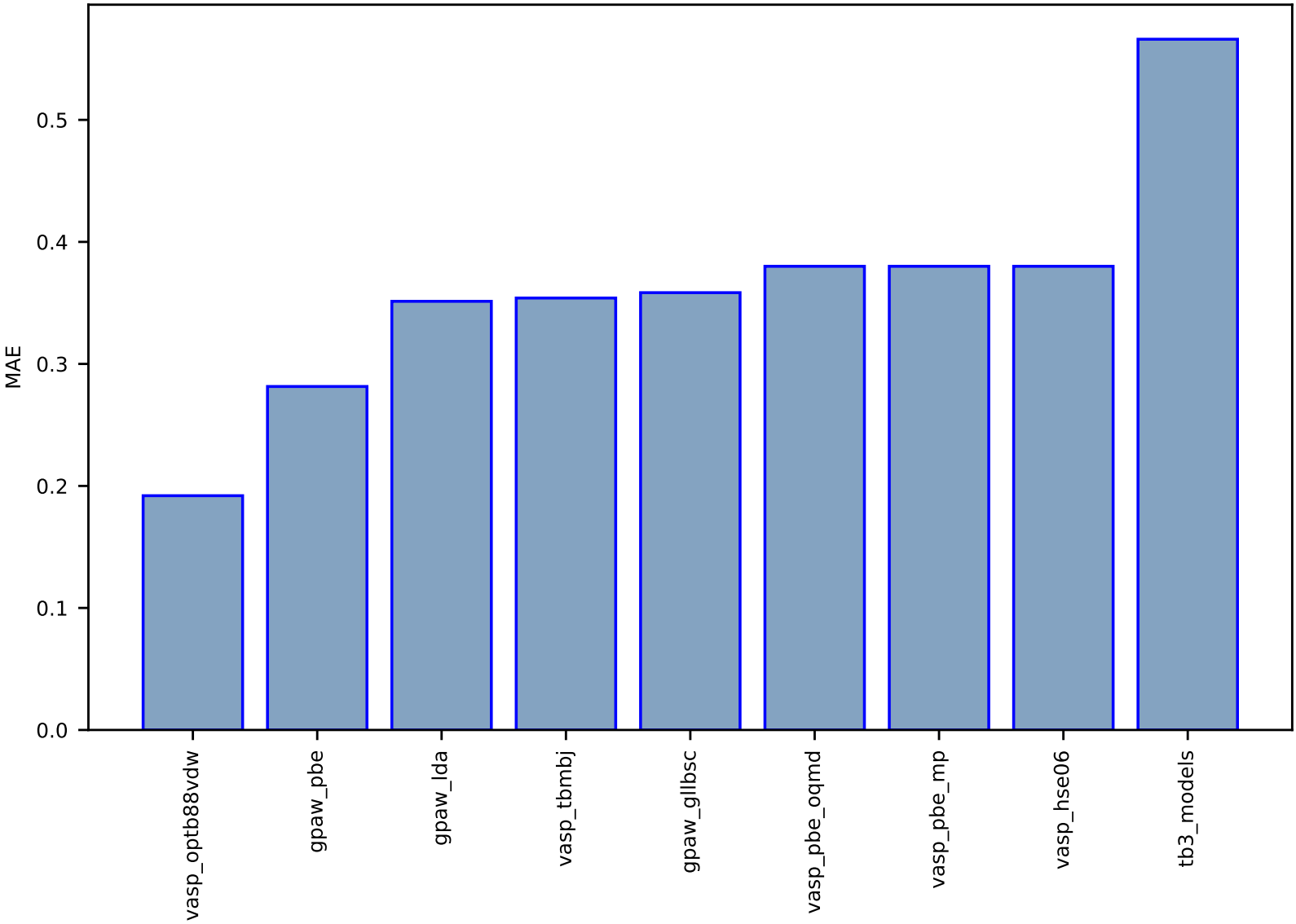


Fig S66:ES-SinglePropertyPrediction-bandgap_JVASP_1183_InP-dft_3d-test-mae.csv.zip

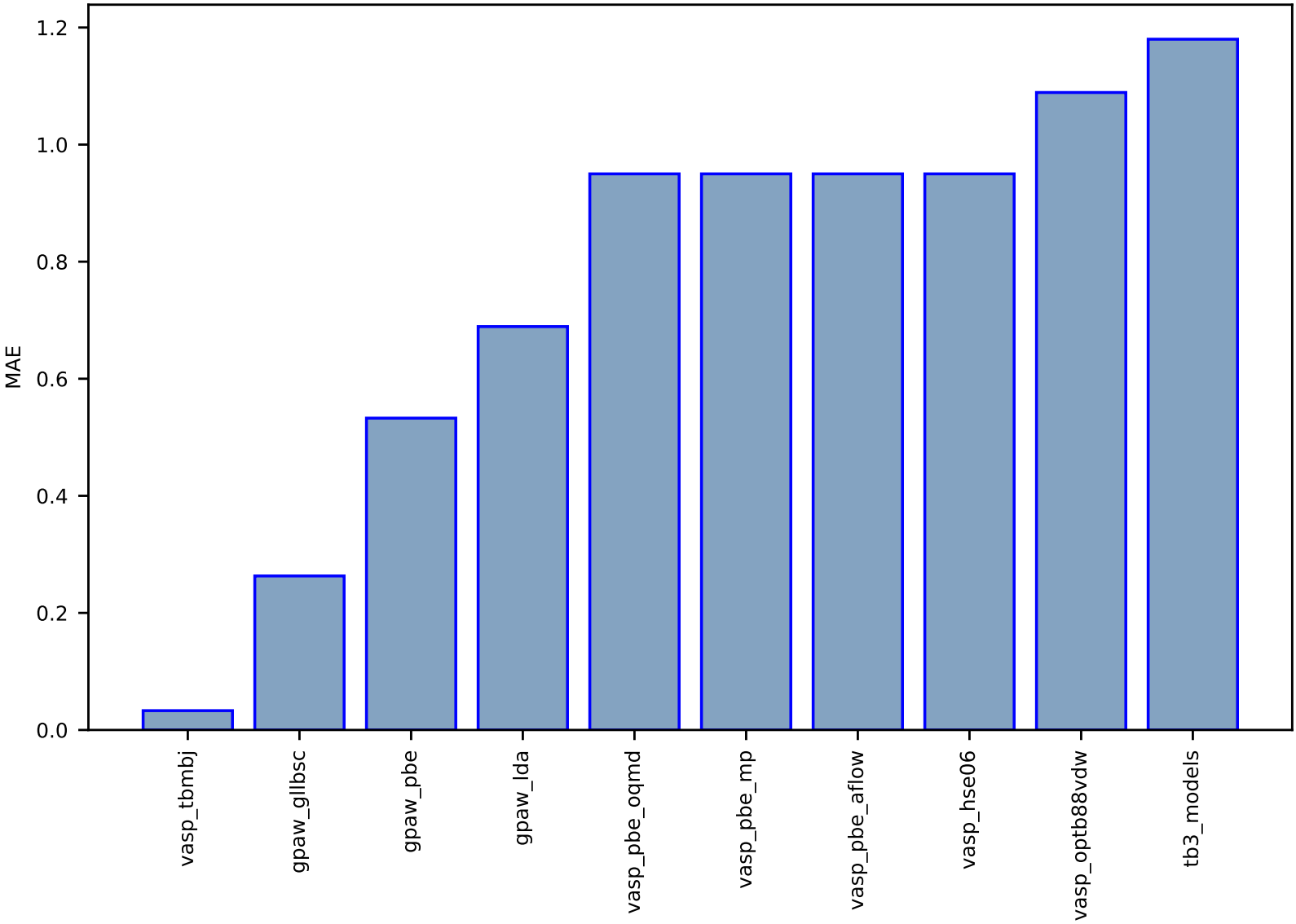


Fig S67:ES-SinglePropertyPrediction-bandgap_JVASP_7678_MgSe-dft_3d-test-mae.csv.zip

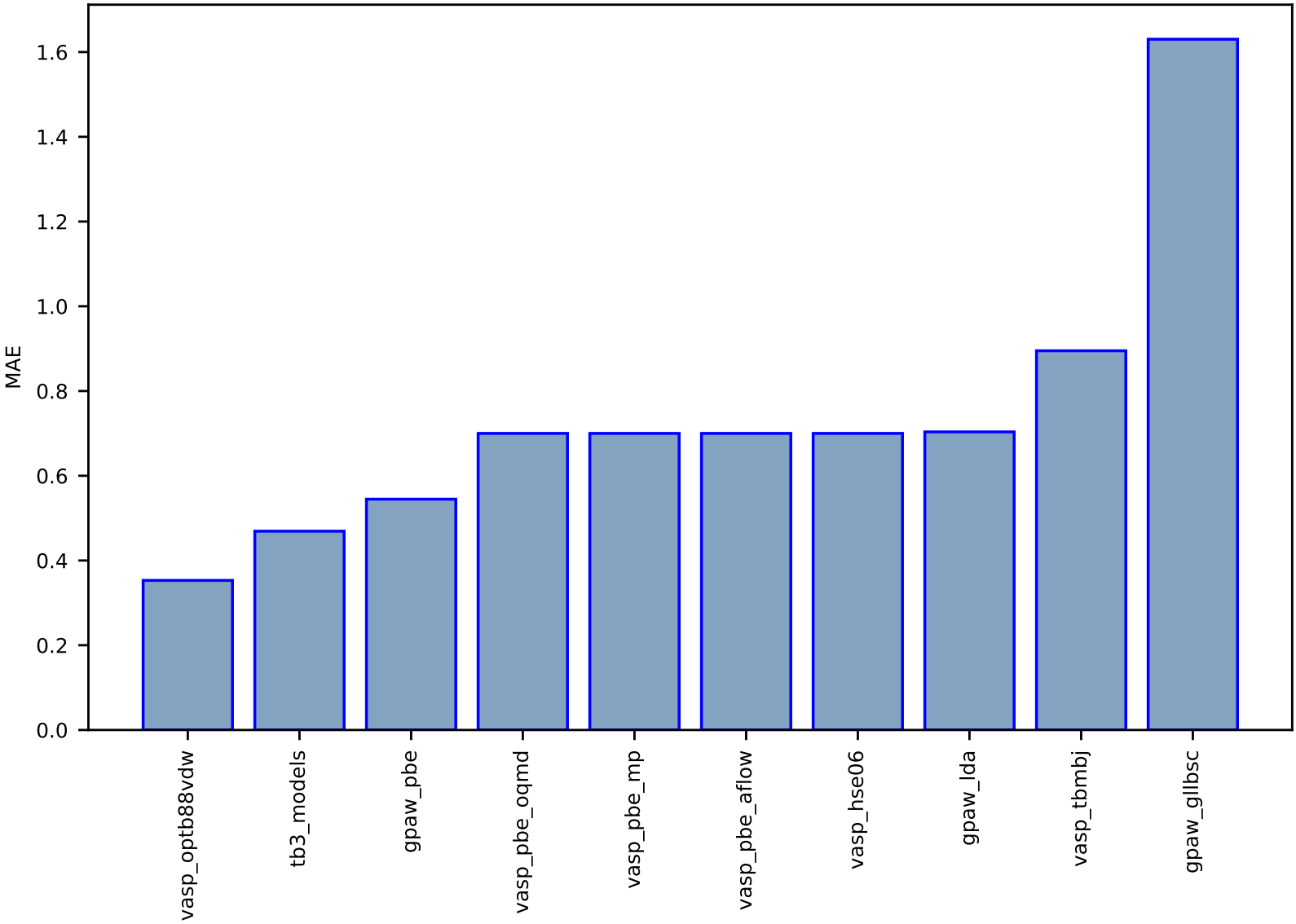


Fig S68:ES-SinglePropertyPrediction-bandgap_JVASP_8082_SrTiO3-dft_3d-test-mae.csv.zip

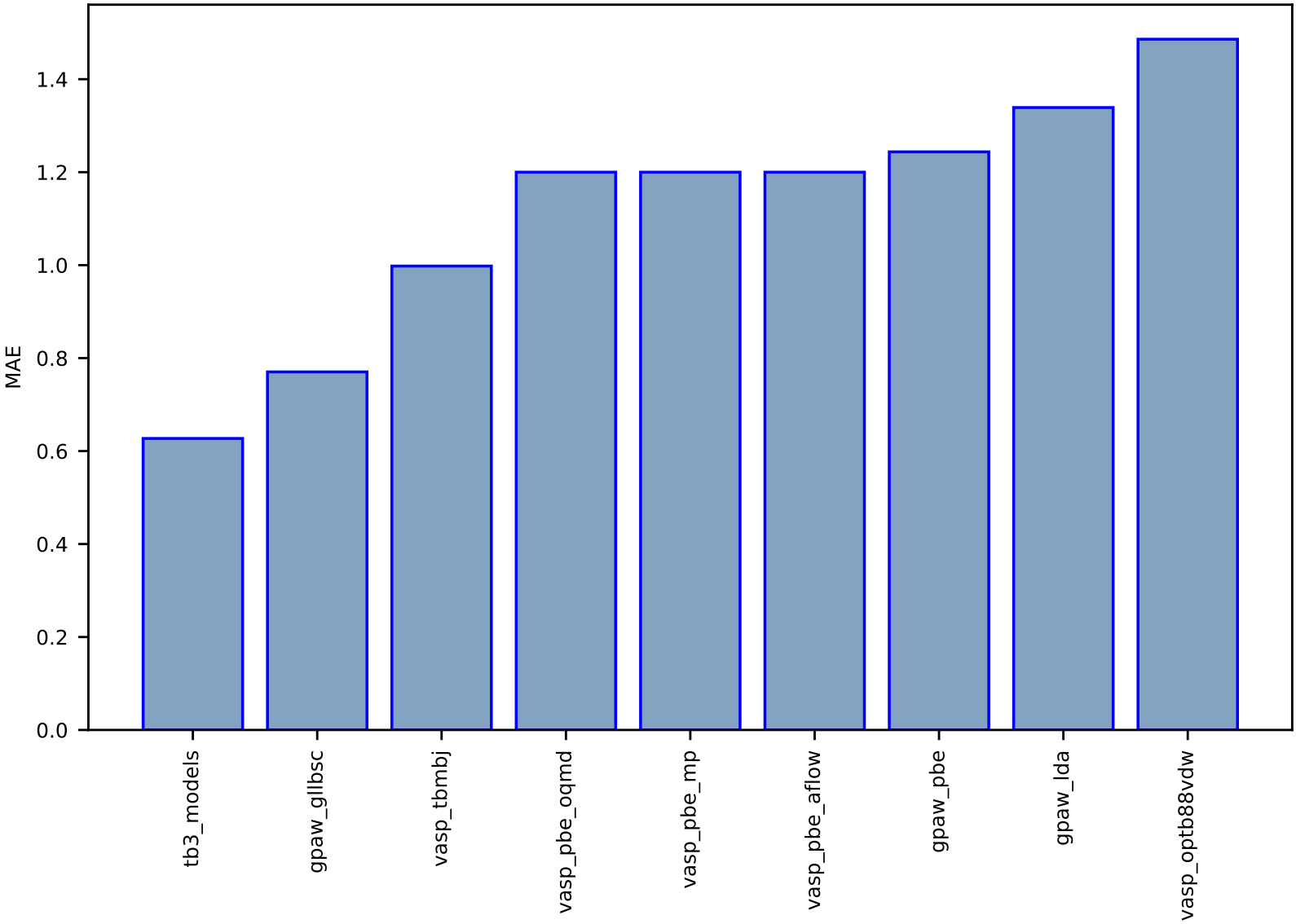


Fig S69:ES-SinglePropertyPrediction-bandgap_JVASP_1954_AgCl-dft_3d-test-mae.csv.zip

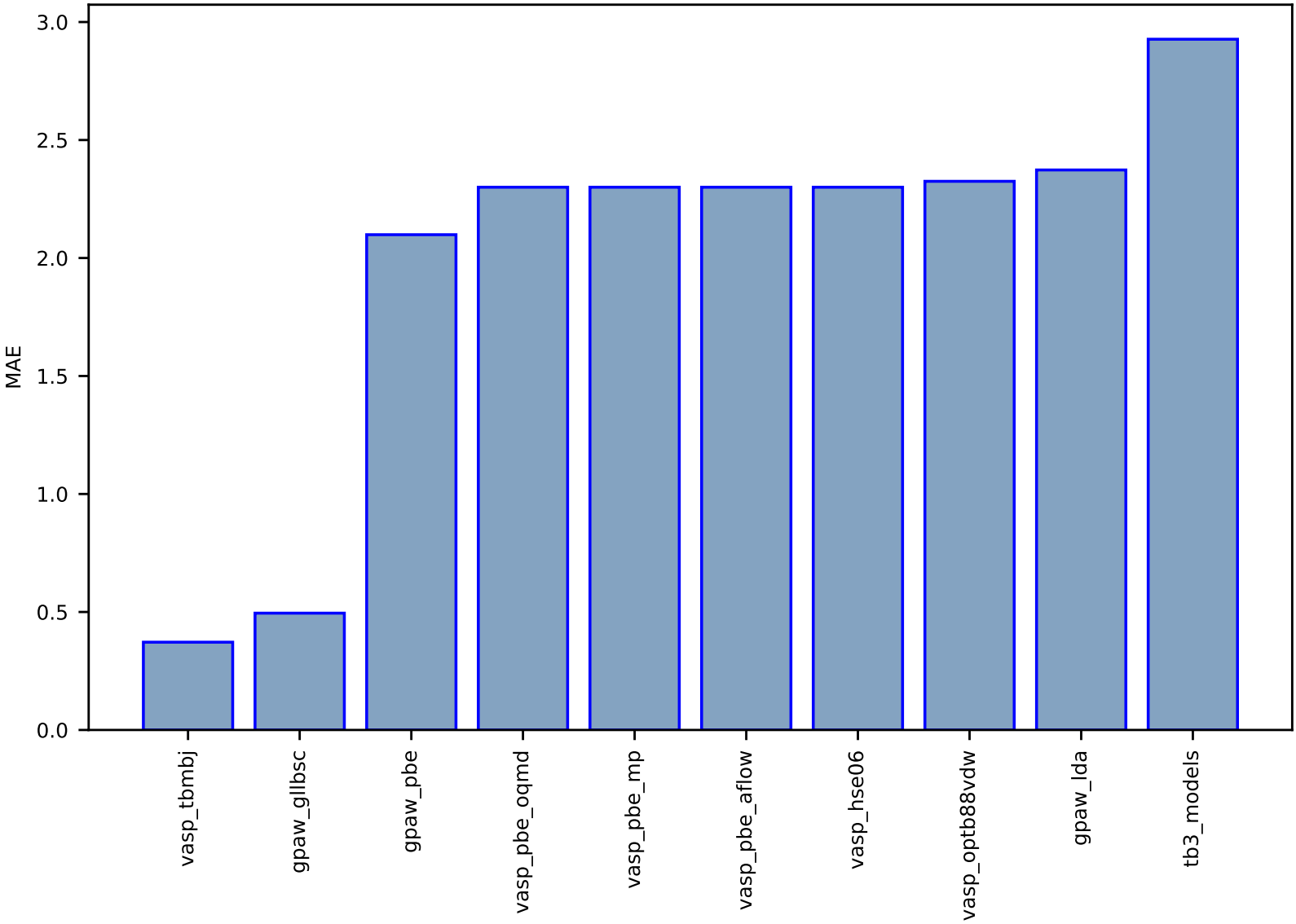


Fig S70:ES-SinglePropertyPrediction-bandgap_JVASP_1702_ZnS-dft_3d-test-mae.csv.zip

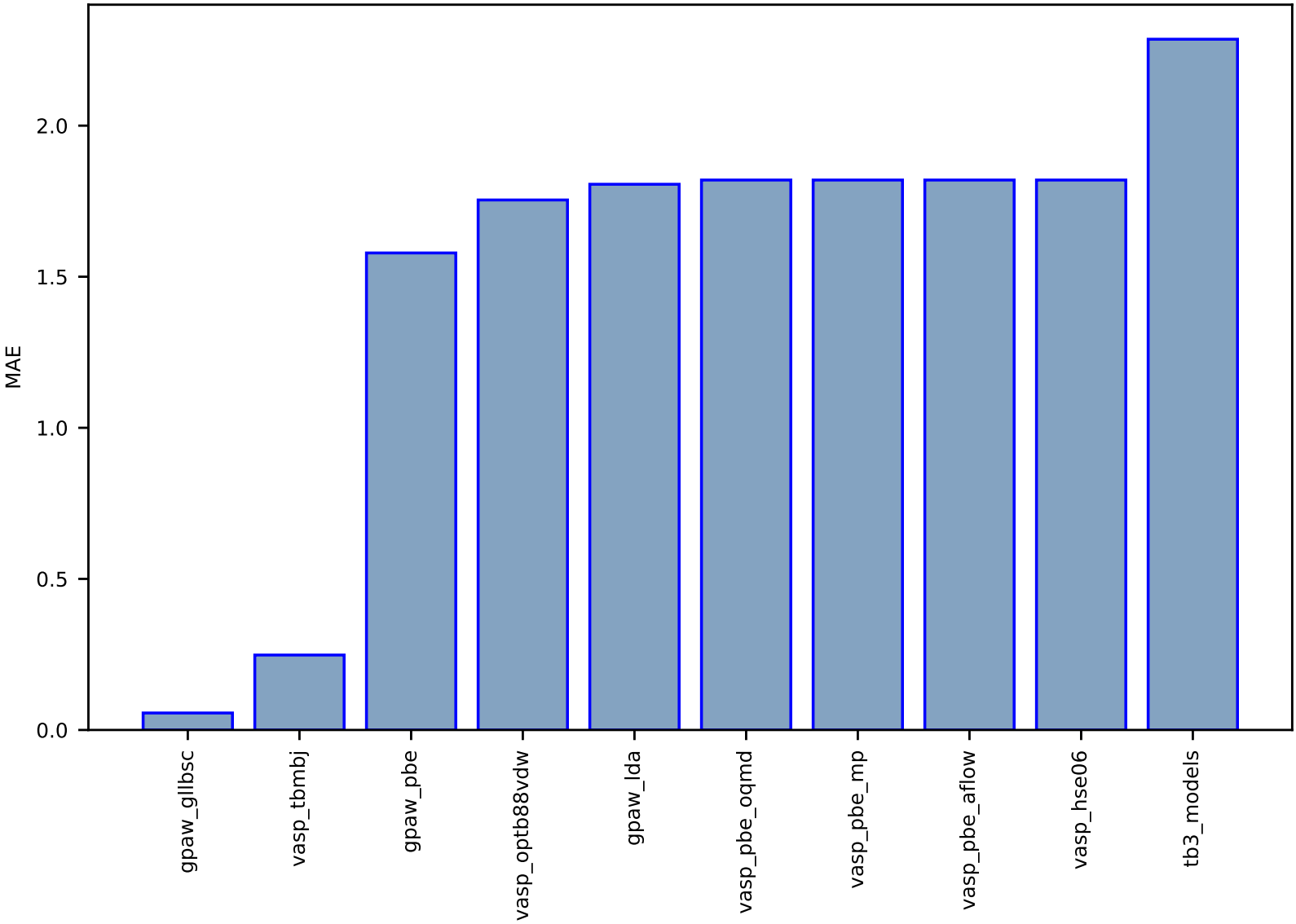


Fig S71:ES-SinglePropertyPrediction-bandgap_JVASP_1453_AlCuO2-dft_3d-test-mae.csv.zip

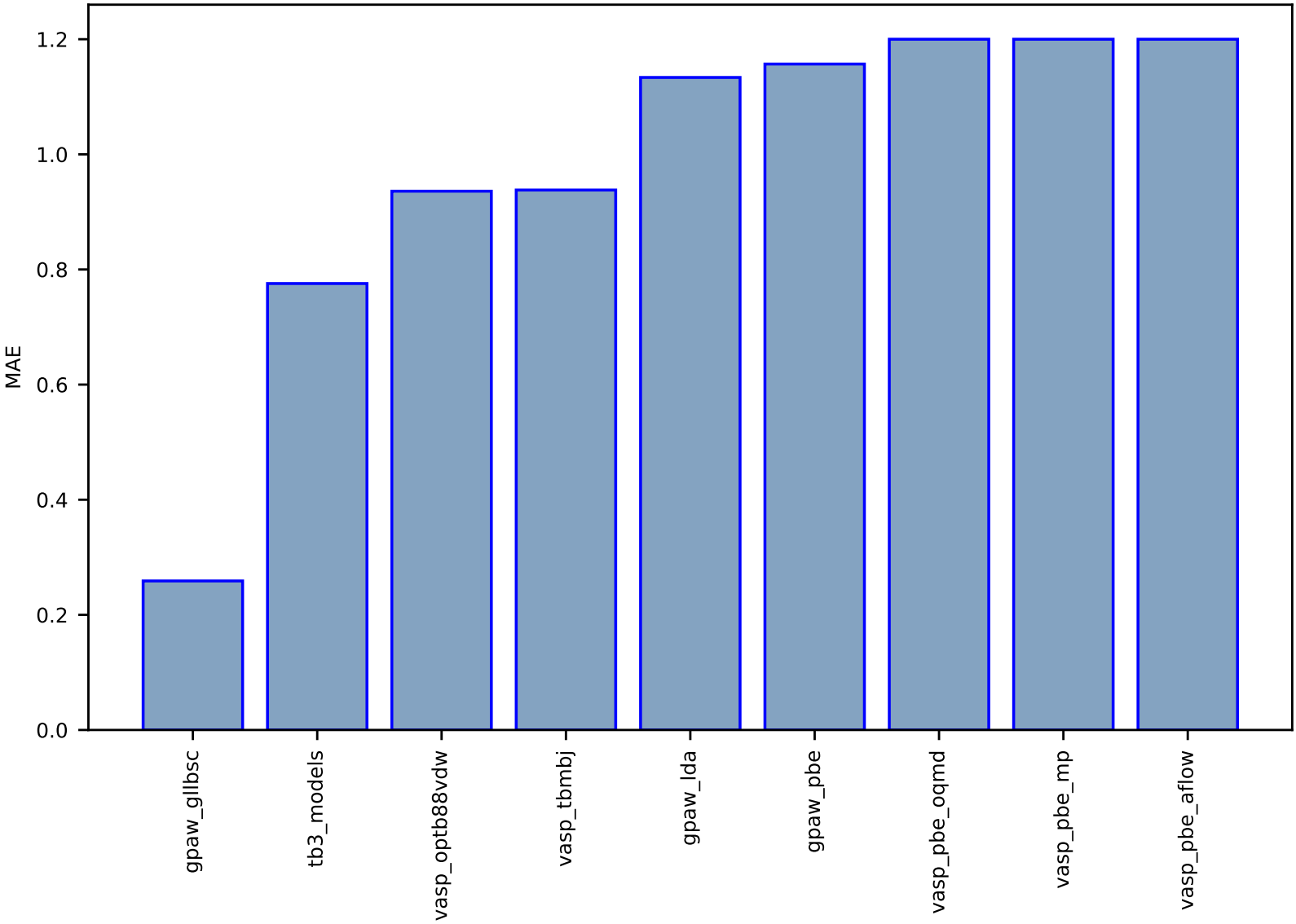


Fig S72:ES-SinglePropertyPrediction-bandgap-dft_3d-test-mae.csv.zip

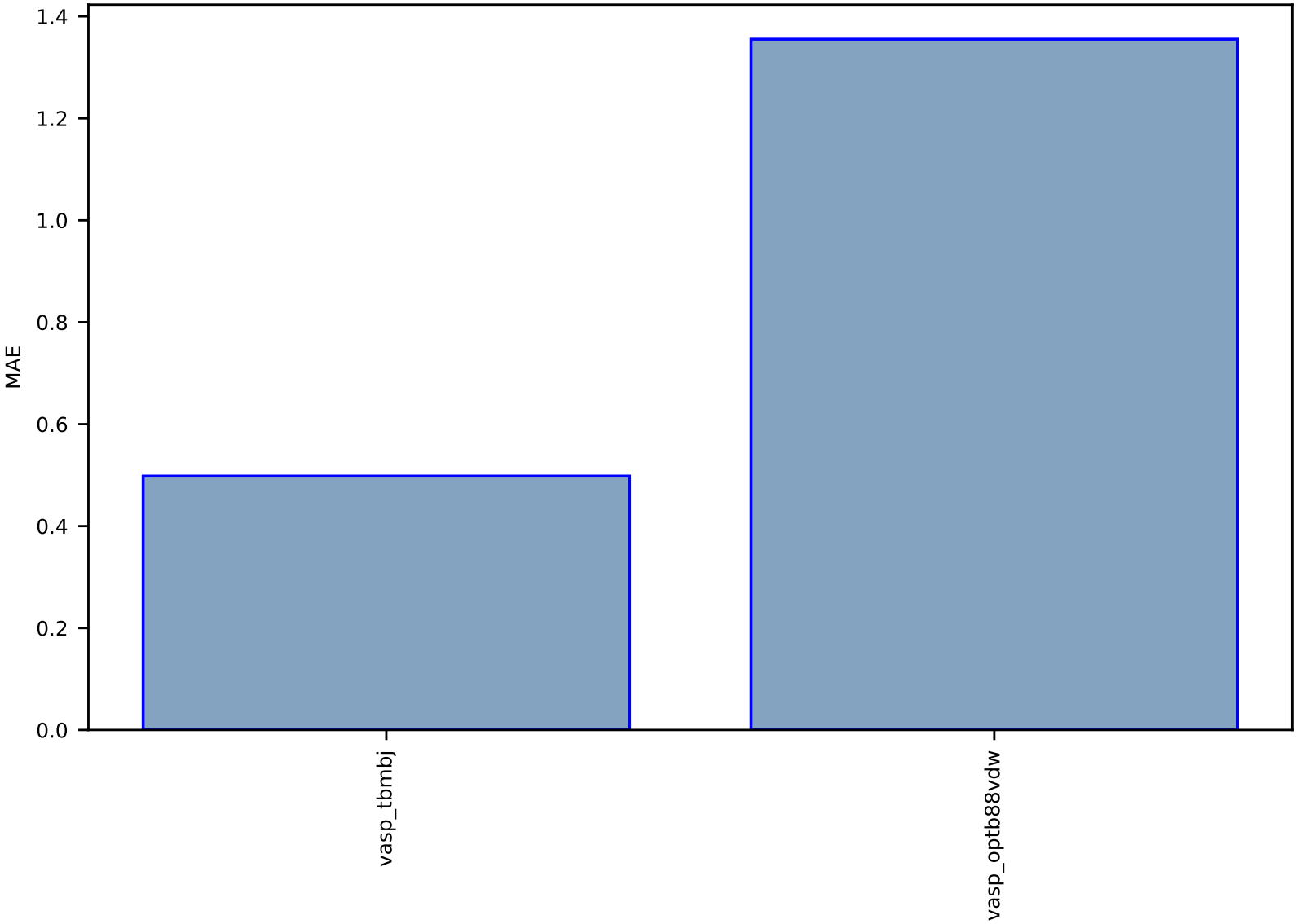


Fig S73:ES-SinglePropertyPrediction-bandgap_JVASP_91_C-dft_3d-test-mae.csv.zip

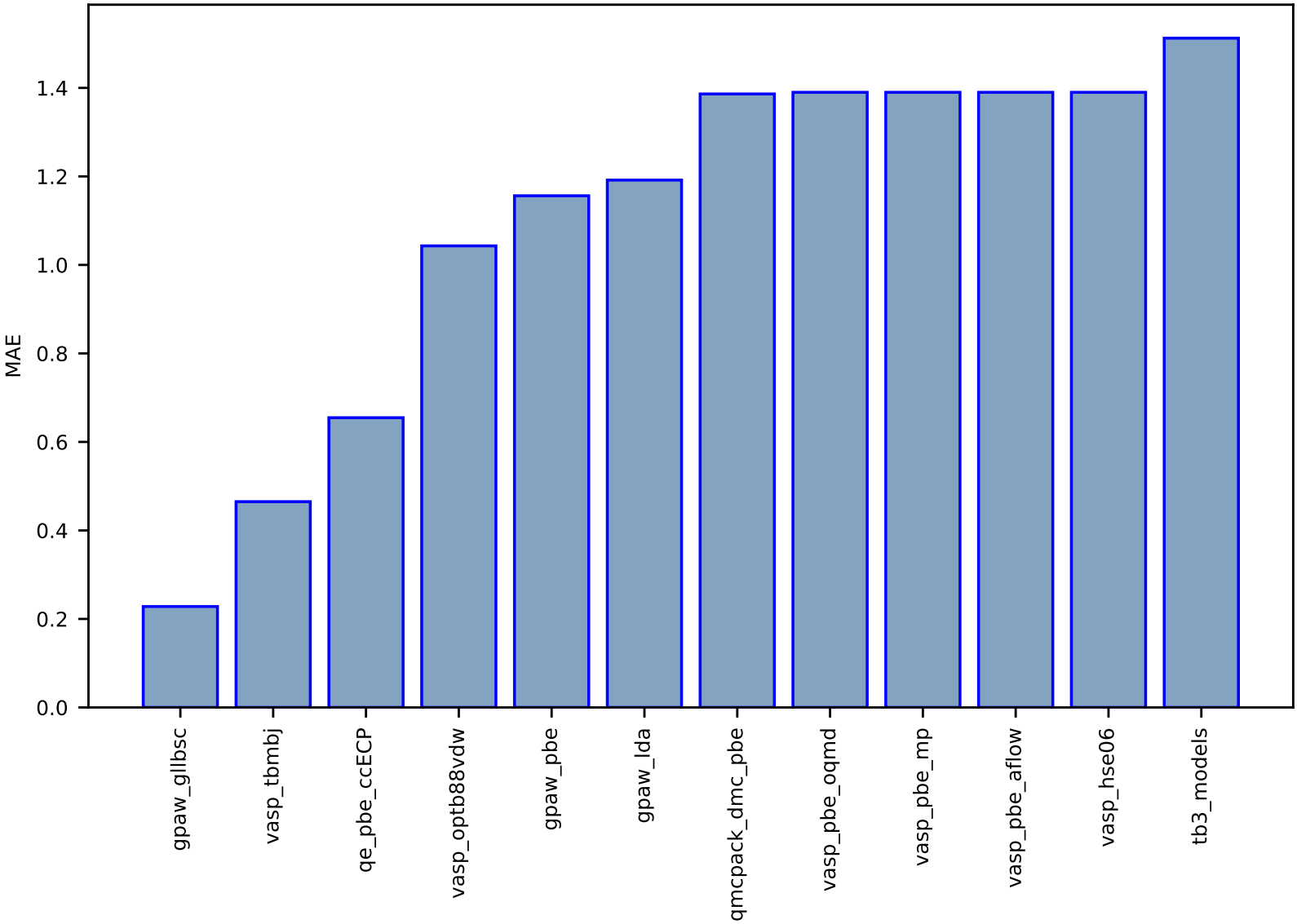


Fig S74:ES-SinglePropertyPrediction-bandgap_JVASP_23_CdTe-dft_3d-test-mae.csv.zip

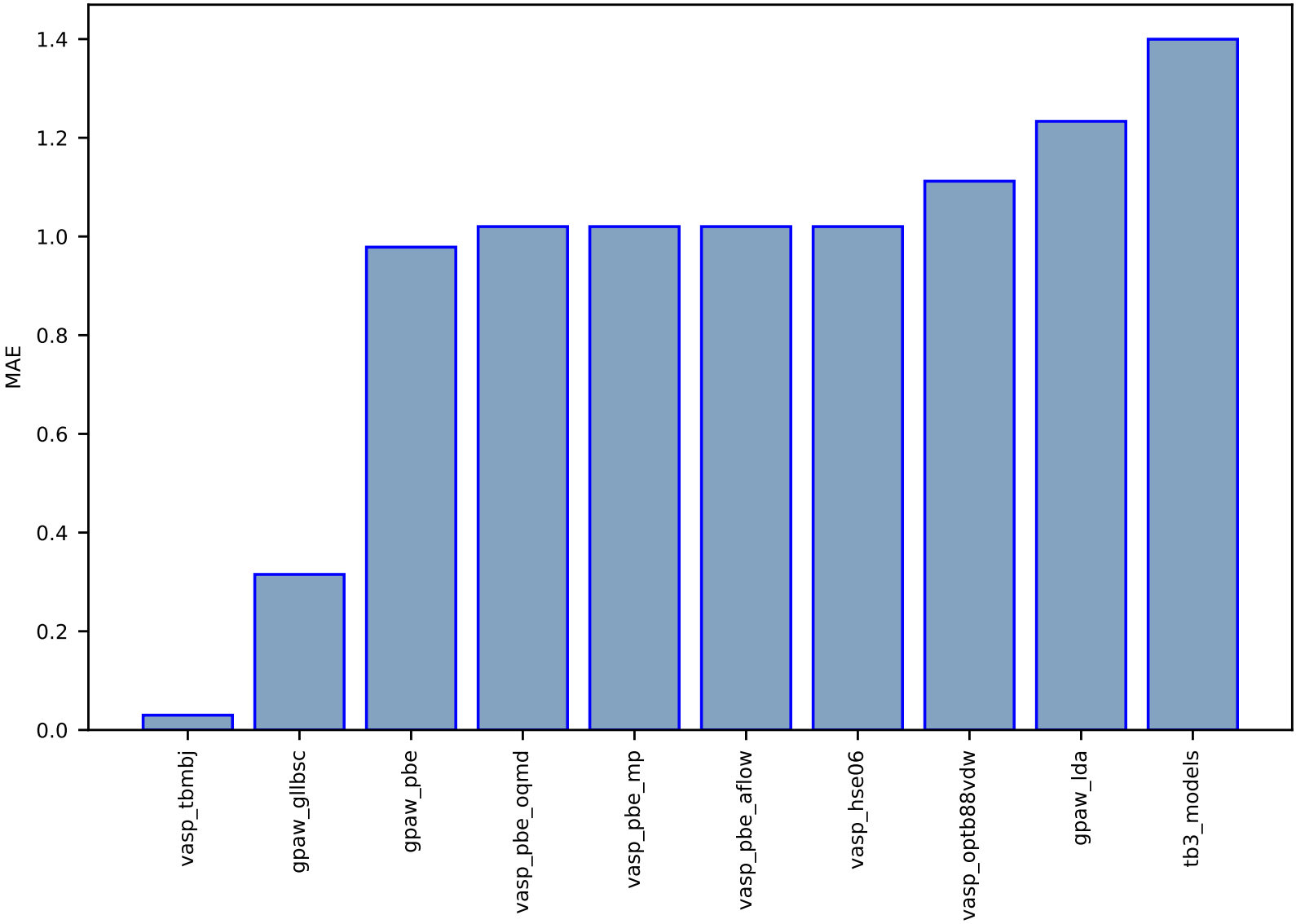


Fig S75:ES-Spectra-dielectric_function_JVASP_266_InP-dft_3d-test-multimae.csv.zip

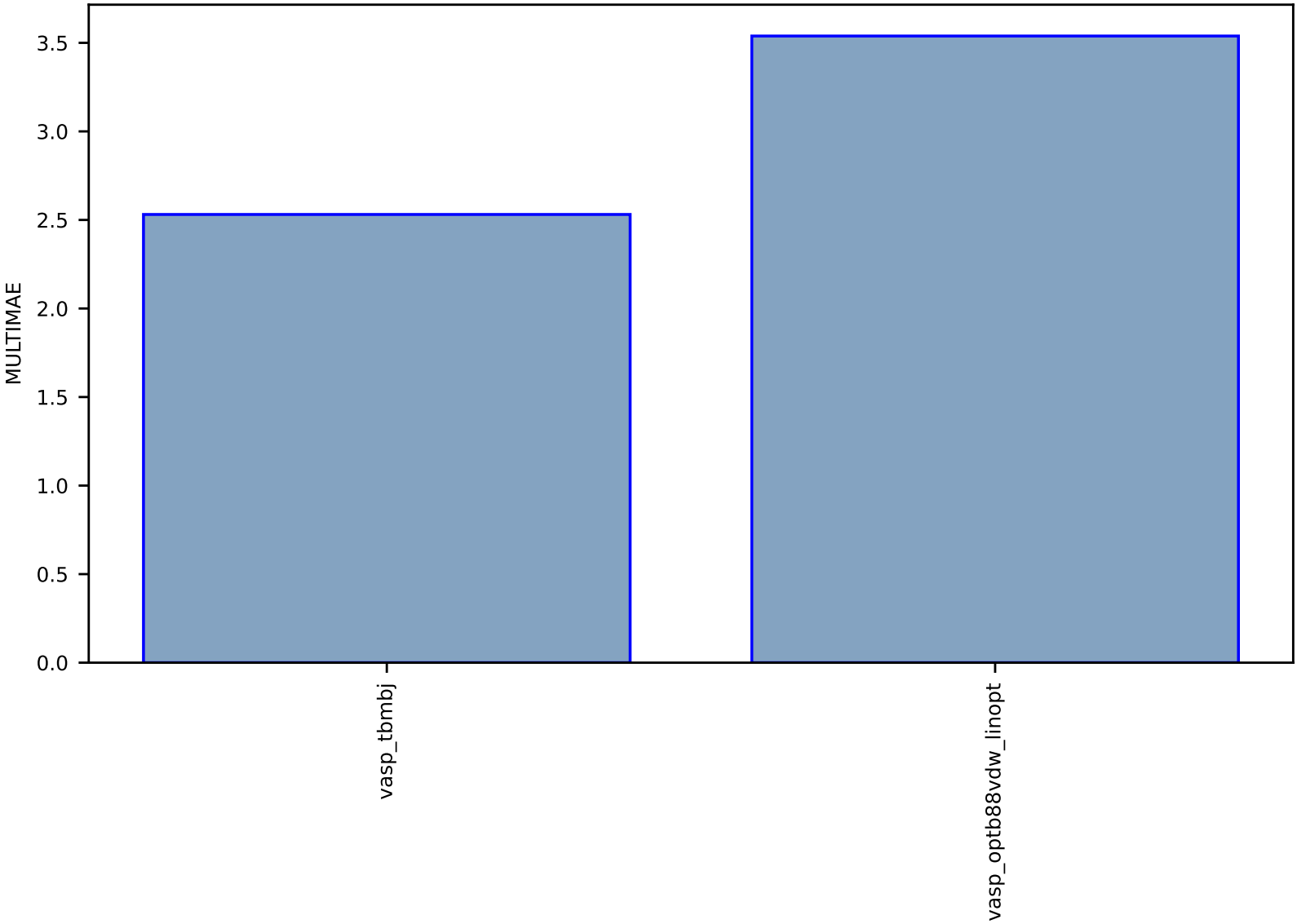


Fig S76:ES-Spectra-dielectric_function_JVASP_1002_Si-dft_3d-test-multimae.csv.zip

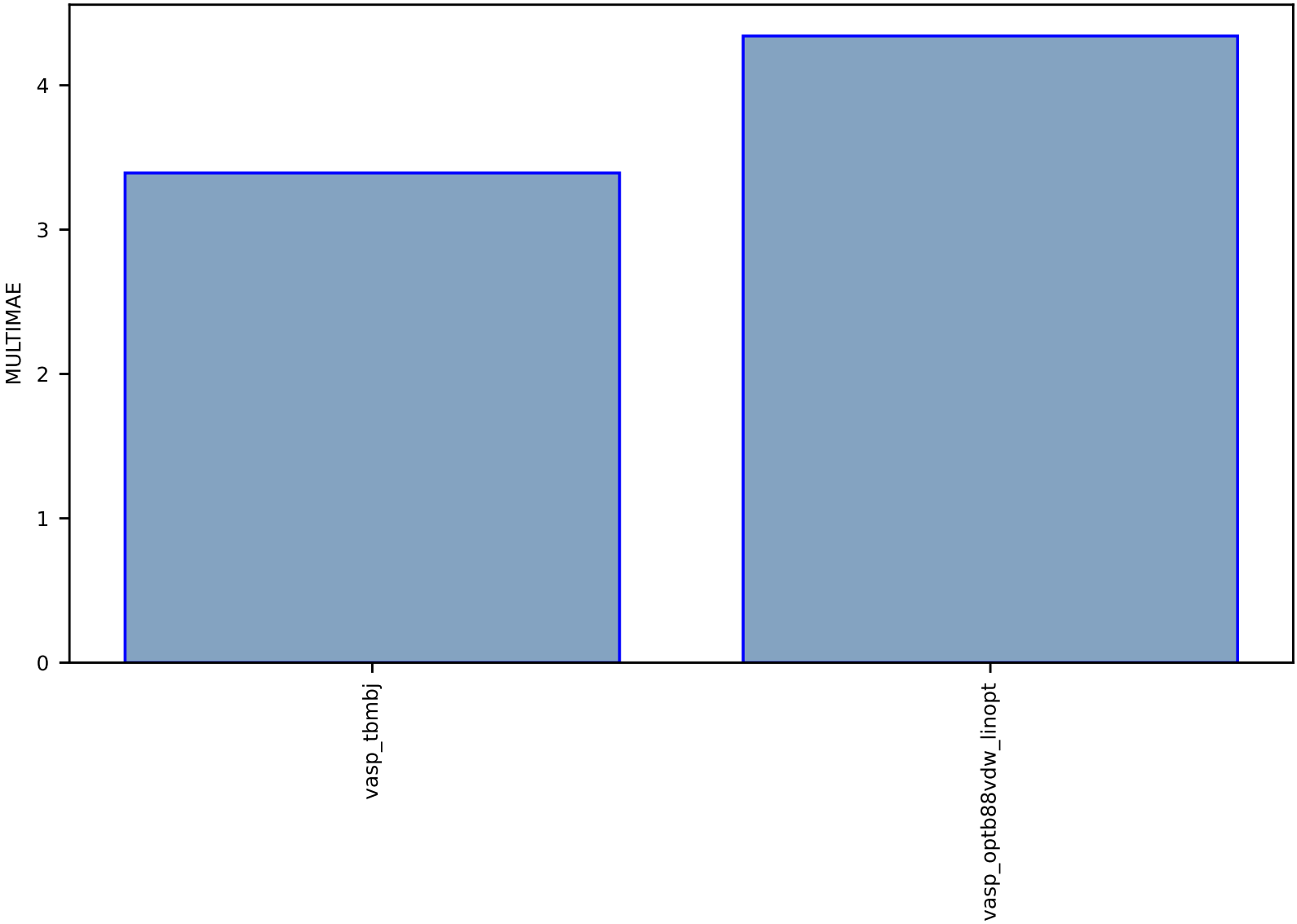


Fig S77:ES-SinglePropertyPrediction-bandgap_JVASP_72_WS2-dft_3d-test-mae.csv.zip

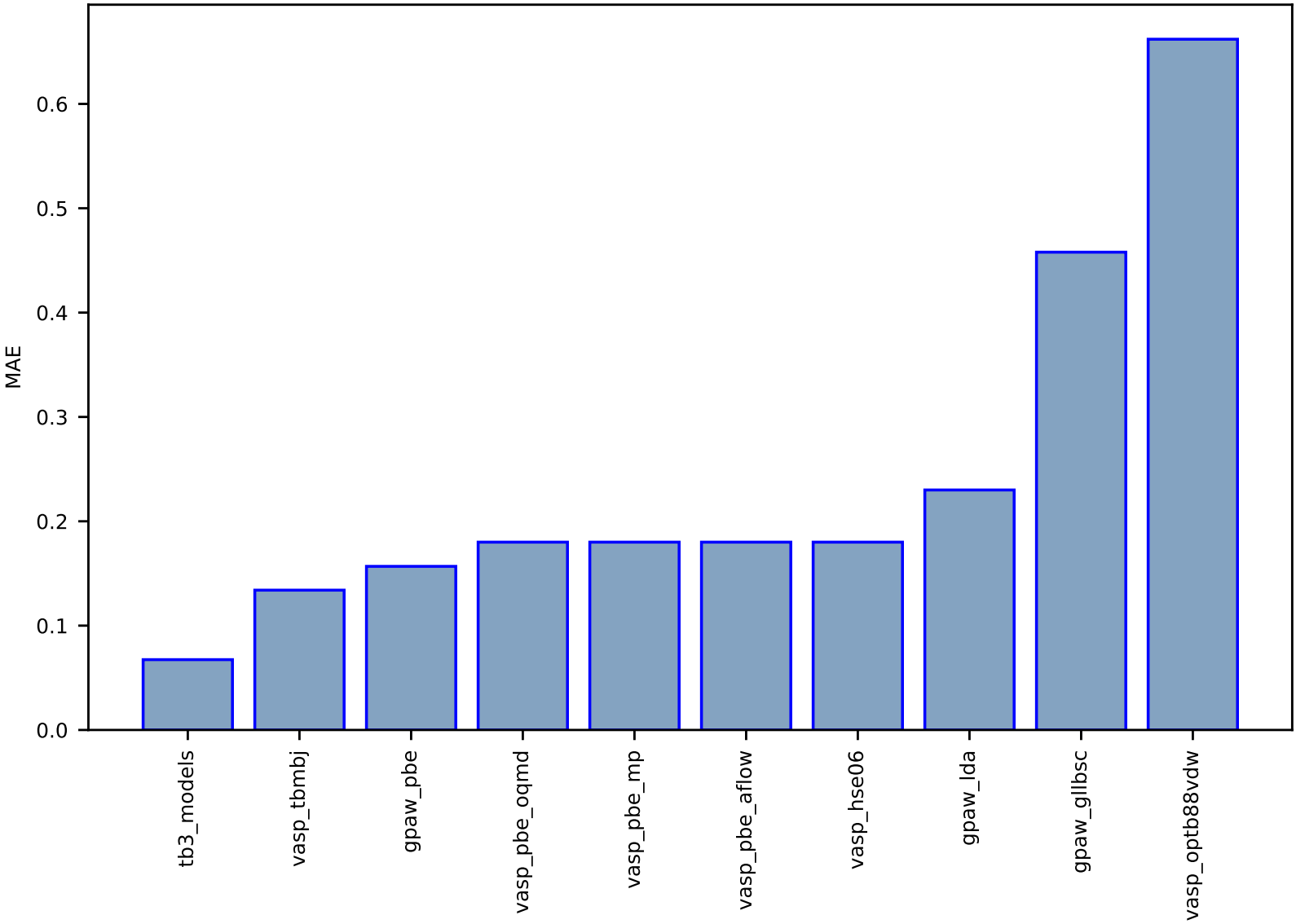


Fig S78:ES-SinglePropertyPrediction-bandgap_JVASP_1201_CuCl-dft_3d-test-mae.csv.zip

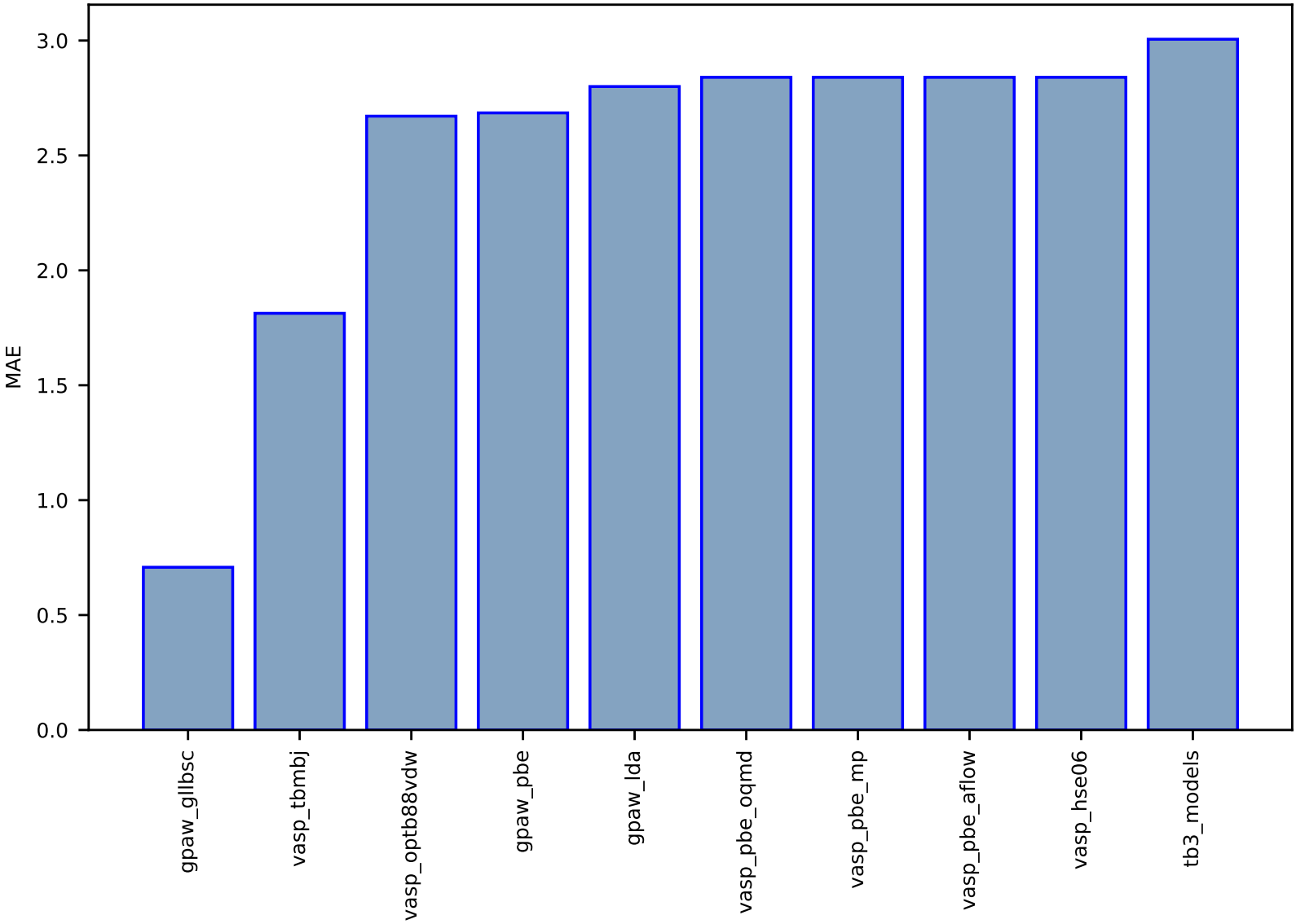


Fig S79:ES-SinglePropertyPrediction-slme_JVASP_7757_CdTe-dft_3d-test-mae.csv.zip

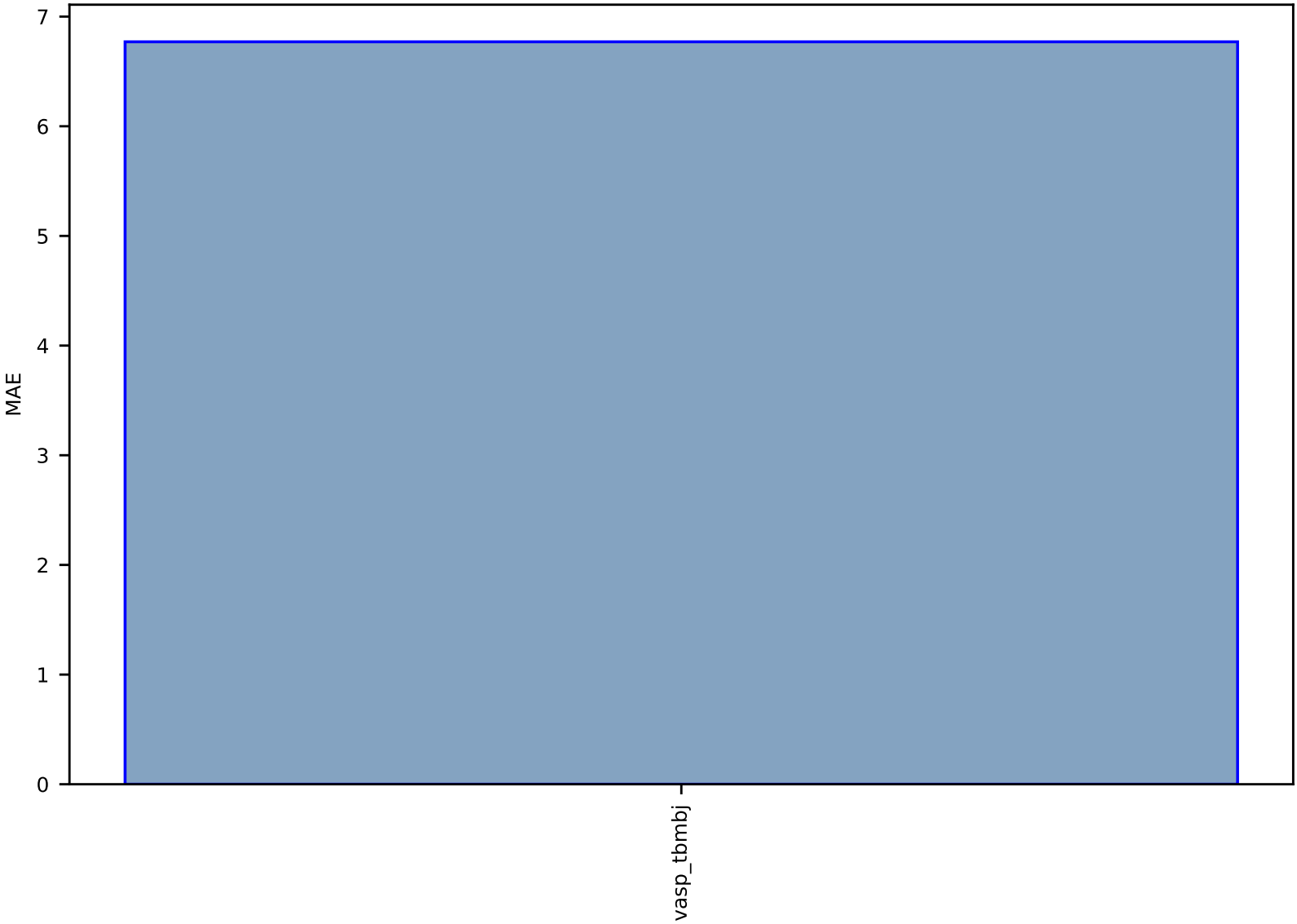


Fig S80:ES-SinglePropertyPrediction-bandgap_JVASP_1315_BaS-dft_3d-test-mae.csv.zip

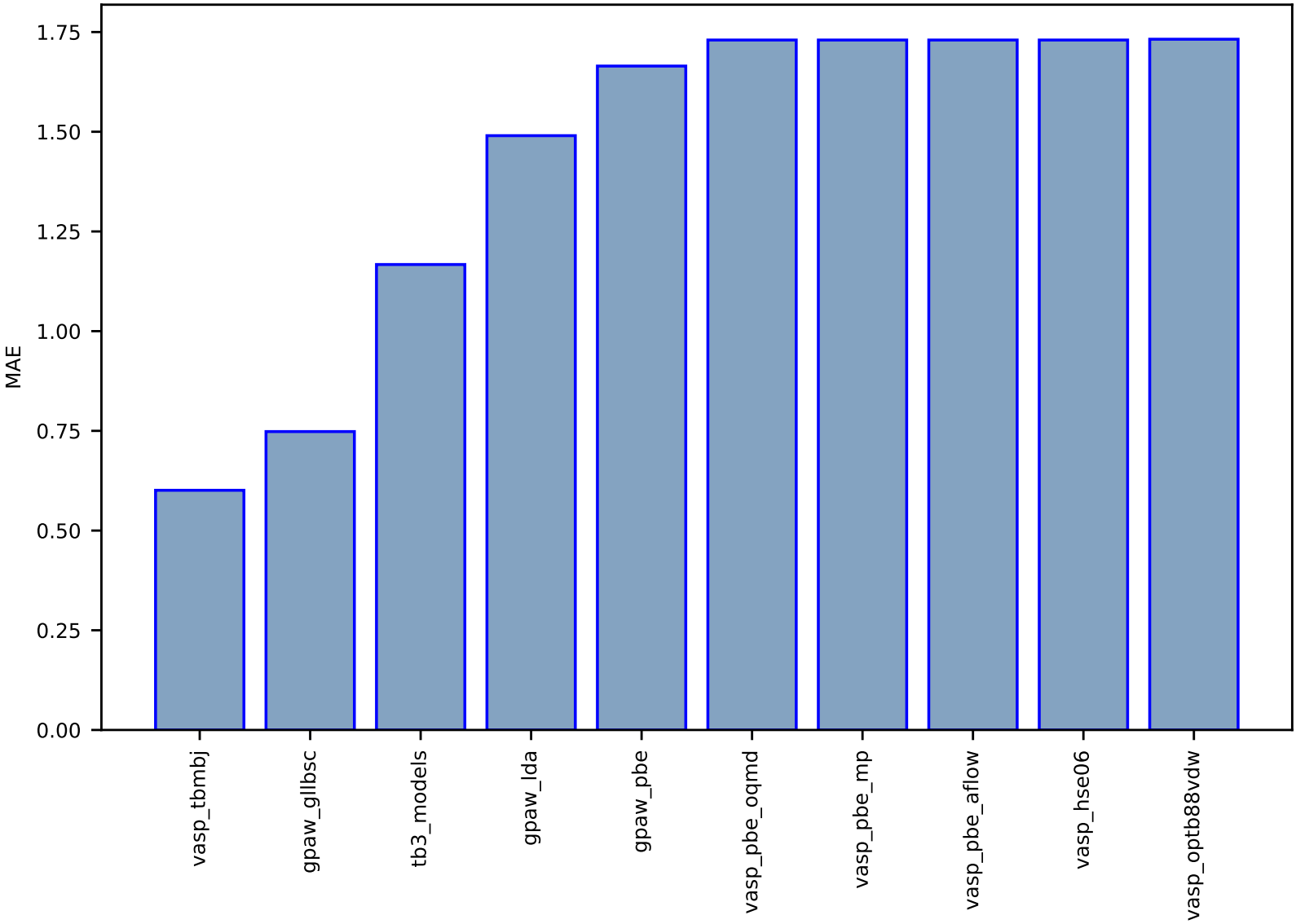


Fig S81:ES-Spectra-dielectric_function_JVASP_890_Ge-dft_3d-test-multimae.csv.zip

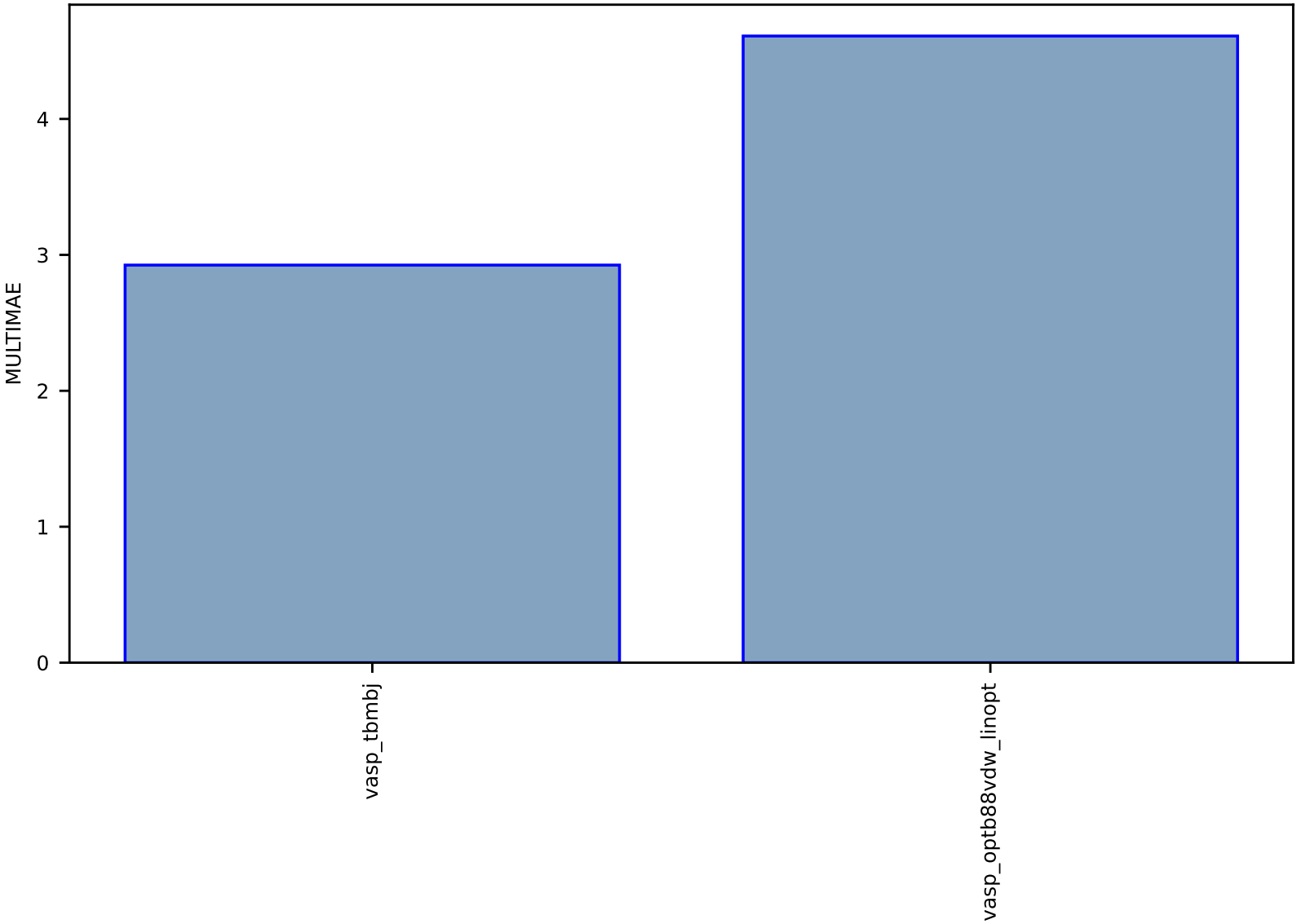


Fig S82:ES-SinglePropertyPrediction-bandgap_JVASP_97_InAs-dft_3d-test-mae.csv.zip

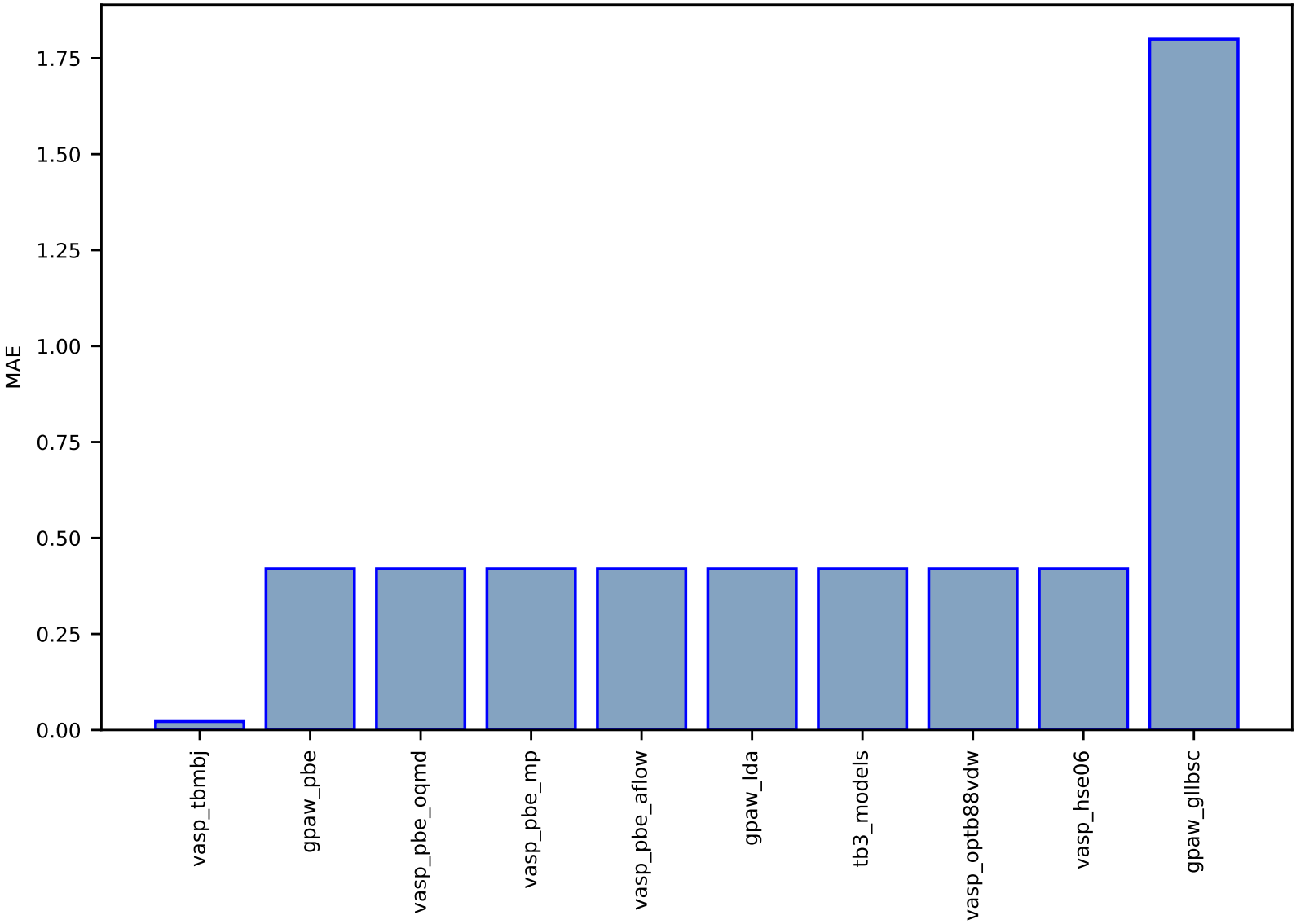


Fig S83:ES-SinglePropertyPrediction-bandgap_JVASP_116_MgO-dft_3d-test-mae.csv.zip

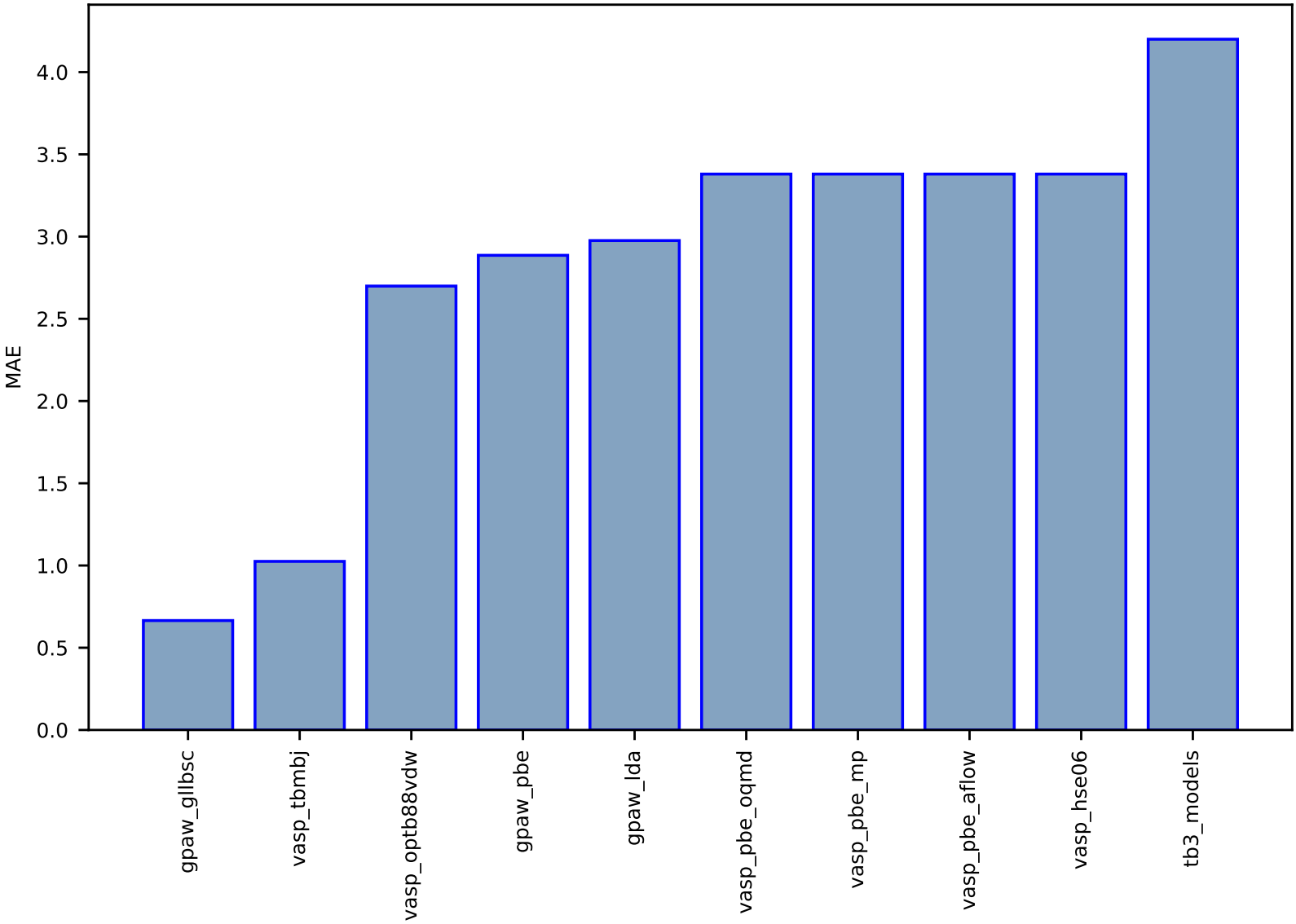


Fig S84:ES-SinglePropertyPrediction-bandgap_JVASP_1192_CdSe-dft_3d-test-mae.csv.zip

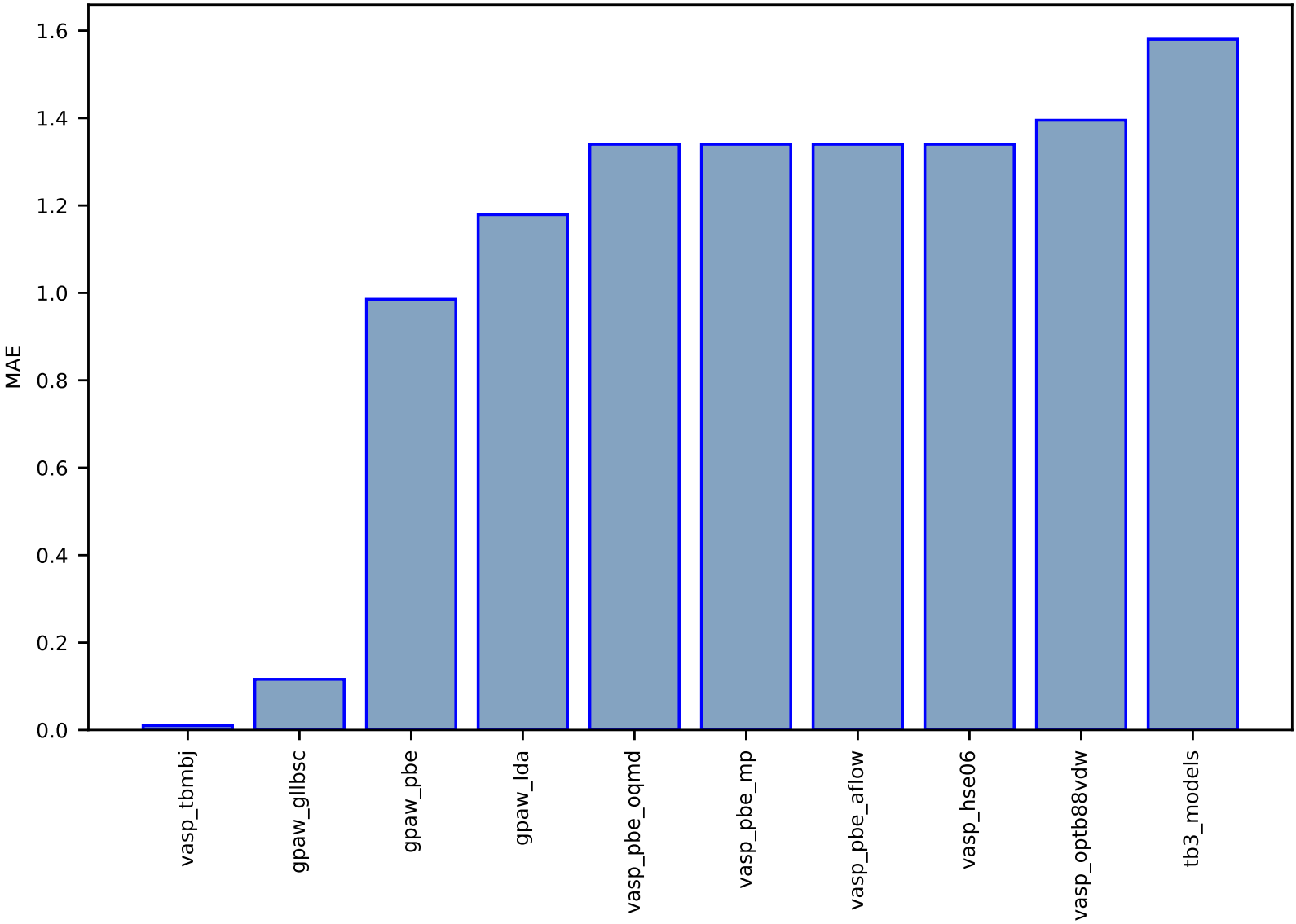


Fig S85:ES-SinglePropertyPrediction-bandgap_JVASP_9147_HfO2-dft_3d-test-mae.csv.zip

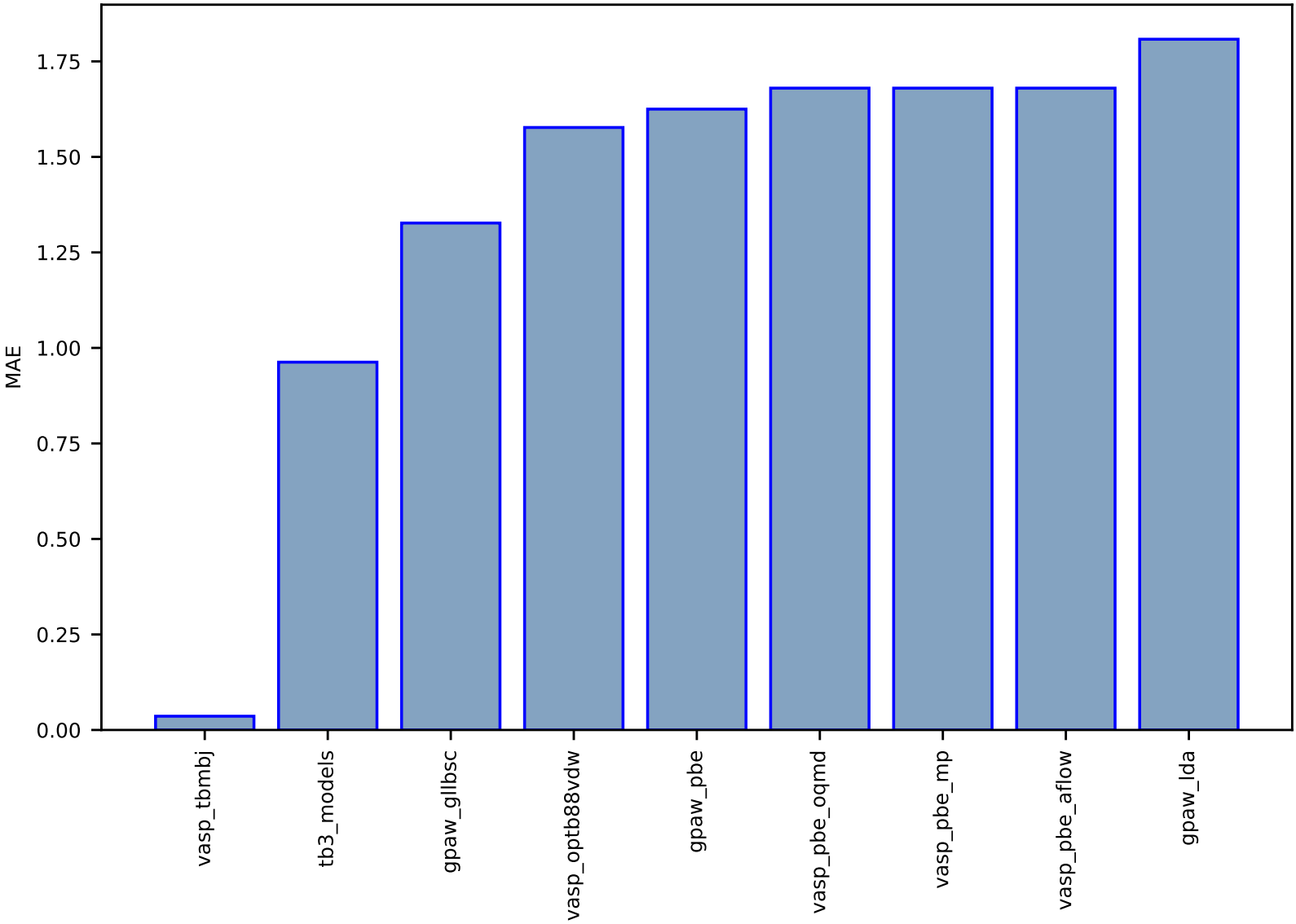


Fig S86:ES-SinglePropertyPrediction-bandgap_JVASP_7860_SnTe-dft_3d-test-mae.csv.zip

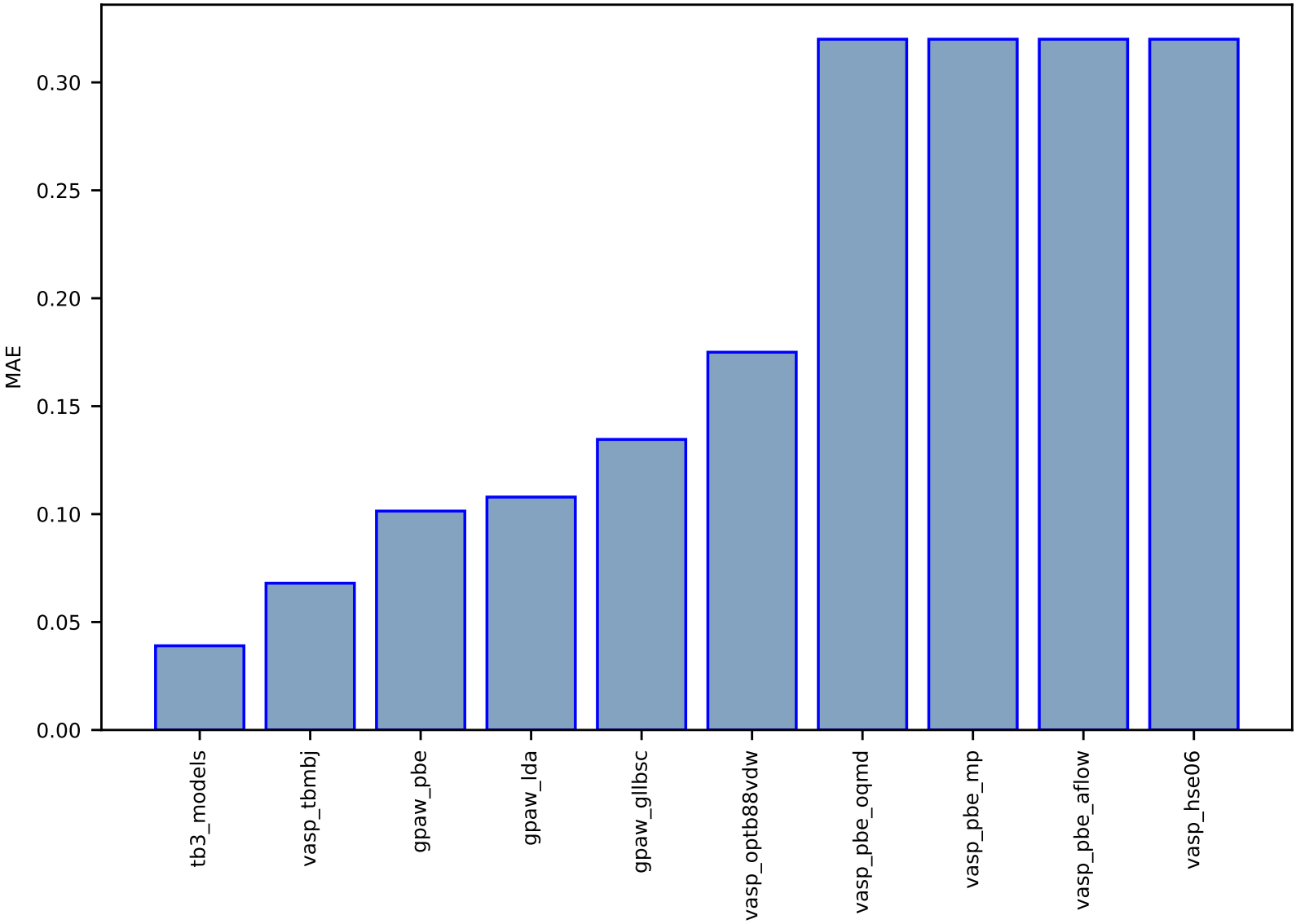


Fig S87:ES-Spectra-dielectric_function_JVASP_1174_GaAs-dft_3d-test-multimae.csv.zip

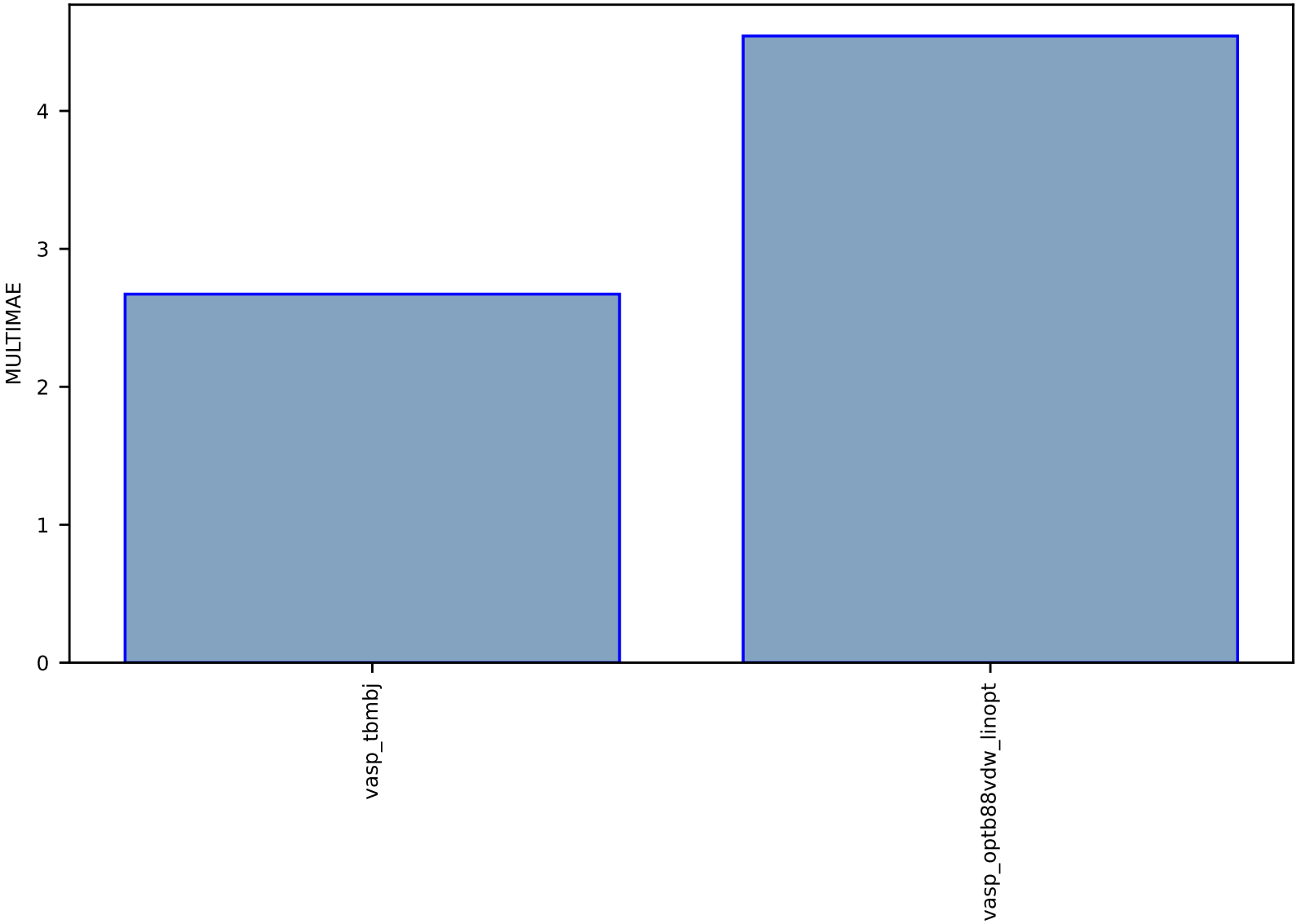


Fig S88:ES-SinglePropertyPrediction-bandgap_JVASP_1174_GaAs-dft_3d-test-mae.csv.zip

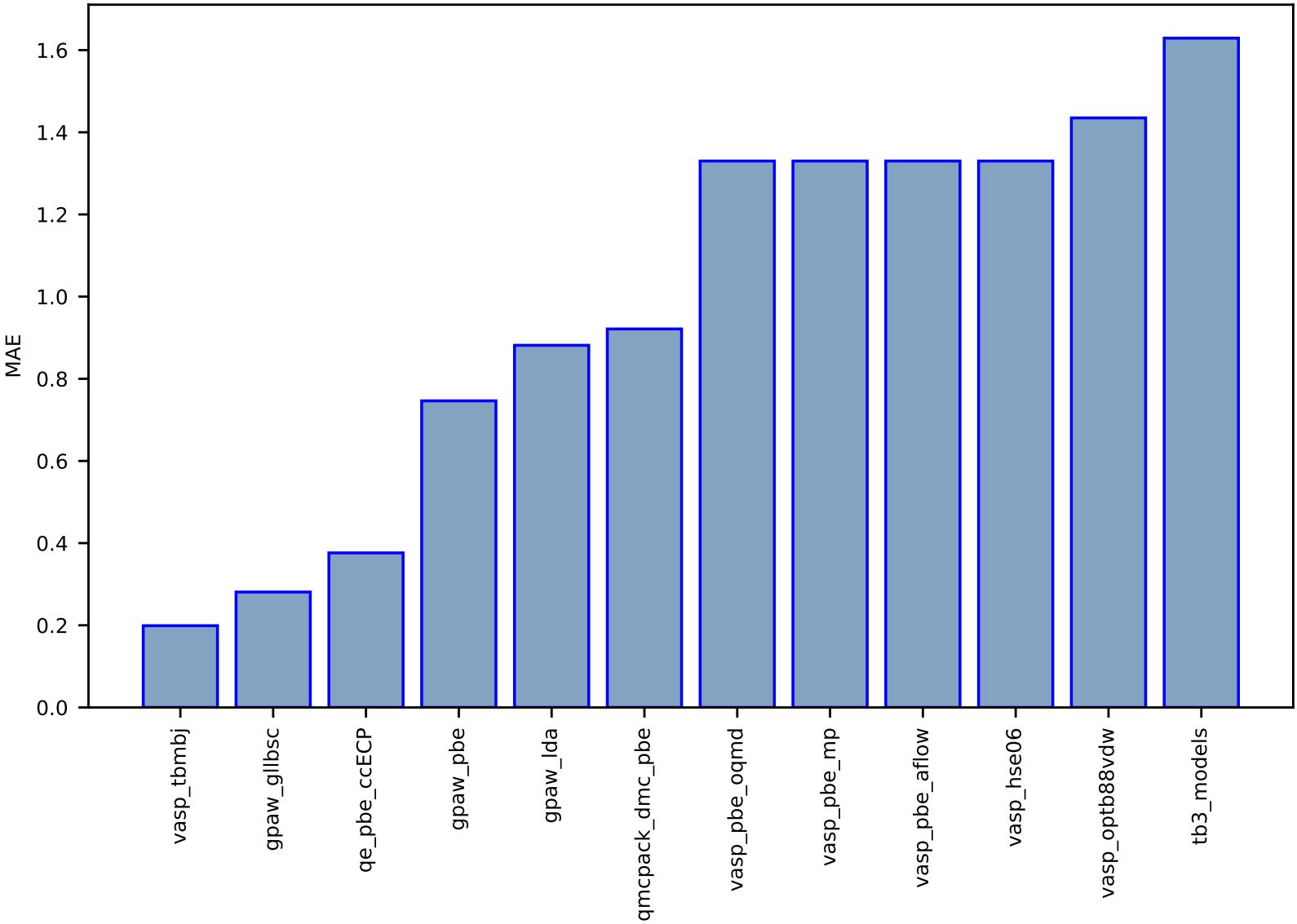


Fig S89:ES-SinglePropertyPrediction-bandgap_JVASP_1408_AlSb-dft_3d-test-mae.csv.zip

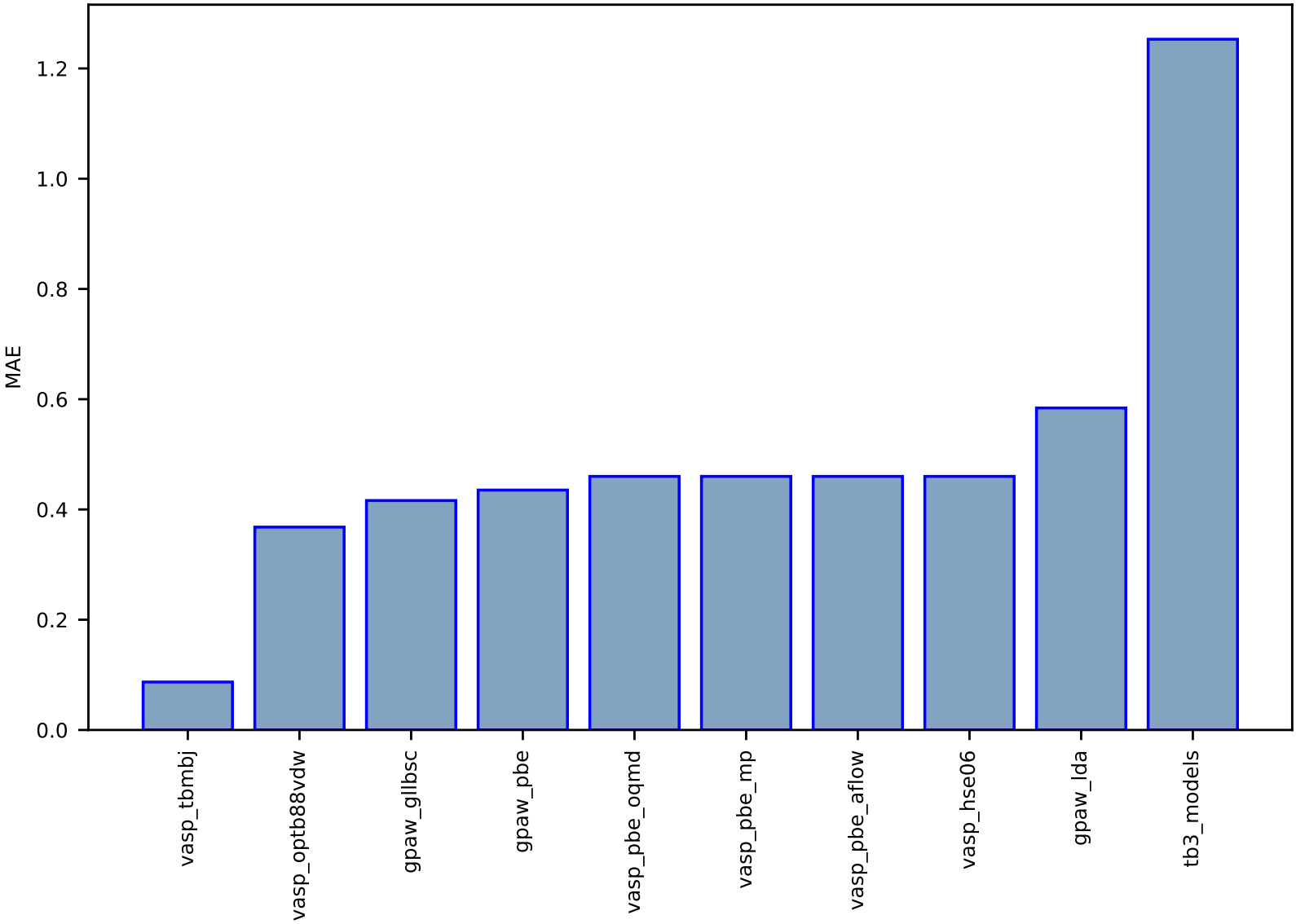


Fig S90:ES-SinglePropertyPrediction-bandgap_JVASP_7630_BAs-dft_3d-test-mae.csv.zip

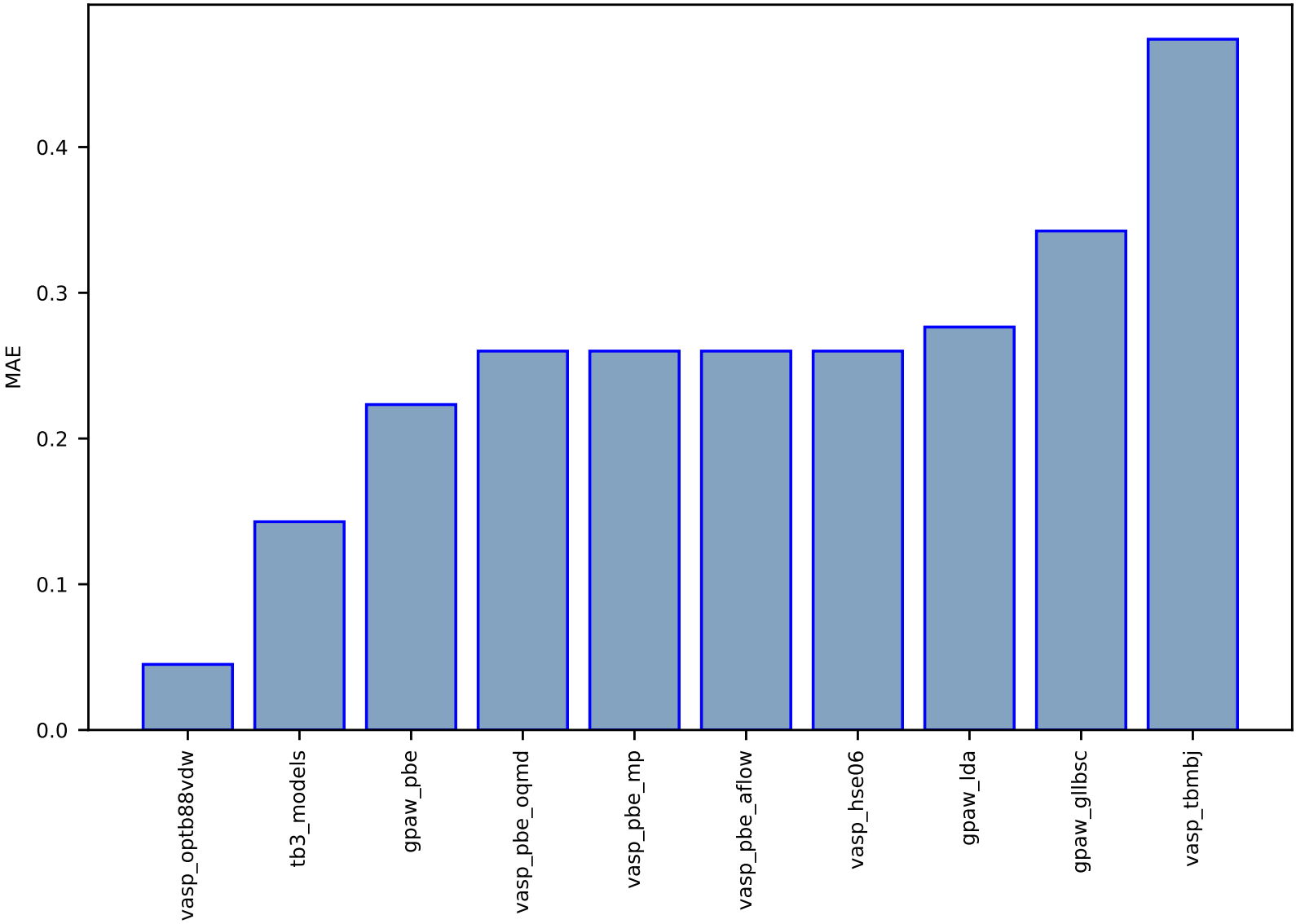


Fig S91:ES-SinglePropertyPrediction-bandgap_JVASP_5_TiO2-dft_3d-test-mae.csv.zip

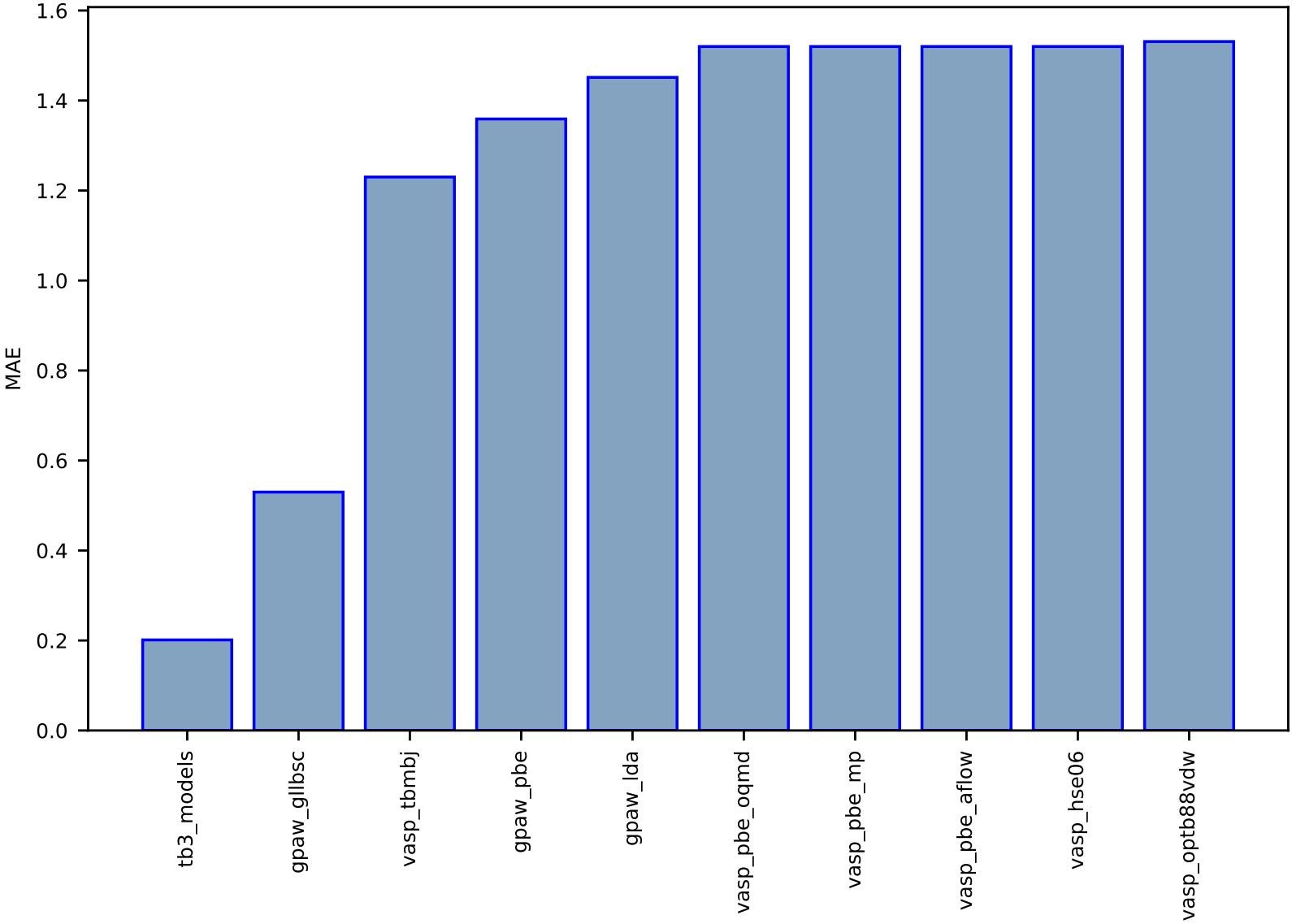


Fig S92:ES-SinglePropertyPrediction-bandgap_JVASP_113_ZrO2-dft_3d-test-mae.csv.zip

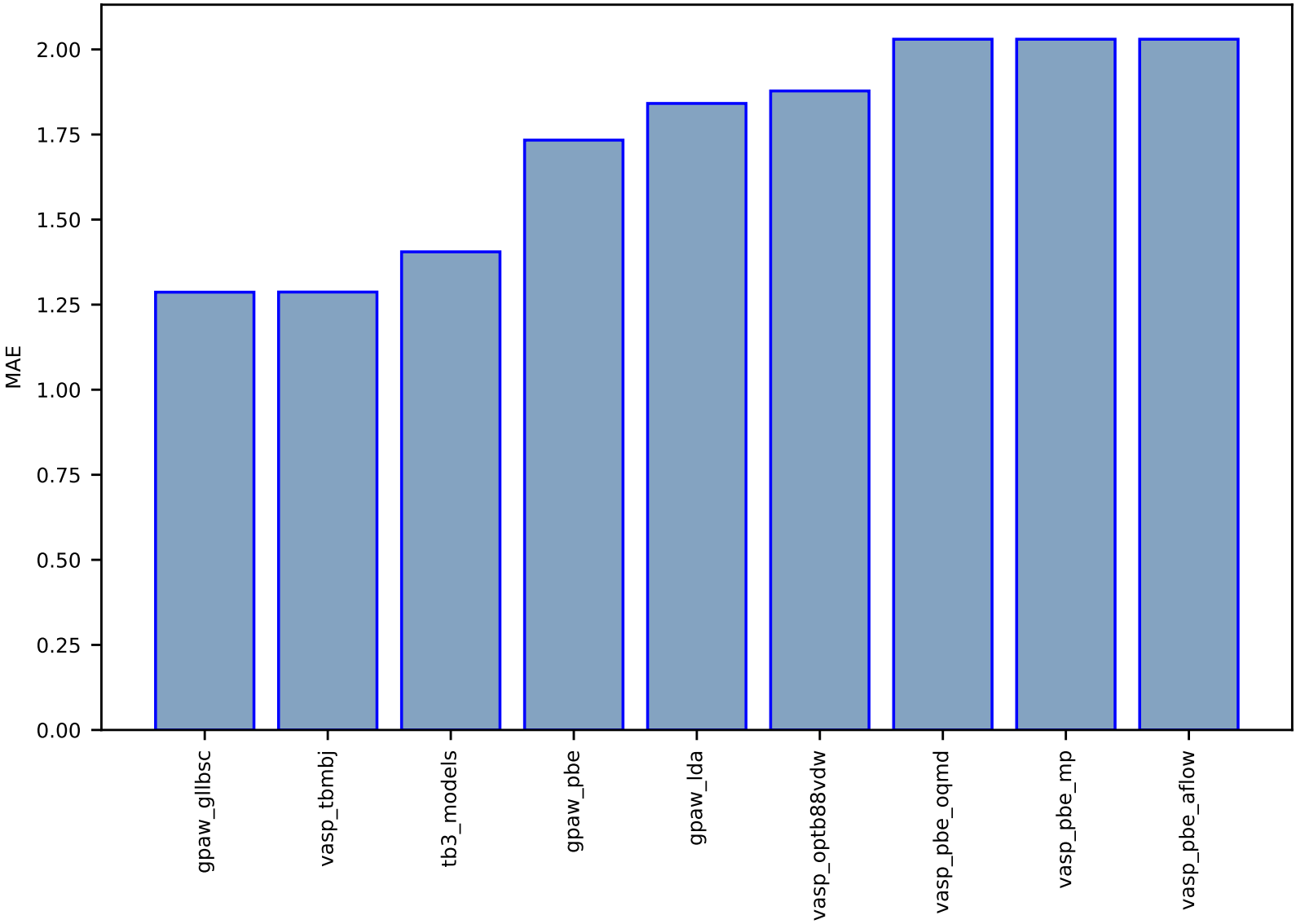


Fig S93:ES-SinglePropertyPrediction-bandgap_JVASP_890_Ge-dft_3d-test-mae.csv.zip

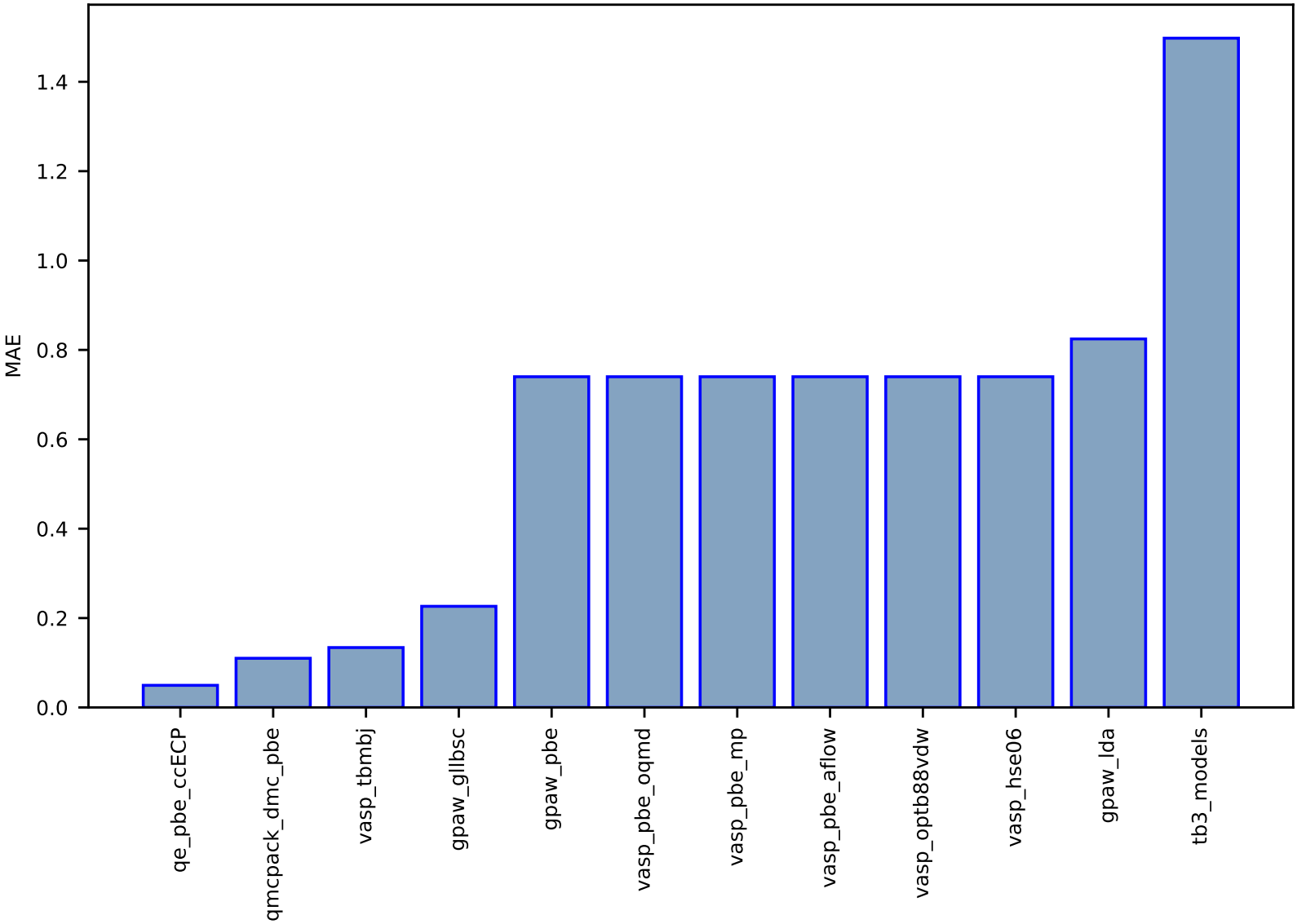


Fig S94:ES-SinglePropertyPrediction-slme_JVASP_1174_GaAs-dft_3d-test-mae.csv.zip

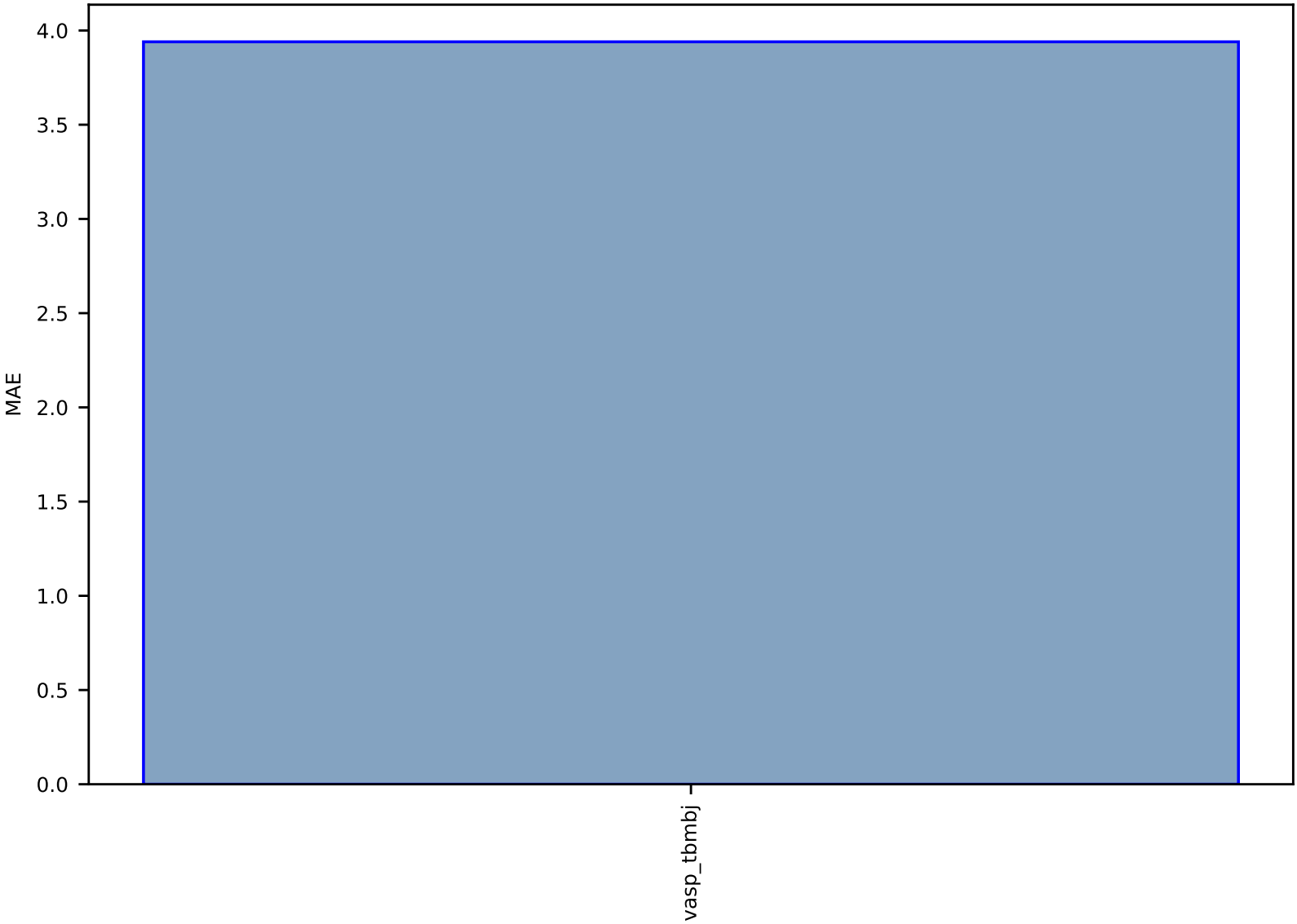


Fig S95:ES-SinglePropertyPrediction-bandgap_JVASP_1327_AlP-dft_3d-test-mae.csv.zip

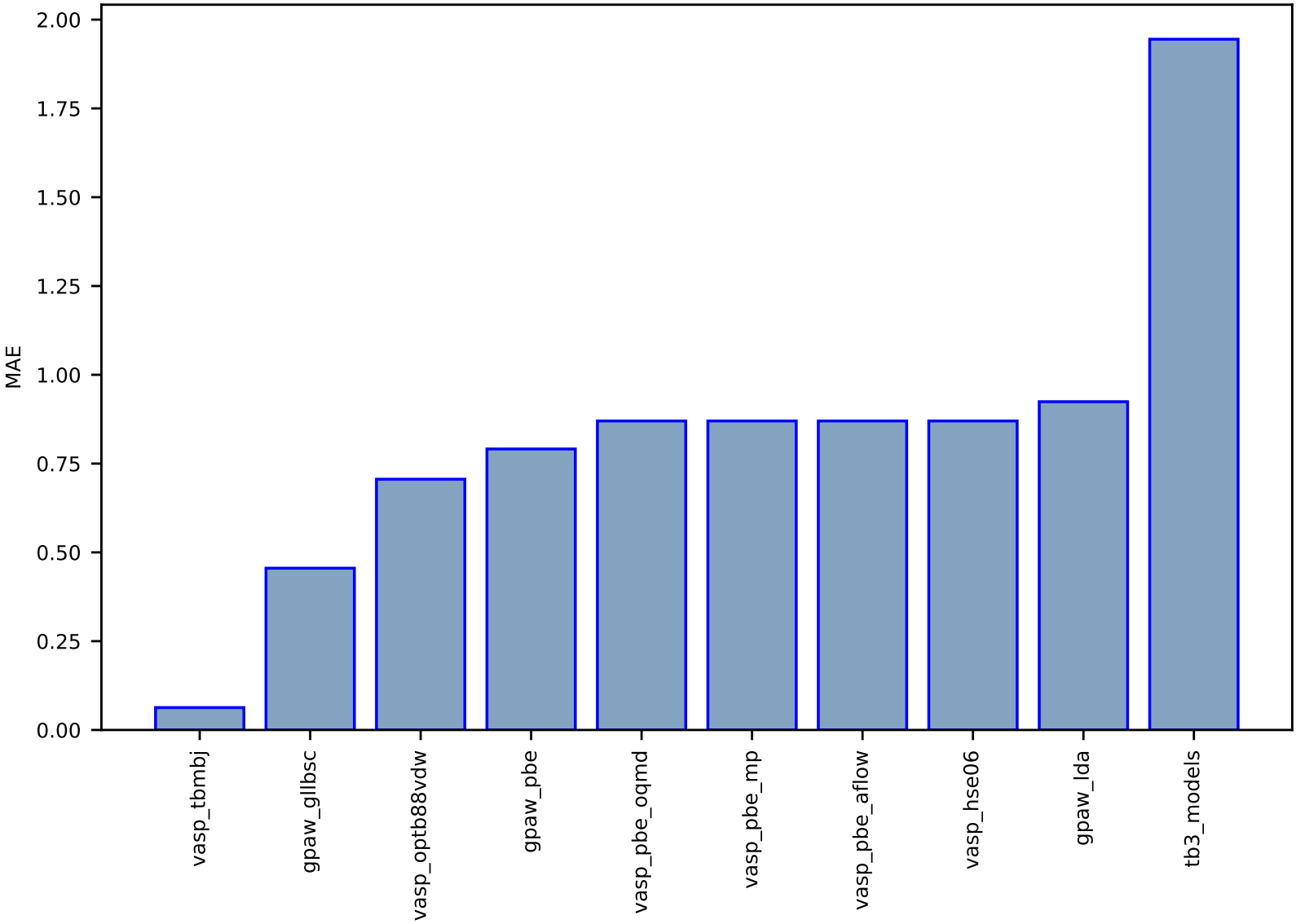


Fig S96:ES-SinglePropertyPrediction-bandgap_JVASP_104_TiO2-dft_3d-test-mae.csv.zip

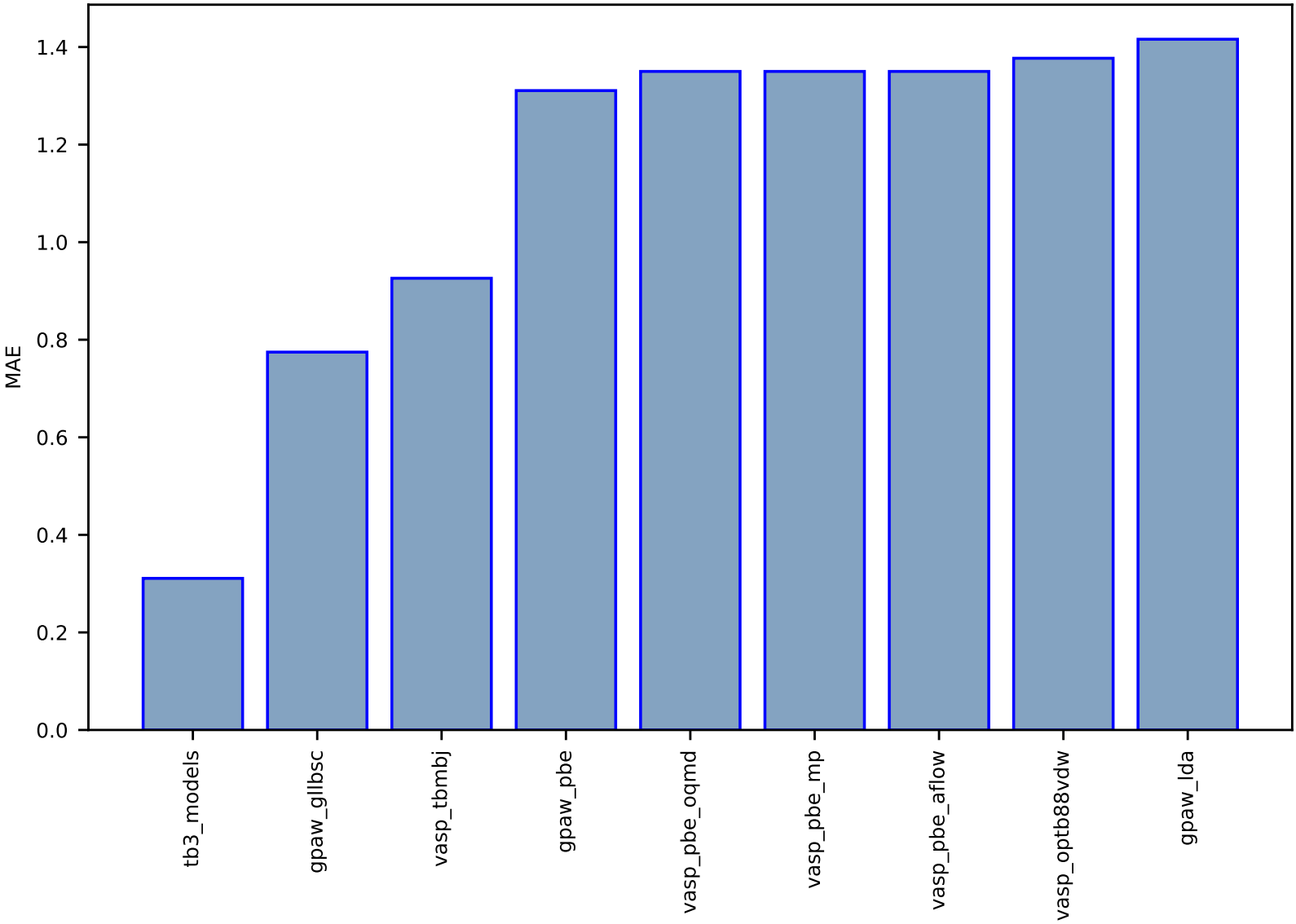


Fig S97:ES-SinglePropertyPrediction-bandgap_JVASP_1312_BP-dft_3d-test-mae.csv.zip

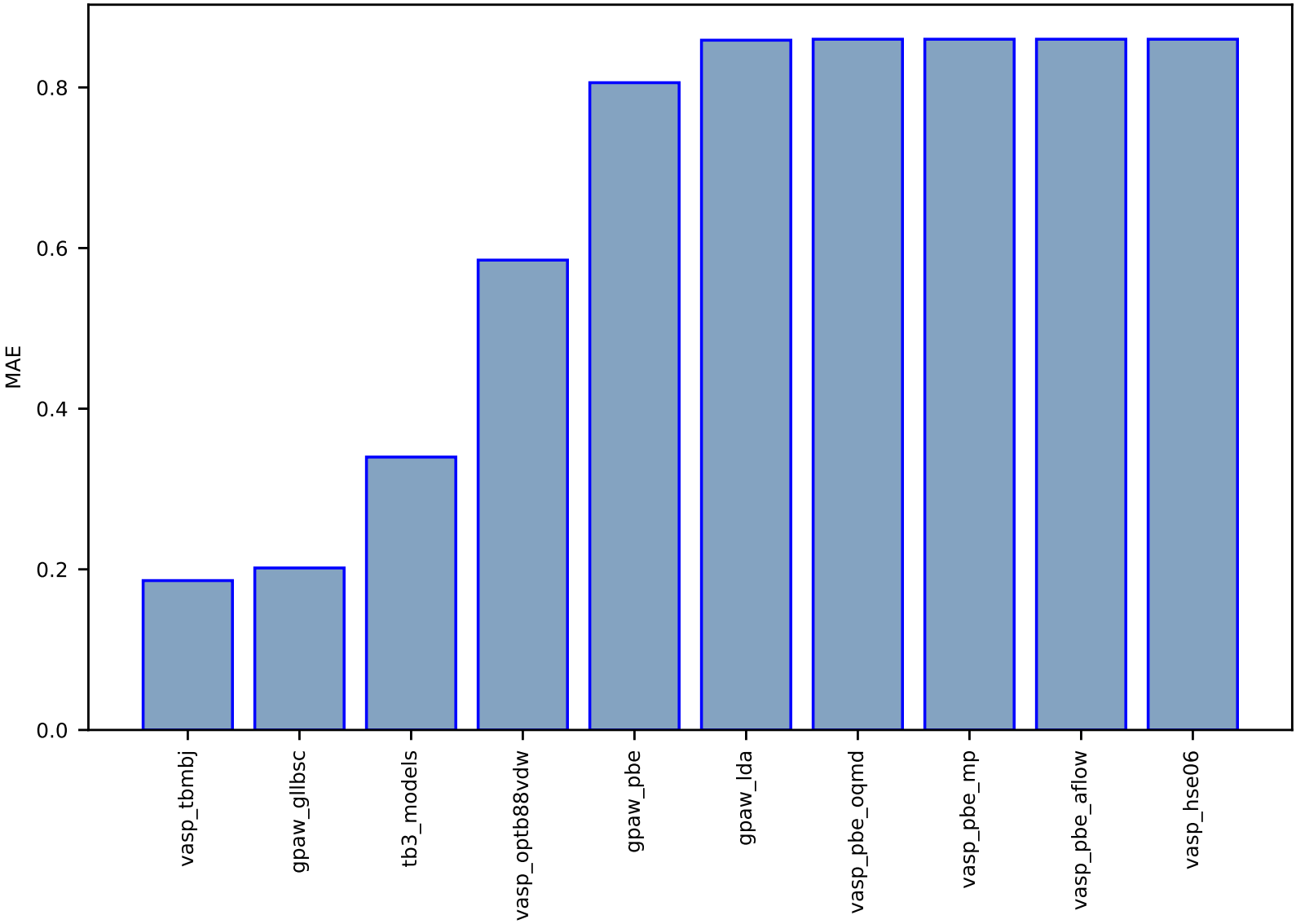


Fig S98:ES-SinglePropertyPrediction-bandgap_JVASP_1145_KCl-dft_3d-test-mae.csv.zip

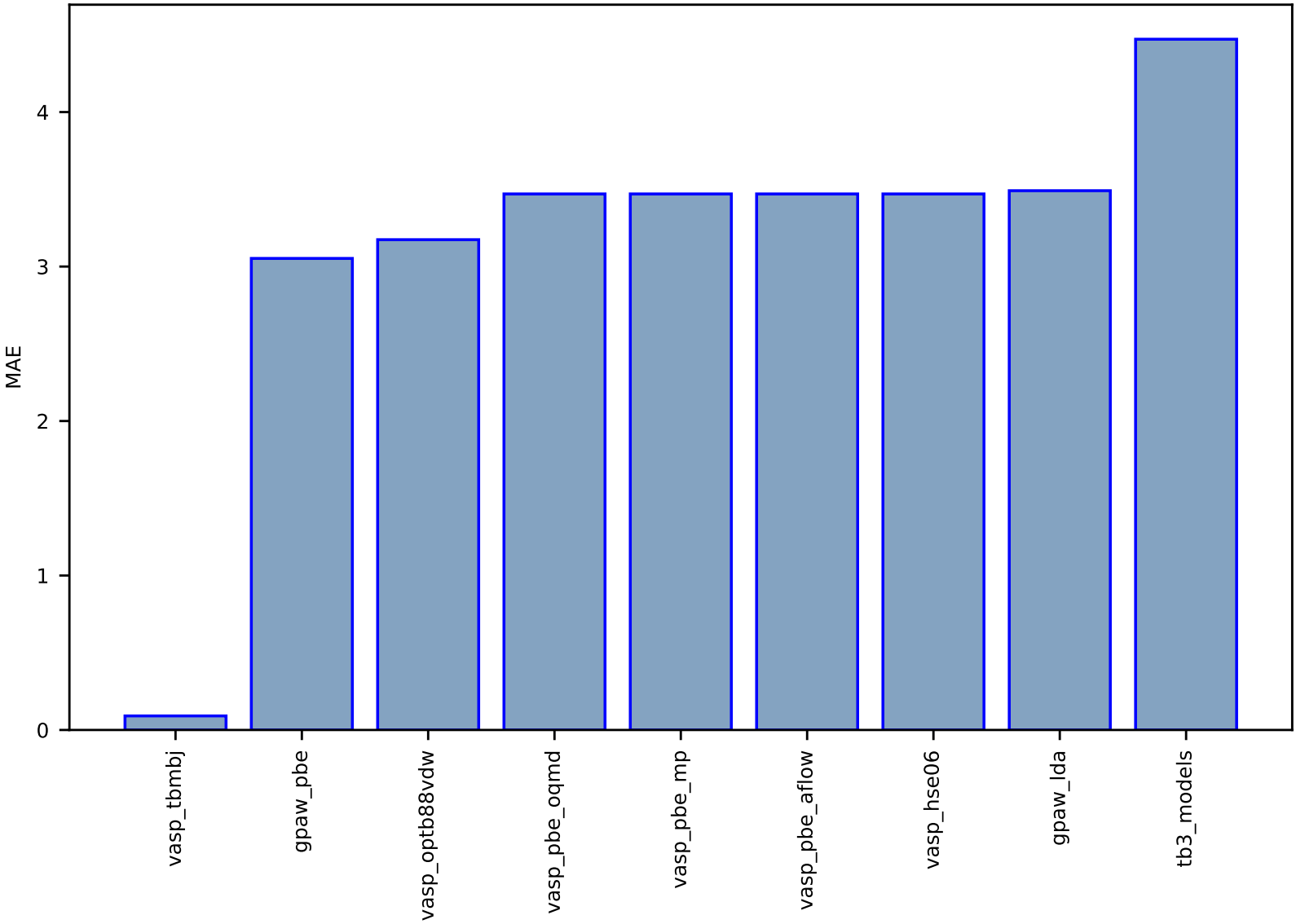


Fig S99:ES-SinglePropertyPrediction-bandgap_JVASP_17_BN-dft_3d-test-mae.csv.zip

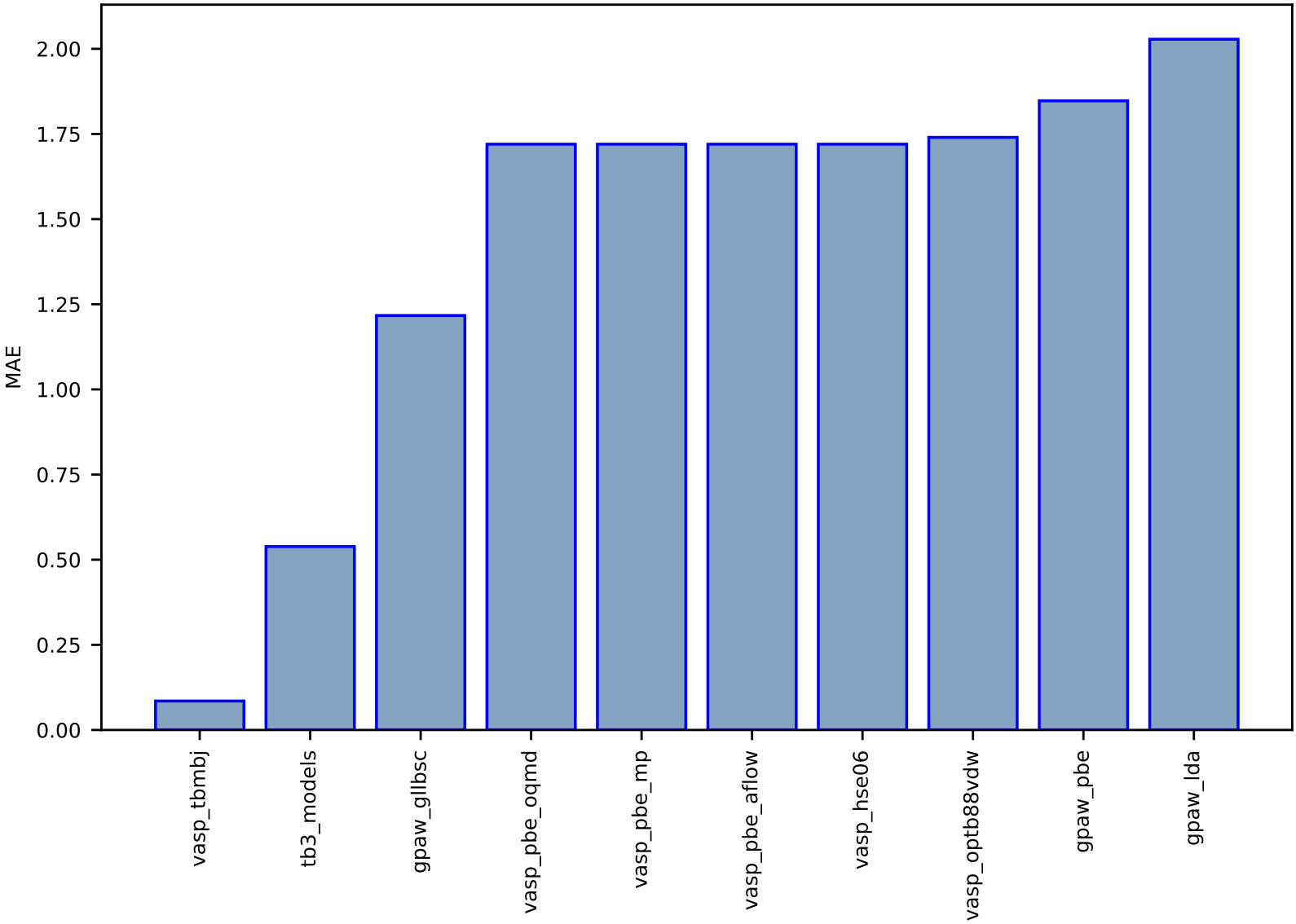


Fig S100:ES-SinglePropertyPrediction-bandgap_JVASP_30_GaN-dft_3d-test-mae.csv.zip

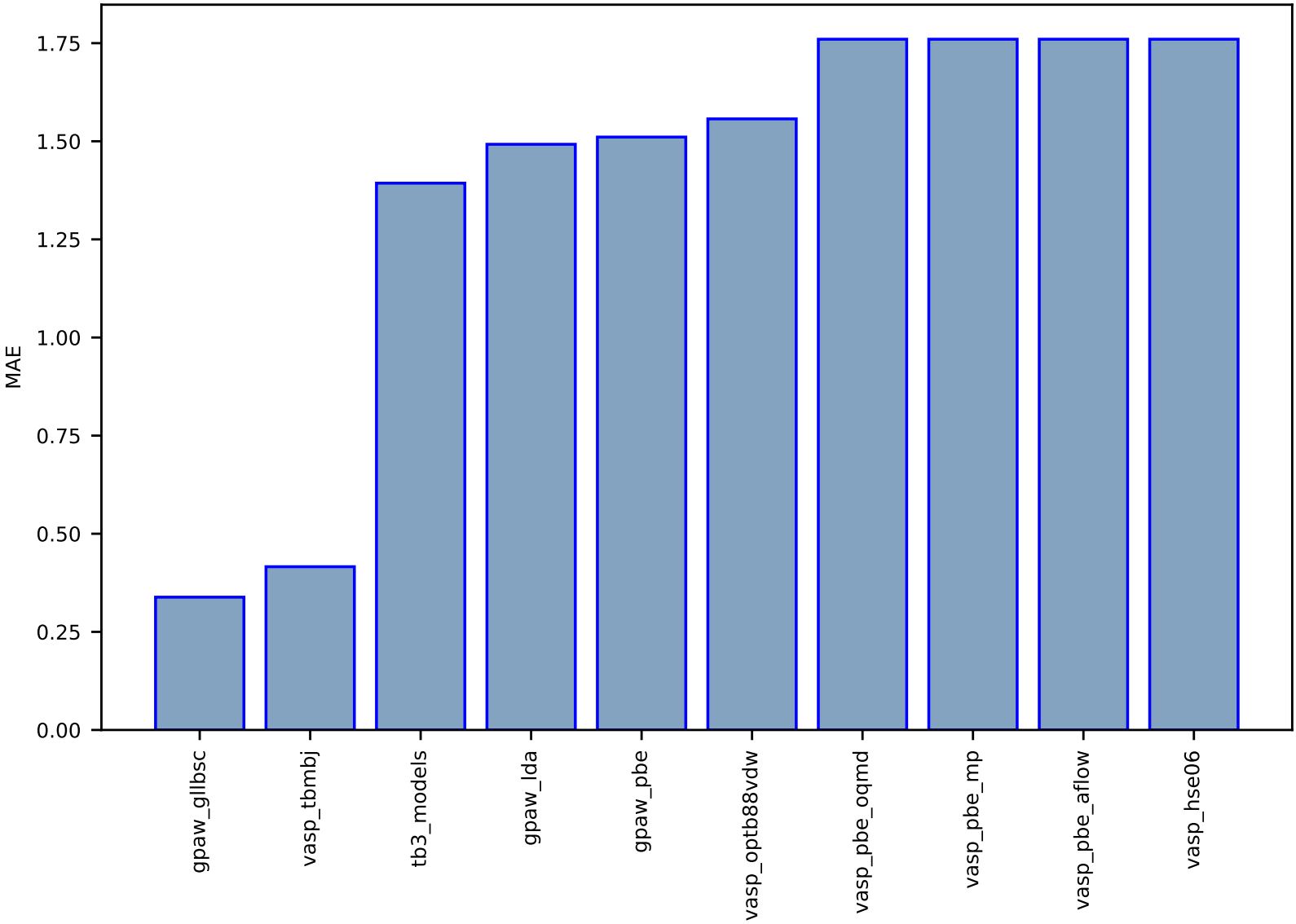


Fig S101:ES-SinglePropertyPrediction-bandgap_JVASP_32_Al2O3-dft_3d-test-mae.csv.zip

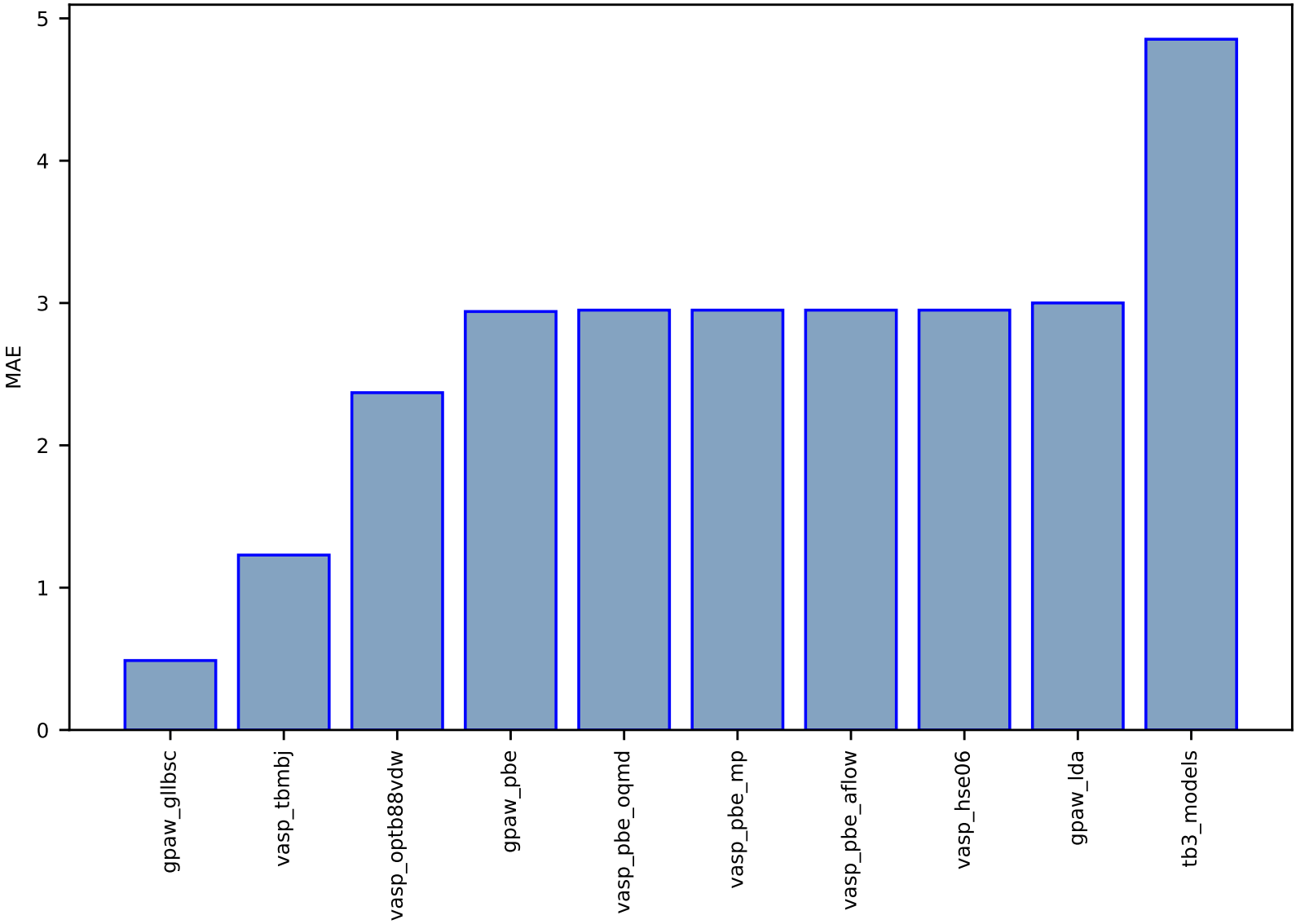


Fig S102:ES-SinglePropertyPrediction-bandgap_JVASP_1294_BaSe-dft_3d-test-mae.csv.zip

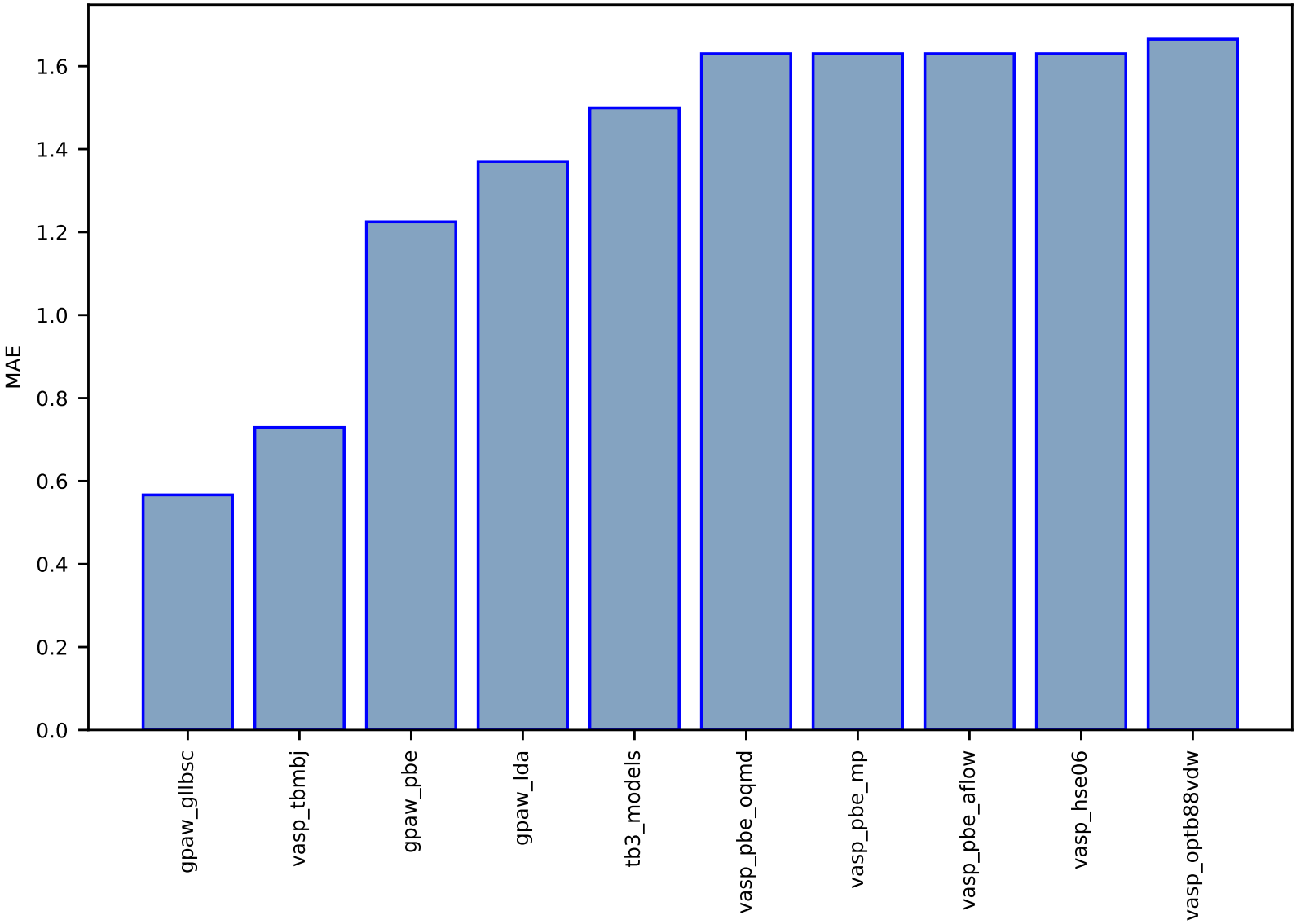


Fig S103:ES-SinglePropertyPrediction-bandgap_JVASP_8003_CdS-dft_3d-test-mae.csv.zip

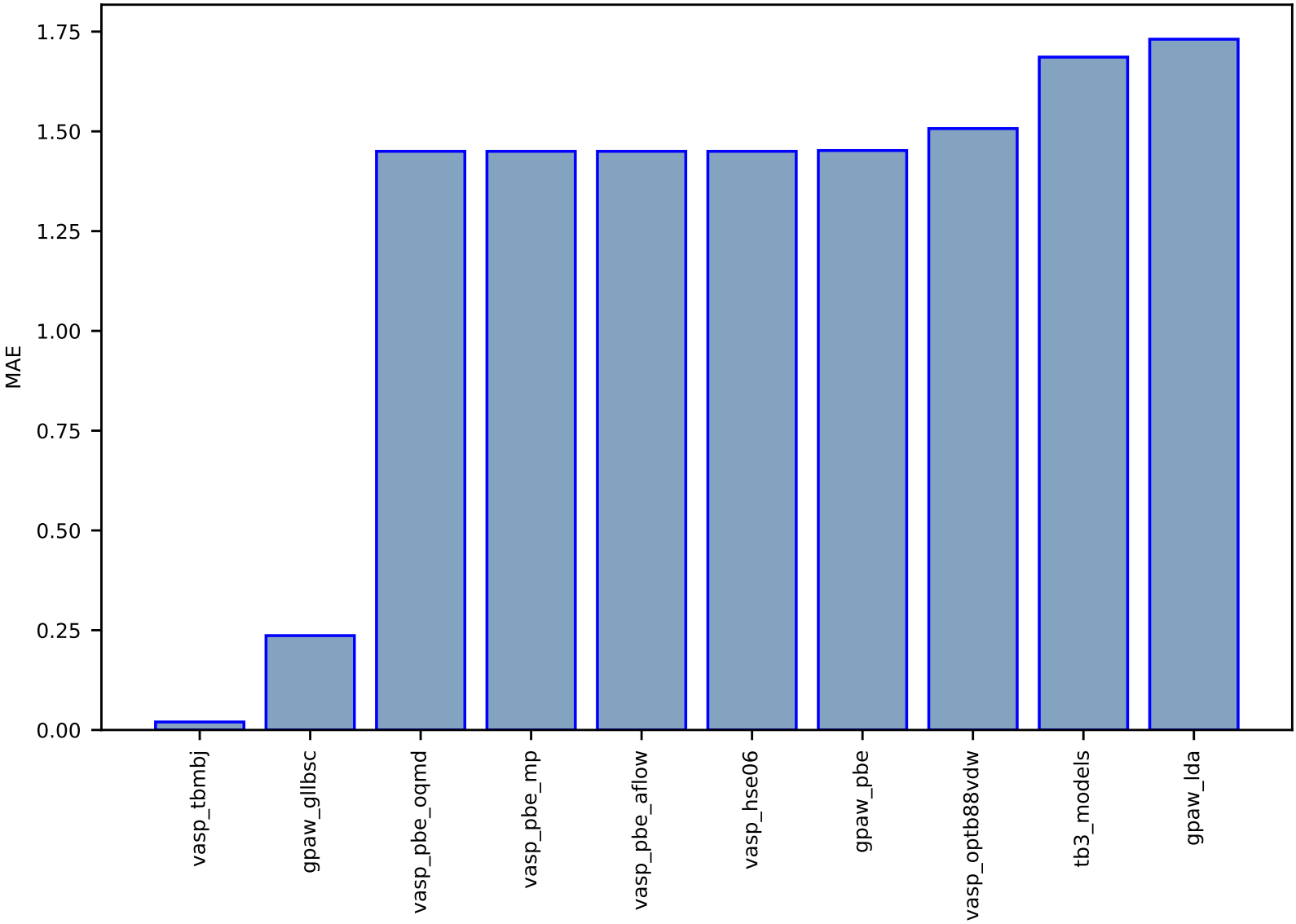


Fig S104:ES-SinglePropertyPrediction-bandgap_JVASP_1180_InN-dft_3d-test-mae.csv.zip

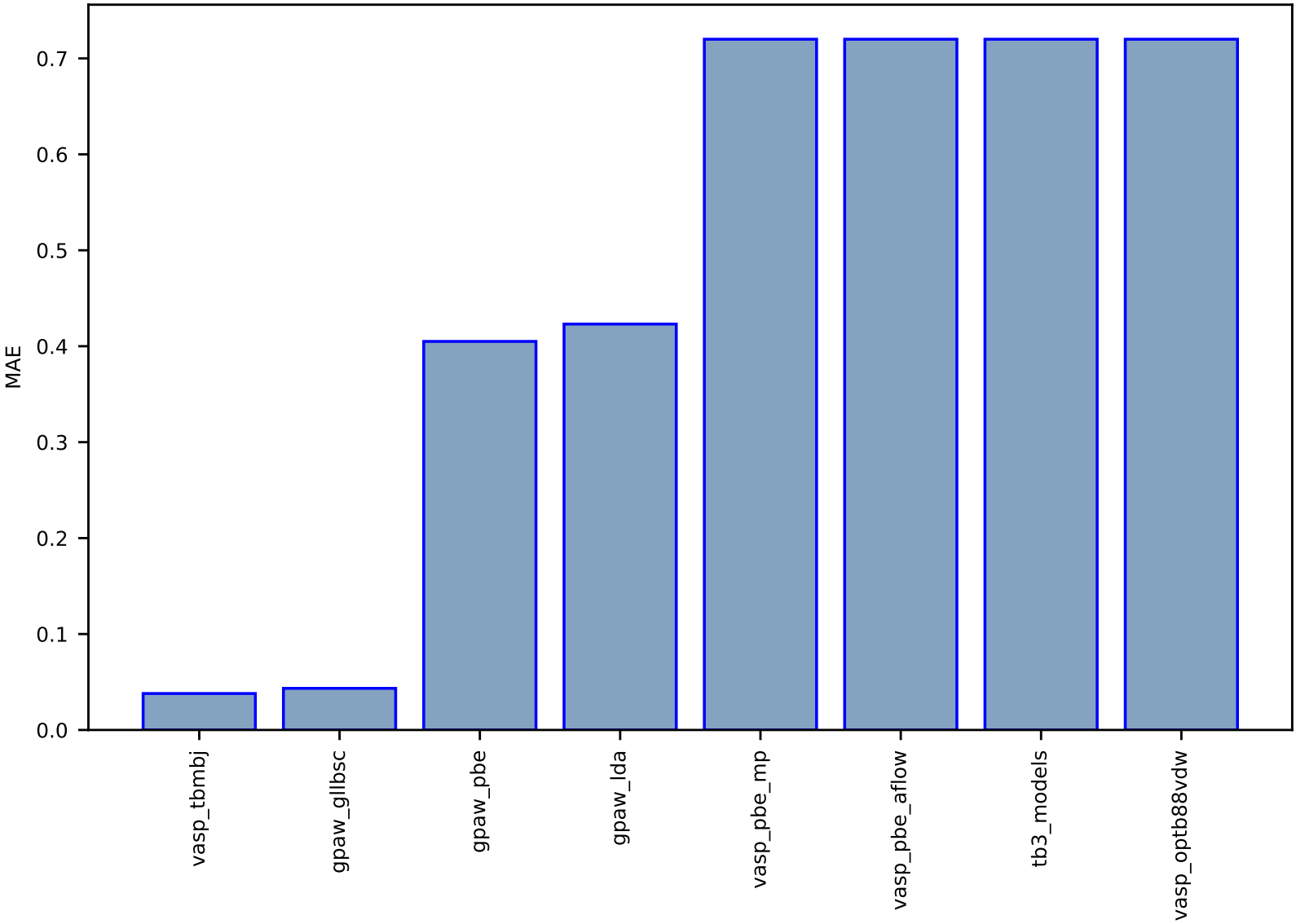


Fig S105:ES-Spectra-dielectric_function-dft_3d-test-multimae.csv.zip

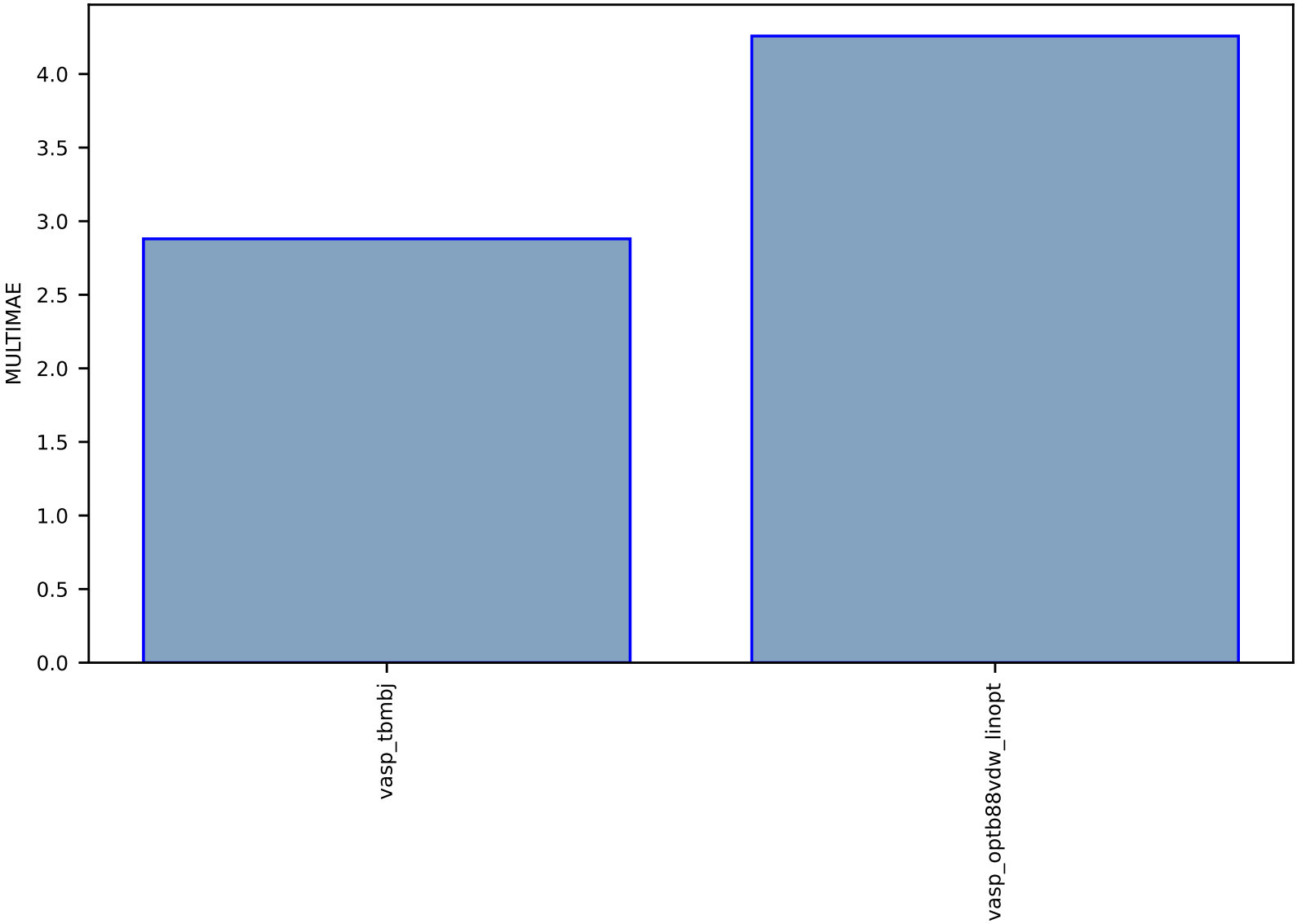


Fig S106:ES-SinglePropertyPrediction-bandgap_JVASP_8566_AgI-dft_3d-test-mae.csv.zip

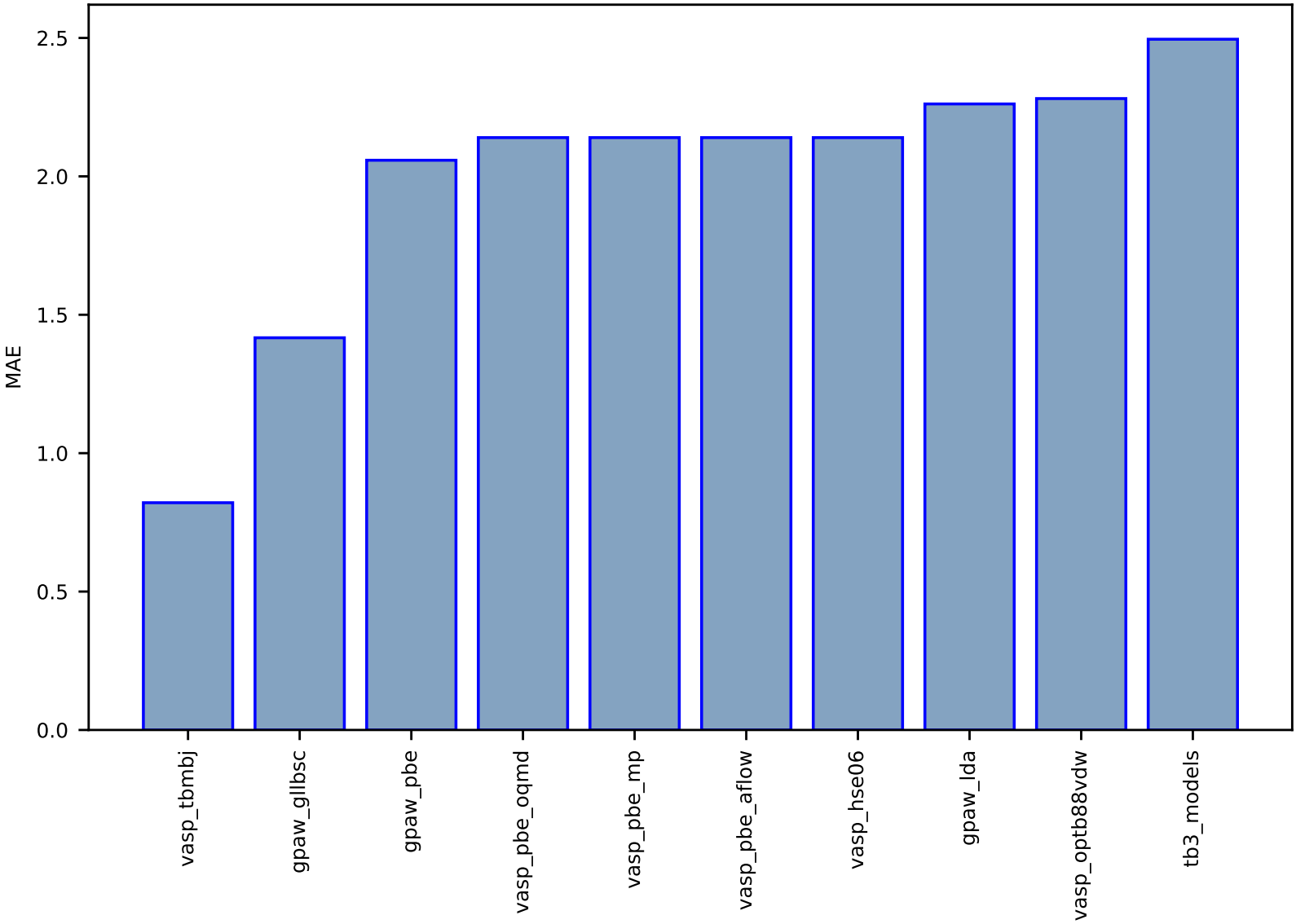


Fig S107:ES-SinglePropertyPrediction-bandgap_JVASP_1267_BaTe-dft_3d-test-mae.csv.zip

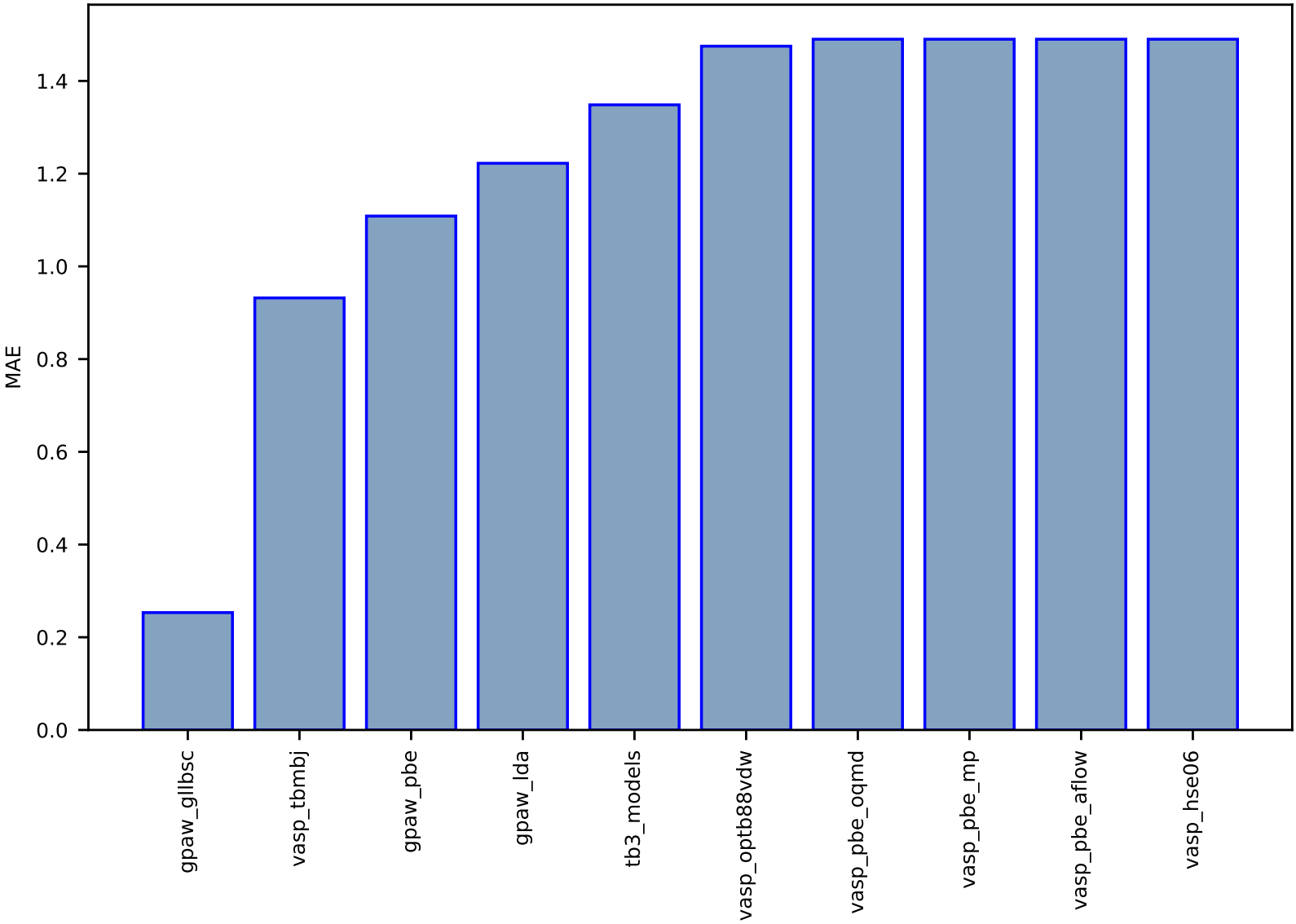


Fig S108:ES-SinglePropertyPrediction-bandgap_JVASP_95_CdS-dft_3d-test-mae.csv.zip

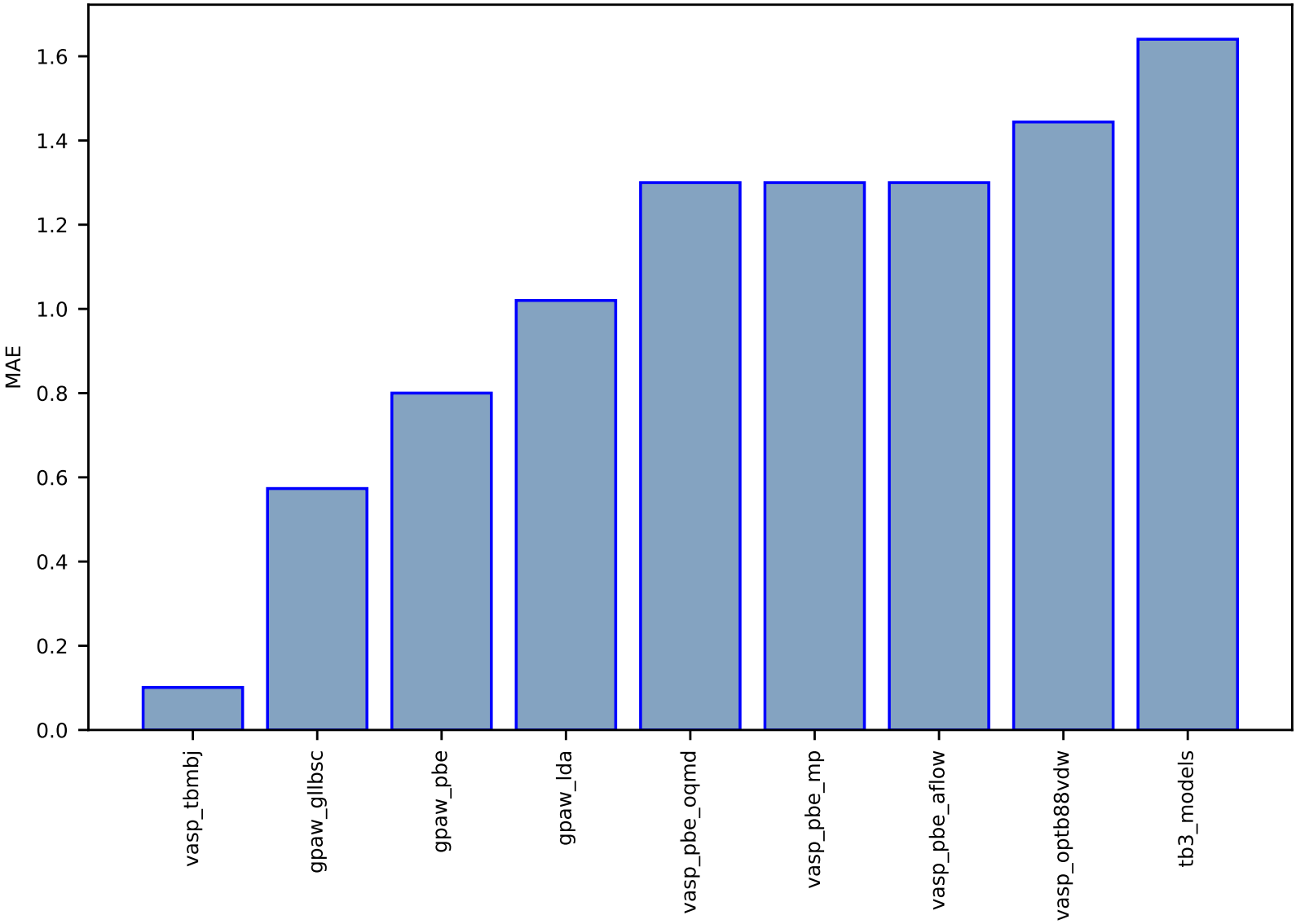


Fig S109:ES-SinglePropertyPrediction-bandgap_JVASP_8583_AgBr-dft_3d-test-mae.csv.zip

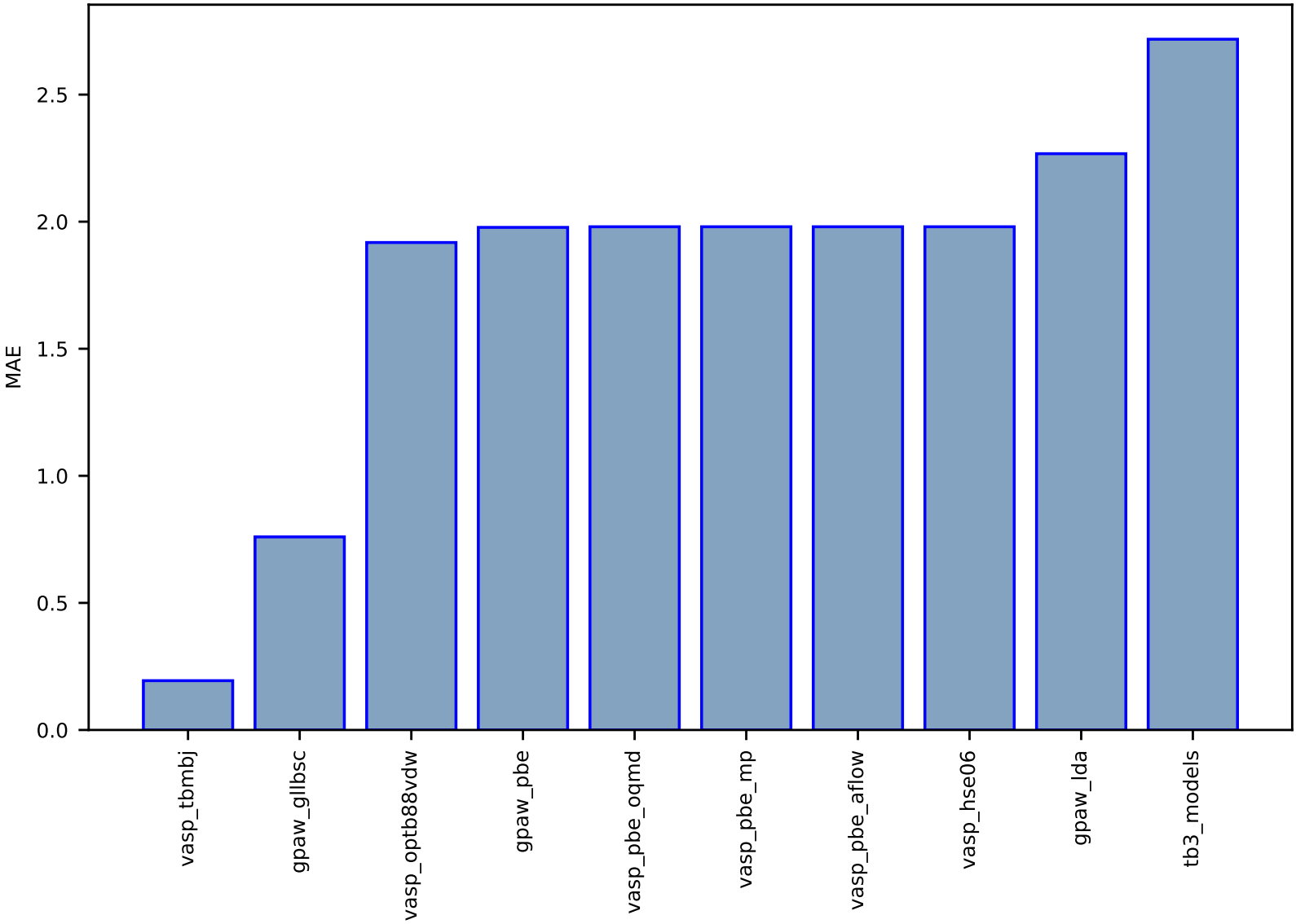


Fig S110:FF-SinglePropertyPrediction-viscosity-lj_2d_liquid-test-mae.csv.zip

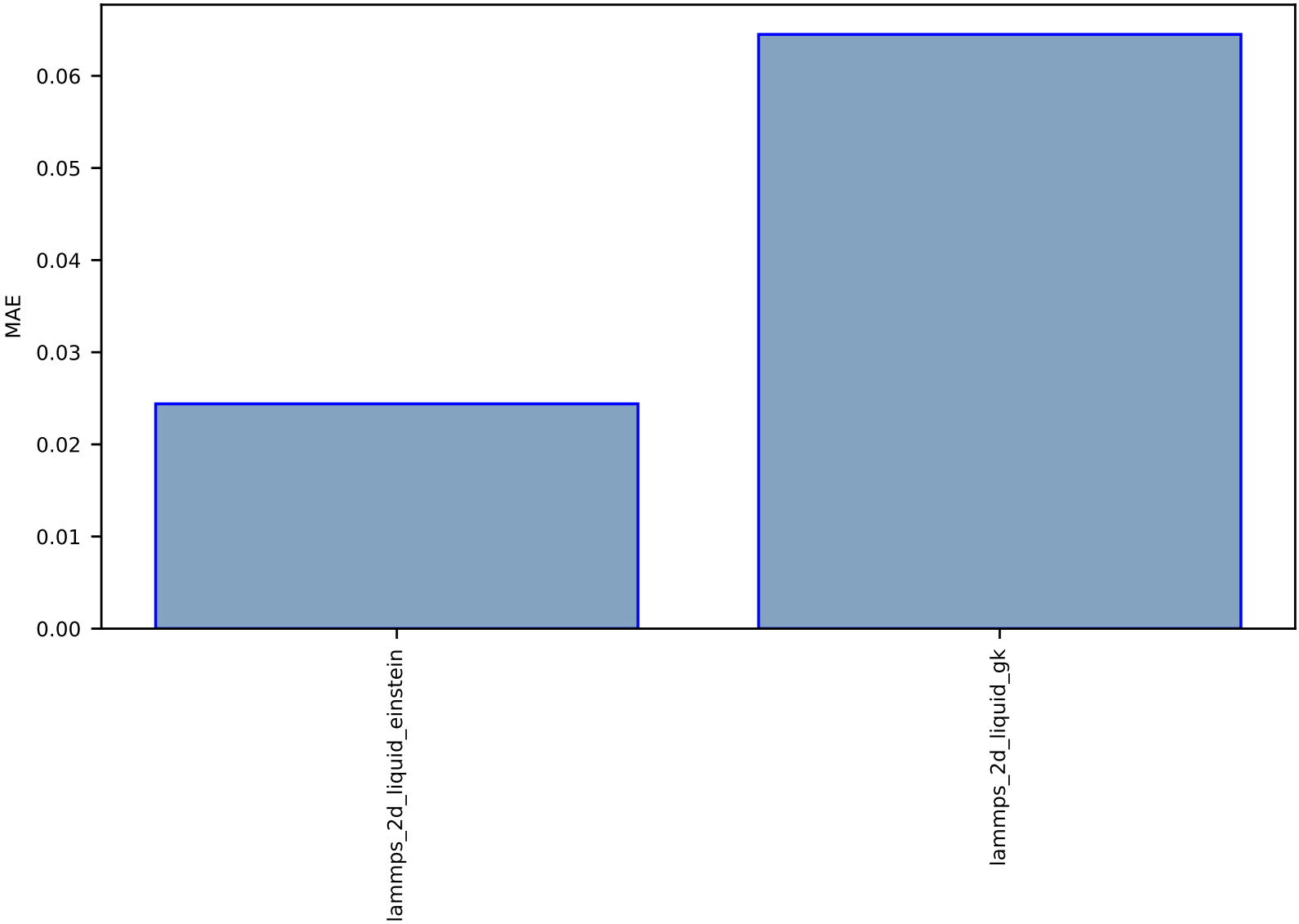


Fig S111:AI-SinglePropertyPrediction-slme-dft_3d-test-mae.csv.zip

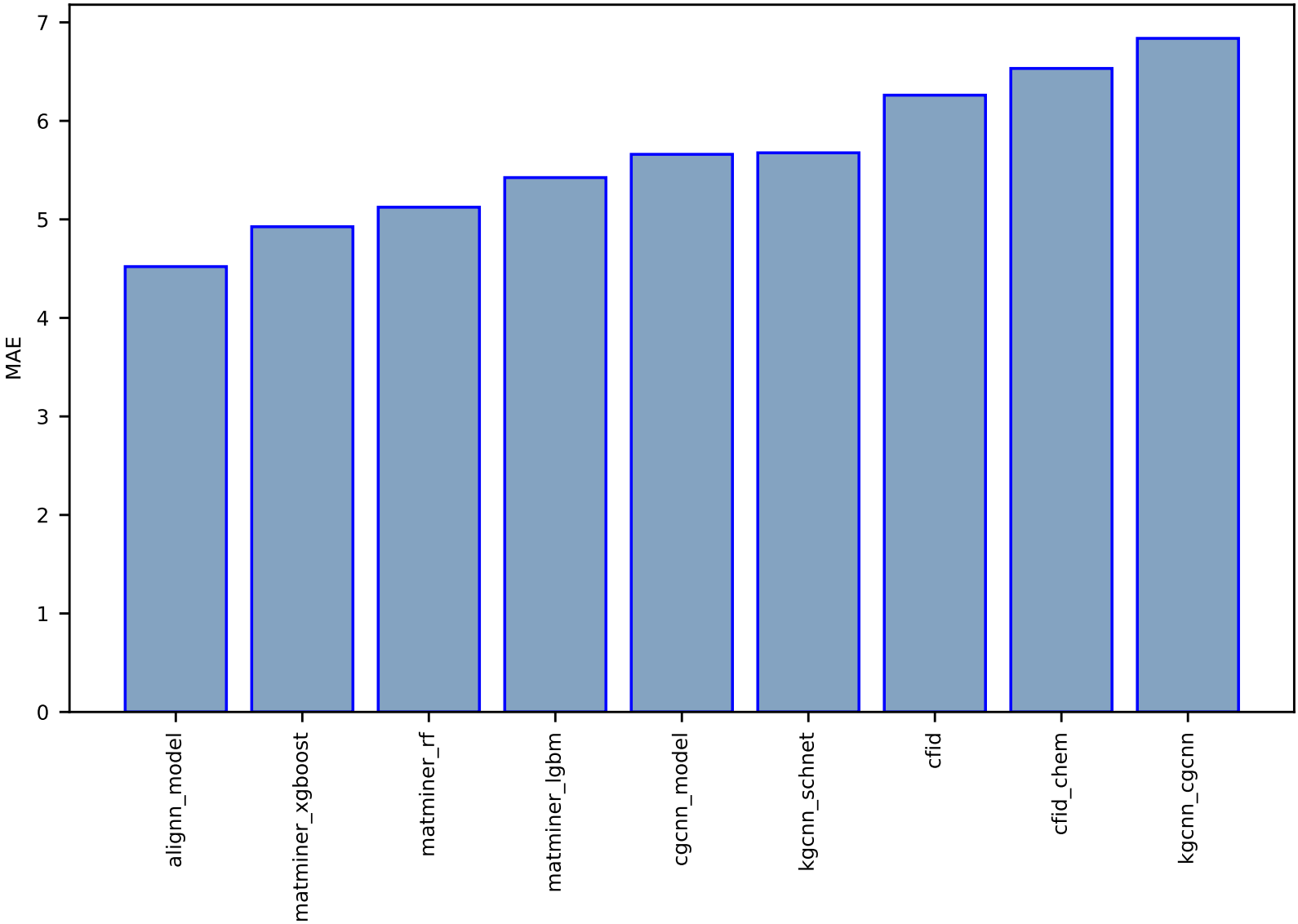


Fig S112:AI-SinglePropertyPrediction-mbj_bandgap-dft_3d-test-mae.csv.zip

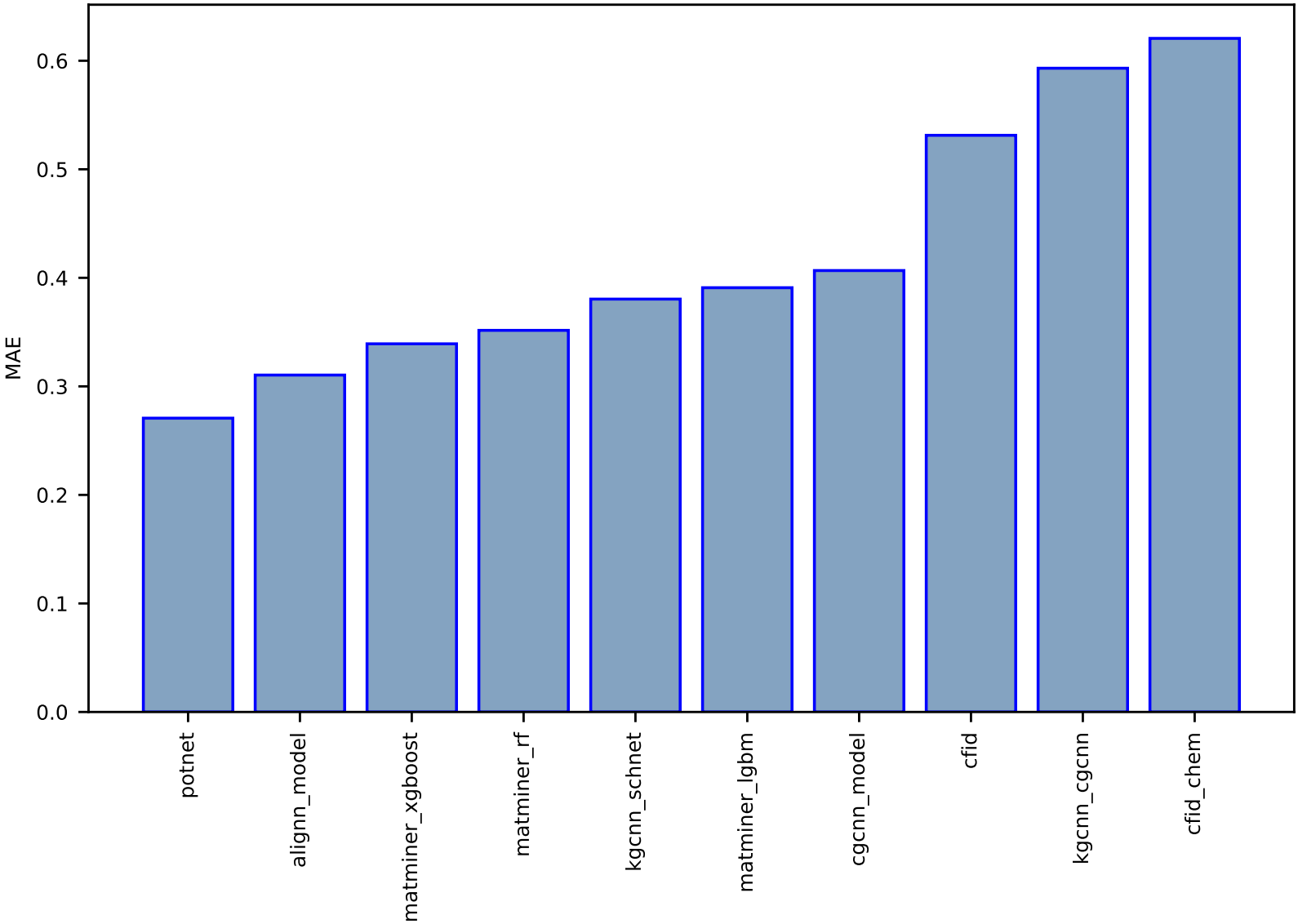


Fig S113:AI-SinglePropertyPrediction-max_efg-dft_3d-test-mae.csv.zip

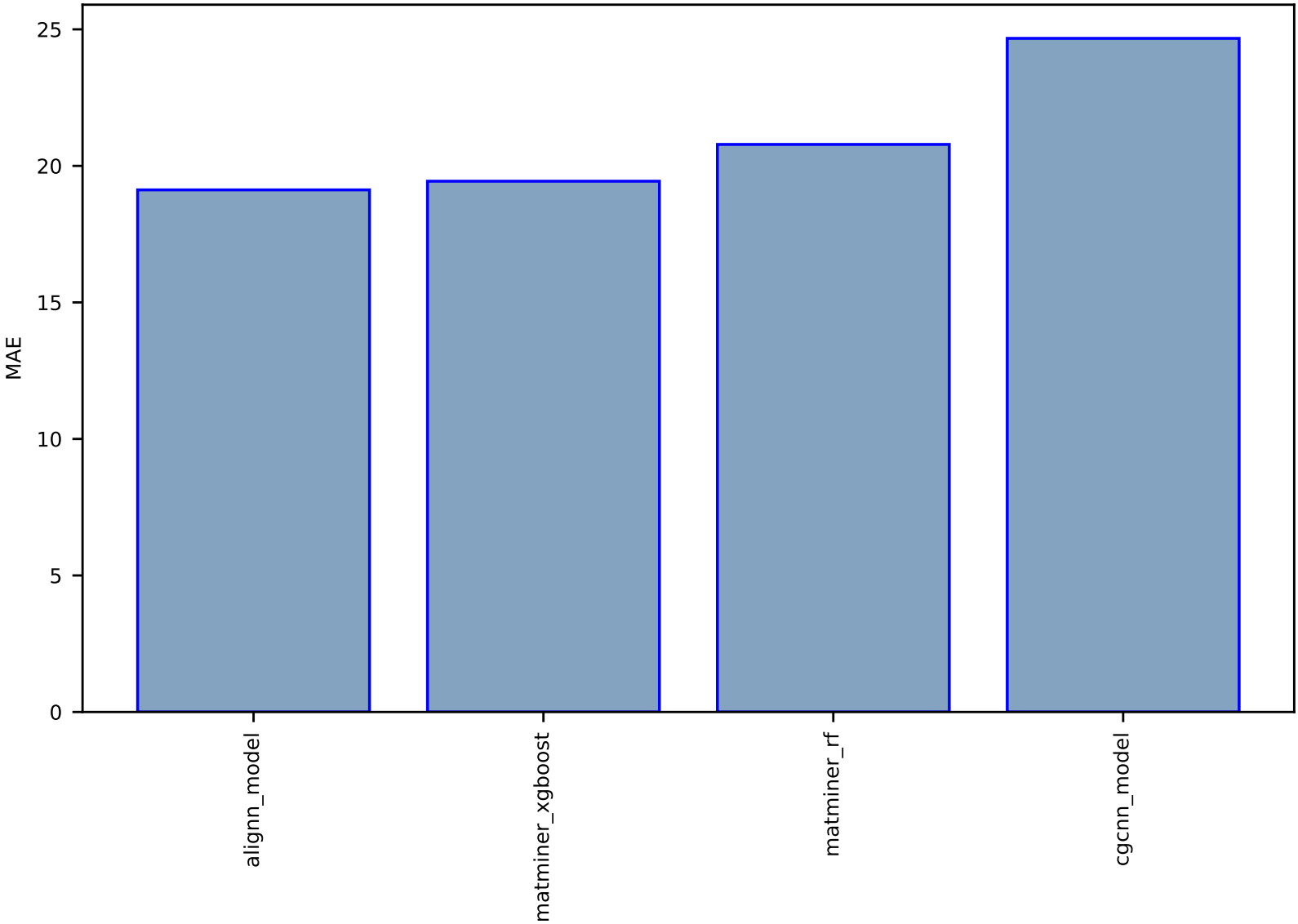


Fig S114:AI-SinglePropertyPrediction-n_Seebeck-dft_3d-test-mae.csv.zip

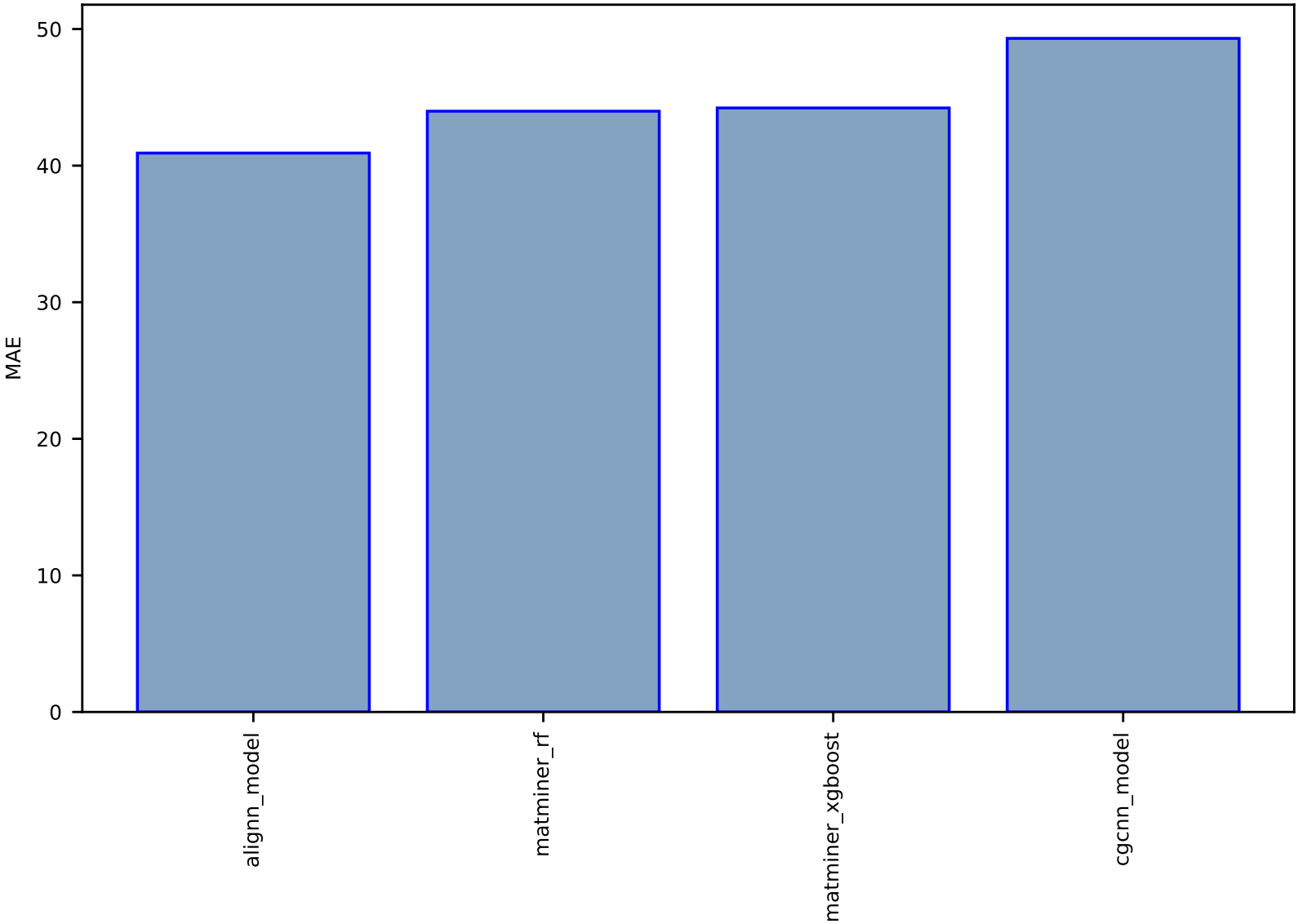


Fig S115:AI-SinglePropertyPrediction-ehull-dft_3d-test-mae.csv.zip

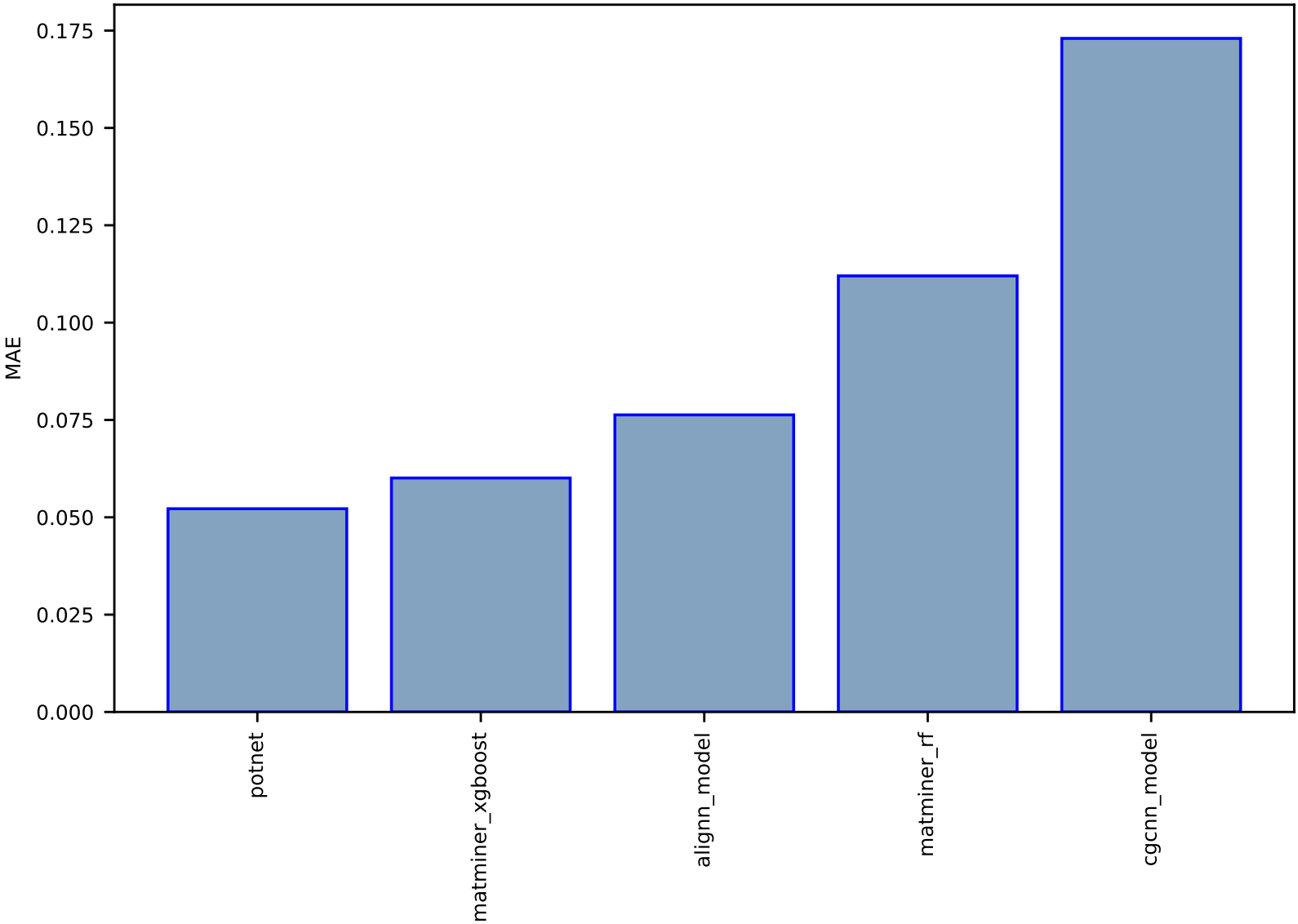


Fig S116:AI-SinglePropertyPrediction-magmom_oszicar-dft_3d-test-mae.csv.zip

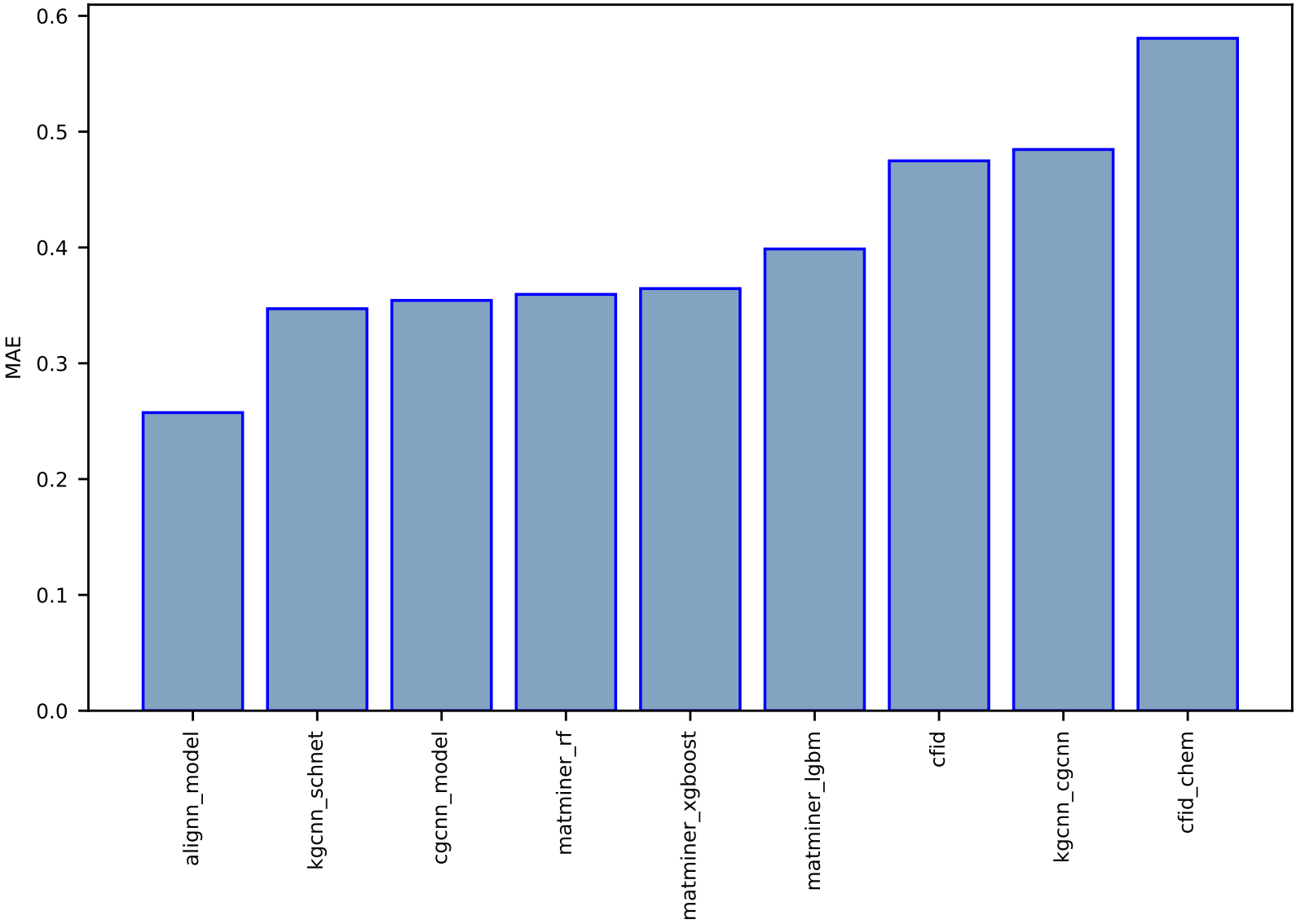


Fig S117:AI-SinglePropertyPrediction-dfpt_piezo_max_dielectric-dft_3d-test-mae.csv.zip

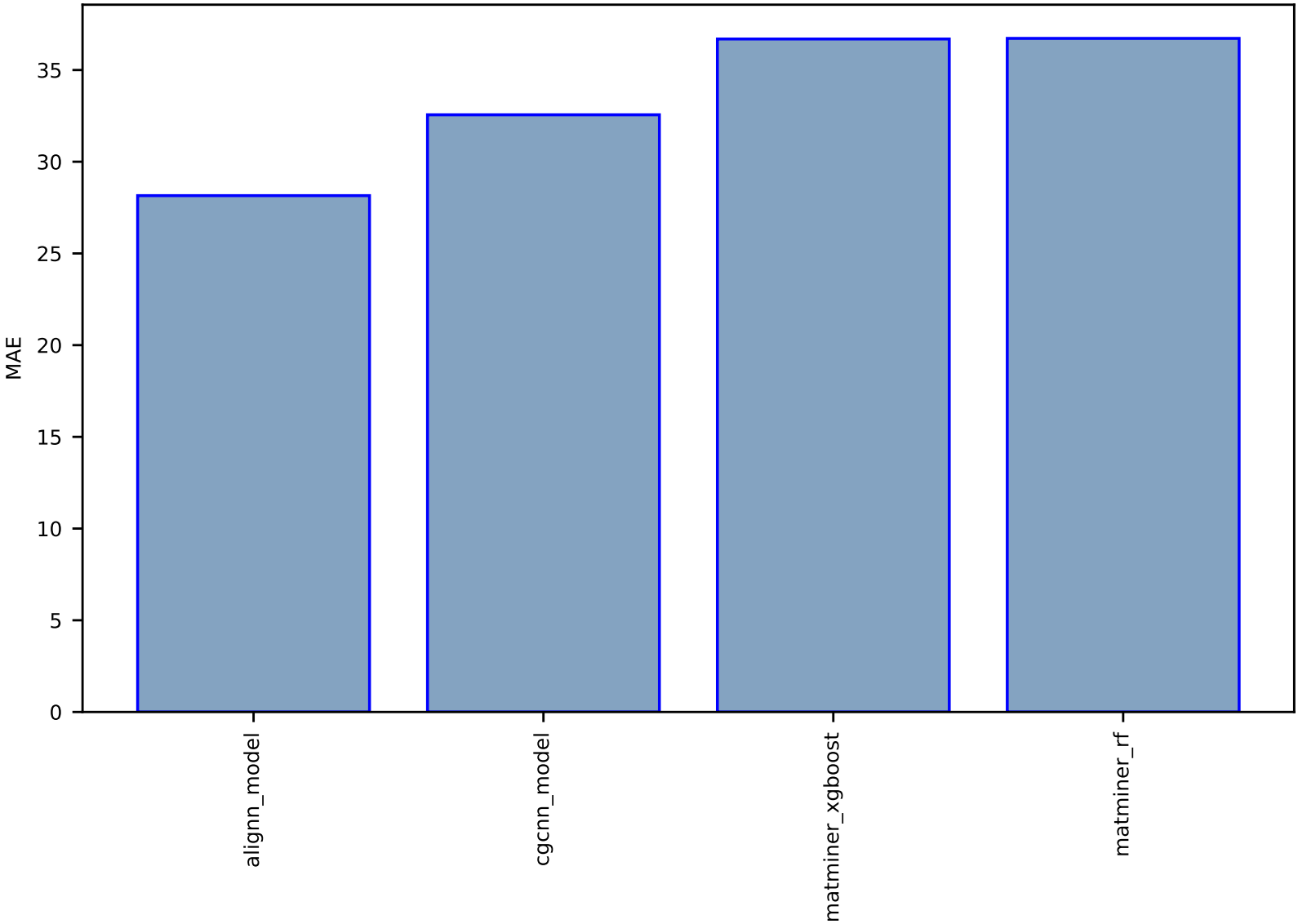


Fig S118:ES-SinglePropertyPrediction-bulk_modulus_JVASP_1002_Si-dft_3d-test-mae.csv.zip

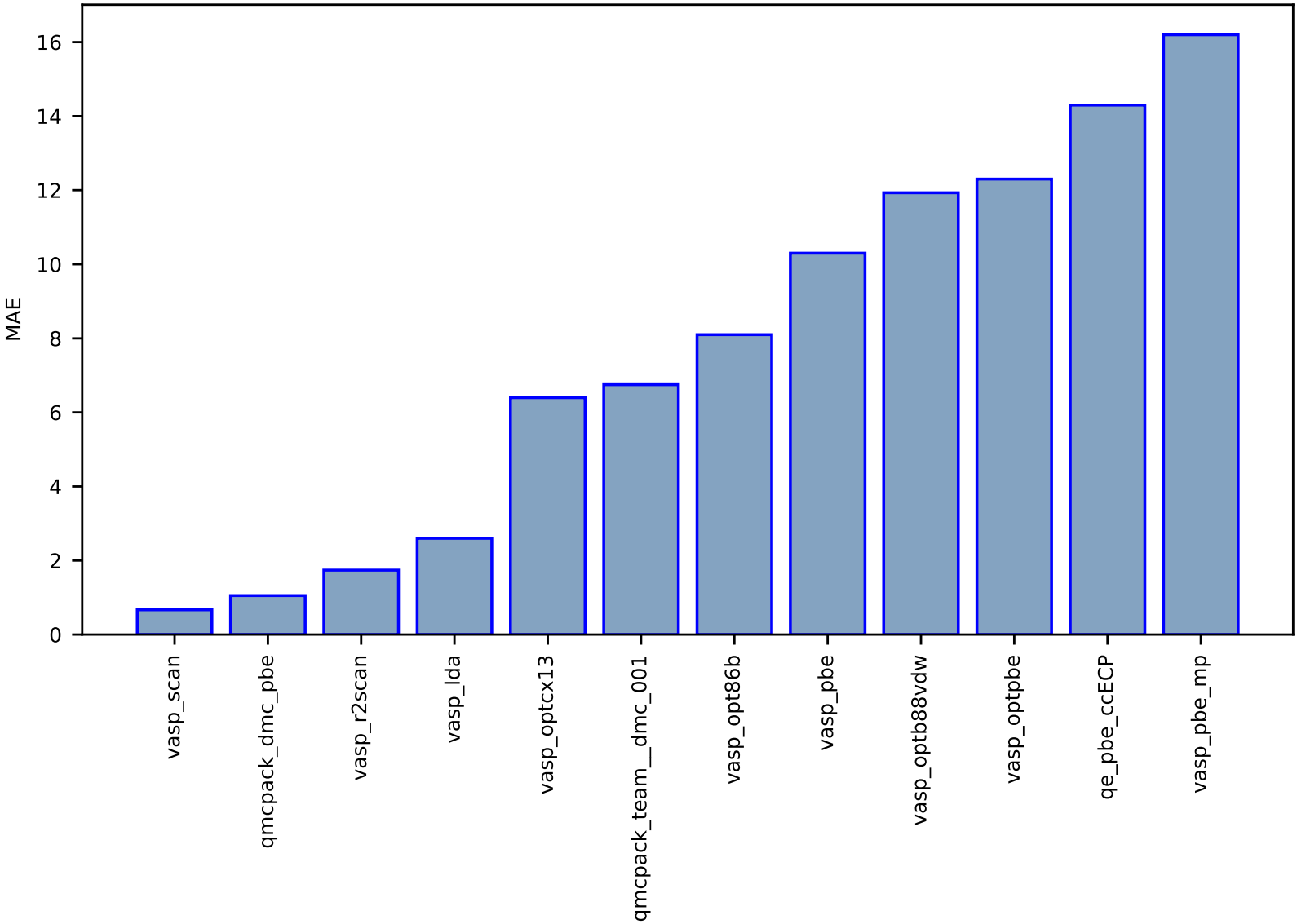


Fig S119:AI-SinglePropertyPrediction-ef-vacancydb_elements-test-mae.csv.zip

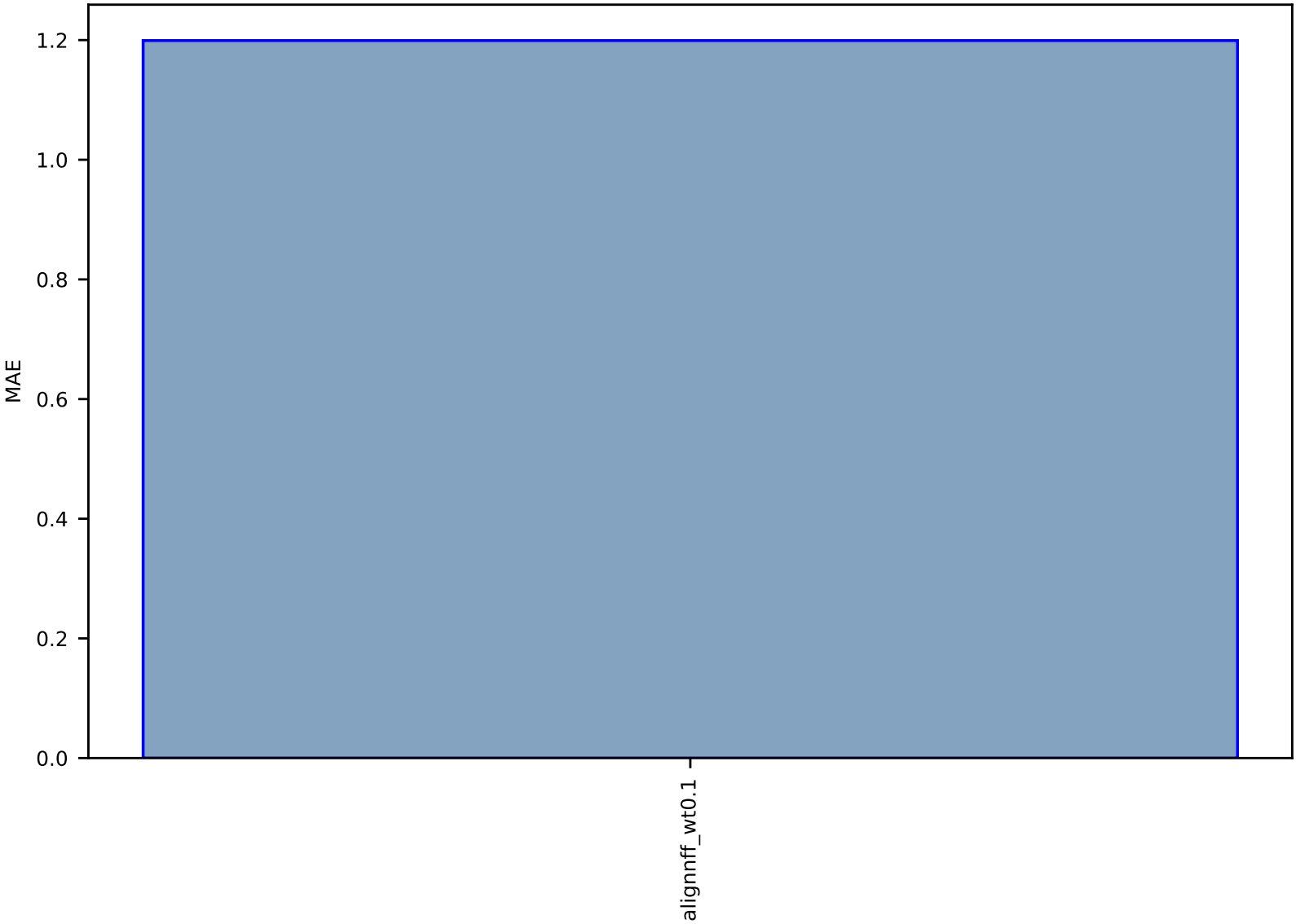


Fig S120:AI-SinglePropertyPrediction-ef-vacancydb_oxides-test-mae.csv.zip

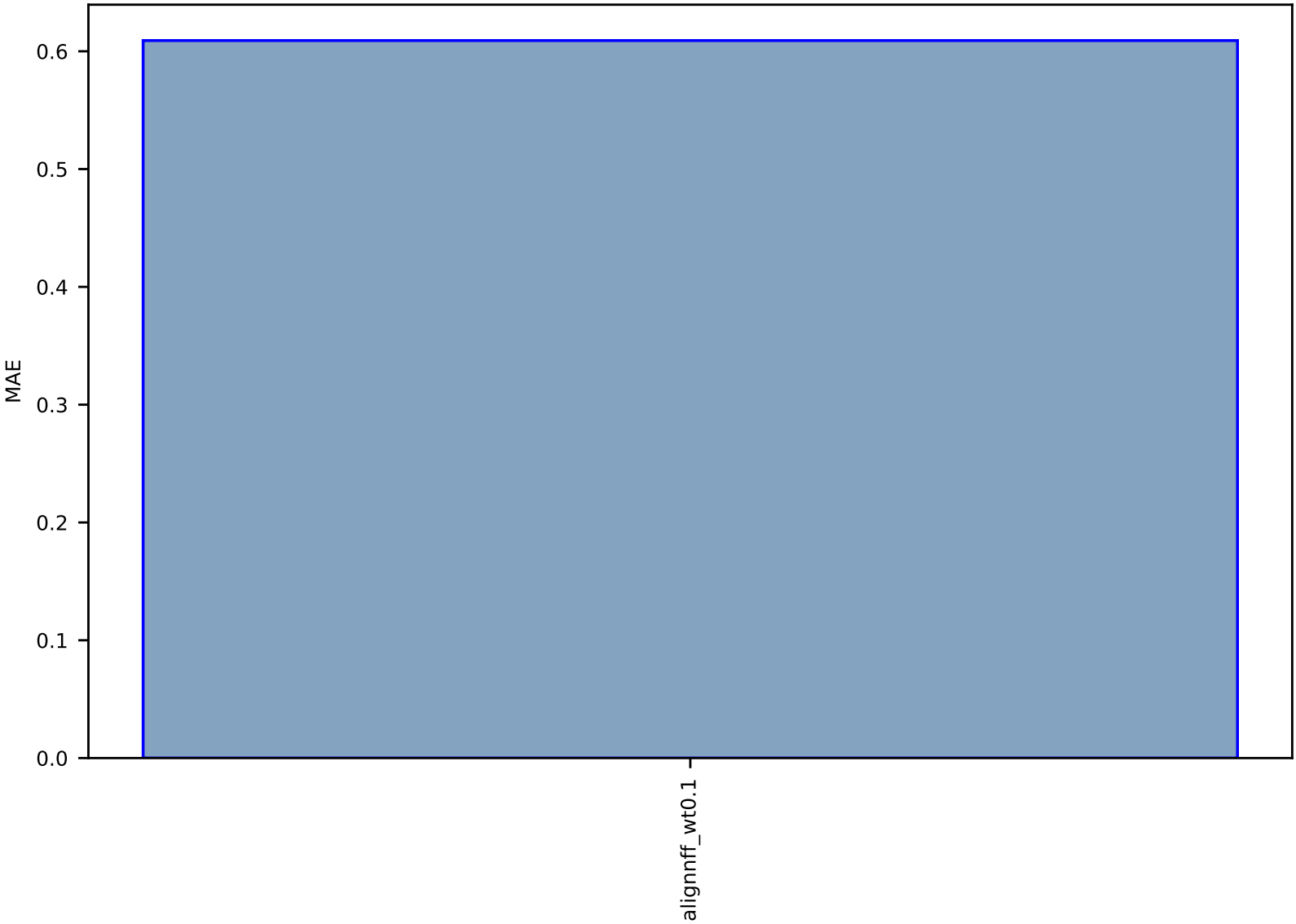


Fig S121:AI-SinglePropertyPrediction-ef-vacancydb_2D-test-mae.csv.zip

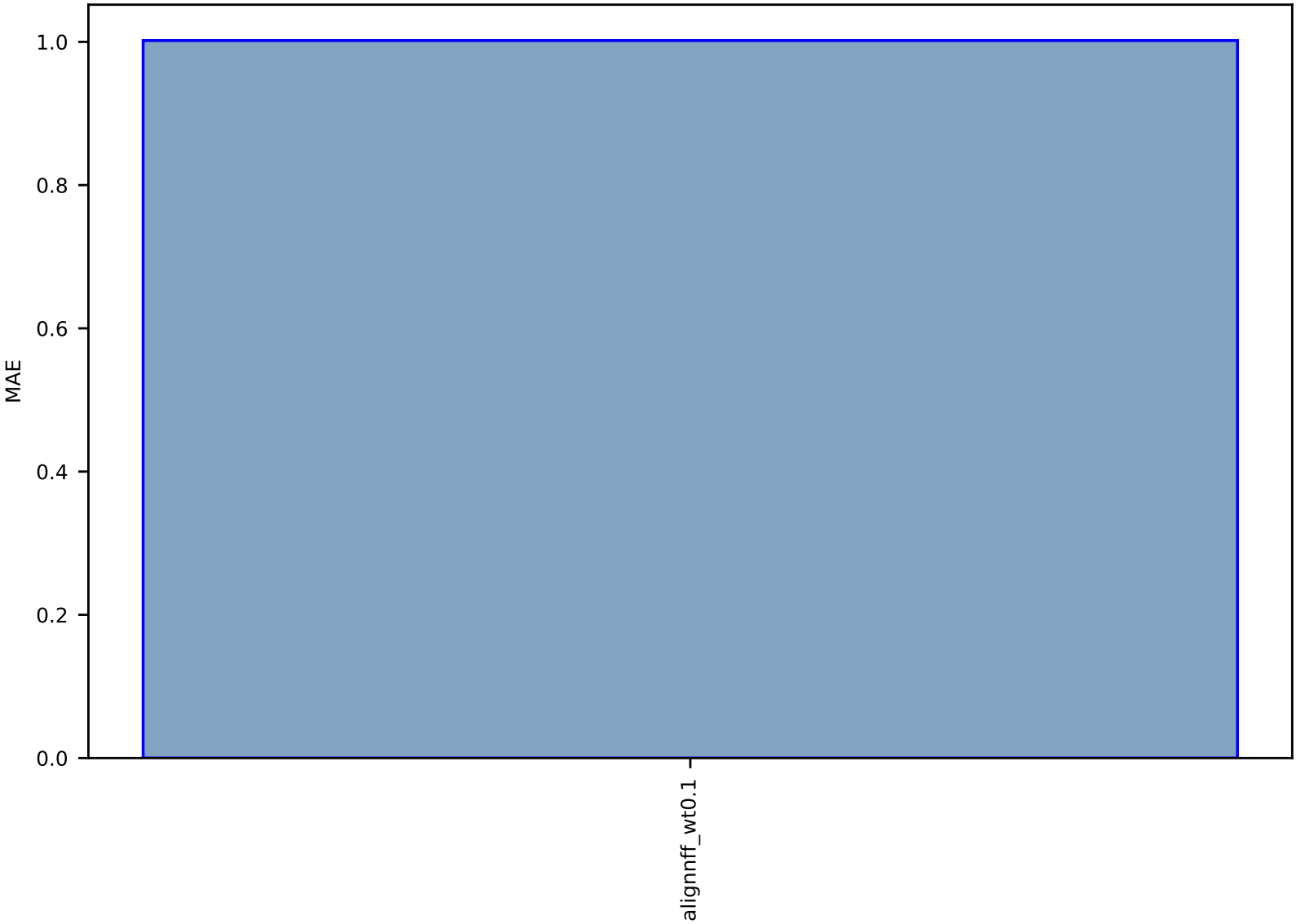


Fig S122:AI-SinglePropertyPrediction-ef-vacancydb-test-mae.csv.zip

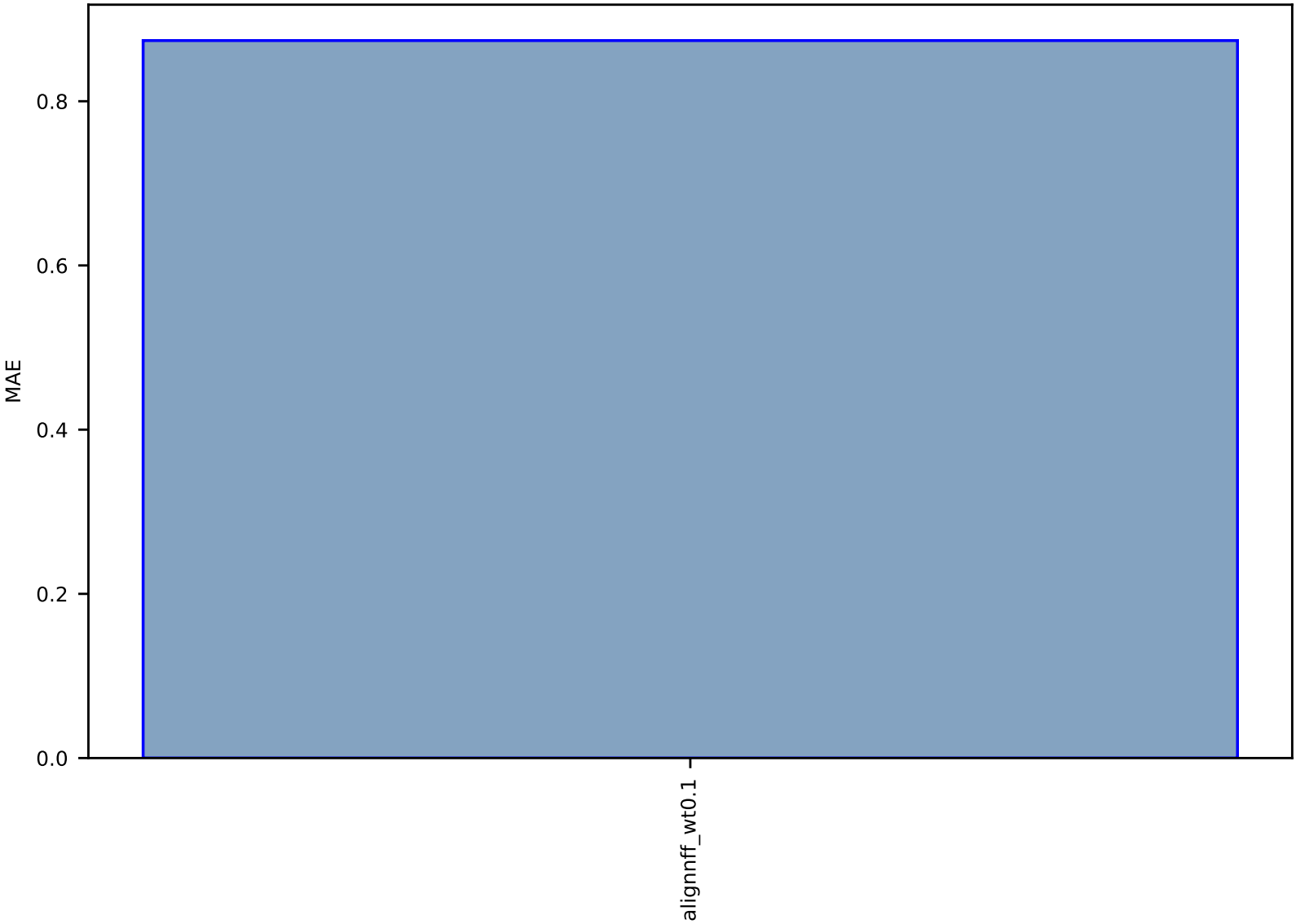


Fig S123:AI-SinglePropertyPrediction-relaxed_energy-ocp100k-test-mae.csv.zip

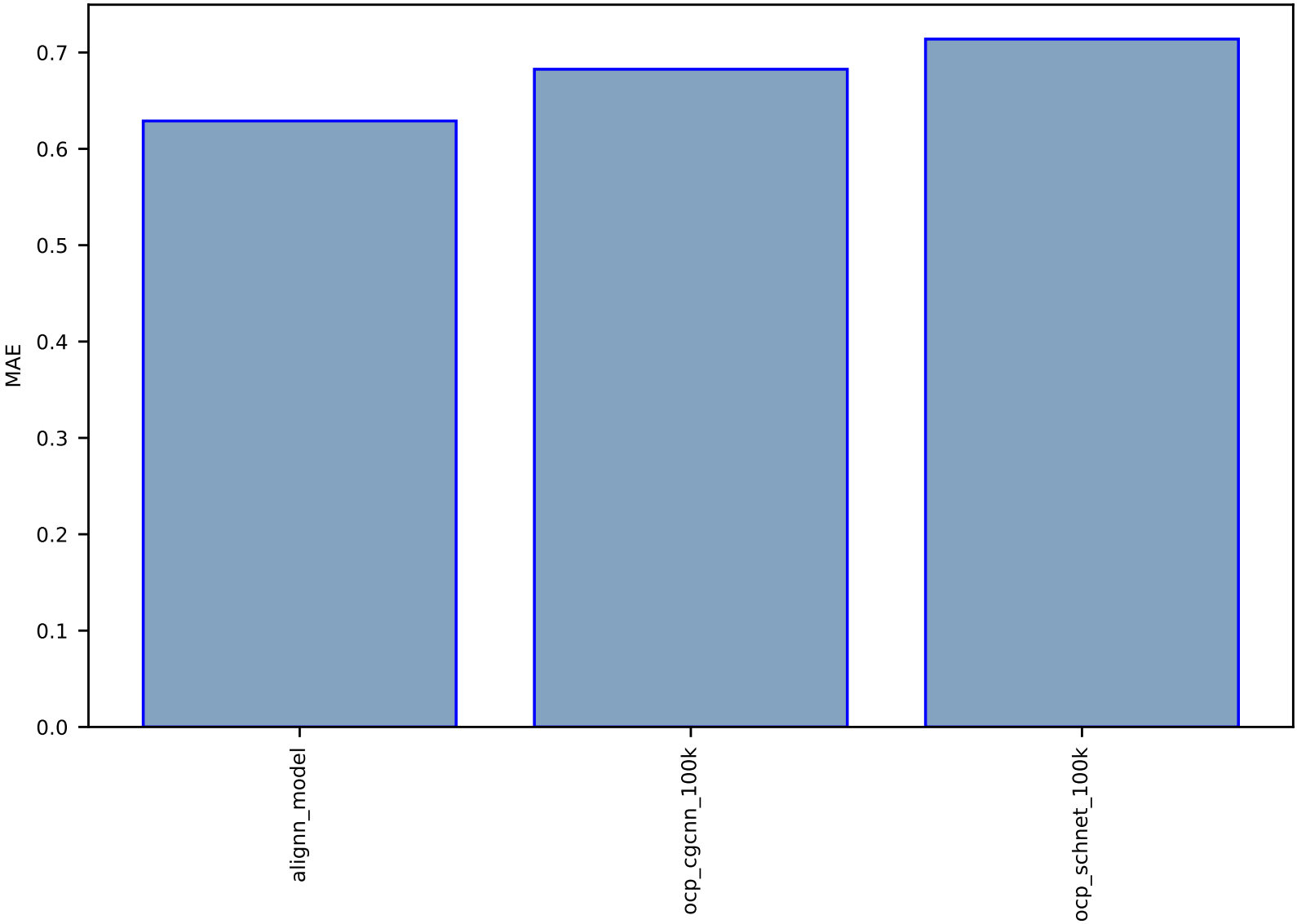


Fig S124:ES-SinglePropertyPrediction-bulk_modulus_JVASP_1174_GaAs-dft_3d-test-mae.csv.zip

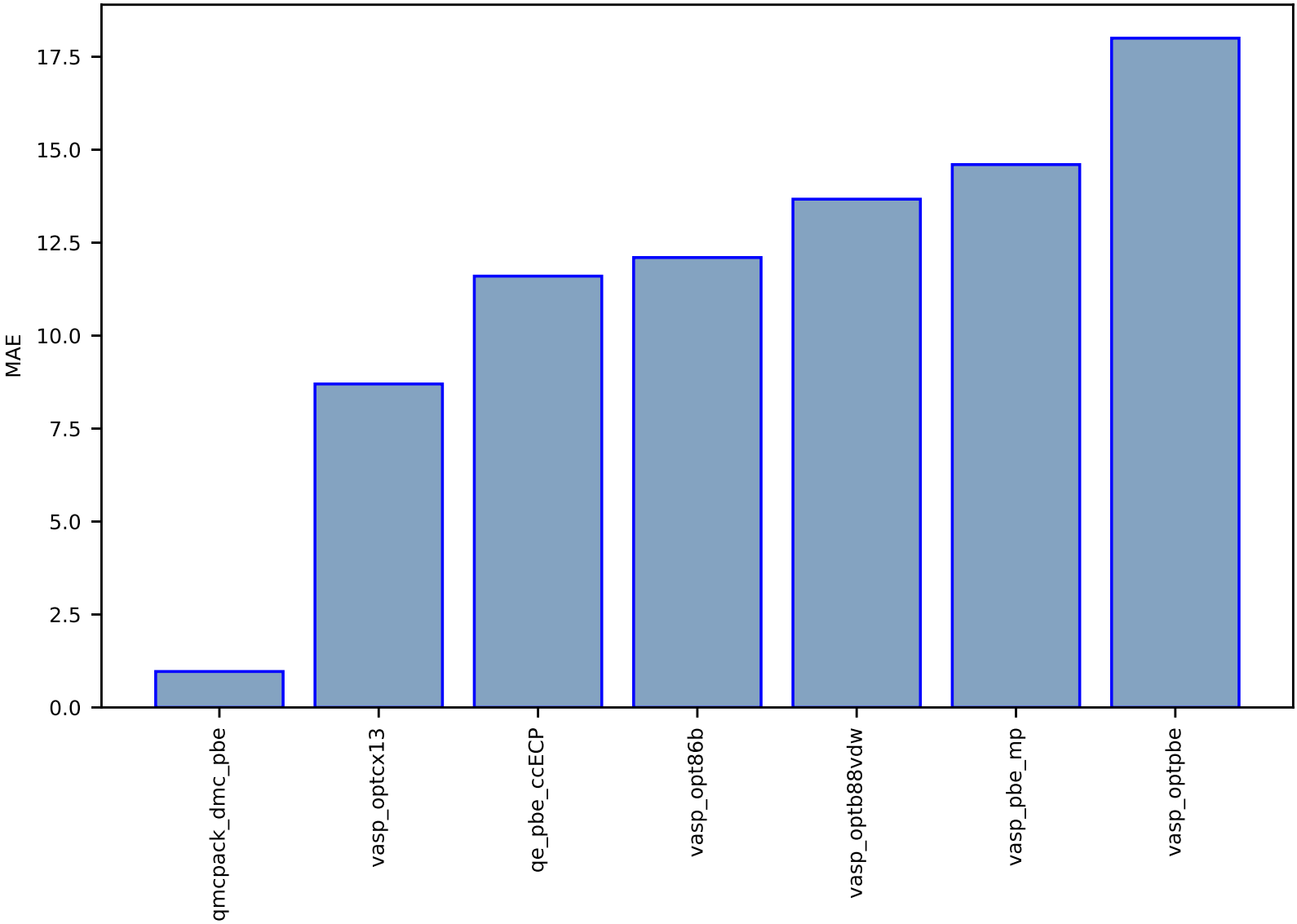


Fig S125:ES-SinglePropertyPrediction-bulk_modulus_JVASP_890_Ge-dft_3d-test-mae.csv.zip

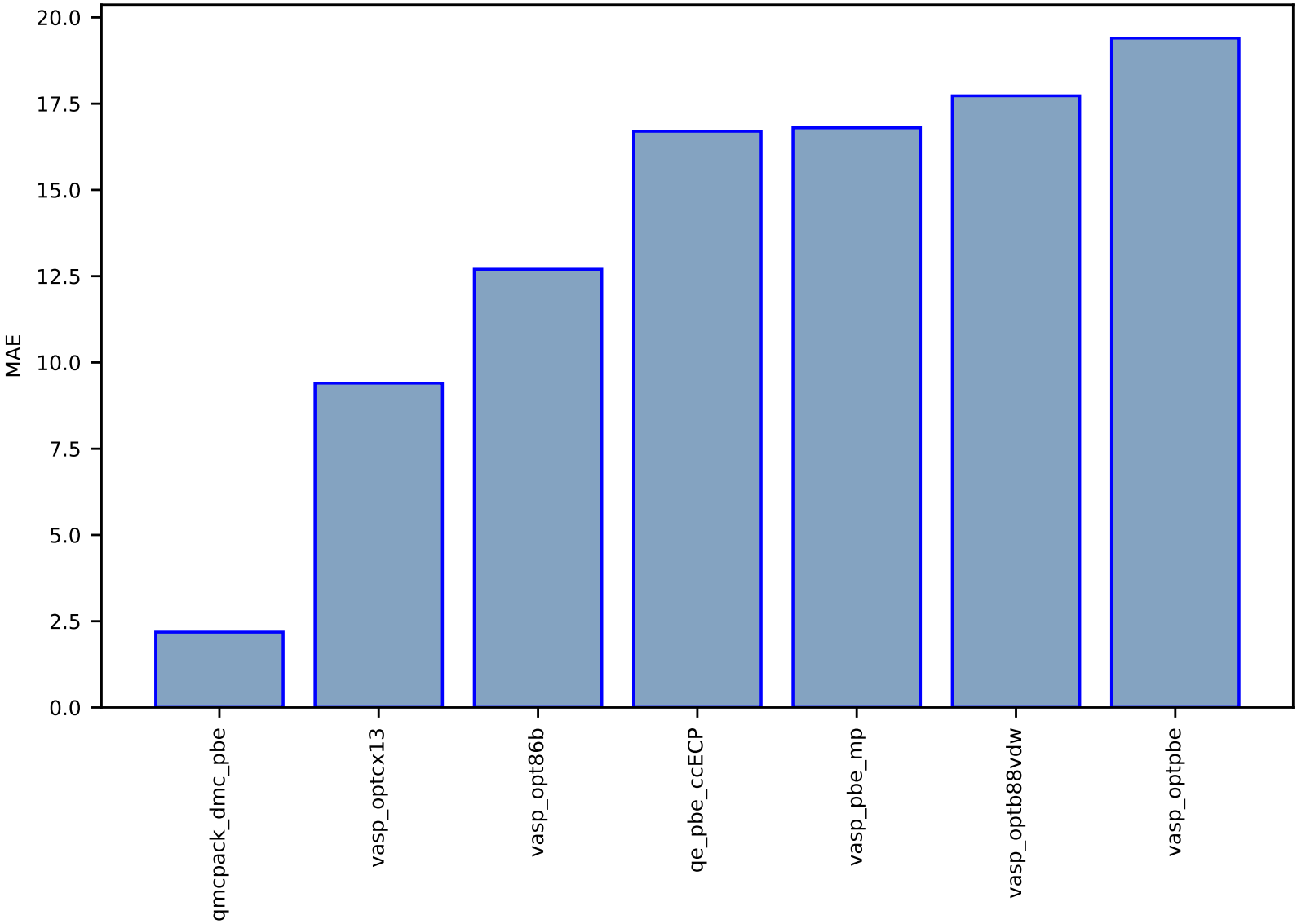


Fig S126:ES-SinglePropertyPrediction-bulk_modulus_JVASP_91_C-dft_3d-test-mae.csv.zip

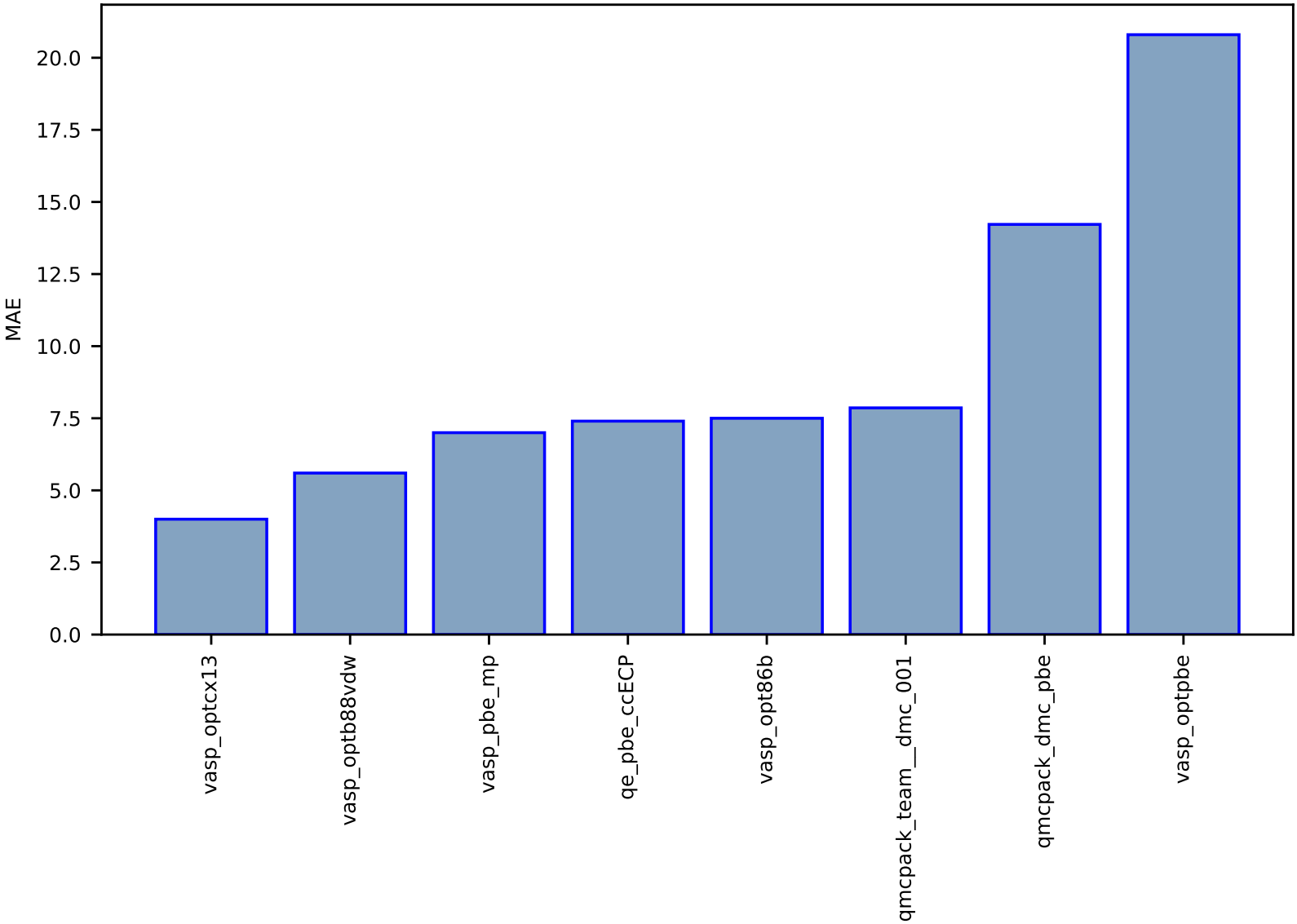


Fig S127:ES-SinglePropertyPrediction-bulk_modulus_JVASP_14813_Rb-dft_3d-test-mae.csv.zip

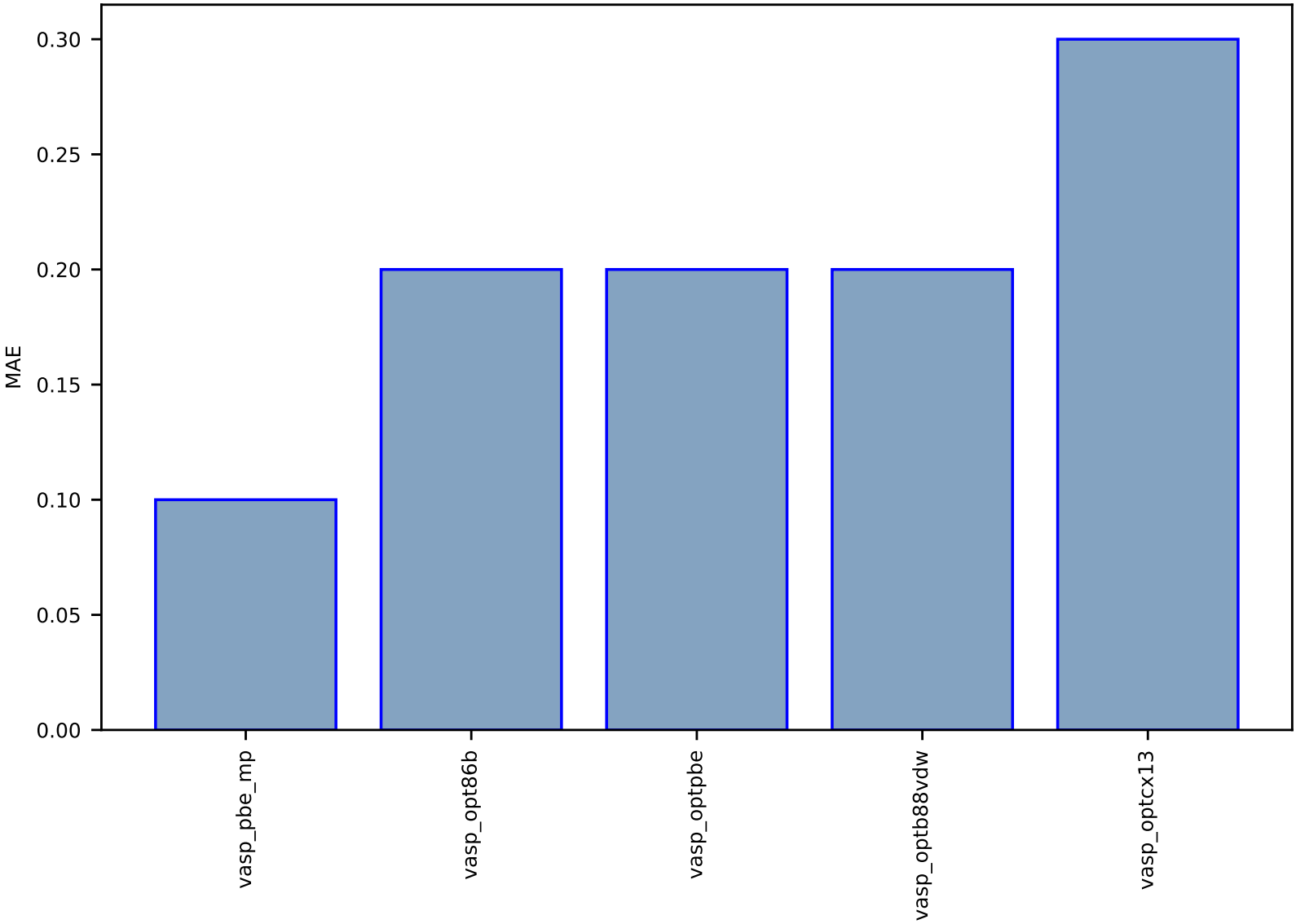


Fig S128:ES-SinglePropertyPrediction-bulk_modulus_JVASP_21208_Sr-dft_3d-test-mae.csv.zip

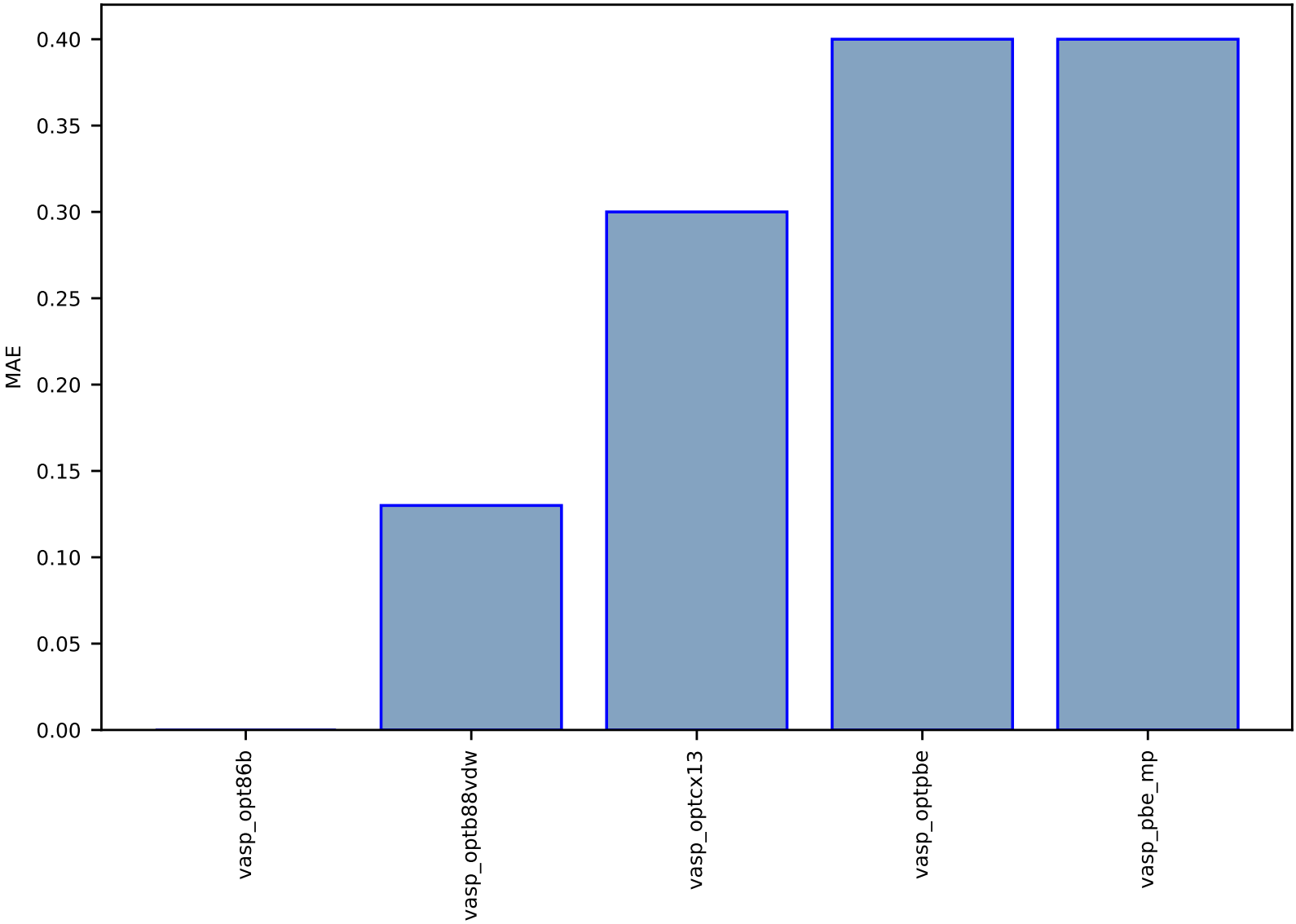


Fig S129:ES-SinglePropertyPrediction-bulk_modulus-dft_3d-test-mae.csv.zip

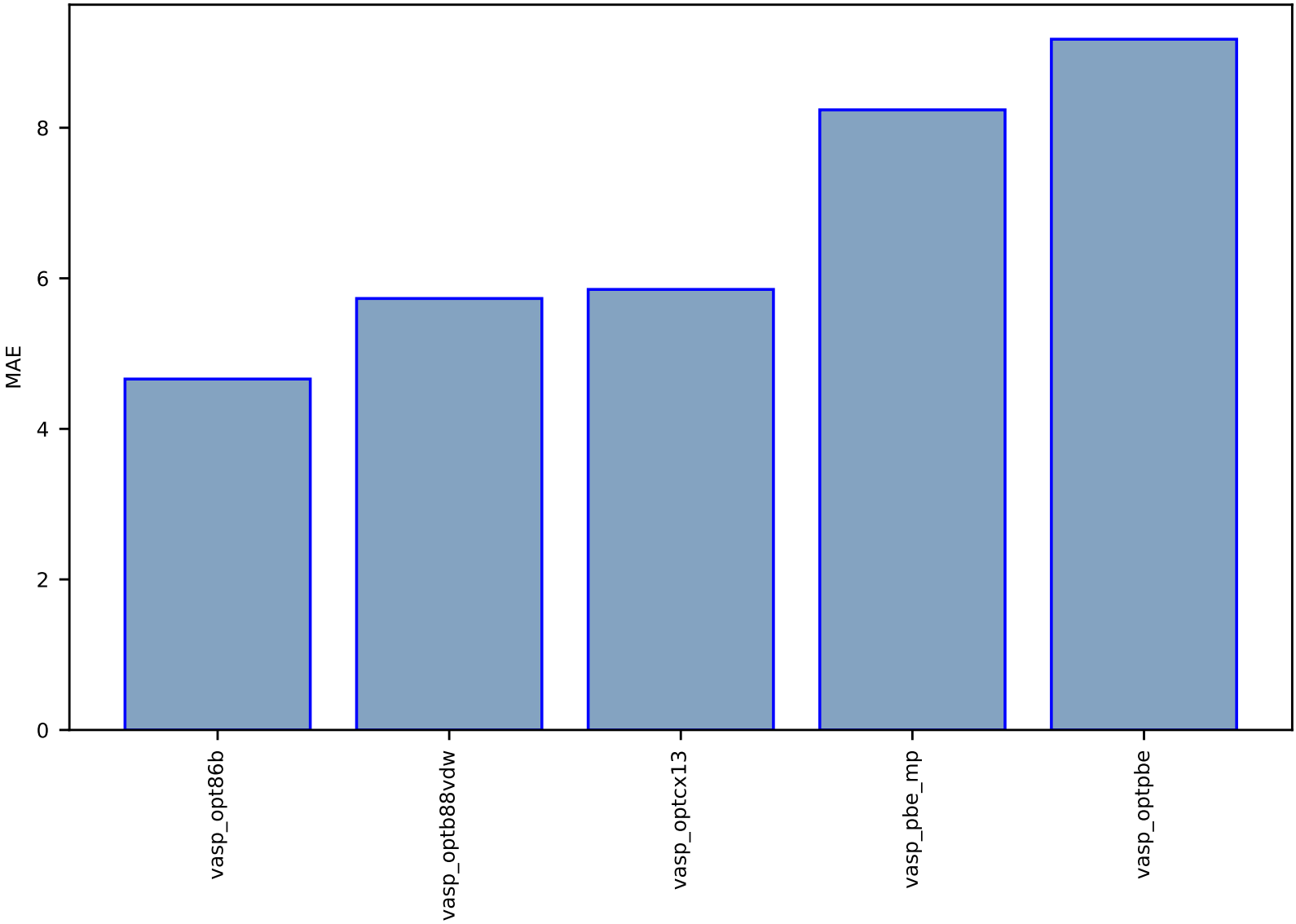


Fig S130:ES-SinglePropertyPrediction-bulk_modulus_JVASP_14606_Ag-dft_3d-test-mae.csv.zip

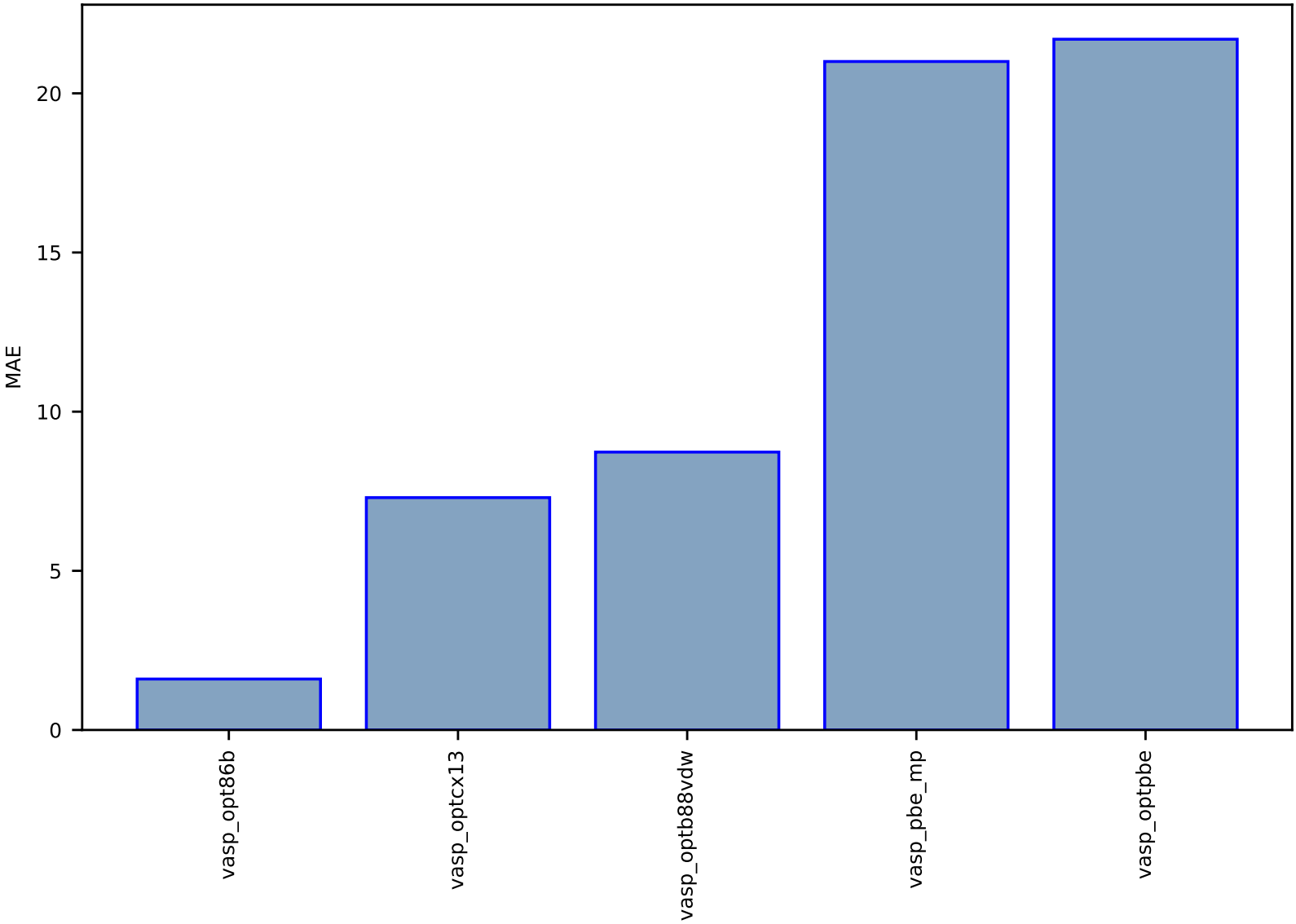


Fig S131:ES-SinglePropertyPrediction-bulk_modulus_JVASP_867_Cu-dft_3d-test-mae.csv.zip

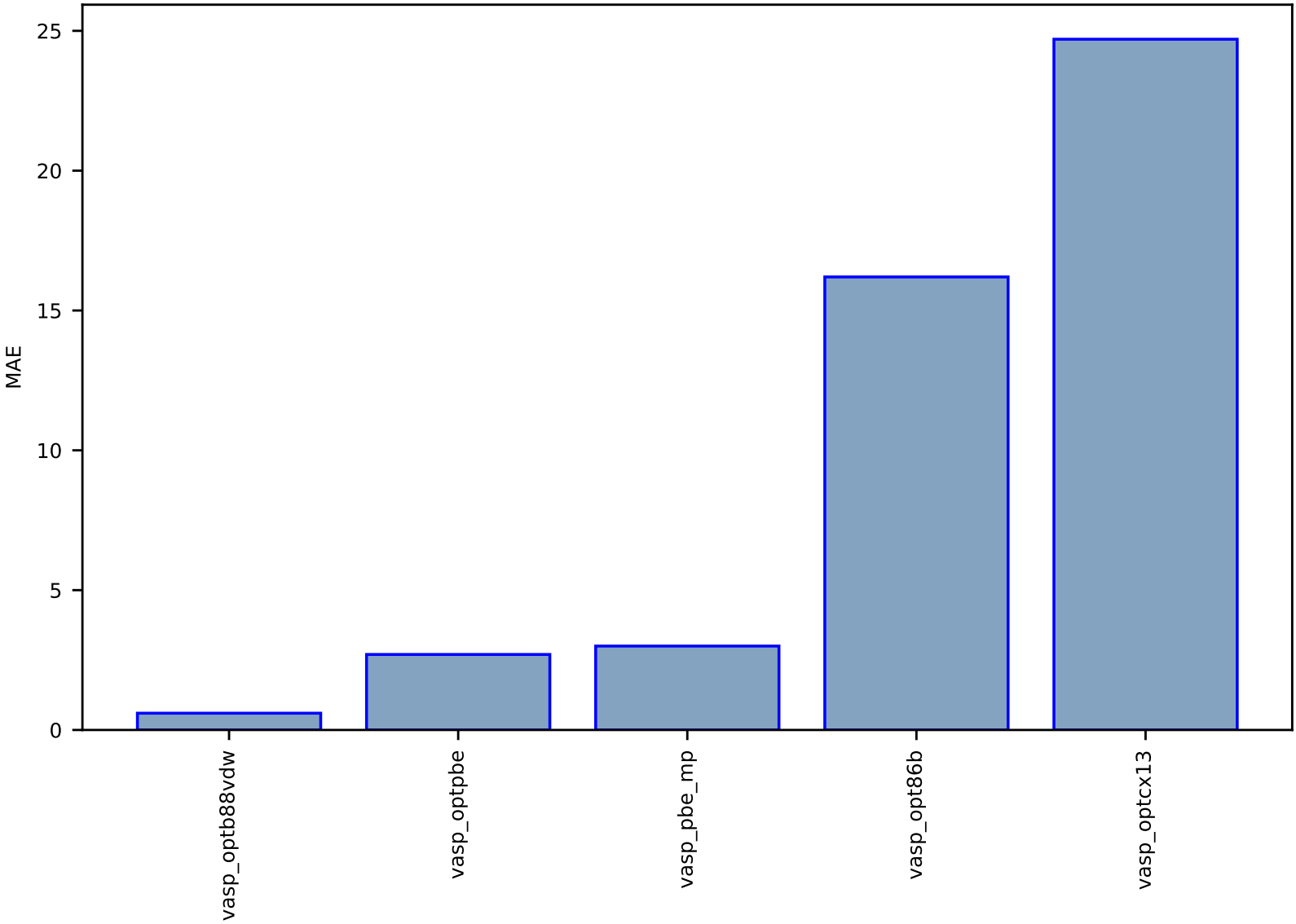


Fig S132:ES-SinglePropertyPrediction-bulk_modulus_JVASP_1130_LiF-dft_3d-test-mae.csv.zip

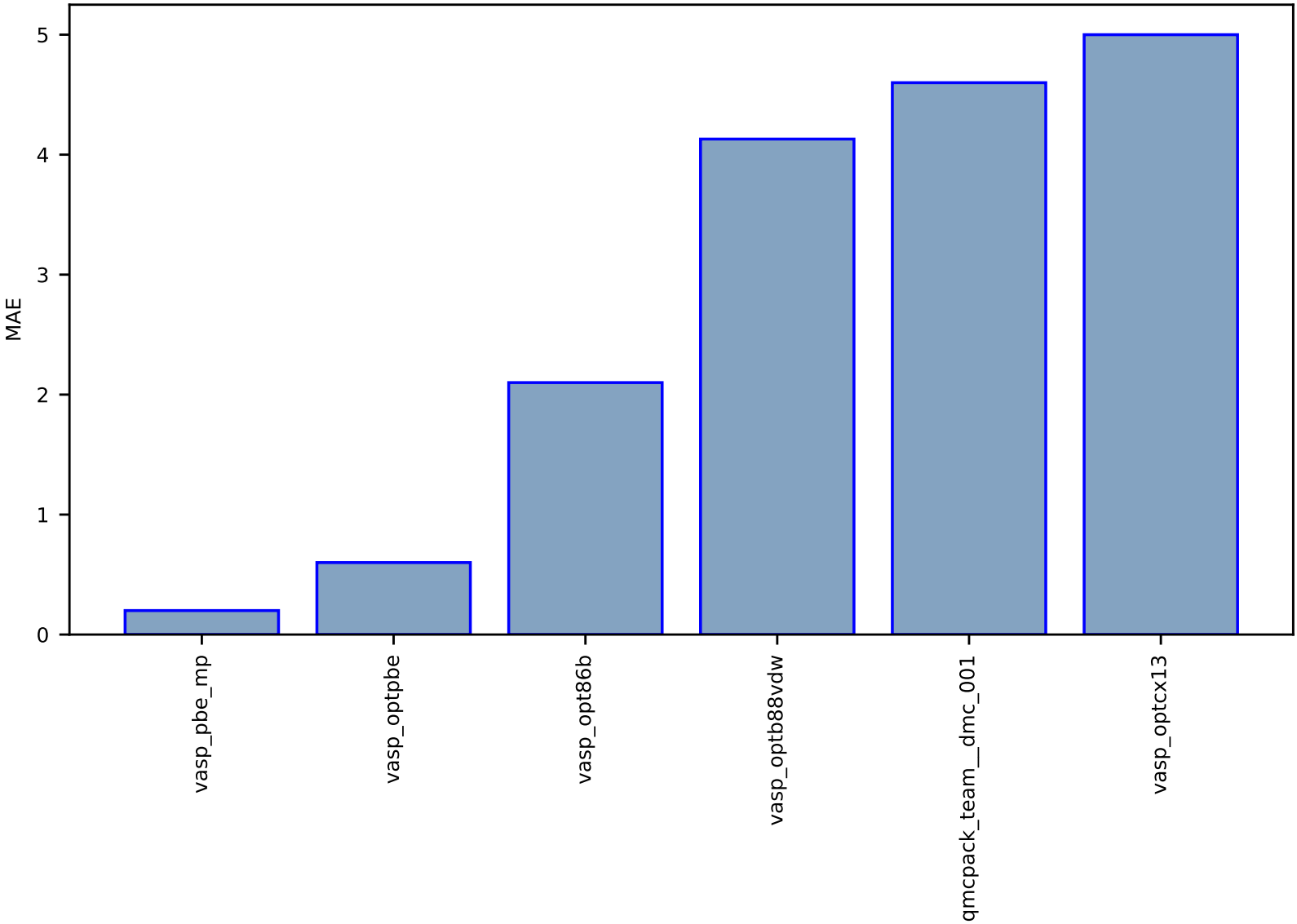


Fig S133:ES-SinglePropertyPrediction-bulk_modulus_JVASP_25114_K-dft_3d-test-mae.csv.zip

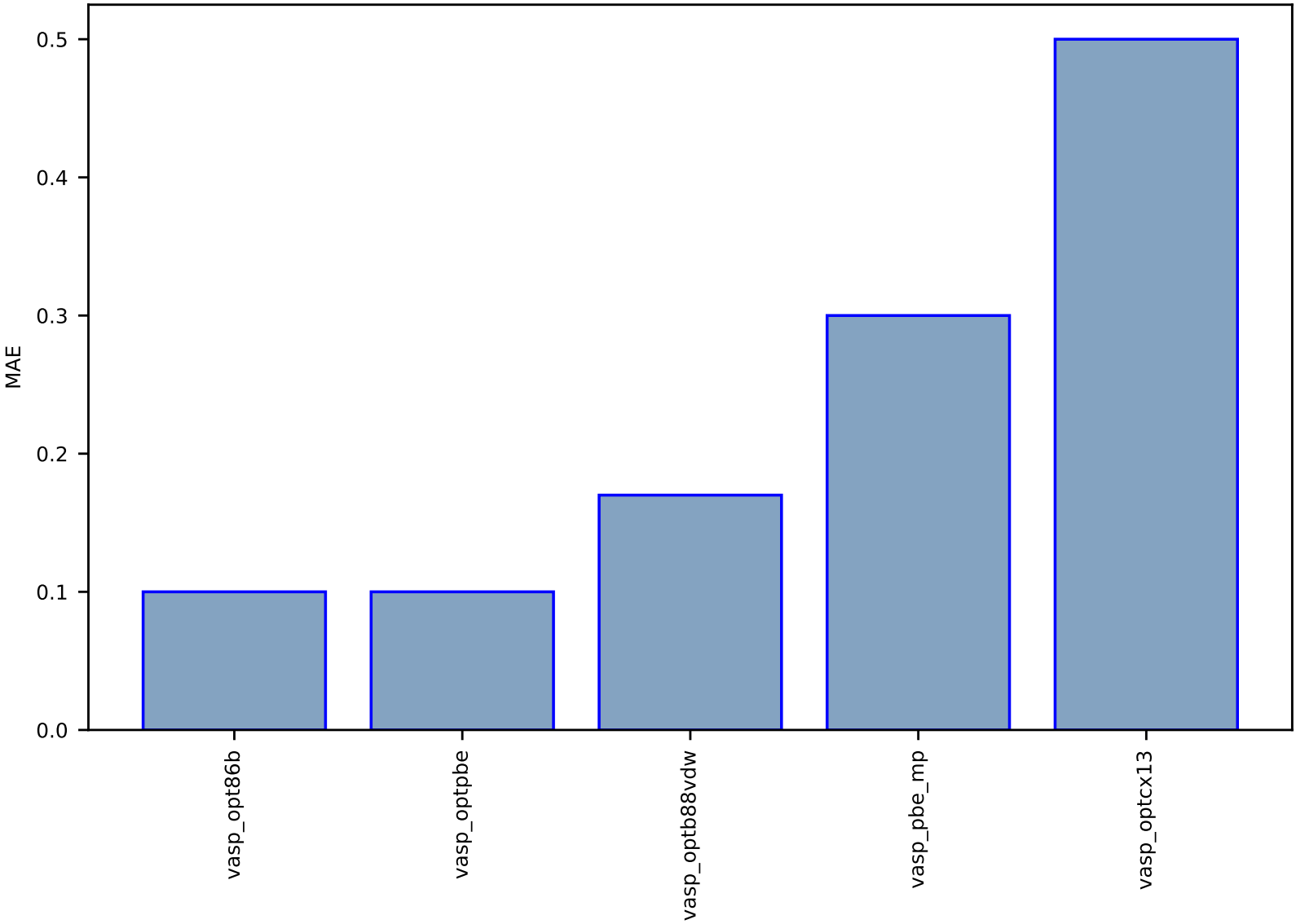


Fig S134:ES-SinglePropertyPrediction-bulk_modulus_JVASP_23862_NaCl-dft_3d-test-mae.csv.zip

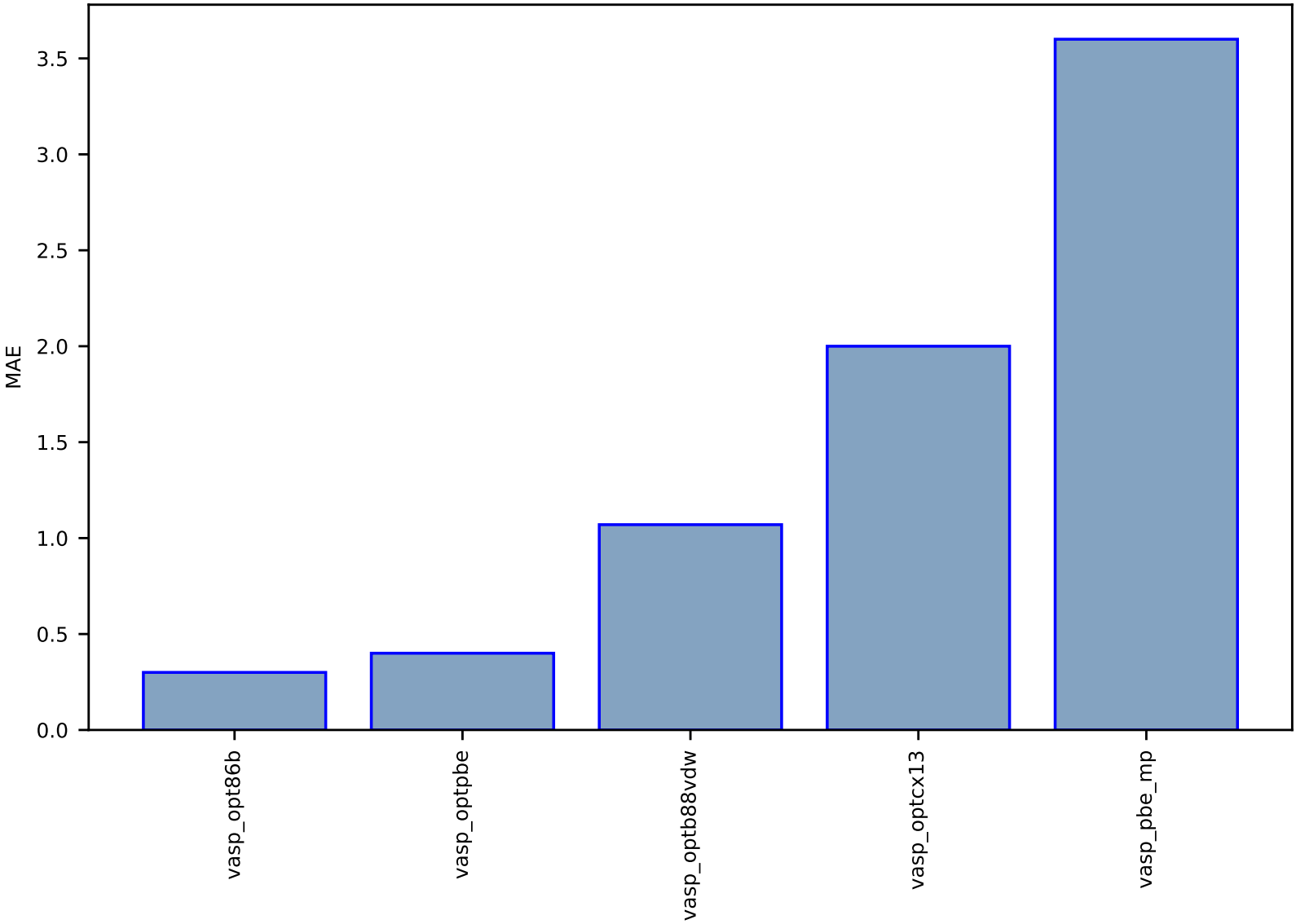


Fig S135:ES-SinglePropertyPrediction-bulk_modulus_JVASP_25065_Li-dft_3d-test-mae.csv.zip

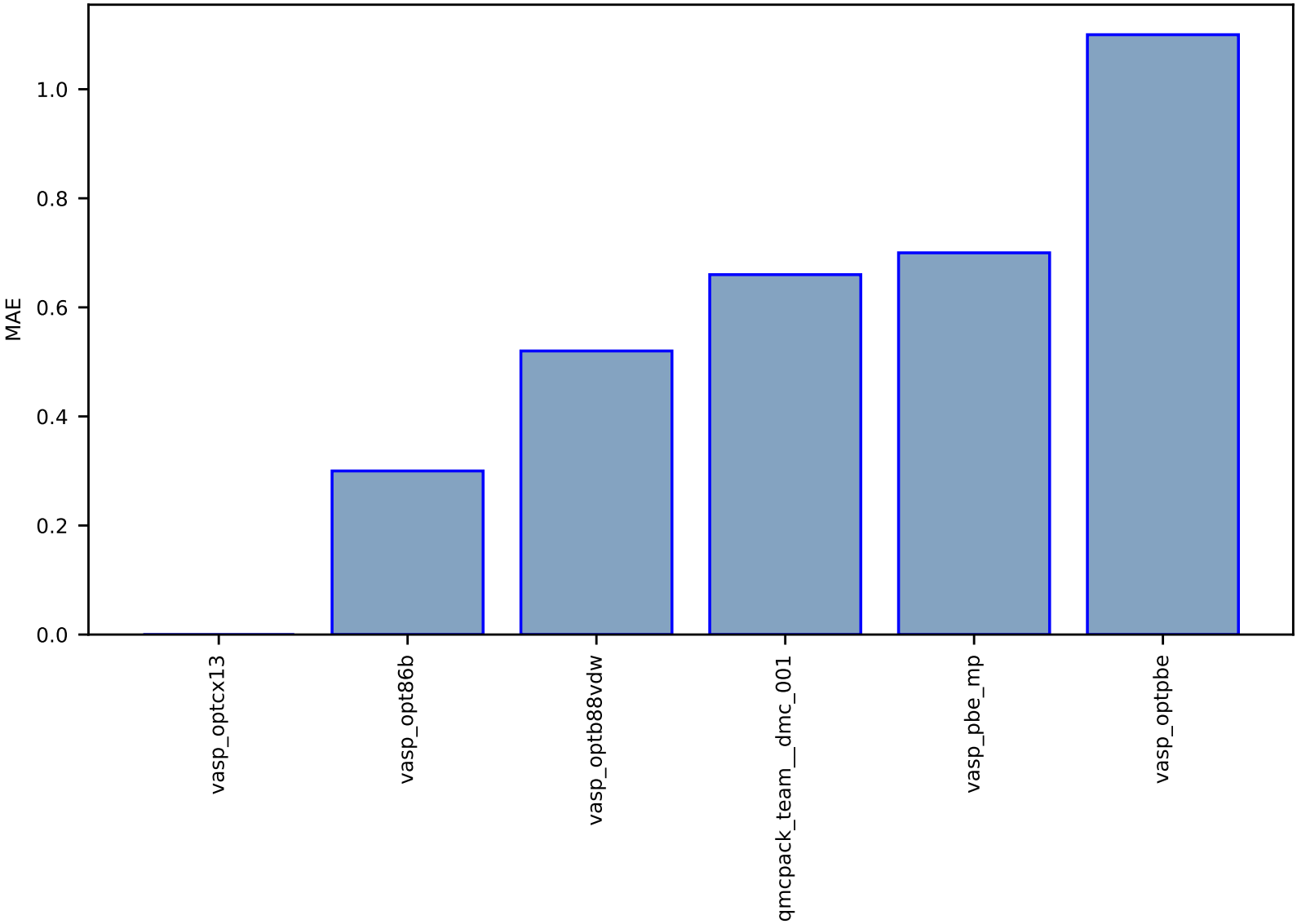


Fig S136:ES-SinglePropertyPrediction-bulk_modulus_JVASP_23864_LiCl-dft_3d-test-mae.csv.zip

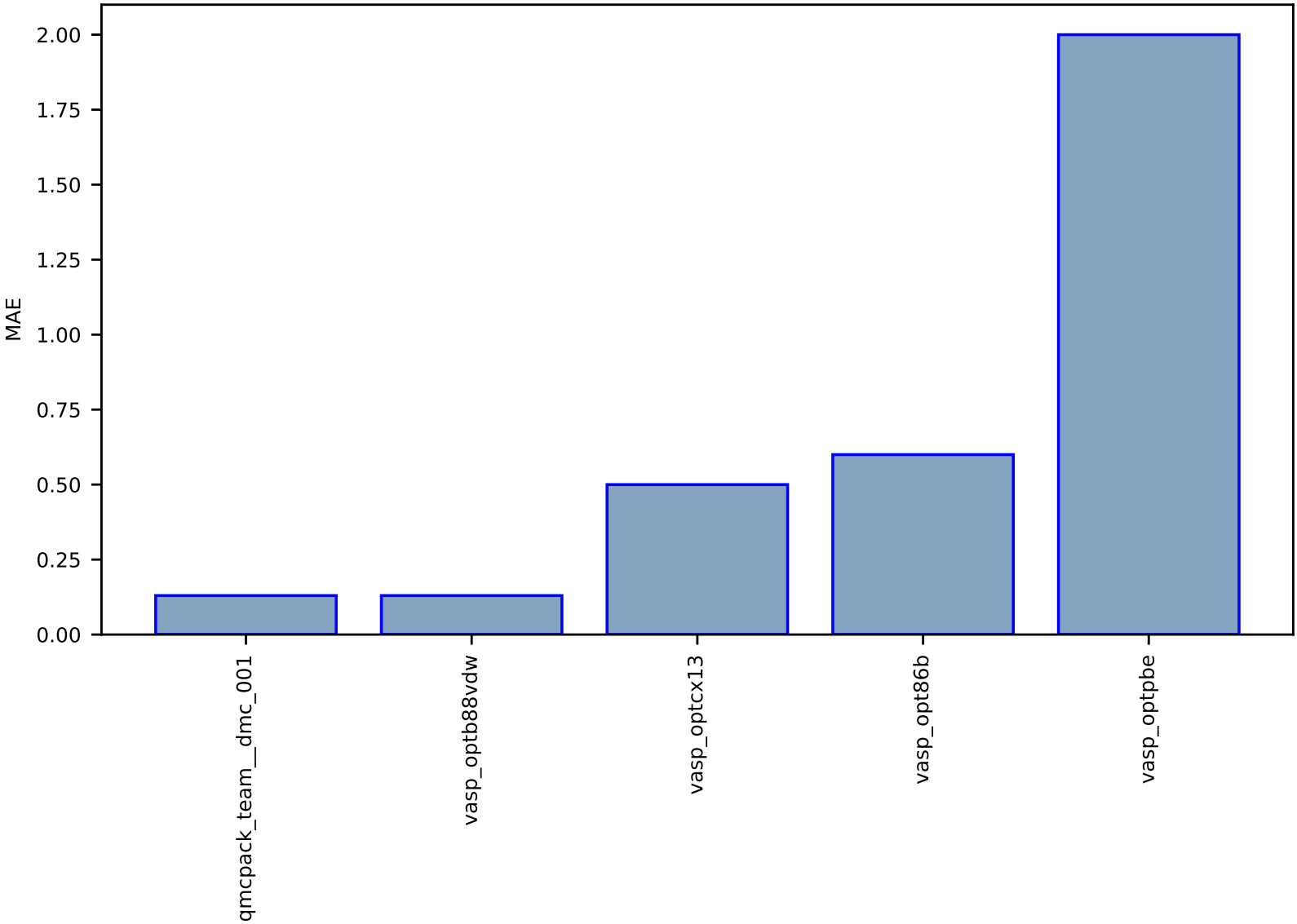


Fig S137:ES-SinglePropertyPrediction-bulk_modulus_JVASP_116_MgO-dft_3d-test-mae.csv.zip

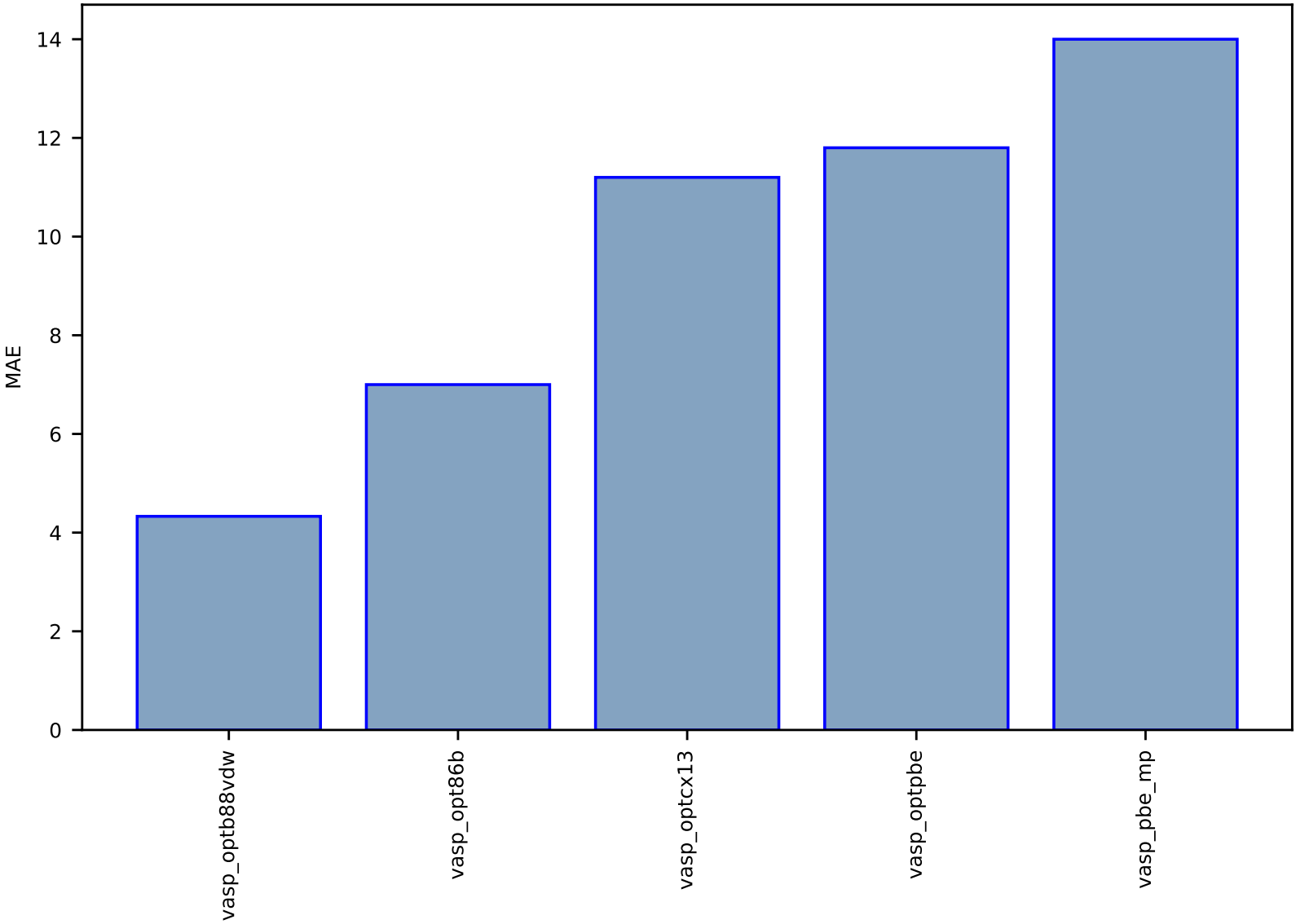


Fig S138:ES-SinglePropertyPrediction-bulk_modulus_JVASP_20326_NaF-dft_3d-test-mae.csv.zip

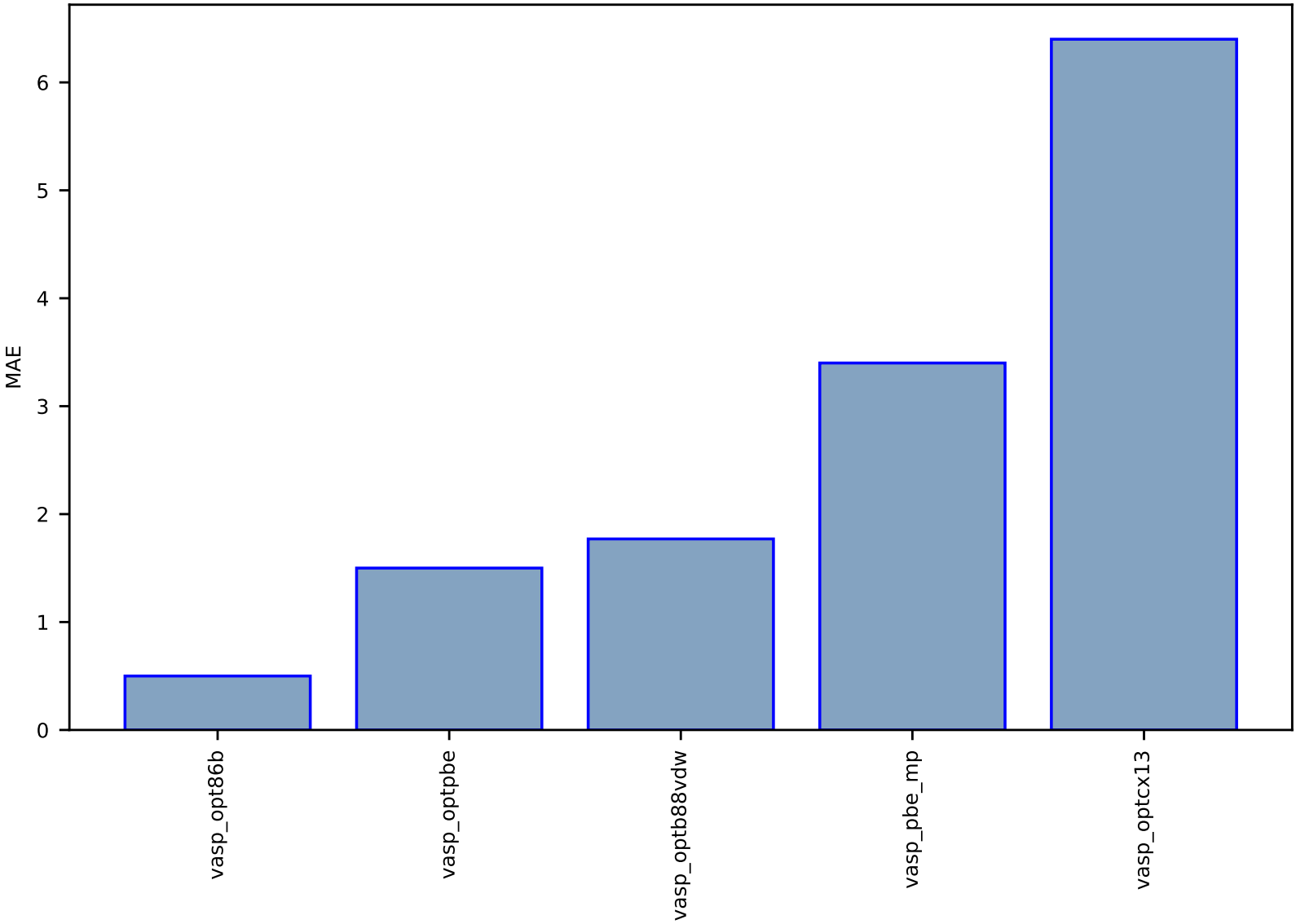


Fig S139:ES-SinglePropertyPrediction-bulk_modulus_JVASP_14604_Ba-dft_3d-test-mae.csv.zip

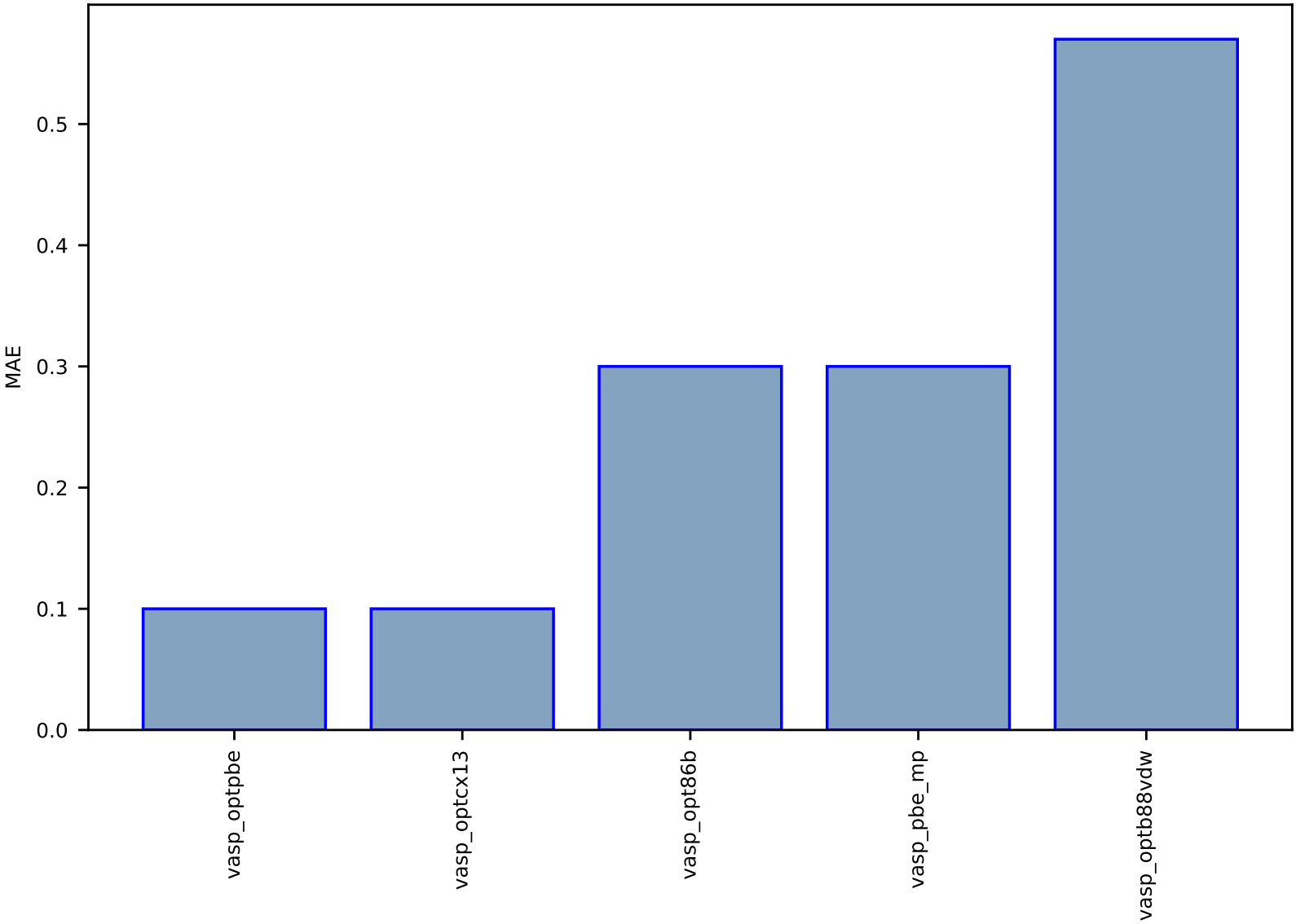


Fig S140:ES-SinglePropertyPrediction-bulk_modulus_JVASP_182_SiC-dft_3d-test-mae.csv.zip

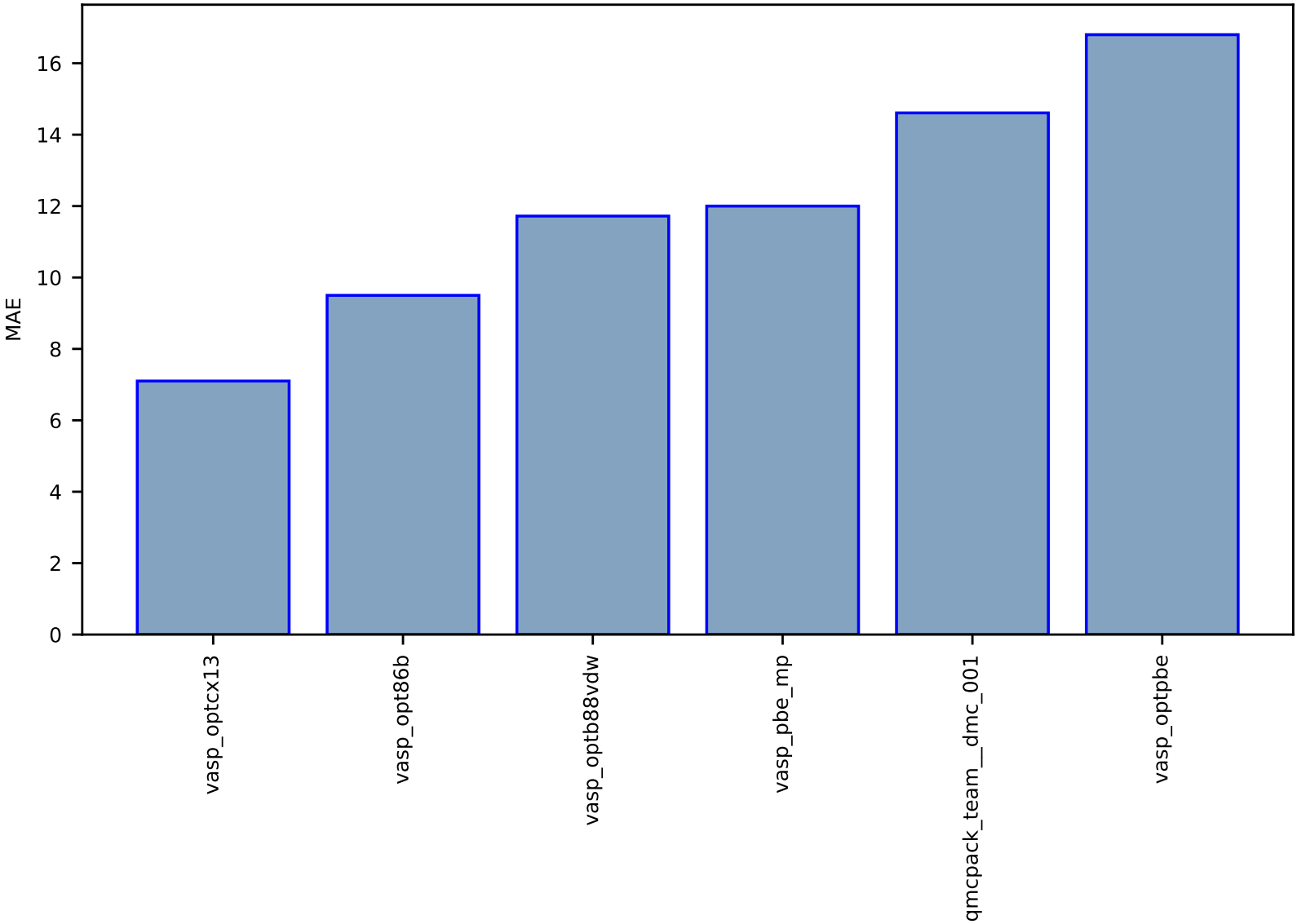


Fig S141:ES-SinglePropertyPrediction-bulk_modulus_JVASP_25180_Ca-dft_3d-test-mae.csv.zip

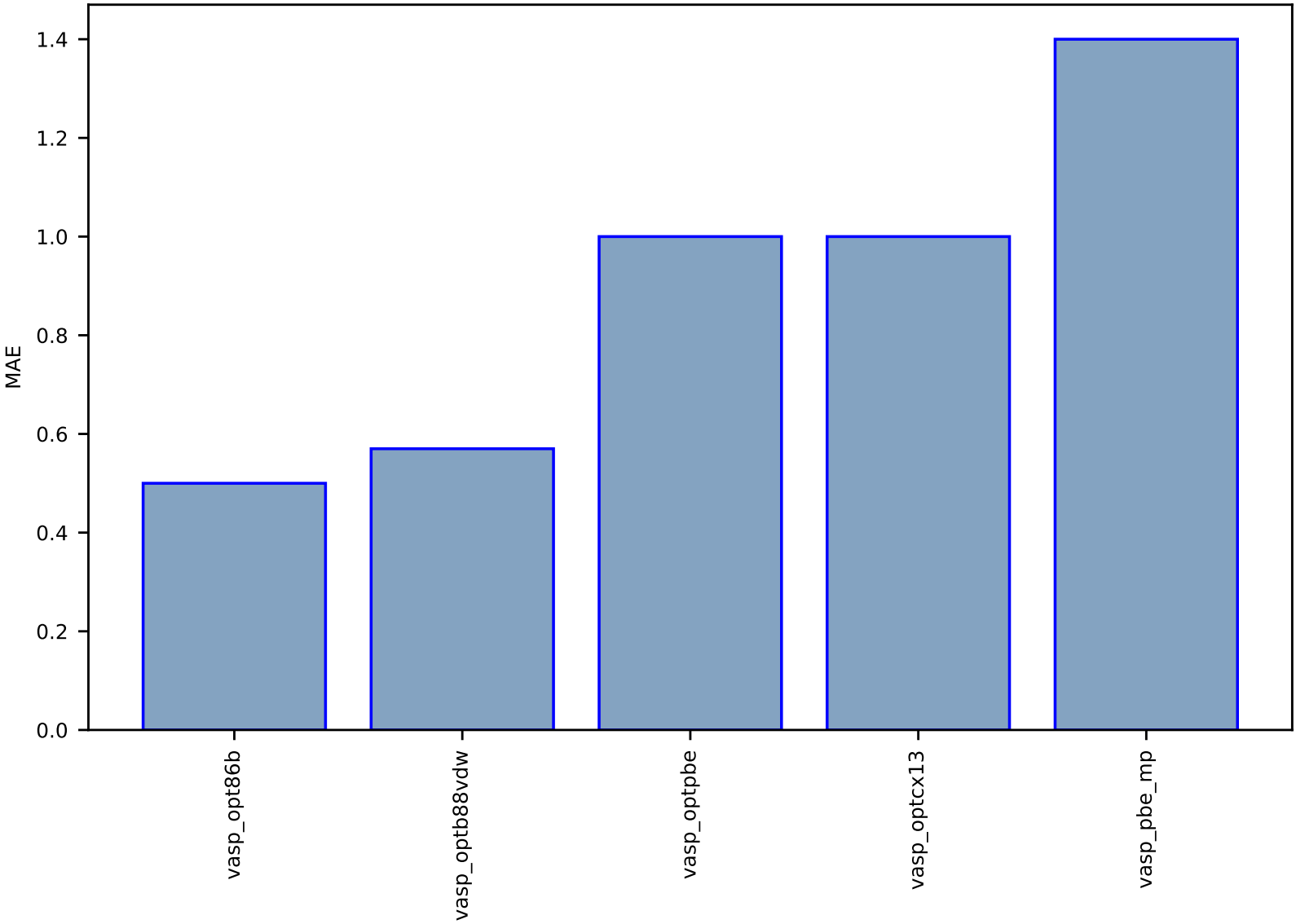


Fig S142:ES-SinglePropertyPrediction-bulk_modulus_JVASP_984_Rh-dft_3d-test-mae.csv.zip

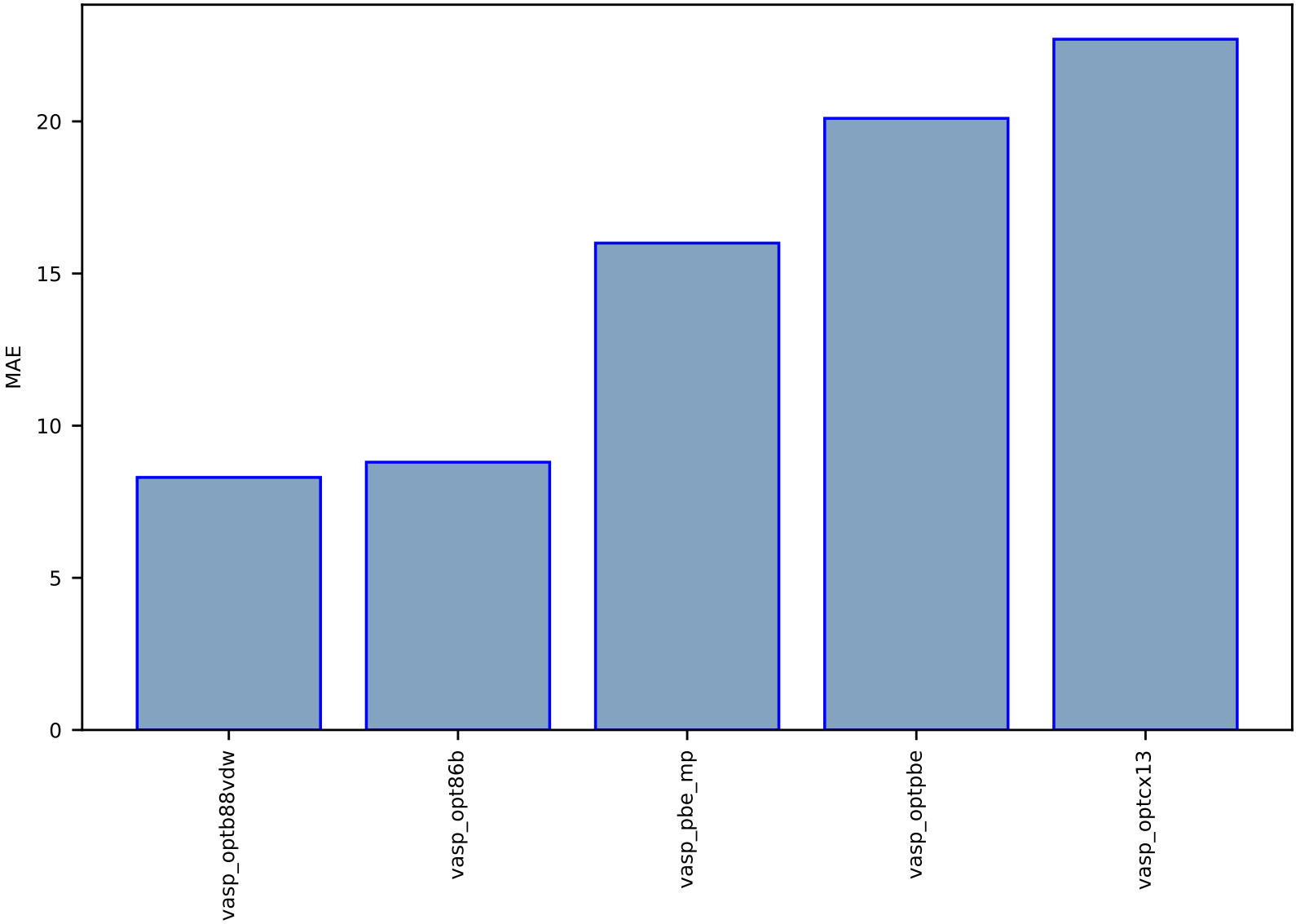


Fig S143:ES-SinglePropertyPrediction-bulk_modulus_JVASP_963_Pd-dft_3d-test-mae.csv.zip

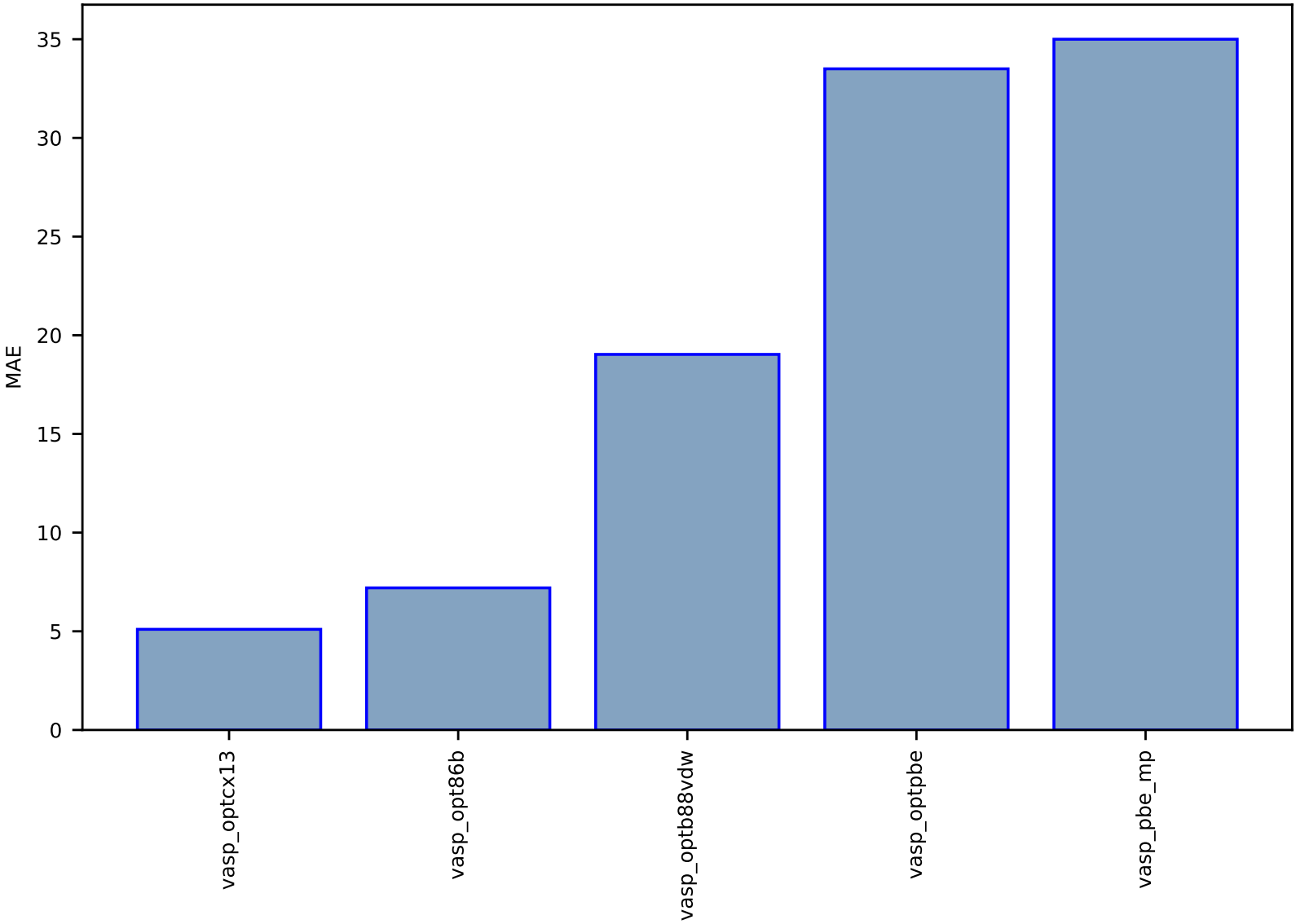


Fig S144:ES-SinglePropertyPrediction-bulk_modulus_JVASP_816_Al-dft_3d-test-mae.csv.zip

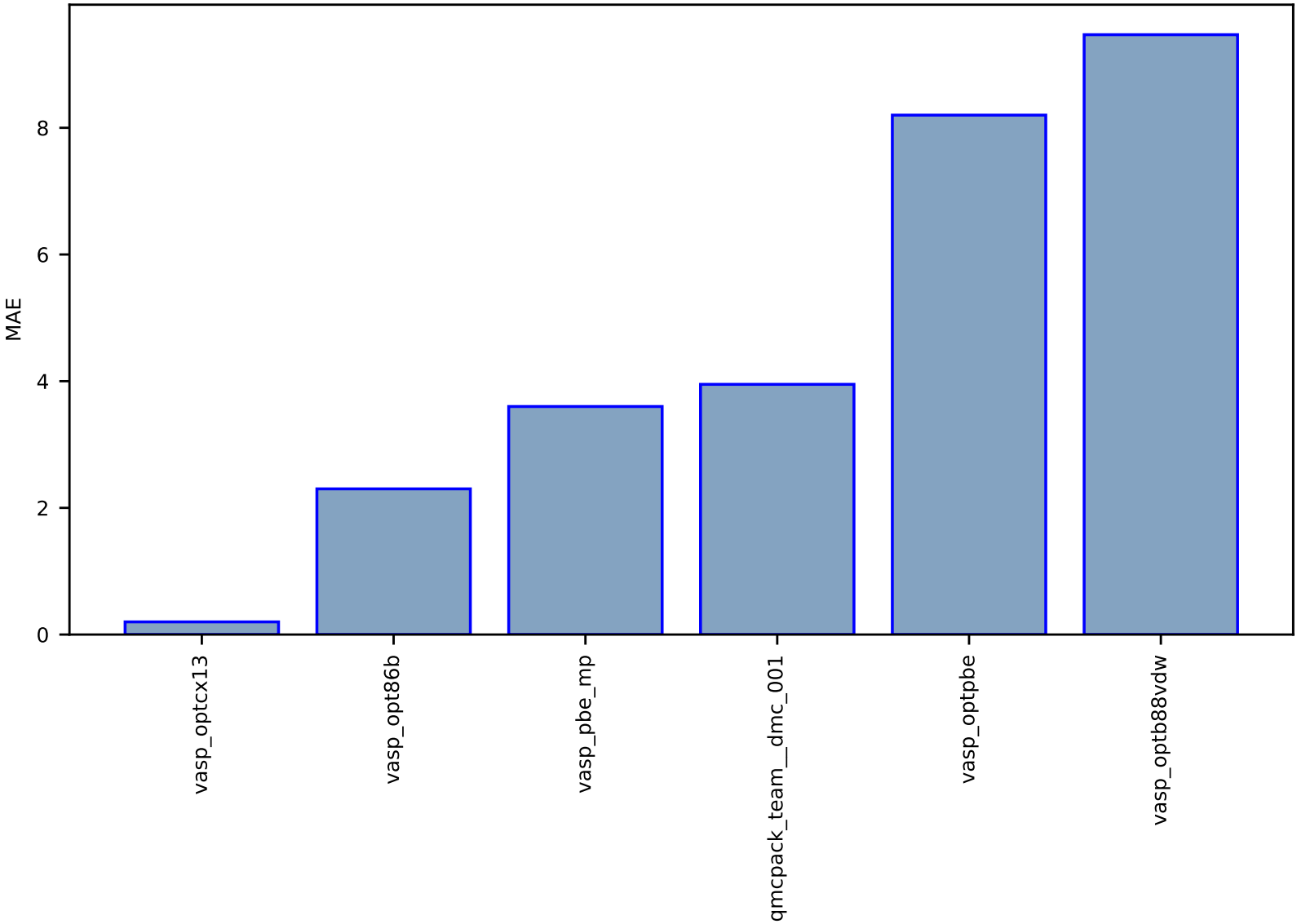


Fig S145:AI-SinglePropertyPrediction-surface_area_m2cm3-hmof-test-mae.csv.zip

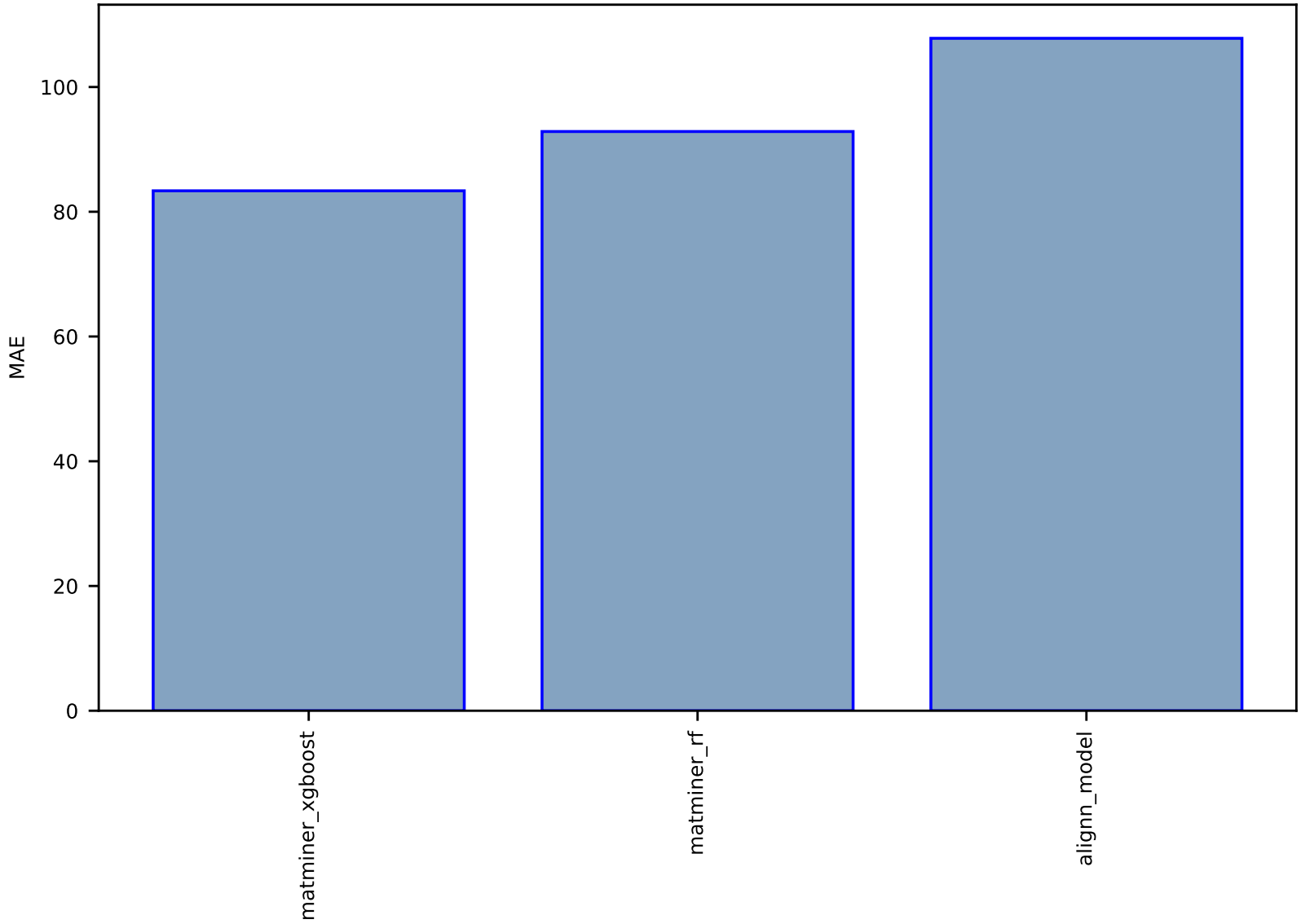


Fig S146:AI-SinglePropertyPrediction-epsy-dft_3d-test-mae.csv.zip

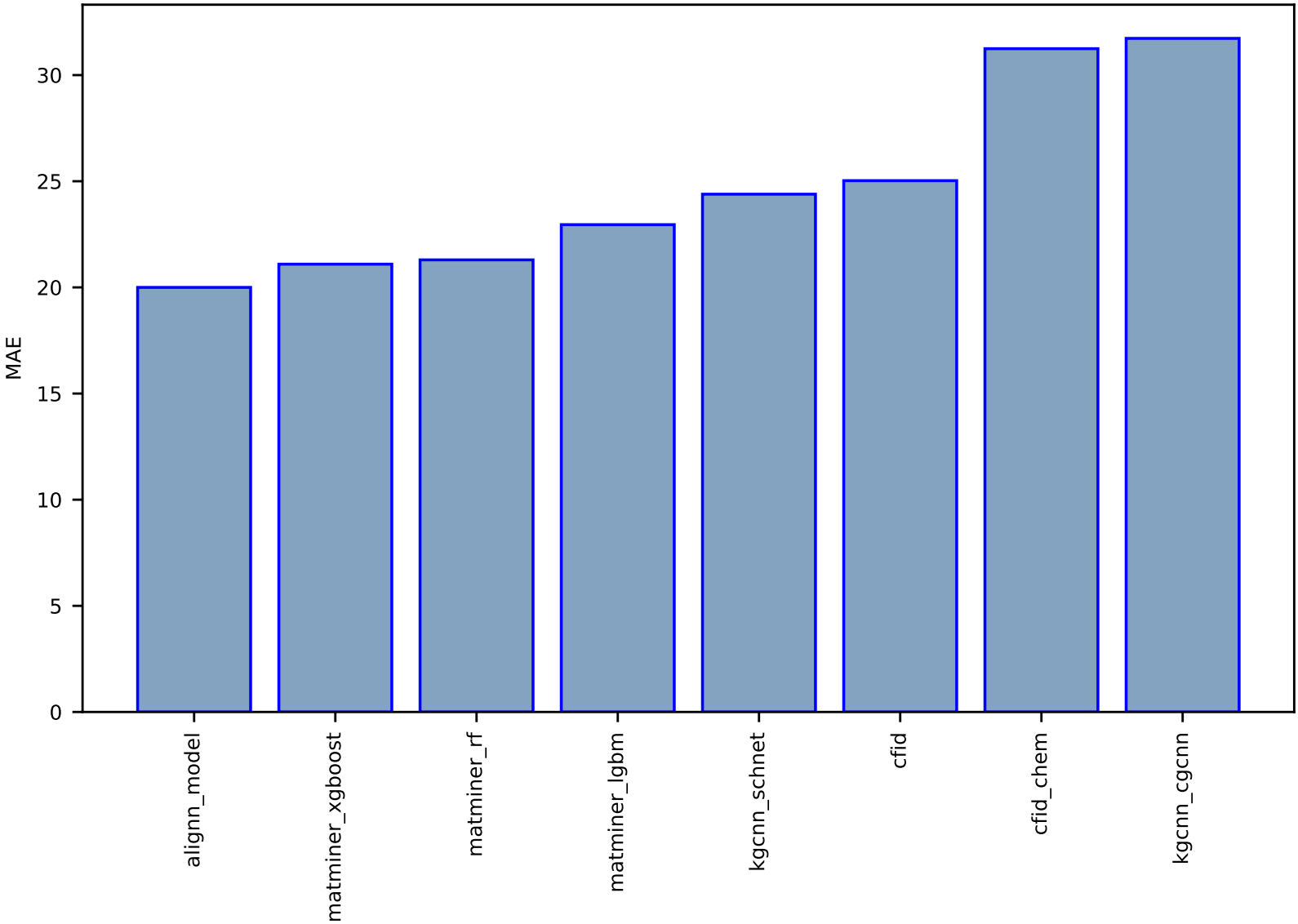


Fig S147:AI-SinglePropertyPrediction-bandgap-qmof-test-mae.csv.zip

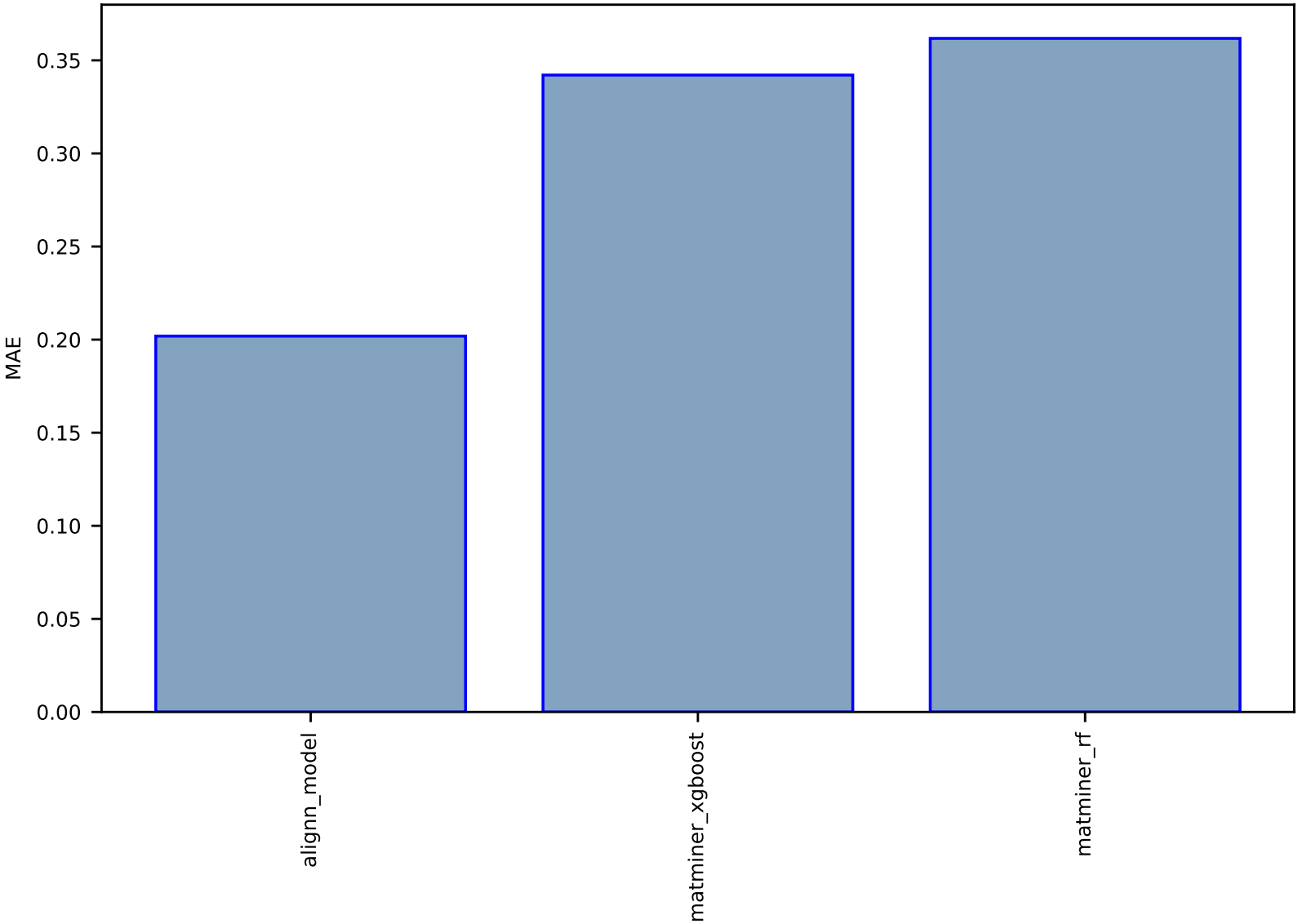


Fig S148:AI-SinglePropertyPrediction-max_co2_adsp-hmof-test-mae.csv.zip

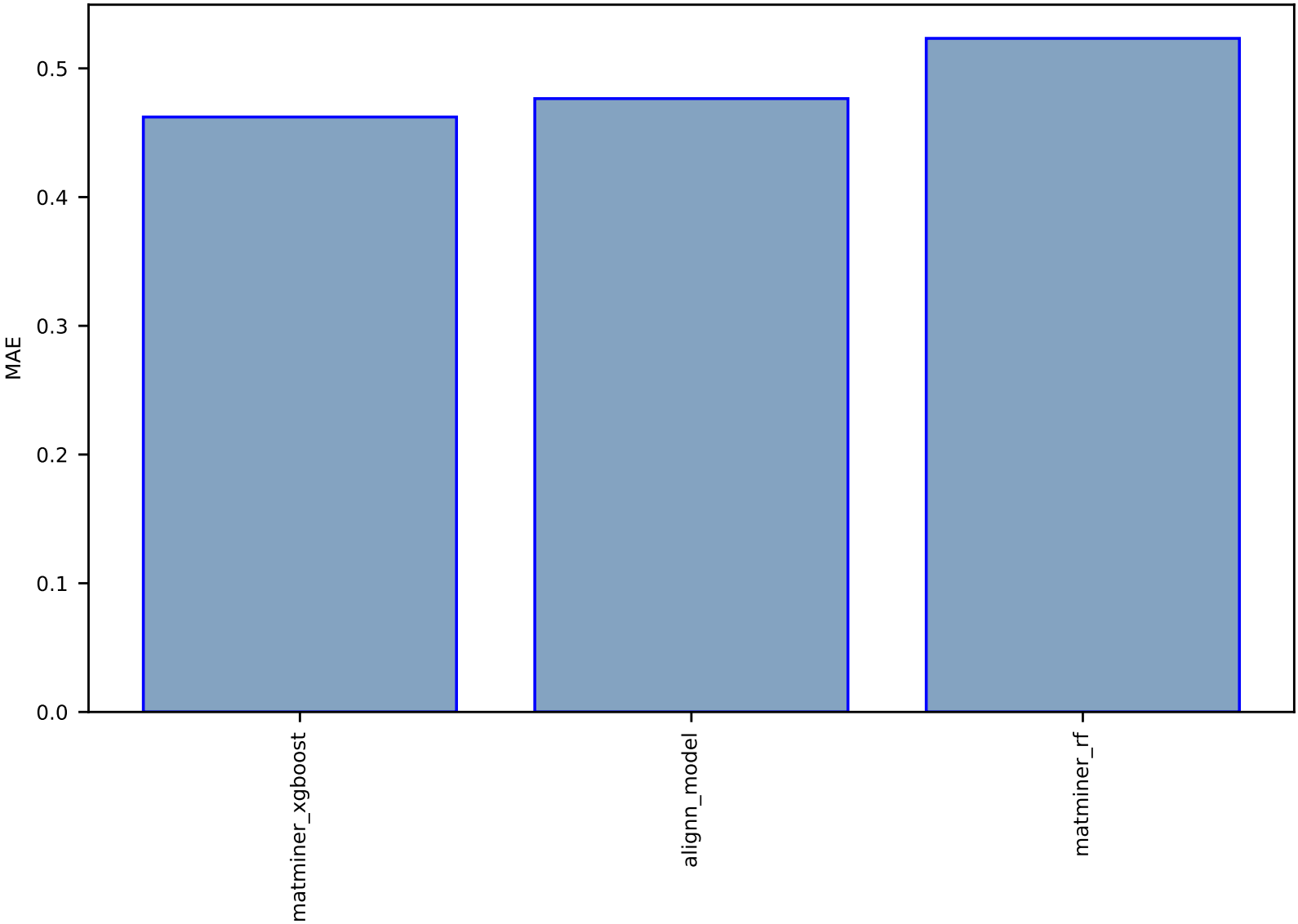


Fig S149:AI-SinglePropertyPrediction-Band_gap_HSE-snumat-test-mae.csv.zip

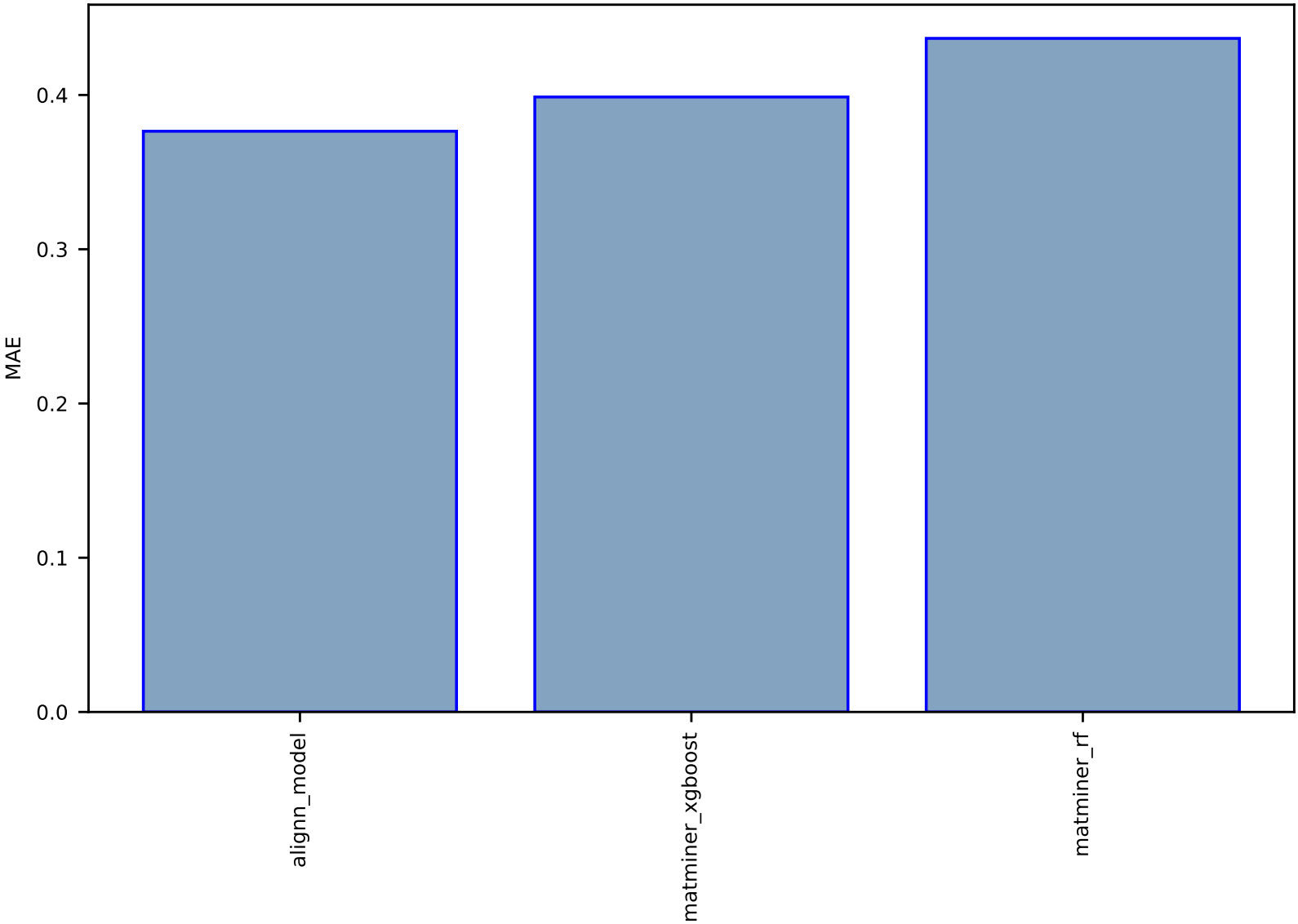


Fig S150:AI-SinglePropertyPrediction-H-qm9_std_jctc-test-mae.csv.zip

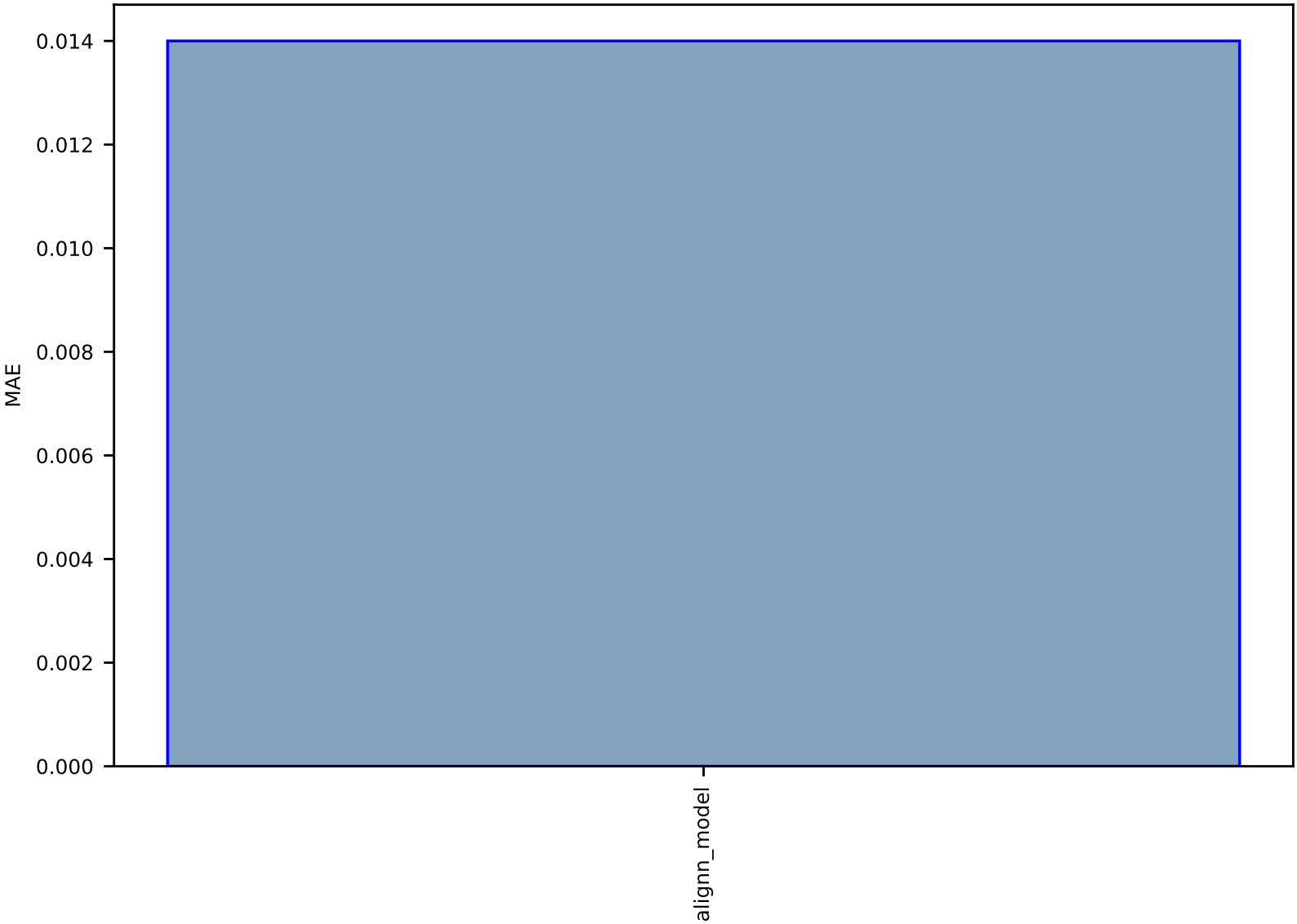


Fig S151:AI-SinglePropertyPrediction-pld-hmof-test-mae.csv.zip

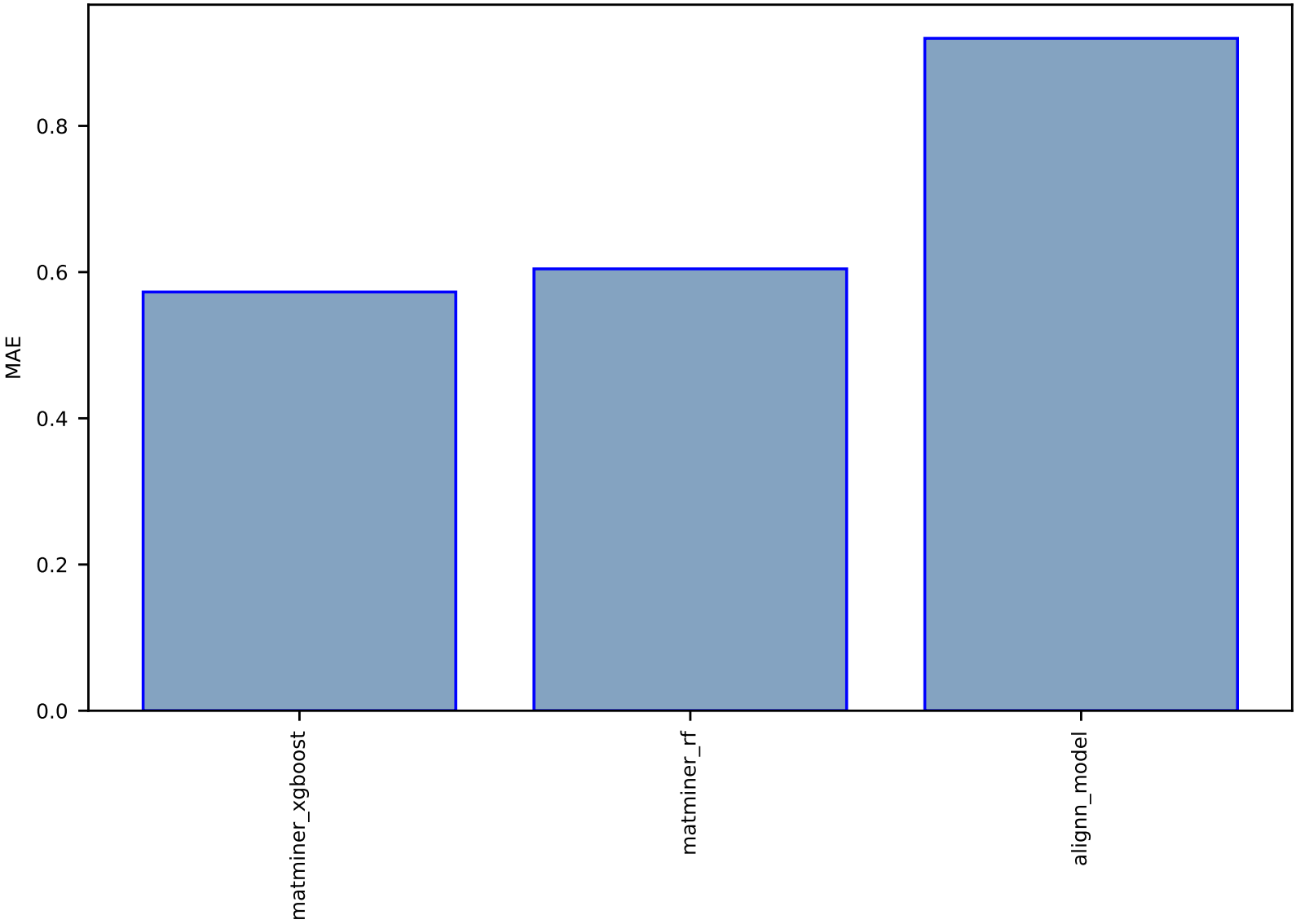


Fig S152:AI-SinglePropertyPrediction-surface_area_m2g-hmof-test-mae.csv.zip

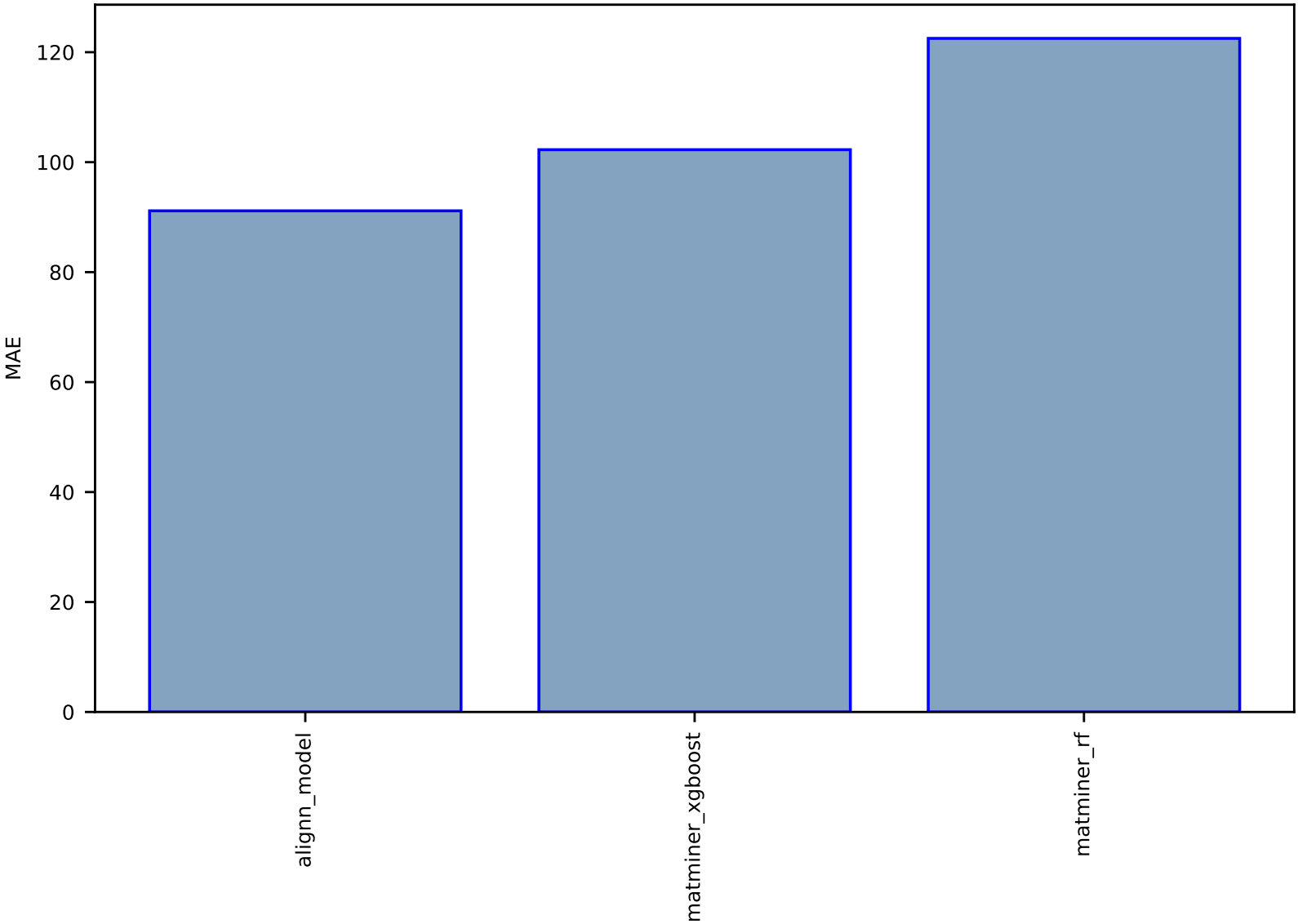


Fig S153:AI-SinglePropertyPrediction-G-qm9_std_jctc-test-mae.csv.zip

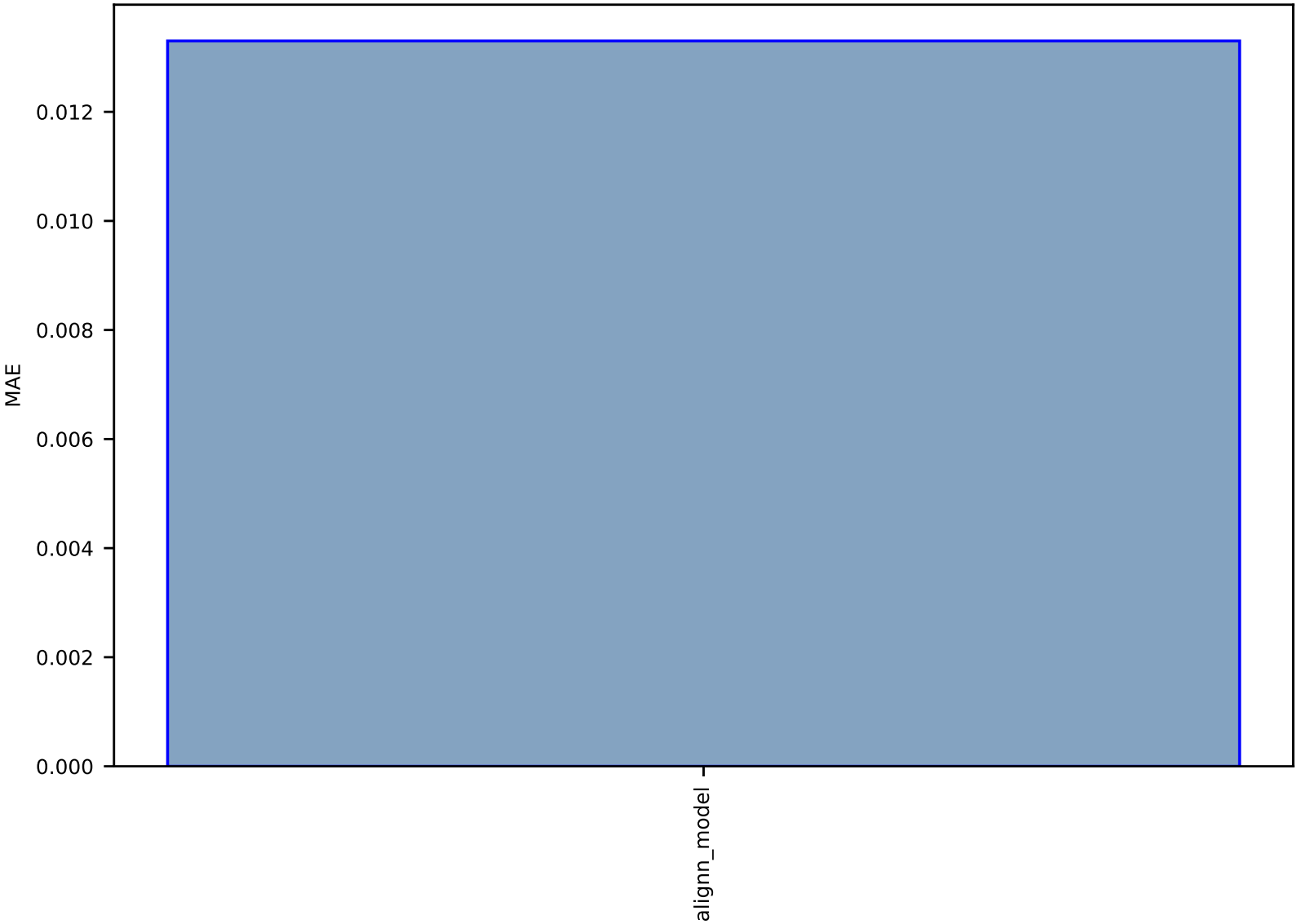


Fig S154:AI-SinglePropertyPrediction-Cv-qm9_std_jctc-test-mae.csv.zip

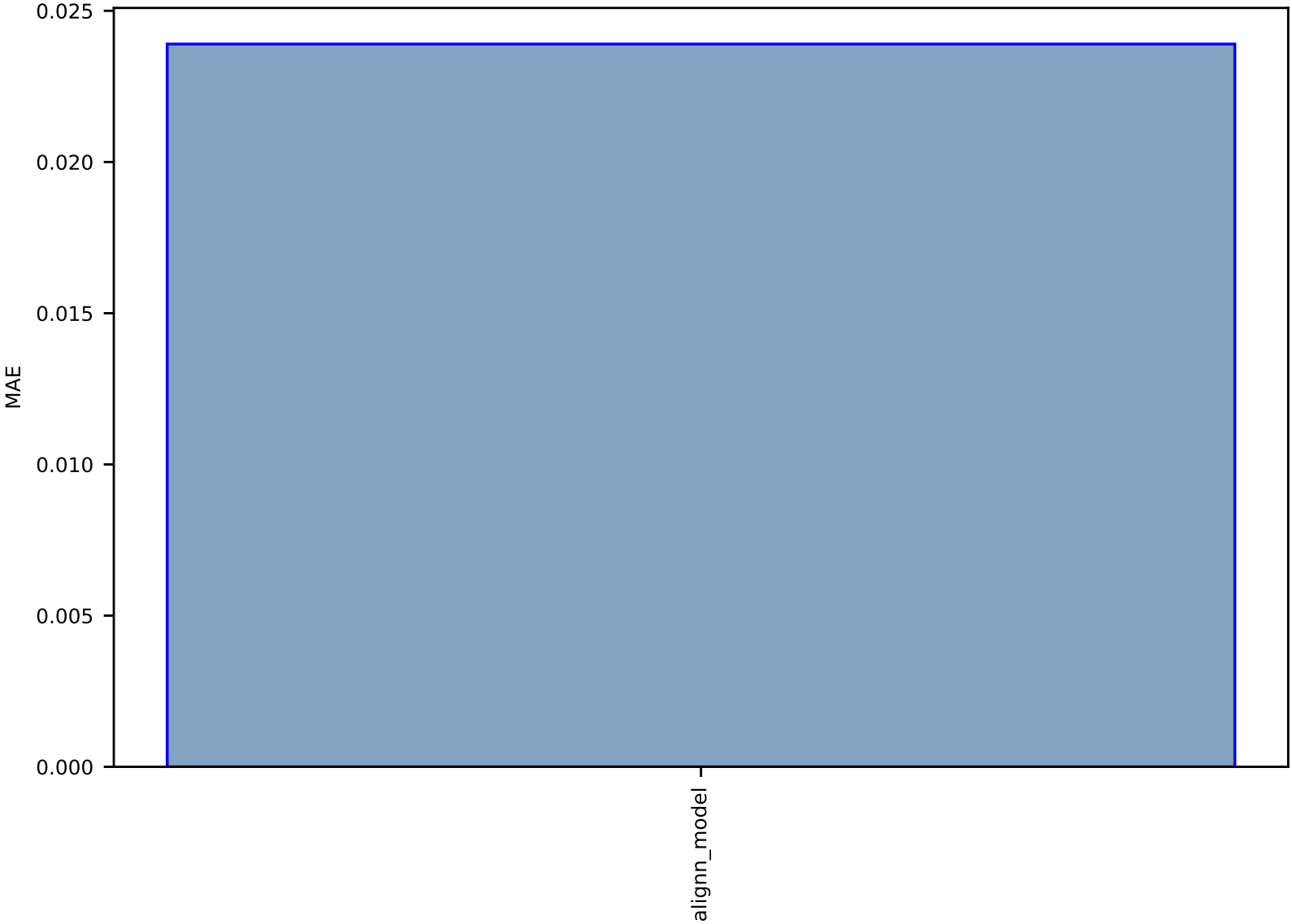


Fig S155:AI-SinglePropertyPrediction-relaxed_energy-ocp_all-test-mae.csv.zip

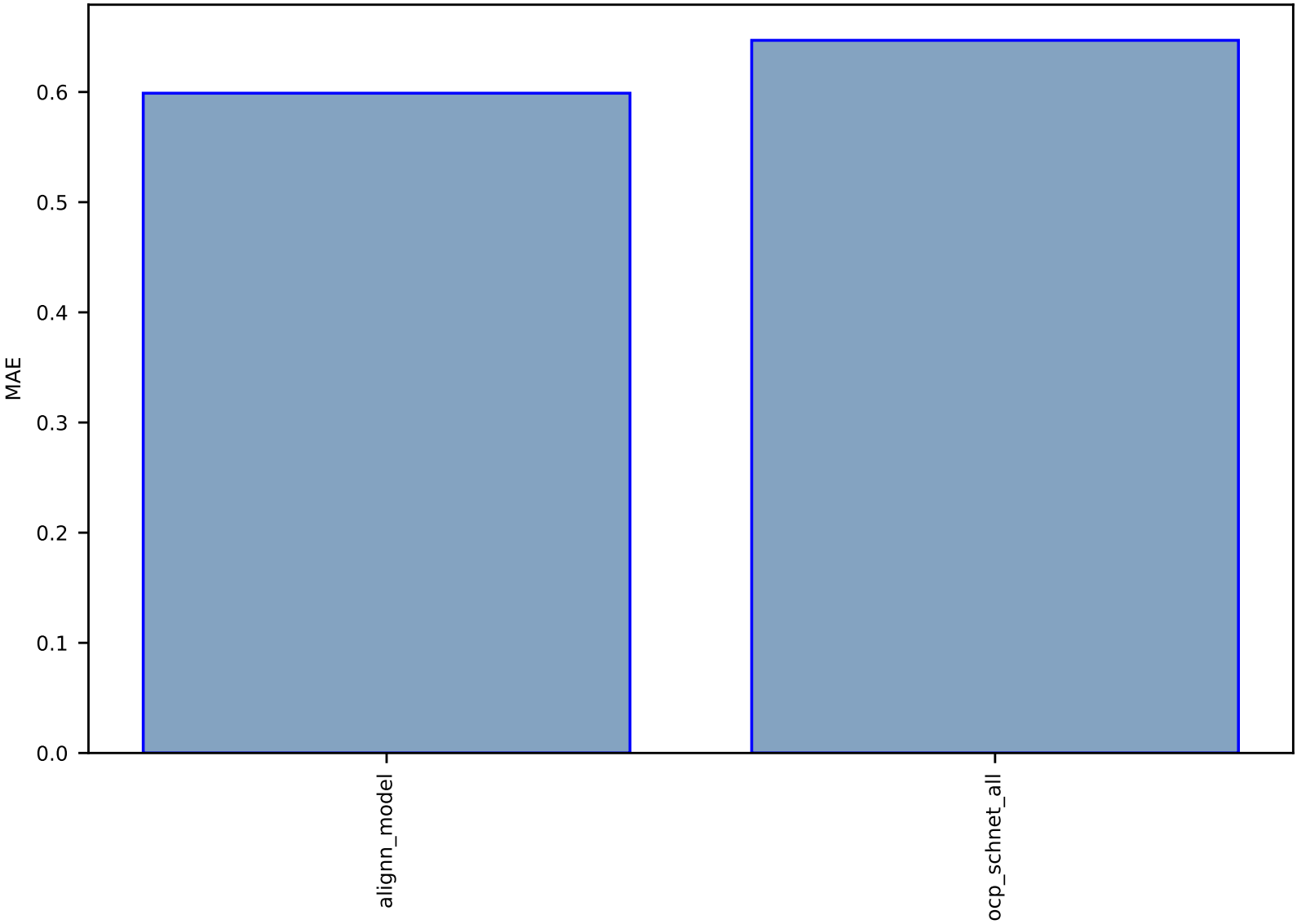


Fig S156:AI-SinglePropertyPrediction-gap-qm9_std_jctc-test-mae.csv.zip

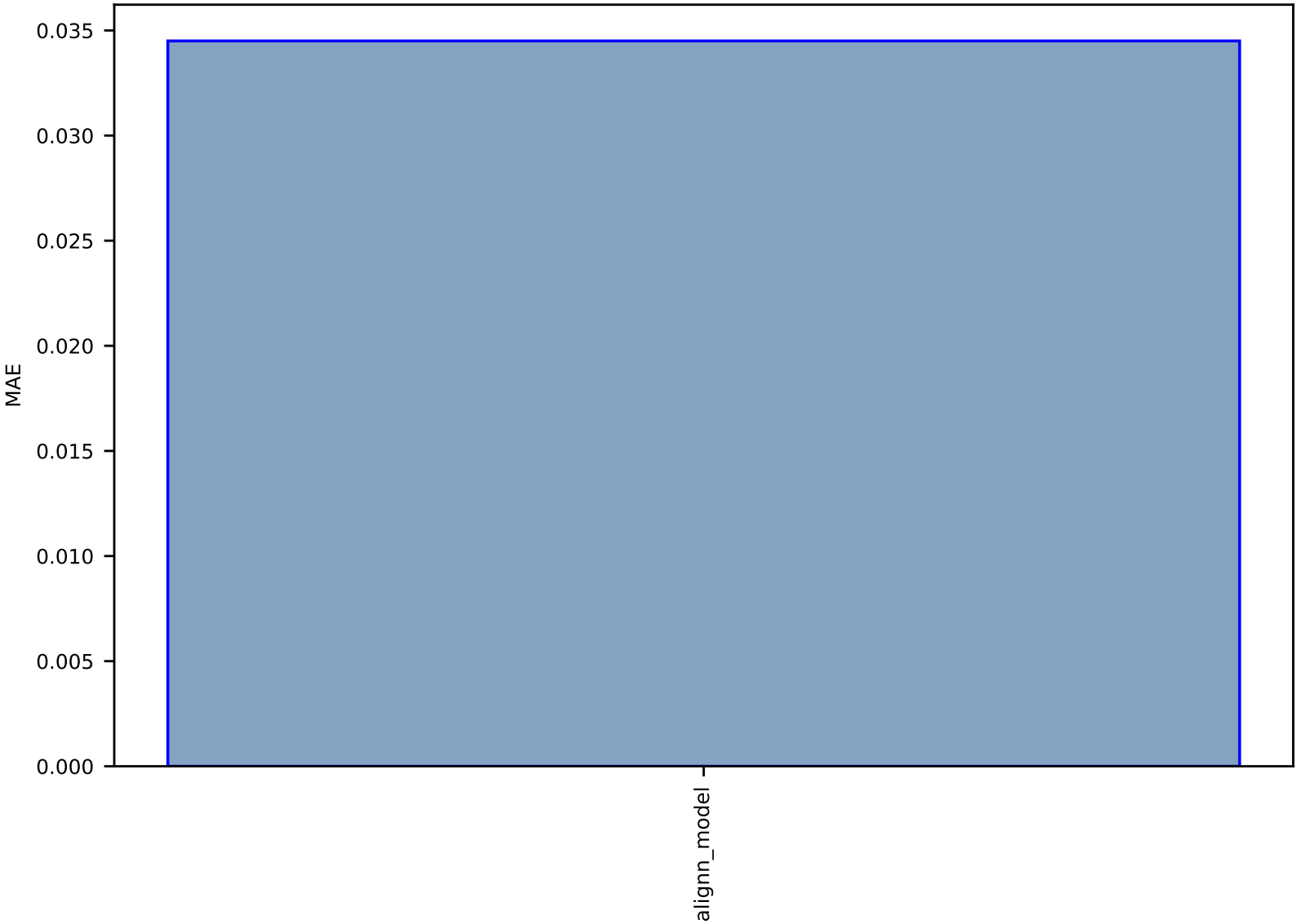


Fig S157:AI-SinglePropertyClass-optb88vdw_bandgap-dft_3d-test-acc.csv.zip

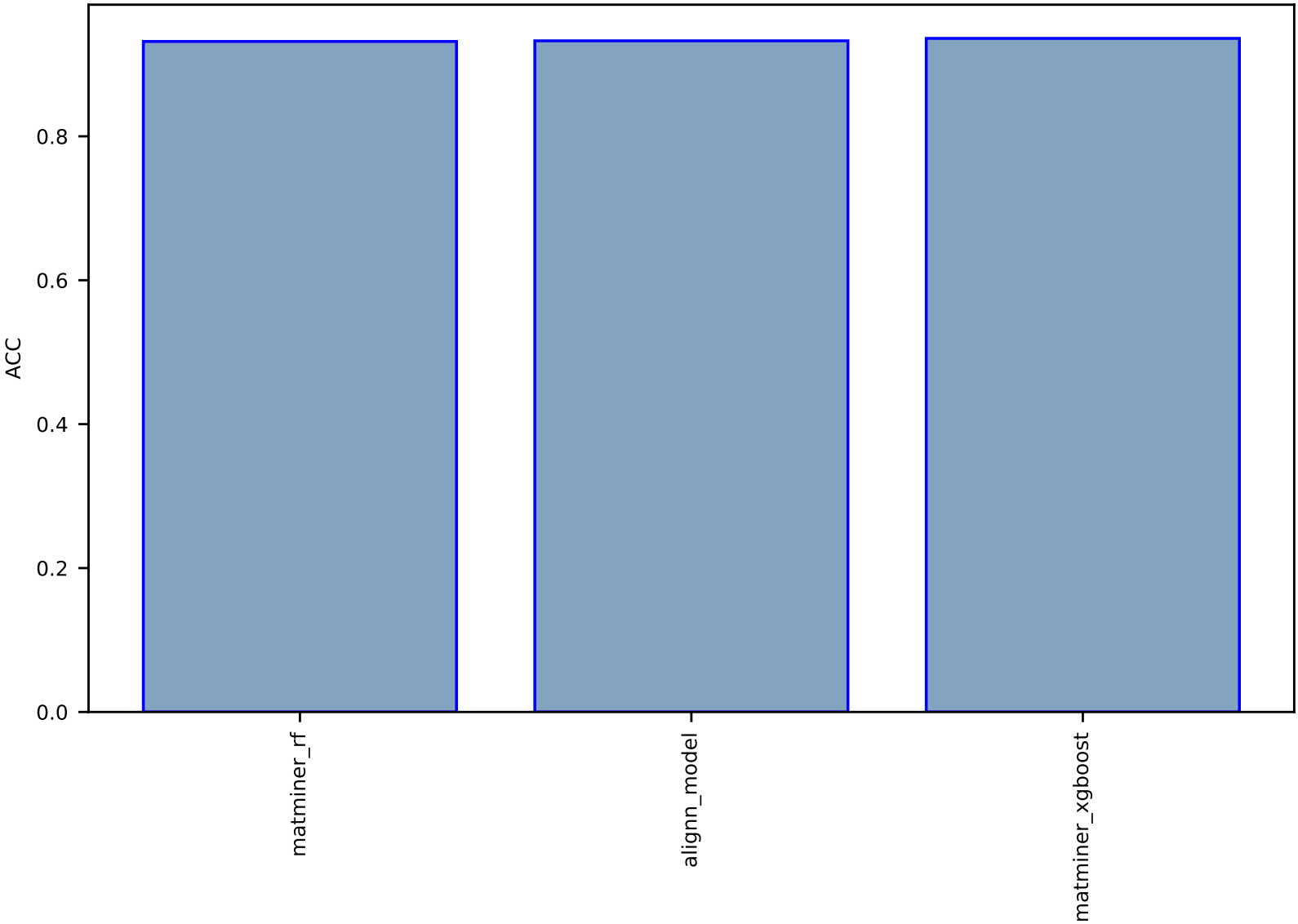


Fig S158:AI-SinglePropertyPrediction-LUMO-qm9_std_jctc-test-mae.csv.zip

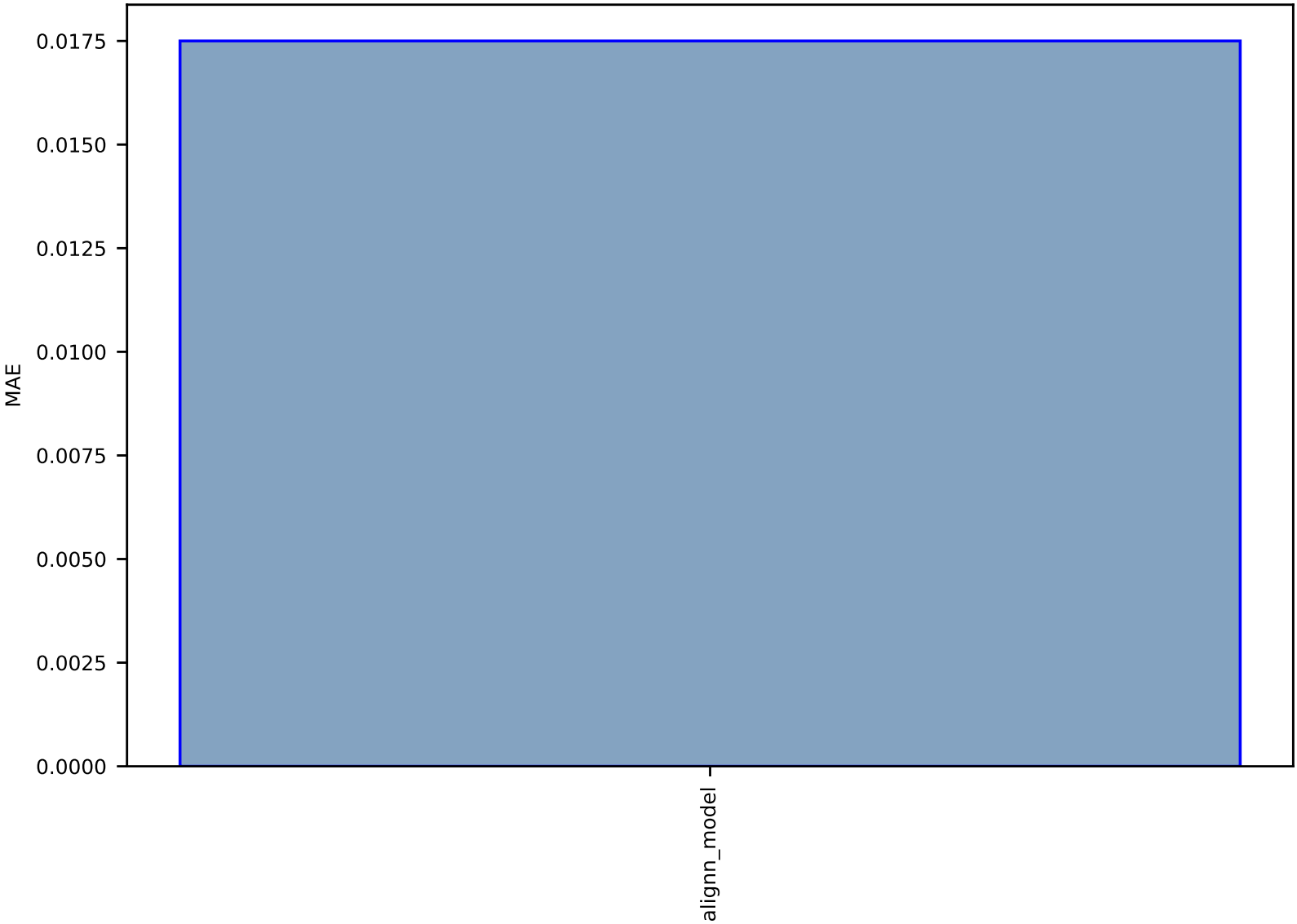


Fig S159:AI-SinglePropertyPrediction-energy_per_atom-qe_tb-test-mae.csv.zip

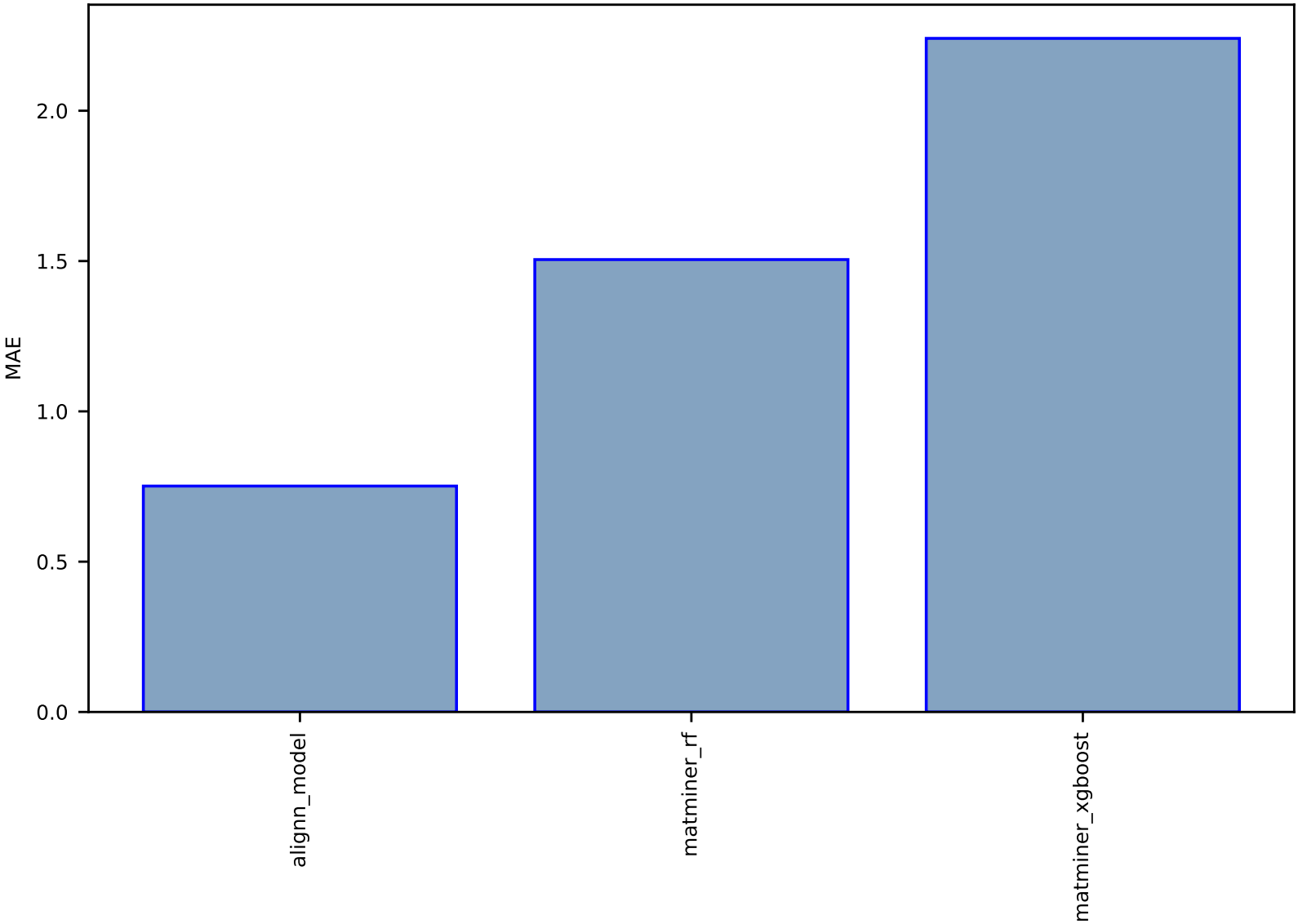


Fig S160:AI-SinglePropertyPrediction-encut-dft_3d-test-mae.csv.zip

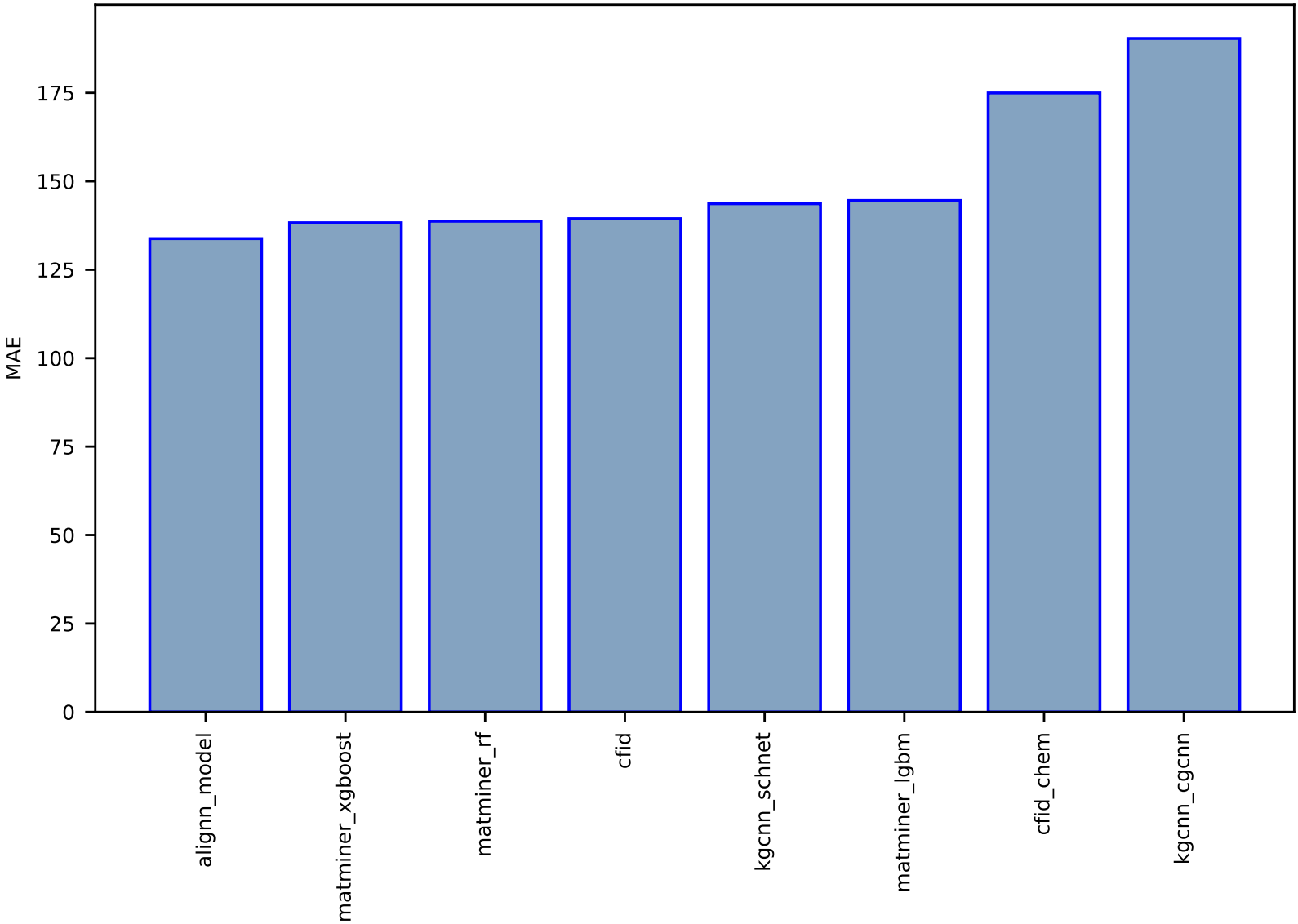


Fig S161:AI-SinglePropertyPrediction-relaxed_energy-ocp10k-test-mae.csv.zip

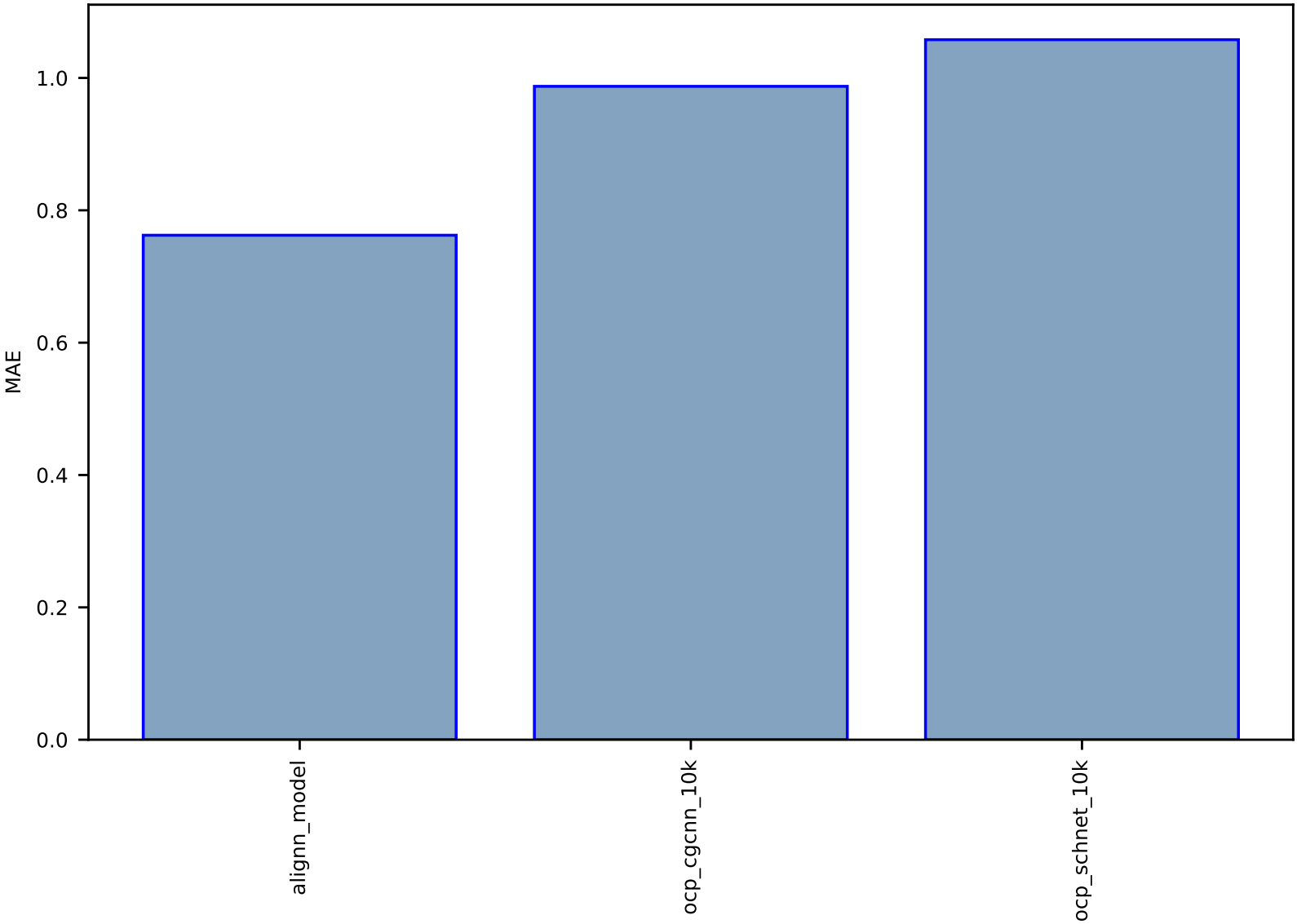


Fig S162:AI-SinglePropertyPrediction-ZPVE-qm9_std_jctc-test-mae.csv.zip

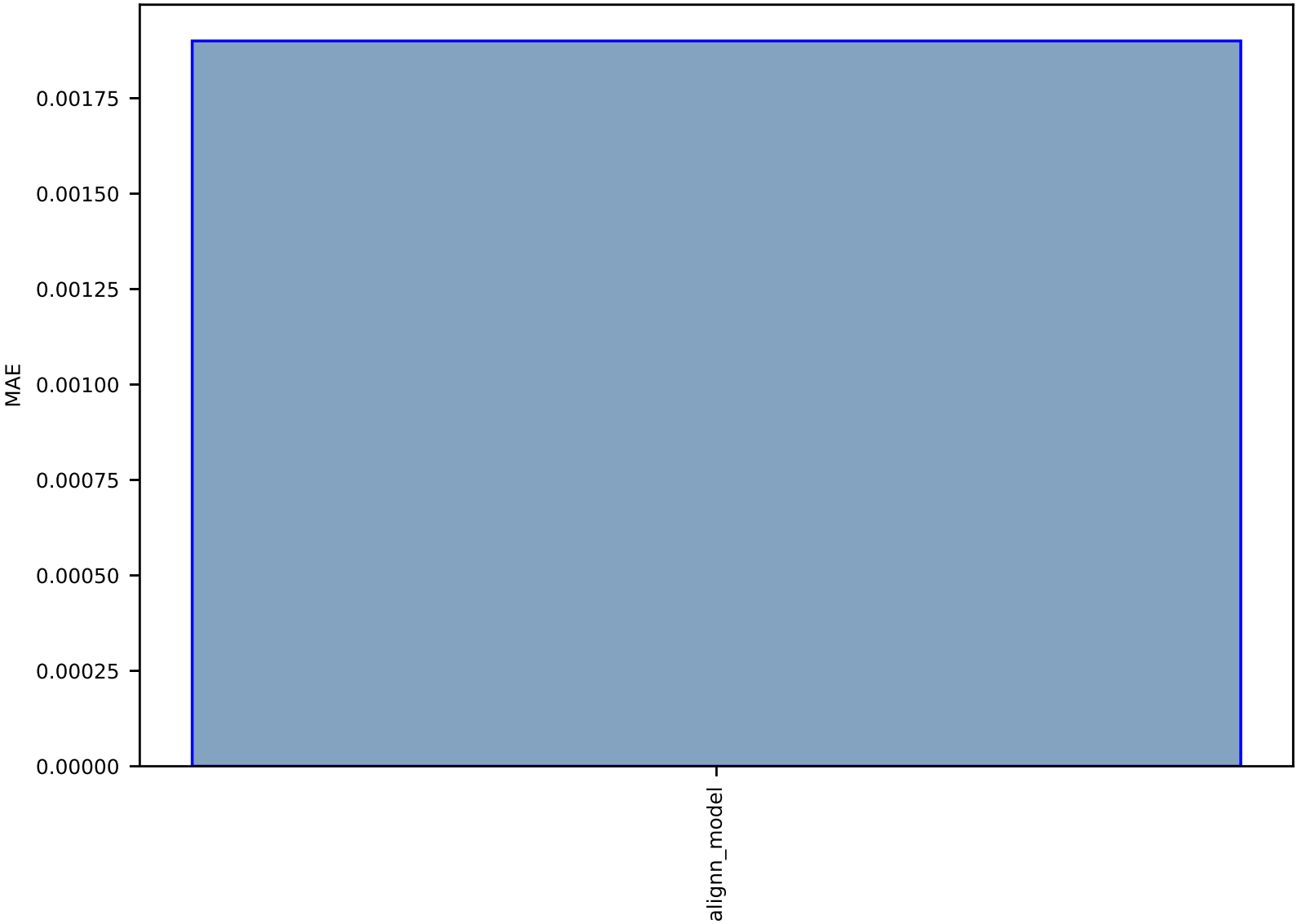


Fig S163:AI-SinglePropertyPrediction-avg_hole_mass-dft_3d-test-mae.csv.zip

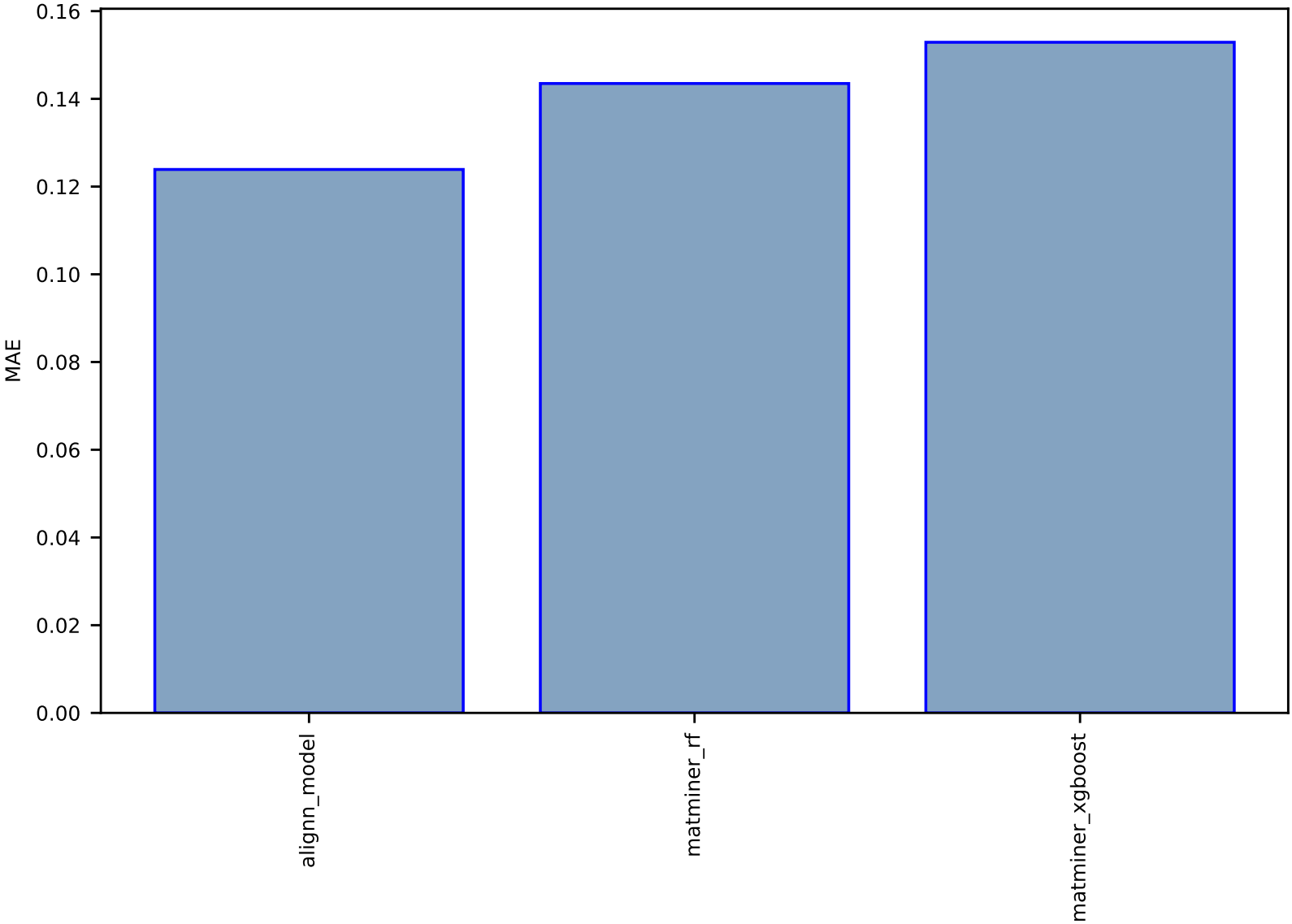


Fig S164:AI-SinglePropertyPrediction-mu-qm9_std_jctc-test-mae.csv.zip

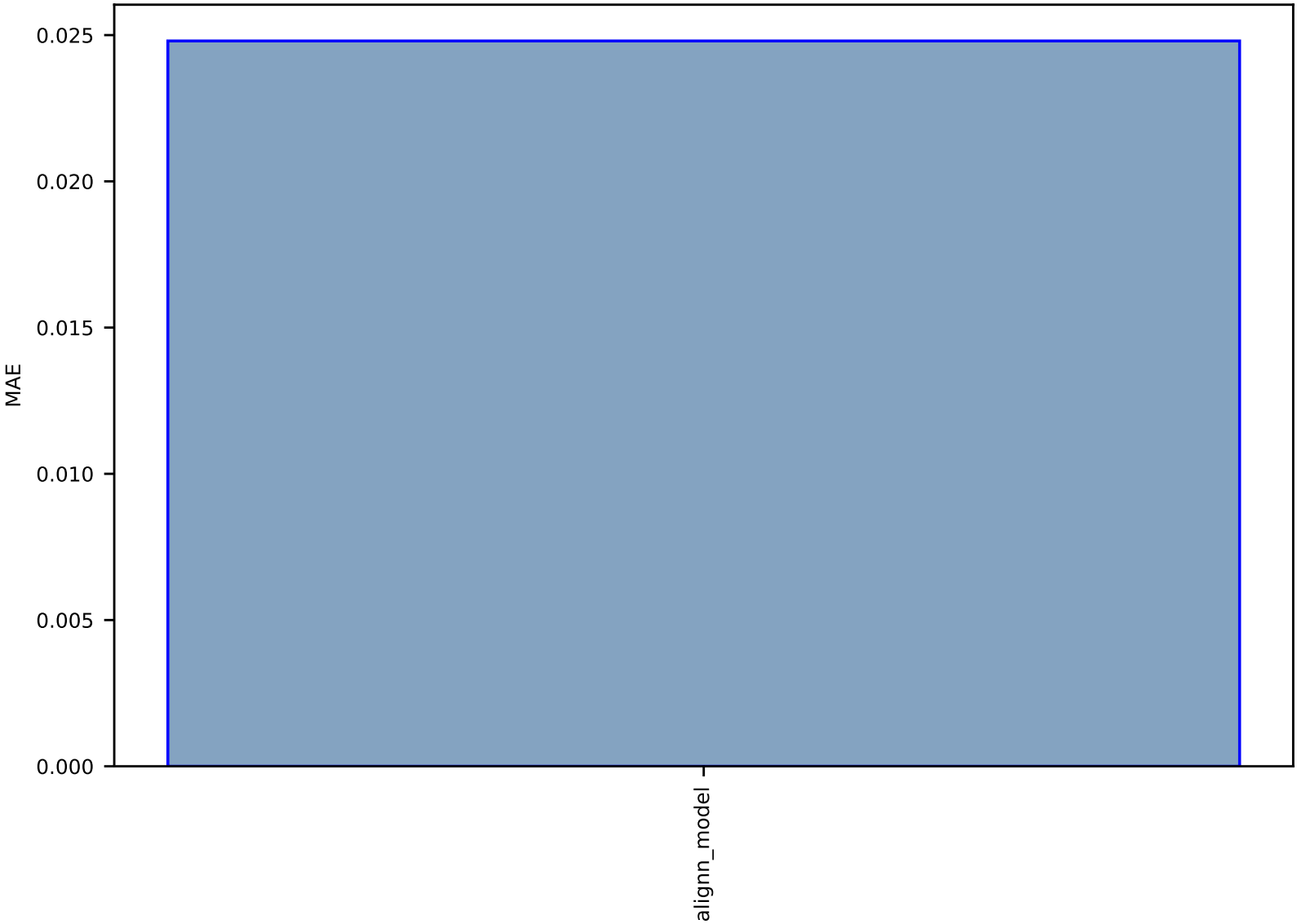


Fig S165:AI-SinglePropertyPrediction-HOMO-qm9_std_jctc-test-mae.csv.zip

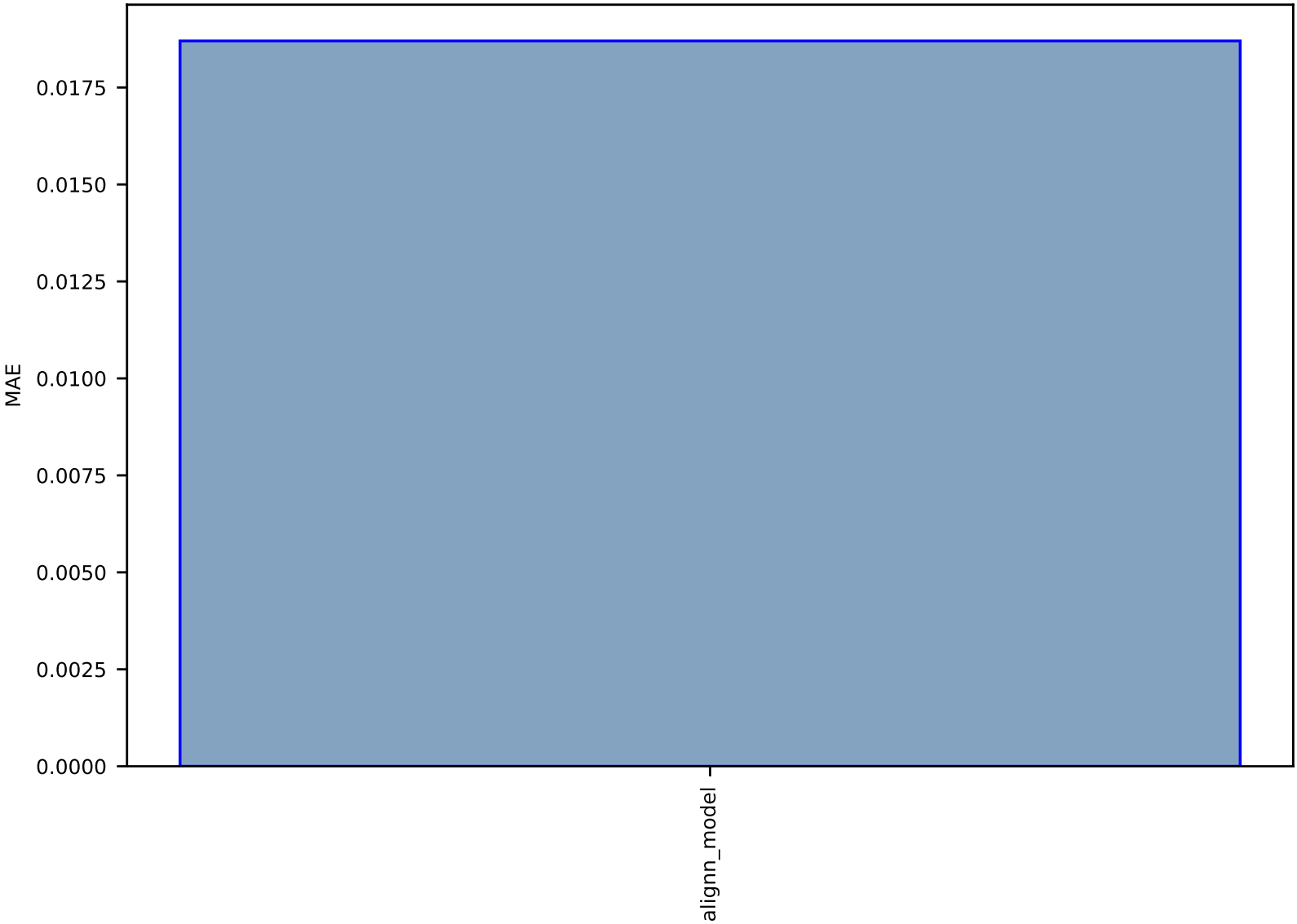


Fig S166:AI-SinglePropertyClass-magmom_oszicar-dft_3d-test-acc.csv.zip

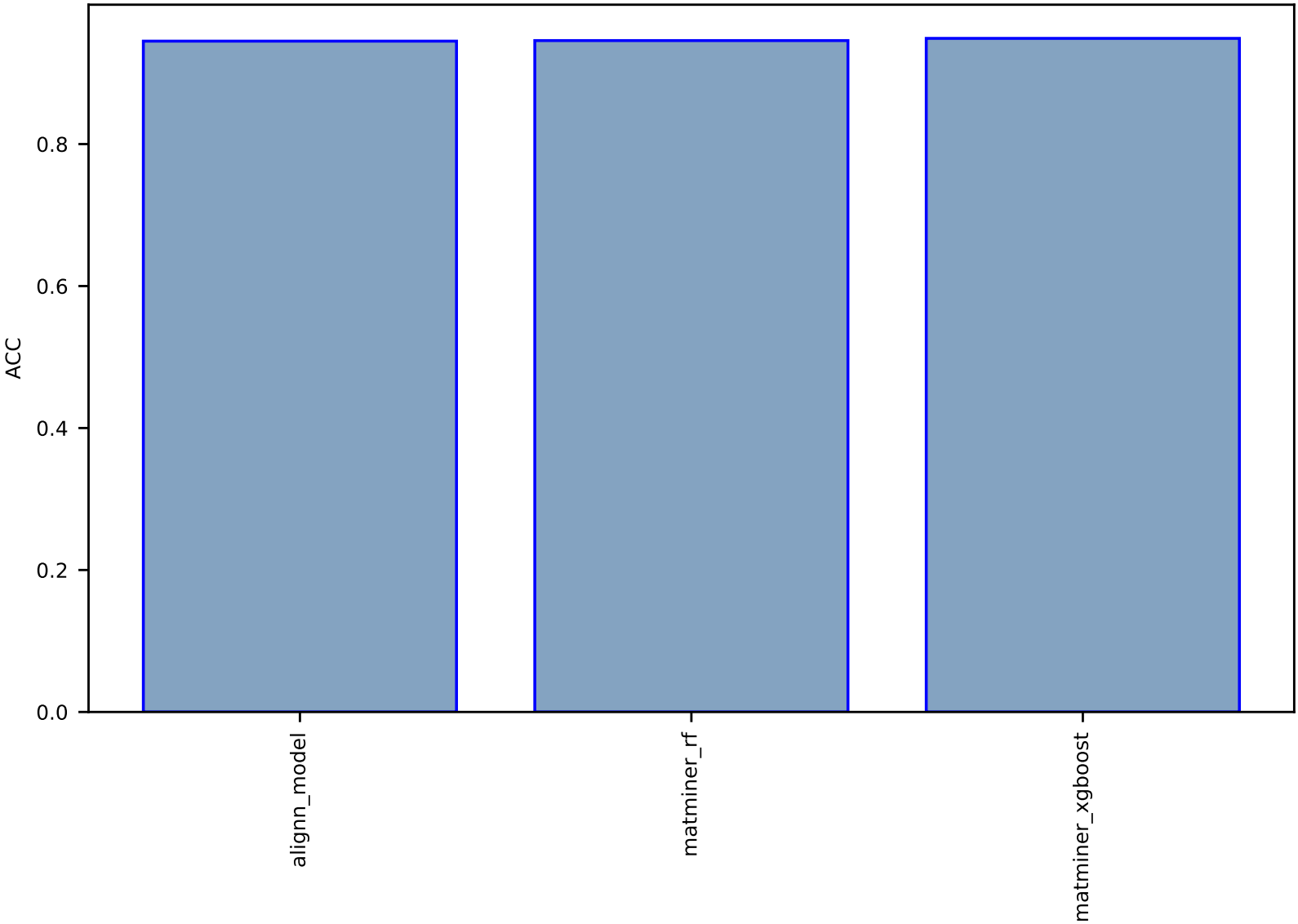


Fig S167:AI-SinglePropertyPrediction-gappbe-megnet-test-mae.csv.zip

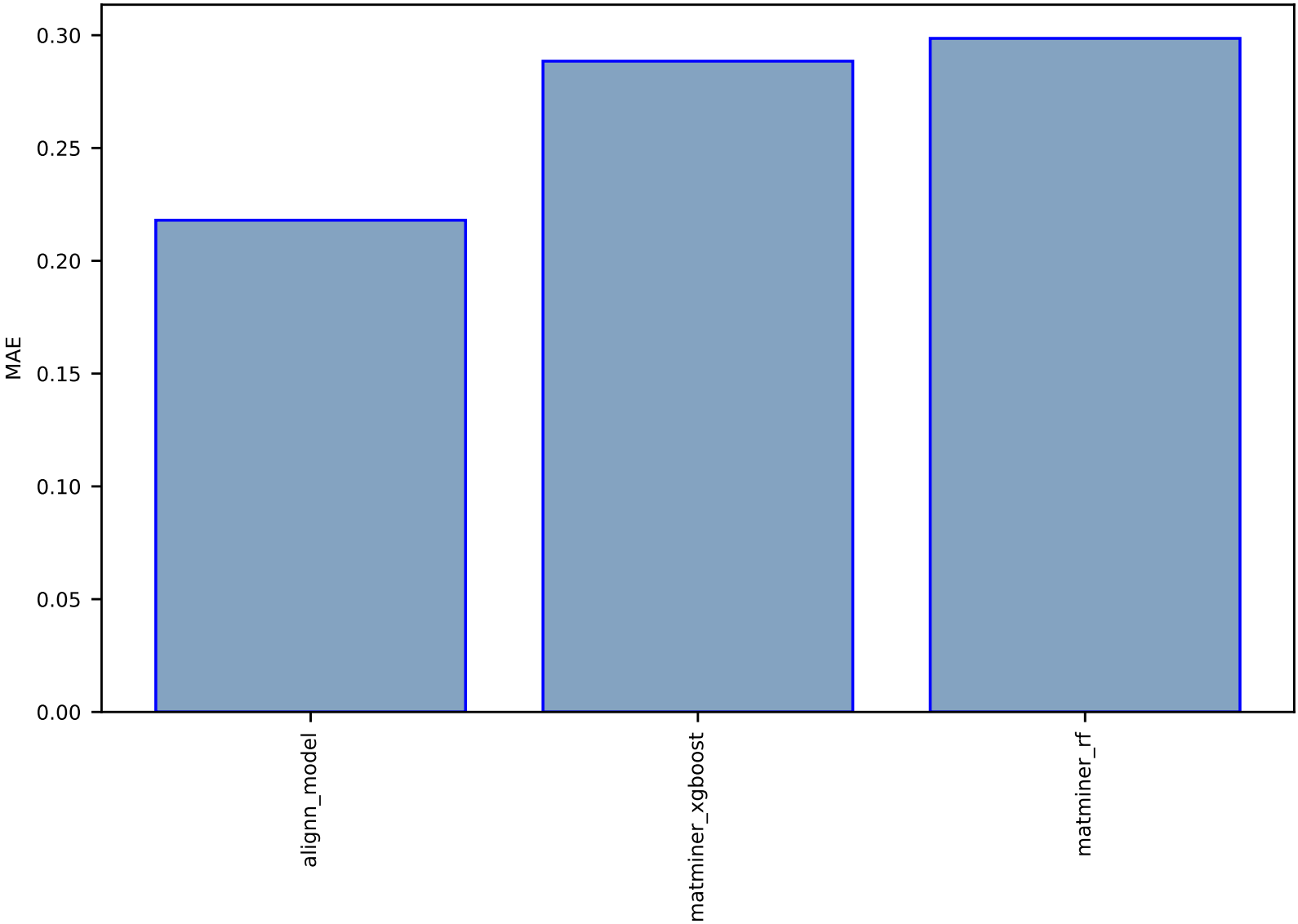


Fig S168:AI-SinglePropertyClass-n_powerfact-dft_3d-test-acc.csv.zip

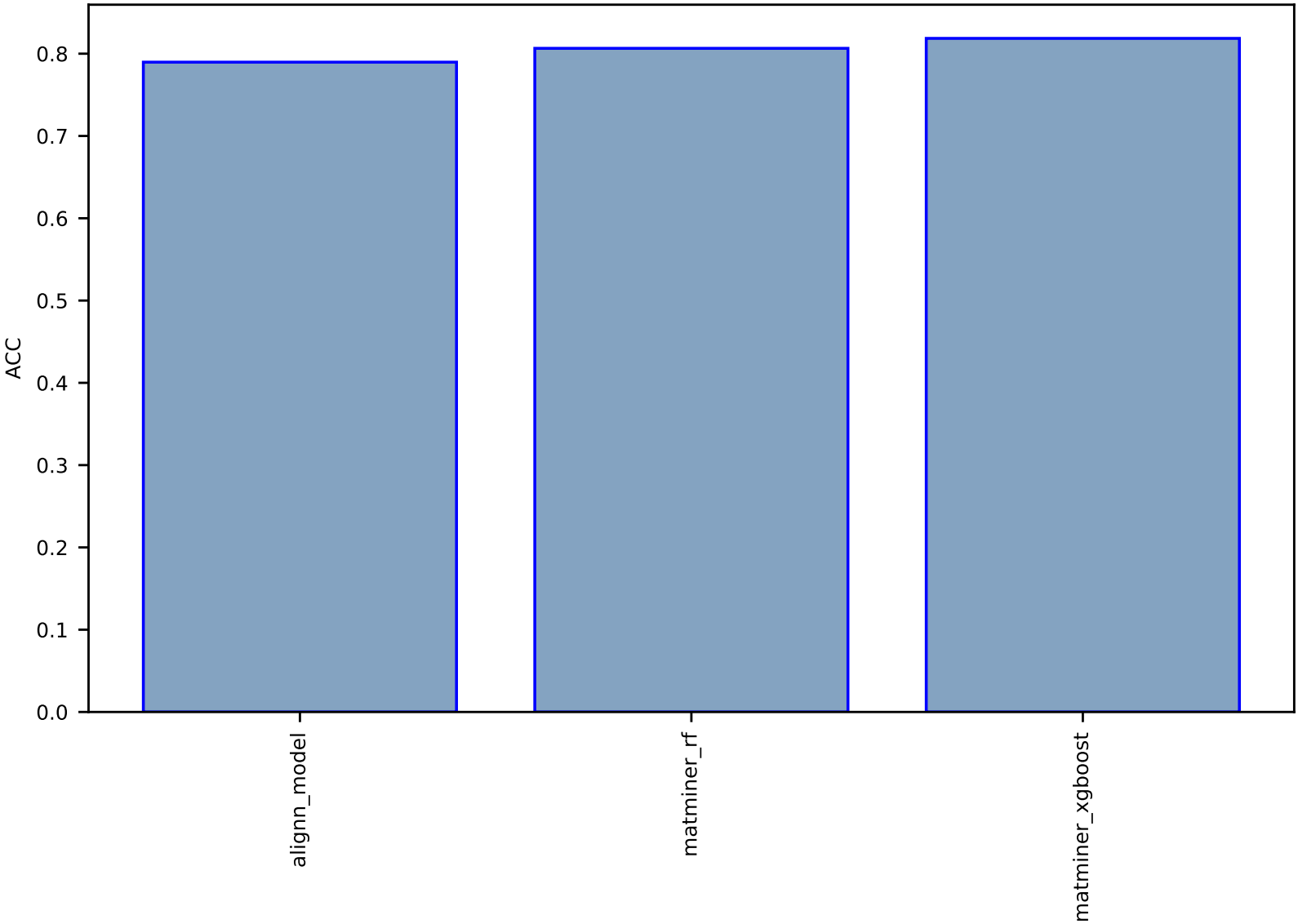


Fig S169:AI-SinglePropertyPrediction-f_enp-qe_tb-test-mae.csv.zip

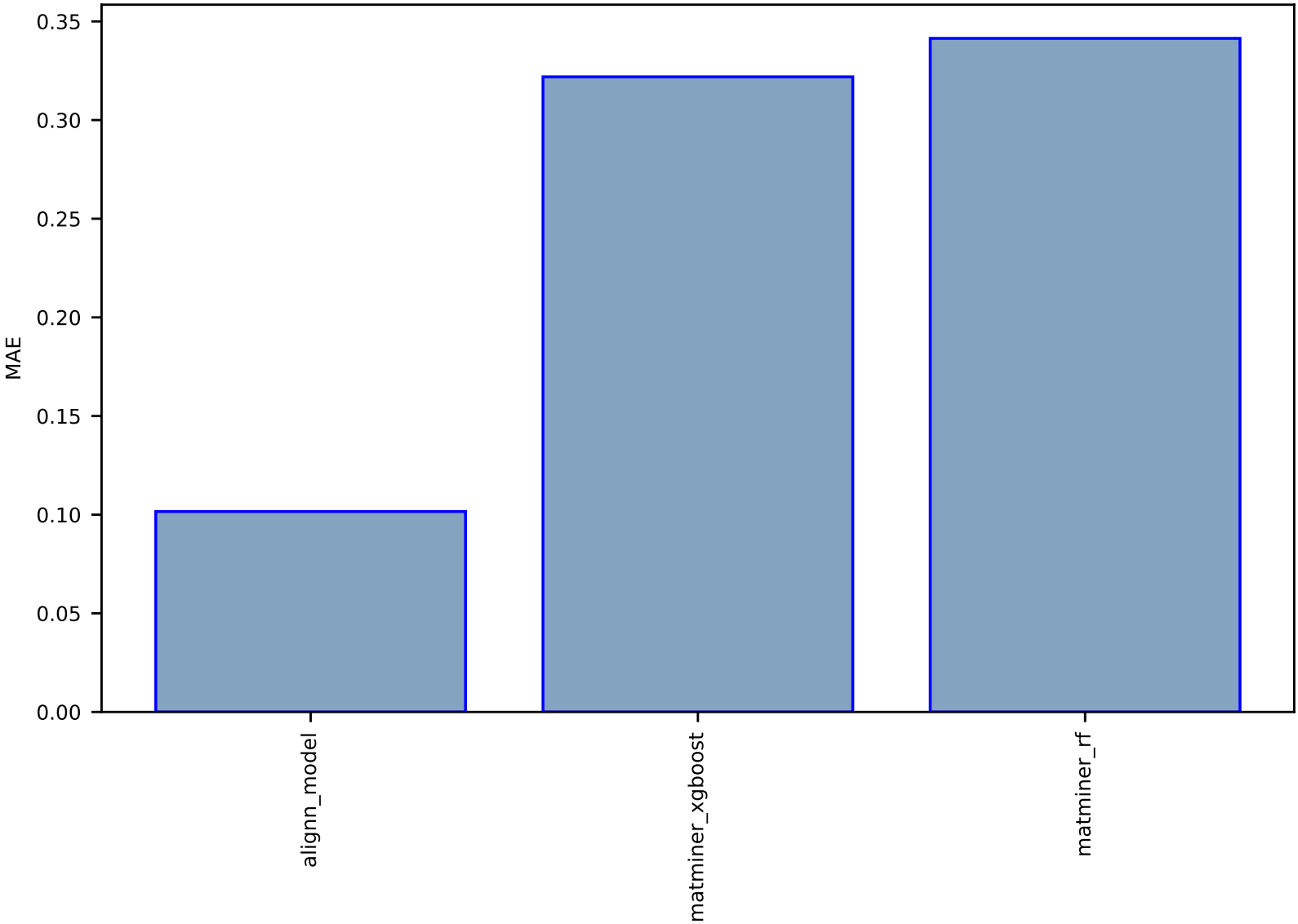


Fig S170:AI-SinglePropertyClass-p_Seebeck-dft_3d-test-acc.csv.zip

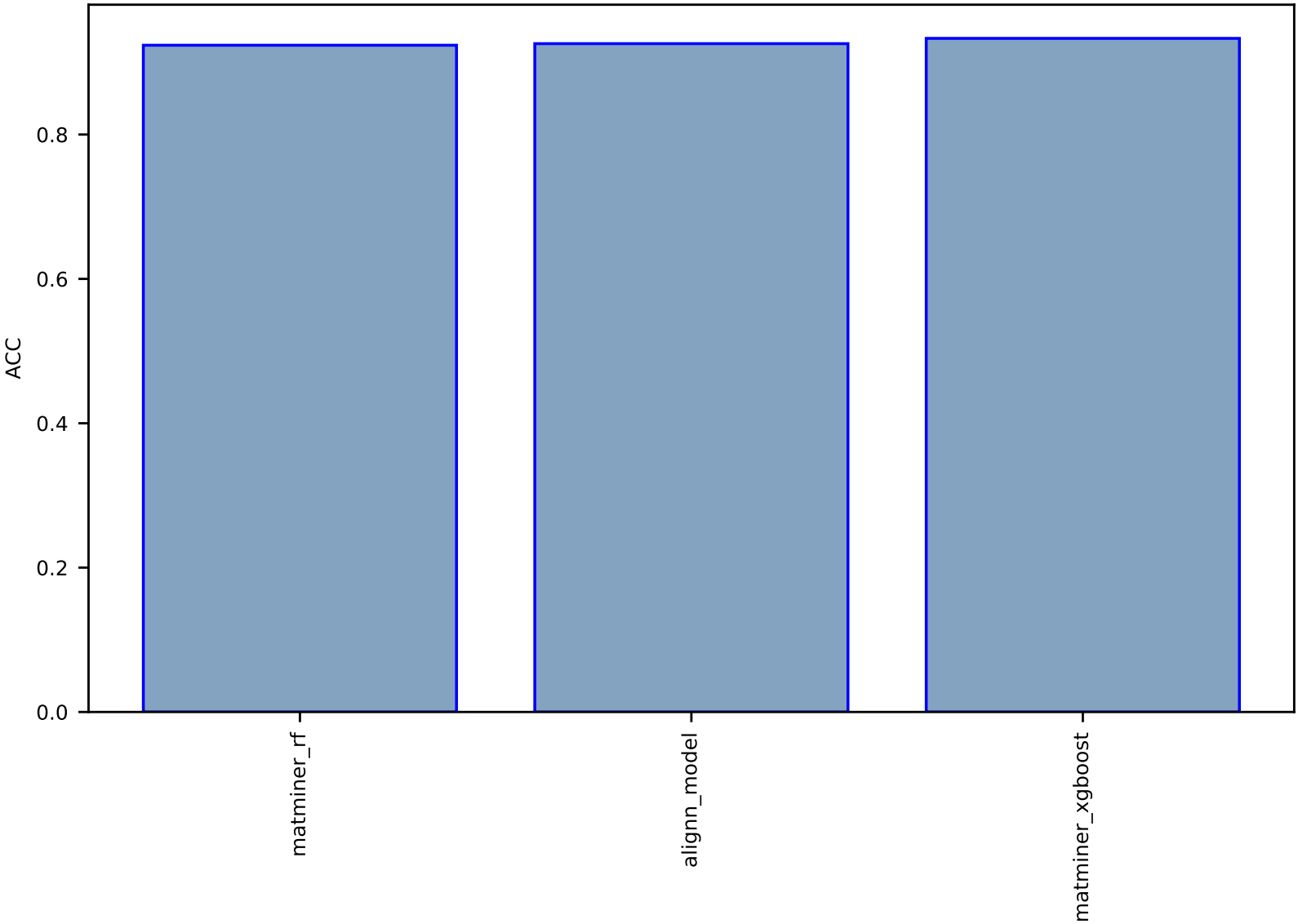


Fig S171:AI-SinglePropertyPrediction-e_form-megnet-test-mae.csv.zip

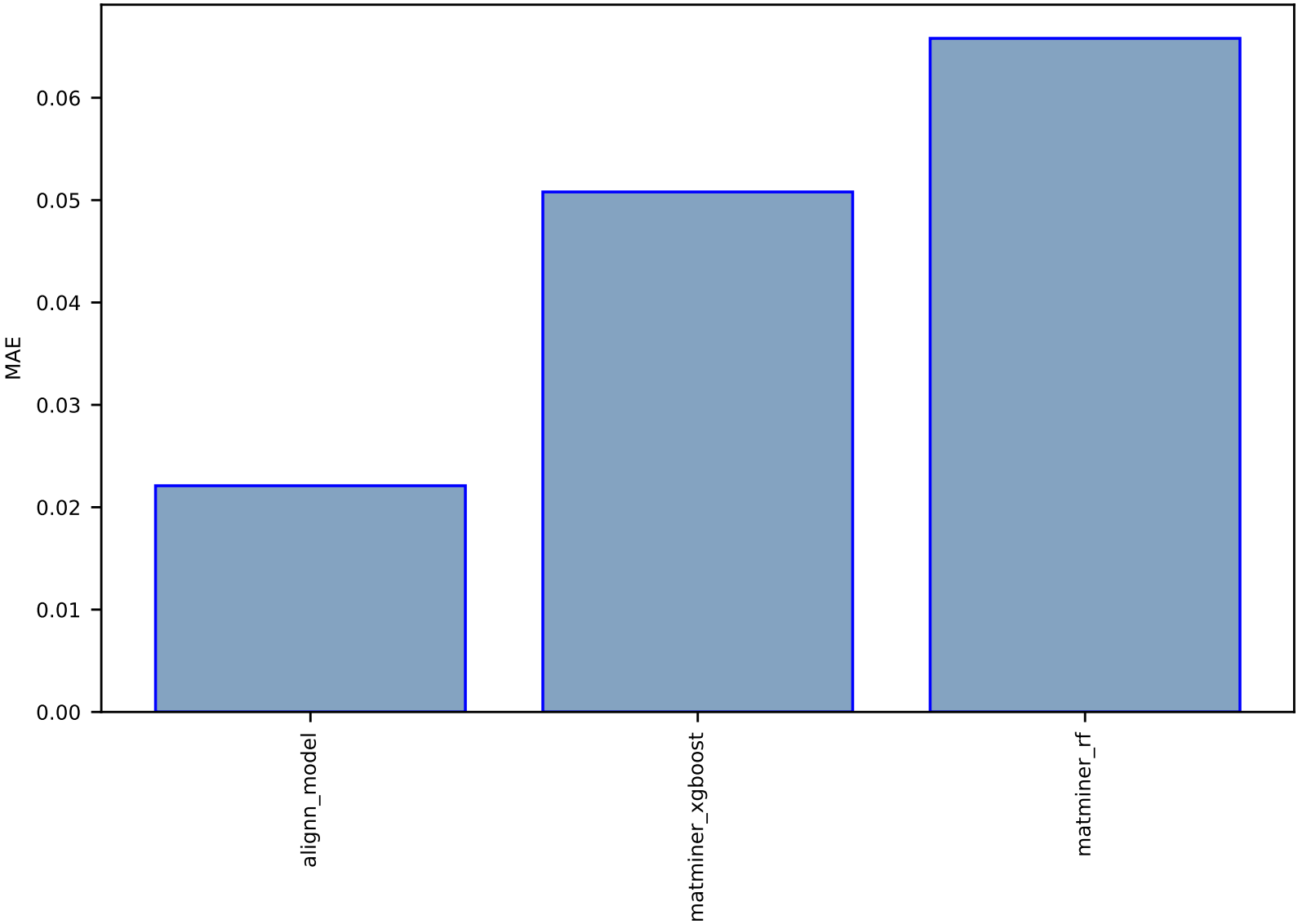


Fig S172:AI-SinglePropertyPrediction-HSE_decomp_energy-halide_peroskites-test-mae.csv.zip

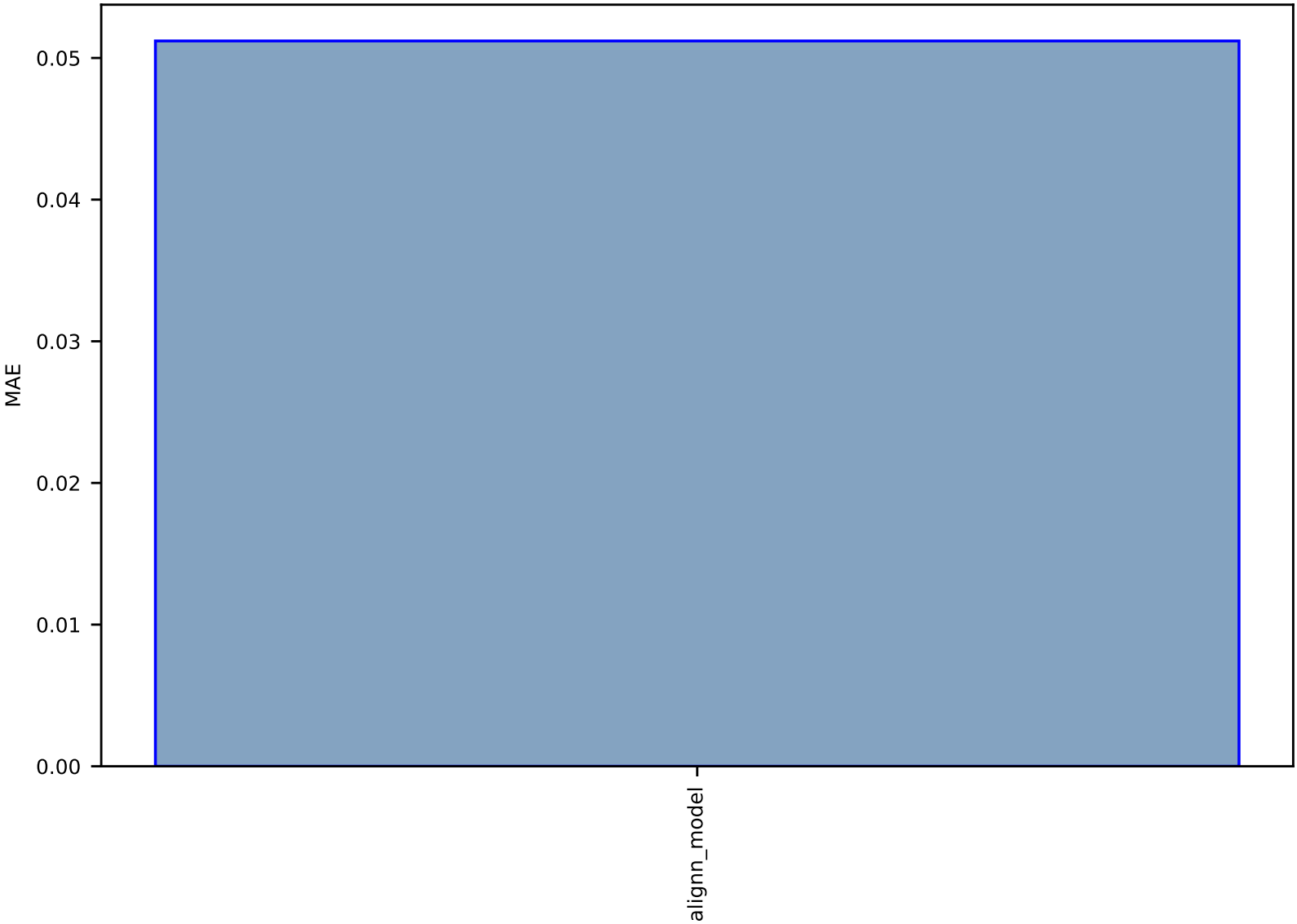


Fig S173:AI-SinglePropertyClass-mbj_bandgap-dft_3d-test-acc.csv.zip

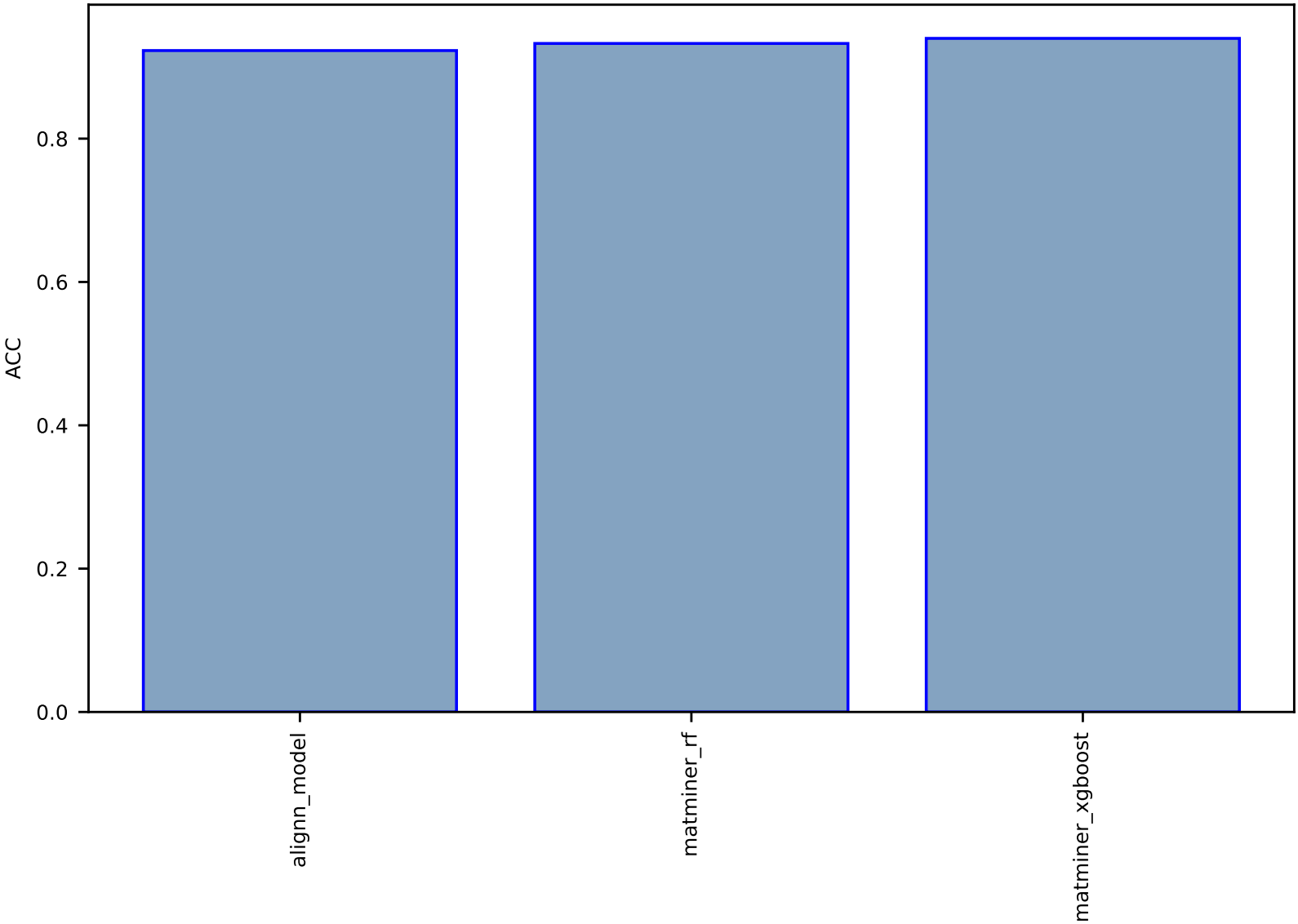


Fig S174:AI-SinglePropertyPrediction-bulk_modulus_kv-dft_3d-test-mae.csv.zip

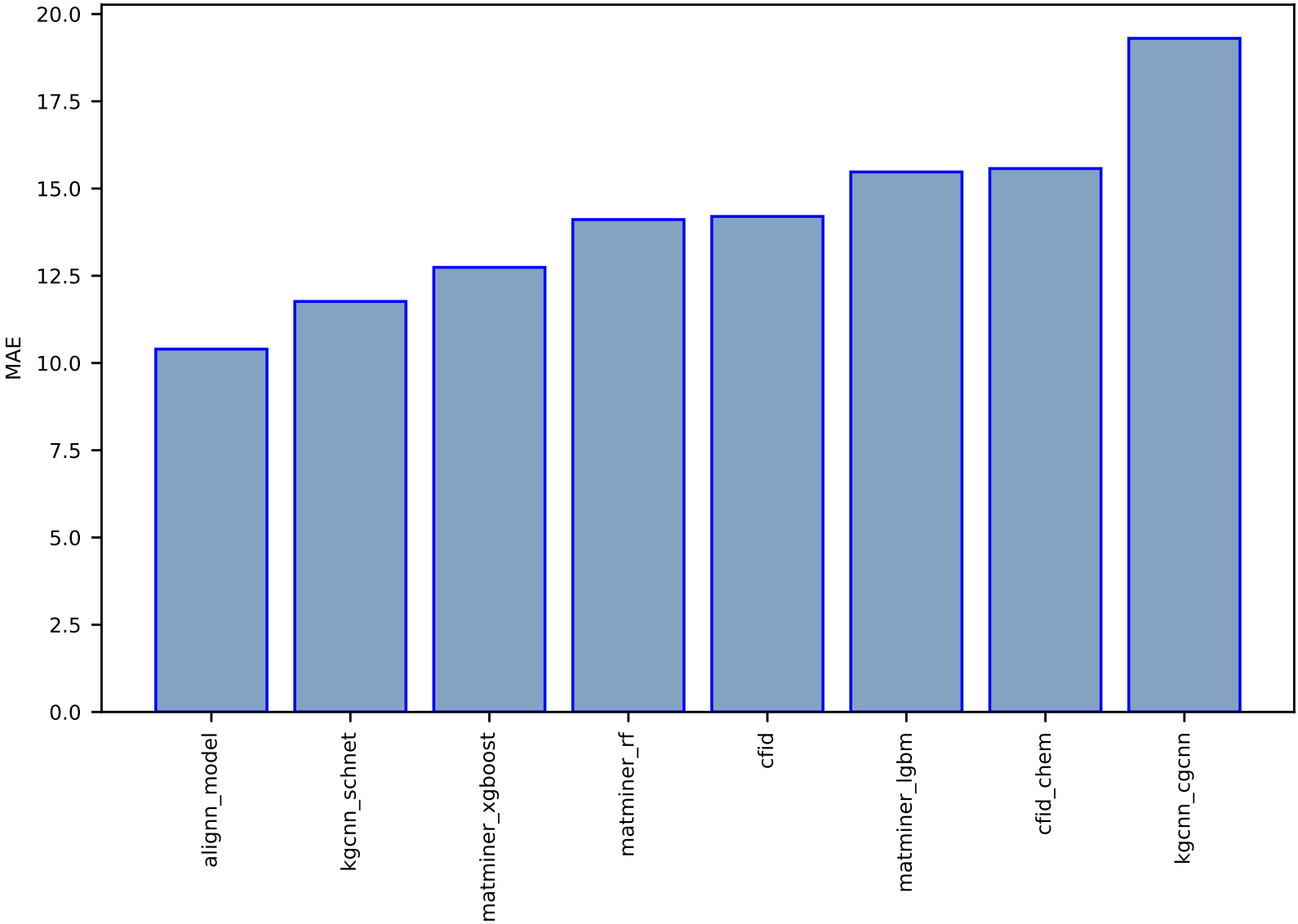


Fig S175:AI-SinglePropertyPrediction-PBE_gap-halide_peroskites-test-mae.csv.zip

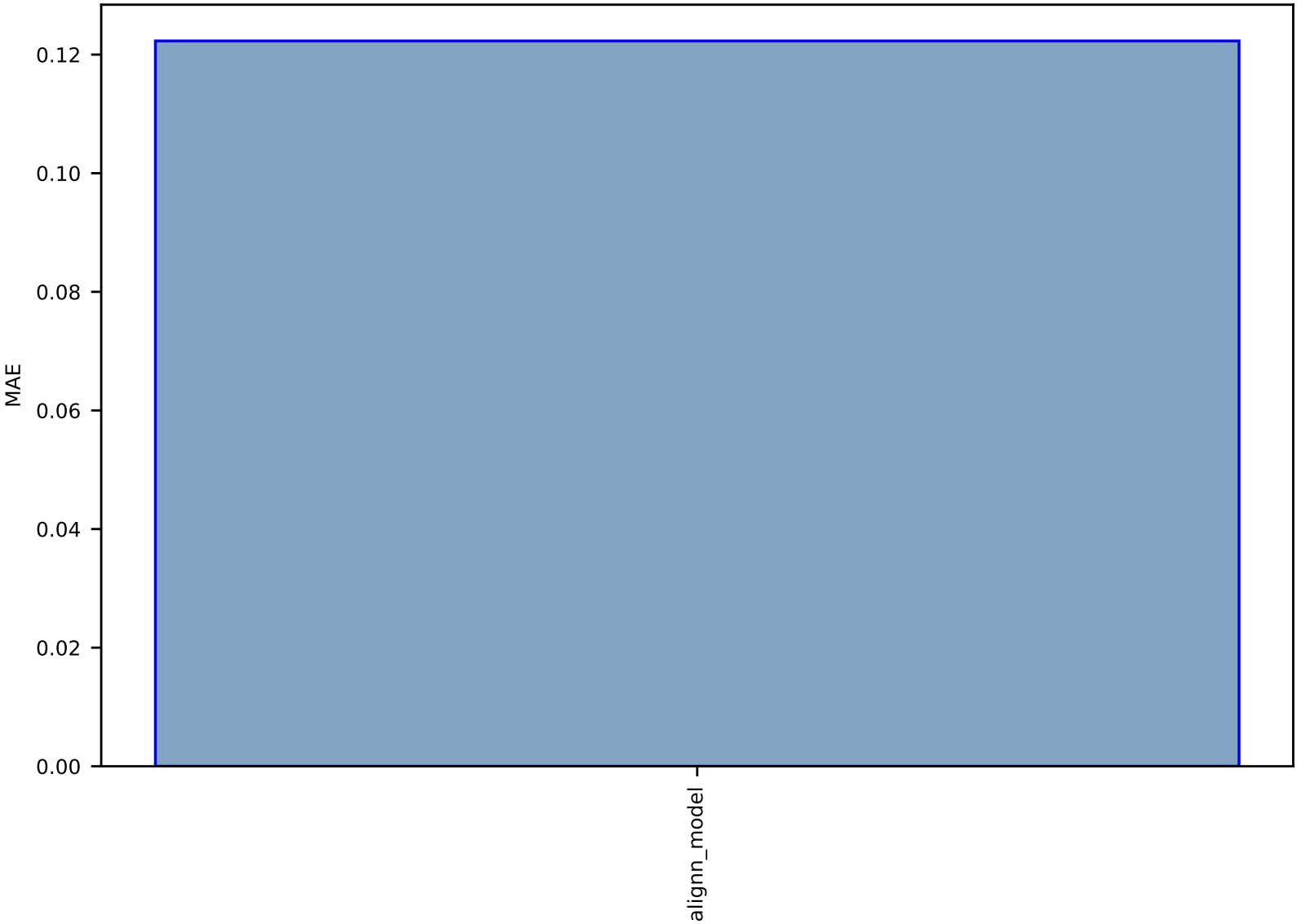


Fig S176:AI-SinglePropertyPrediction-void_fraction-hmof-test-mae.csv.zip

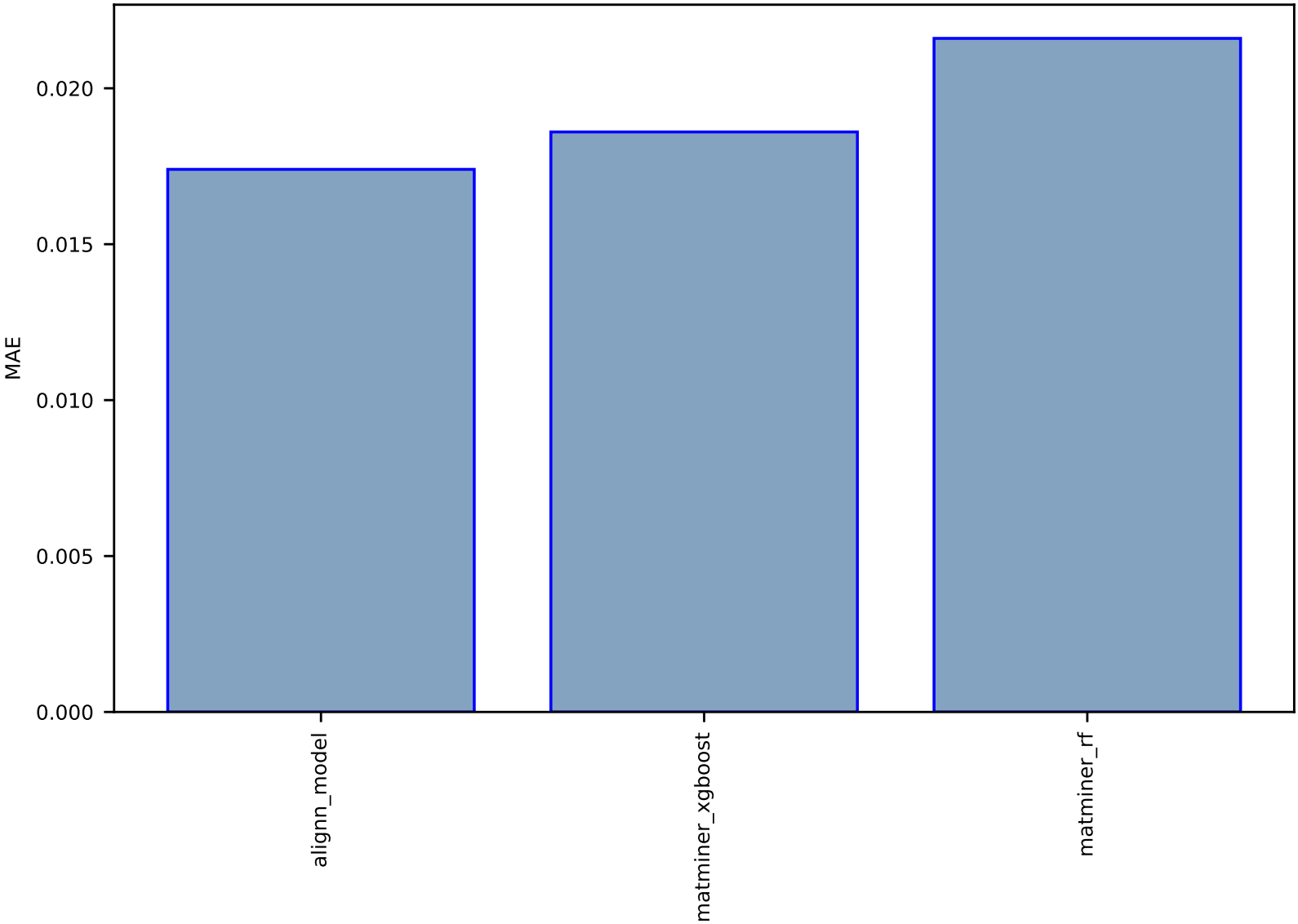


Fig S177:AI-SinglePropertyPrediction-ead-tinnet_N-test-mae.csv.zip

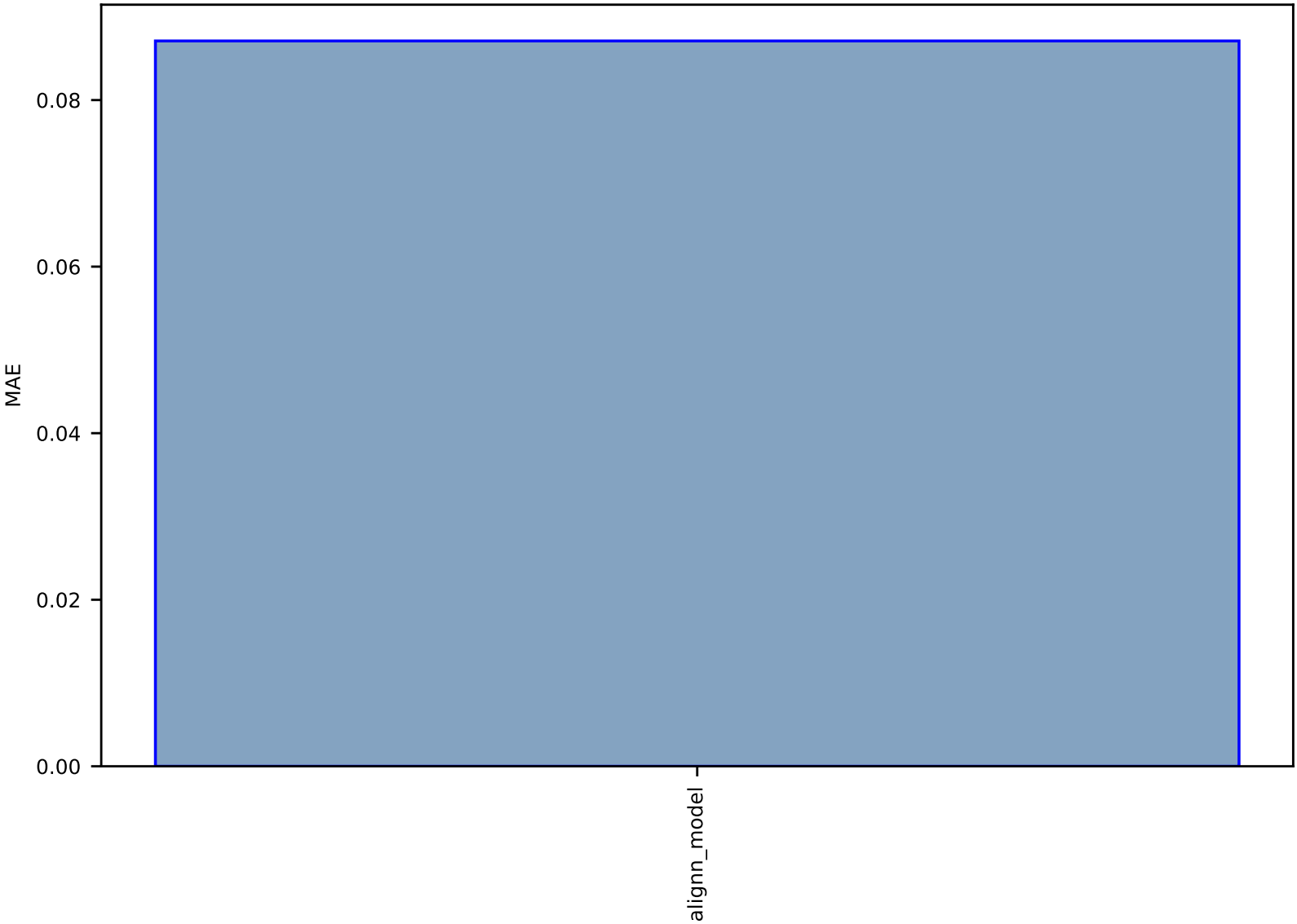


Fig S178:AI-SinglePropertyPrediction-avg_elec_mass-dft_3d-test-mae.csv.zip

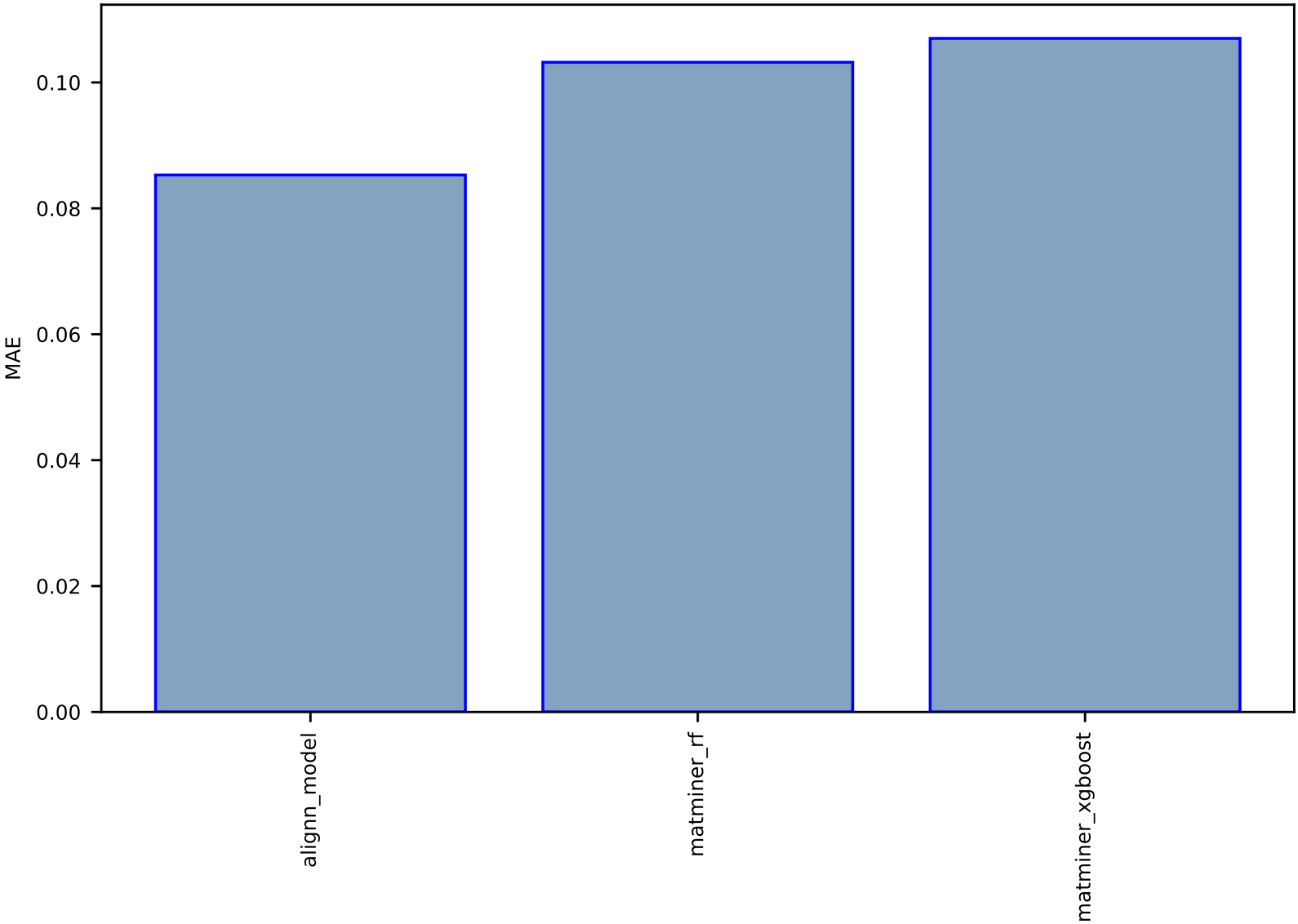


Fig S179:AI-SinglePropertyPrediction-U0-qm9_std_jctc-test-mae.csv.zip

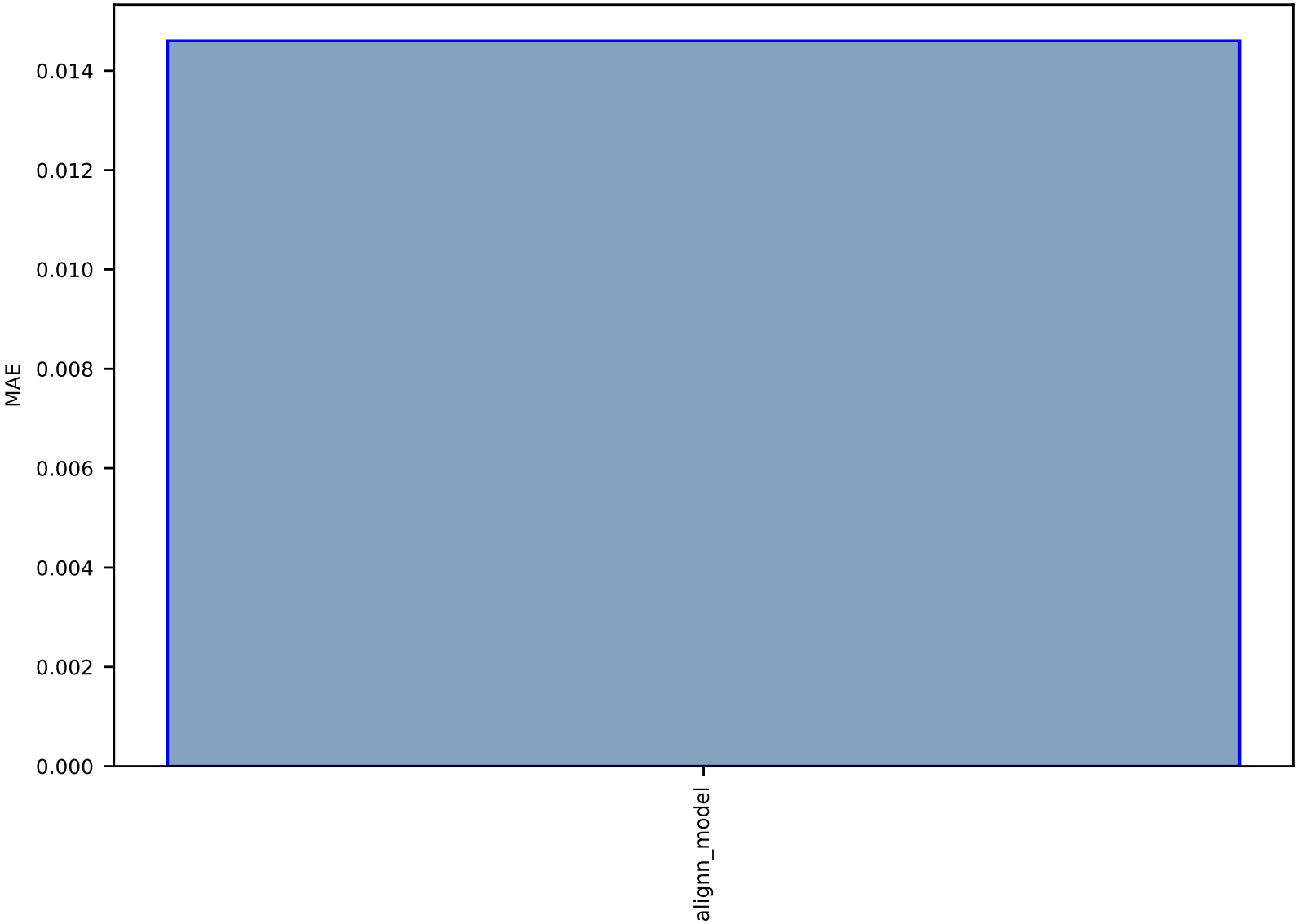


Fig S180:AI-SinglePropertyClass-spillage-dft_3d-test-acc.csv.zip

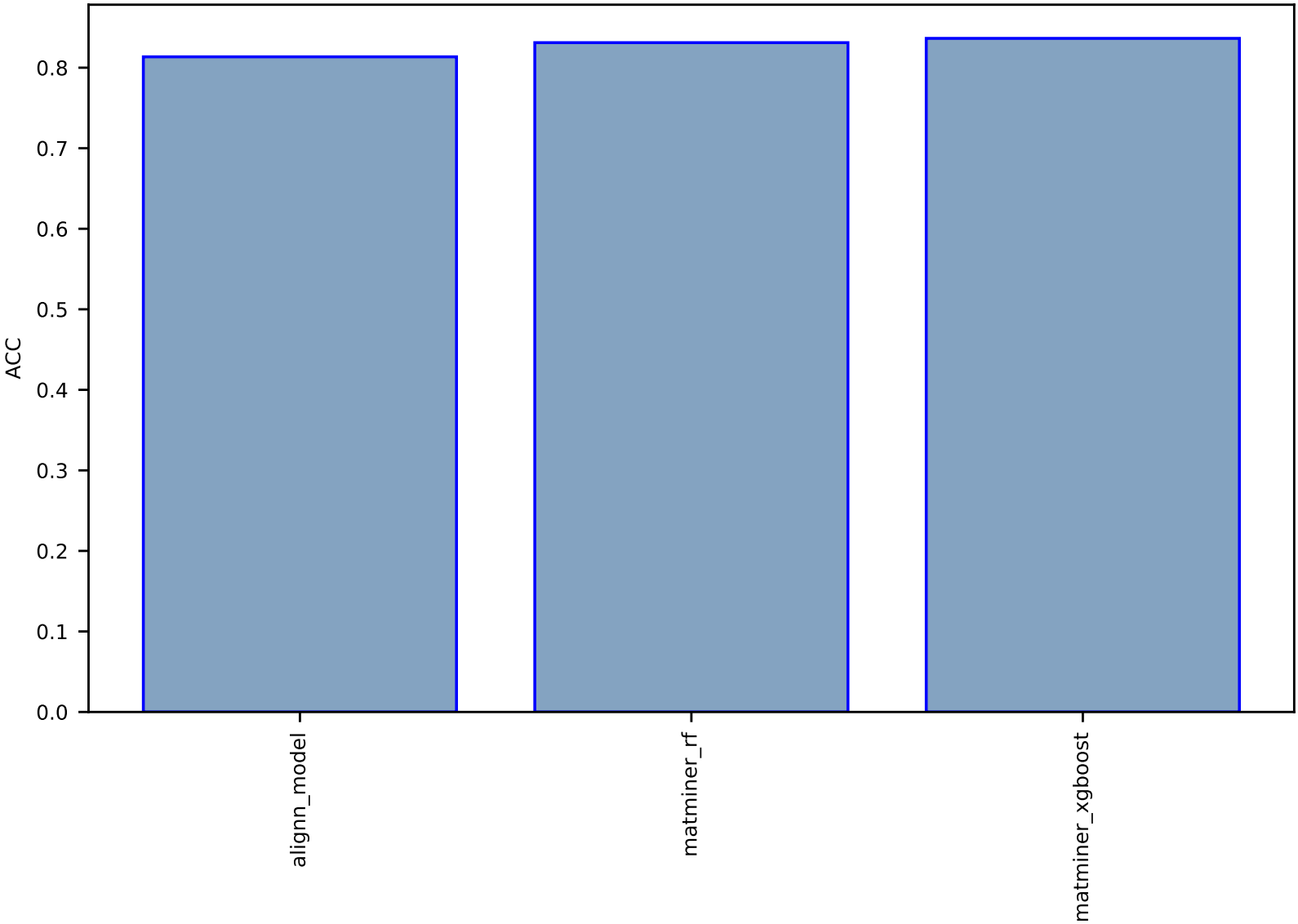


Fig S181:AI-SinglePropertyPrediction-alpha-qm9_std_jctc-test-mae.csv.zip

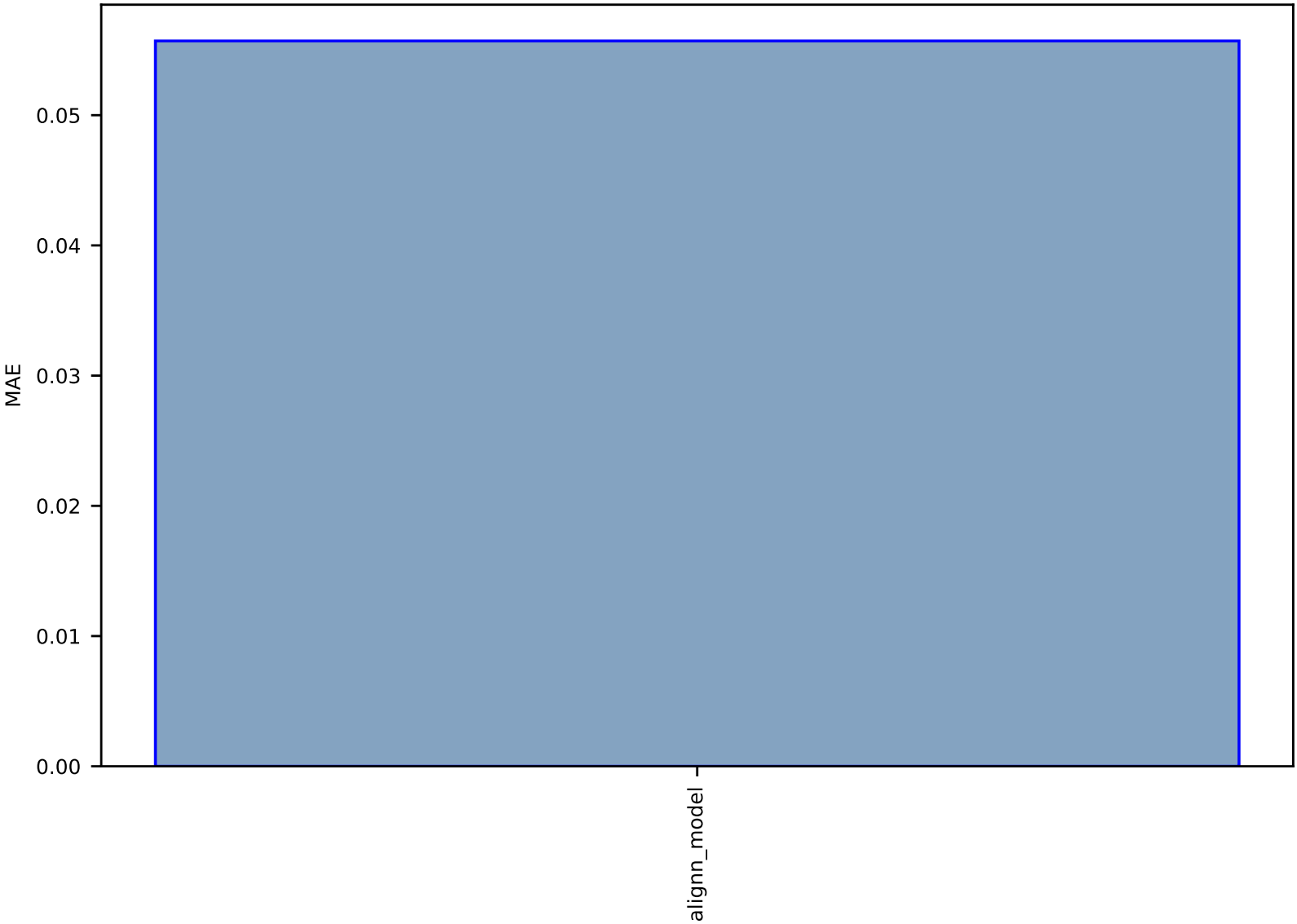


Fig S182:AI-SinglePropertyPrediction-PBE_decomp_energy-halide_peroskites-test-mae.csv.zip

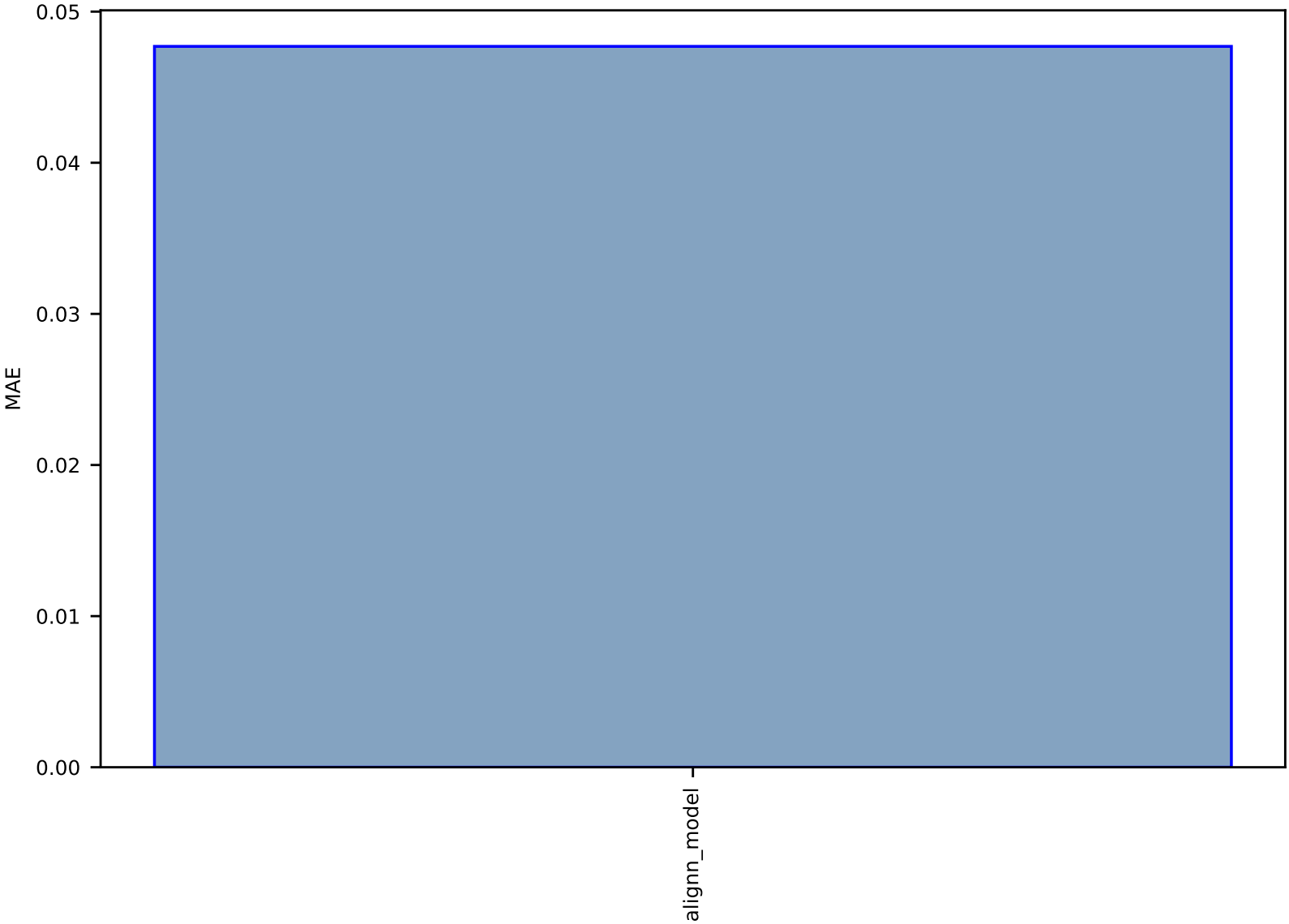


Fig S183:AI-SinglePropertyPrediction-R2-qm9_std_jctc-test-mae.csv.zip

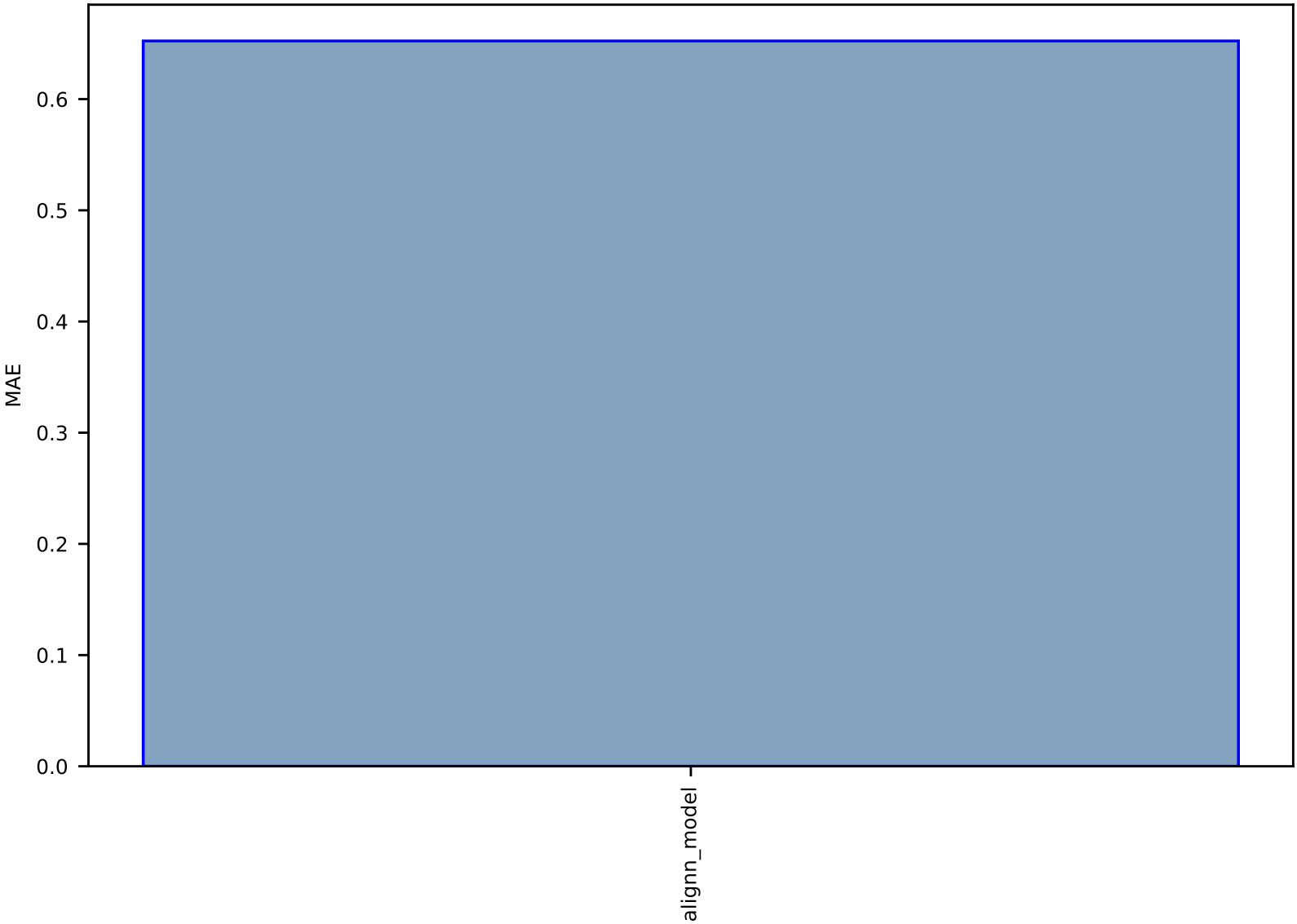


Fig S184:AI-SinglePropertyPrediction-min_co2_adsp-hmof-test-mae.csv.zip

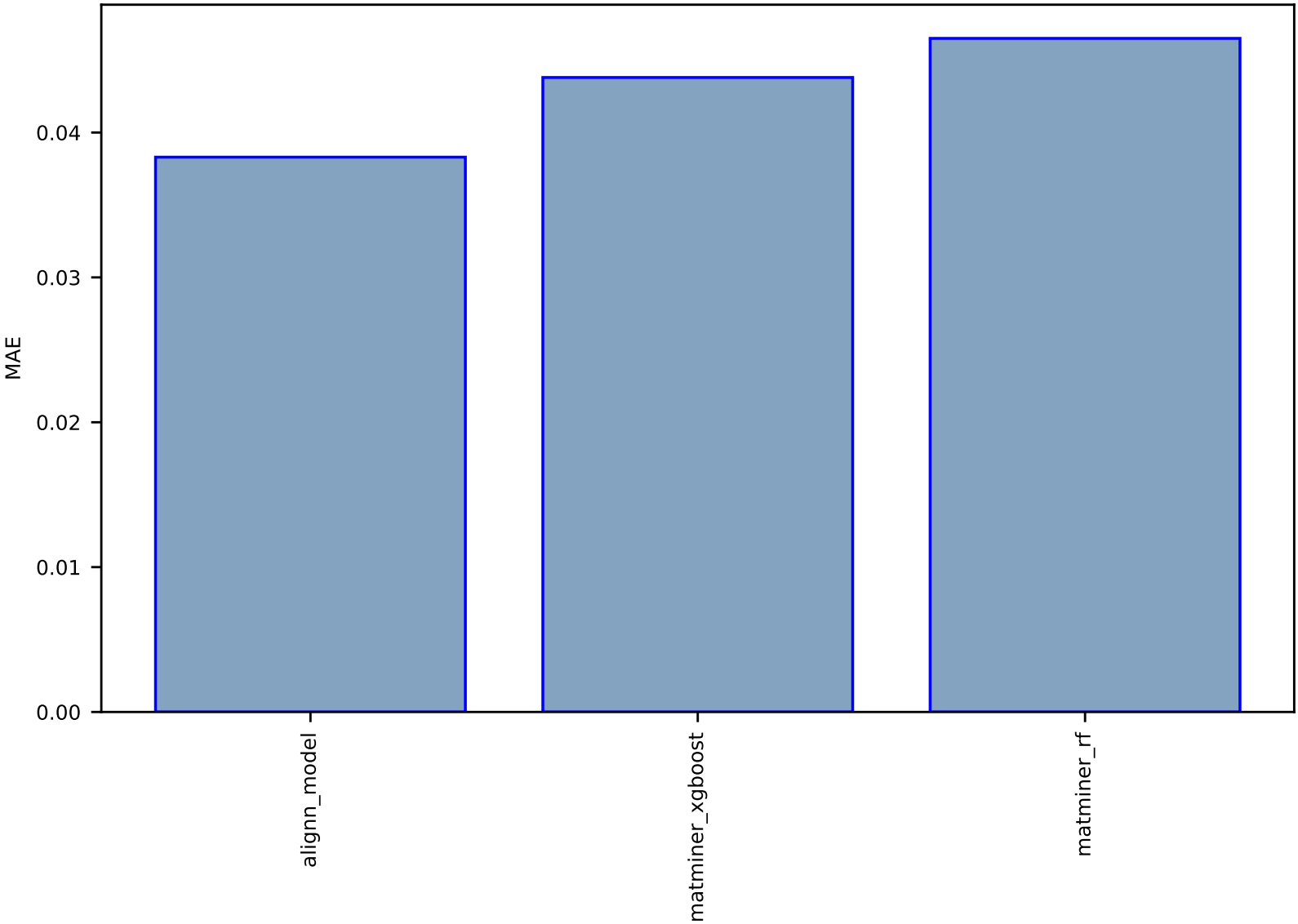


Fig S185:AI-SinglePropertyClass-slme-dft_3d-test-acc.csv.zip

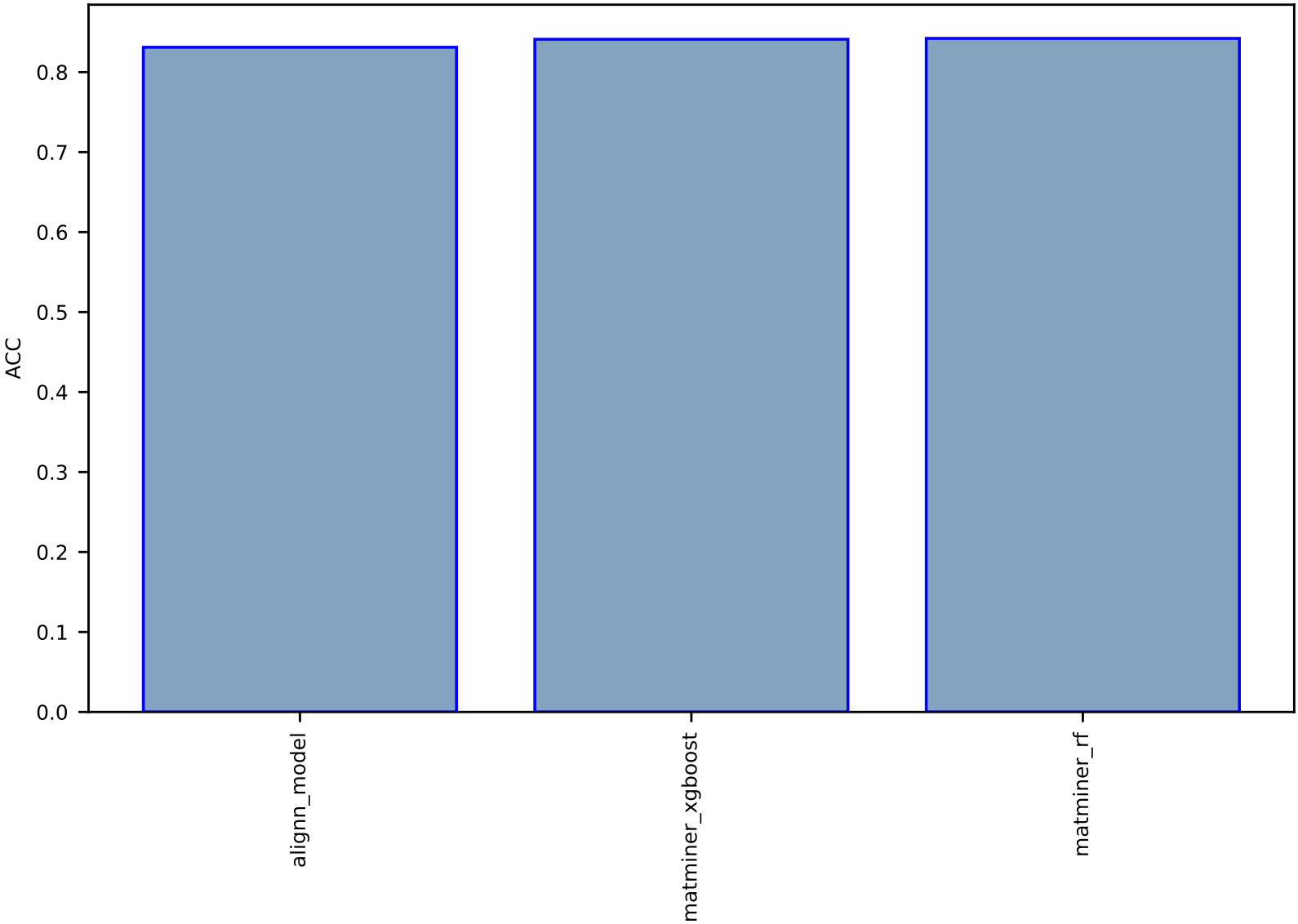


Fig S186:AI-SinglePropertyPrediction-HSE_gap-halide_peroskites-test-mae.csv.zip

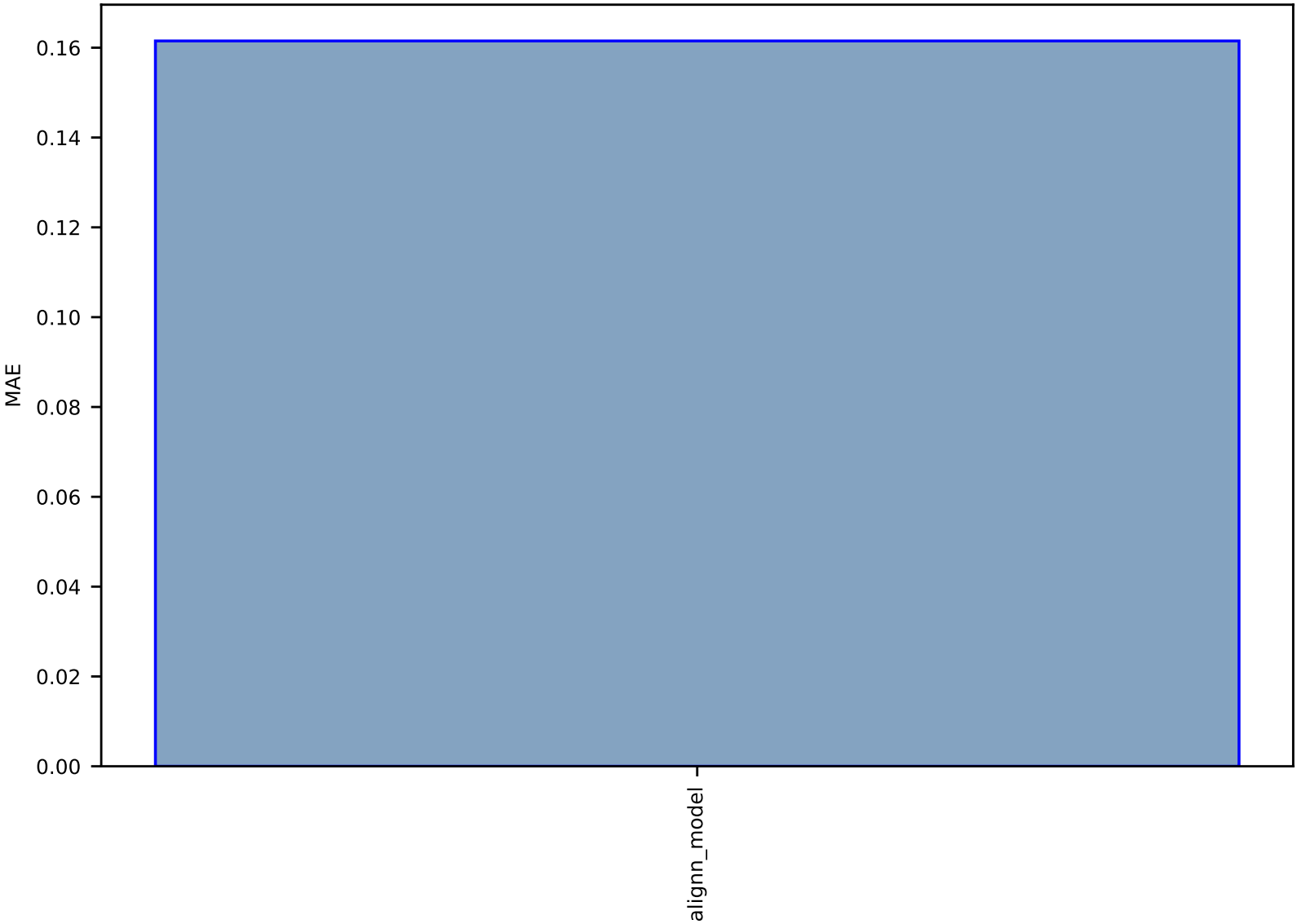


Fig S187:AI-Spectra-ph_dos-edos_pdos-test-multimae.csv.zip

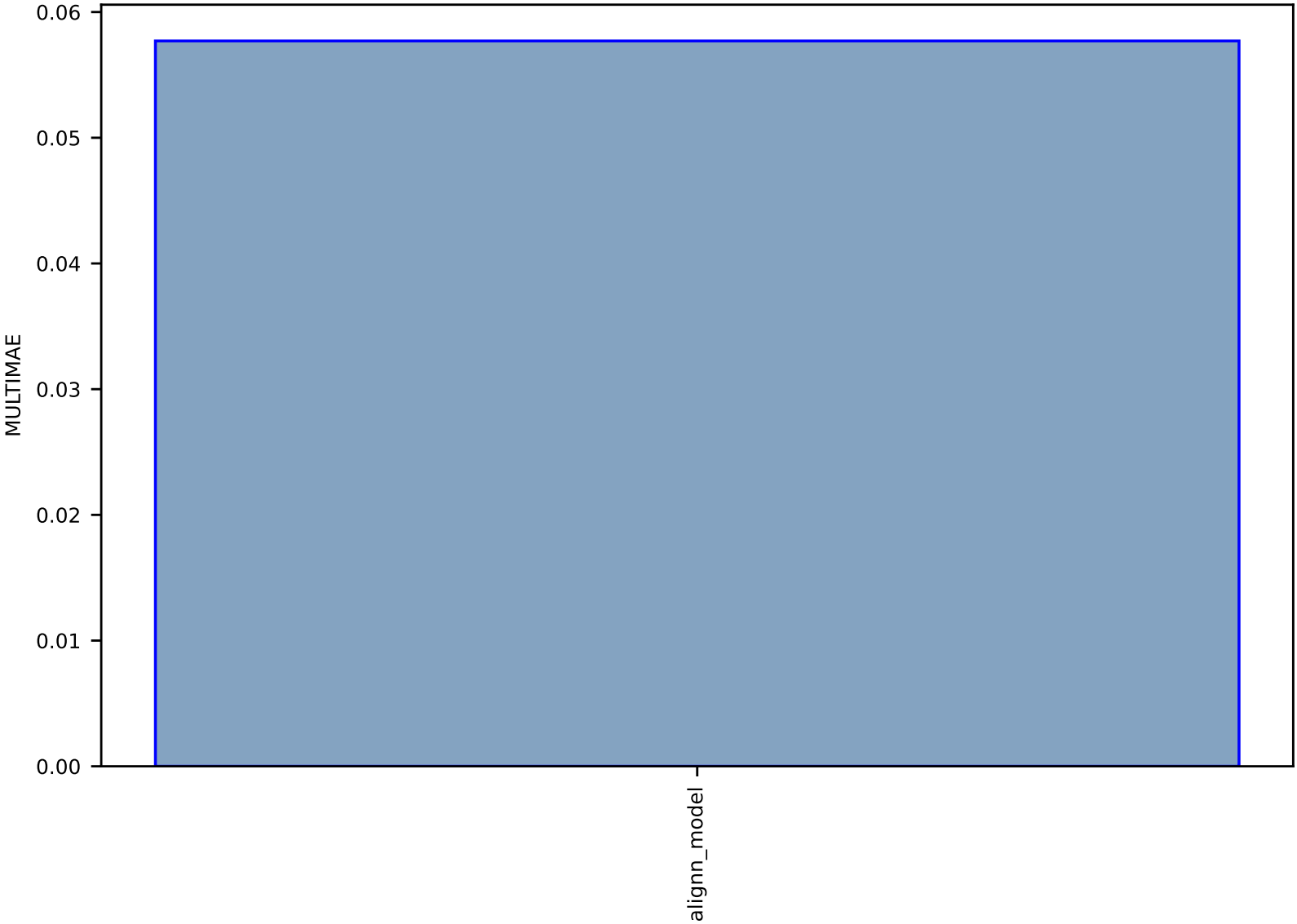


Fig S188:AI-SinglePropertyPrediction-Ref_ind-halide_peroskites-test-mae.csv.zip

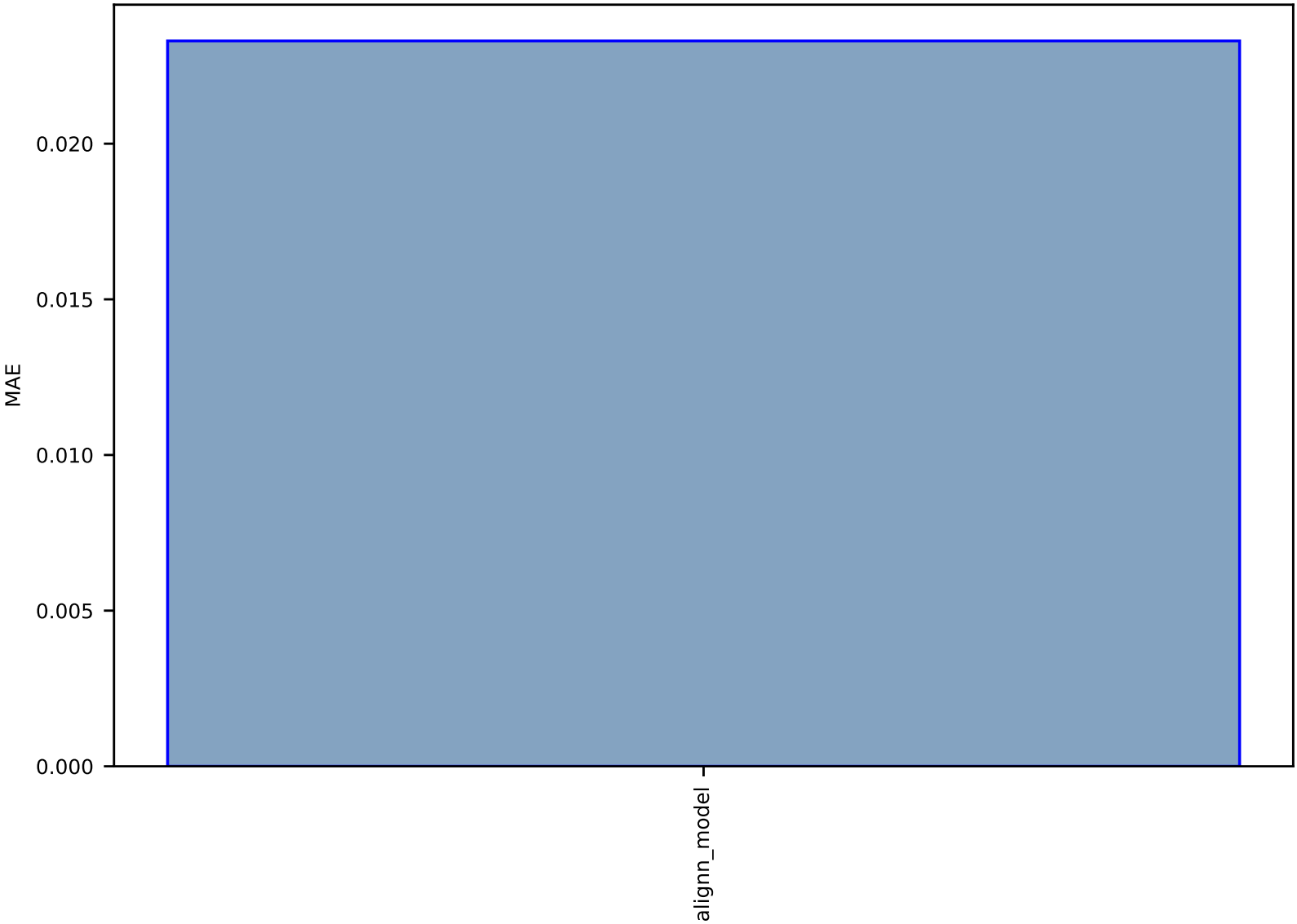


Fig S189:AI-SinglePropertyPrediction-U-qm9_std_jctc-test-mae.csv.zip

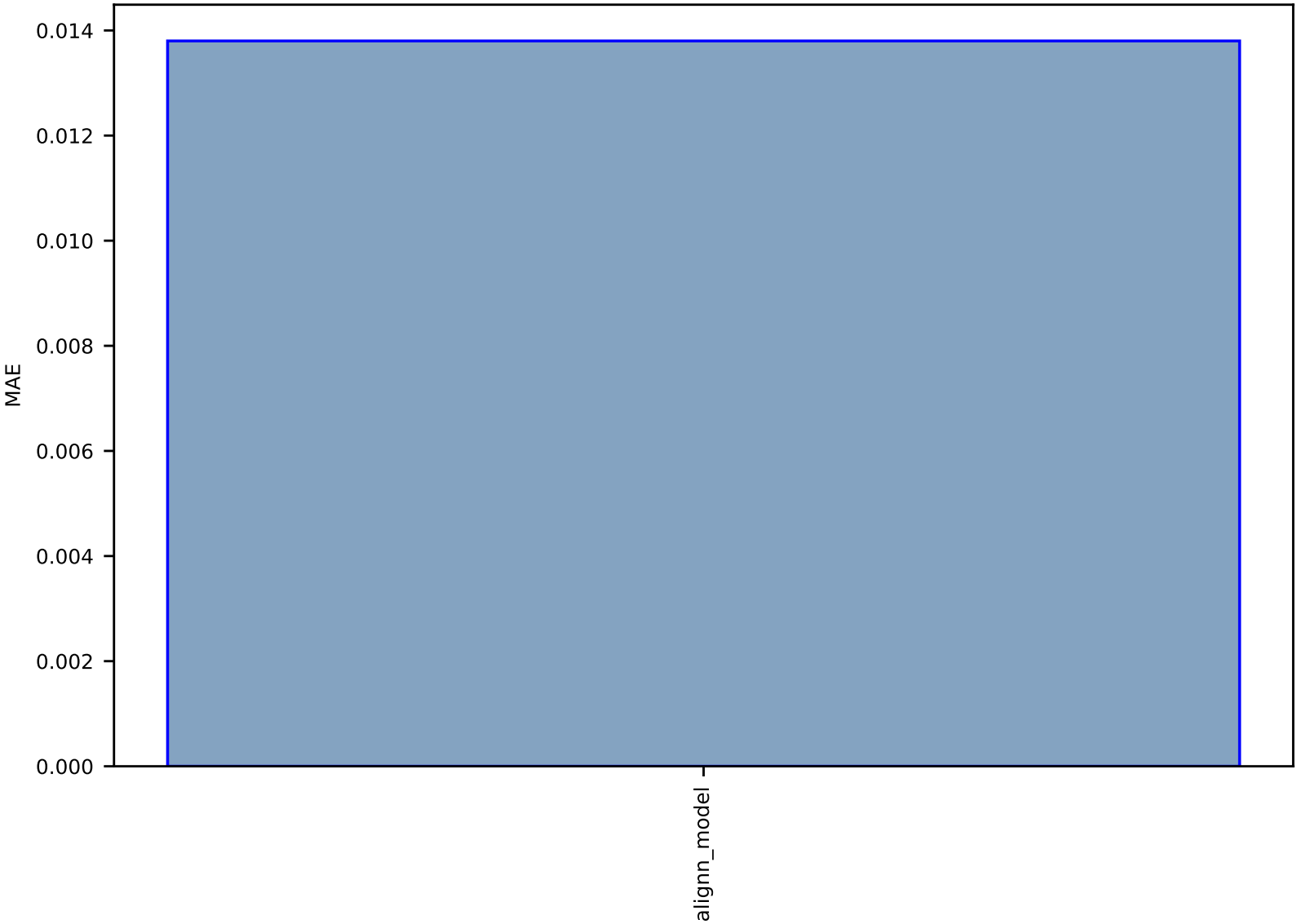


Fig S190:AI-SinglePropertyPrediction-lcd-hmof-test-mae.csv.zip

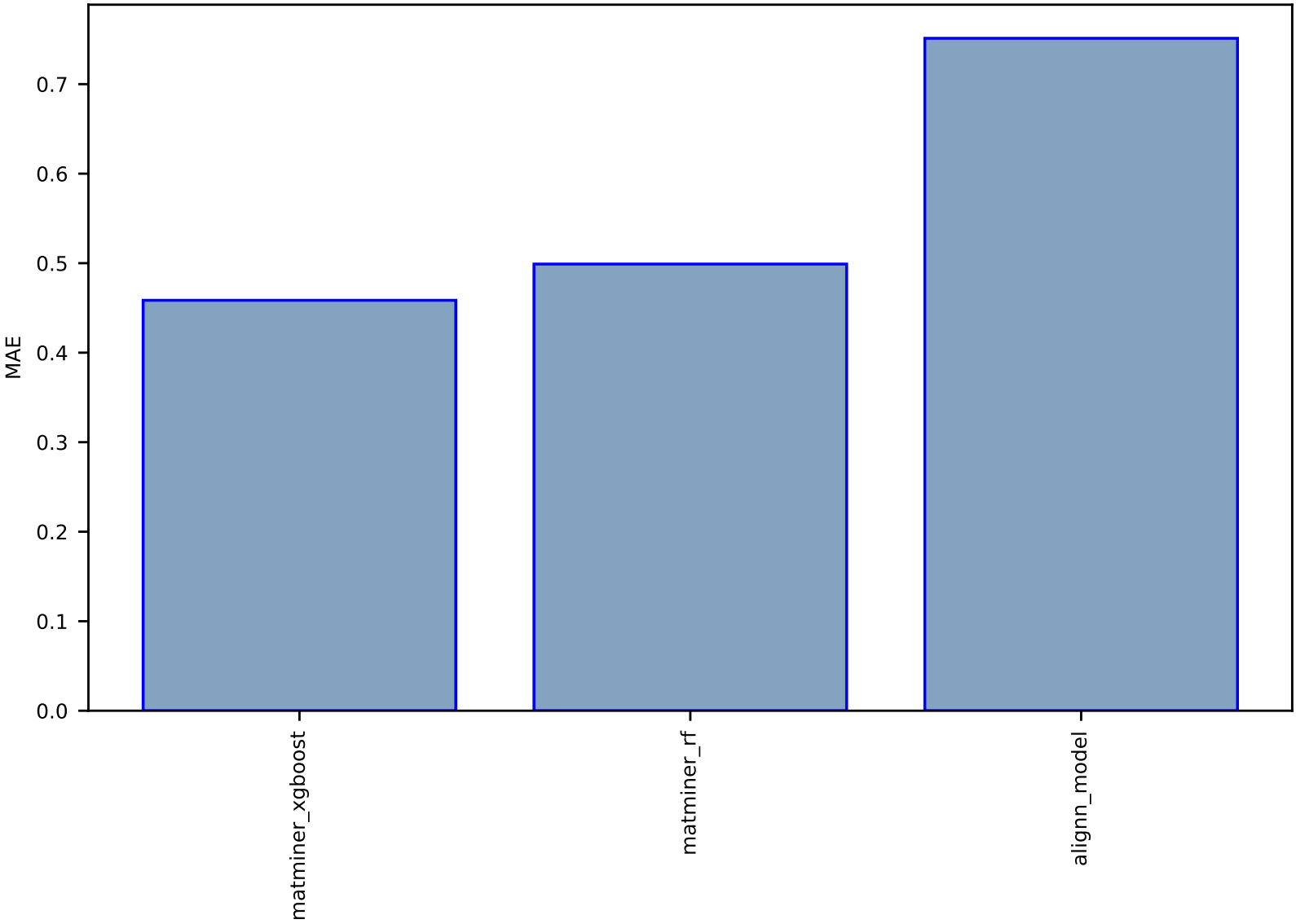


Fig S191:AI-TextGen-text-arxiv_gen-test-rouge.csv.zip

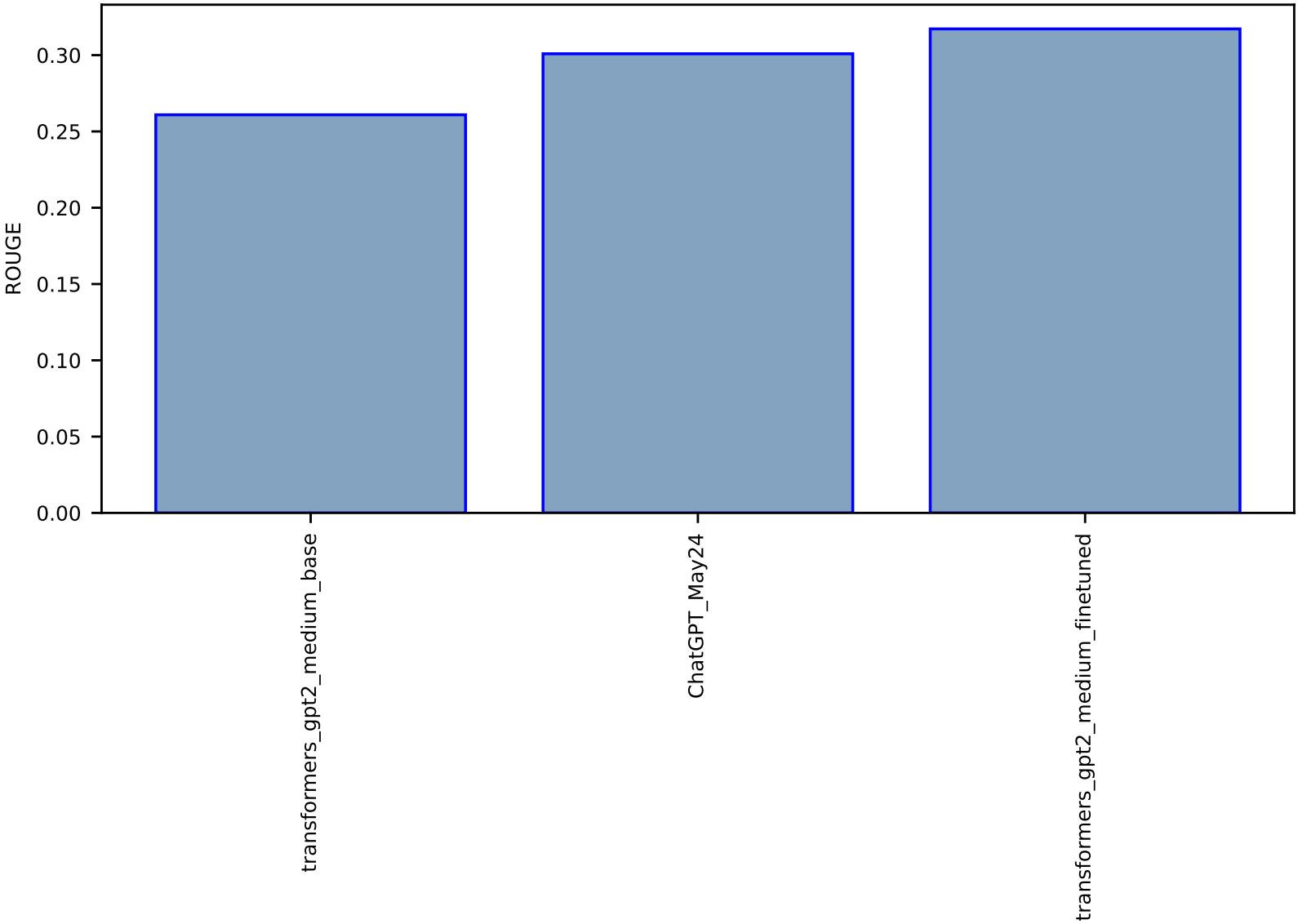


Fig S192:FF-SinglePropertyPrediction-deltaF-biobench-test-mae.csv.zip

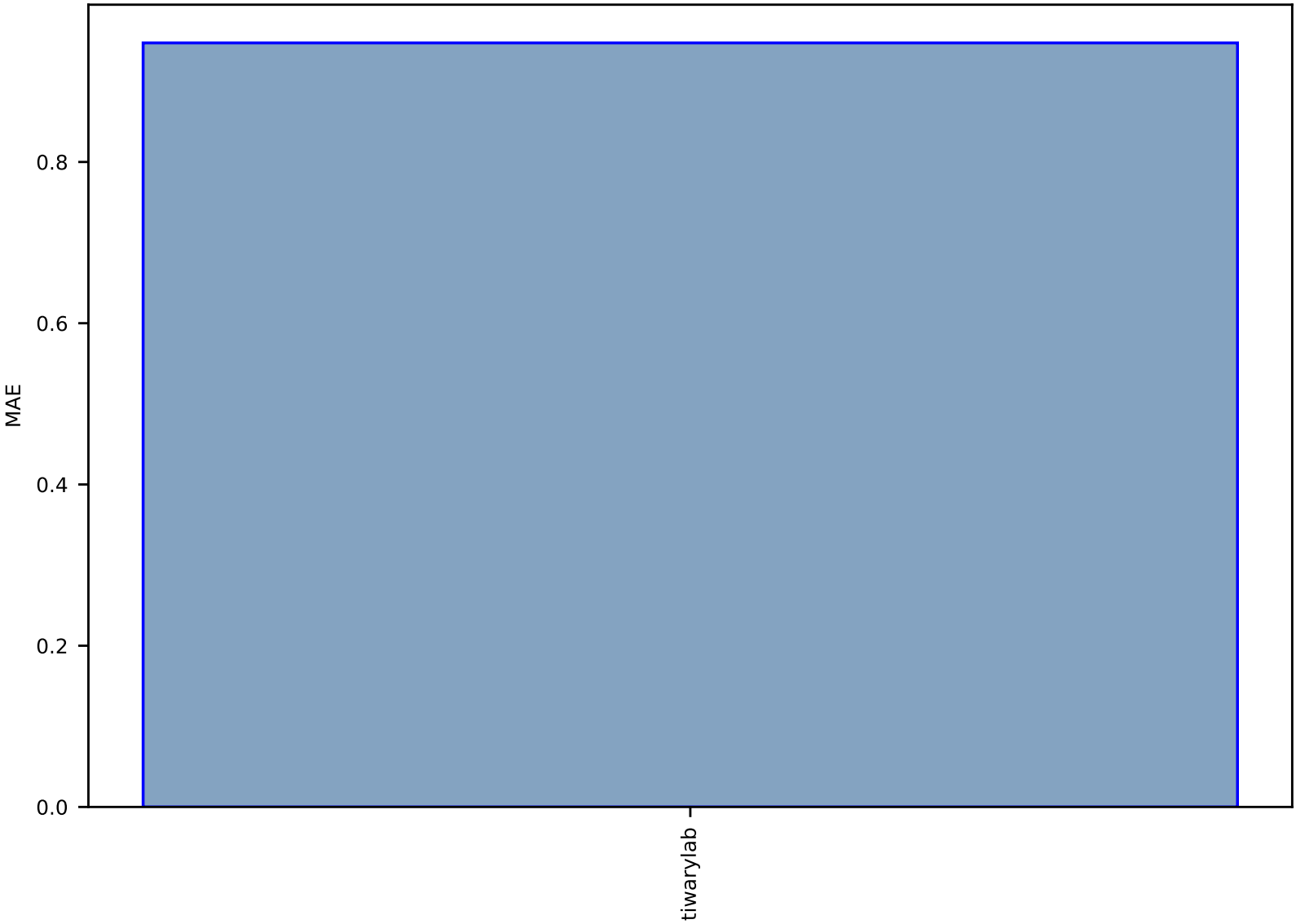


Fig S193:FF-SinglePropertyPrediction-left_handed_population-biobench-test-mae.csv.zip

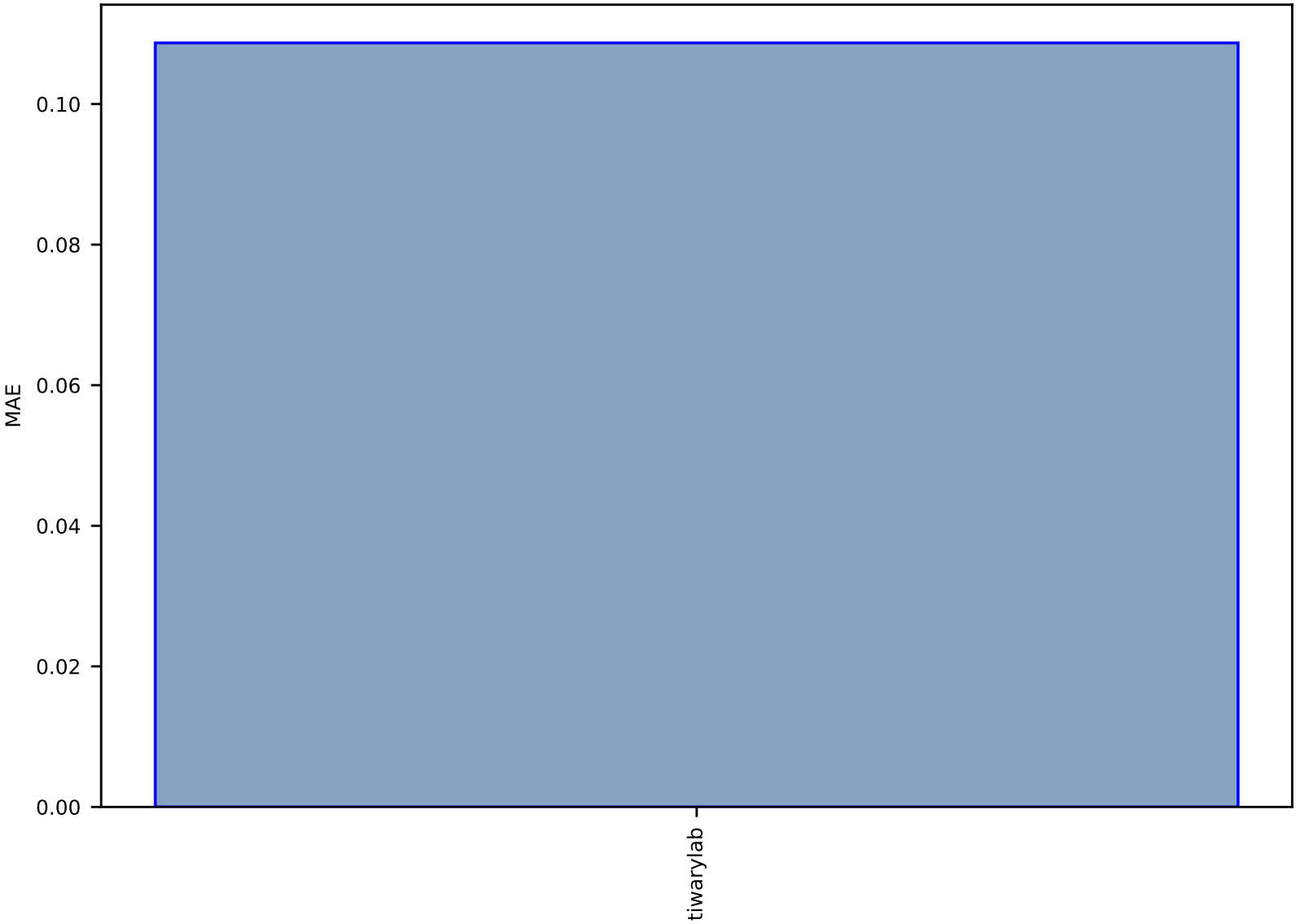


Fig S194:FF-SinglePropertyPrediction-right_handed_population-biobench-test-mae.csv.zip

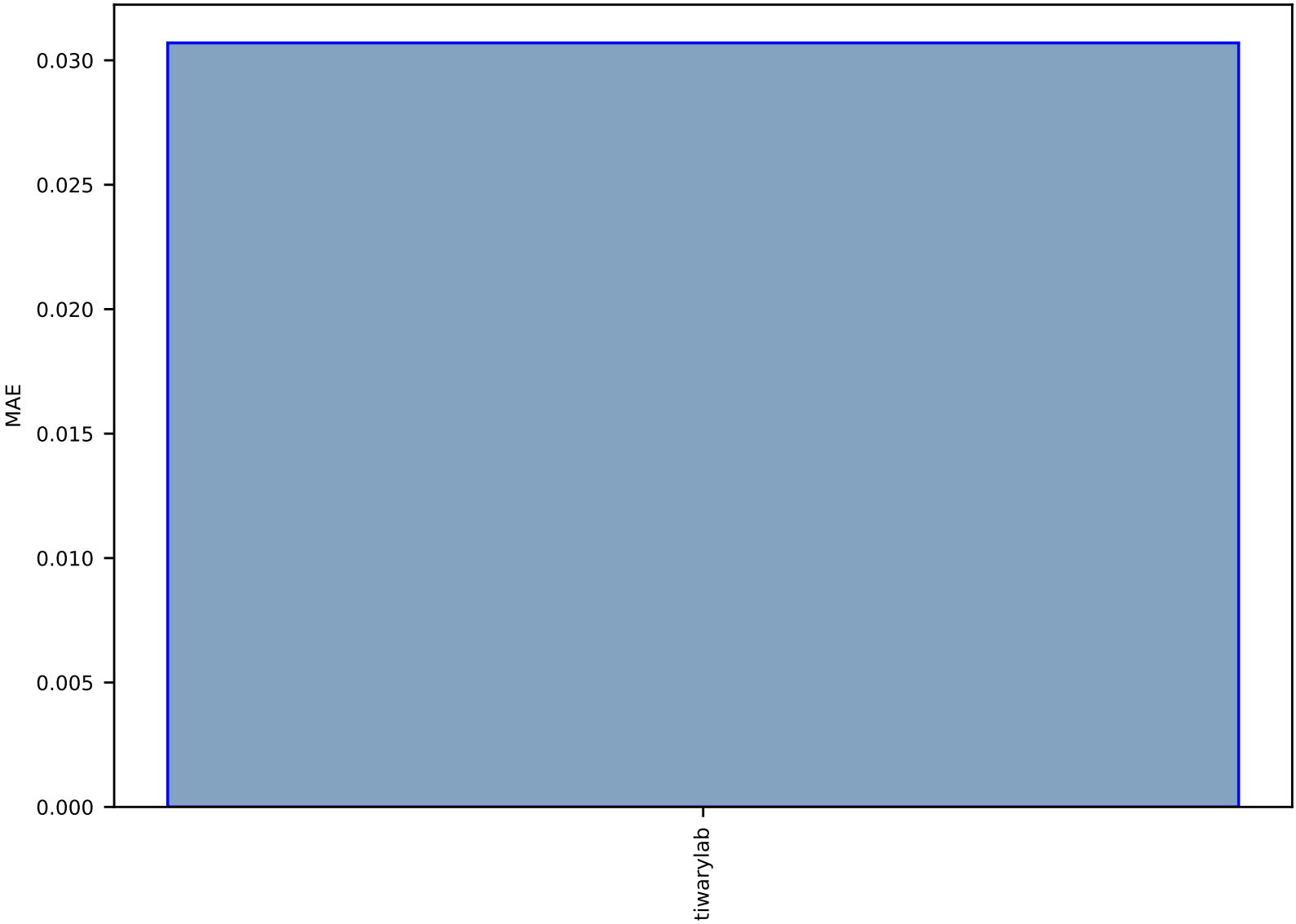


Fig S195:ES-SinglePropertyPrediction-Tc_supercon-dft_3d-test-mae.csv.zip

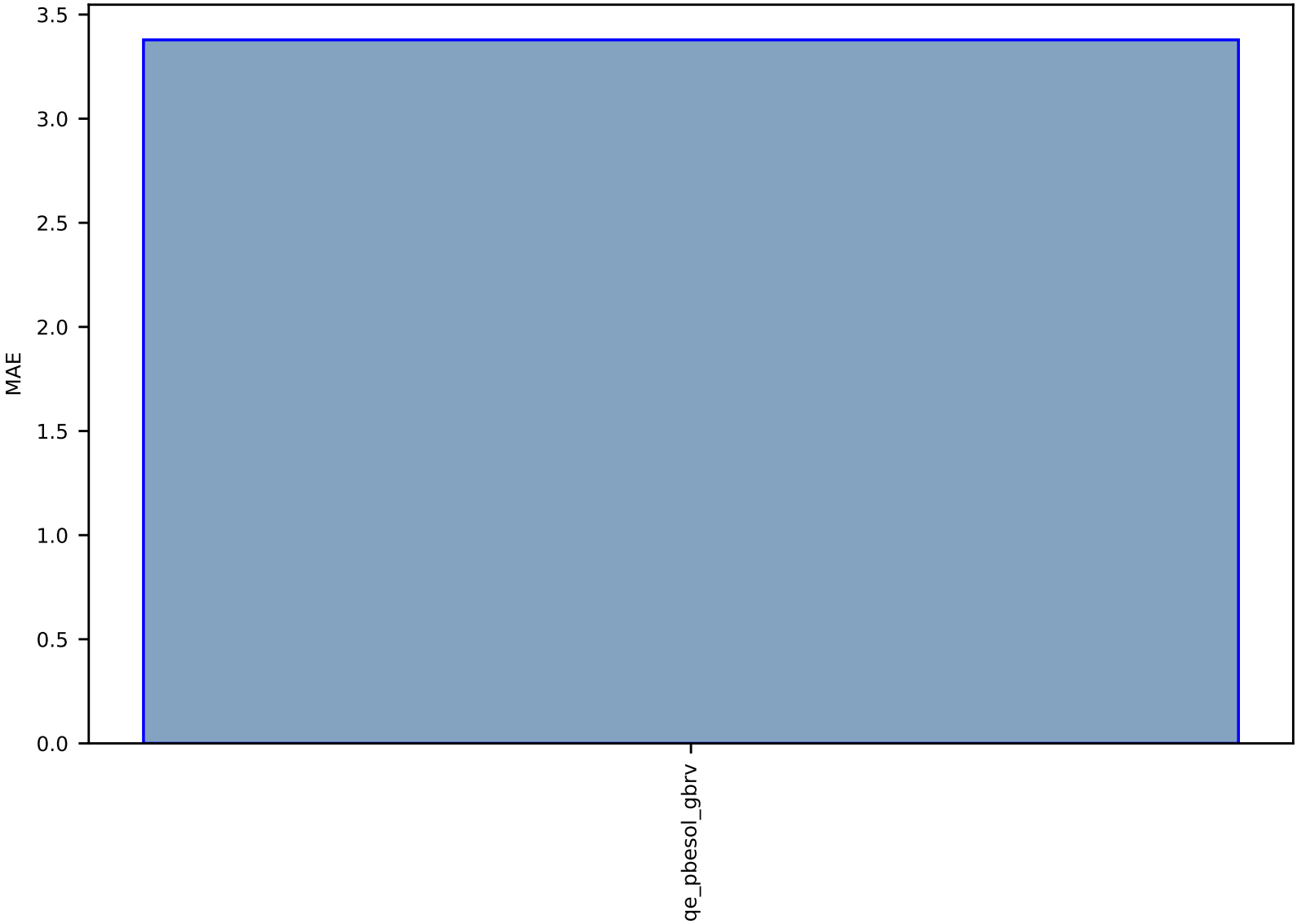


Fig S196:ES-SinglePropertyPrediction-Tc_supercon_JVASP_816_Al-dft_3d-test-mae.csv.zip

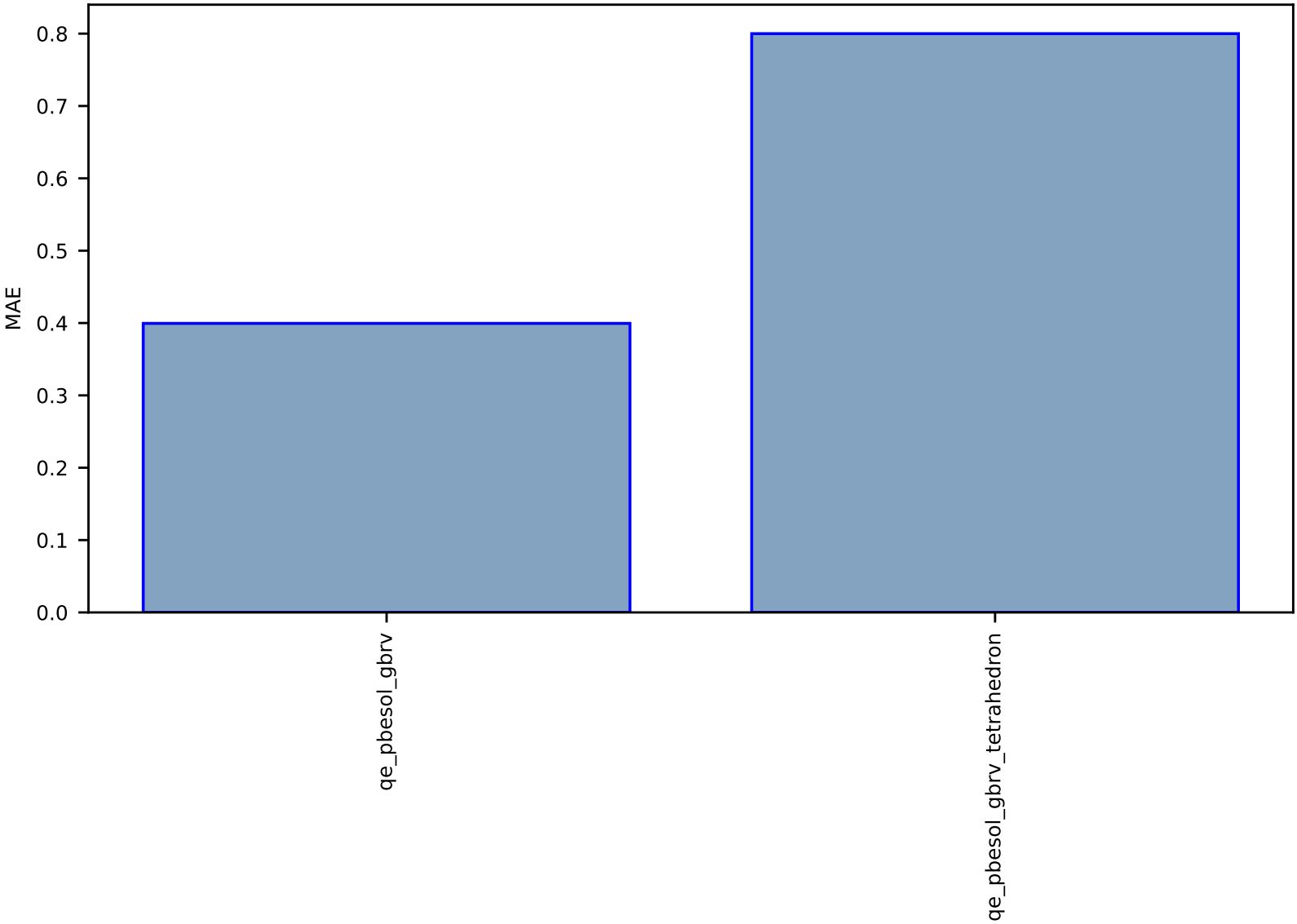


Fig S197:ES-SinglePropertyPrediction-Tc_supercon_JVASP_961_Pb-dft_3d-test-mae.csv.zip

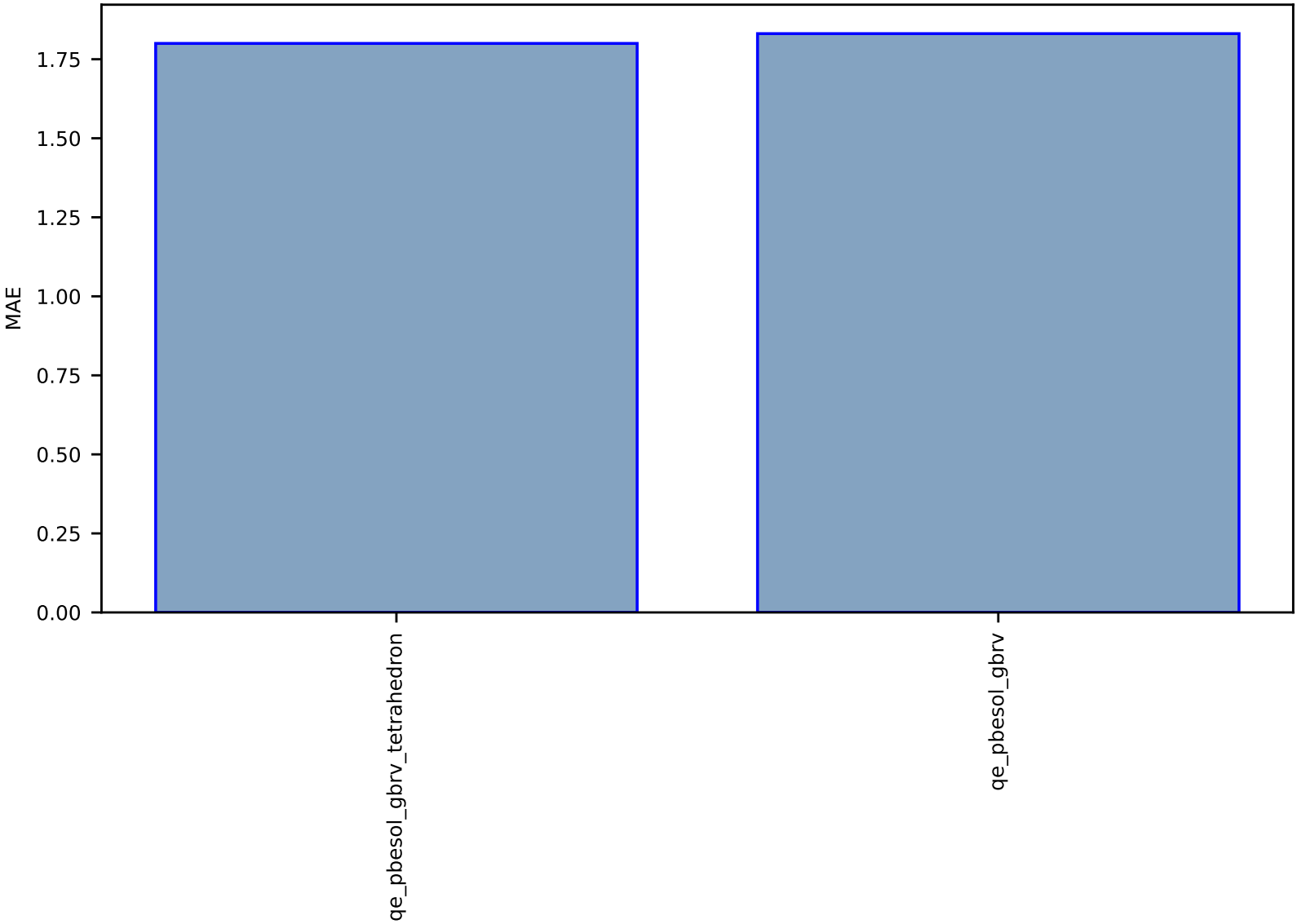


Fig S198:ES-SinglePropertyPrediction-Tc_supercon_JVASP_19889_NbC-dft_3d-test-mae.csv.zip

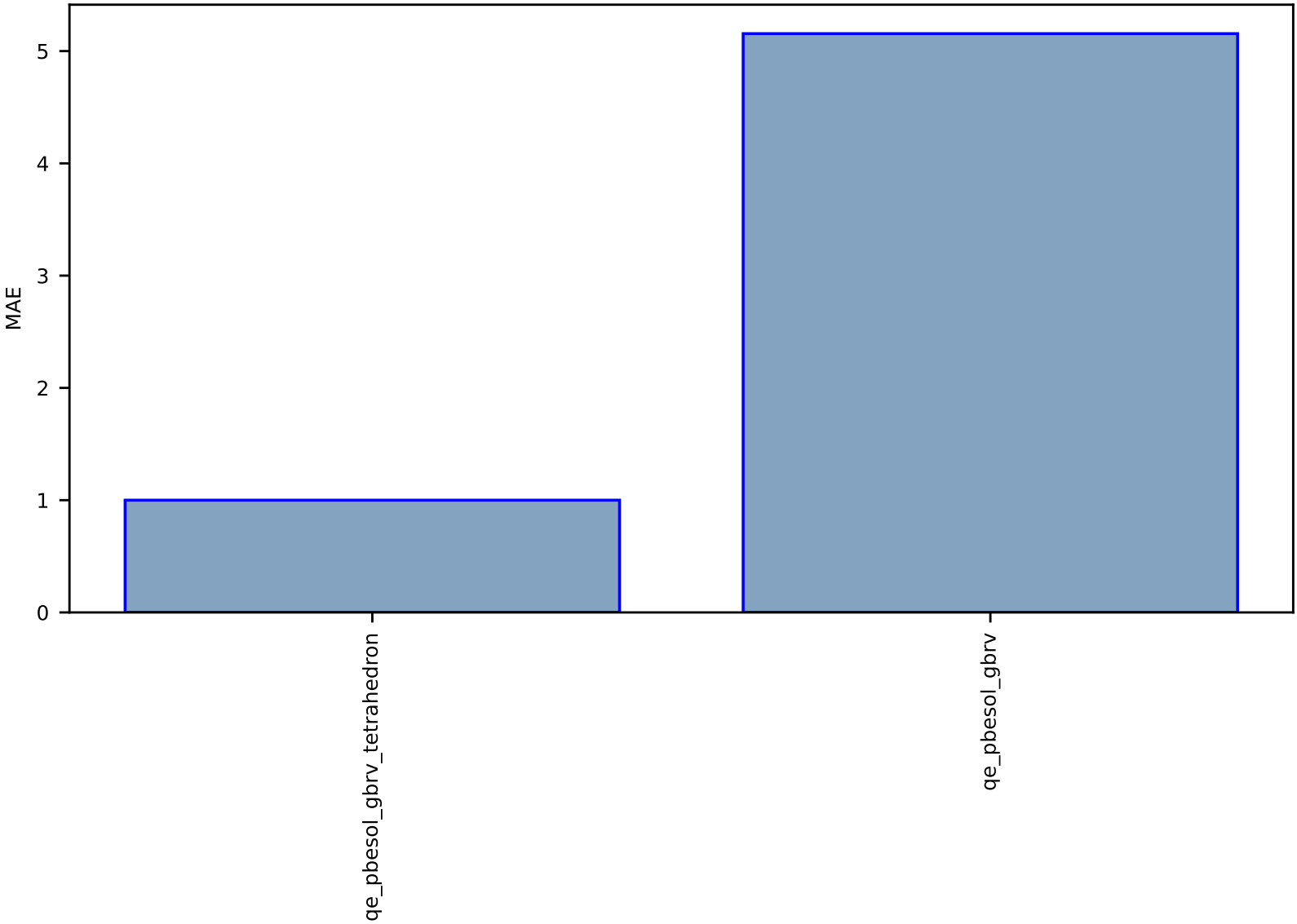


Fig S199:ES-SinglePropertyPrediction-Tc_supercon_JVASP_20620_YB6-dft_3d-test-mae.csv.zip

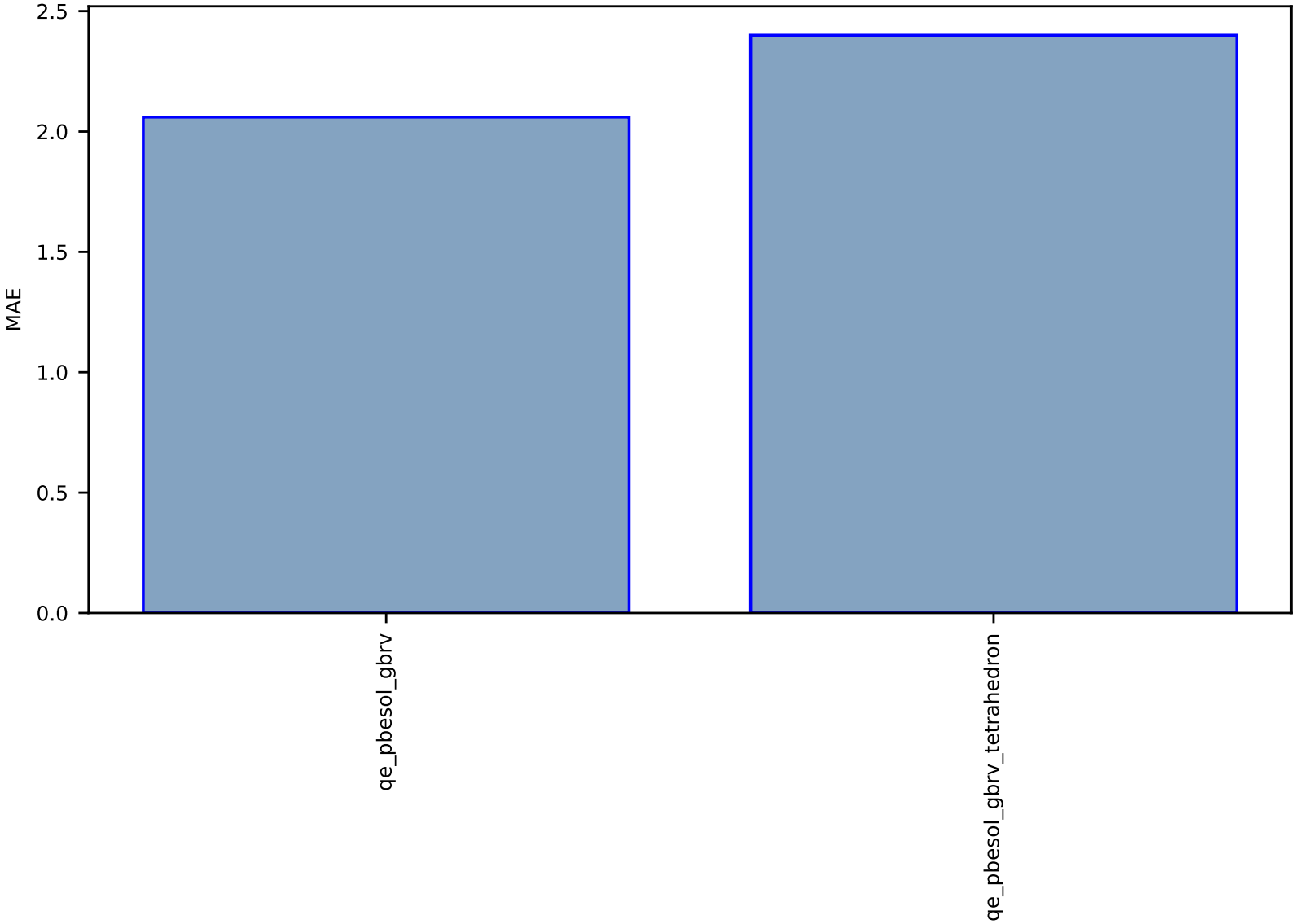


Fig S200:ES-SinglePropertyPrediction-Tc_supercon_JVASP_11981_Nb3Al-dft_3d-test-mae.csv.zip

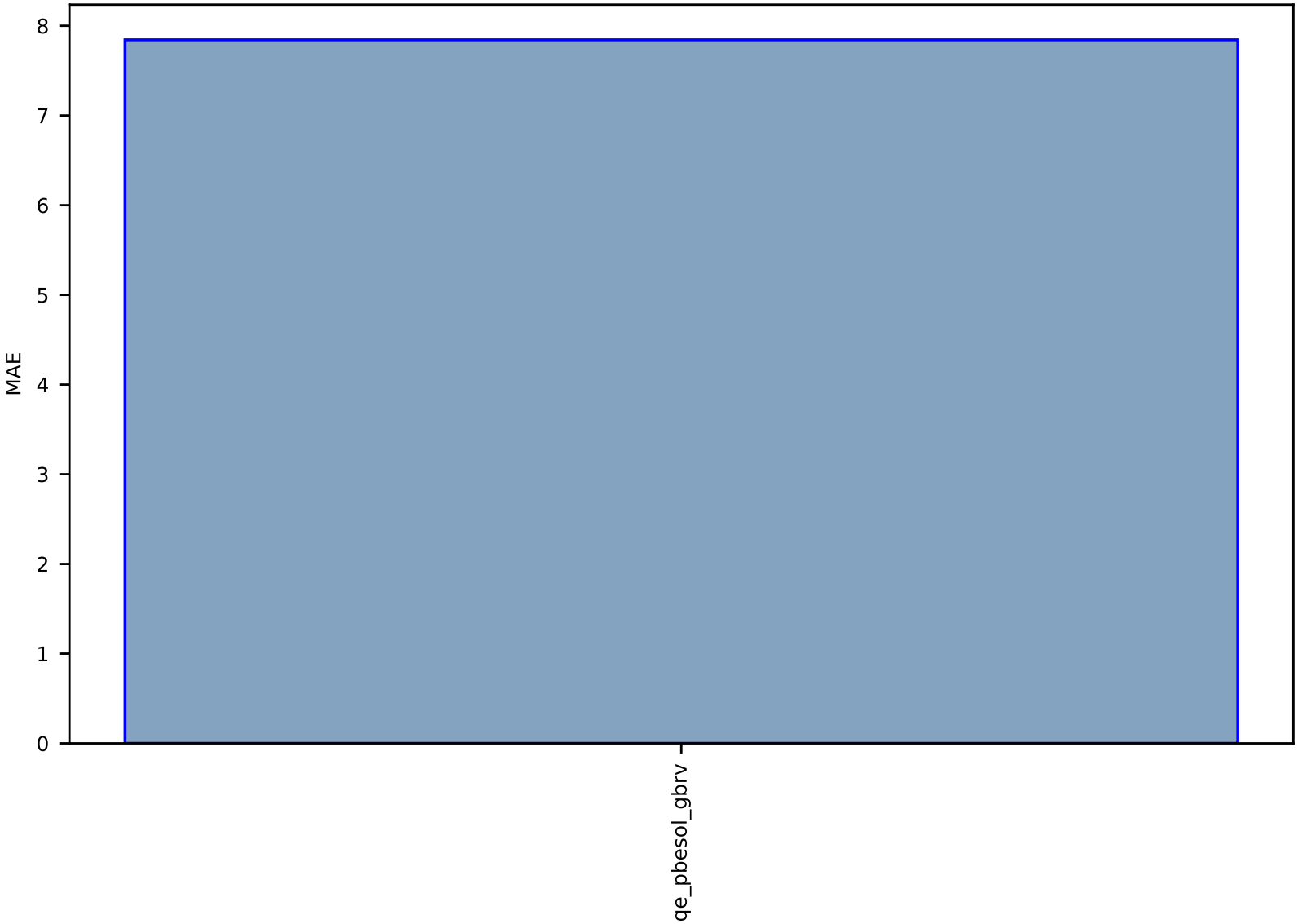


Fig S201:ES-SinglePropertyPrediction-Tc_supercon_JVASP_36335_NbN-dft_3d-test-mae.csv.zip

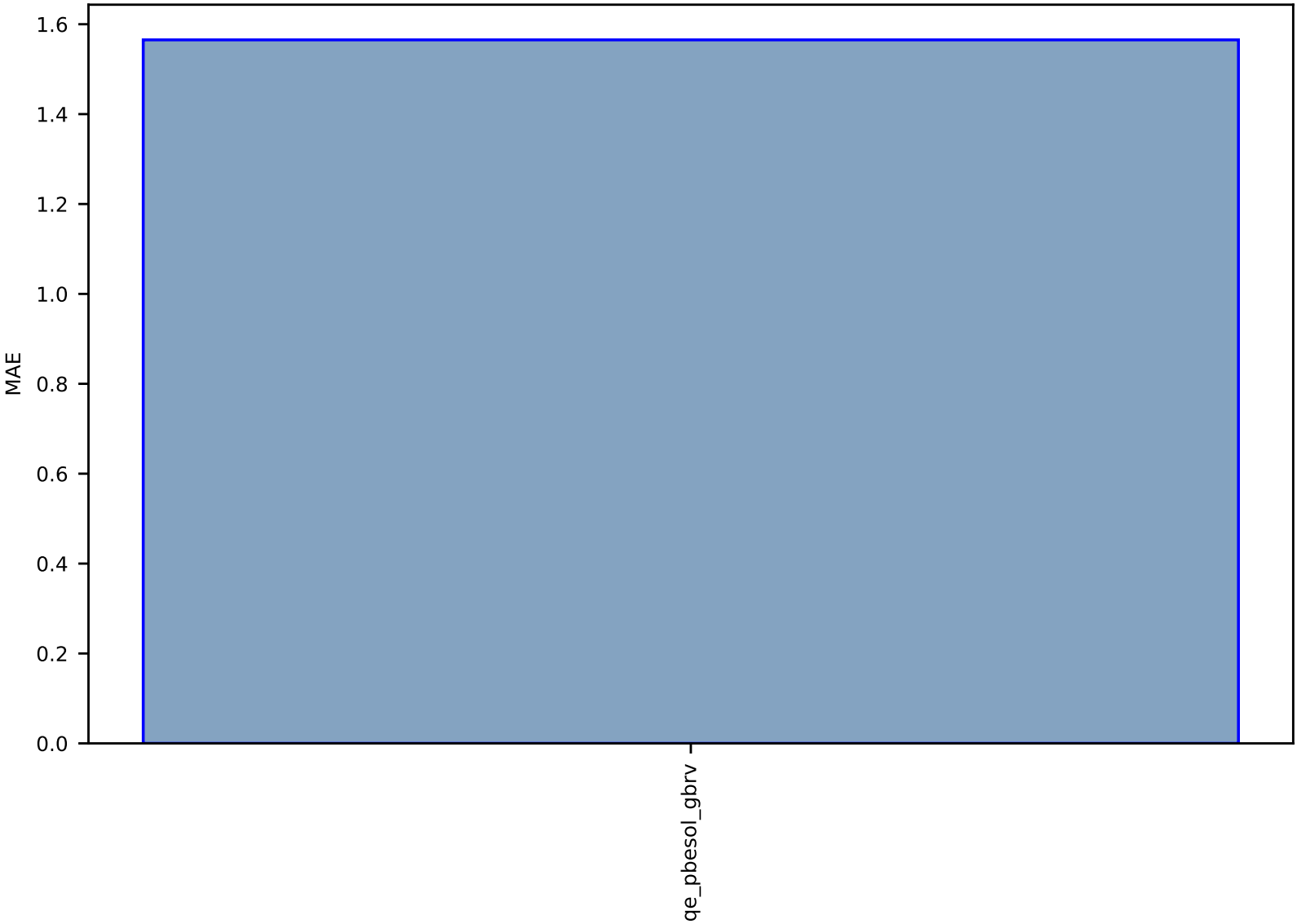


Fig S202:ES-SinglePropertyPrediction-Tc_supercon_JVASP_1014_Ta-dft_3d-test-mae.csv.zip

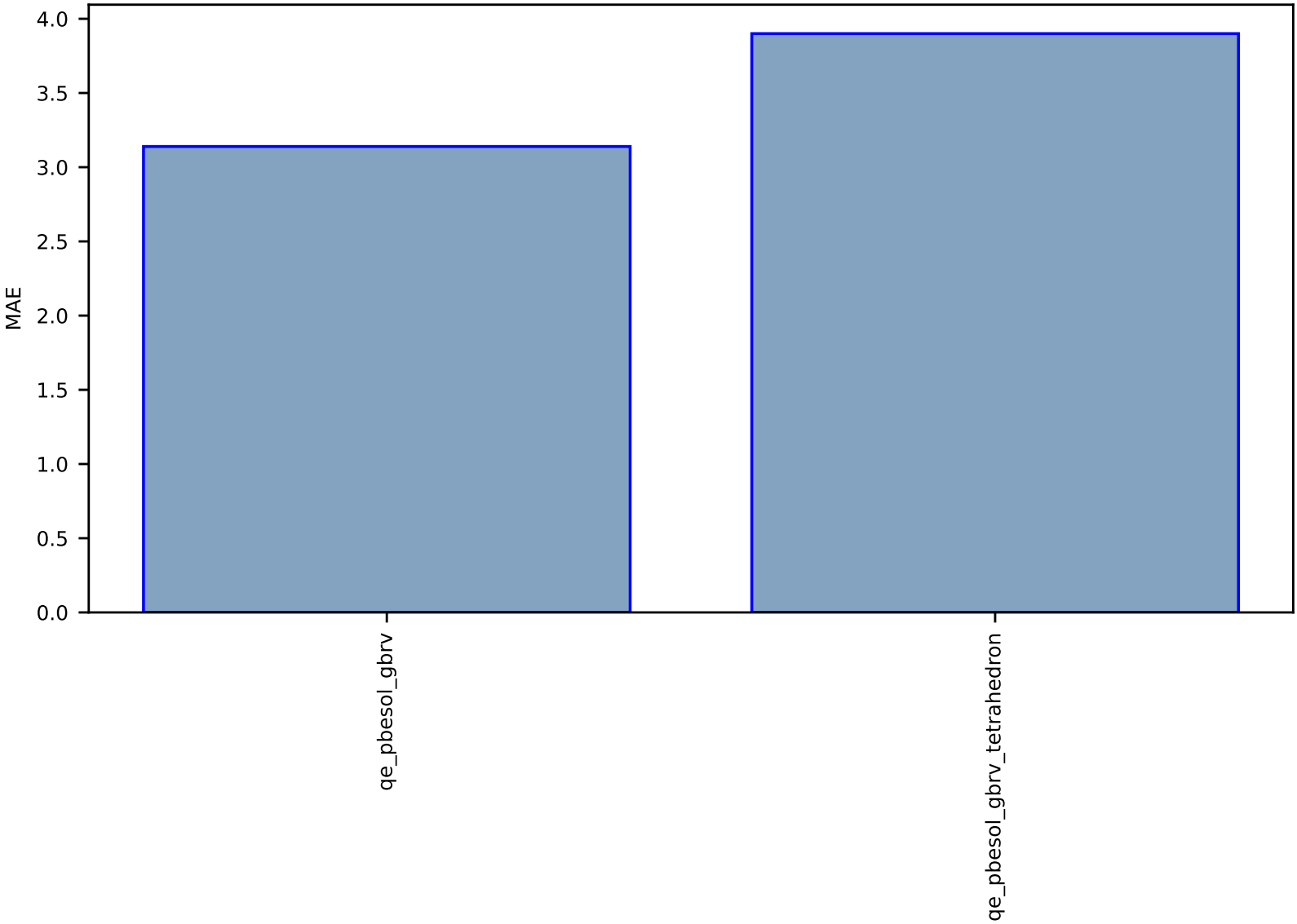


Fig S203:ES-SinglePropertyPrediction-Tc_supercon_JVASP_14960_V3Si-dft_3d-test-mae.csv.zip

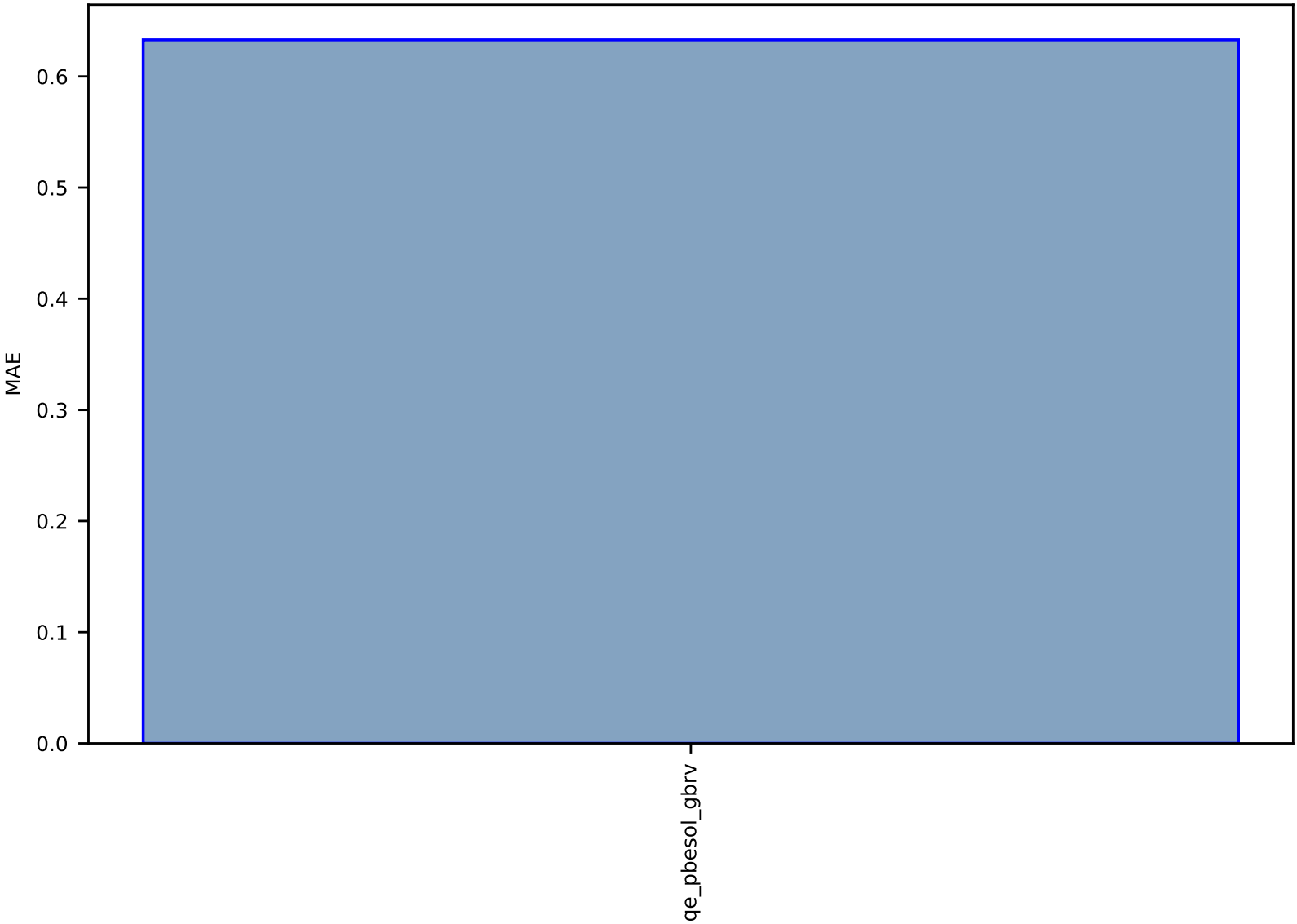


Fig S204:ES-SinglePropertyPrediction-Tc_supercon_JVASP_655_NbSe2-dft_2d-test-mae.csv.zip

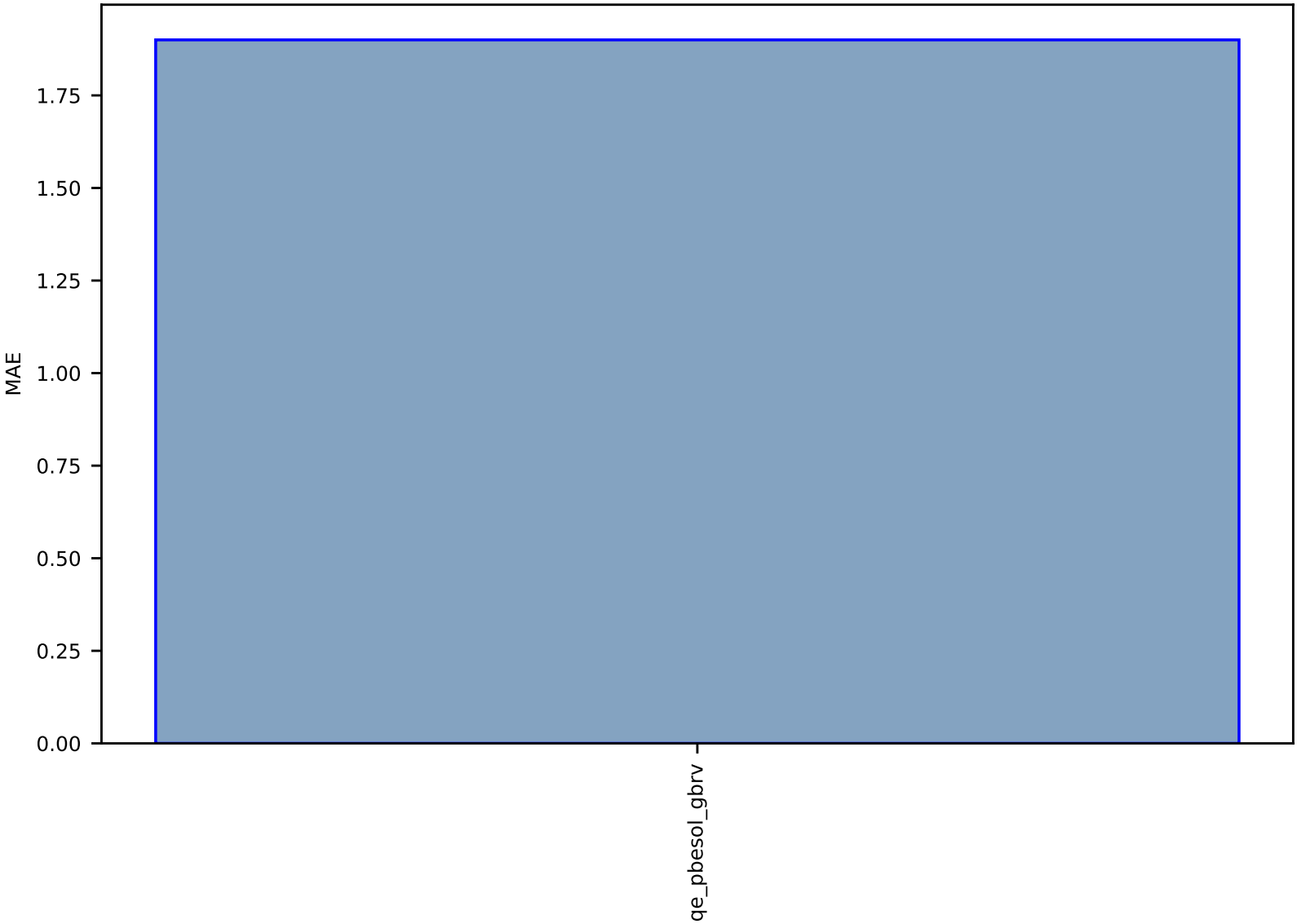


Fig S205:ES-SinglePropertyPrediction-Tc_supercon_JVASP_19679_ZrN-dft_3d-test-mae.csv.zip

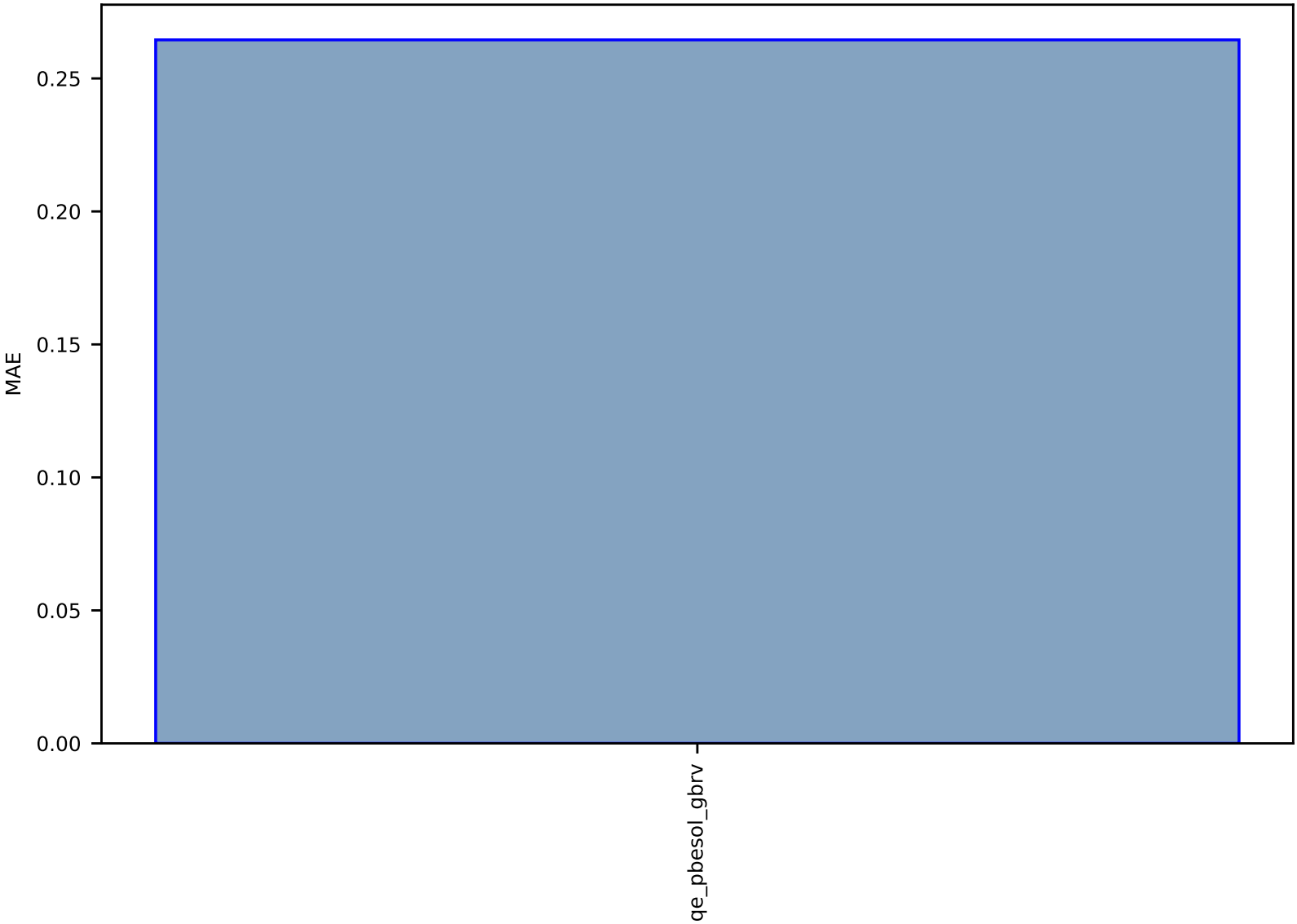


Fig S206:ES-SinglePropertyPrediction-Tc_supercon_JVASP_15938_Nb3Si-dft_3d-test-mae.csv.zip

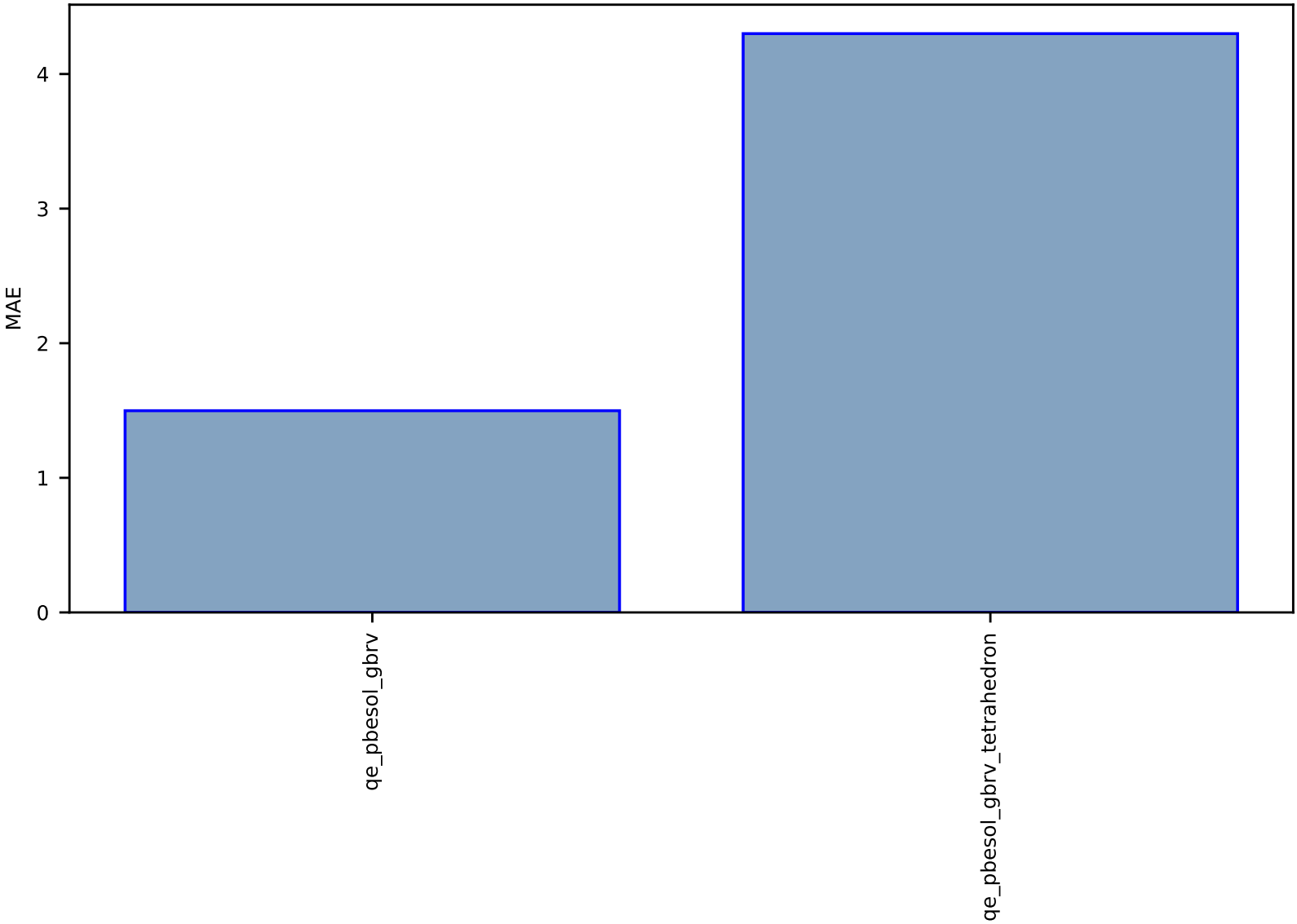


Fig S207:ES-SinglePropertyPrediction-Tc_supercon_JVASP_14837_V-dft_3d-test-mae.csv.zip

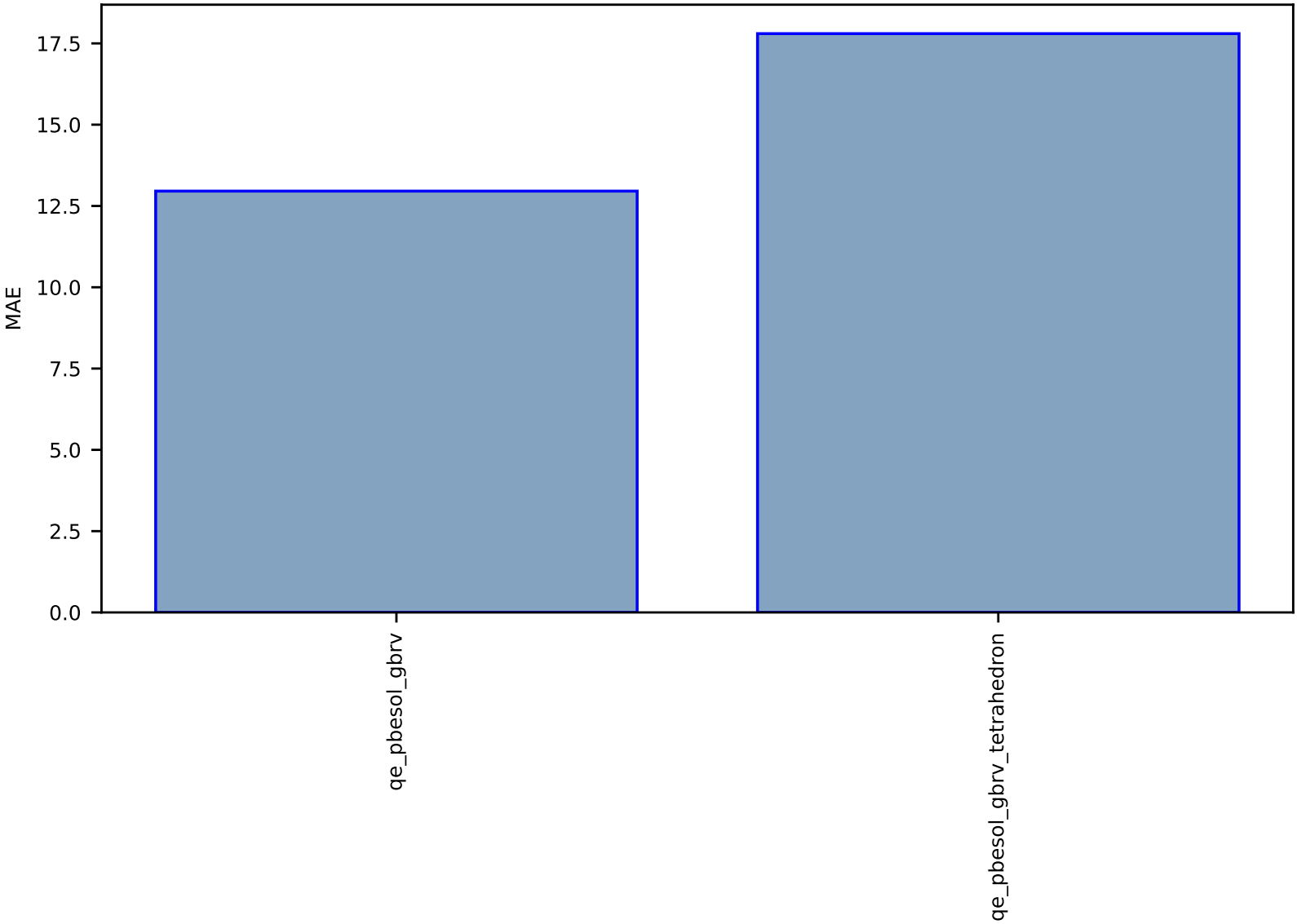


Fig S208:ES-SinglePropertyPrediction-Tc_supercon_JVASP_934_Nb-dft_3d-test-mae.csv.zip

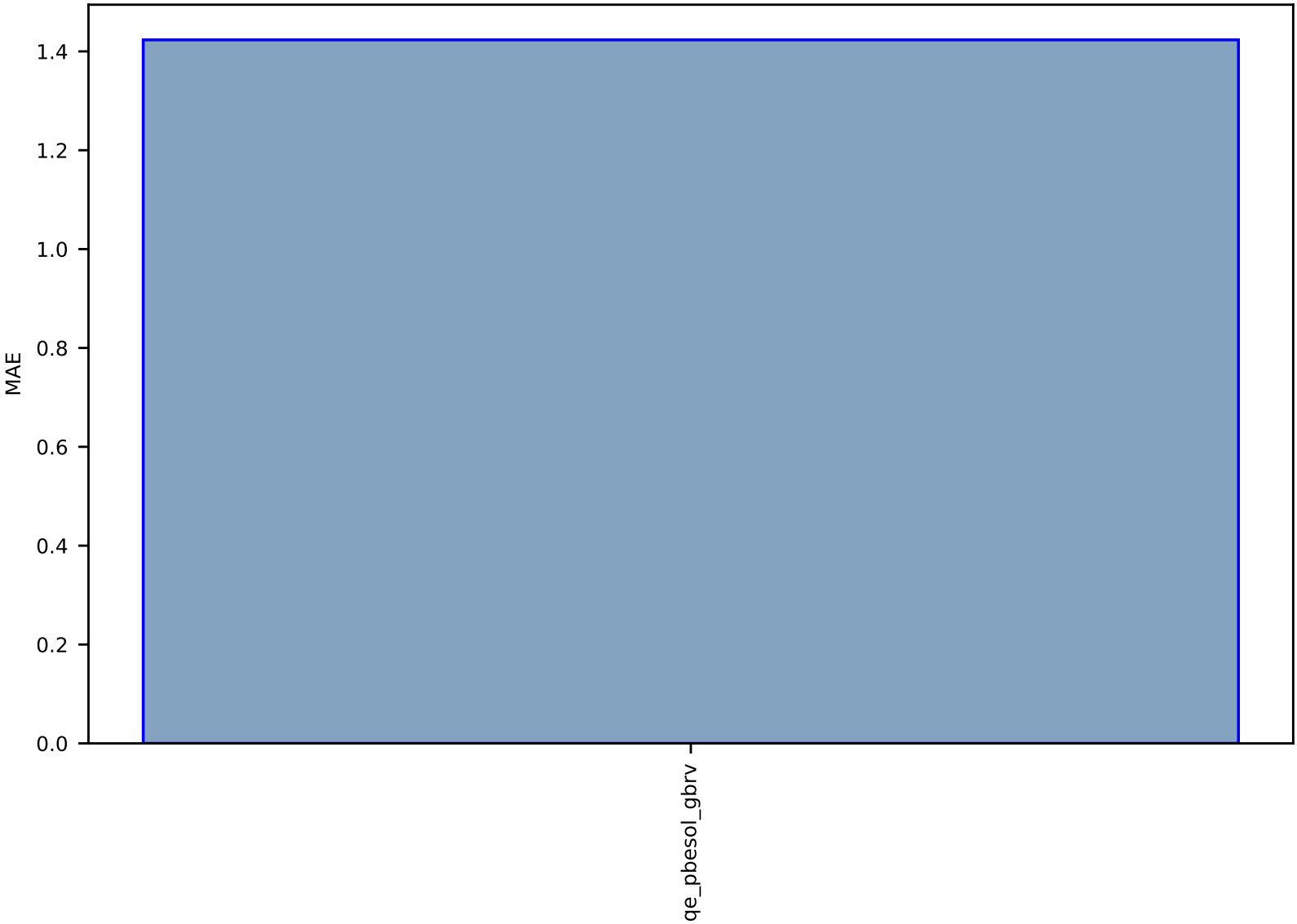


Fig S209:ES-SinglePropertyPrediction-Tc_supercon_JVASP_1151_MgB2-dft_3d-test-mae.csv.zip

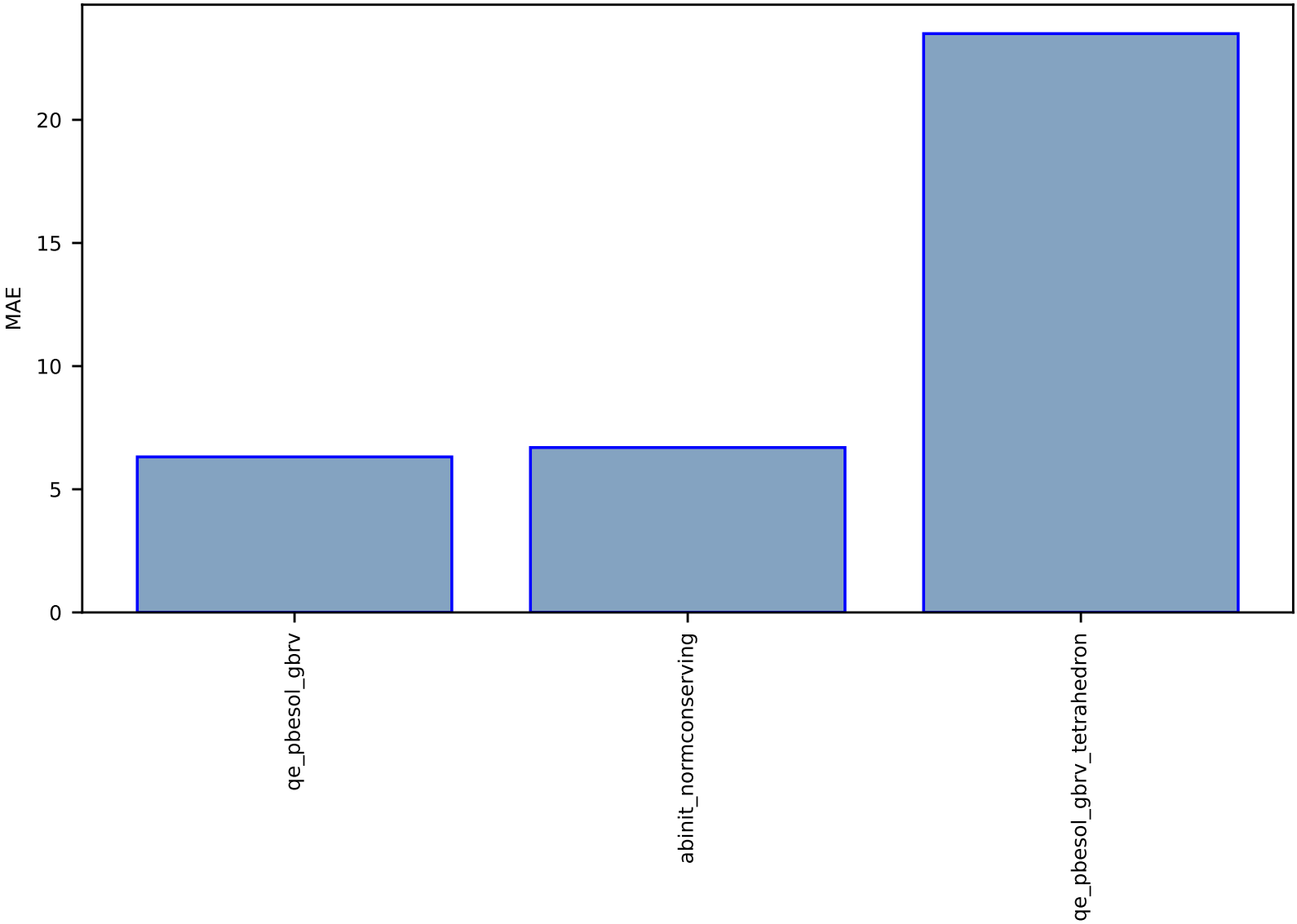


Fig S210:ES-SinglePropertyPrediction-Tc_supercon_JVASP_14492_NbO-dft_3d-test-mae.csv.zip

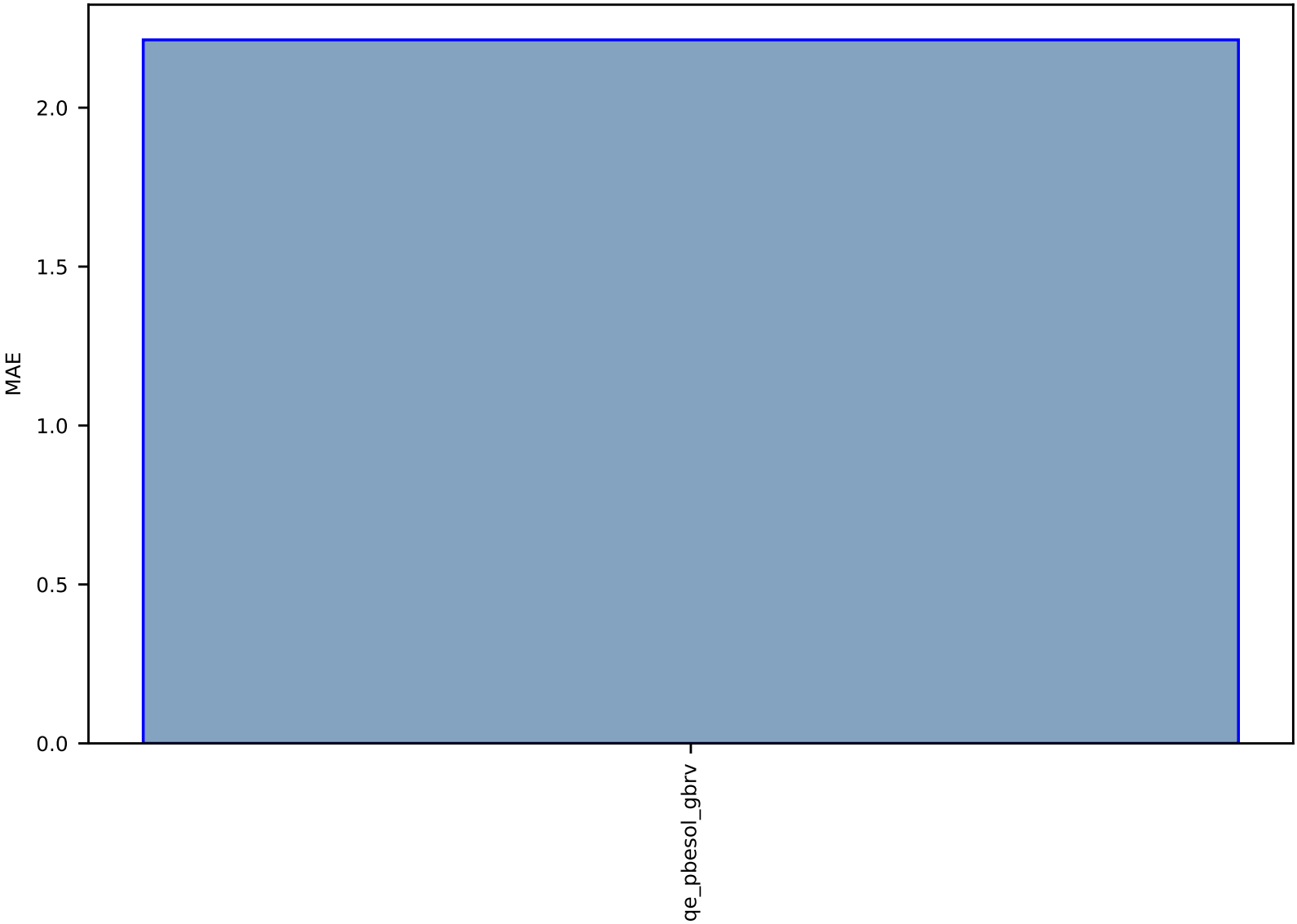


Fig S211:ES-SinglePropertyPrediction-Tc_supercon_JVASP_646_NbS2-dft_2d-test-mae.csv.zip

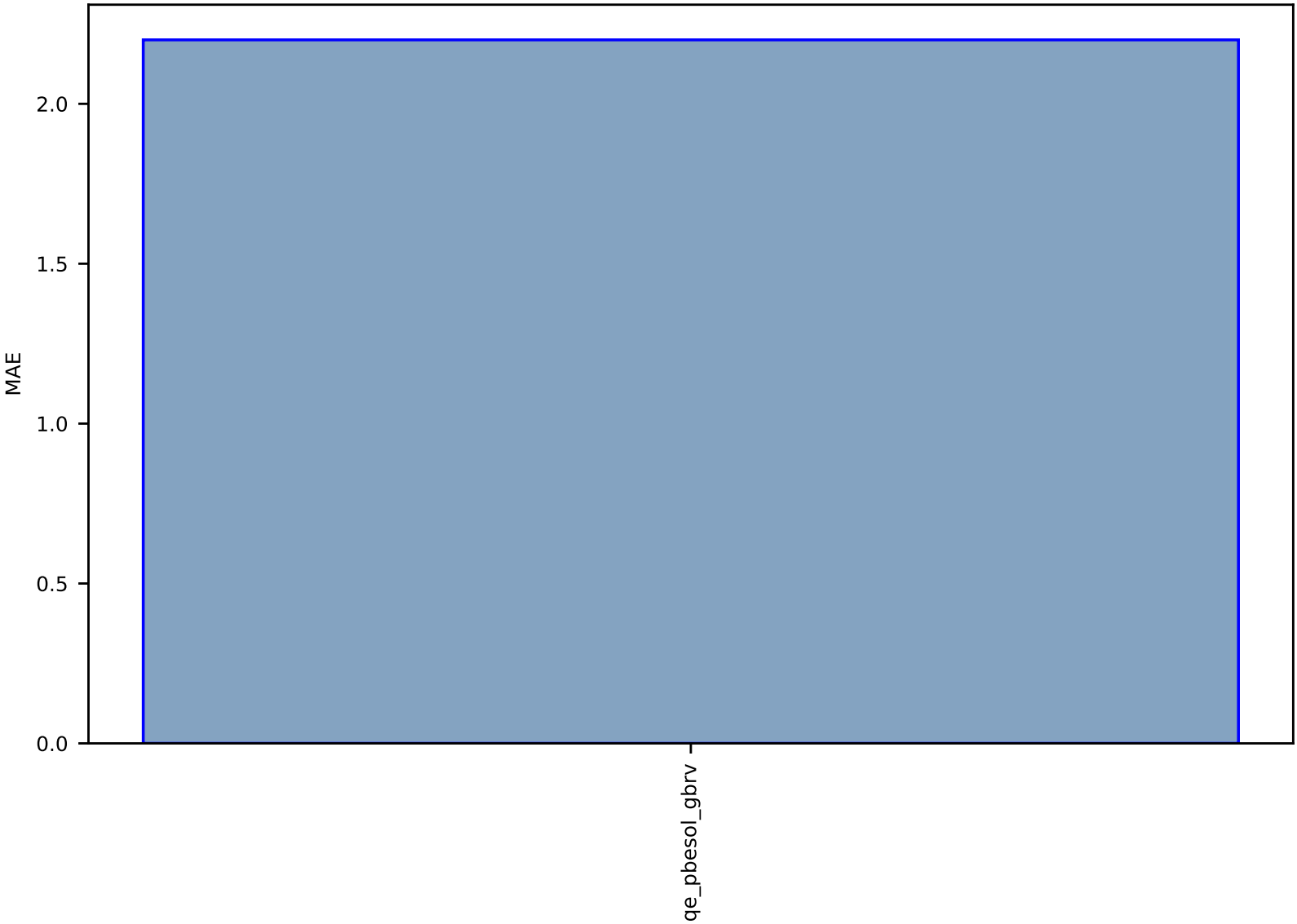


Fig S212:EXP-SinglePropertyPrediction-Tc_supercon_JVASP_1151_MgB2-dft_3d-test-mae.csv.zip

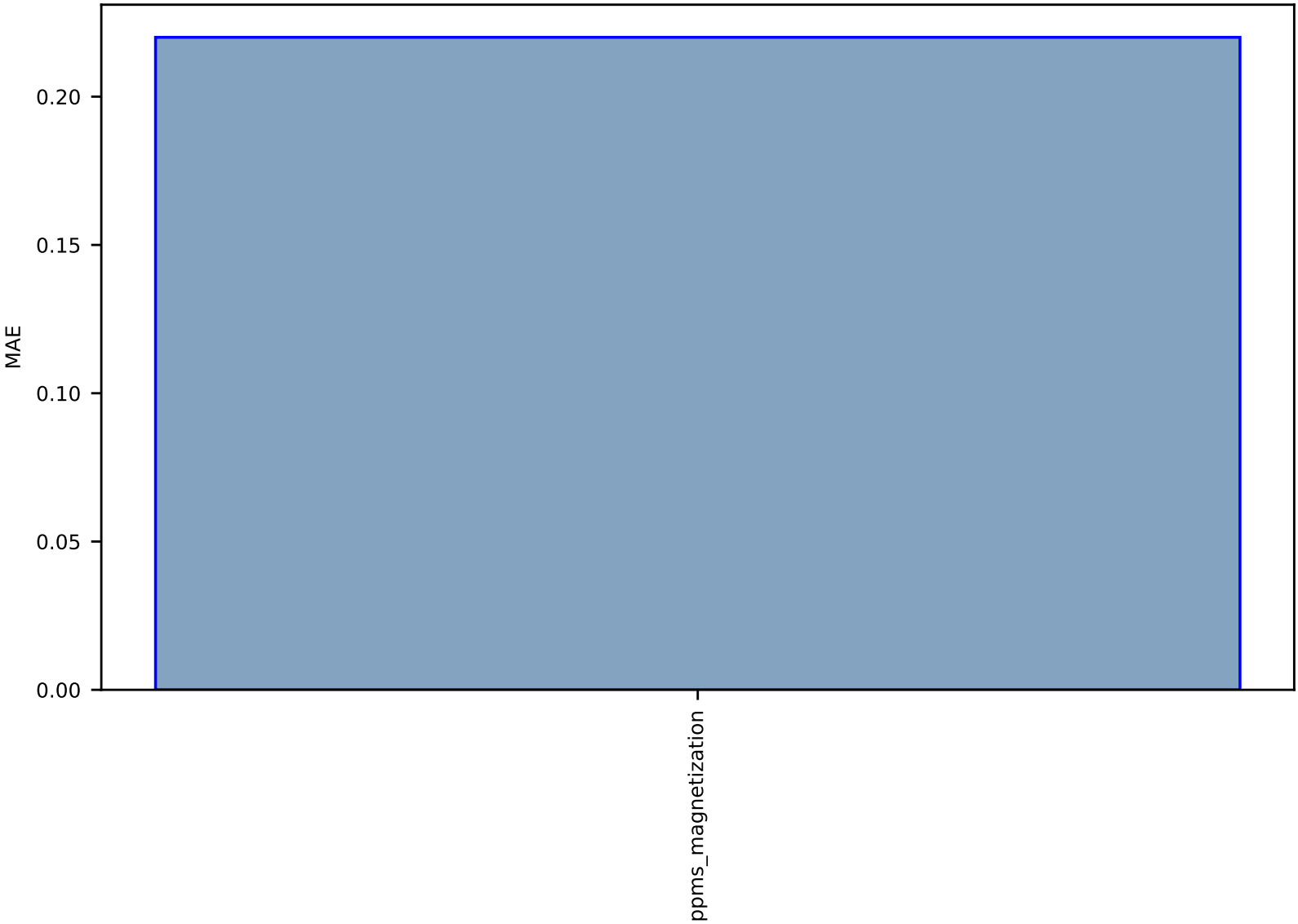


Fig S213:AI-SinglePropertyPrediction-formation_energy-mxene275-test-mae.csv.zip

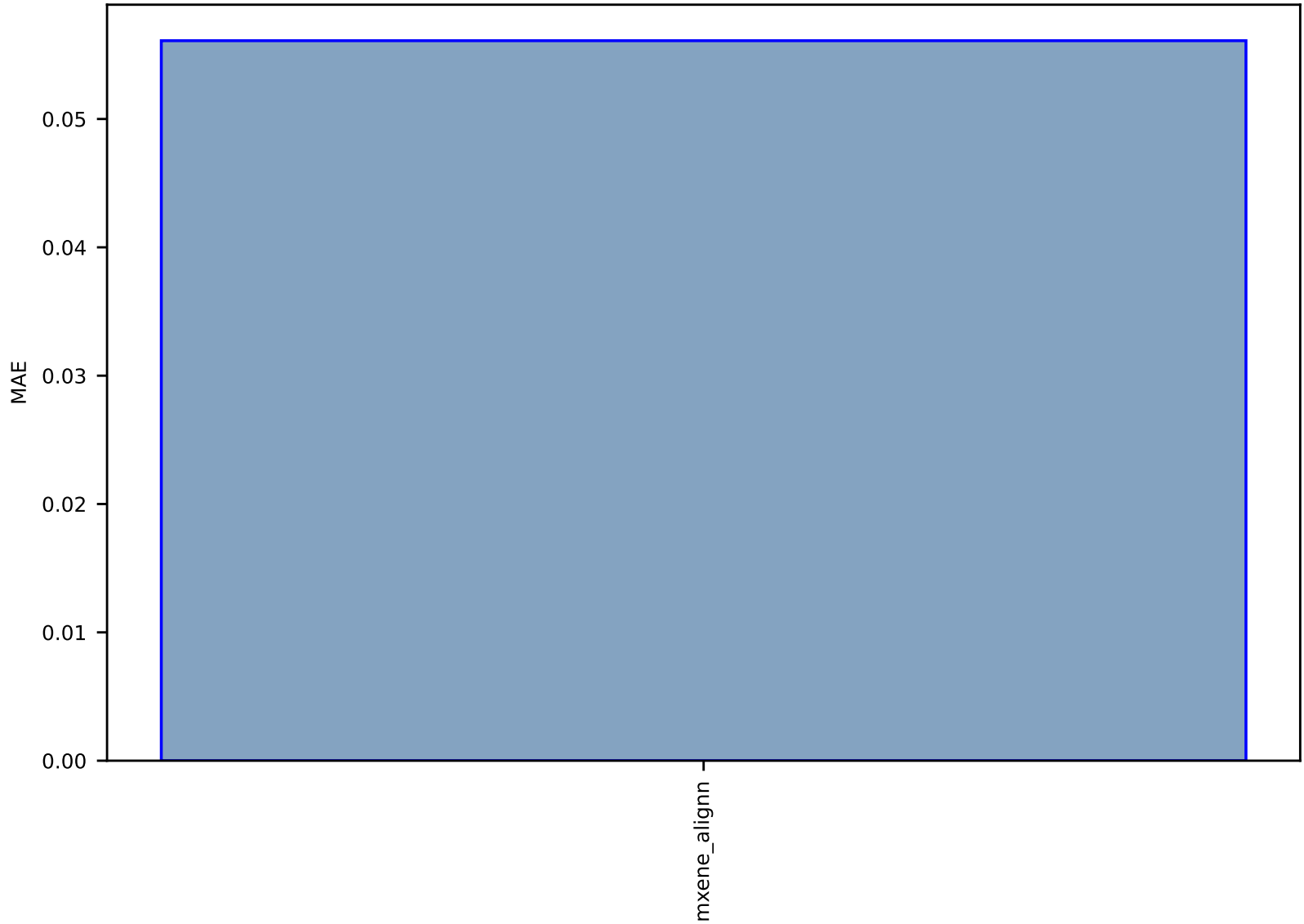


Fig S214:ES-SinglePropertyPrediction-bulk_modulus_JVASP_7809_ZrO2-dft_3d-test-mae.csv.zip

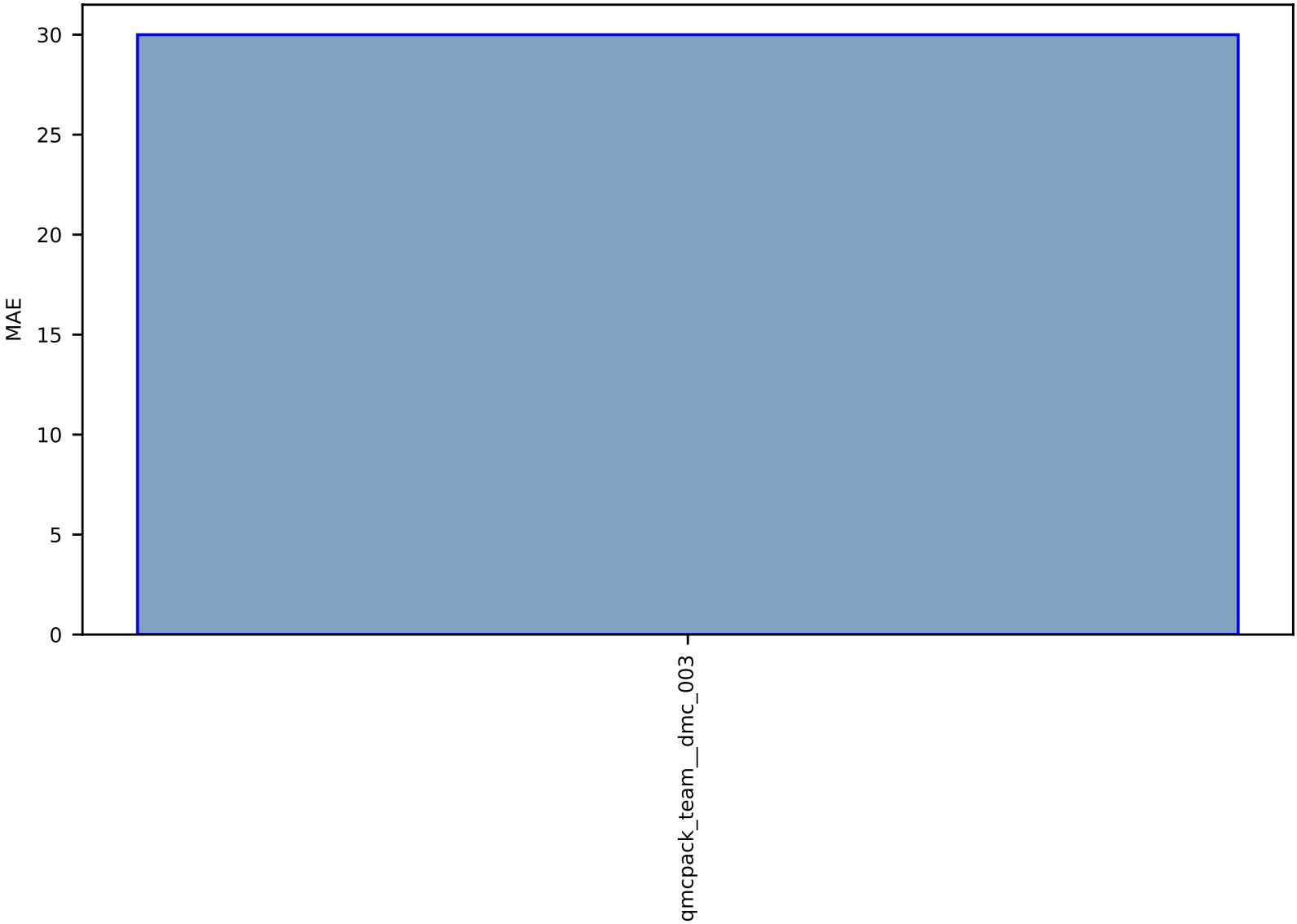


Fig S215:ES-SinglePropertyPrediction-bulk_modulus_JVASP_9147_HfO2-dft_3d-test-mae.csv.zip

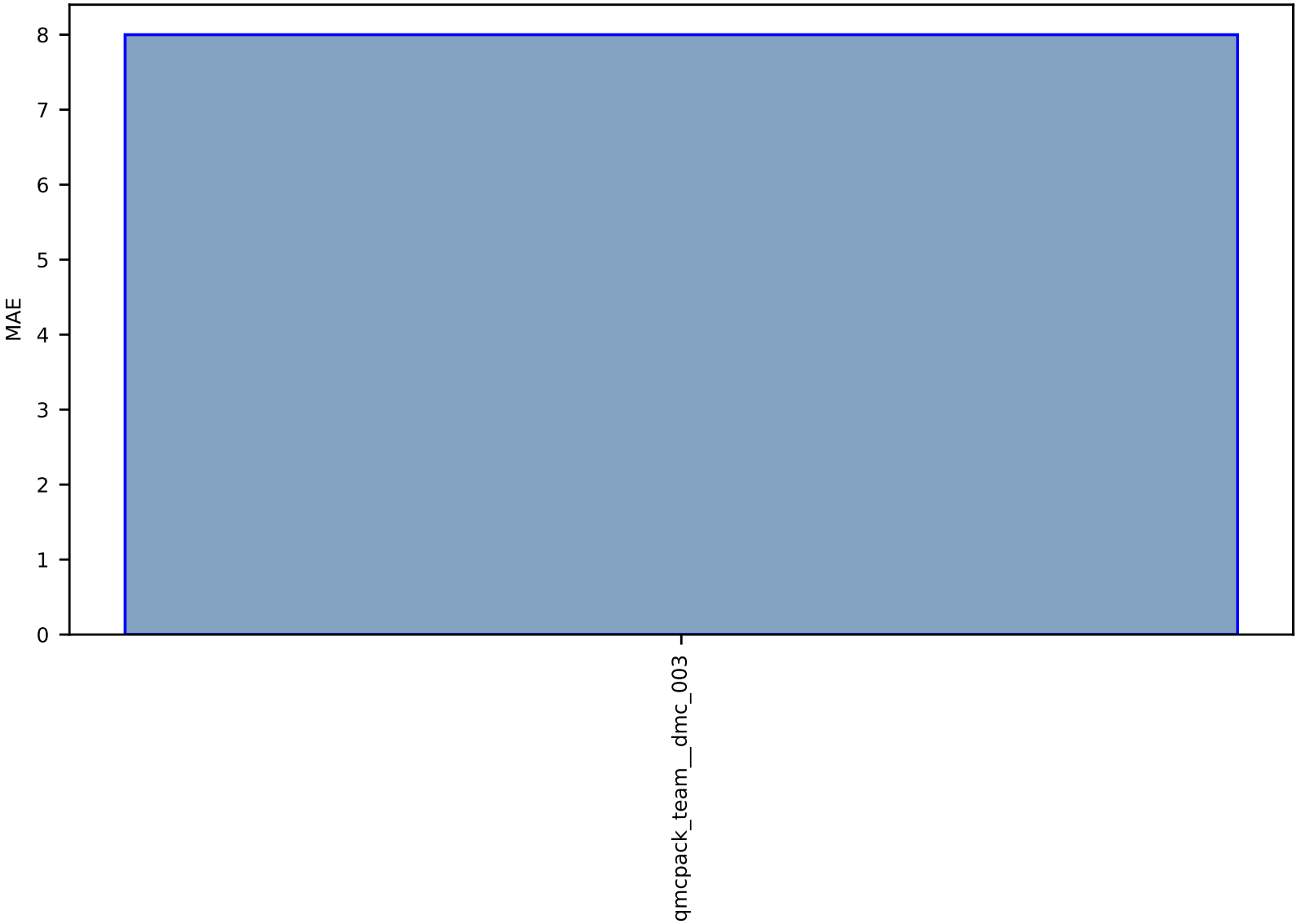


Fig S216:ES-SinglePropertyPrediction-bulk_modulus_JVASP_14590_Hf-dft_3d-test-mae.csv.zip

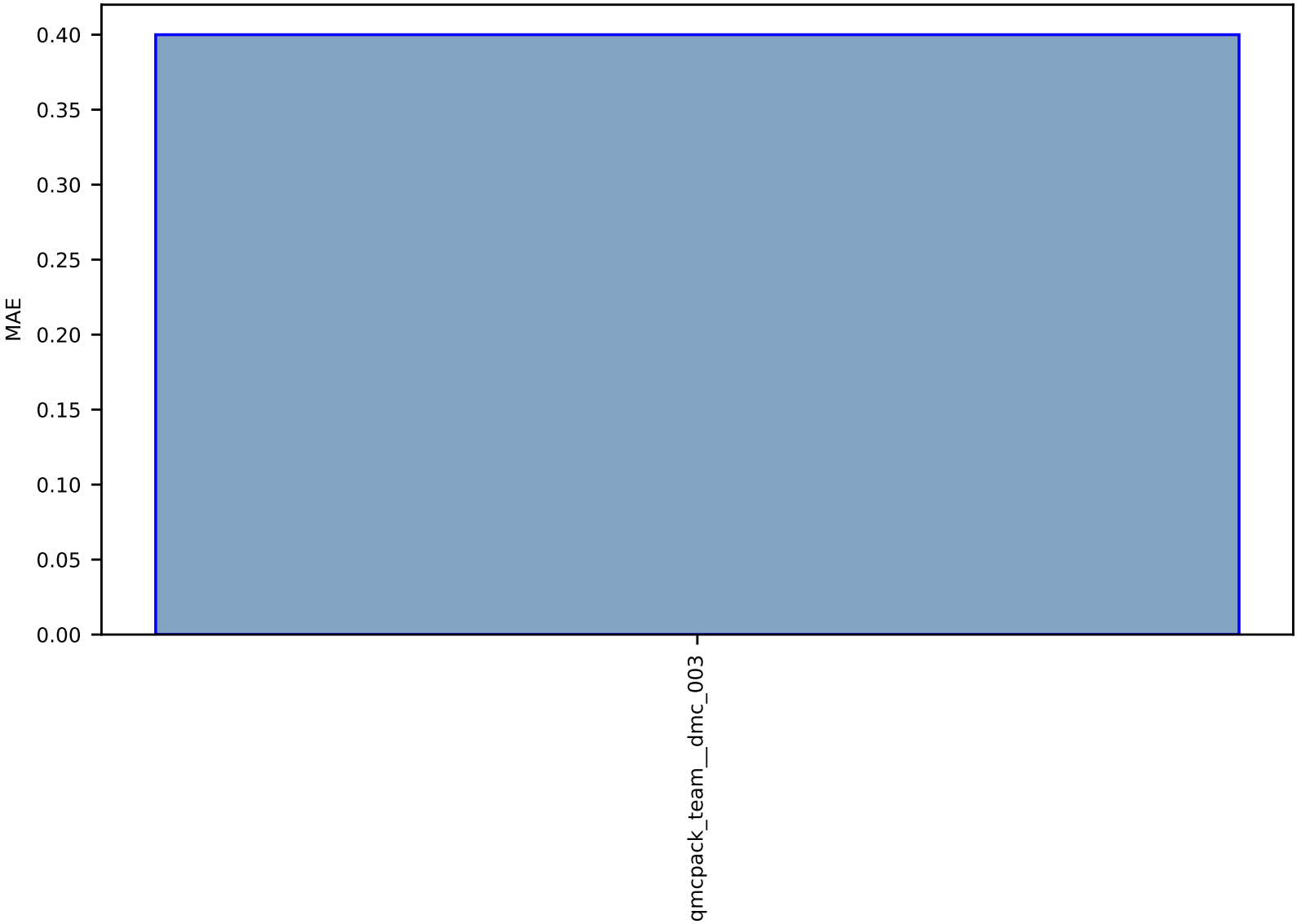


Fig S217:ES-SinglePropertyPrediction-bulk_modulus_JVASP_7871_NiO-dft_3d-test-mae.csv.zip

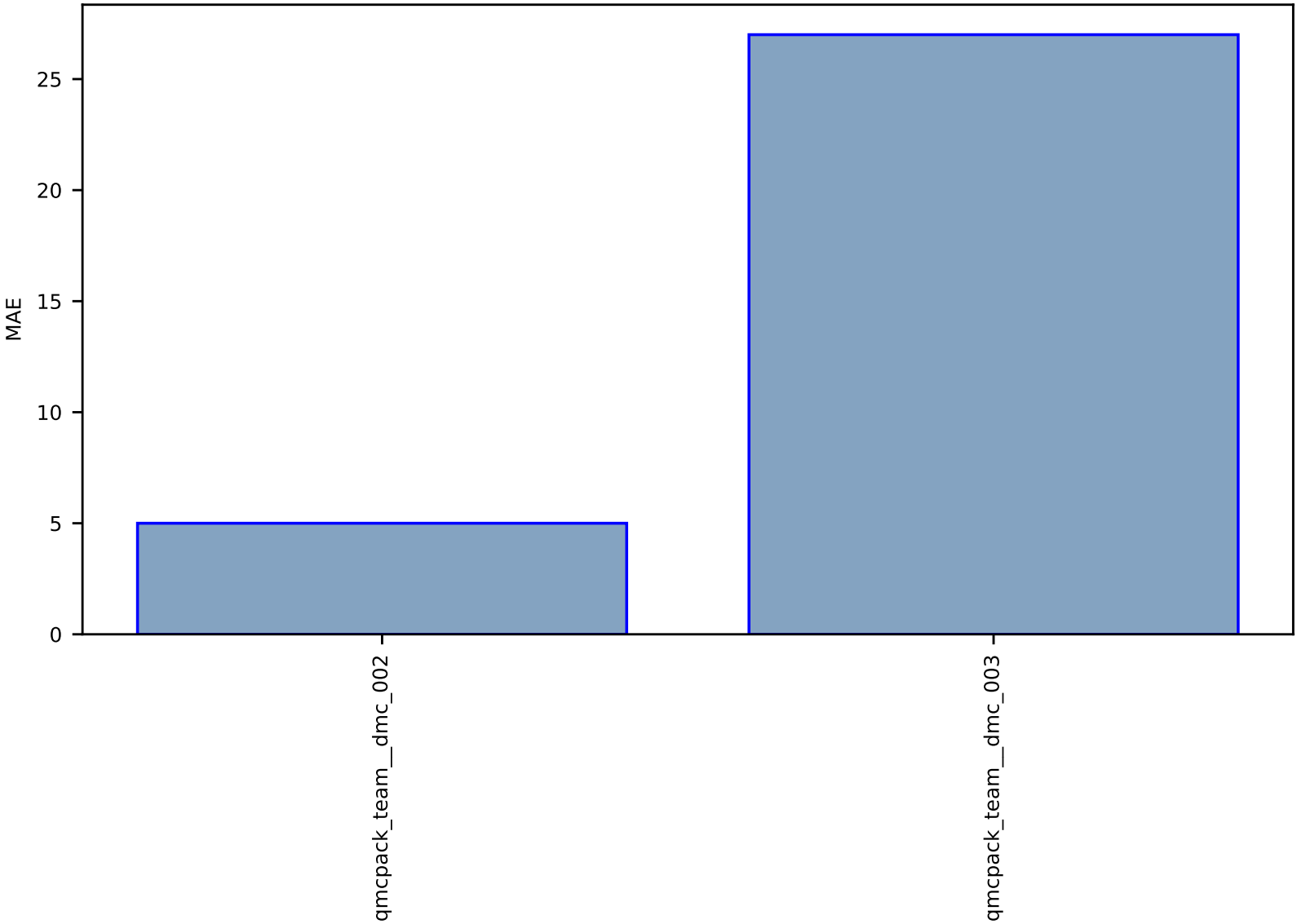


Fig S218:ES-SinglePropertyPrediction-bulk_modulus_JVASP_14612_Zr-dft_3d-test-mae.csv.zip

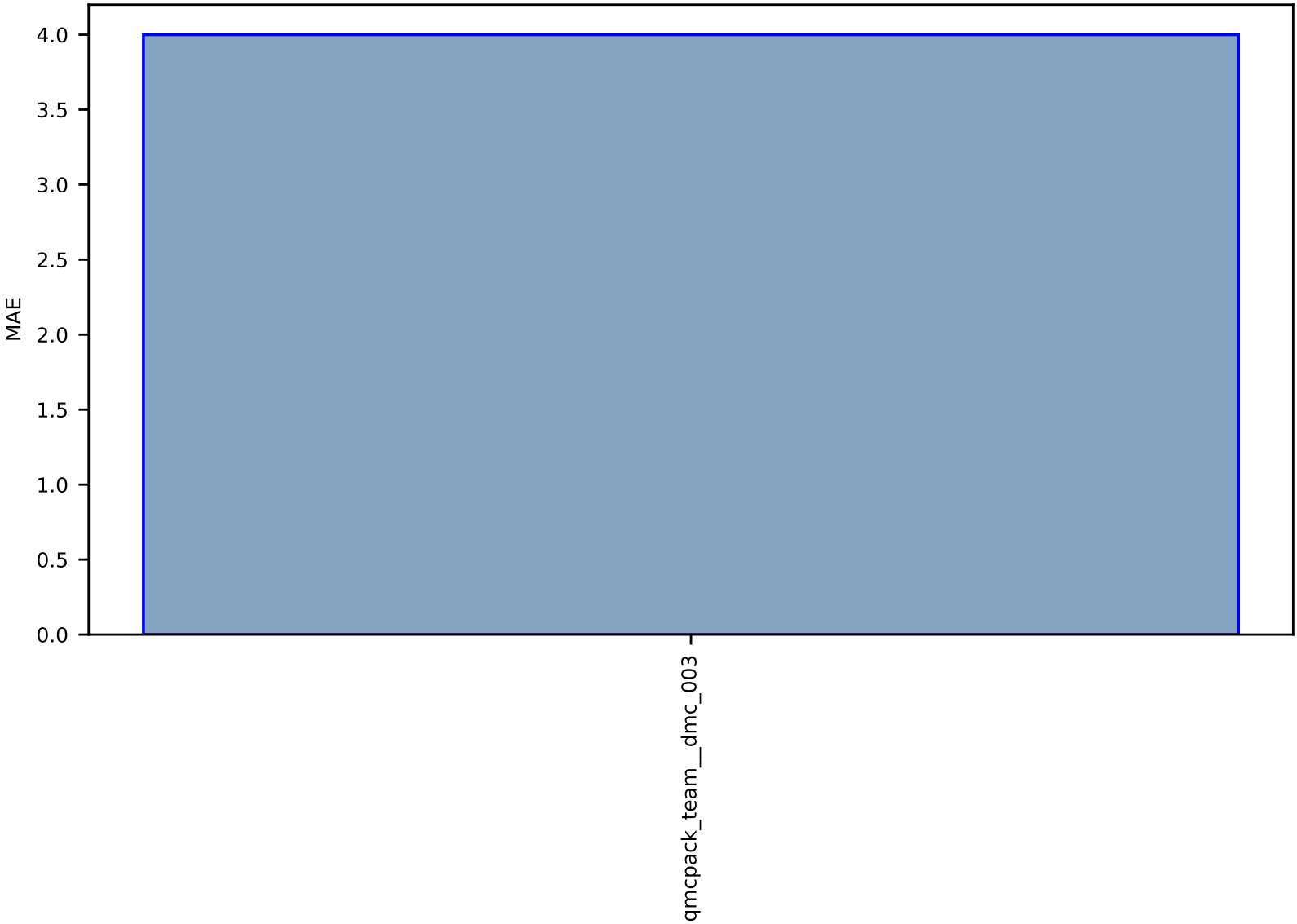


Fig S219:ES-SinglePropertyPrediction-bulk_modulus_JVASP_34249_HfO2-dft_3d-test-mae.csv.zip

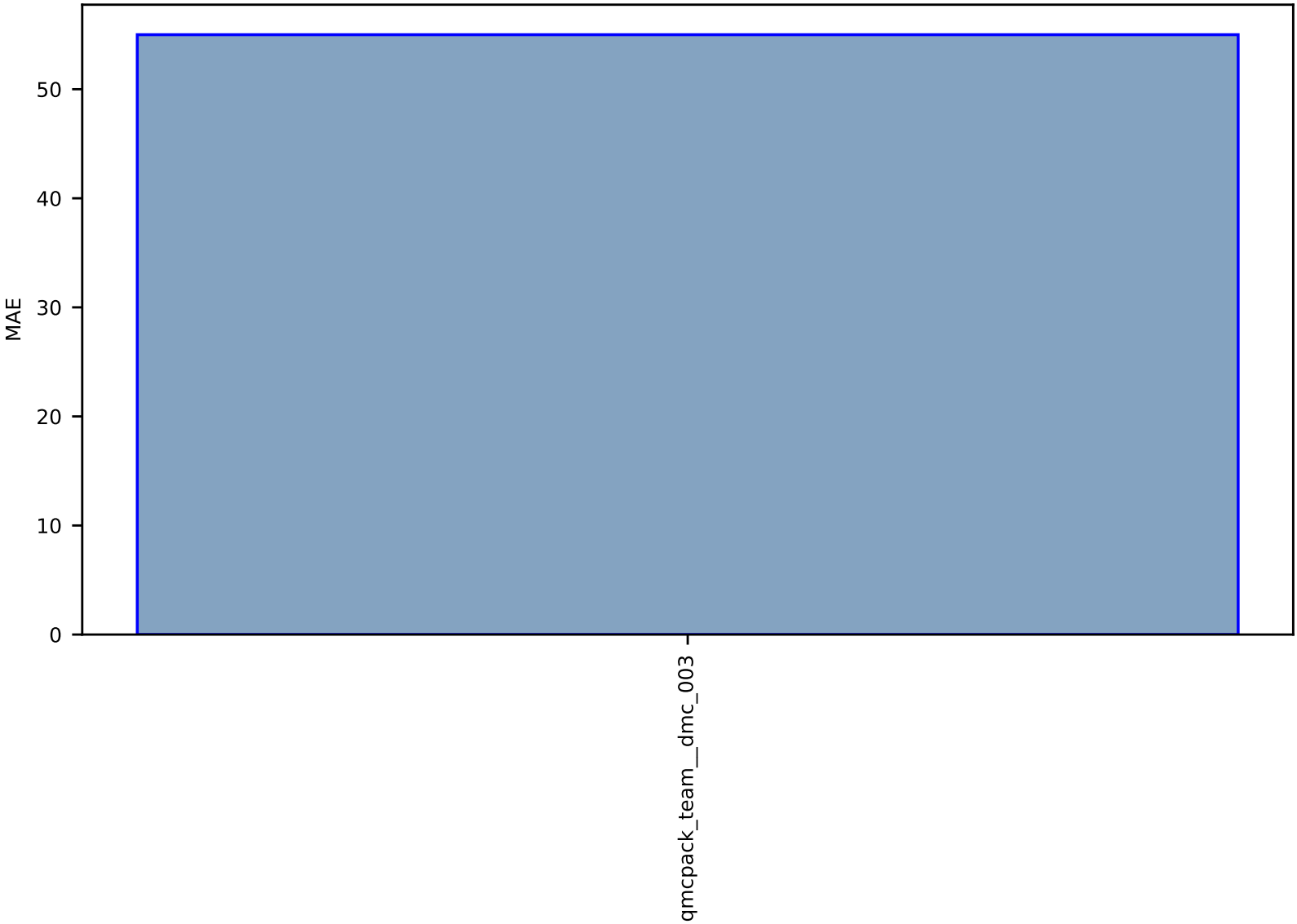


Fig S220:ES-SinglePropertyPrediction-bulk_modulus_JVASP_350_ZrO2-dft_3d-test-mae.csv.zip

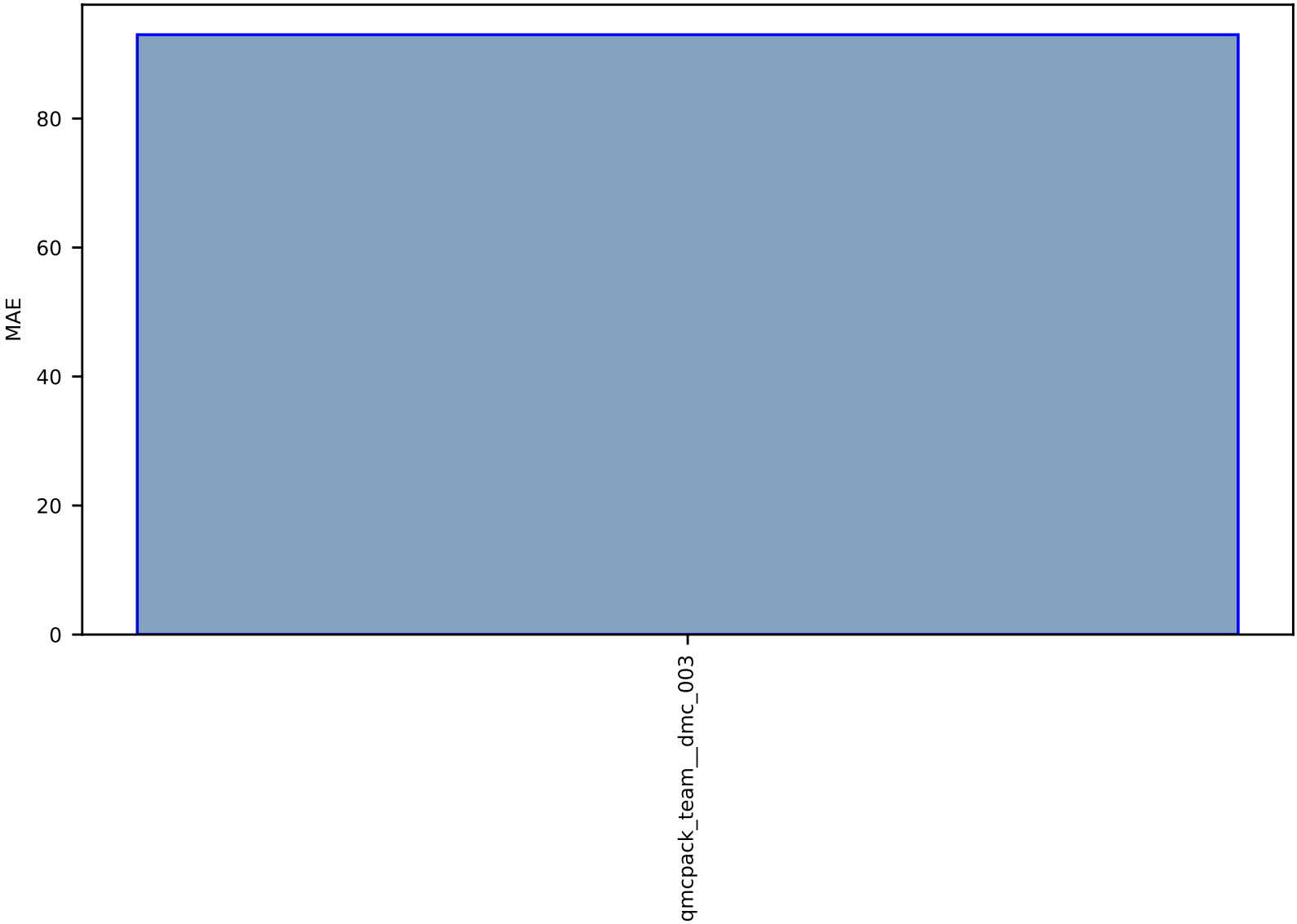


Fig S221:ES-SinglePropertyPrediction-bulk_modulus_JVASP_113_ZrO2-dft_3d-test-mae.csv.zip

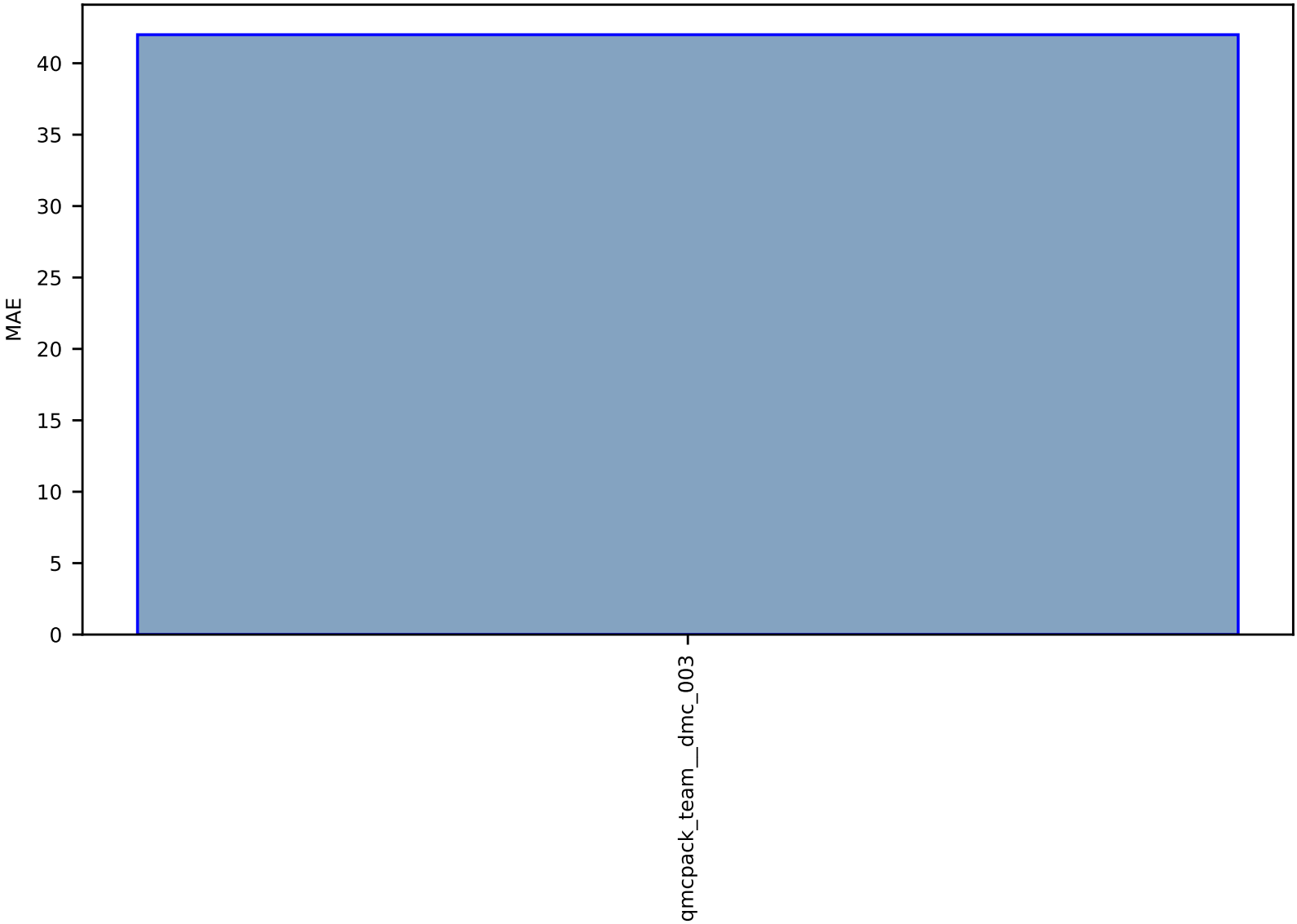


Fig S222:ES-SinglePropertyPrediction-epsx_JVASP_1393_GaP-dft_3d-test-mae.csv.zip

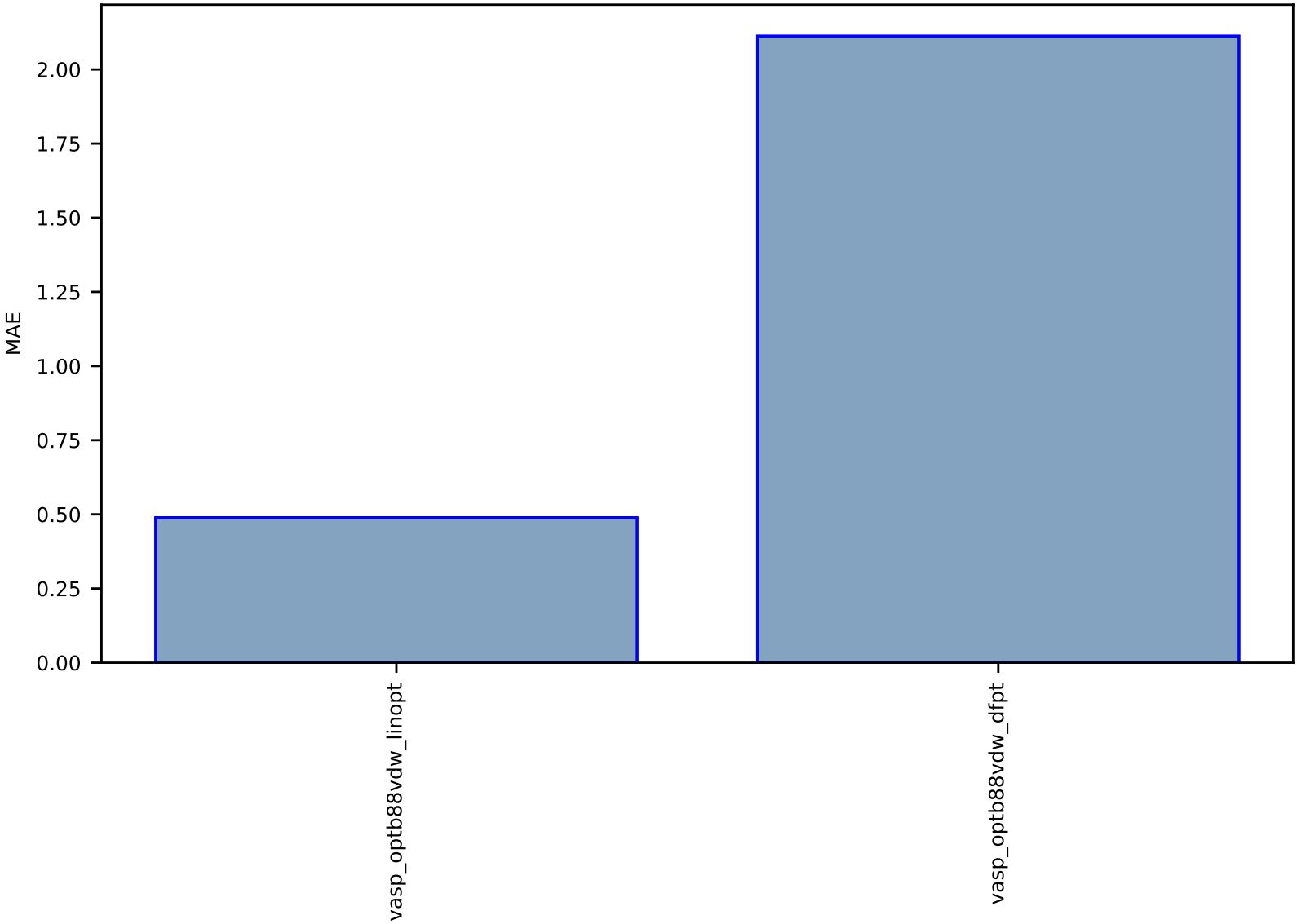


Fig S223:ES-SinglePropertyPrediction-epsx_JVASP_72_WS2-dft_3d-test-mae.csv.zip

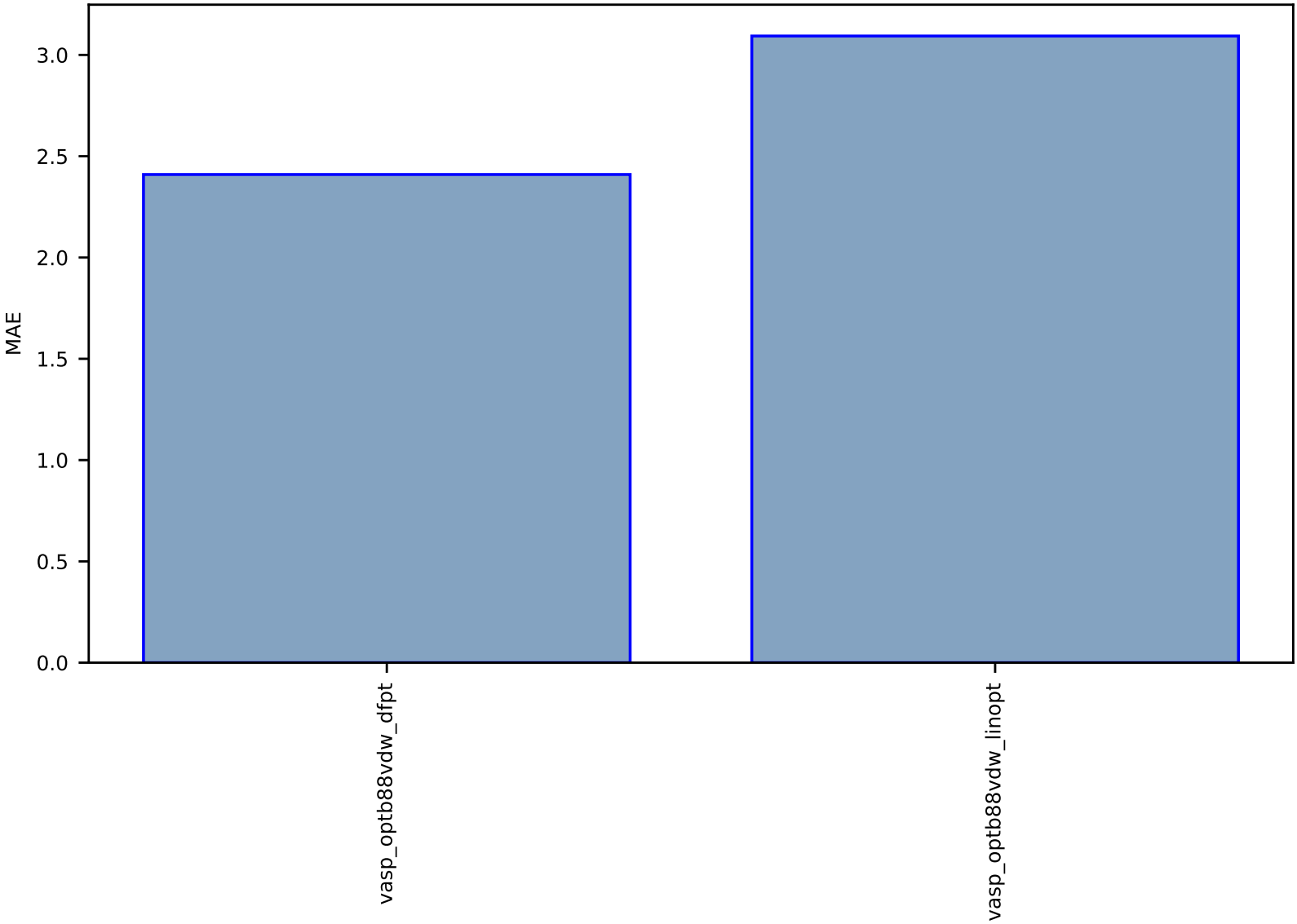


Fig S224:ES-SinglePropertyPrediction-epsx_JVASP_1702_ZnS-dft_3d-test-mae.csv.zip

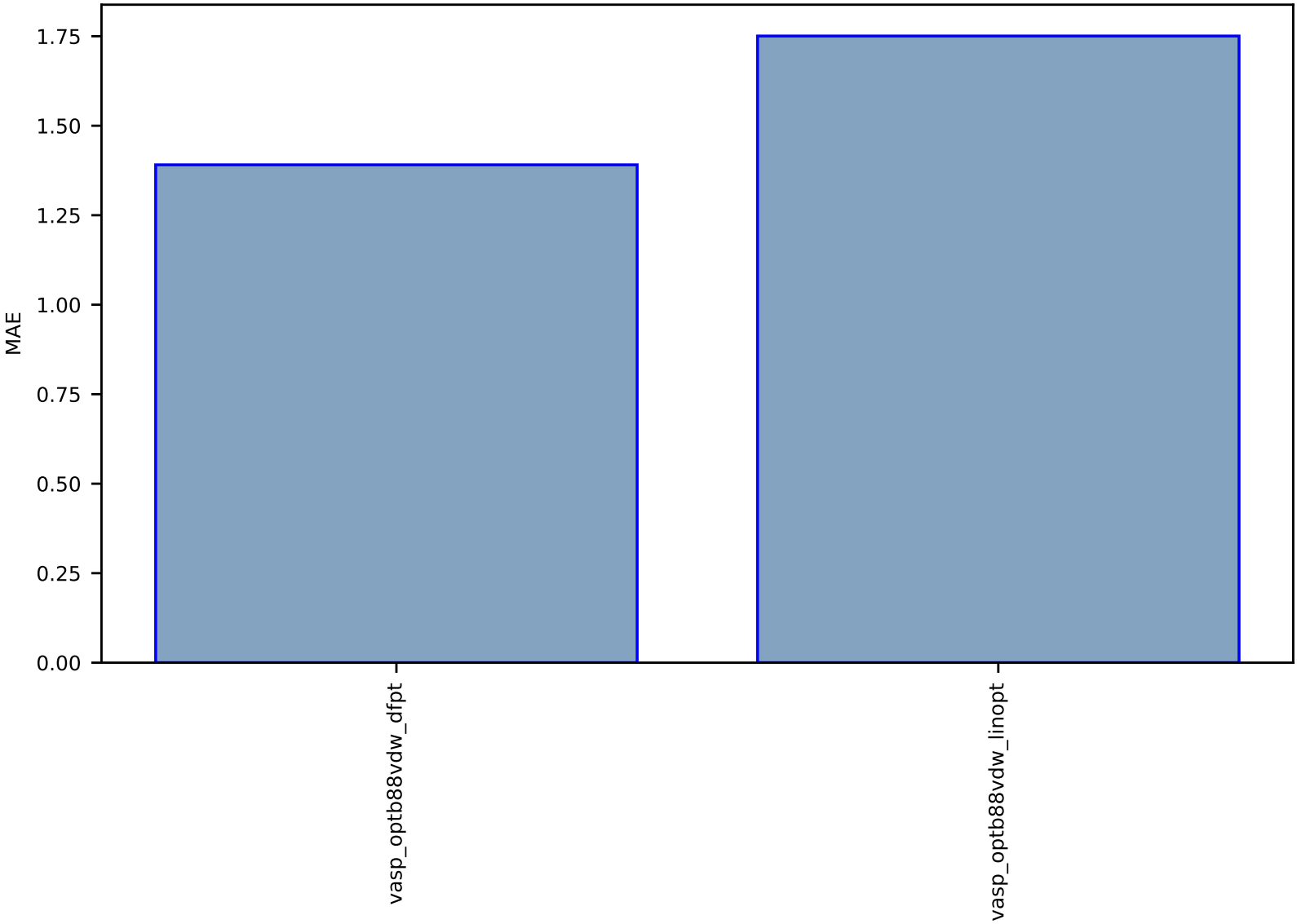


Fig S225:ES-SinglePropertyPrediction-epsx_JVASP_1408_AlSb-dft_3d-test-mae.csv.zip

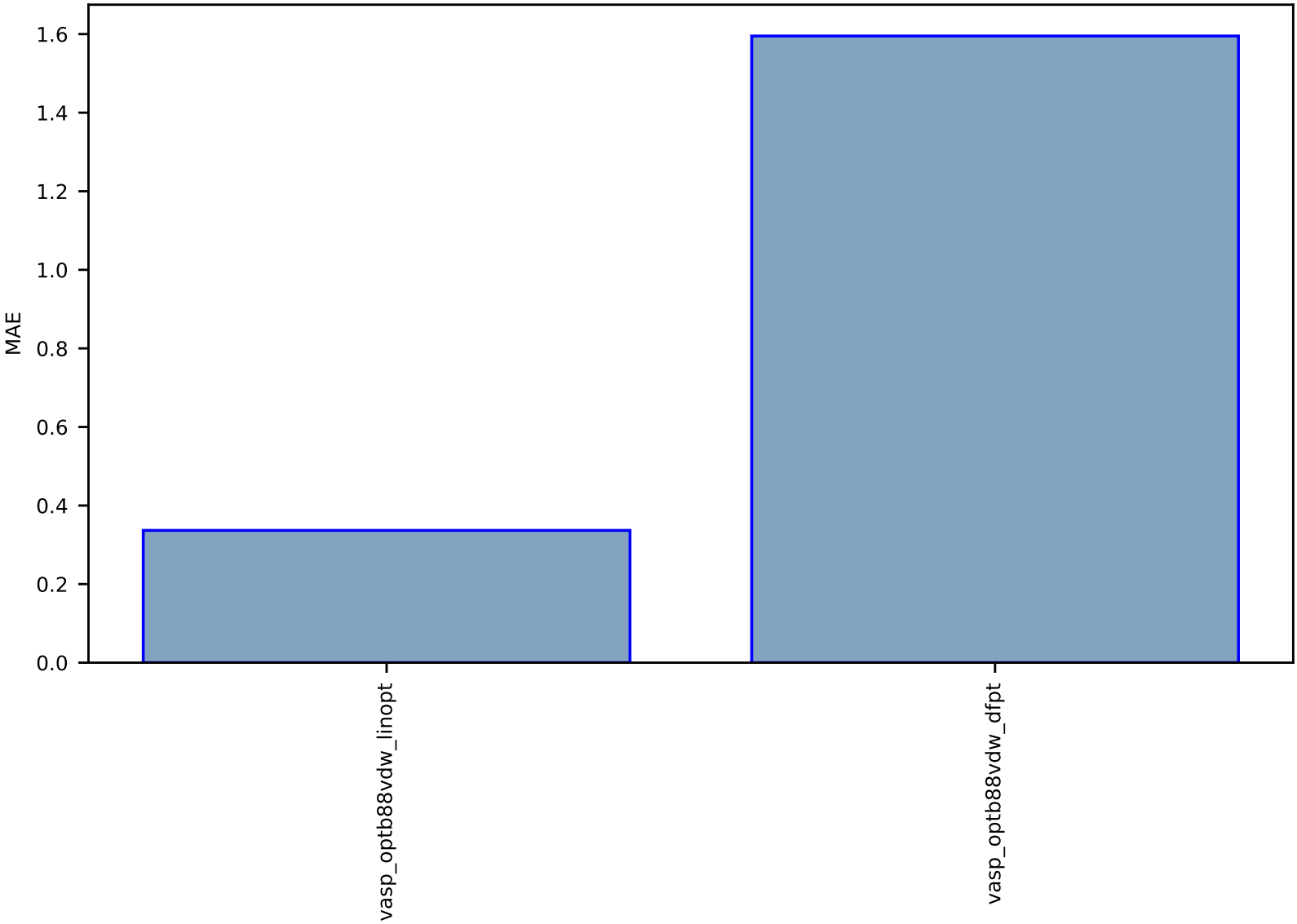


Fig S226:ES-SinglePropertyPrediction-epsx_JVASP_75_WSe2-dft_3d-test-mae.csv.zip

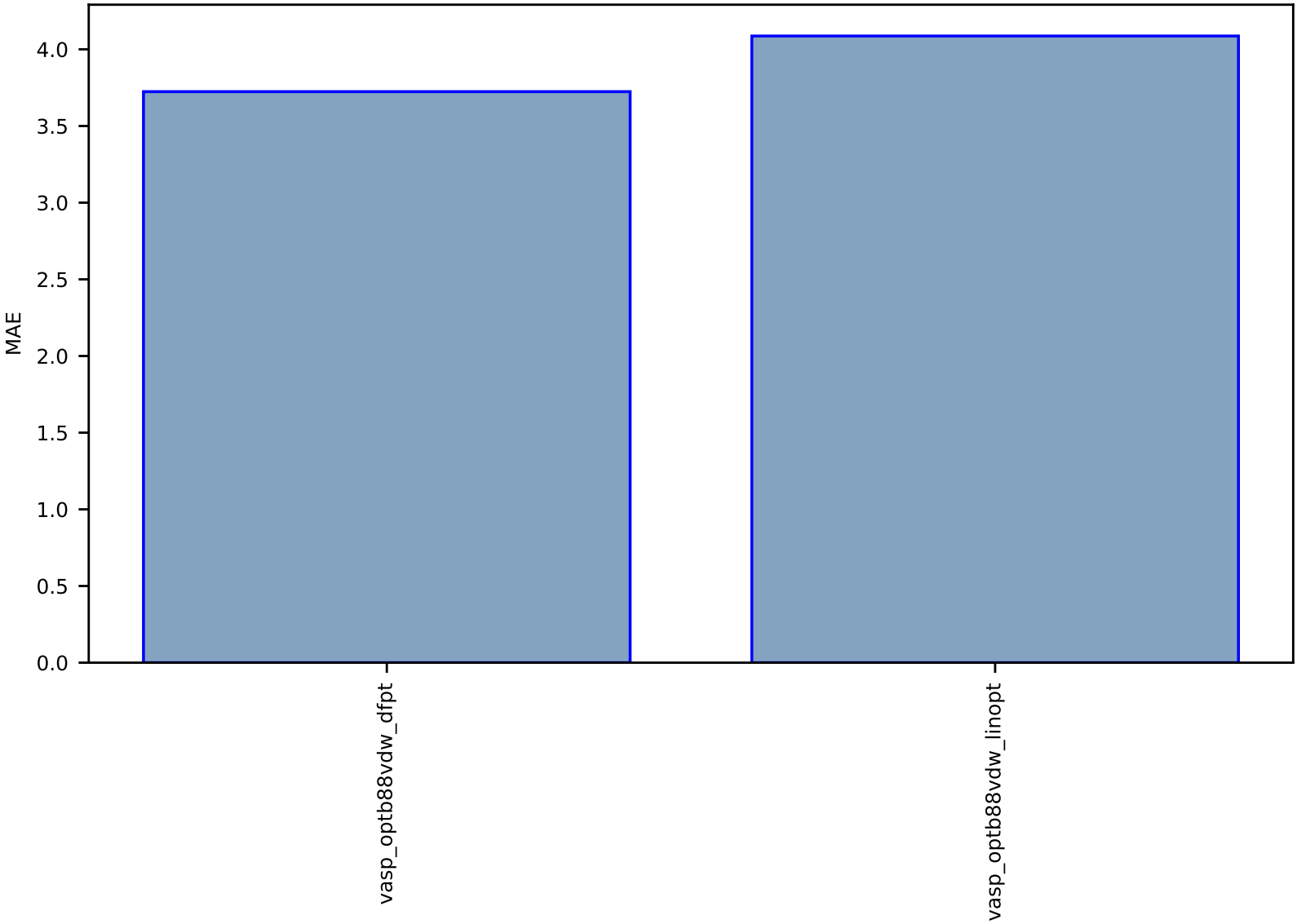


Fig S227:ES-SinglePropertyPrediction-epsx_JVASP_2376_ZnSiP2-dft_3d-test-mae.csv.zip

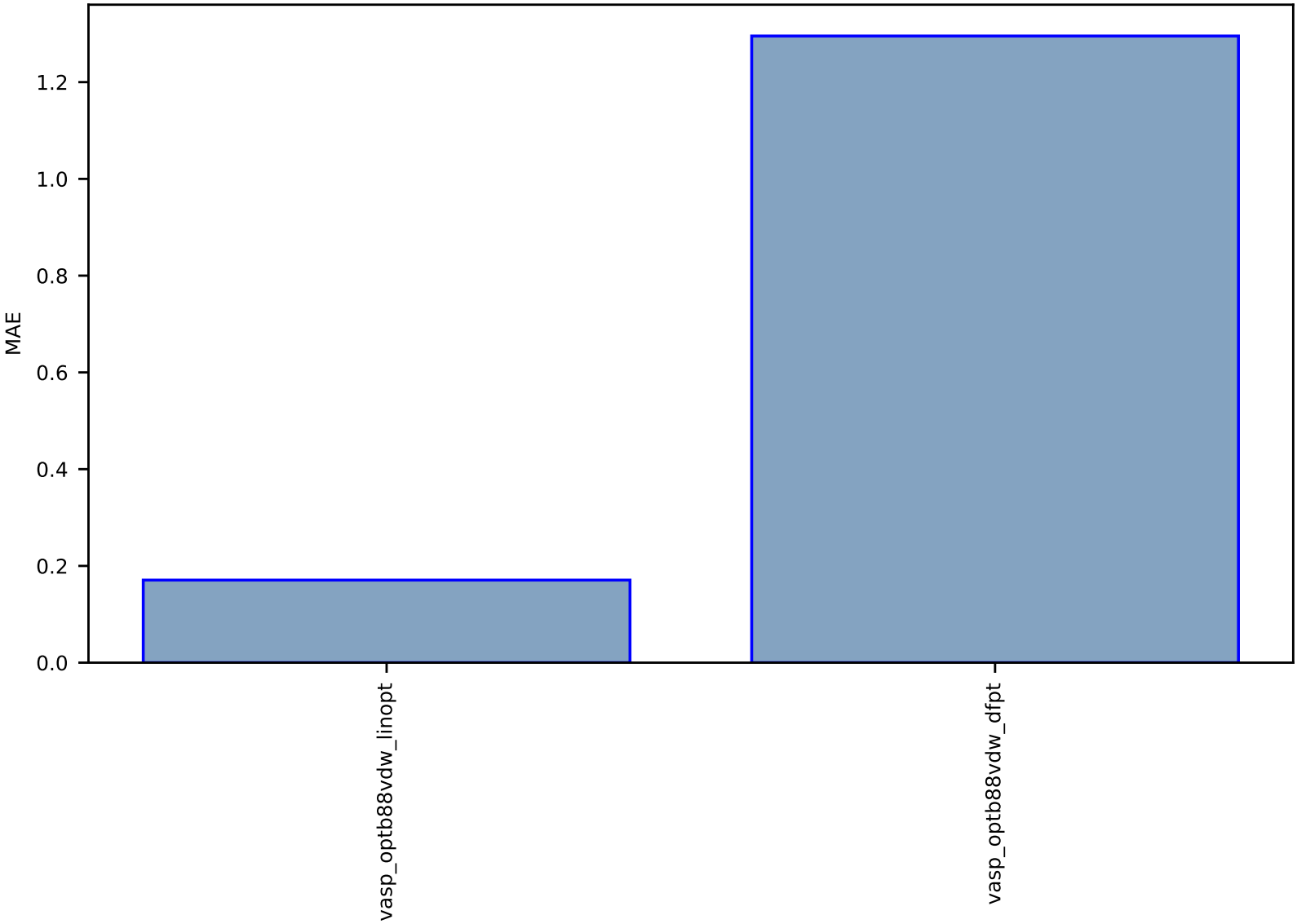


Fig S228:ES-SinglePropertyPrediction-epsx_JVASP_182_SiC-dft_3d-test-mae.csv.zip

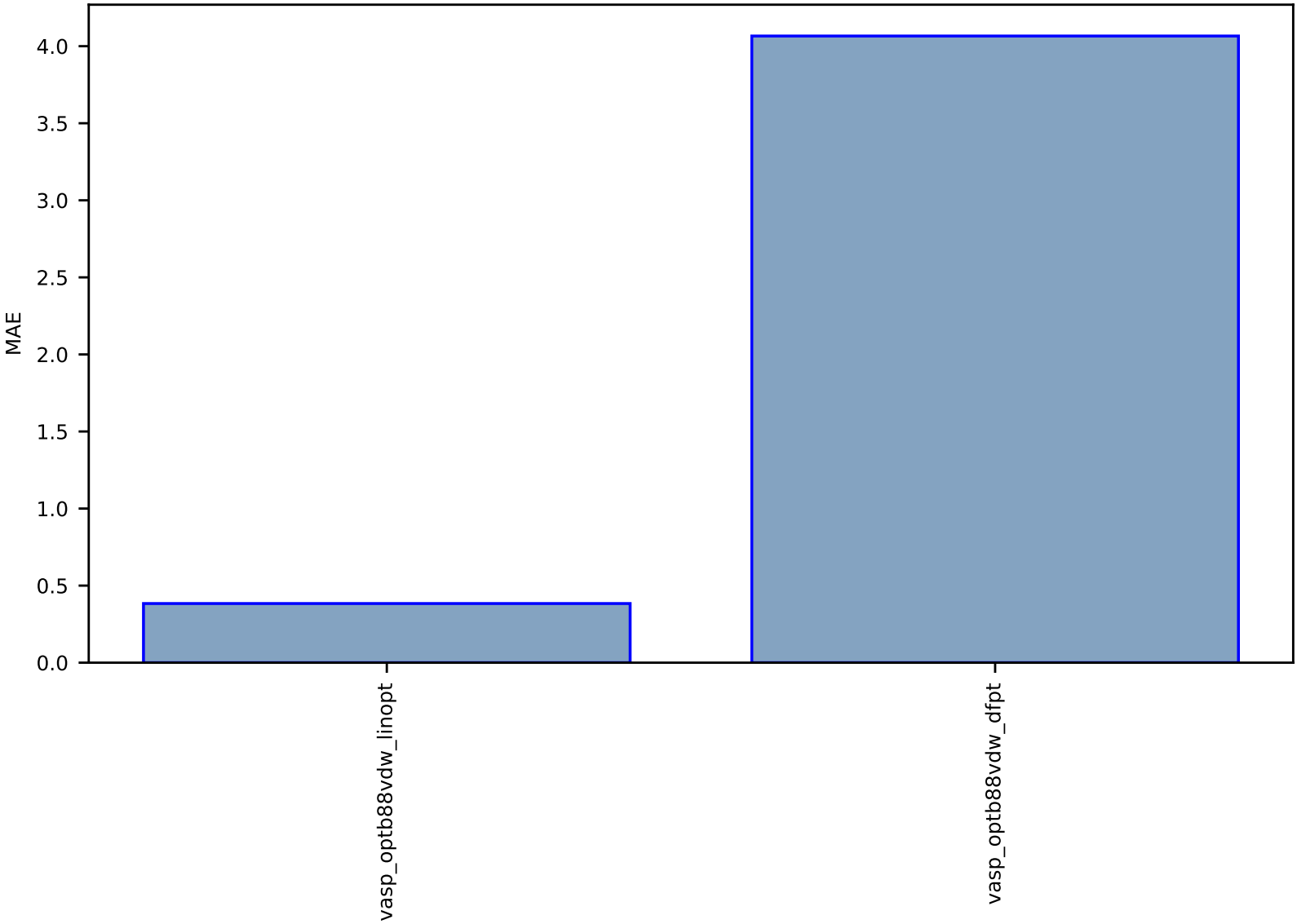


Fig S229:ES-SinglePropertyPrediction-epsx_JVASP_60_Te2Mo-dft_3d-test-mae.csv.zip

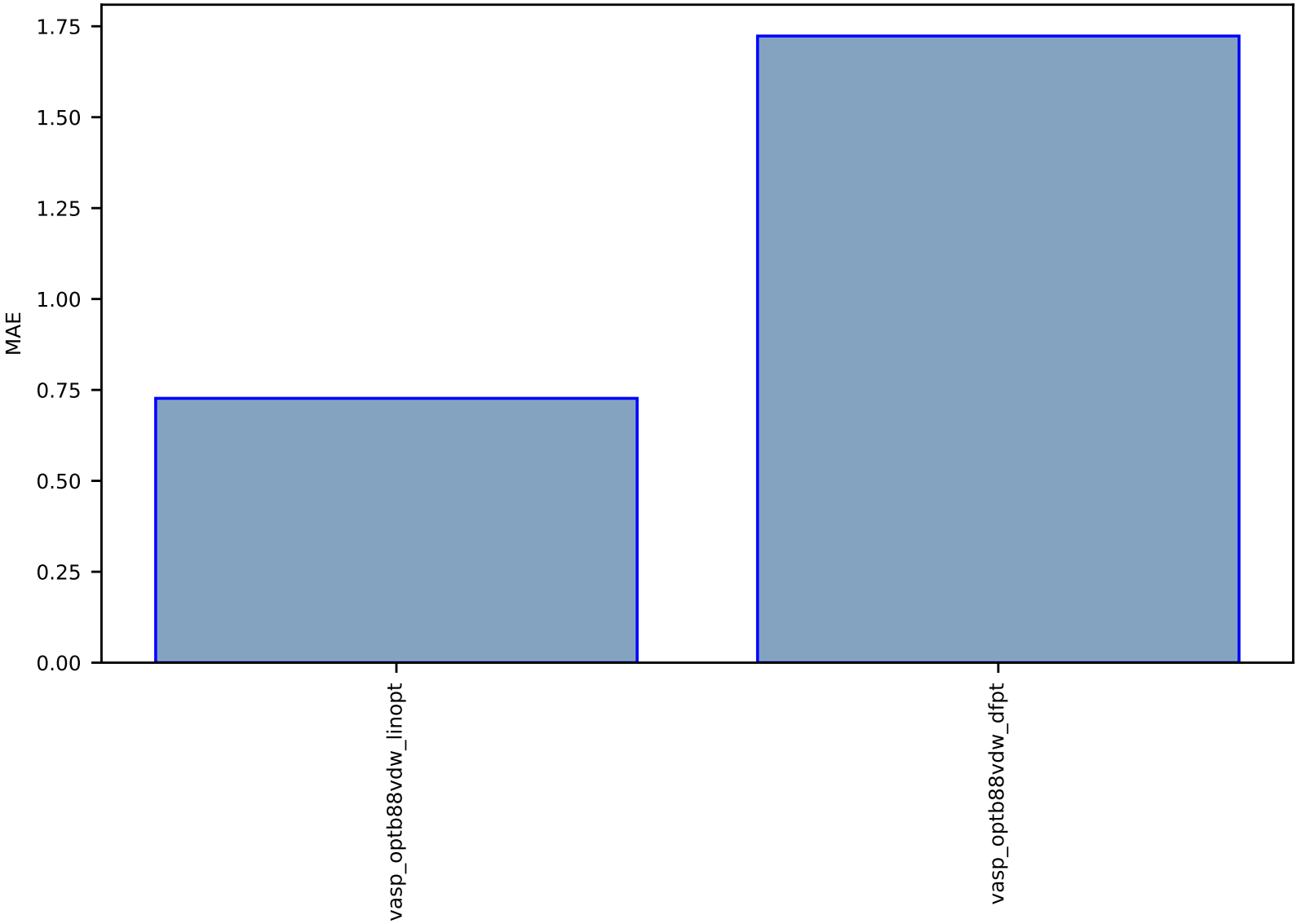


Fig S230:ES-SinglePropertyPrediction-epsx_JVASP_57_MoSe2-dft_3d-test-mae.csv.zip

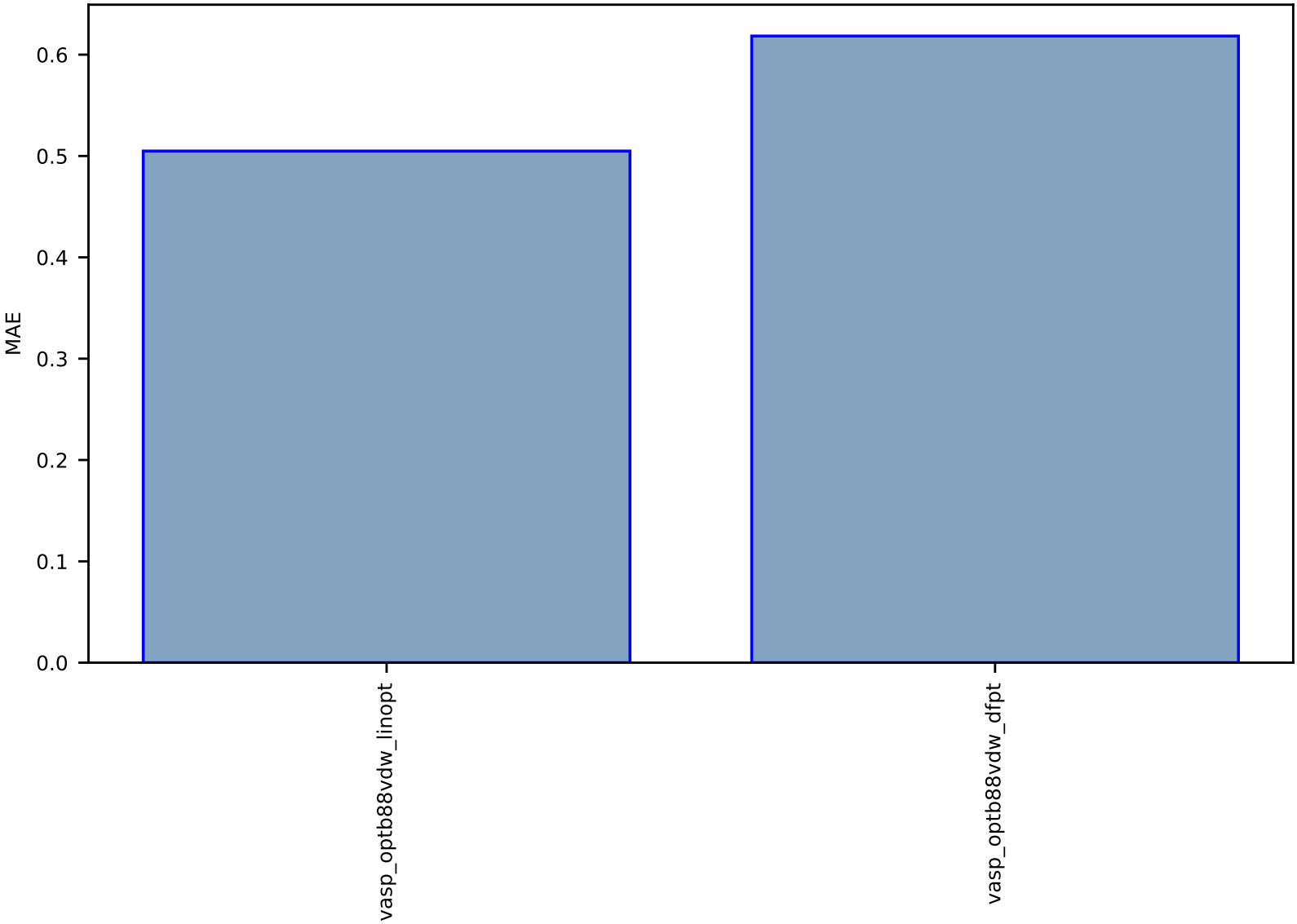


Fig S231:ES-SinglePropertyPrediction-epsx_JVASP_1327_AlP-dft_3d-test-mae.csv.zip

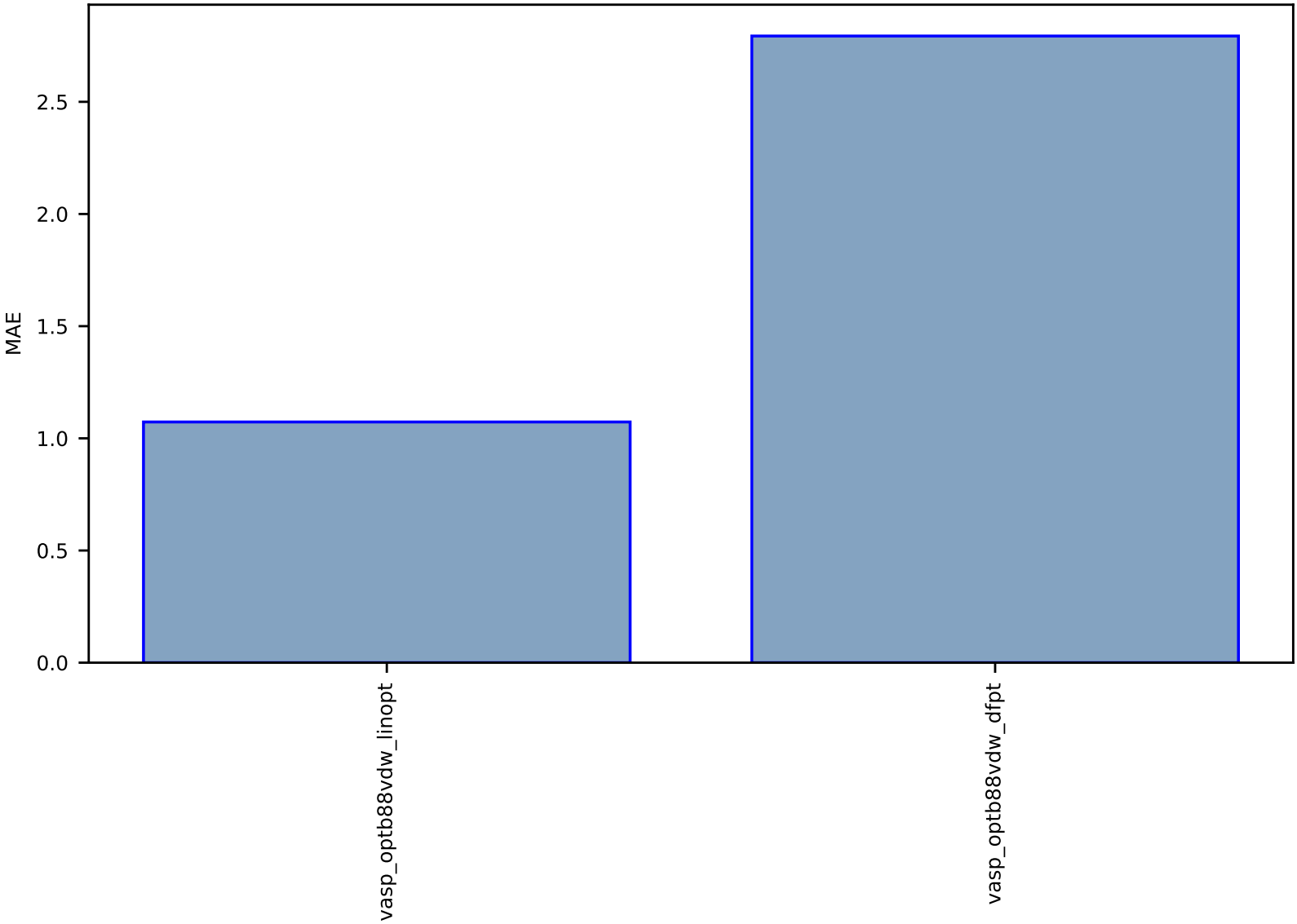


Fig S232:ES-SinglePropertyPrediction-epsx_JVASP_1312_BP-dft_3d-test-mae.csv.zip

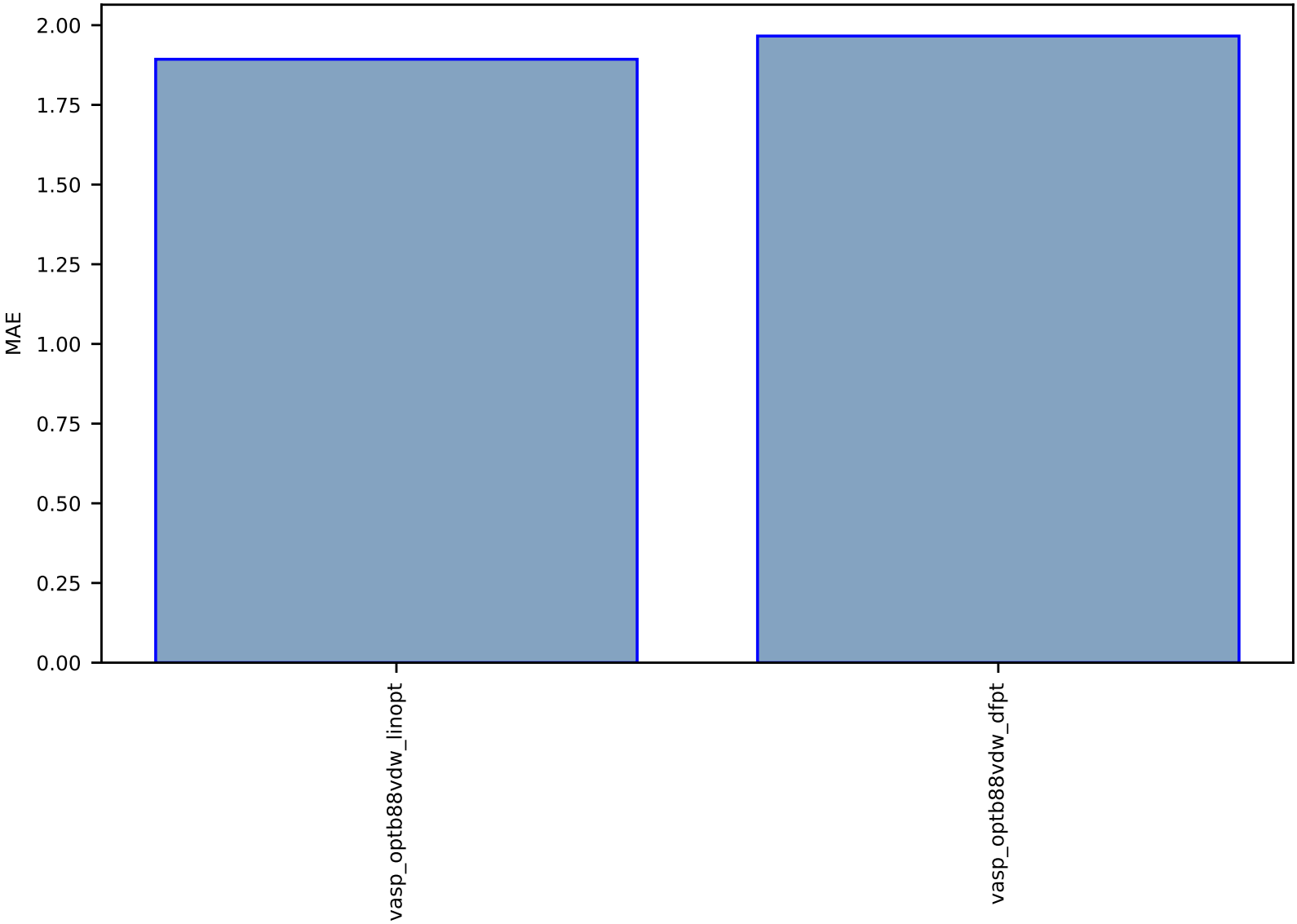


Fig S233:ES-SinglePropertyPrediction-epsx_JVASP_8041_HgTe-dft_3d-test-mae.csv.zip

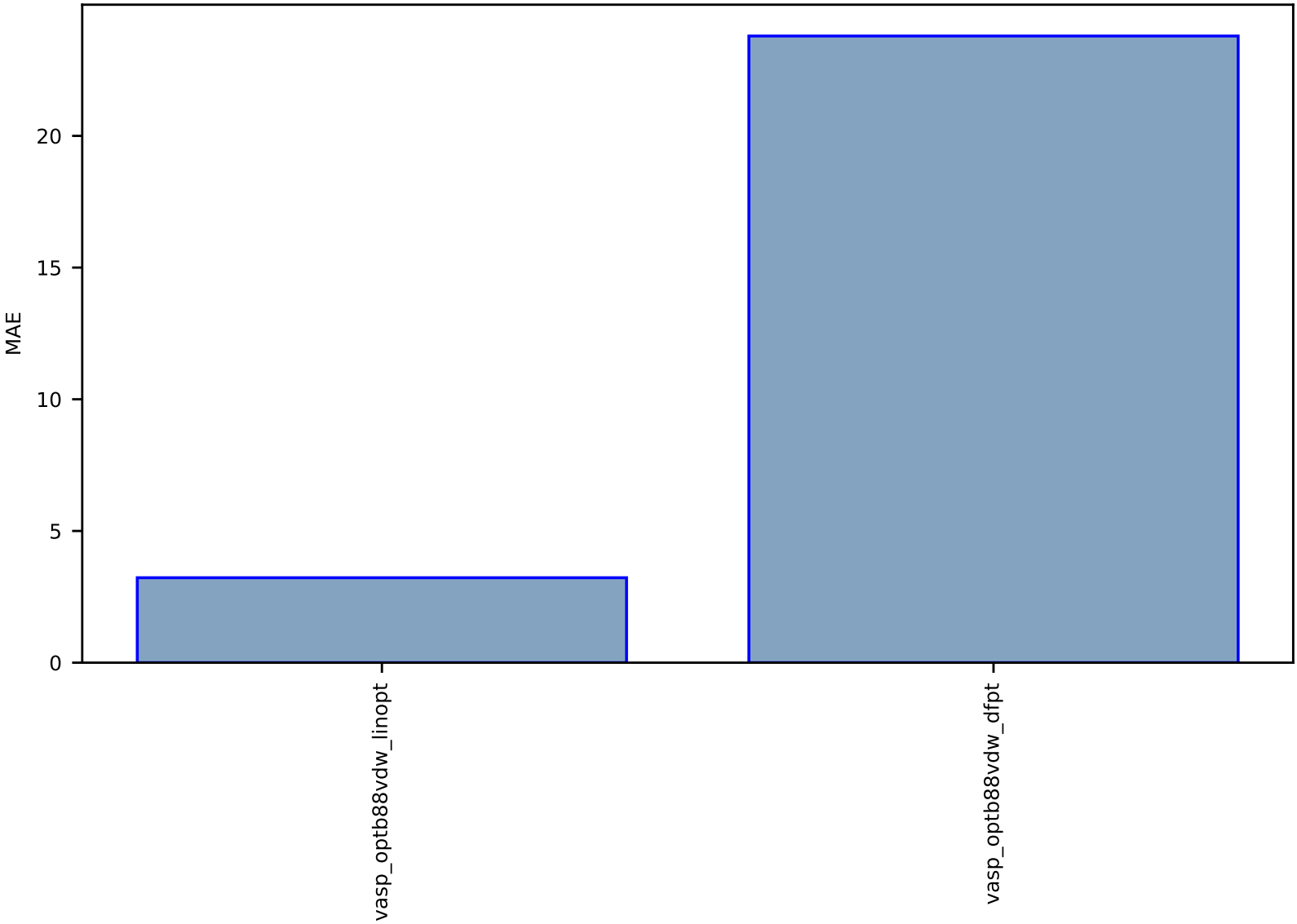


Fig S234:ES-SinglePropertyPrediction-epsx_JVASP_54_MoS2-dft_3d-test-mae.csv.zip

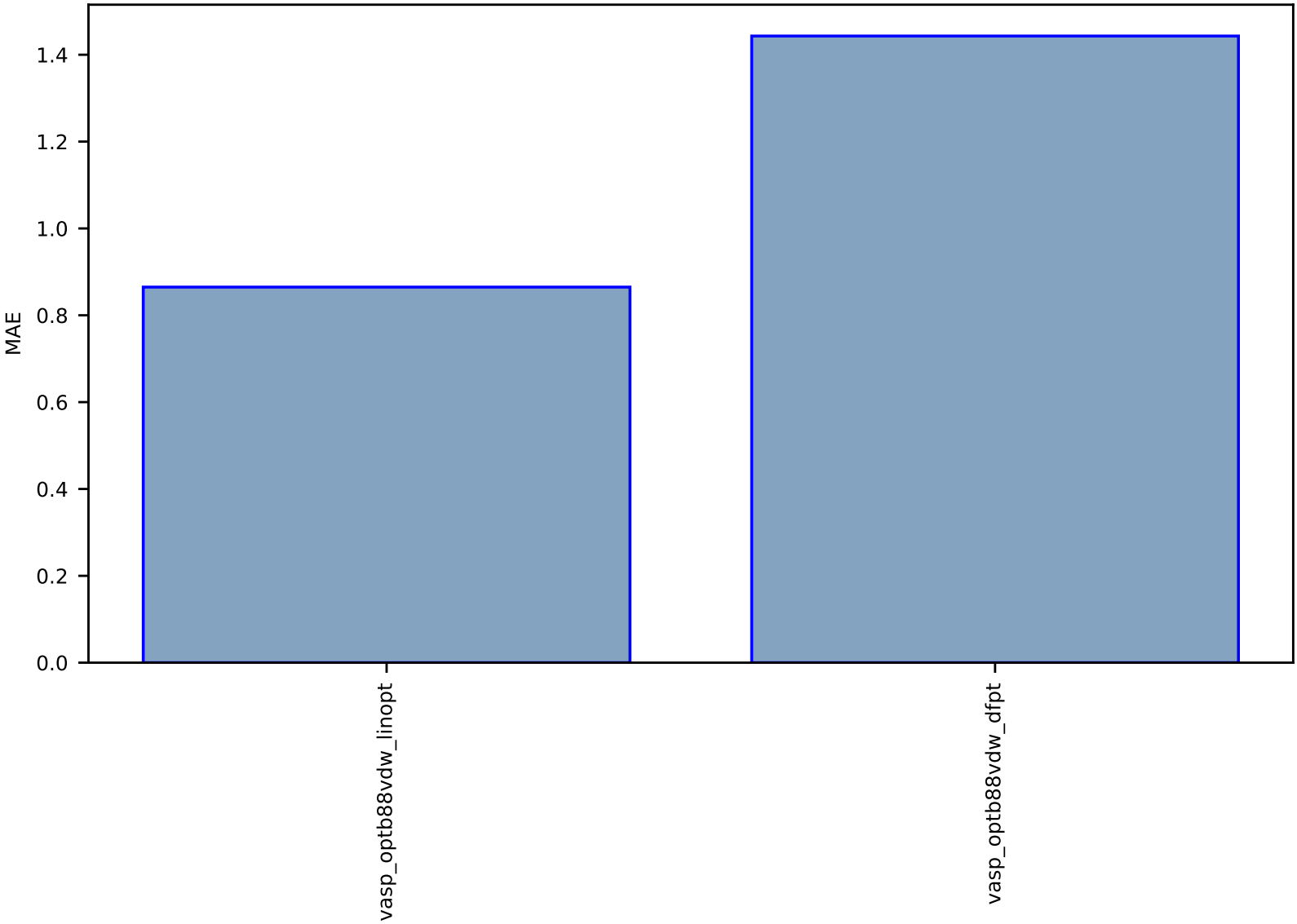


Fig S235:ES-SinglePropertyPrediction-epsx_JVASP_2355_ZnGeP2-dft_3d-test-mae.csv.zip

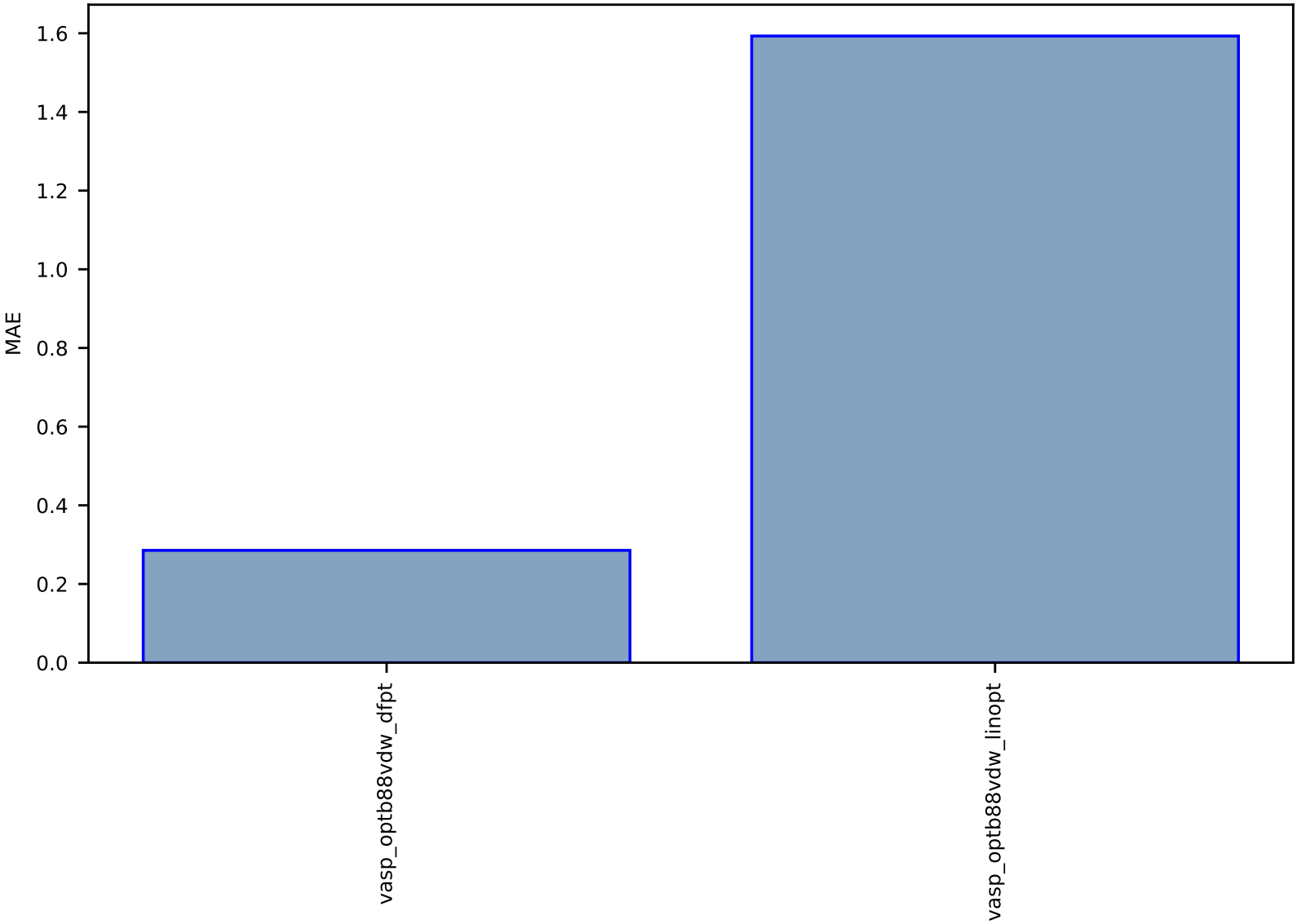


Fig S236:ES-SinglePropertyPrediction-epsx-dft_3d-test-mae.csv.zip

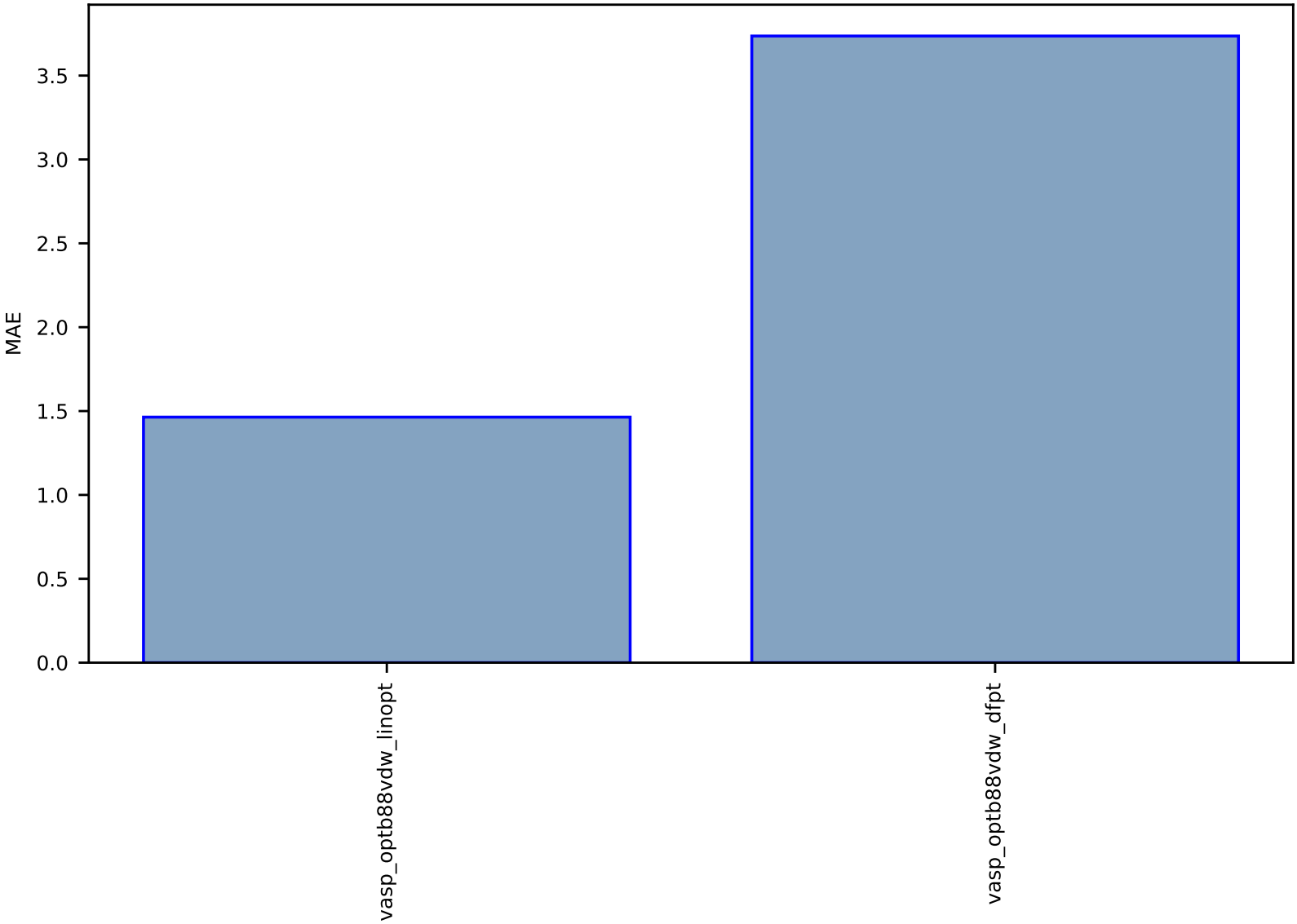


Fig S237:ES-SinglePropertyPrediction-epsx_JVASP_17_BN-dft_3d-test-mae.csv.zip

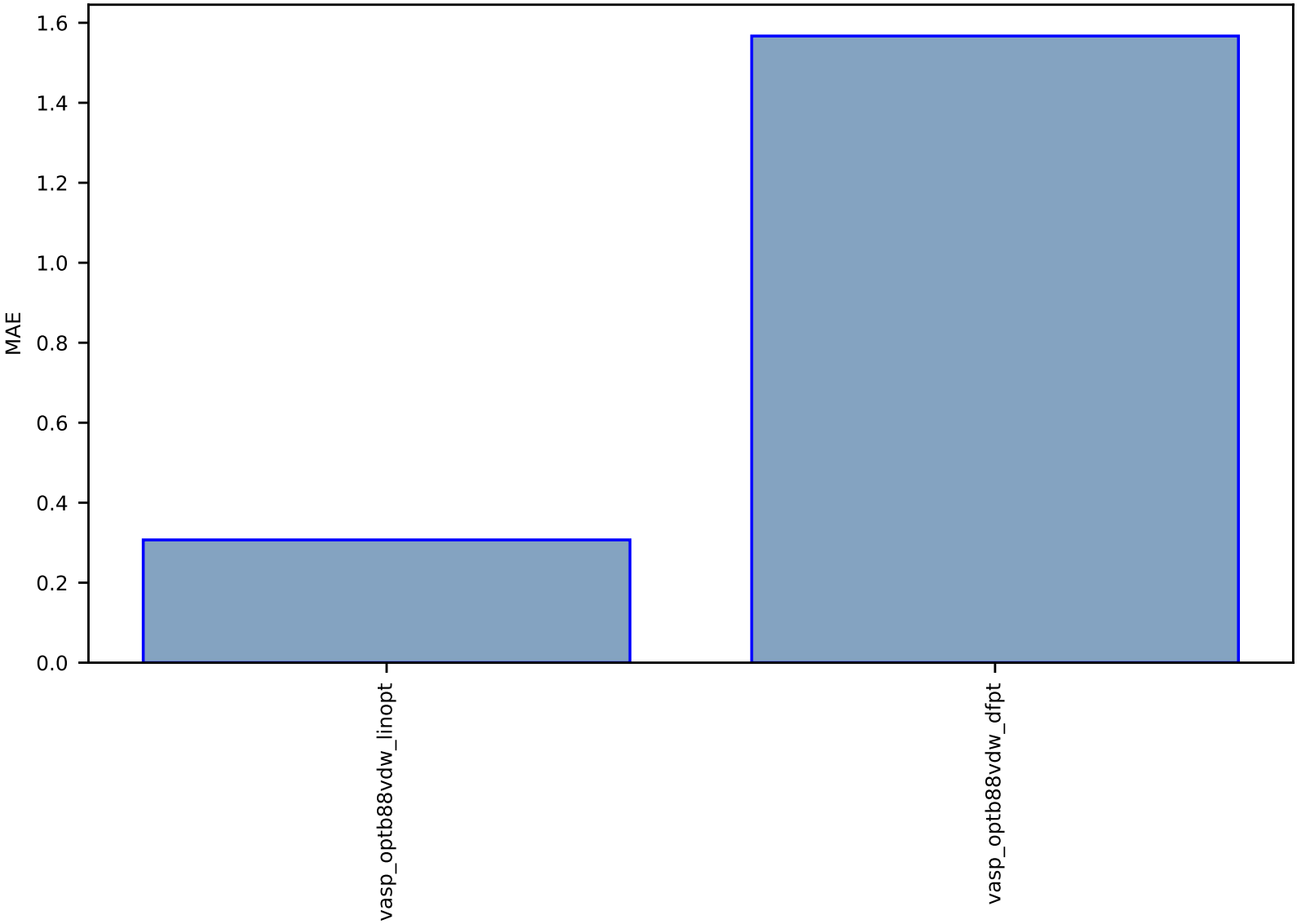


Fig S238:ES-SinglePropertyPrediction-epsx_JVASP_23_CdTe-dft_3d-test-mae.csv.zip

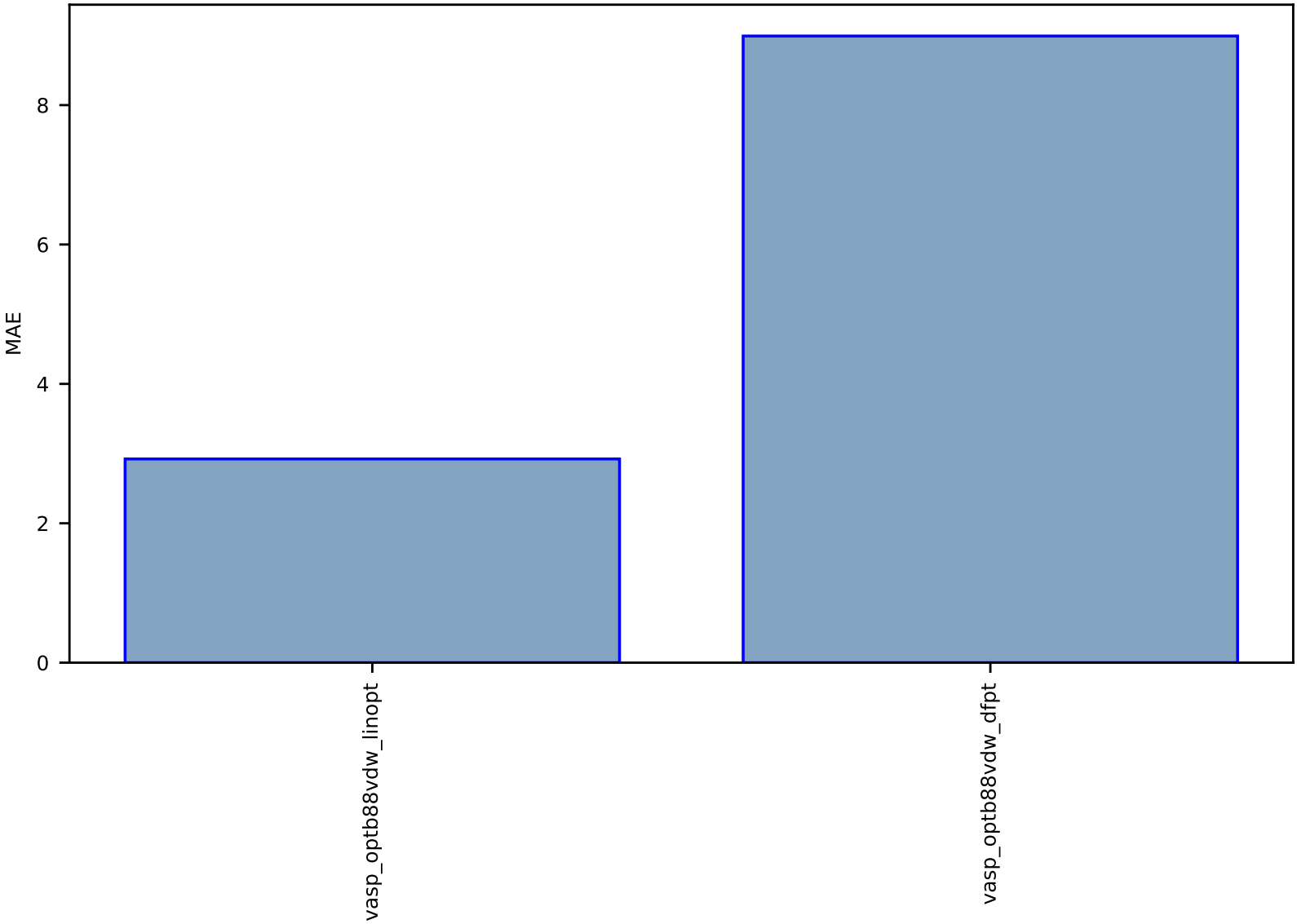


Fig S239:AI-SinglePropertyPrediction-final_energy-qe_tb-test-mae.csv.zip

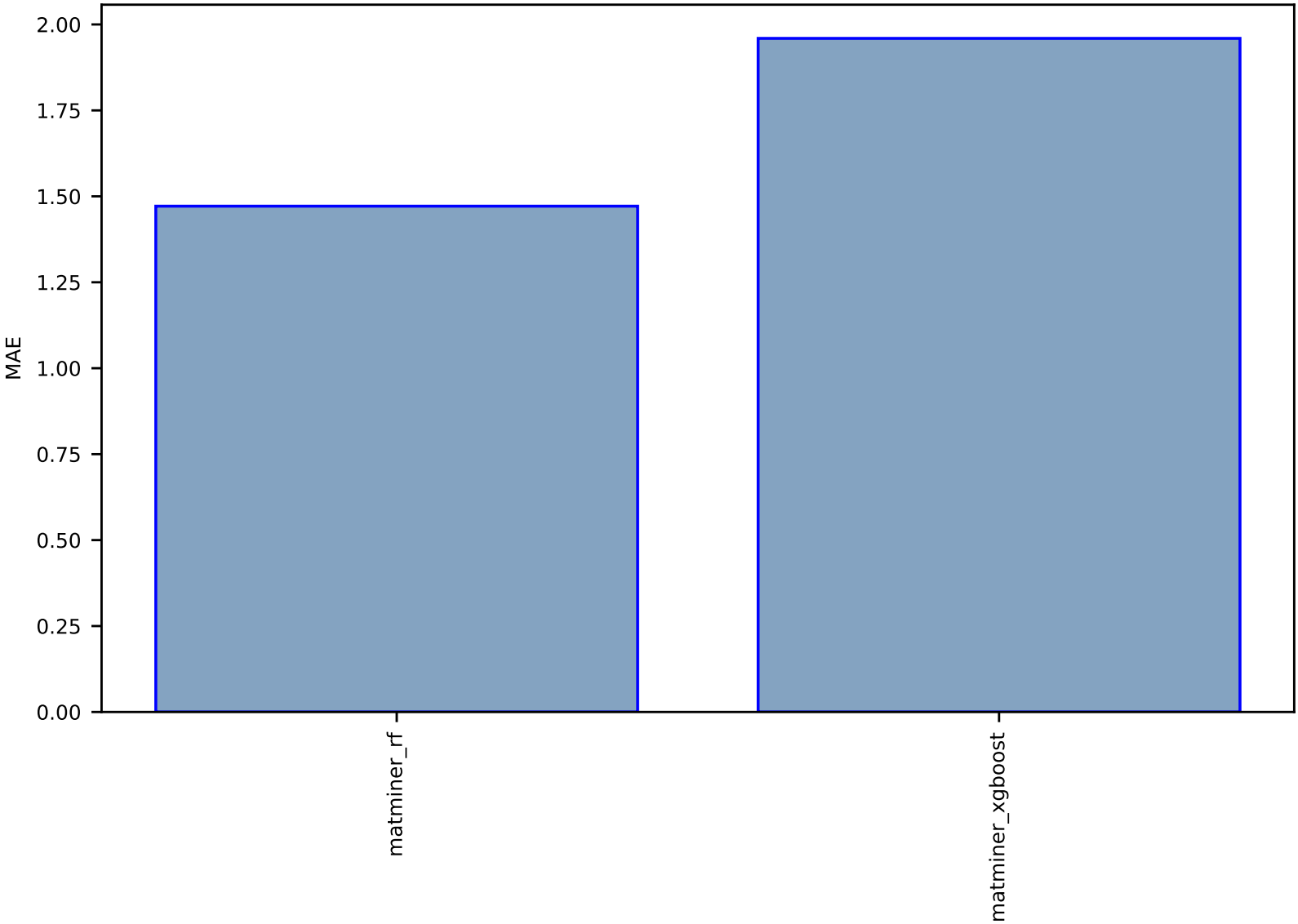


Fig S240:AI-TextSummary-text-arxiv_summary-test-rouge.csv.zip

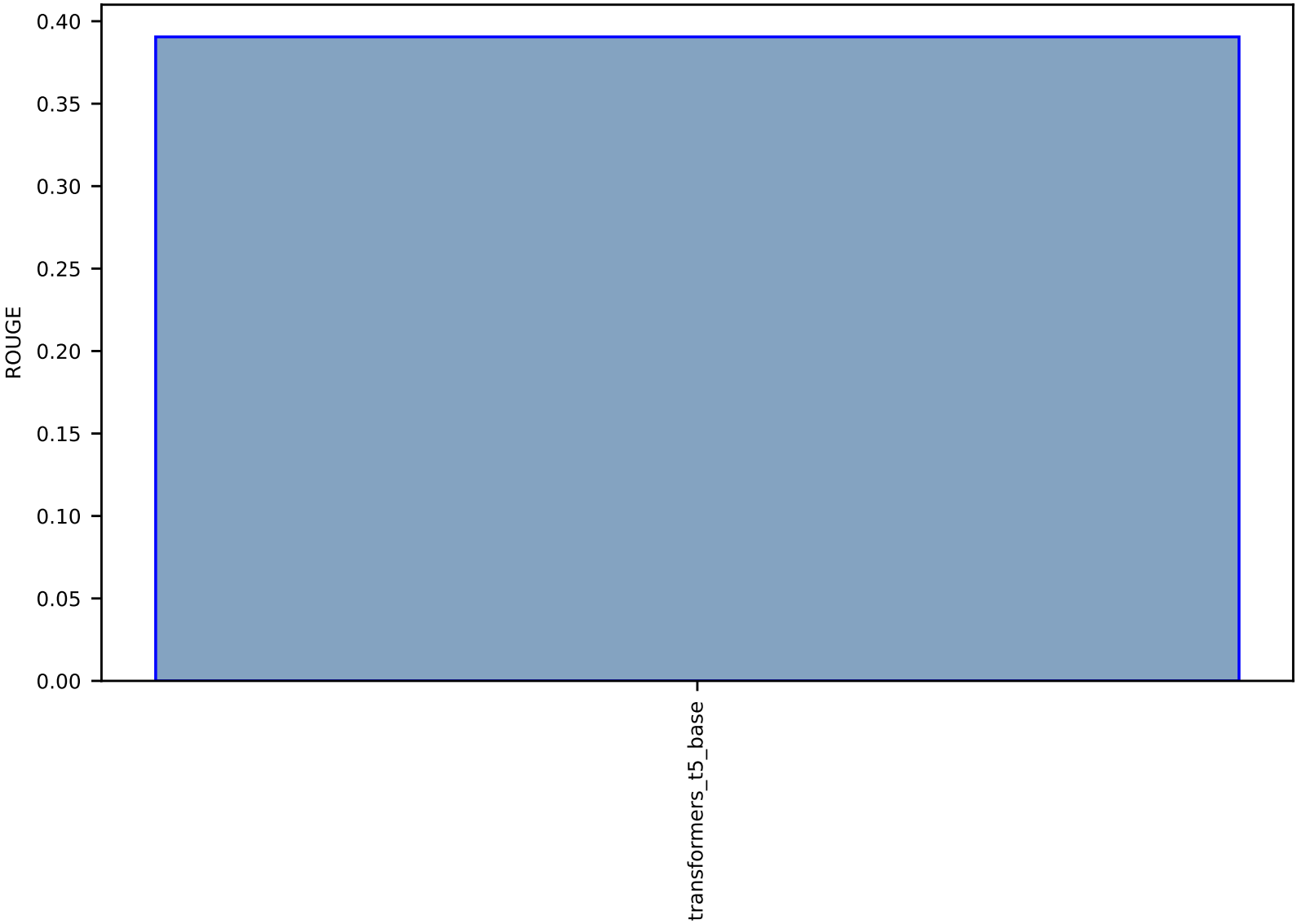


Fig S241:EXP-Spectra-XRD_JVASP_19821_MgB2-dft_3d-test-multimae.csv.zip

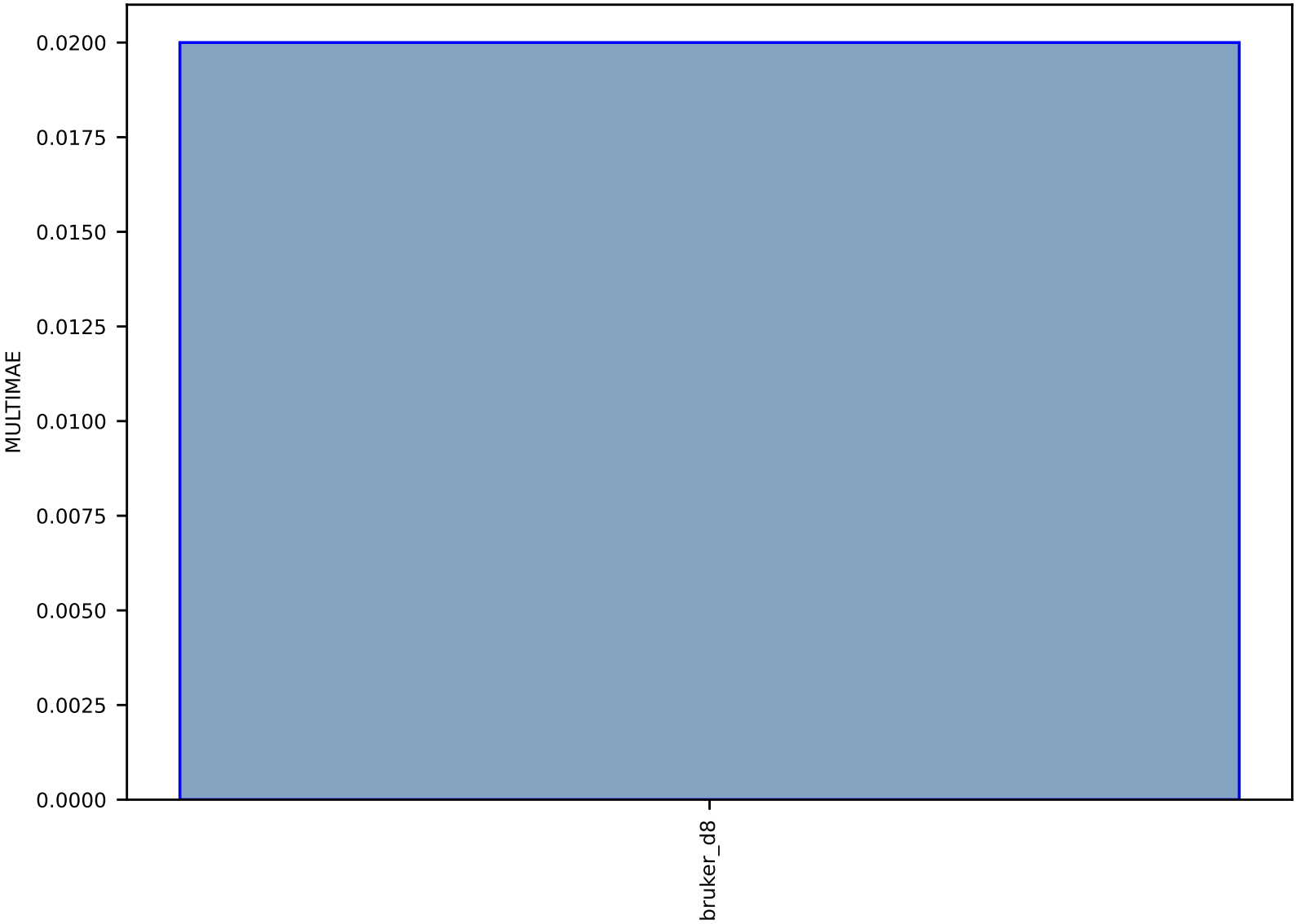


Fig S242:ES-SinglePropertyPrediction-bulk_modulus_JVASP_25213_Kr-dft_3d-test-mae.csv.zip

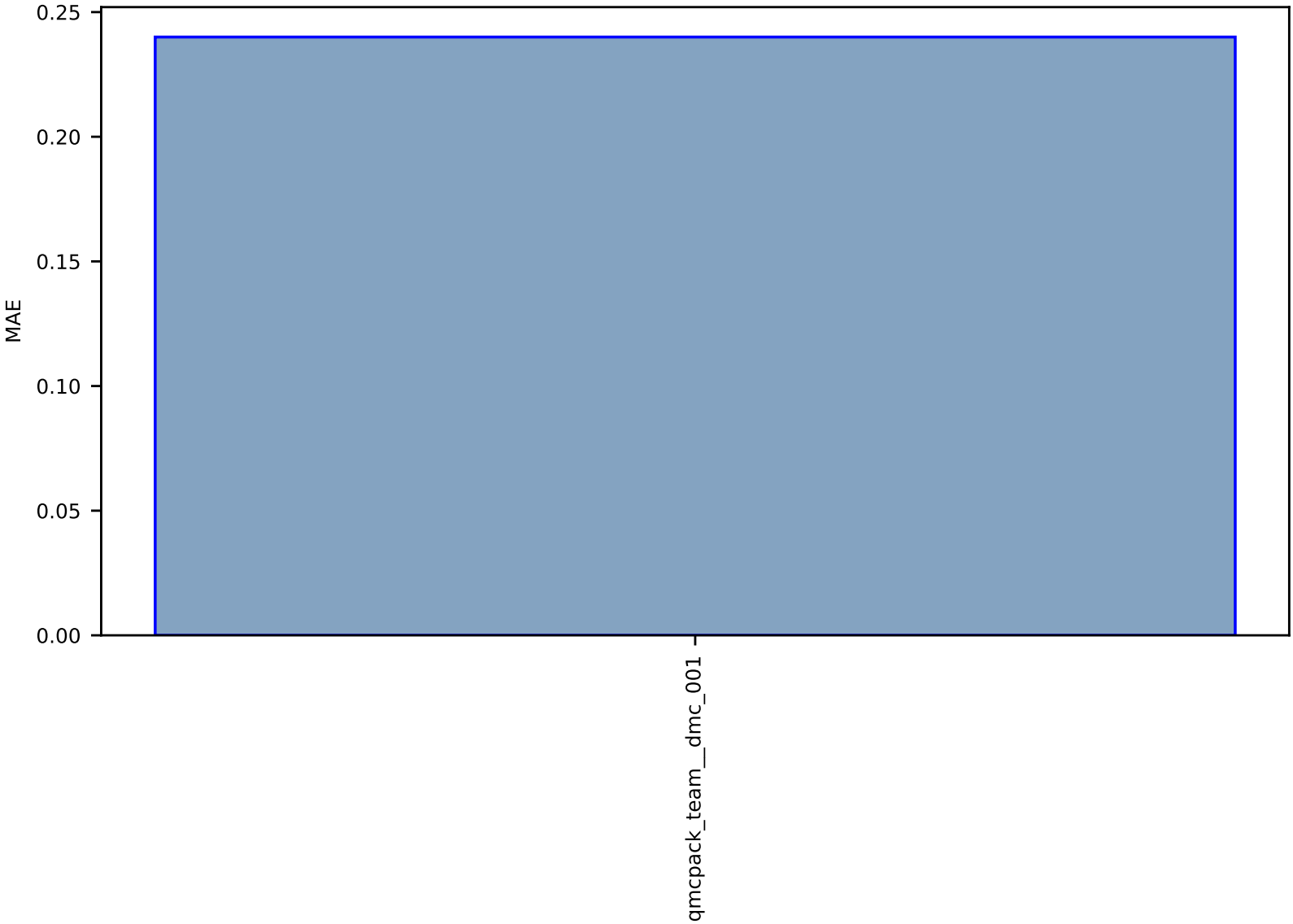


Fig S243:ES-SinglePropertyPrediction-bulk_modulus_JVASP_79204_BN-dft_3d-test-mae.csv.zip

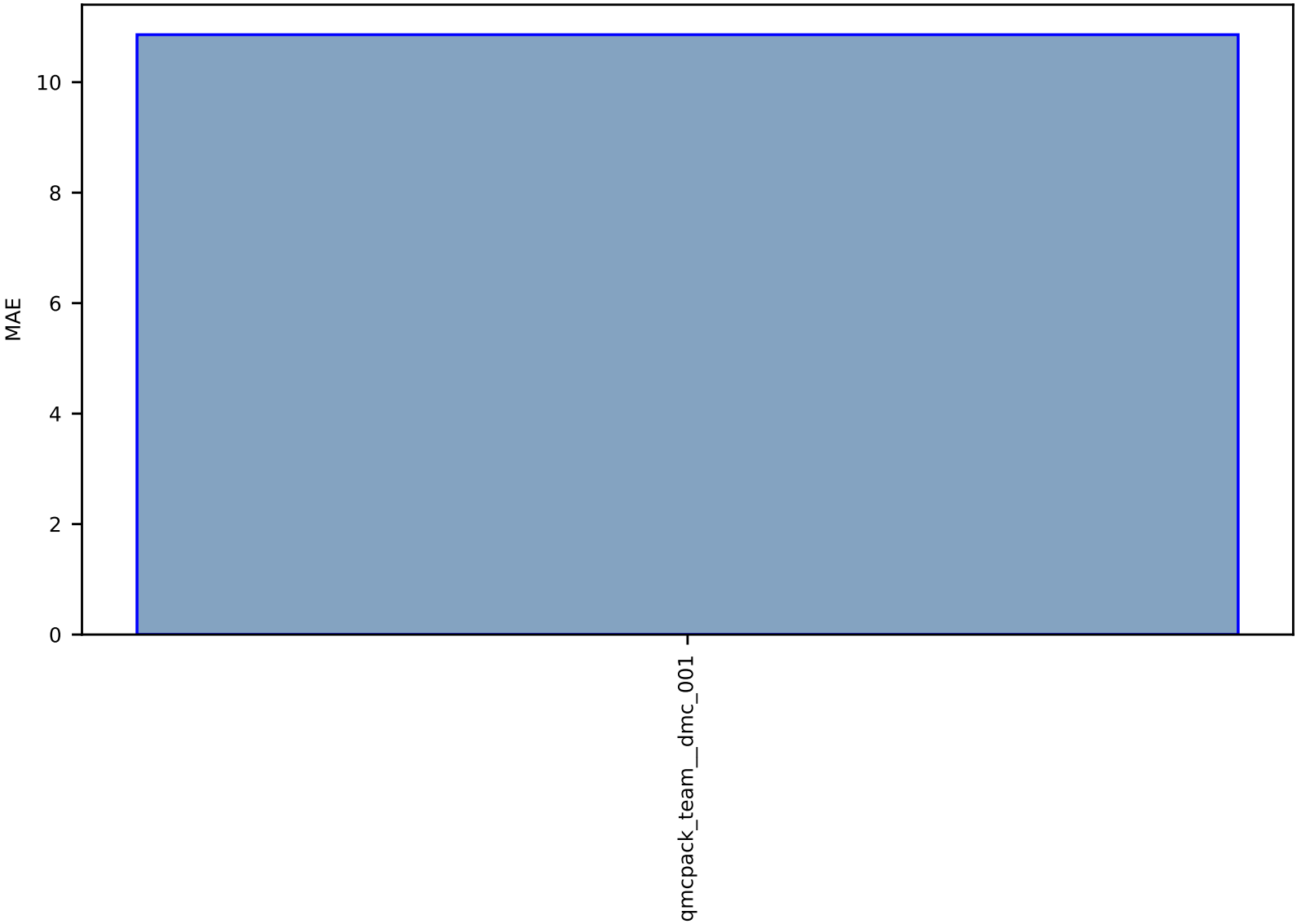


Fig S244:ES-SinglePropertyPrediction-bulk_modulus_JVASP_834_Be-dft_3d-test-mae.csv.zip

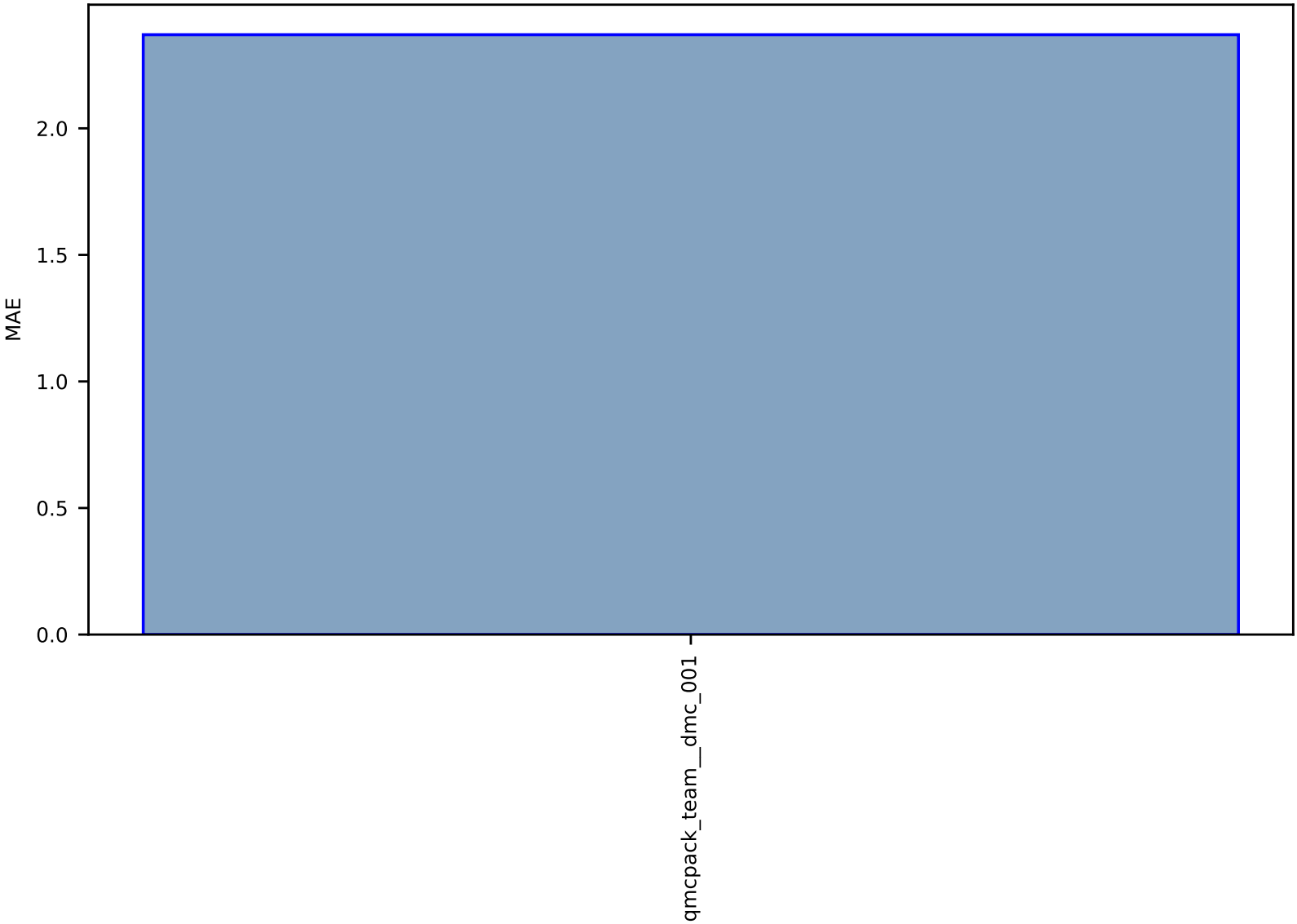


Fig S245:ES-SinglePropertyPrediction-bulk_modulus_JVASP_819_Ar-dft_3d-test-mae.csv.zip

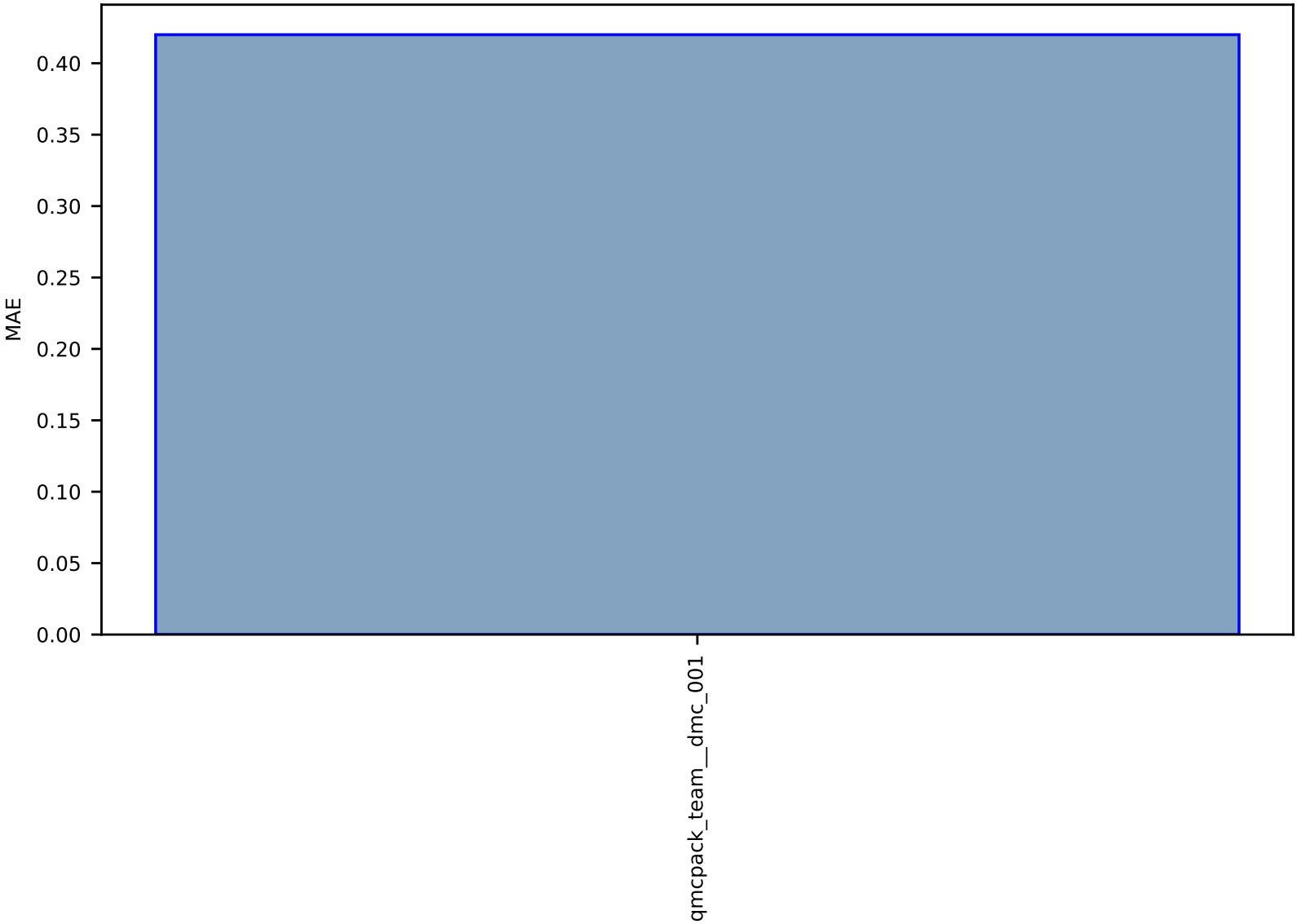


Fig S246:ES-SinglePropertyPrediction-bulk_modulus_JVASP_1312_BP-dft_3d-test-mae.csv.zip

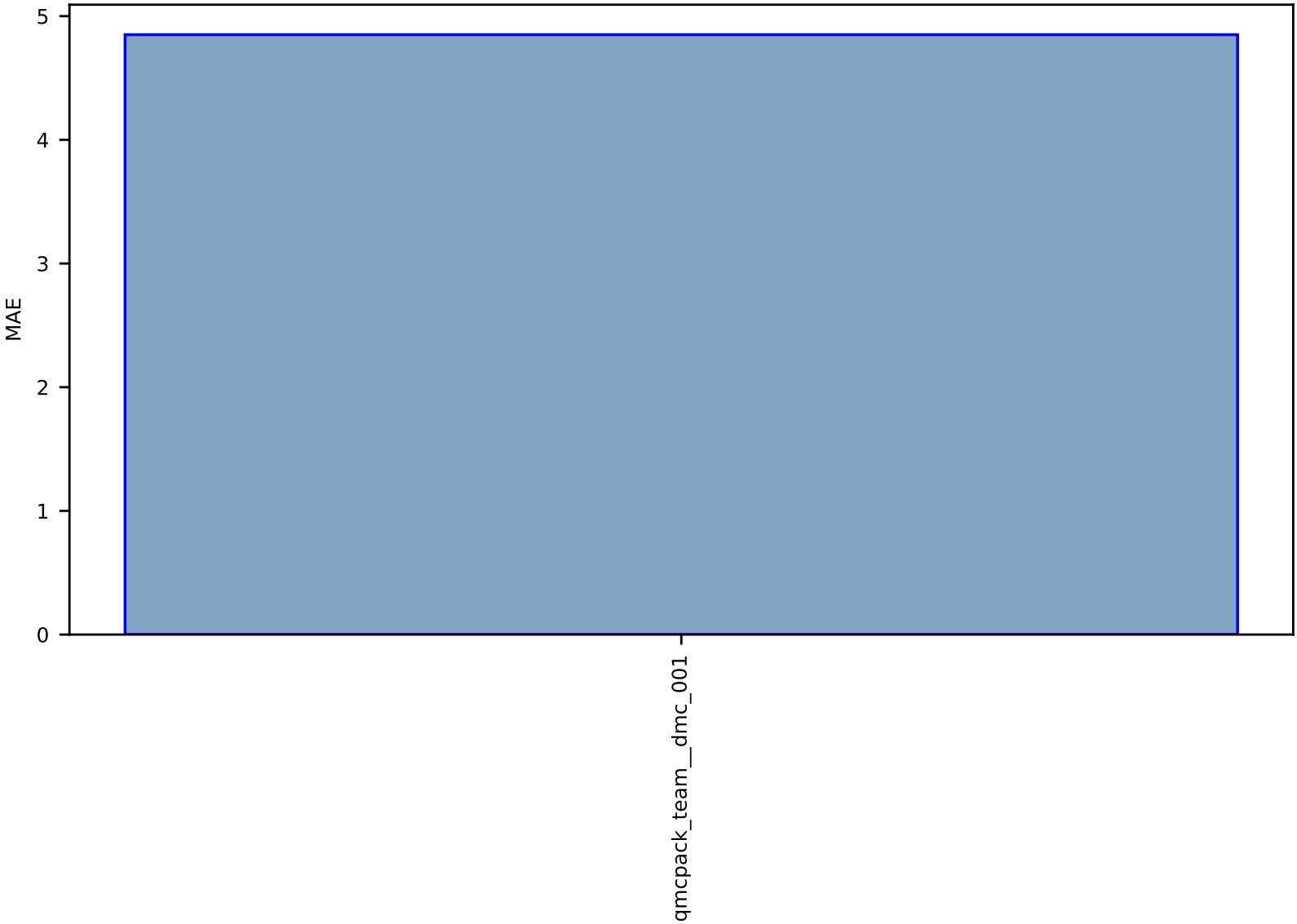


Fig S247:ES-SinglePropertyPrediction-bulk_modulus_JVASP_25248_Xe-dft_3d-test-mae.csv.zip

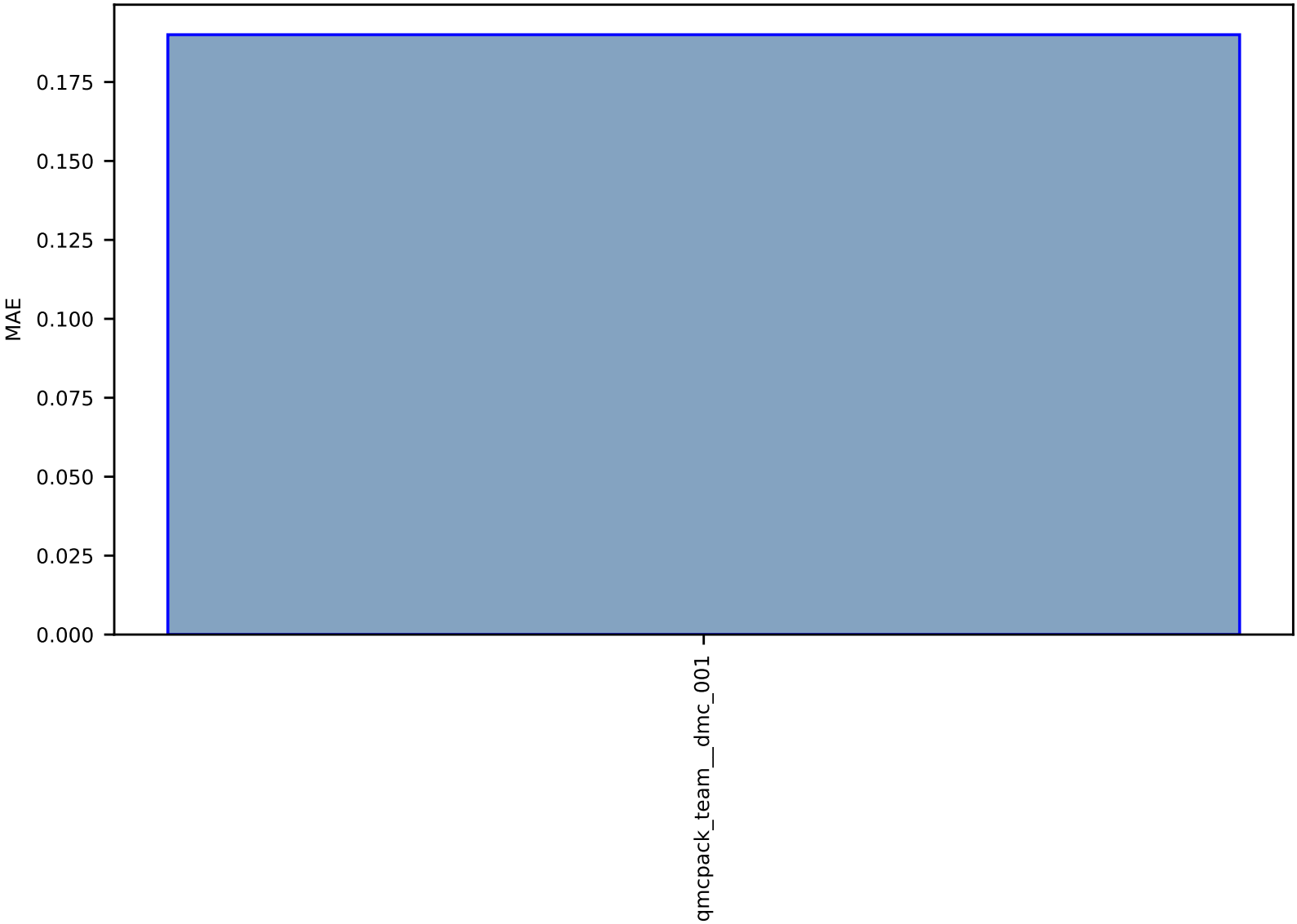


Fig S248:ES-SinglePropertyPrediction-ef-vacancydb-test-mae.csv.zip

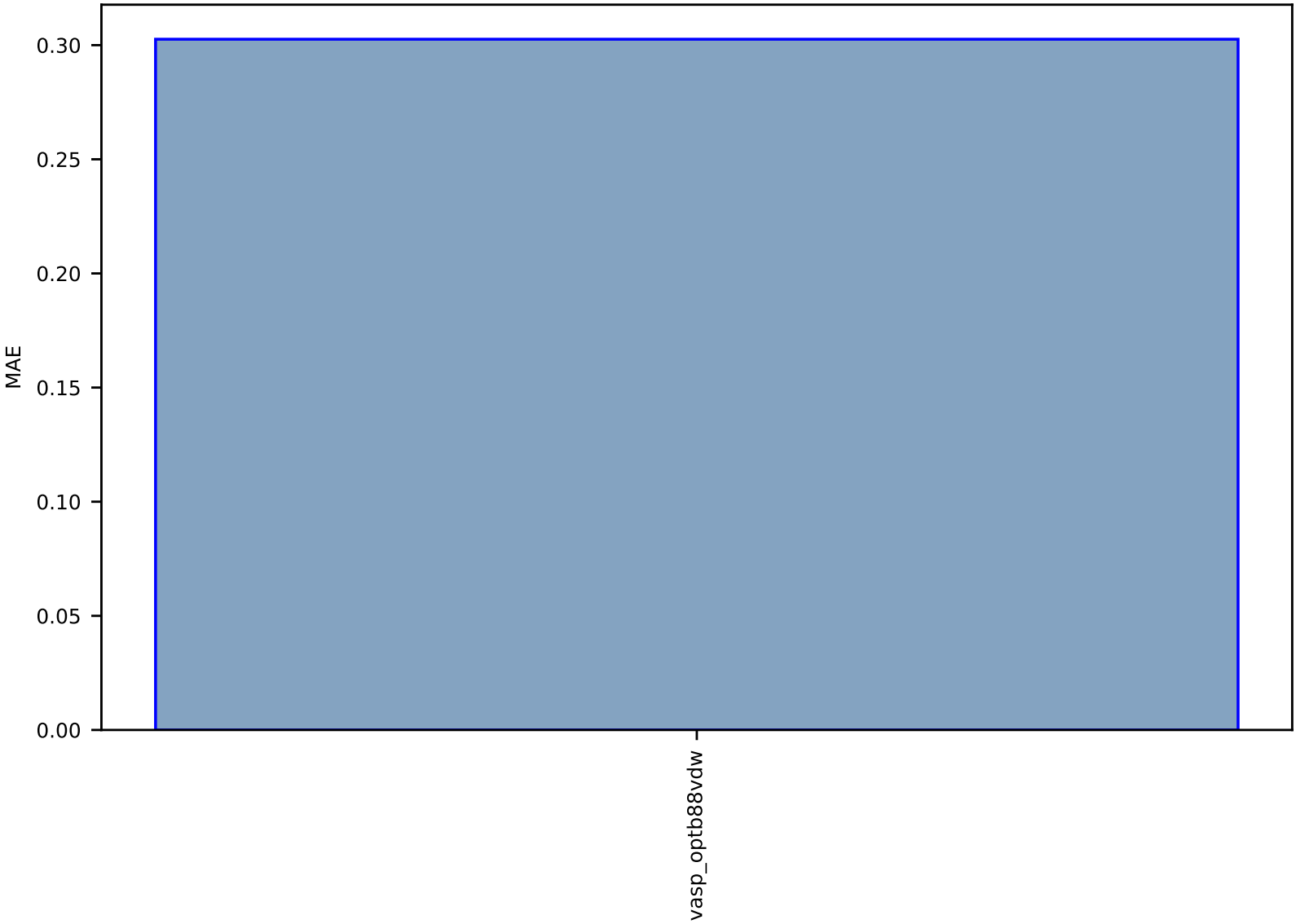


Fig S249:ES-SinglePropertyPrediction-max_piezoelec_JVASP_3450_TiPbO3-dft_3d-test-mae.csv.zip

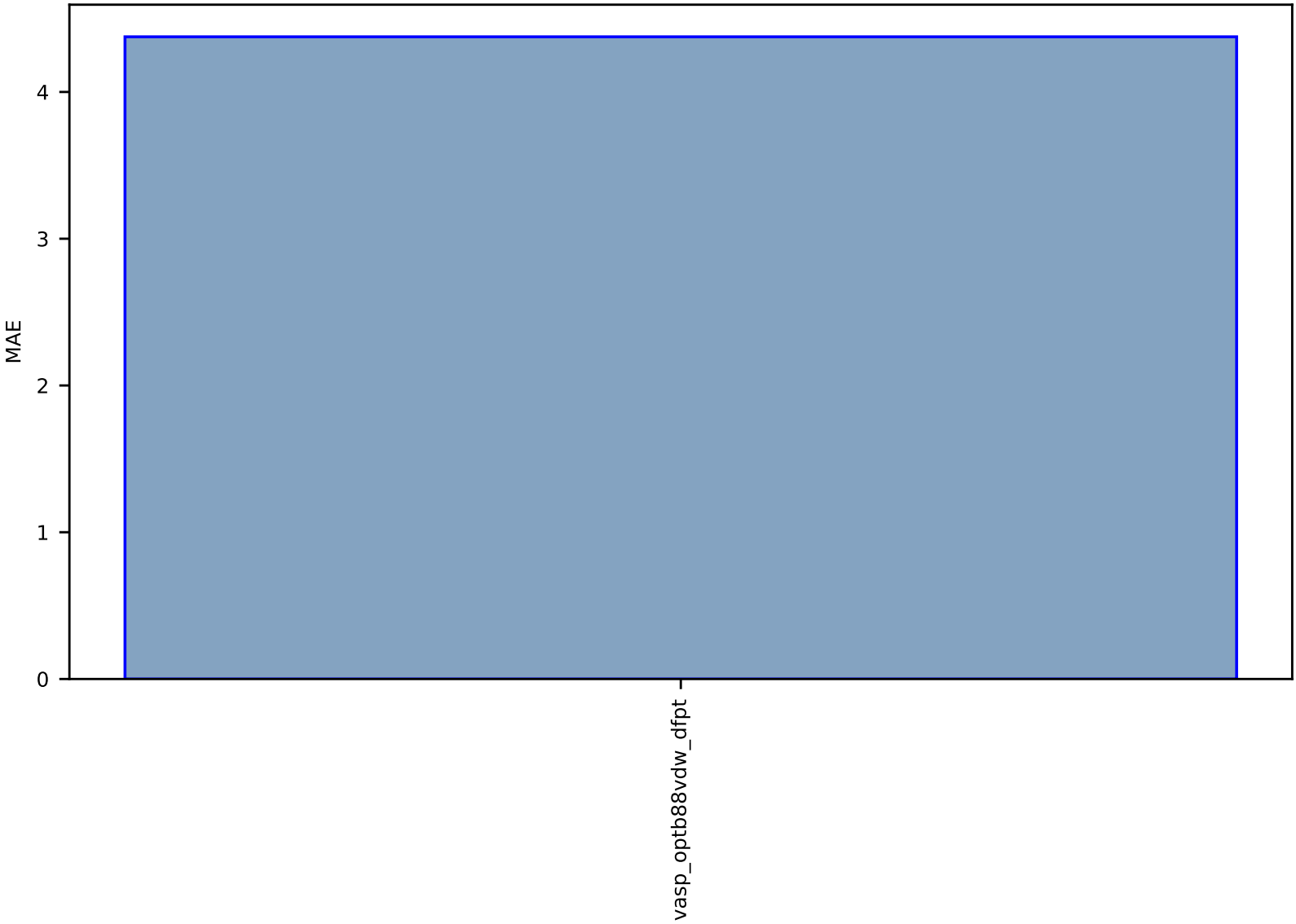


Fig S250:ES-SinglePropertyPrediction-max_piezoelec_JVASP_57695_BN-dft_3d-test-mae.csv.zip

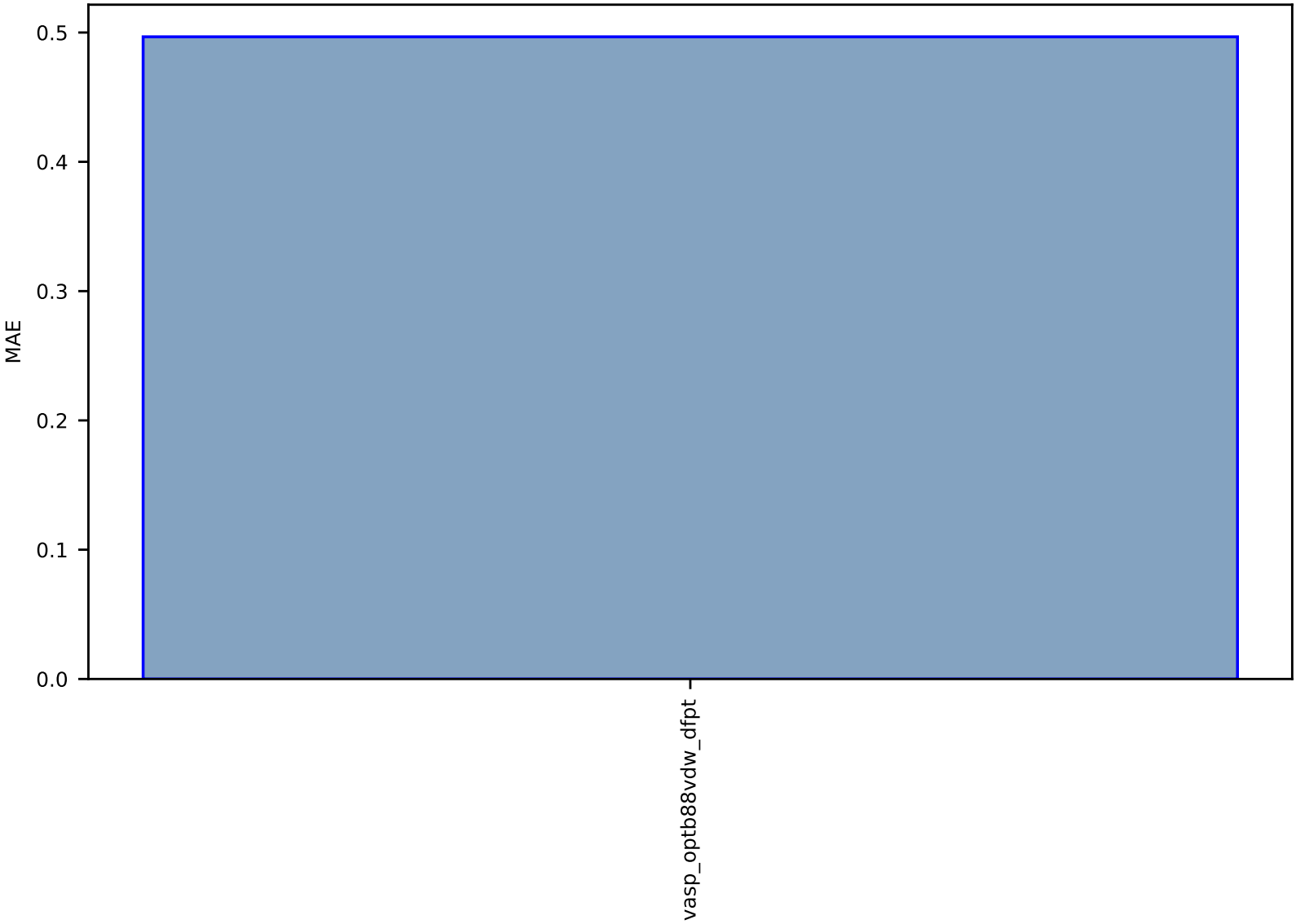


Fig S251:ES-SinglePropertyPrediction-max_piezoelec_JVASP_8047_ZnSe-dft_3d-test-mae.csv.zip

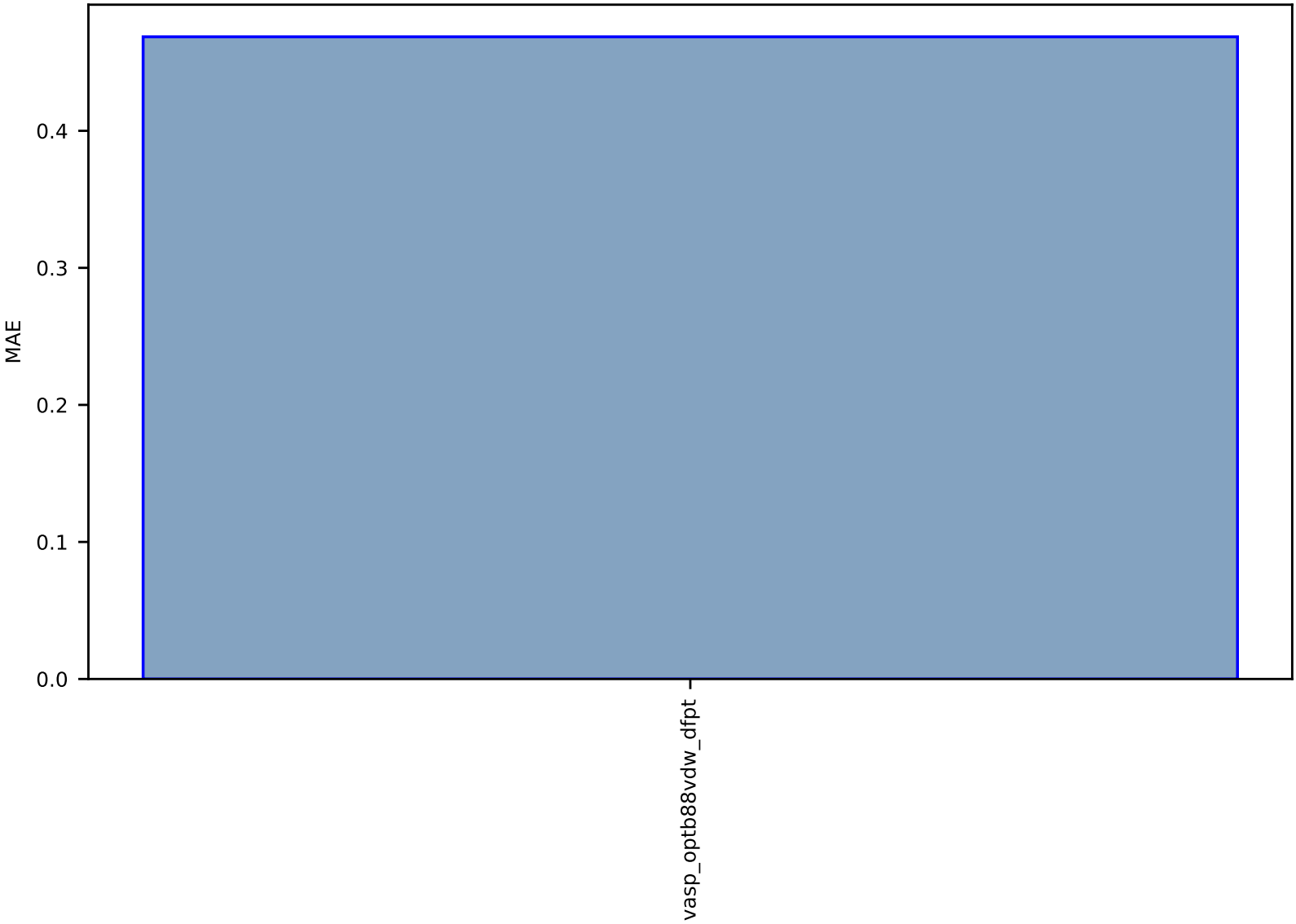


Fig S252:ES-SinglePropertyPrediction-max_piezoelec_JVASP_7648_ZnS-dft_3d-test-mae.csv.zip

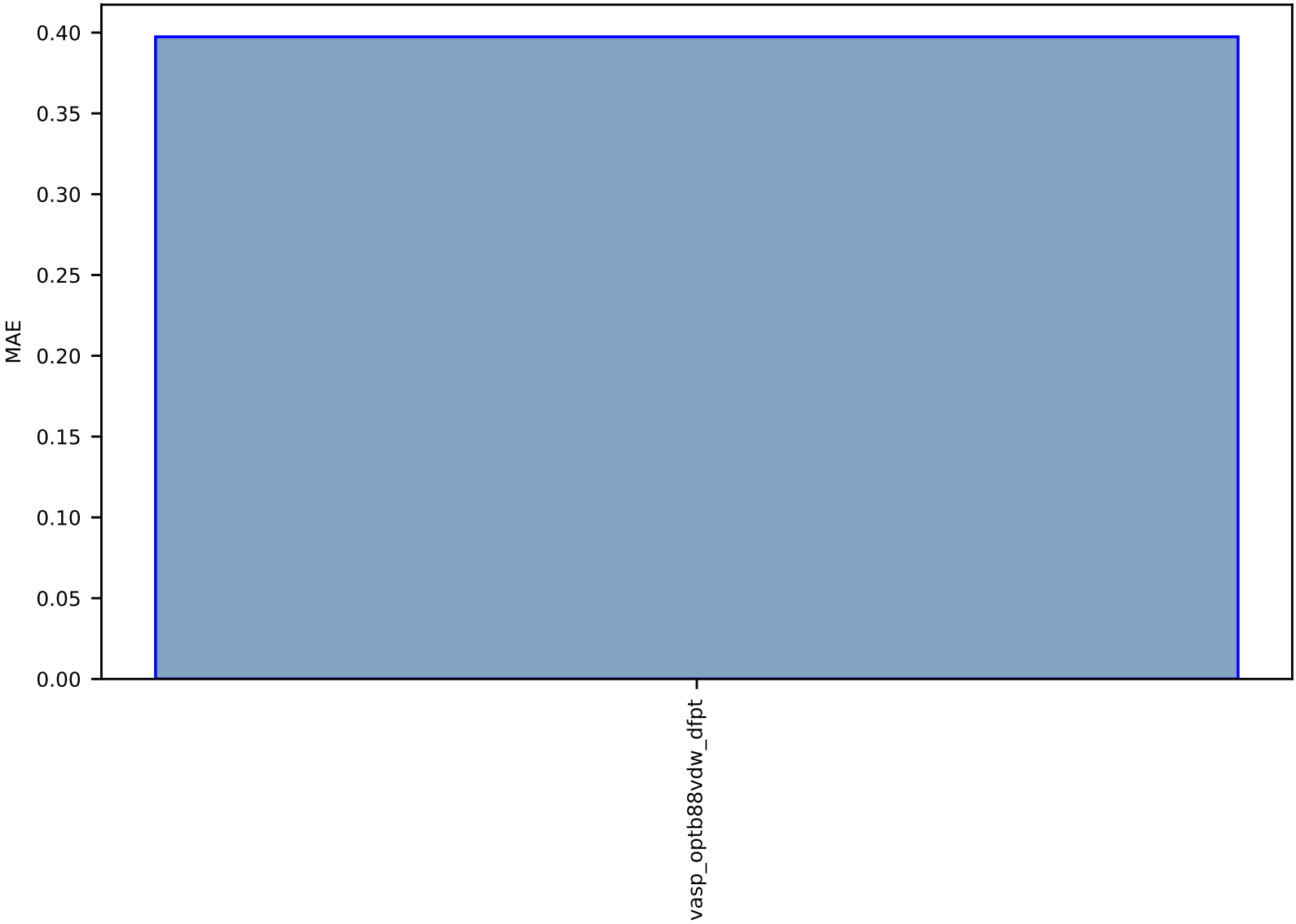


Fig S253:ES-SinglePropertyPrediction-max_piezoelec_JVASP_1408_AlSb-dft_3d-test-mae.csv.zip

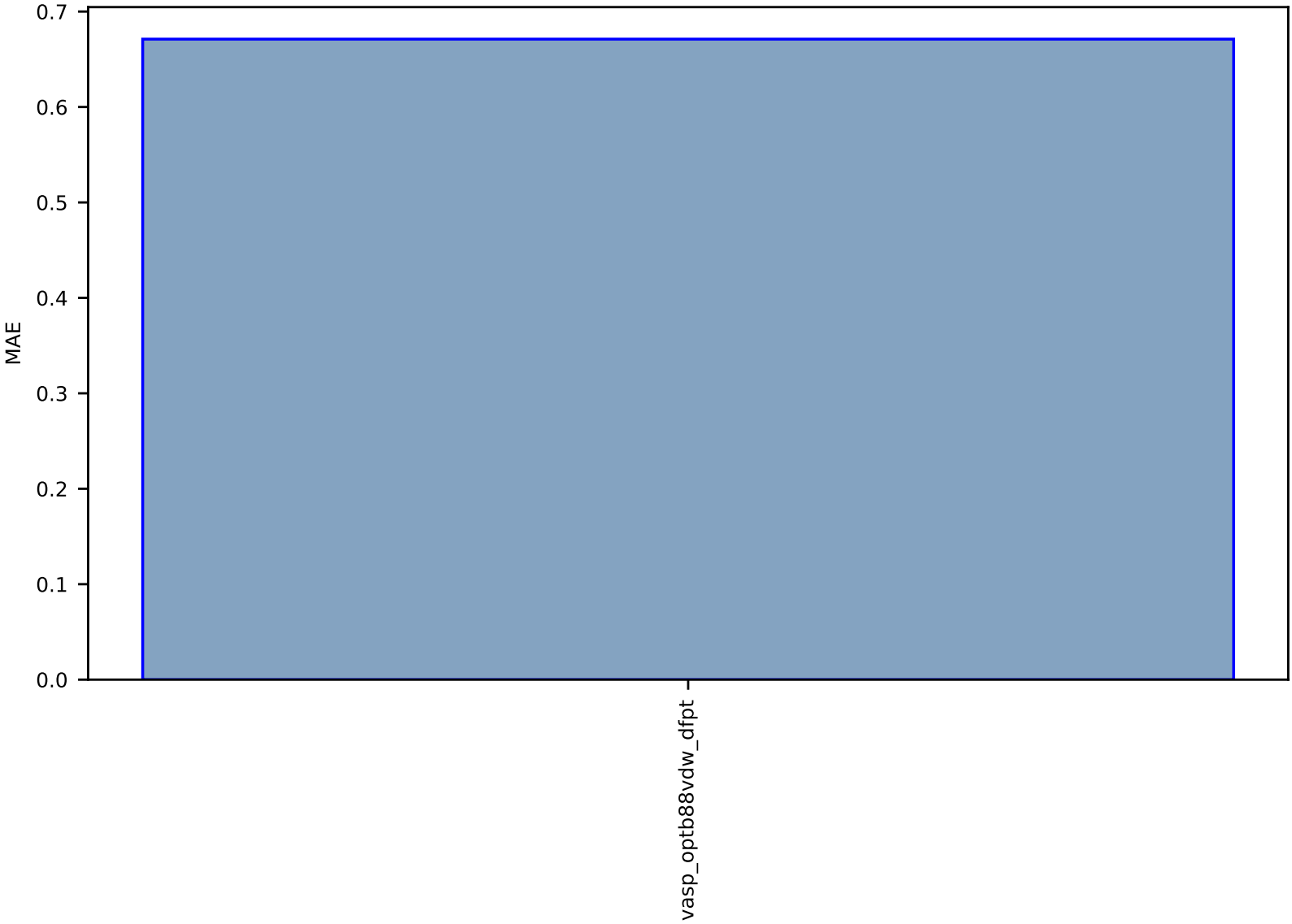


Fig S254:ES-SinglePropertyPrediction-max_piezoelec_JVASP_20778_BeO-dft_3d-test-mae.csv.zip

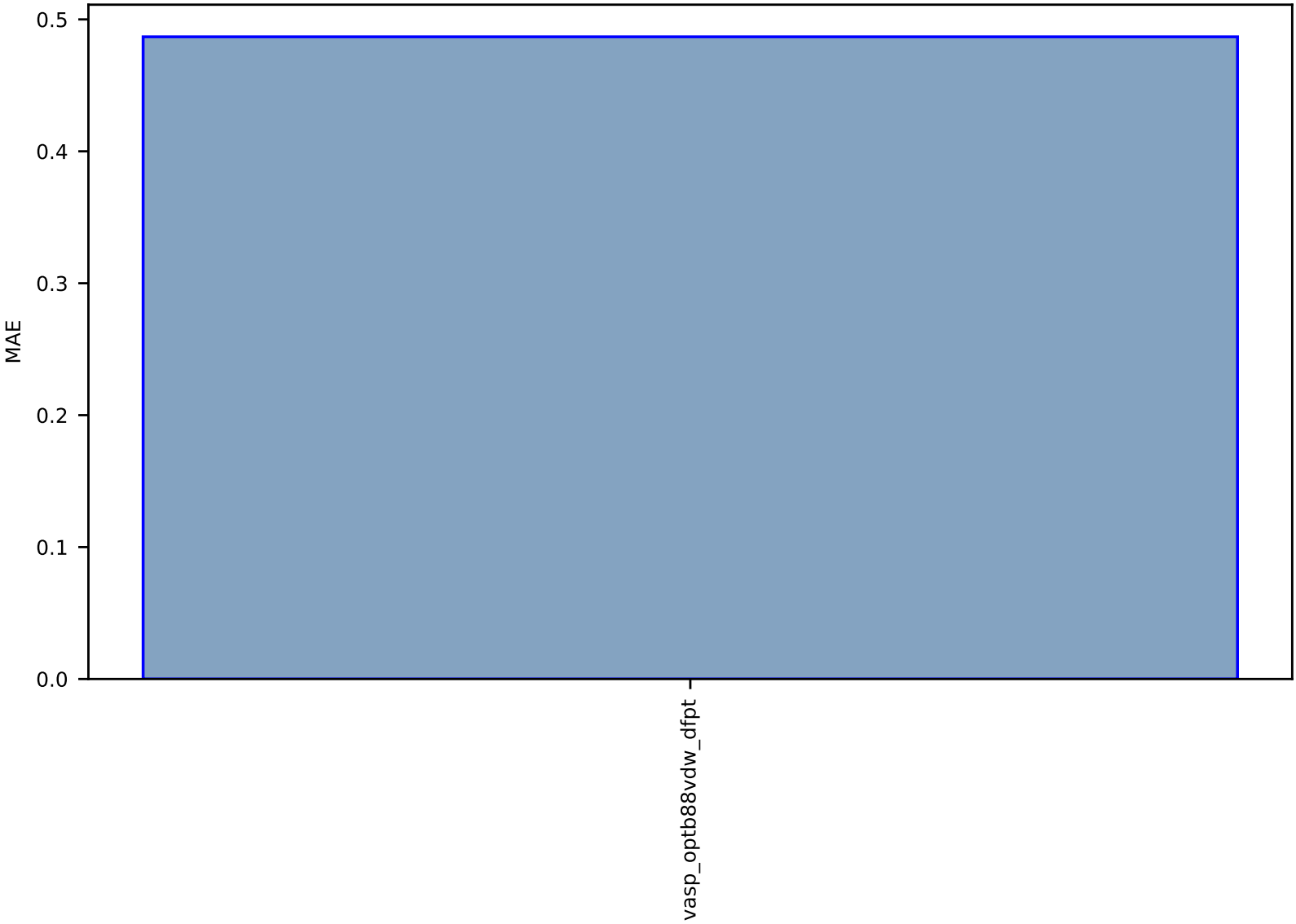


Fig S255:ES-SinglePropertyPrediction-max_piezoelec_JVASP_1240_LiNbO3-dft_3d-test-mae.csv.zip

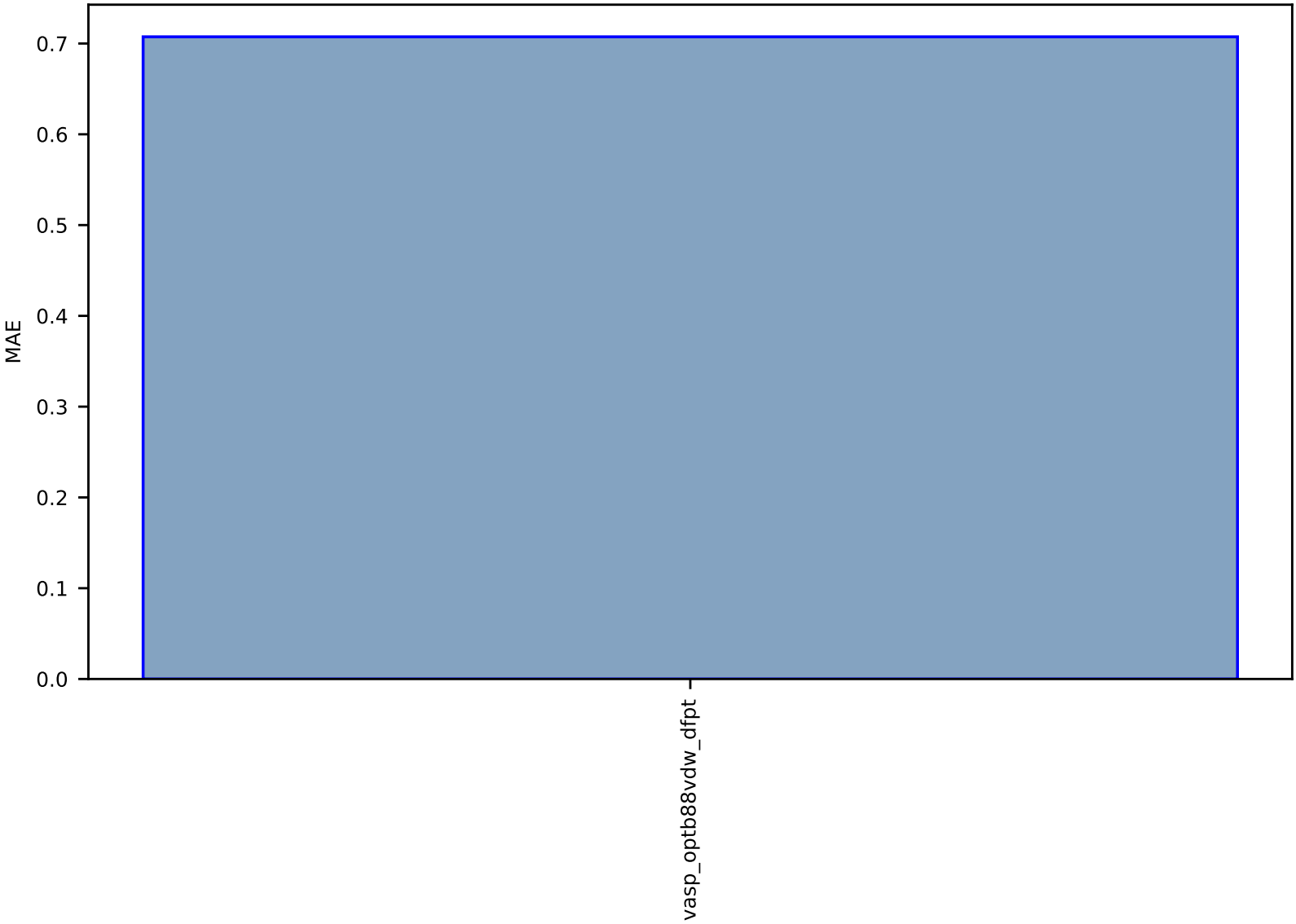


Fig S256:ES-SinglePropertyPrediction-max_piezoelec_JVASP_39_AlN-dft_3d-test-mae.csv.zip

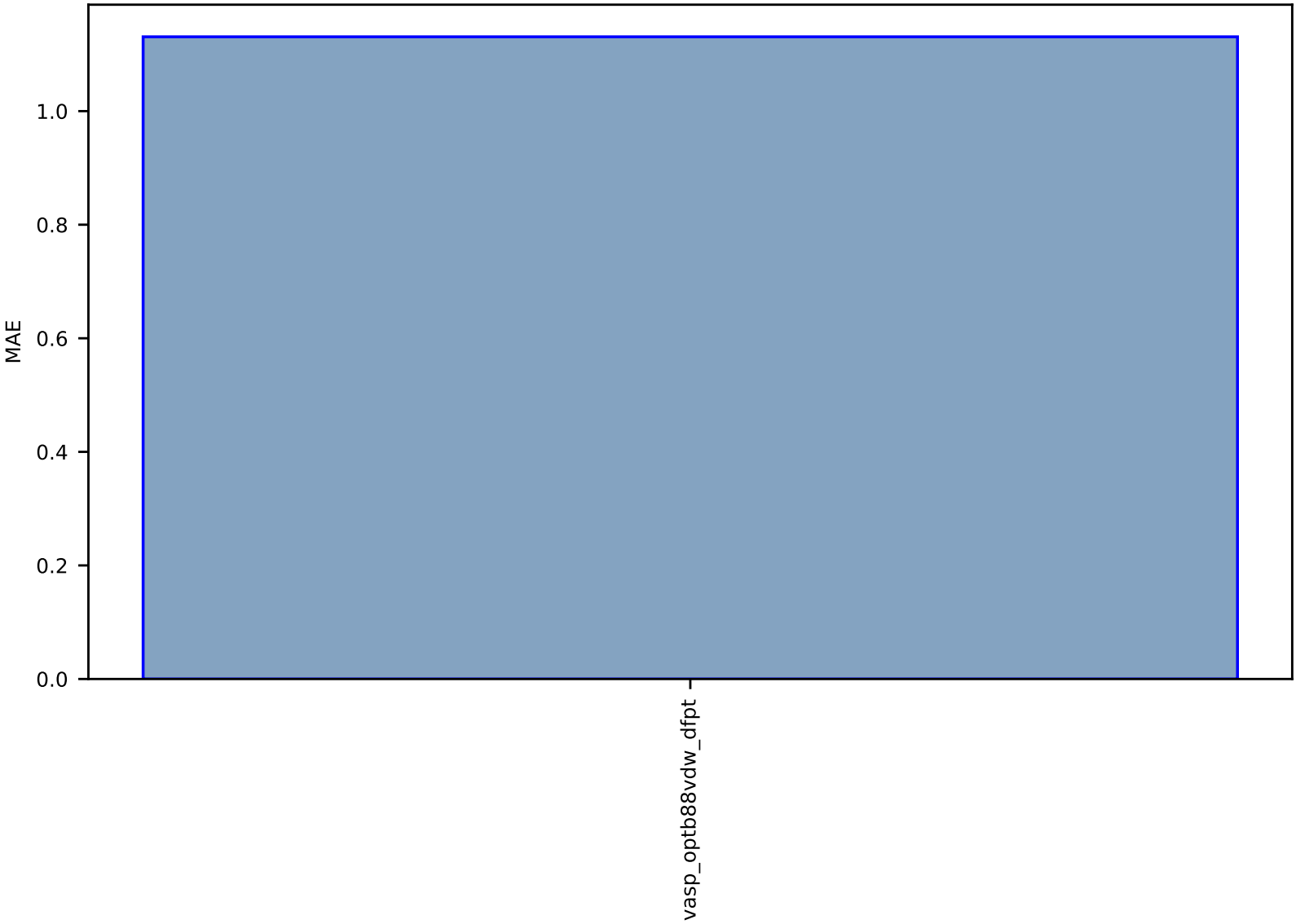


Fig S257:ES-SinglePropertyPrediction-max_piezoelec_JVASP_41_SiO2-dft_3d-test-mae.csv.zip

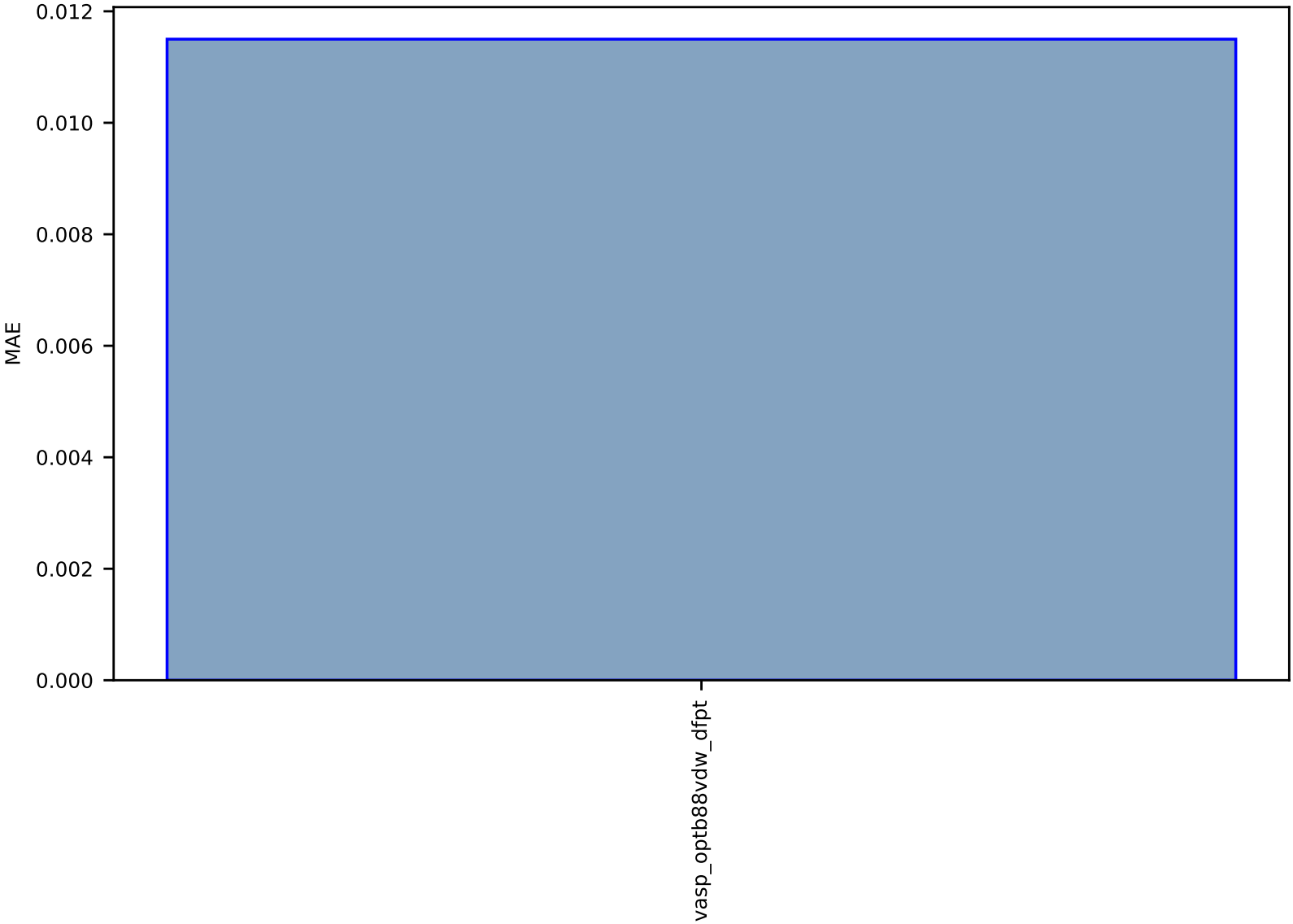


Fig S258:ES-SinglePropertyPrediction-max_piezoelec_JVASP_35711_GaSb-dft_3d-test-mae.csv.zip

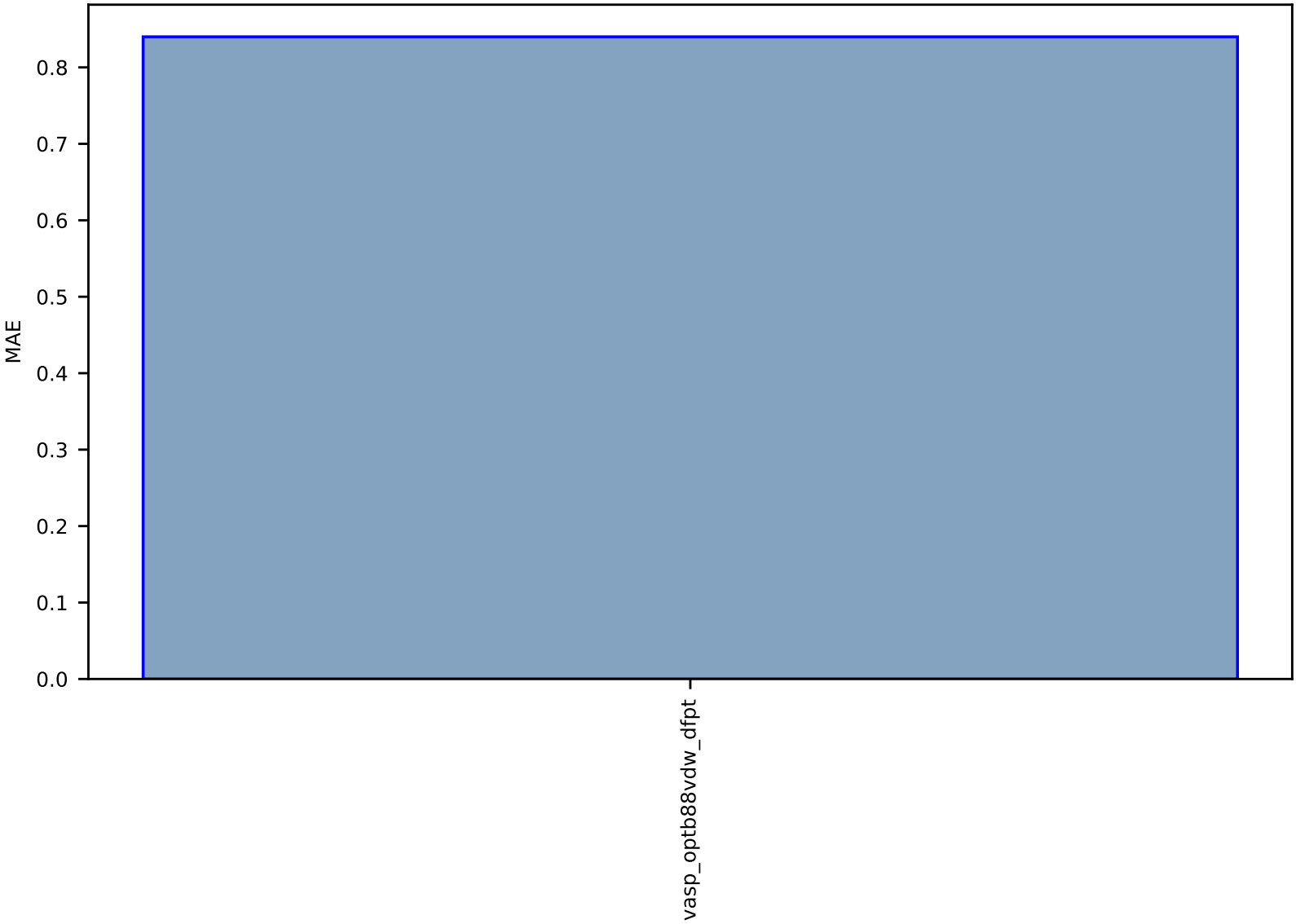


Fig S259:ES-SinglePropertyPrediction-max_piezoelec_JVASP_30_GaN-dft_3d-test-mae.csv.zip

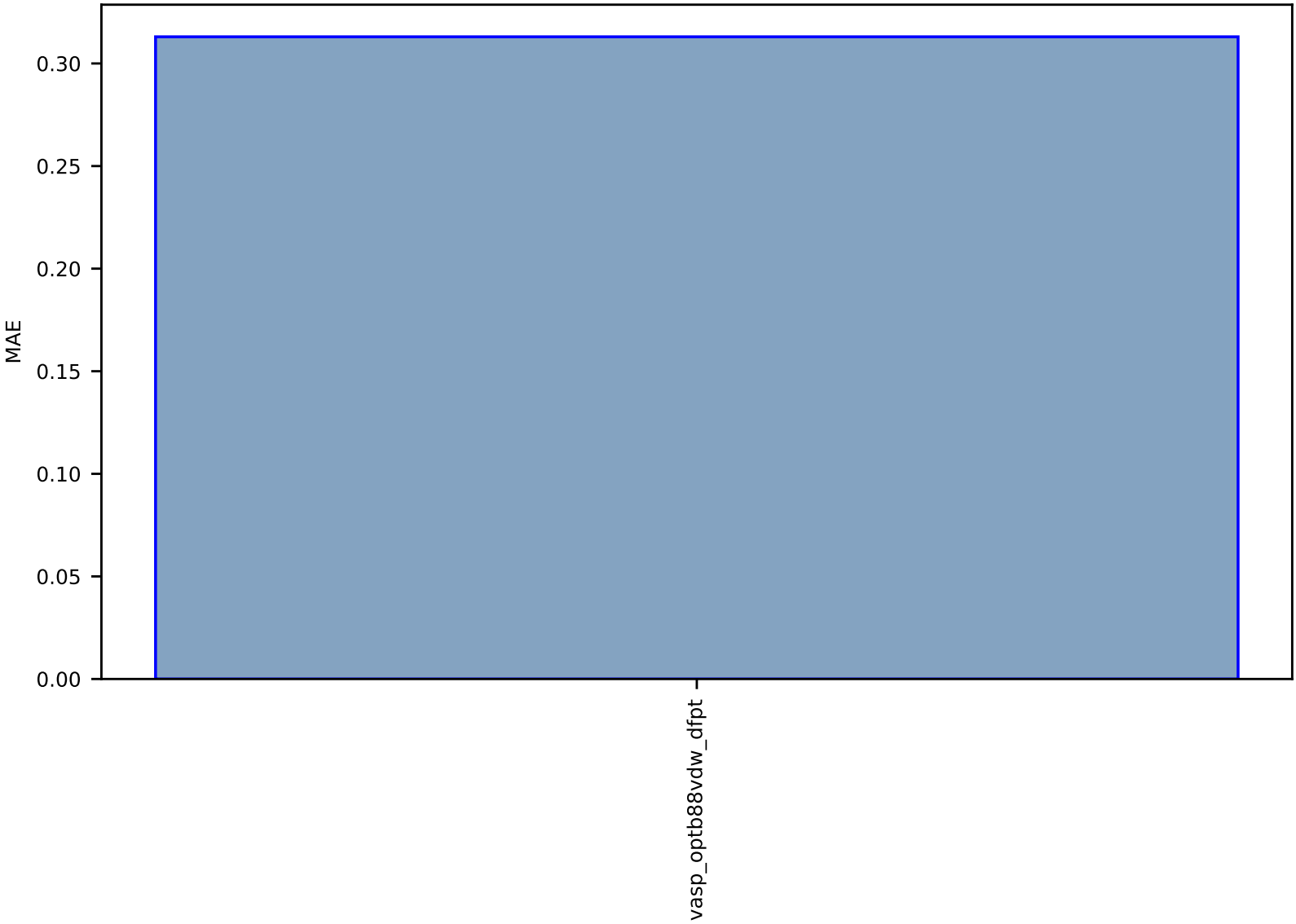


Fig S260:ES-SinglePropertyPrediction-max_piezoelec_JVASP_1195_ZnO-dft_3d-test-mae.csv.zip

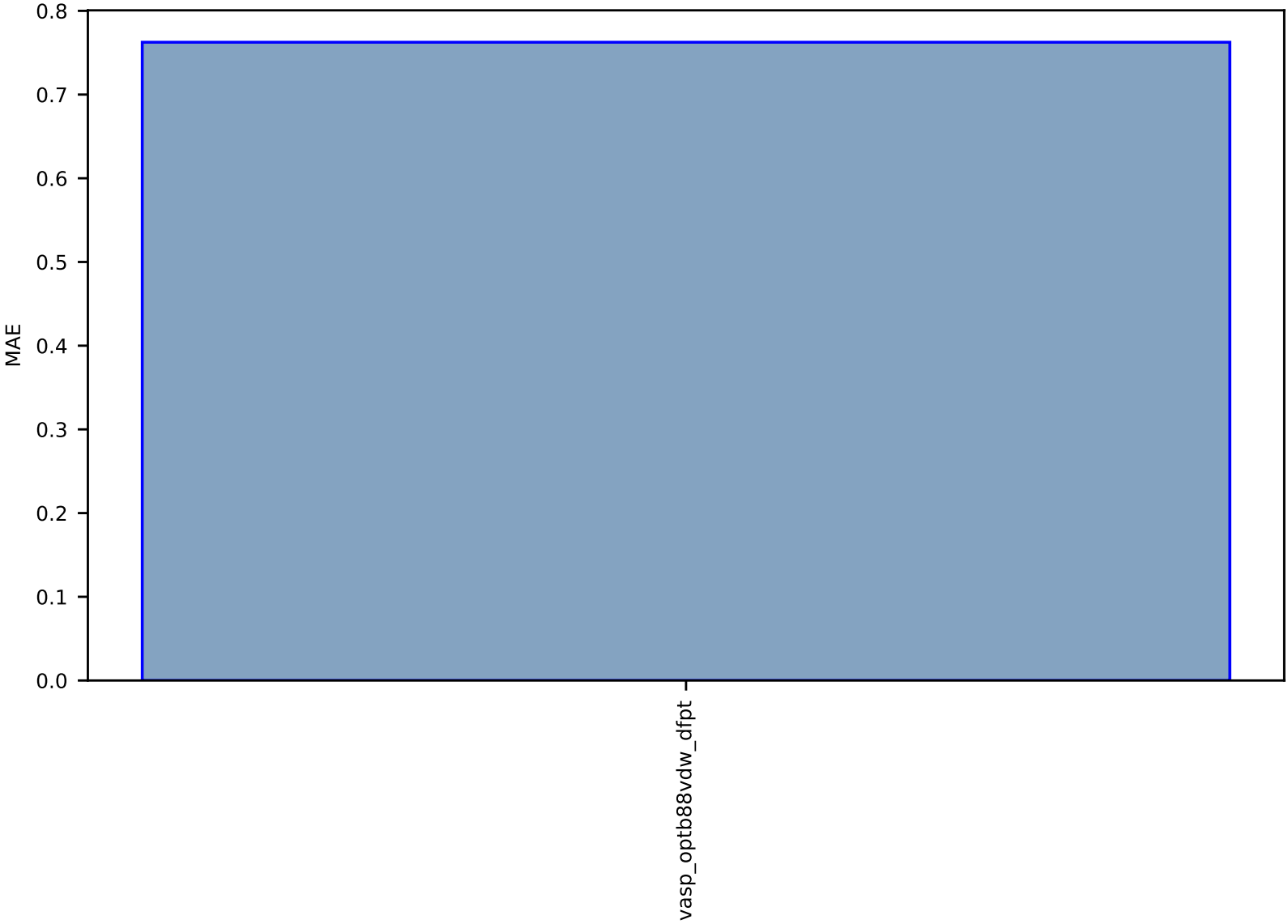


Fig S261:ES-SinglePropertyPrediction-max_piezoelec_JVASP_1180_InN-dft_3d-test-mae.csv.zip

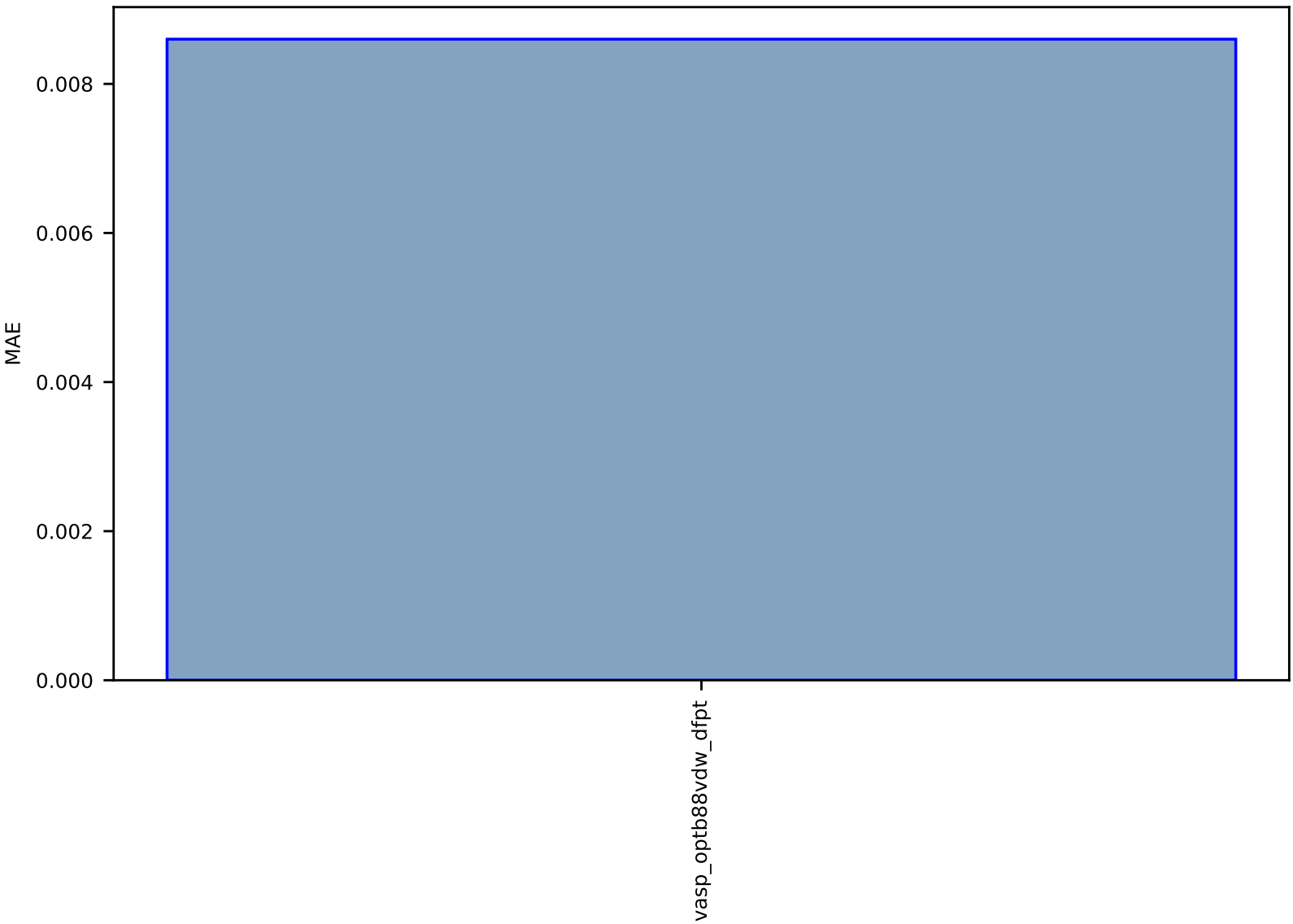


Fig S262:ES-SinglePropertyPrediction-max_piezoelec_JVASP_1327_AlP-dft_3d-test-mae.csv.zip

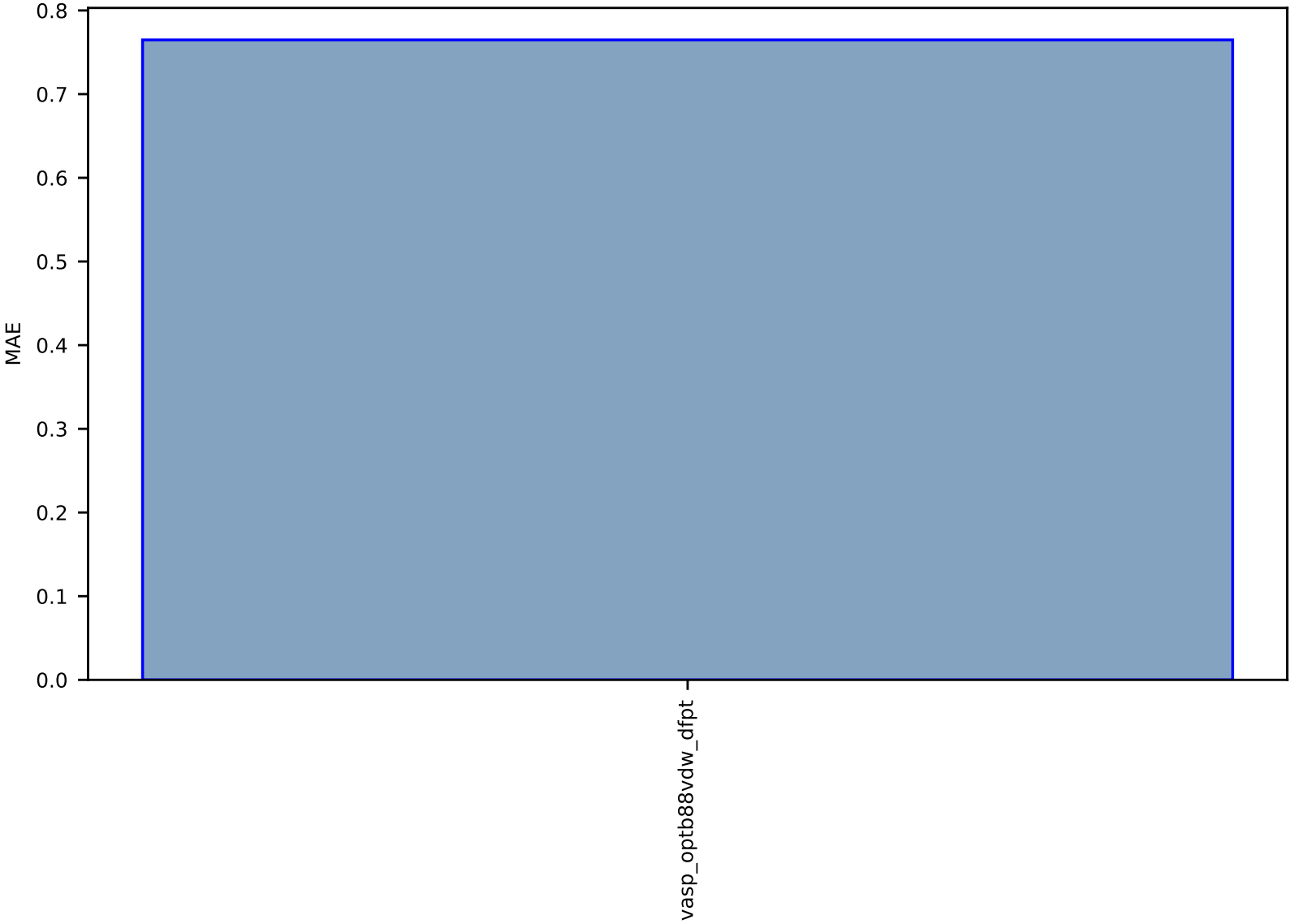


Fig S263:ES-SinglePropertyPrediction-max_piezoelec_JVASP_110_BaTiO3-dft_3d-test-mae.csv.zip

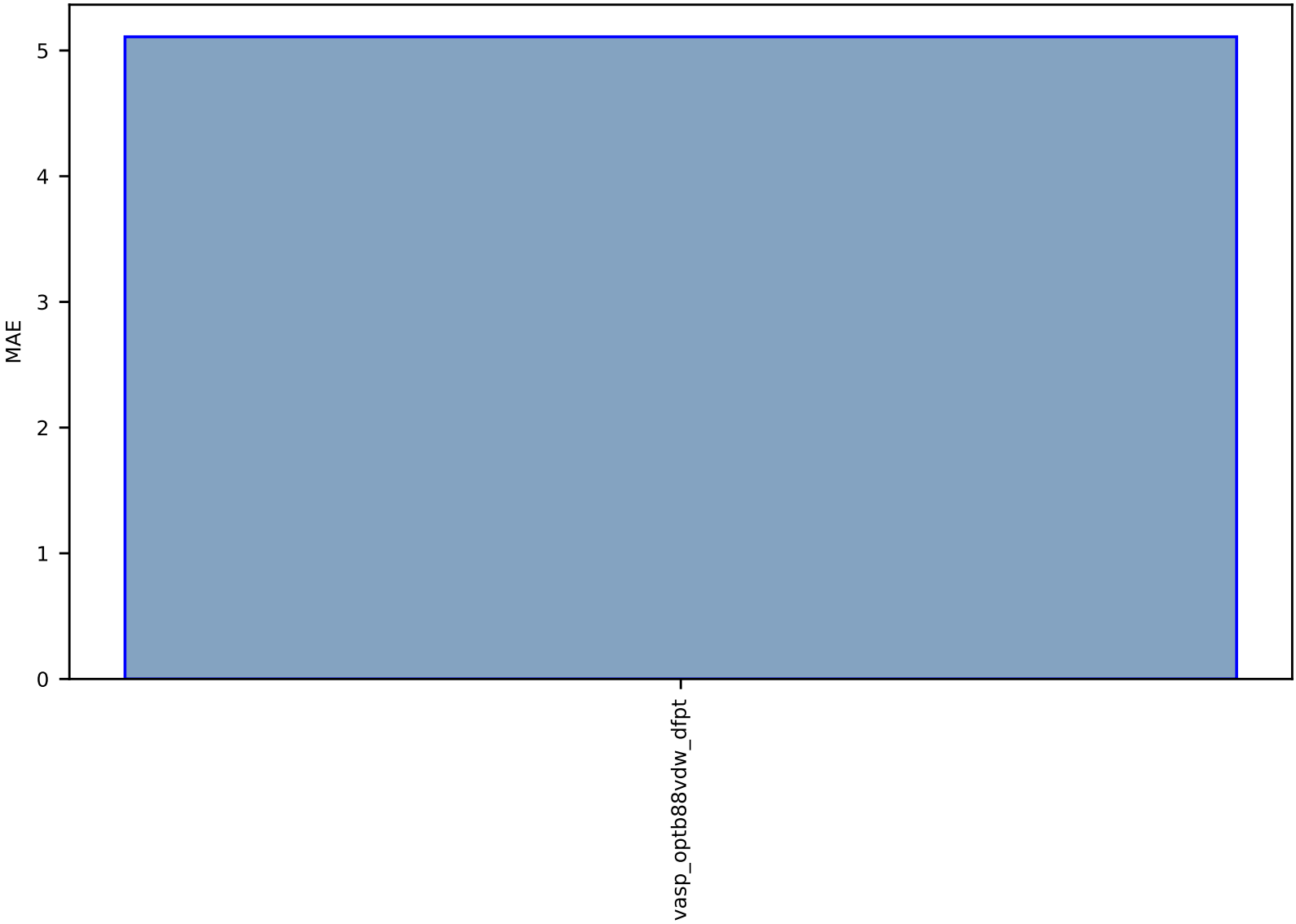


Fig S264:ES-SinglePropertyPrediction-max_piezoelec_JVASP_1372_AlAs-dft_3d-test-mae.csv.zip

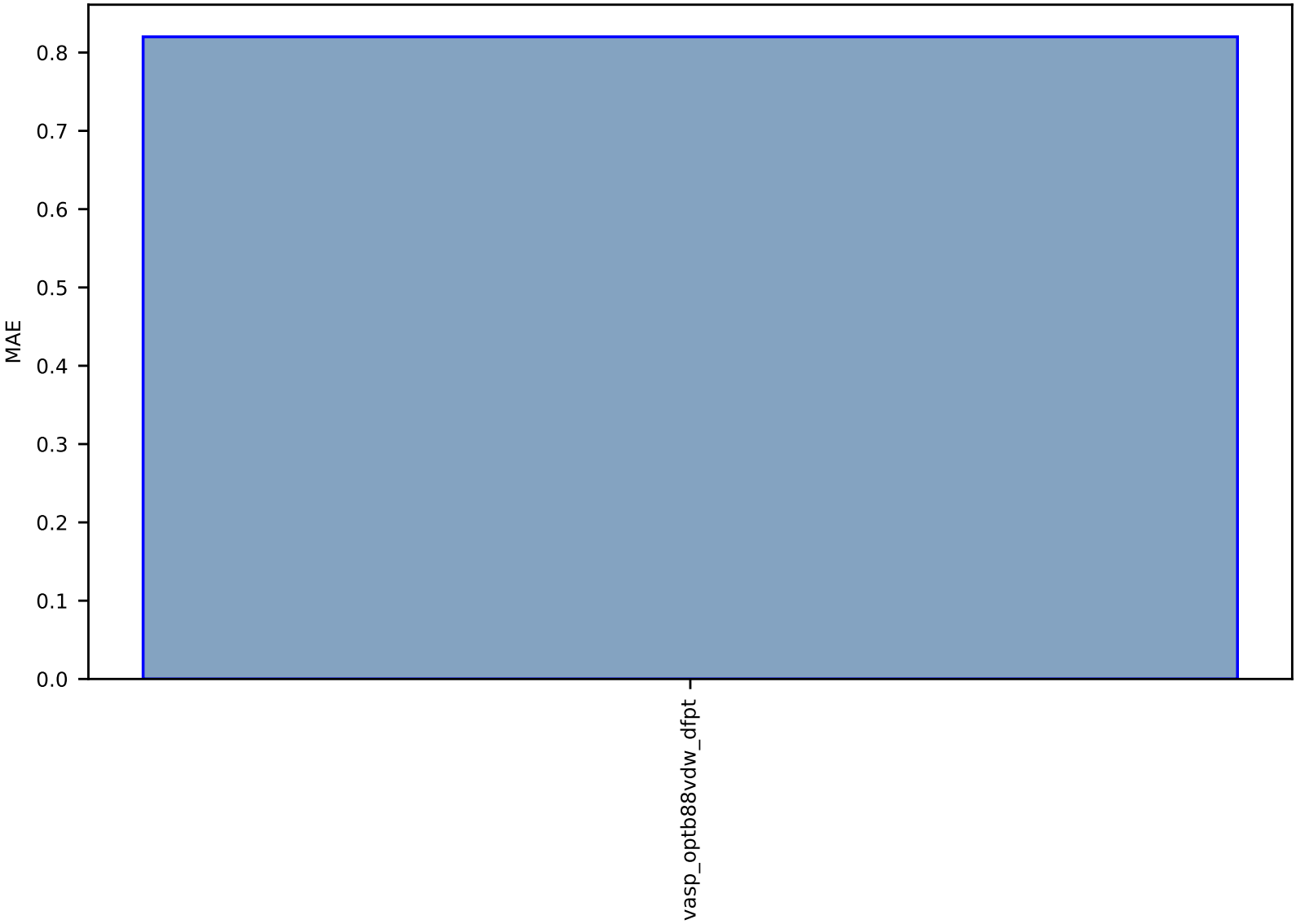


Fig S265:ES-SinglePropertyPrediction-max_piezoelec-dft_3d-test-mae.csv.zip

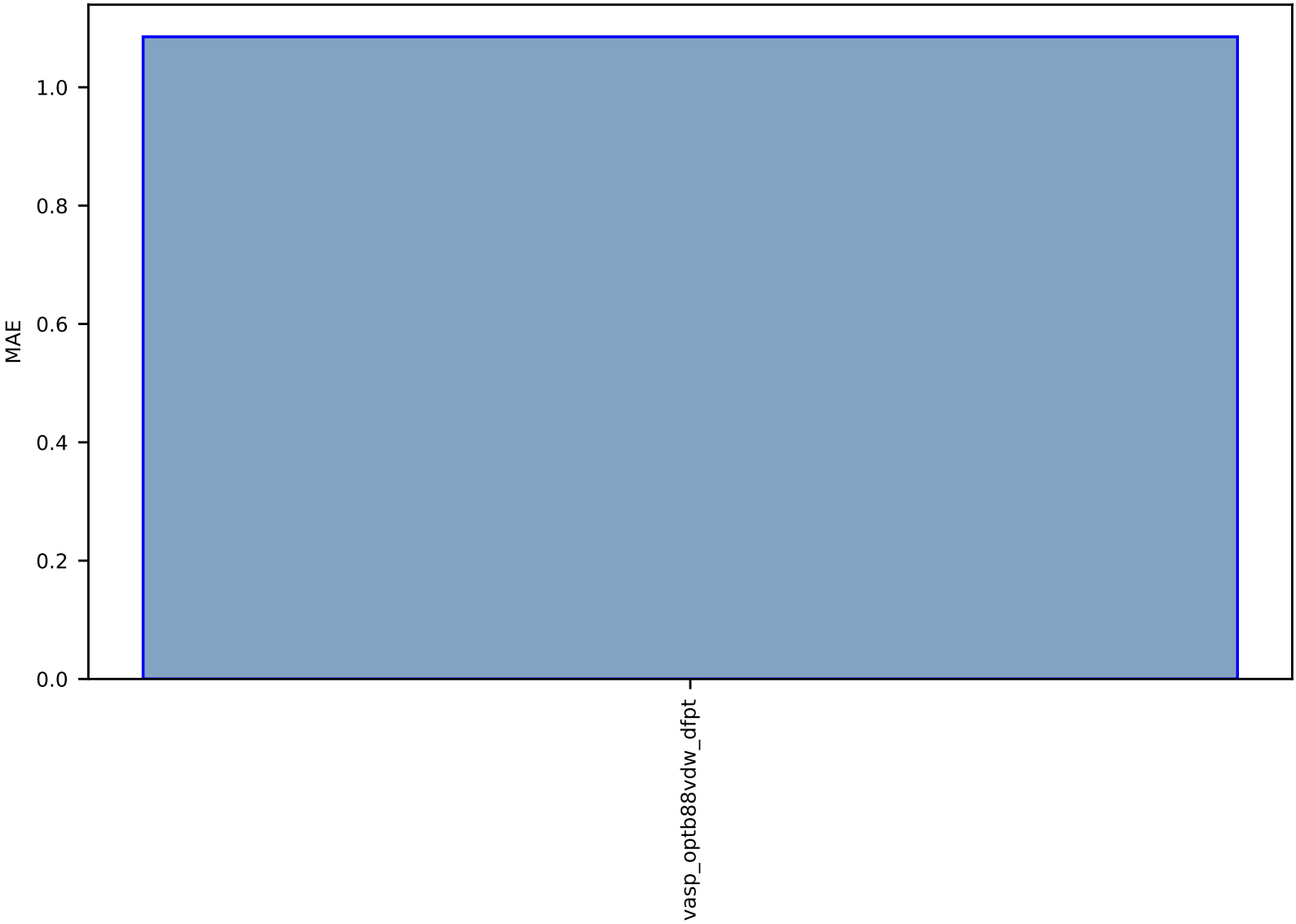


Fig S266:ES-SinglePropertyPrediction-bulk_modulus_JVASP_20793_Y2O3-dft_3d-test-mae.csv.zip

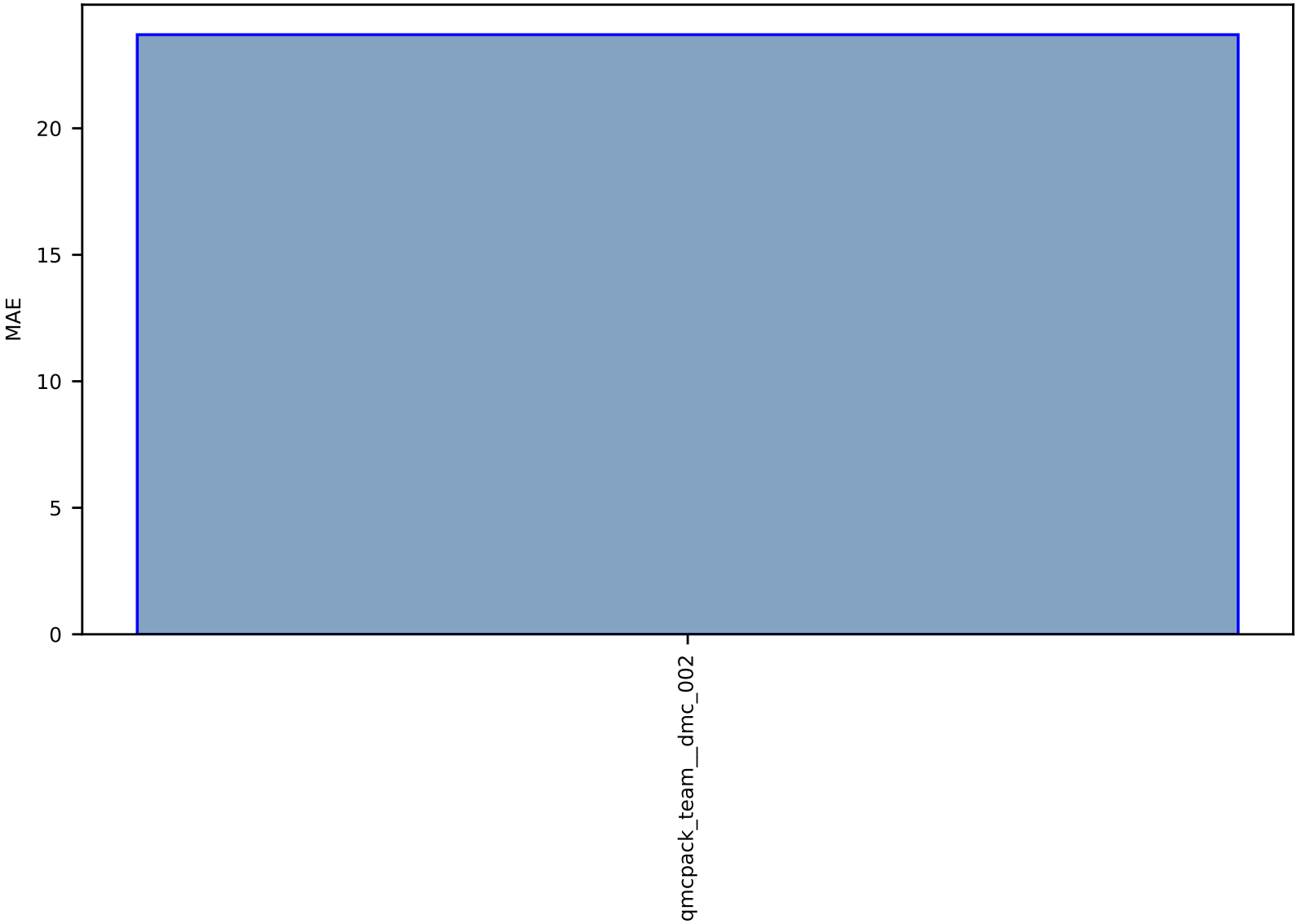


Fig S267:ES-SinglePropertyPrediction-bulk_modulus_JVASP_92796_La2O3-dft_3d-test-mae.csv.zip

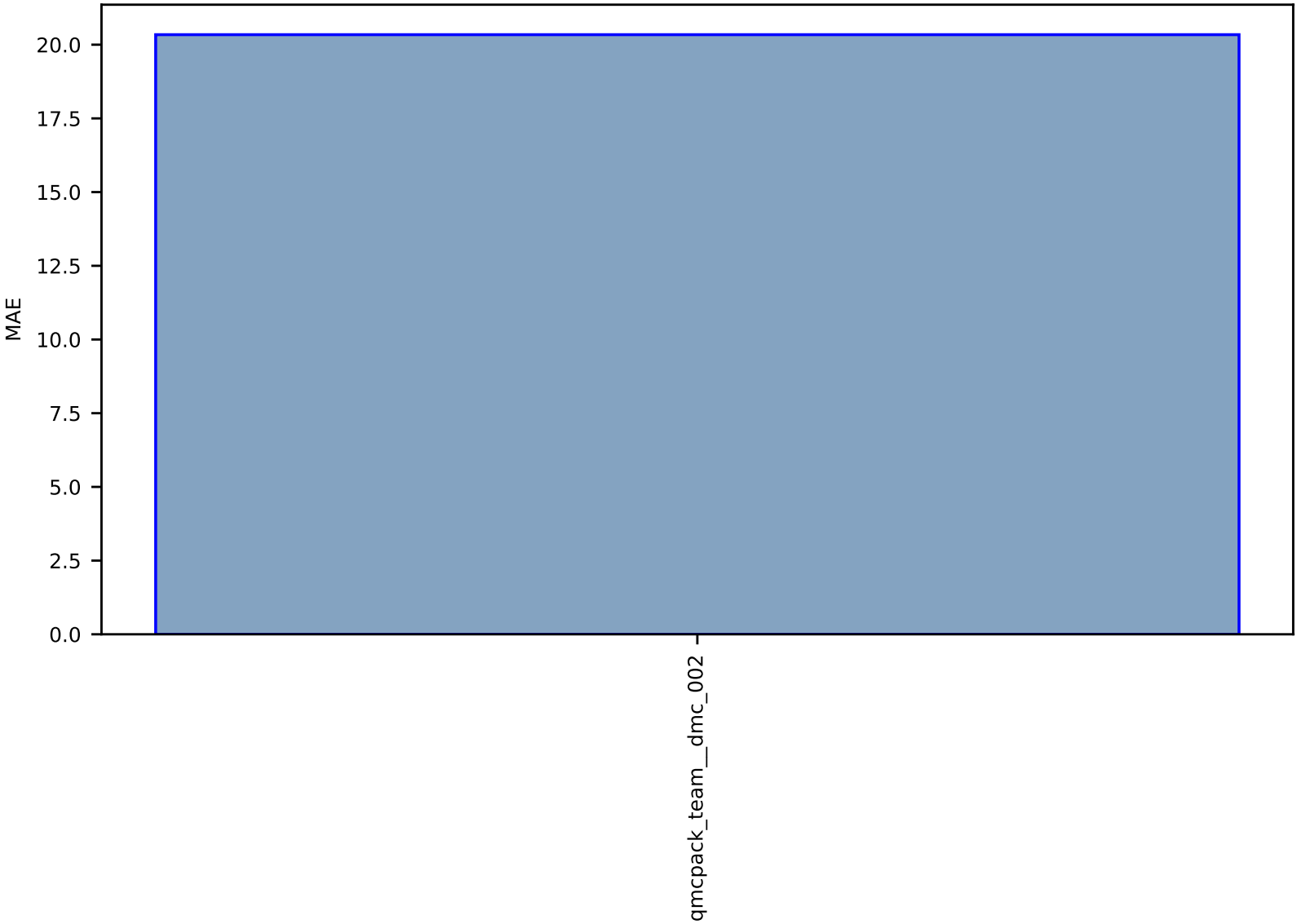


Fig S268:ES-SinglePropertyPrediction-bulk_modulus_JVASP_87128_CoO-dft_3d-test-mae.csv.zip

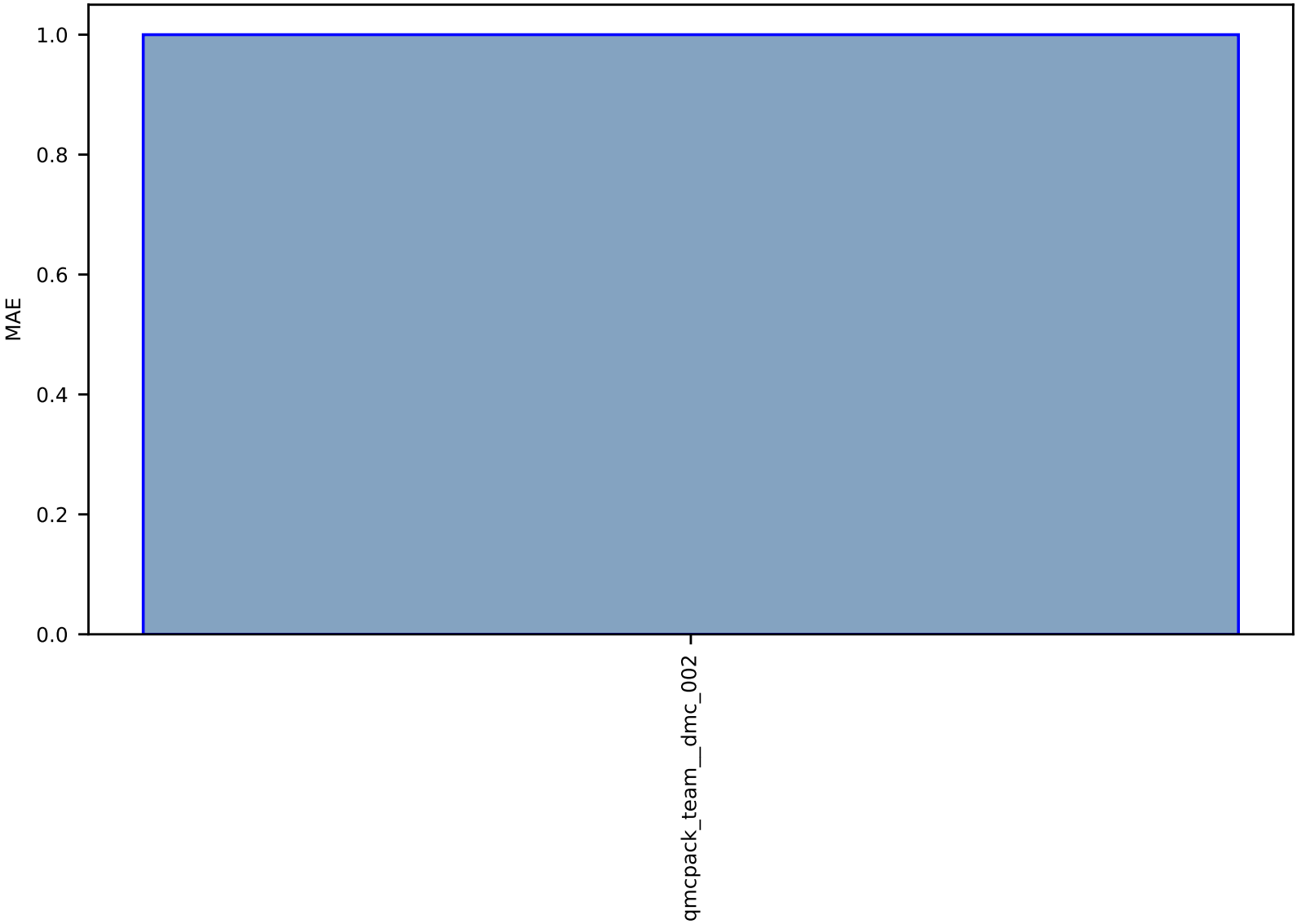


Fig S269:ES-SinglePropertyPrediction-bulk_modulus_JVASP_20290_Sc2O3-dft_3d-test-mae.csv.zip

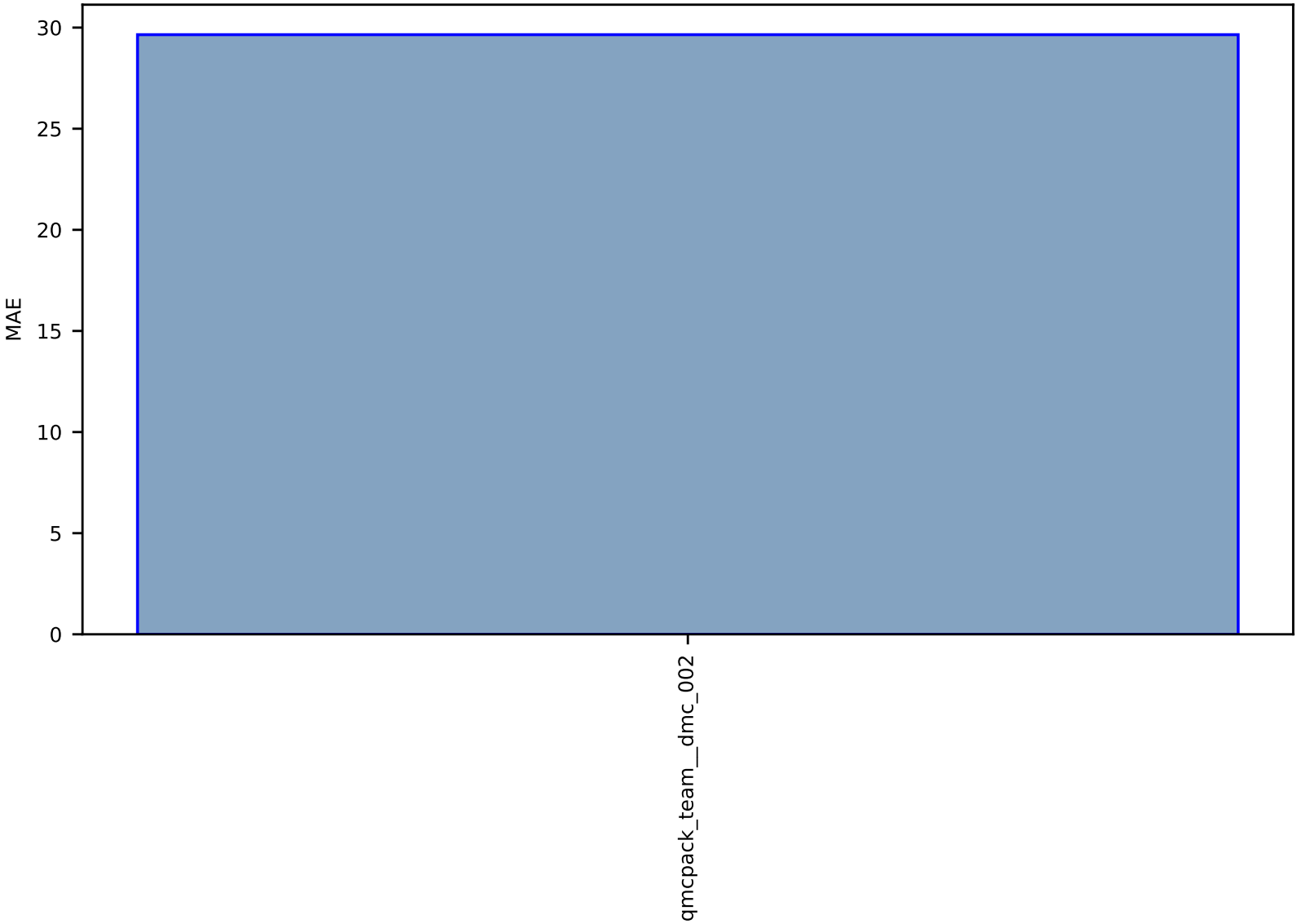


Fig S270:ES-SinglePropertyPrediction-bulk_modulus_JVASP_1405_CaO-dft_3d-test-mae.csv.zip

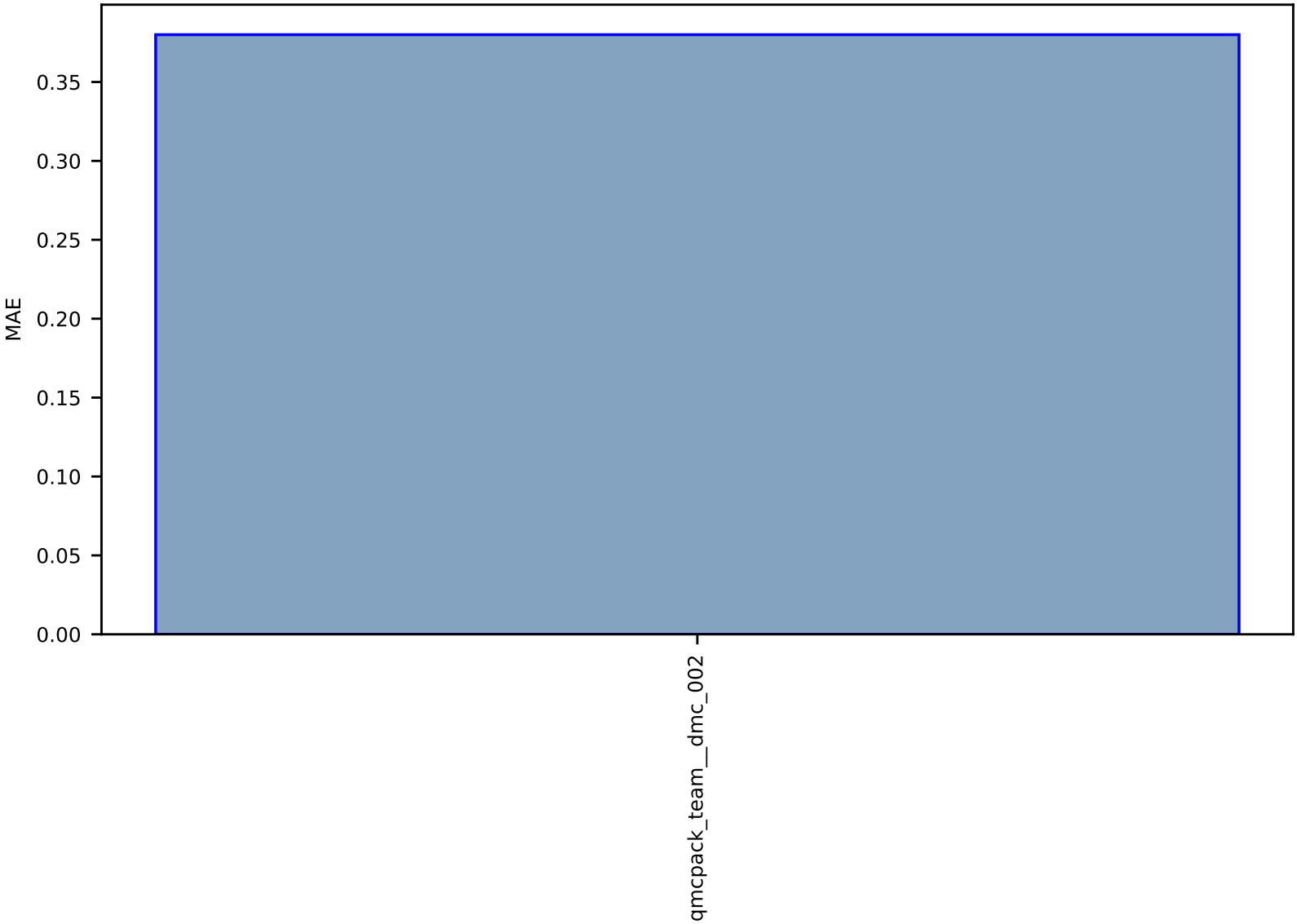


Fig S271:ES-SinglePropertyPrediction-bulk_modulus_JVASP_7870_MnO-dft_3d-test-mae.csv.zip

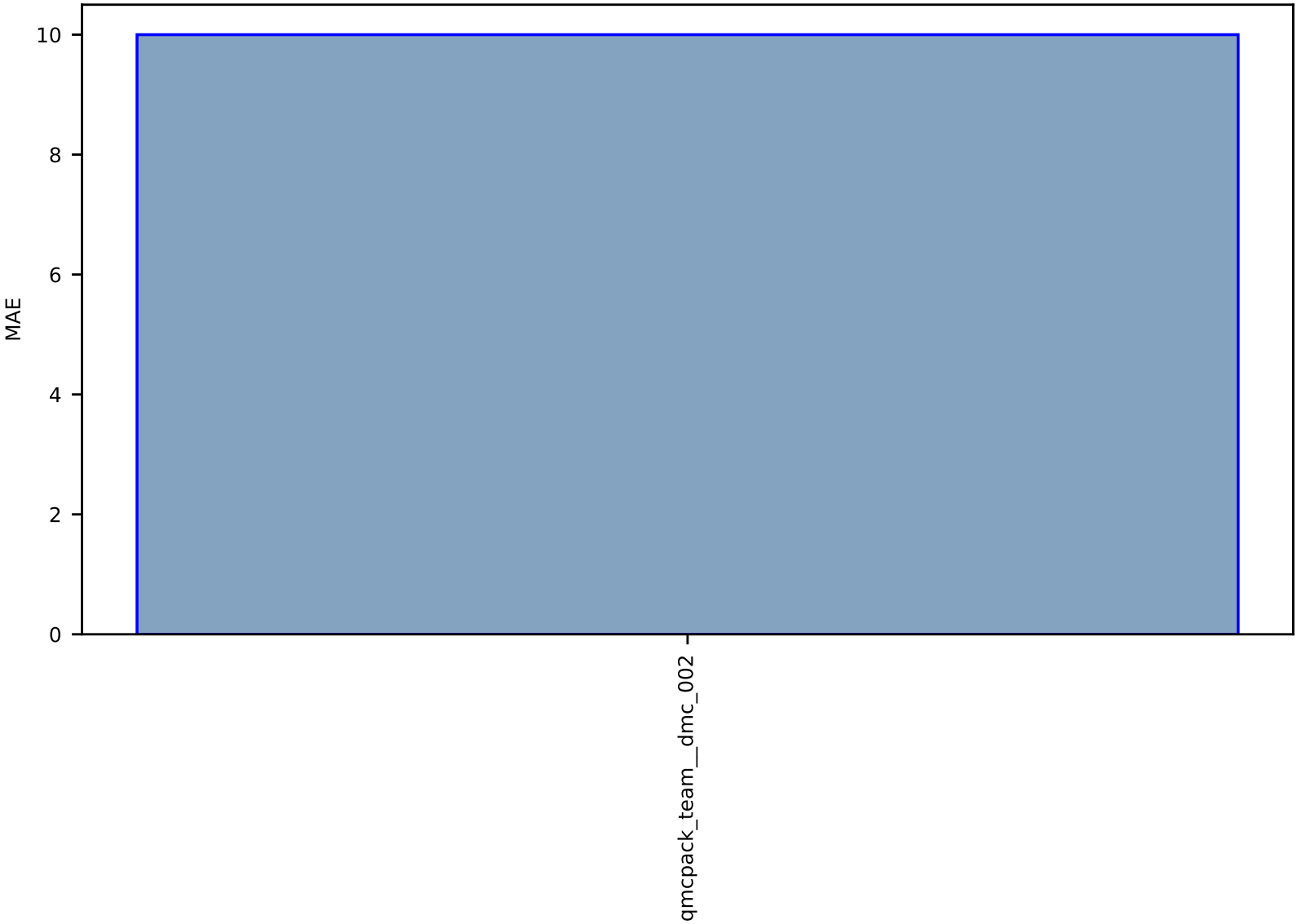


Fig S272:ES-SinglePropertyPrediction-bulk_modulus_JVASP_1306_BaO-dft_3d-test-mae.csv.zip

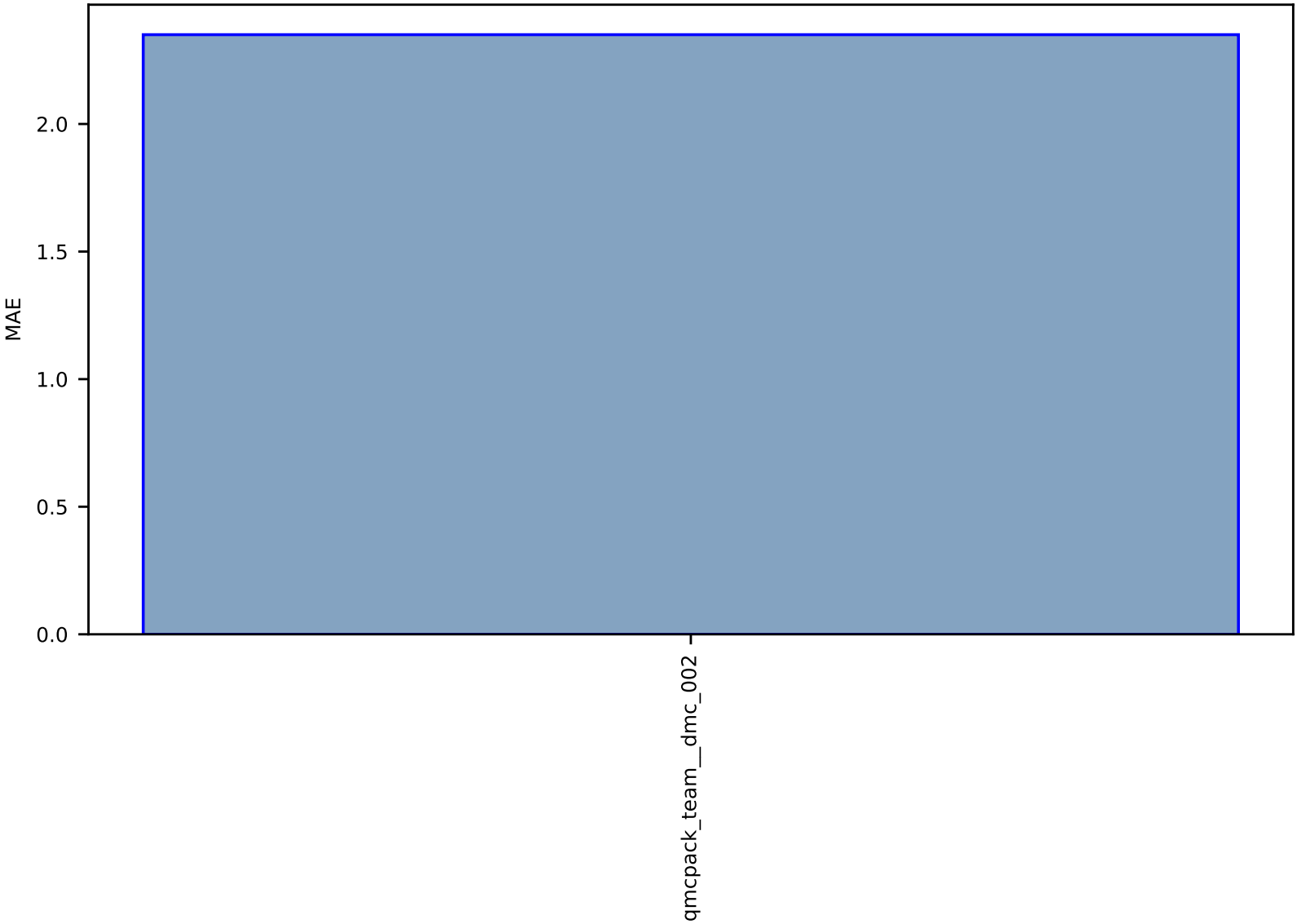


Fig S273:ES-SinglePropertyPrediction-bulk_modulus_JVASP_1390_SrO-dft_3d-test-mae.csv.zip

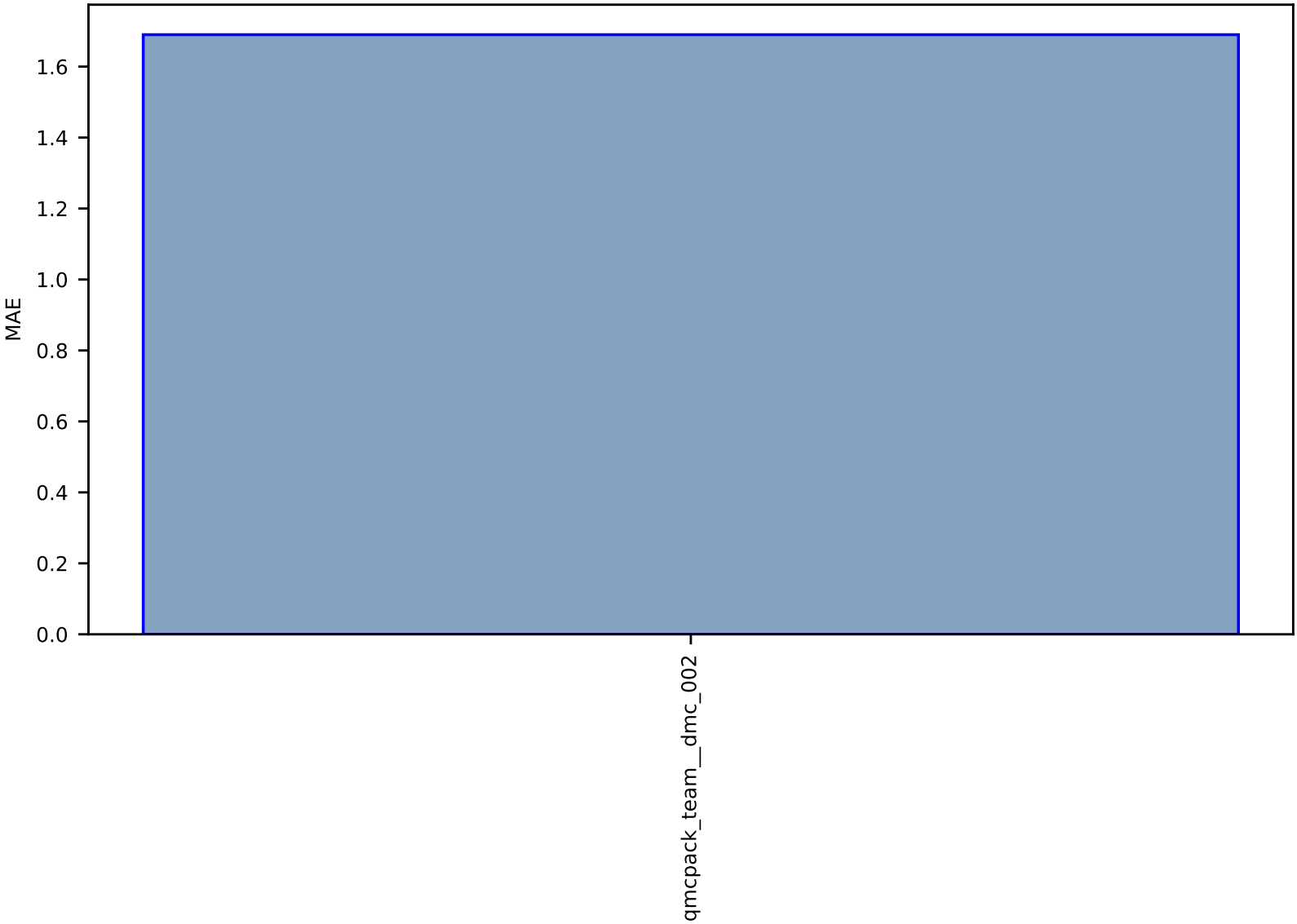


Fig S274:AI-TokenClass-labels-mat_scholar_ner-test-acc.csv.zip