Supplementary Information

Towards Accurate and Scalable High-throughput MOF Adsorption Screening: Merging Classical Force Fields and Universal Machine Learned Interatomic Potentials

Satyanarayana Bonakala[a], Mohammad Wahiduzzaman[a], Taku Watanabe[b], Karim Hamzaoui[b], and Guillaume Maurin[a,c*]

[a]ICGM, Univ. Montpellier, CNRS, ENSCM, 34293 Montpellier, France

[b]Matlantis Corporation, Otemachi, Chiyoda-ku, 100-0004 Tokyo, Japan

[c]Institut Universitaire de France, 75005 Paris, France

Email: guillaume.maurin1@umontpellier.fr

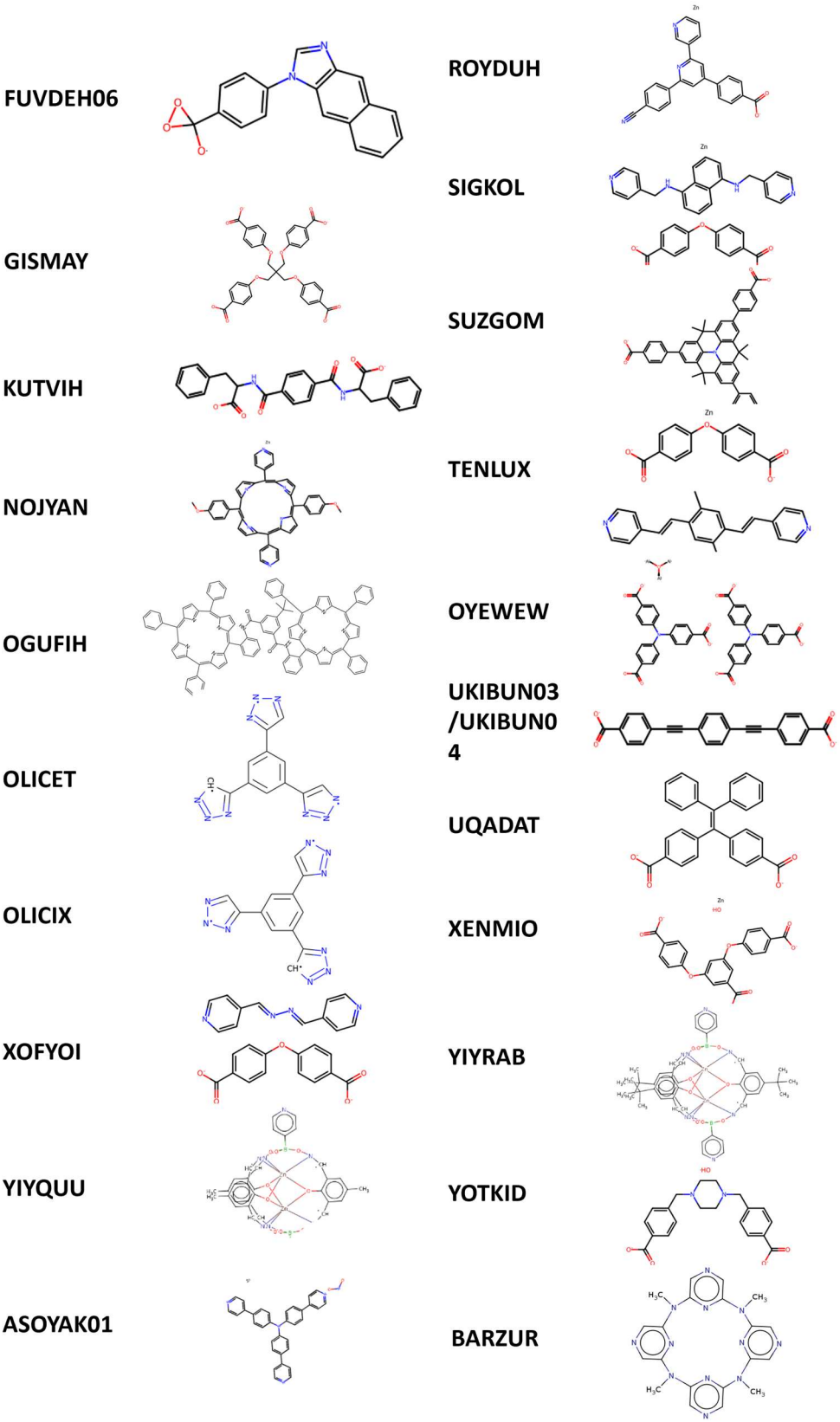


Fig. S1 CSD refcodes and corresponding linker structures of MOFs evaluated during the synthetic feasibility assessment. The left column lists the MOF CSD refcodes, while the right column illustrates the associated organic linker molecules extracted using the MOFid toolkit. These linkers were analyzed for complexity, and synthetic accessibility.

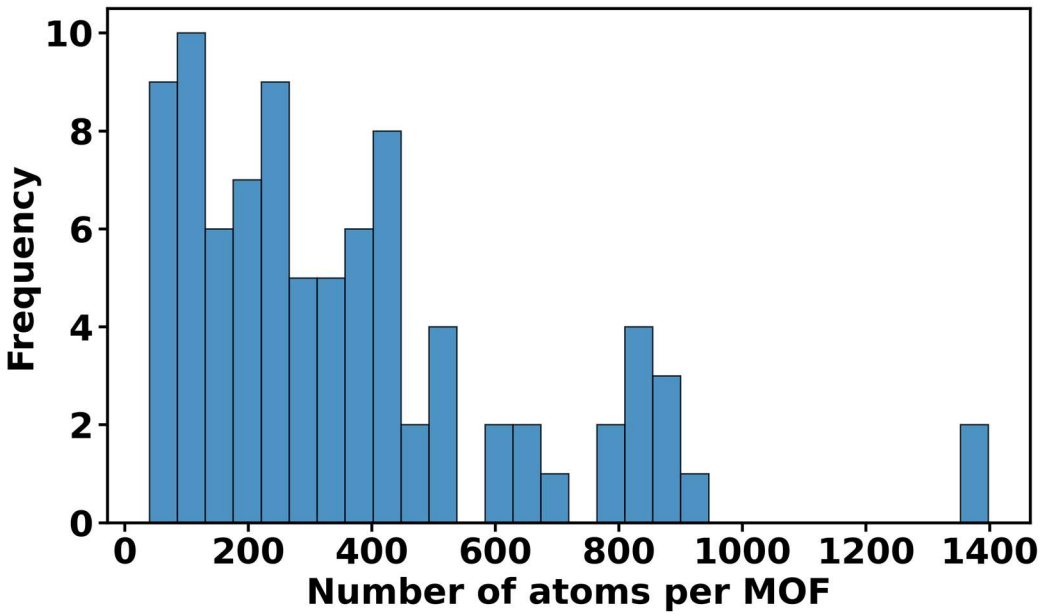


Fig. S2 Distribution of the number of atoms per MOF in the benchmarking dataset of 88 structures.

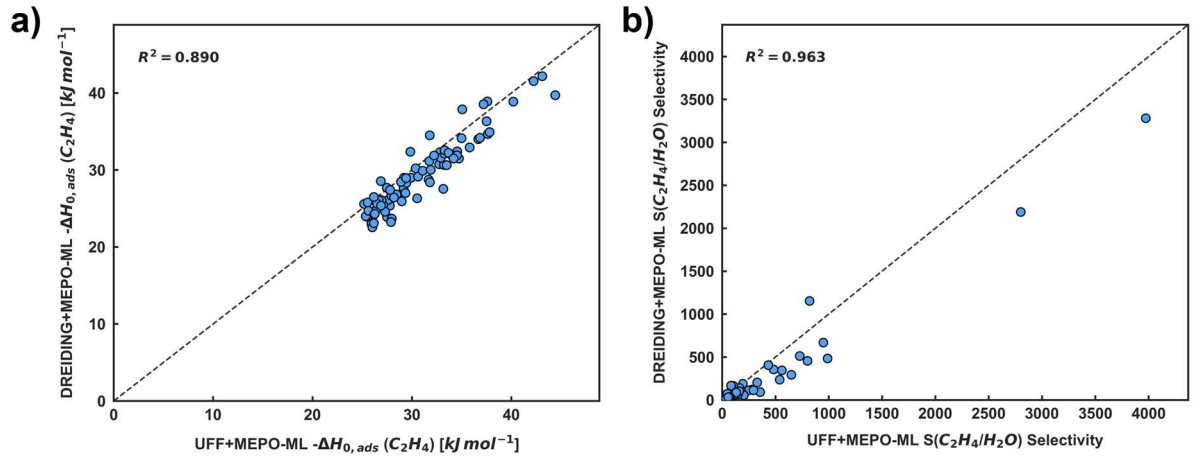


Fig. S3. Comparison of ethylene adsorption performance predicted using UFF- and DREIDING-based workflows. Parity plots showing the correlation between **(a)** ethylene adsorption enthalpy ($-\Delta H_{0,ads}(C_2H_4)$) and **(b)** ideal $S(C_2H_4/H_2O)$ selectivity across the 88 MOF subset. Both workflows utilize the same MEPO-ML charges for electrostatic interactions. The high R^2 values (0.89 and 0.96, respectively).

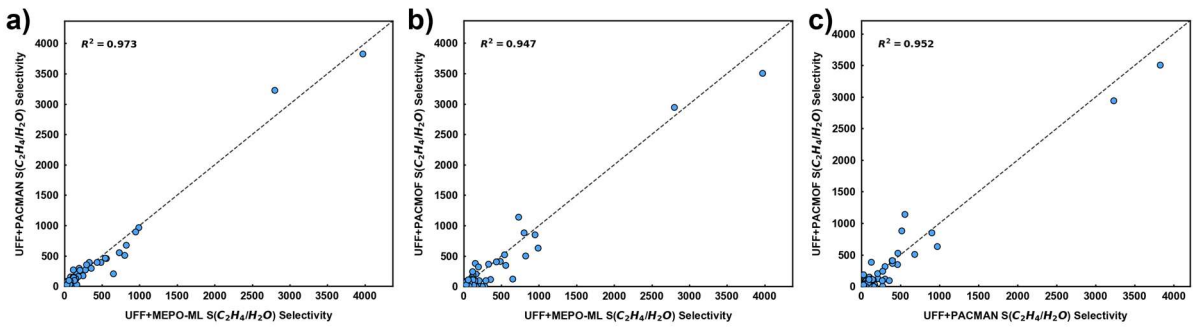


Fig. S4. Comparison of ideal $S(C_2H_4/H_2O)$ selectivity using different ML-based partial charges. Parity plots showing the correlation of ideal ethylene/water selectivity calculated using UFF paired with three different ML charge schemes for the 88 MOF subset: (a) PACMAN vs. MEPO-ML charges, (b) PACMOF vs. MEPO-ML charges, and (c) PACMOF vs. PACMAN charges.

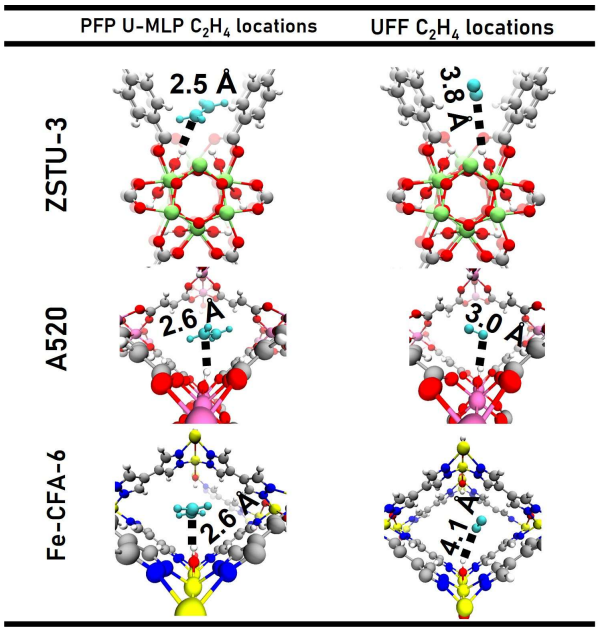


Fig. S5. Ethylene (cyan) adsorption configurations in three outlier MOFs (red crosses in Figure 4e of the main text), comparing predictions from PFP u-MLIP (left) and UFF (right). Distances indicate the shortest framework–ethylene contacts, highlighting the stronger binding geometries captured by PFP u-MLIP compared to UFF. Ethylene hydrogen atoms are omitted in FF simulations as it is described with Fisher et al.,[1] united-atom model.

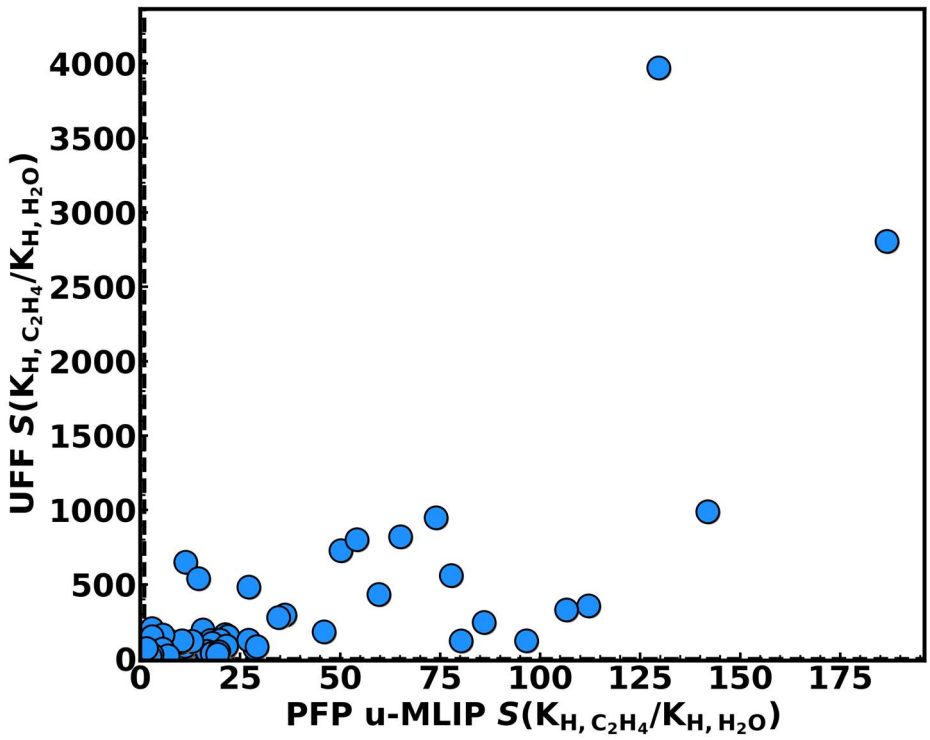


Fig. S6. Comparison of ethylene/water ideal selectivity $S(C_2H_4/H_2O)= K_{H,C_2H_4}/K_{H,H_2O}$, predicted from UFF and PFP u-MLIP Widom insertion simulations for the identified 88 MOF subset. Each point represents one MOF. Dashed lines indicate $S(C_2H_4/H_2O)=1$.

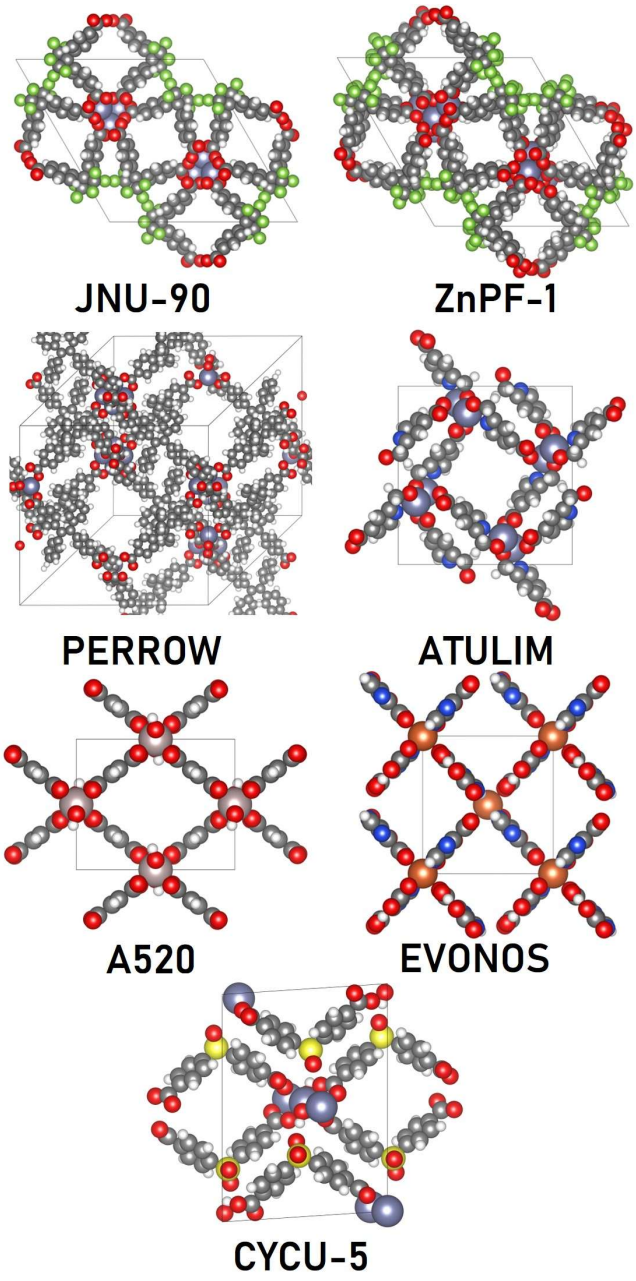


Fig. S7. Crystal structures of selected top-performing MOFs identified in Figure 4f.

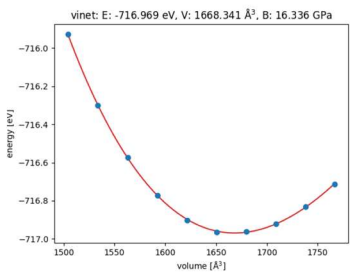
Table S1. Comparison of calculated ethylene adsorption enthalpies ($-\Delta H_{0,ads}(C_2H_4)$) with experimental isosteric heat of adsorption (Q_{st}) values reported in the literature.

MOF	$-\Delta H_{0,ads}(C_2H_4)$ (kJ·mol^{-1})	Experimental Q_{st} (kJ·mol^{-1})
A520	32.7	30.50[2]
MIL-53ht (Al)	26.1	30.50[3]
$Zn_3(Im)_6$	31.7	27.66[4]
Zn(hba)	29.2	26.10[5]

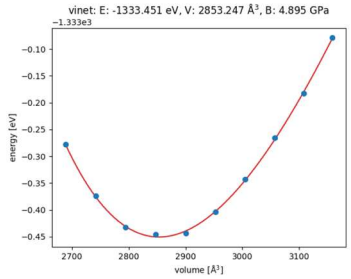
Table S2. Adsorption enthalpy and volume variations of the seven top-performing MOFs identified in Figure 4f. Shown are the differences in ethylene adsorption enthalpies ($\Delta(\Delta H_{0,ads}(C_2H_4))$) between rigid and cell-optimized MOF structures, together with the corresponding percentage volume change ($\Delta V\%$). $\Delta(\Delta H_o,ads(C_2H_4))$ is defined as $\Delta H_{0,ads}(C_2H_4)$ of the rigid MOF minus $\Delta H_{0,ads}(C_2H_4)$ of the cell-optimized MOF. $\Delta V\%$ is calculated as: $\Delta V\% = V(\text{rigid MOF}) - V(\text{cell-optimized MOF}) \times 100 / V(\text{rigid MOF})$

MOF	$\Delta(\Delta H_{0,ads}(C_2H_4))$ (kJ·mol^{-1})	$\Delta V\%$
JNU-90	-0.32	-2.2
ZnPF-1	-1.88	-4.3
PERROW	0.68	0.2
ATULIM	-0.23	-1.6
A520	-0.39	19
EVONOS	-4.13	2.2
CYCU-5	-0.49	1.5

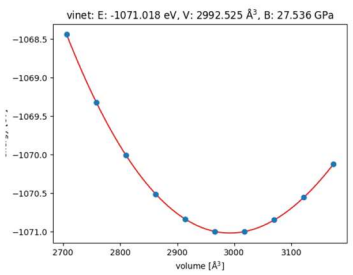
ATULIM_full



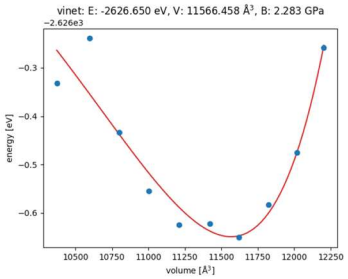
AVIMIC_full



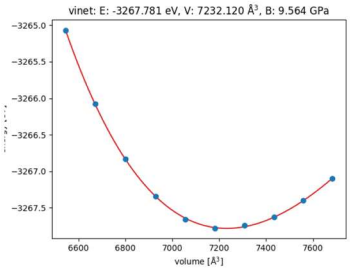
AWAREV_full



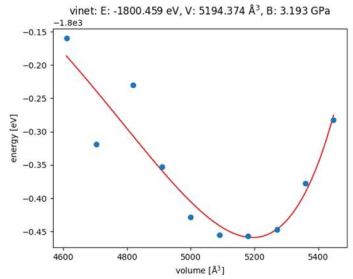
BIJZIG_full



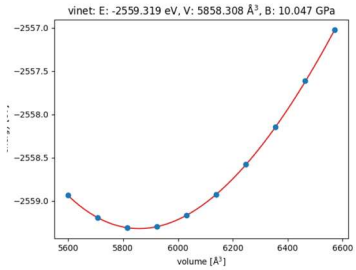
BODPUH_full



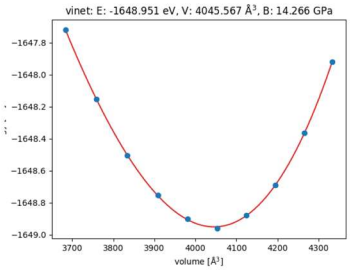
BOMCOX_full



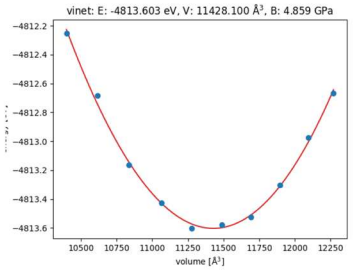
BUVWOF02_full



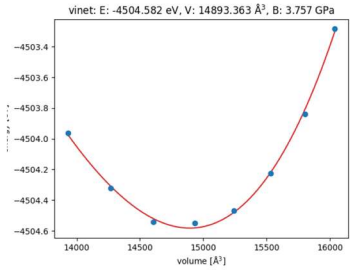
CAWVIH_SL



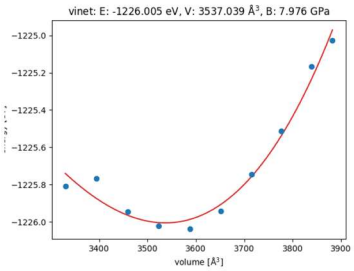
CETTUT_full



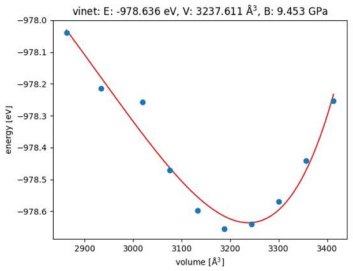
DAQMAN_full



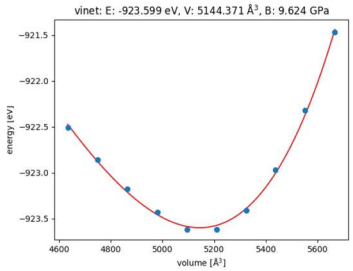
DORCIX02_P1



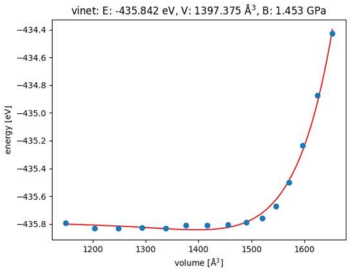
ECAHAT_full



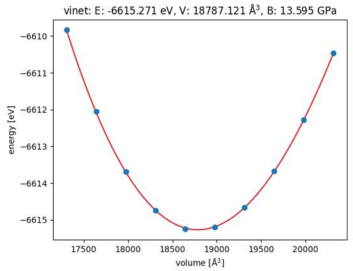
EGEJIK_manual



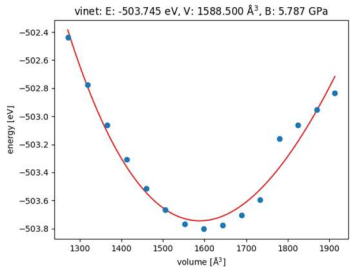
EGELUY01_manual



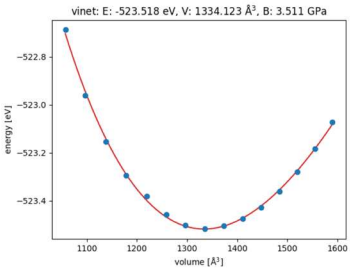
EHETER02_full



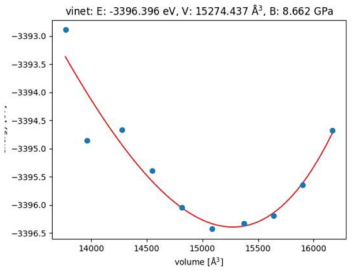
ETEDEN_full



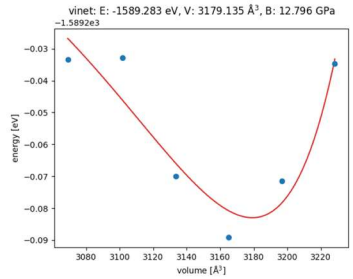
FAQDUA_full



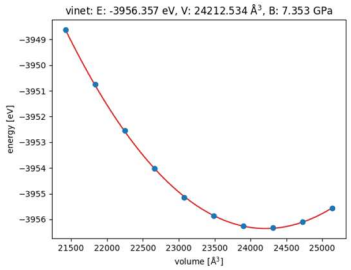
FERHAN_full



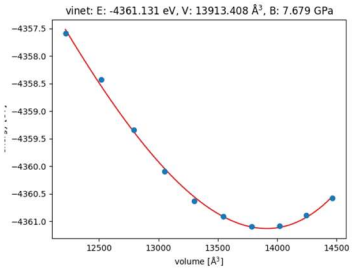
FUVFUZ_full



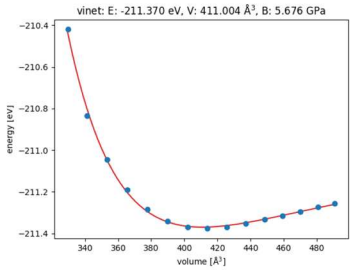
GITVIP01_full



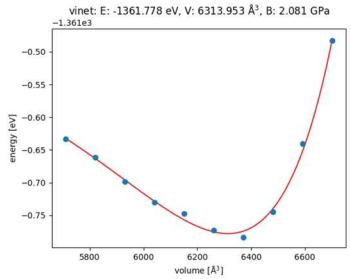
GOVSER_full



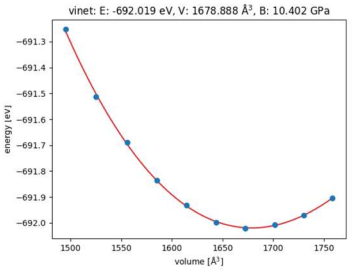
HAMXEC_full



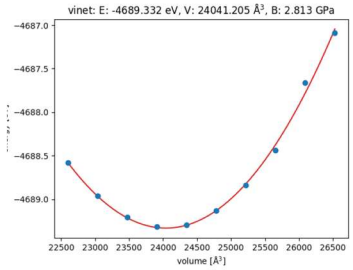
HIFVOI_full



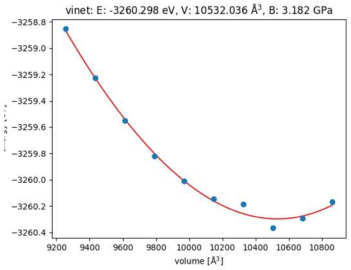
HONCEU_P1



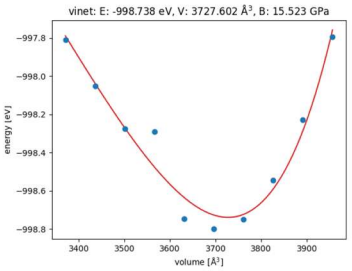
ICICAB_P1



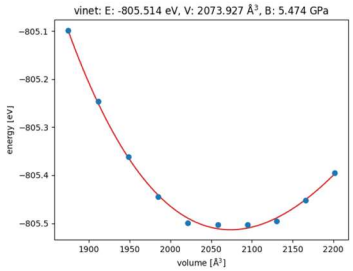
KULWAR_full



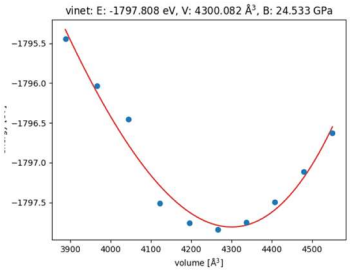
KUXSIH02_P1



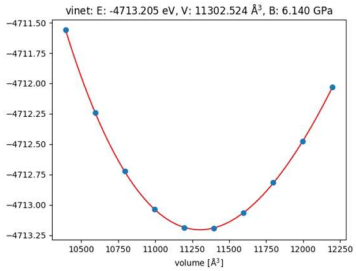
LARHUK_full



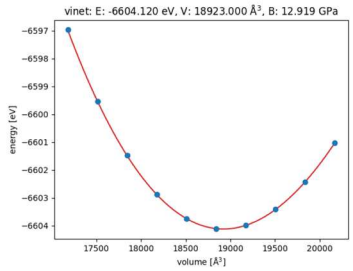
LEDBAB_full



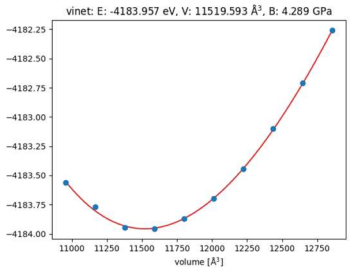
LEVPAF_full



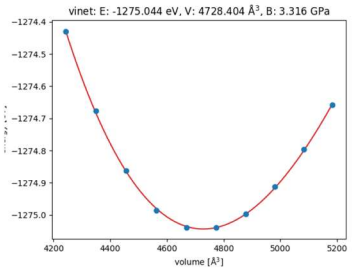
MECWIB_full



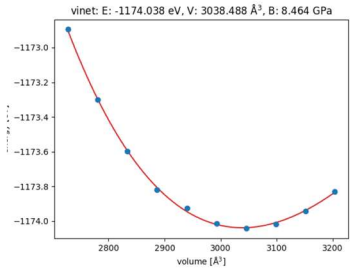
MUFTUD_full



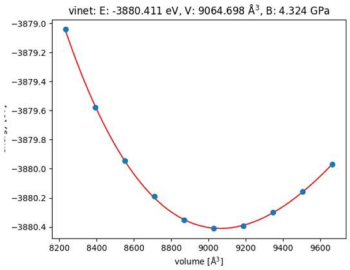
NEGTEB_full



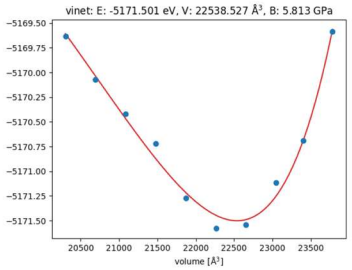
NOPGUY_P1



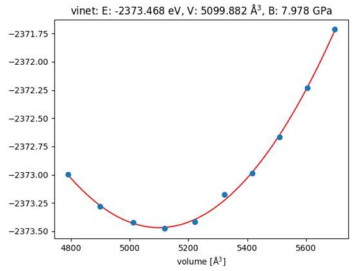
OPOQES_full



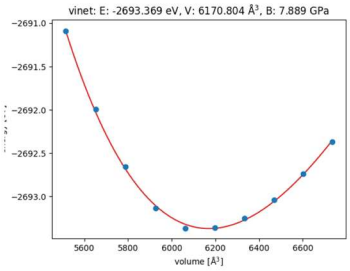
PERROW_full



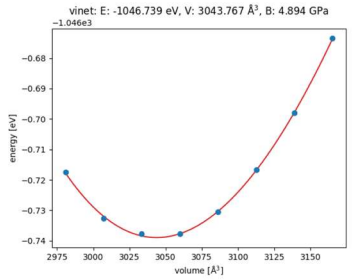
PETTAK_full



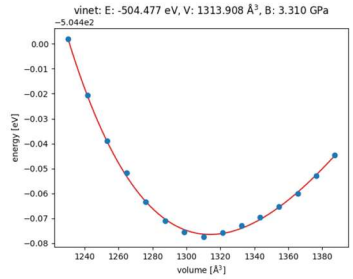
POGMOQ_full



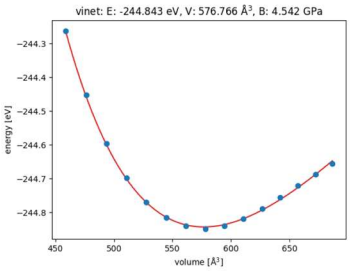
POKXIA_full



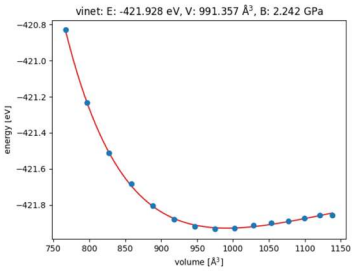
QAQTEJ_full



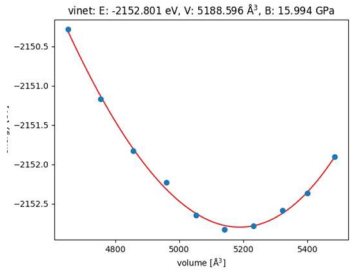
QUQPOI_full



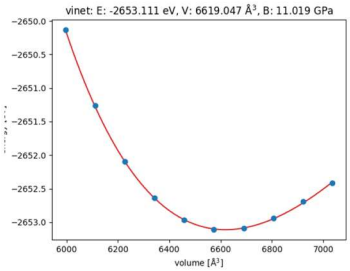
REBQIB_full



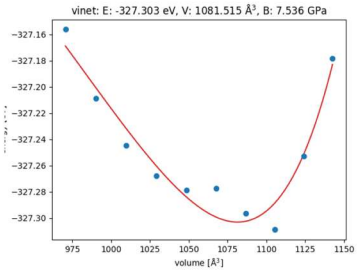
RETDAZ_full



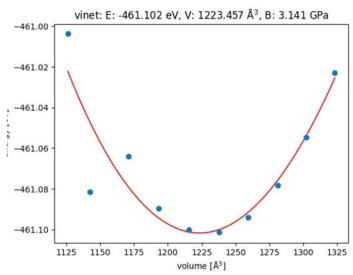
ROMFIM_P1



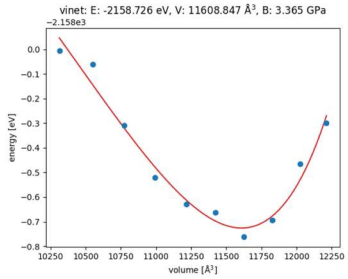
RUYGEZ_full



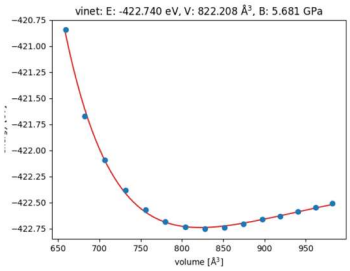
SEFKOG_full



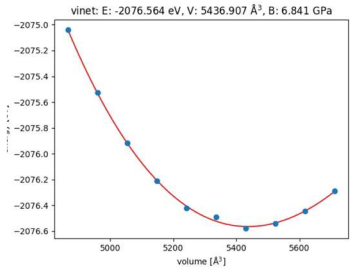
SEPNEJ_full



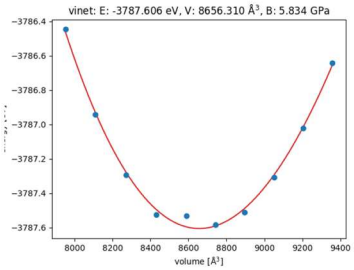
SETQEP_clean



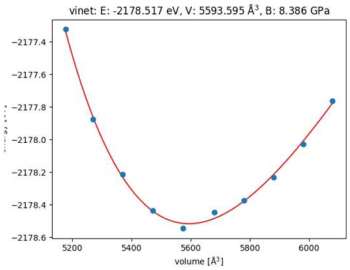
SIGNUU_full



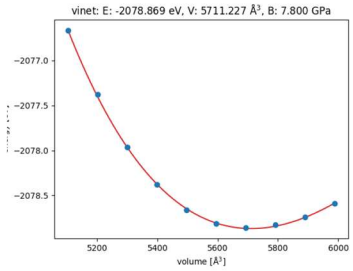
SODCAQ_P1



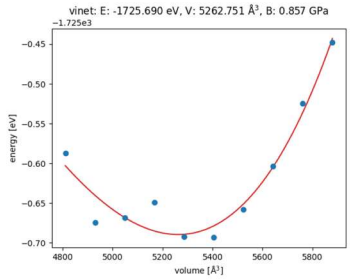
TENLOR_full



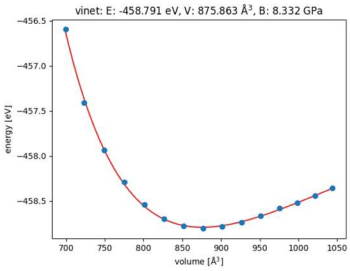
TENMAE_full



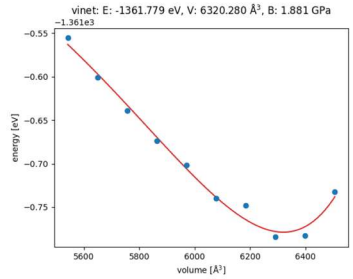
UQADEX_full



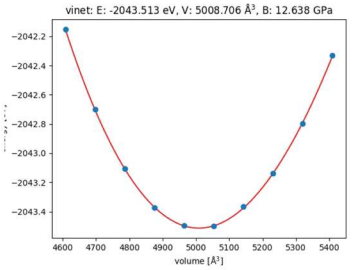
UVAPEN01_full



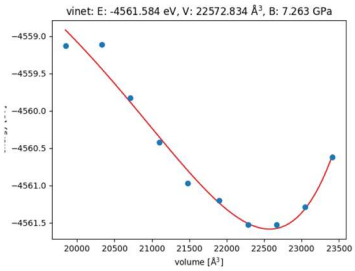
VEJYOZ_full



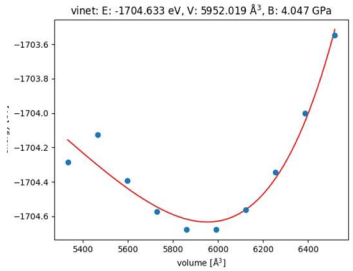
VIZRIH_full



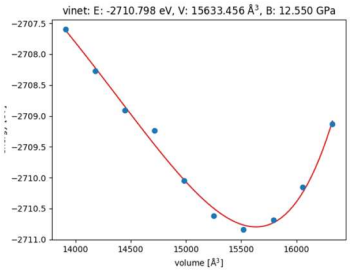
VOLQOD_full



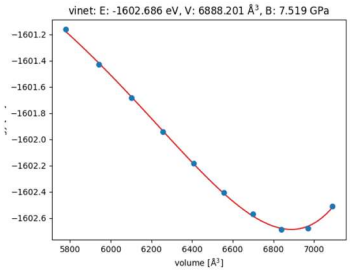
VUDTOH_P1



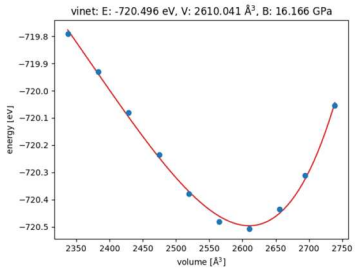
WAKHUO_full



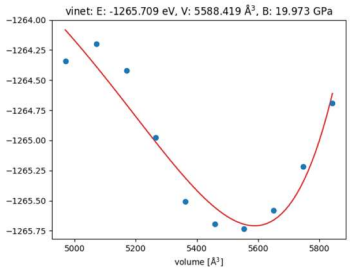
WORKUL_full



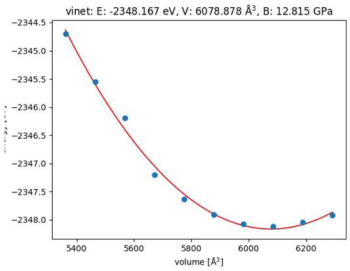
WOSQAZ_full



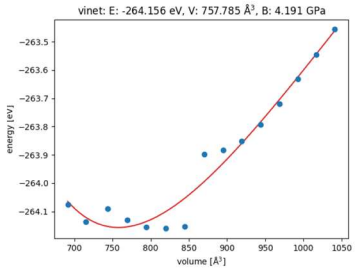
XADCOW_P1



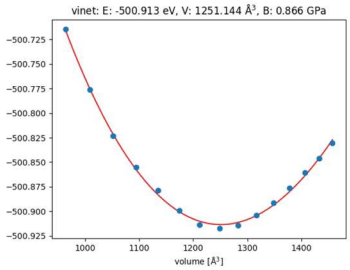
XEGXUF_full



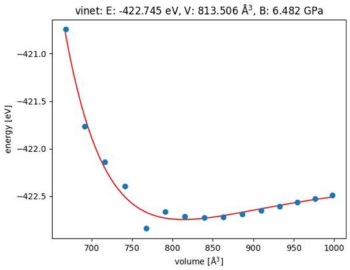
XOHLIS01_full



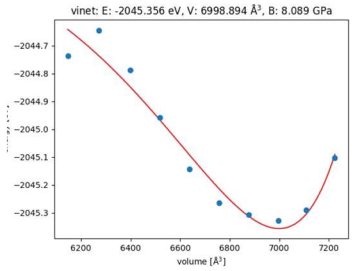
XOWJEB_full



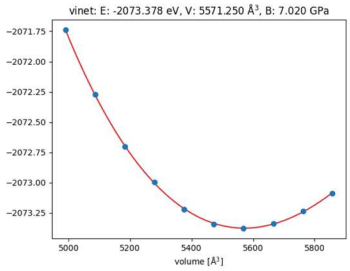
YAXBOP_SL



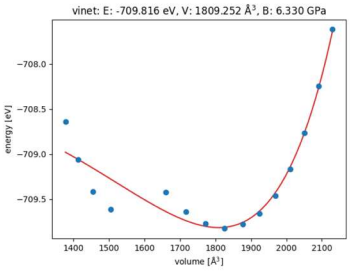
ZAVBUX_full



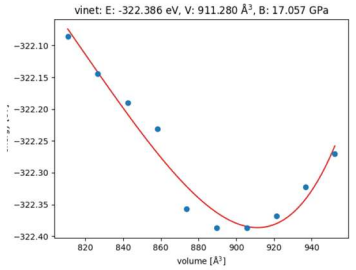
ZETLER_full



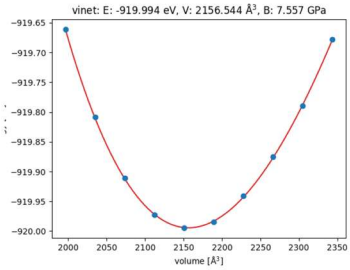
qmof-05dc5b6



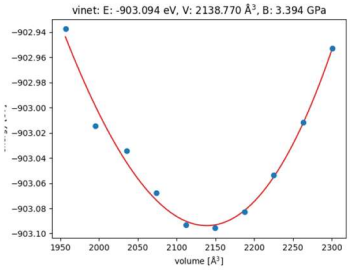
qmof-1985e04



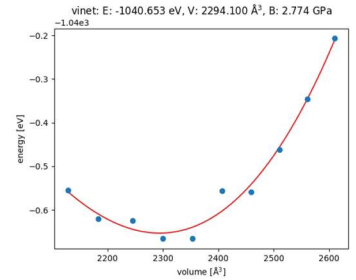
qmof-2f0e952



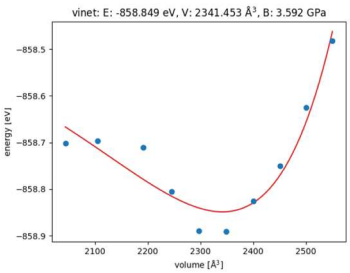
qmof-345d728



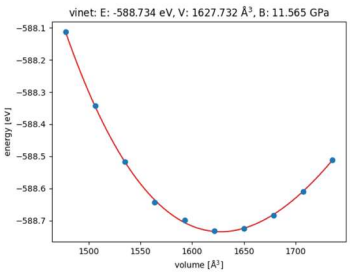
qmof-39fec9e



qmof-41f5a4b



qmof-53872ef



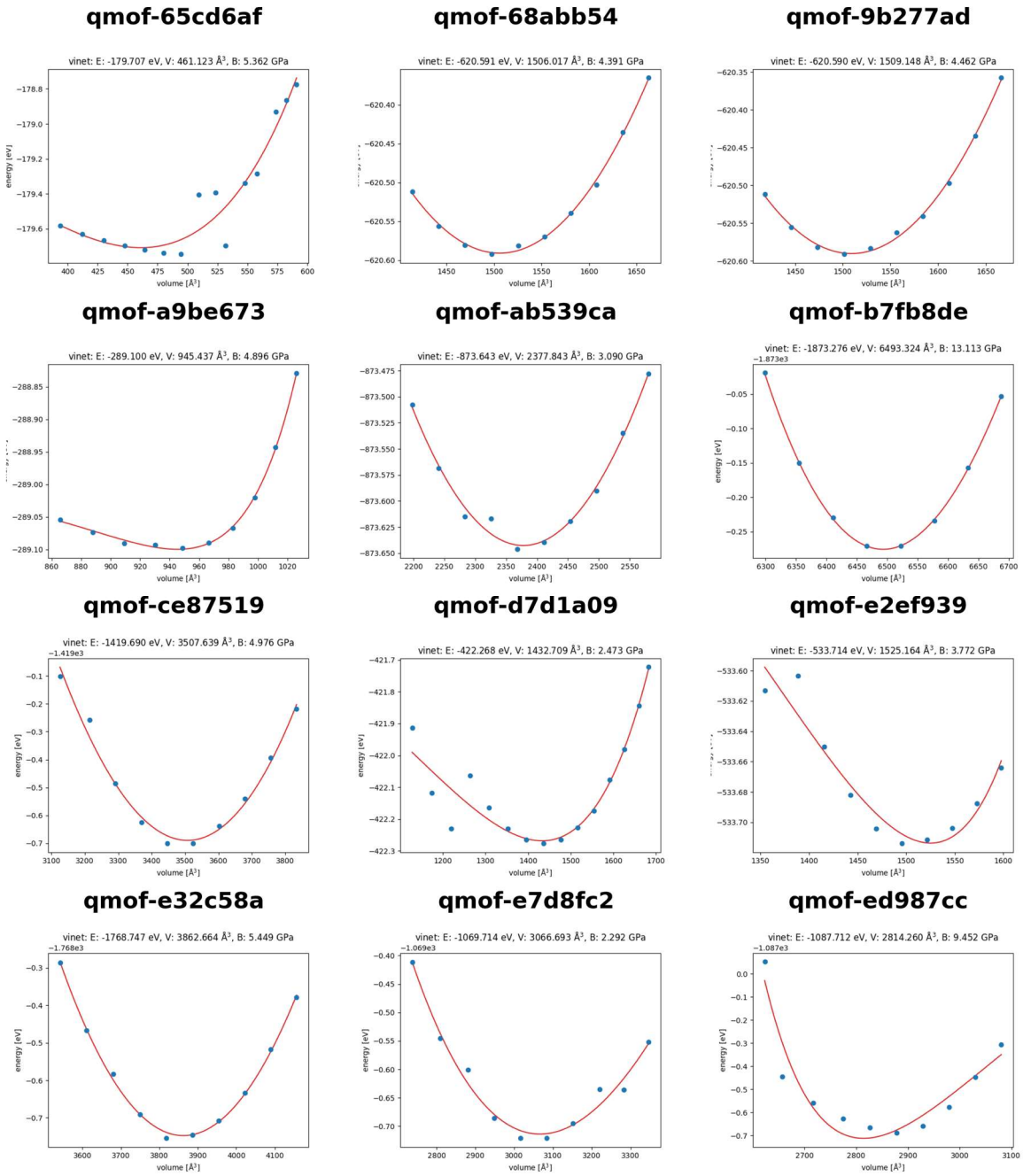


Figure S8. Energy-volume (E-V) curves utilized for the calculation of bulk moduli for the shortlisted MOFs. The blue markers represent the calculated data points, while the solid red lines indicate the corresponding fits using the Vinet equation of state (EOS). Each subplot is labeled with the respective MOF reference and the resulting bulk modulus (B) in GPa.

References

1 M. Fischer, F. Hoffmann and M. Fröba, *ChemPhysChem*, 2010, **11**, 2220–2229.

2 J. Peng, Y. Sun, Y. Wu, Z. Lv and Z. Li, *Ind. Eng. Chem. Res.*, 2019, **58**, 8290–8295.

3 R. P. P. L. Ribeiro, B. C. R. Camacho, A. Lyubchyk, I. A. A. C. Esteves, F. J. A. L. Cruz and J. P. B. Mota, *Microporous and Mesoporous Materials*, 2016, **230**, 154–165.

4 S. Guo, H.-Z. Li, Z.-W. Wang, Z.-Y. Zhu, S.-H. Zhang, F. Wang and J. Zhang, *Inorg. Chem. Front.*, 2022, **9**, 2011–2015.

5 Y. Zhang, S. Deng, X. Cui, J. Yue, H. Huang, C. Xue, H. Yang and L. Gan, *Chem. Sci.*, 2025, **16**, 19381–19388.