



Cite this: DOI: 10.1039/d5cp04871k

Computational modelling workflows for metal–organic polyhedra in The World Avatar

 Patrick W. V. Butler,^{id a} Arravind Subramanian,^b Simon D. Rihm,^{ac} Ari F. Fischer,^{bd} Tej S. Choksi,^{id bd} Aleksandar Kondinski,^e Sebastian Mosbach,^{ab} Jethro Akroyd^{id ab} and Markus Kraft^{id *abf}

Metal–organic polyhedra (MOPs) have exceptional potential for host–guest chemistry, but their discovery is hindered by the large combinatorial space of their building units. Computational screening offers a powerful workflow for efficiently screening large sets of MOPs; however, reliable results depend on accurate modelling of the geometries and properties of the MOPs. In this paper, we extend our previous workflow, which assembled computation-ready MOPs using purely geometric operations, by incorporating post-assembly computational modelling. We benchmark a range of methods, including machine learning interatomic potentials (MLIPs) and tight-binding DFT, against 85 experimentally resolved MOP structures. The results show that geometry optimisation significantly refines the initial assembled MOP structures in terms of cavity and pore properties when compared against experimental structures, with MLIPs found to achieve good accuracy and excellent convergence rates. We applied the most reliable method to the entire dataset, integrating the resulting data within The World Avatar through the OntoMOPs ontology and enabling natural language querying. Lastly, we demonstrate the utility of this refined dataset by screening for MOPs with the potential to act as hosts for a urea guest molecule.

 Received 15th December 2025,
 Accepted 2nd May 2026

DOI: 10.1039/d5cp04871k

rsc.li/pccp

1 Introduction

Metal–organic polyhedra (MOPs) have gained considerable attention as a versatile class of cage-like nanostructures featuring well-defined cavities, yielding significant potential in applications ranging from gas separation and storage to catalysis and sensing.^{1–7} These properties of MOPs are comparable to those of extended framework materials such as metal–organic frameworks (MOFs), but the discrete nature of MOPs gives advantages in solution-based preparation and processing.^{8,9} MOPs are typically constructed from a pair of metal and organic chemical building units (CBUs) that self-assemble following specific topological patterns, also called assembly models, to

produce discrete polyhedral architectures. Despite the considerable potential, developing materials based on MOPs is a significant challenge due in large part to the immense combinatorial space of the constituent metal-based and organic-based chemical building units (CBUs), which self-assemble to form MOPs.

Efficiently exploring this space is crucial for identifying promising candidates for applications. Yet, MOP discovery has historically relied on labour-intensive experimental trial-and-error approaches, which severely limit the number of structures that can be assessed. Computational strategies offer a more scalable alternative based on rapidly screening large numbers of hypothetical structures before prioritising a tractable subset for experimental validation.^{10–12} The recent development of machine learning interatomic potentials (MLIPs), which can accurately and efficiently model systems with hundreds to thousands of atoms, has prompted considerable interest in large-scale computational screening of materials.^{13–16} This has become especially prominent with the progress of universal MLIPs, which are pre-trained on large, comprehensive datasets and can then be used in various modelling workflows without further training.^{17–23} However, standard computational workflows for estimating MOP properties relevant for applications are still not clearly defined and the reliability of current methods remains unclear. Moreover,

^a Department of Chemical Engineering and Biotechnology, University of Cambridge, Philippa Fawcett Drive, Cambridge, CB3 0AS, UK. E-mail: mk306@cam.ac.uk

^b CARES, Cambridge Centre for Advanced Research and Education in Singapore, 1 Create, Way, CREATE Tower, #05-05, 138602, Singapore

^c CMFG, GRIPS – Grunderinnenzentrum Pirmasens, Delaware Avenue 1–3, 66953, Pirmasens, Germany

^d School of Chemistry, Chemical Engineering and Biotechnology, Nanyang Technological University, 62 Nanyang Drive, 637459, Singapore

^e Institute of Physical and Theoretical Chemistry, Graz University of Technology, Stremayrgasse 9, Graz, 8010, Austria

^f MIT, Chemical Engineering, 77 Massachusetts Avenue, Room E17-504, Cambridge, MA, 02139, USA



modelling results of MOPs are generally published as independent datasets, without direct connection to the initial data sources or simulation parameters. Recently, we developed a workflow for assembling computation-ready MOPs from known CBUs and assembly models.^{24,25} The results demonstrated that the purely geometric assembly process yielded structures with cavity and pore sizes in good agreement with the experimentally reported values.²⁶ Moreover, through the development of the OntoMOPs ontology, the data were instantiated into The World Avatar,²⁷ a platform for interconnected data and knowledge models, allowing for advanced semantic reasoning based on the relationships between CBUs, MOPs, assembly models, and properties.^{28,29} While the study provided a diverse set of computation-ready structures and an overview of basic geometric properties, the data on application-specific properties and the workflows for evaluating the MOPs for applications were limited.

The purpose of this paper is thus to identify a reliable computational workflow for refining the MOP geometries and estimating properties relevant for downstream screening for applications. We demonstrate geometry optimisations of the assembled MOP structures using various methods, including tight-binding density functional theory (DFT) and machine learning interatomic potentials (MLIPs). By comparing to a set of experimental structures, we identify methods that efficiently yield significant, quantified refinement of the MOP geometries. Subsequently, we scale the post-assembly modelling to the complete dataset of 1435 MOPs, with the results fully ontologised and integrated into The World Avatar, enabling a semantic description of the dataset and natural language querying.

2 Methods

In this work, we extend our geometric MOP assembly workflow by including post-assembly optimisation methods. In total, we consider a dataset of 85 experimentally characterised MOPs, used for benchmarking the computational methods, and 1435 hypothetical MOPs. The construction of the dataset and the details of the methods used are described in the following.

2.1 Dataset

The computation-ready MOPs were extracted from our previously published dataset.²⁴ These structures were generated by assembling combinations of one metal CBU and one organic CBU into MOPs according to defined assembly models. The ontological description of the assembly models in OntoMOPs specifies a set of generic building units (GBUs), their assembly positions, and their connecting points. Only pairs of CBUs that satisfy the CBU requirements are assembled, with our MOP Assembler code calculating the required transformations for each molecule to align the CBU centre with the CBU centre and the CBU binding sites with the connecting points. Once correctly oriented, the assembler calculates a scaling factor based on the size of the CBUs to expand or contract the distances

from the centroid, preventing collisions and ensuring the binding sites of neighbouring CBUs form the appropriate bonds. The set of CBUs and assembly models was curated from experimentally known structures, thereby revealing the immediate chemical space accessible. The experimental structures used to create this dataset were selected based on the definition of MOPs as coordination-driven assemblies of building units analogous to those of MOFs (*i.e.* metal secondary building units and organic linkers). Due to this definition, the dataset does not include dynamically self-assembled supramolecular cages, such as those formed by metal-catalysed imine condensation. Other than being experimentally characterised, there were no restrictions on permanent porosity or other properties. Subsequent studies have added additional assembly models and CBUs, and simplified access to the OntoMOPs knowledge graph with natural language querying and an object graph mapper that functions as a Python interface.^{25,26,28}

2.2 Computational methods

Calculations of the MOPs were conducted using a variety of methods, including universal MLIPs, tight-binding DFT, and Kohn–Sham DFT. Geometry optimisations using MLIPs were performed using the corresponding Atomic Simulation Environment (ASE) calculators and the FIRE optimisation method.^{30,31} The maximum number of steps was set to 5000 and the force convergence criterion was set to 0.01 eV Å⁻¹. Further details of the MACE,¹⁷ Orb,²¹ and UMA²² universal MLIPs used can be found in the associated publications. For the MACE models, we used MACE-MPA-0, MACE-OMAT-0, and MACE-MATPES-r2SCAN-0. The orb model used is the orb_v3_-conservative_inf_omat model featuring conservative forces and infinite neighbour atom lists. Lastly, the UMA models used are of the small 1p1 checkpoint version with the task type as either OMAT (Open Materials dataset)^{32,33} or ODAC (Open Direct-Air-Capture dataset).³⁴ In all cases, models that were trained on datasets that did not explicitly include dispersion contributions were augmented by summing the calculator output with the torch-dftd calculator, which implements the D3(BJ) dispersion correction method.^{35–37} The models combined with the D3 dispersion correction were MACE-MPA-0, MACE-OMAT-0, MACE-MATPES-r2SCAN-0, Orb-V3-con-inf-omat, and UMA-s-1p1(OMAT).

The tight-binding DFT calculations were performed using the GFN2-xTB method implemented in the xTB software package (version 6.5.0).³⁸ The standard ANCOpt optimiser and default settings were used except that the maximum number of SCF iterations was increased to 750. For these calculations and those using B97-3c below, the total charge of the MOPs was set according to the values extracted from the OntoMOPs knowledge graph (28 out of 85 MOPs have non-zero overall charge). All methods were applied consistently for charged and non-charged MOPs. Calculations that were performed using the GFN2-xTB(DMF) method additionally included implicit solvation using the ALPB method with the DMF parameters.

The DFT calculations at the B97-3c level were performed using the ORCA software package (version 4.2.1).^{39,40} The particulars of



the basis set and functional used in the B97-3c method are detailed in the original report.⁴¹ In all cases, the convergence of the single point calculations was determined by an energy tolerance of 1.0×10^{-6} .

3 Results and discussion

In the following sections, we describe the results from the evaluation of the various computational methods and applying them at scale. The first section details the evaluation benchmark, comparing the errors in optimised geometric properties against those from experimental structures. As a baseline, we used the initial MOP Assembler errors. After this, we scale the geometry optimisations to the hypothetical structures extracted from the OntoMOPs dataset and assess the influence of geometry optimisations on the distributions of MOP properties. We further demonstrate exploiting the optimised structural data as input to advanced modelling workflows and for identifying potential hosts for a urea molecule. Lastly, we illustrate exploration of the ontologised calculation data using Marie, a natural language agent interfaced with the MOPs knowledge graph.

3.1 Benchmarking MOP geometry optimisations

To evaluate the performance of our computational workflow and its ability to reproduce experimentally observed geometries, we curated a benchmark set of 85 MOP structures extracted from the previous OntoMOPs dataset with reported single-crystal X-ray diffraction (SCXRD) structures. This set covers a diverse range of MOPs comprising a distinct combination of one metal CBU, one organic CBU, and an assembly

model. In total, 29 metal CBUs, 47 organic CBUs, and 13 assembly models are represented in the set, with the most common metal CBU being 4-planar Cu_2 , which occurs in 30 MOPs, and the most common organic CBU being trimesic acid (benzene-1,3,5-tricarboxylic acid), which occurs in 13 MOPs. The distribution of the assembly models is shown in Fig. 1, with $(4\text{-planar})_{12}(2\text{-bent})_{24}$ being the most common with 35 structures. These experimental structures serve as a benchmark to assess the robustness and accuracy of the computational methods across the OntoMOPs dataset. Previously, we have shown that the MOP Assembler yields good agreement with experimental structures when applied to known MOPs.²⁶ Our aim here is thus to see whether subsequent geometry optimisation of the MOP Assembler structures improves this agreement.

We focused on three key geometric features of MOPs relevant to their host-guest functionality and thereby applications: the radius of the largest sphere that can fit in the MOP cavity (inner sphere radius), the average diameter of windows through which guests could enter or exit the MOP, and the mass-weighted radius of gyration, which is a measure of the atom distribution. These were calculated for both the experimental structures and the corresponding optimized structures, with the window diameters being calculated using *pywindow*.⁴² We then quantified the accuracy of the computed geometries by comparing these properties against their experimental counterparts and calculating the percentage error. The results for each method and metric are summarised in Fig. 2.

The results show that in general the geometry optimisations yielded reduced errors in all metrics, with errors in some cases more than halved. For the inner sphere diameter, we find that the MOP Assembler structures achieve an excellent percentage MAE (%MAE) of 4.3%, which is improved by a few methods,

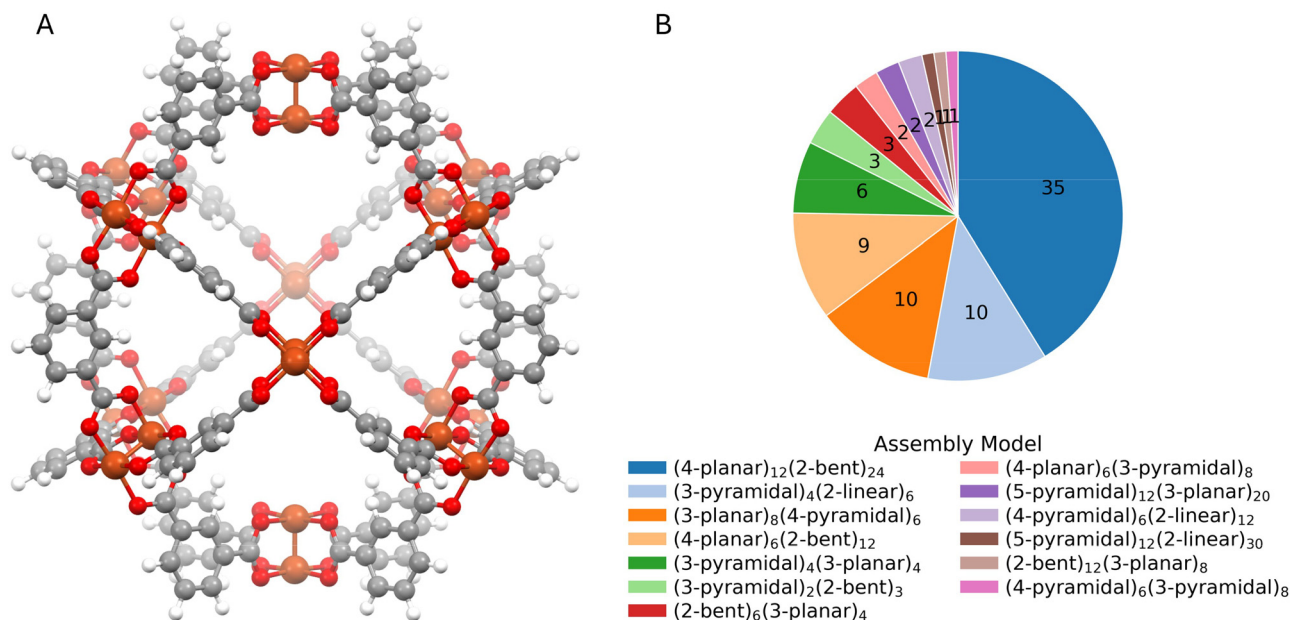


Fig. 1 Example MOP featuring the most common assembly model and metal CBU in the experimental benchmark, $(4\text{-planar})_{12}(2\text{-bent})_{24}$ and Cu_2 , respectively (A), and the number of MOP experimental structures for each assembly model (B).



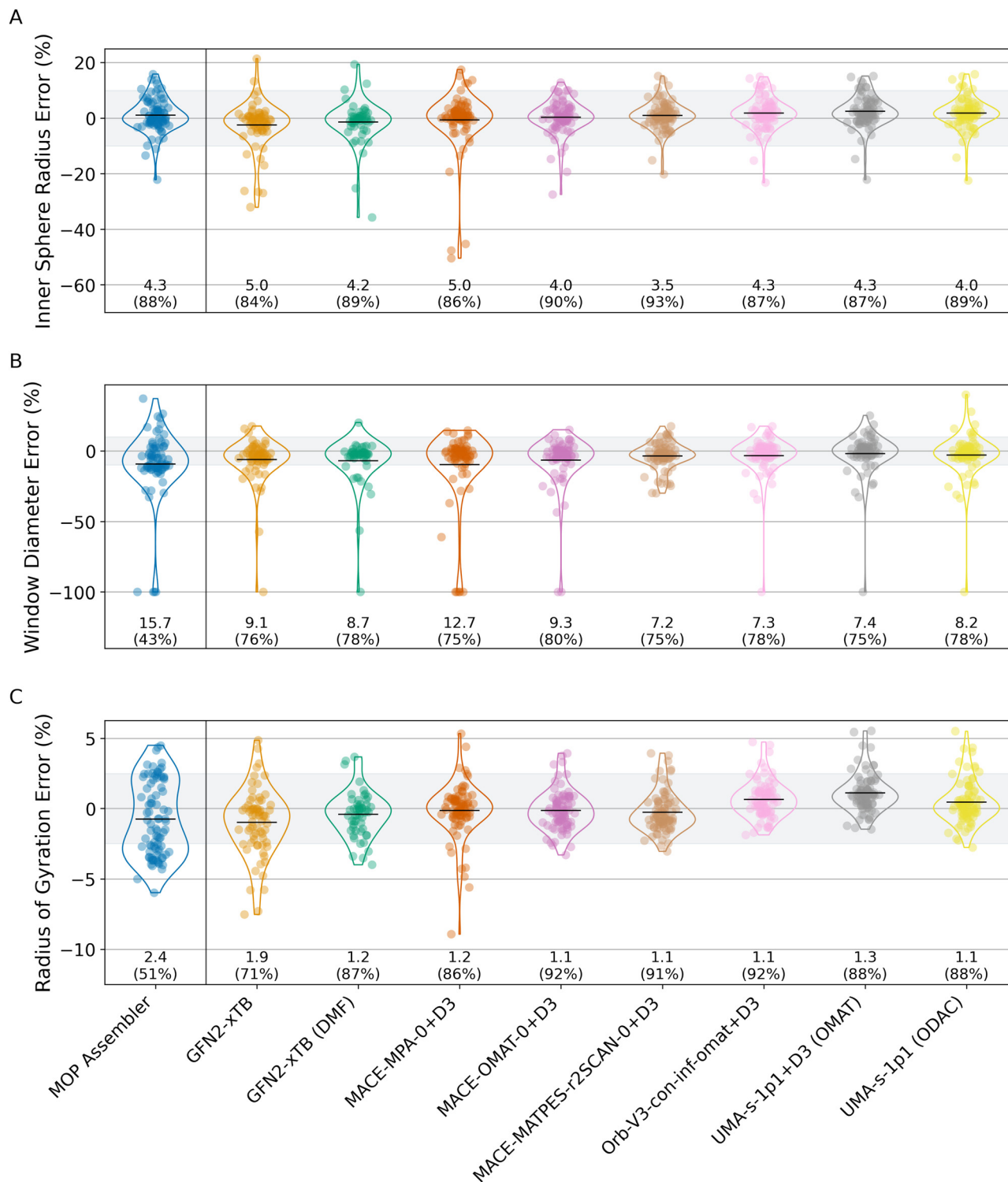


Fig. 2 Violin plots for each method illustrating the percent error in terms of the inner sphere radius (A), average window diameter (B), and mass-weighted radius of gyration (C) when comparing the MOP Assembler structure optimised with the corresponding method against the experimental values. The distributions include up to 85 structures with unconverged or failed optimisations omitted. The means of the distributions are indicated by the black bars with the percent MAEs underneath each distribution. The proportion of data points within the shaded region is displayed in brackets below the MAE.

namely MACE-MATPES-r2SCAN-0 + D3 (3.5%), UMA-s-1p1(ODAC) (4.0%), MACE-OMAT-0 + D3 (4.0%), and GFN2-xTB(DMF) (4.2%).

In absolute terms, this reduces the MAE in the inner sphere volumes from 221.1 Å³ with the MOP Assembler to 152.5 Å³



following geometry optimisation with MACE-MATPES-r2SCAN-0 + D3 (see Table S1). Despite the general improvement, we observe increased errors following optimisations for a small number of structures across multiple methods. Inspecting these results, we find that these typically occur for MOPs with larger, flexible organic CBUs, which when optimised in the gas phase contract inwards. For example, the MOP with the CSD refcode XIYJUM (Fig. S11), which is a tetrahedral MOP featuring a 3-planar 1,3,5-tris(4-carboxyphenyl)benzene CBU, has an average error of -18.6% following optimisation in the gas phase compared to -7.7% for the initial MOP Assembler structure. Notably, comparing the GFN2-xTB results, the error is considerably reduced when using implicit solvation (-8.24% vs. -32.07%). However, we do not find a general trend of error with CBU size (Fig. S5).

In the other two metrics, we observe larger improvements following geometry optimisations. For example, the percent error in the average window diameter is nearly halved with all methods compared to the MOP Assembler. Again the MACE-MATPES-r2SCAN-0 + D3 model performs best, yielding a %MAE of 7.2% (MAE = 0.29 \AA). However, we note that there are multiple occurrences of the optimisations leading to an elimination of all windows and a -100% percent error. Upon inspection, we find that these are structures containing the compact trimesic acid CBU, which typically yield window diameters that are already at the limits of detectability. Indeed, by excluding structures with an average window diameter less than 1.8 \AA , all but one method (MACE-MPA-0 + D3) no longer result in such large errors (see Fig. S1) and with this filter the best performing method is GFN2-xTB(DMF), which returns a %MAE of 4.7% (MAE = 0.23 \AA).

We similarly observe large improvements in the radius of gyration error following geometry optimisation with most methods. In this case, the best performing methods are MACE-OMAT-0 + D3, MACE-MATPES-r2SCAN-0 + D3, Orb-v3-con-inf-omat + D3, and UMA-s1p1 (ODAC), which yield %MAEs of 1.0% (MAE = $0.117\text{--}0.123 \text{ \AA}$). GFN2-xTB(DMF), MACE-MPA-0 + D3, and UMA-s1p1 + D3 (OMAT) also return low %MAEs between 1.2 and 1.3% (MAE = $0.118\text{--}0.140 \text{ \AA}$). The largest errors are found following optimisation with GFN2-xTB without implicit solvation (%MAE = 1.9 , MAE = 0.196 \AA), further indicating the importance of solvation effects, which can be implicitly captured in the parameterisation of the MLIPs due to inclusion of experimentally reported structures in the training data.

An additional analysis comparing the error distributions of charged and uncharged MOPs finds that the charged MOPs generally have higher errors (Fig. S6–S8). However, this trend is observed across all methods, including GFN2-xTB methods and the MOP assembler, suggesting that it is not introduced by the optimisations or the modelling of the charges but rather a result of the structural complexity within the charged subset, which features many of the larger polynuclear metal building units, including $[\text{Zr}_3\text{O}(\text{OH})_3(\text{C}_5\text{H}_5)_3]$, $[\text{V}_6\text{O}_6(\text{OCH}_3)_9(\text{SO}_4)]$, and $[\text{WV}_5\text{O}_{11}]$. By contrast, the uncharged subset predominantly features bimetallic paddle-wheel metal units.

A further important consideration is the convergence of the calculations, which is presented in Fig. 3. This shows that all

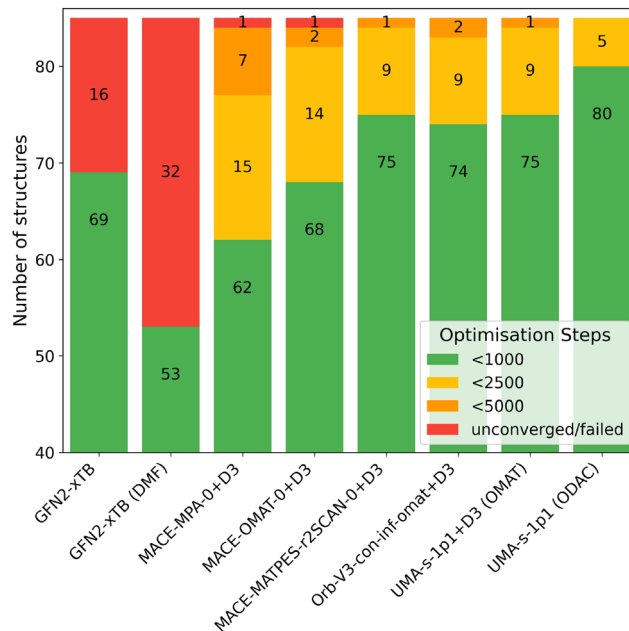


Fig. 3 Summary of convergence statistics for the geometry optimisations using each method. The GFN2-xTB optimisations were performed using the xTB program, while the ML potential simulations were performed in ASE using the FIRE optimiser and a max step limit of 5000.

methods achieve good success rates, but with the MLIPs typically achieving excellent results. Across the methods, the UMA models show the highest success rates, with the UMA-s1p1(ODAC) model successfully converging all 85 structures. By contrast, the GFN2-xTB(DMF) method failed to calculate an optimised structure for 32 MOPs, typically due to numerical instabilities. Interestingly, there were half as many failed optimisations without implicit solvation (*i.e.* GFN2-xTB). Analysing the individual results, the most difficult structure to optimise was the MOP with the CSD refcode GECQEN (Fig. S12), which was failed by 2 methods (GFN2-xTB and MACE-OMAT-0 + D3) and took more than 2500 steps to converge for three other methods. As could be expected, this is one of the largest MOPs in the set having a $(4\text{-planar})_{12}(2\text{-bent})_{24}$ assembly model and containing 960 atoms with many of these atoms in flexible side-chain groups.

Overall, the results of the benchmark show that we can significantly improve the agreement to experimental values by fast geometry optimisations using tight-binding DFT or MLIPs. This comparison has limitations. Most significantly, the experimental values are derived from crystal structures determined by X-ray diffraction at varying finite temperatures and states of solvation, whereas the calculations were performed at 0 K on isolated MOPs in the gas phase or with implicit solvation in the case of GFN2-xTB(DMF). Consequently, the target structural features exhibit experimental variations, arising from finite temperature effects and explicit solvent interactions, which are not accounted for by the optimisations and create uncertainty in the target values. Moreover, crystal packing effects are not included in the isolated MOP calculations. Therefore, differences



of a few percent between computational methods should be interpreted cautiously. However, the consistency with which the optimised structures converge towards experimental values across multiple independent methods, and the clear improvement over the unoptimised MOP Assembler structures, provides confidence that the optimisations provide benefit in improving estimates of observed properties. Additionally, the aim of these optimisations is to serve an initial refinement stage that can provide input for subsequent, more complex modelling that accounts for the neglected effects. For example, finite-temperature and solvent effects could be included by using the geometry optimised structures as input to molecular dynamics simulations, yielding average and standard deviation values for the geometric properties. This approach has been shown to be effective for modelling similar cage molecules.⁴³

To estimate the influence of the neglected experimental effects on the observed errors, we repeated the analysis by comparing against the isolated experimental structures optimised using the same methods (Fig. S9). These results show that the %MAE values are reduced across all methods and properties: inner sphere radius by 0.8 to 3.1%, window diameter by 1.3 to 4.2%, and radius of gyration by 0.2 to 1.6%. However, the general trends are largely the same as those observed when comparing against the experimental data, with GFN2-xTB(DMF) and MACE-MATPES-r2SCAN-0 + D3 consistently returning the lowest errors.

A further limitation of modelling isolated MOPs is that functional properties arising from crystal packing, namely extrinsic porosity, cannot be estimated, yet are crucial for solid-state applications of MOPs. Although the current study focused on cavity properties relevant for screening applications in solution, such as drug delivery and catalysis, the methods applied in this study could equally be applied to periodic structures, and this represents a natural extension of the workflow: using the optimised isolated MOPs as input to generate plausible crystal packings. Accurately predicting the crystal structures of MOPs is not a trivial problem; however, there have been promising developments in the crystal structure prediction of similar supramolecular assemblies.⁴⁴

3.2 MOP geometry optimisations at scale

To assess the scalability of the post-assembly calculations, we next created a workflow extracting a large set of hypothetical MOPs from the OntoMOPs knowledge graph and performing a geometry optimisation on each. To perform these optimisations, we selected the MACE-MATPES-r2SCAN-0 + D3 method, which was found to be one of the most reliable and accurate methods in the previous section. We then compared the inner sphere radius, average window diameter, and radius of gyration before and after optimisation along with the calculated root-mean-square deviations (RMSDs) between the initial and final atomic coordinates.

As shown in Fig. 4a, following optimisation, the inner sphere radius had a mean change of -0.55% with a standard deviation of 5.11% . This is consistent with the previous results that showed the MOP Assembler structures typically already

have realistic cavity sizes, and hence the change is relatively small. However, there is a noticeable tendency for the optimisations to decrease the cavity size and there are examples of up to 39% decrease in the inner sphere radius. Upon inspection of these results, we find that this is a result of the larger organic CBUs contracting inwards under the gas-phase geometry optimisation with the larger percent changes being observed for structures with small initial cavities (less than 2.5 \AA inner sphere radius).

Considering the average window diameter, the optimisations returned an average change of $+1.24\%$. However, there is considerable variation with a notable standard deviation of 15.06% and several structures having a change of greater than 50% . Similar to previous results, we find that these are generally structures with small initial window diameters.

The radius of gyration is also stable on average with a mean change between the initial MOP assembler structure and geometry optimised structure of $+0.70\%$ with a relatively narrow standard deviation of 2.32% . The stability in the radius of gyration, despite larger variations in window metrics, indicates that the optimisations primarily alter local features rather than the overall MOP shape, which remains consistent. This is even true for the structure with the maximum change (-9.84%), which, when visualised, clearly preserves the $(4\text{-planar})_6(3\text{-pyramidal})_8$ assembly model (Fig. S13).

Finally, RMSD values between the atom coordinates of pre- and post-optimised structures averaged 0.55 \AA with a standard deviation of 0.33 \AA . The small average RMSD further supports that the geometry optimisations are refining the MOP Assembler structures. However, there is a notably long tail on the RMSD distribution with a significant number of structures having RMSD values greater than 1.0 \AA and up to 2.36 \AA . Visualising these structures, it is apparent that the high RMSDs result from flexible side chain groups that often undergo significant rearrangement, even though the core topology remains stable (Fig. S14). This is not unexpected since the MOP Assembler uses CBU geometries extracted from crystal structures, which are stabilised by intermolecular interactions that are not present during the geometry optimisation of the isolated MOP.

3.3 Advanced computational workflows

The results from the previous section indicate that MACE-MATPES-r2SCAN-0 + D3 can reliably optimise the initial MOP Assembler structures and, thus, form the basis of more elaborate computational workflows that extend beyond simple structural relaxation. For example, although the MLIP models do not provide explicit information on electronic properties, the relaxed geometries they yield can be used as input for subsequent single-point electronic structure calculations. Importantly, this approach allows accessing electronic properties without the need to perform full geometry optimisations using computationally intense electronic structure methods, which is often a major computational bottleneck.

To illustrate this approach, we extracted the MACE-MATPES-r2SCAN-0 + D3 optimised structures of MOPs featuring the



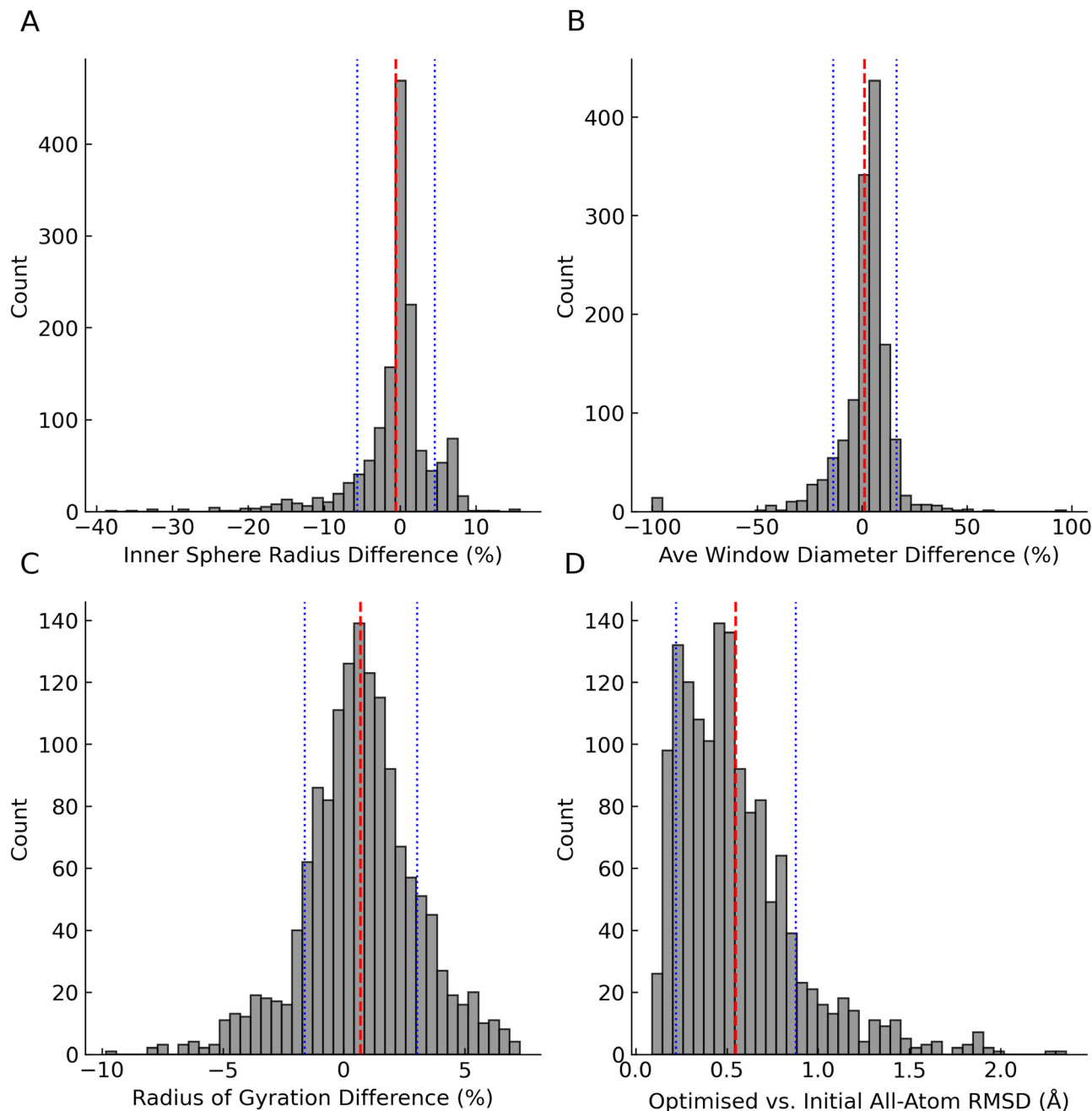


Fig. 4 The results from comparing geometry optimisations of the 1435 MOP structures from the OntoMOPs dataset using MACE-MATPES-r2SCAN-0 + D3 against the initial MOP Assembler structures in terms of the change in the inner sphere radius (A), average window diameter (B), radius of gyration (C), and the RMSD in atomic positions (D). The mean of each distribution (red dashed line) and one standard deviation (blue dotted line) are indicated.

[Zr₃O(OH)₃(C₅H₅)₃] metal building unit and used these as input for single-point calculations at the DFT level using the B97-3c method. From these calculations, we could investigate the distribution of electronic properties among these MOPs, such as the HOMO–LUMO gap energies, which is shown in Fig. 5. Crucially, the computational cost of gaining this insight was only on average 8.11 CPU hours and could be completed on average in 10.6 wall-clock minutes with 76 CPU cores. A set of six structures were able to be geometry-optimised at the DFT level and a comparison of the gap energies (Fig. S10) shows a

low correlation ($R^2 = 0.34$) but good agreement in ranking ($r_s = 0.89$). Although it is important to emphasise that this is not a benchmarking of this approach or combination of methods, this demonstration highlights the potential to efficiently screen for electronic properties in the dataset using the generally good structures from the lower level geometry optimisations.

3.4 Host–guest screening

Having demonstrated the efficiency of our computational approach to yield good approximations of experimental MOPs



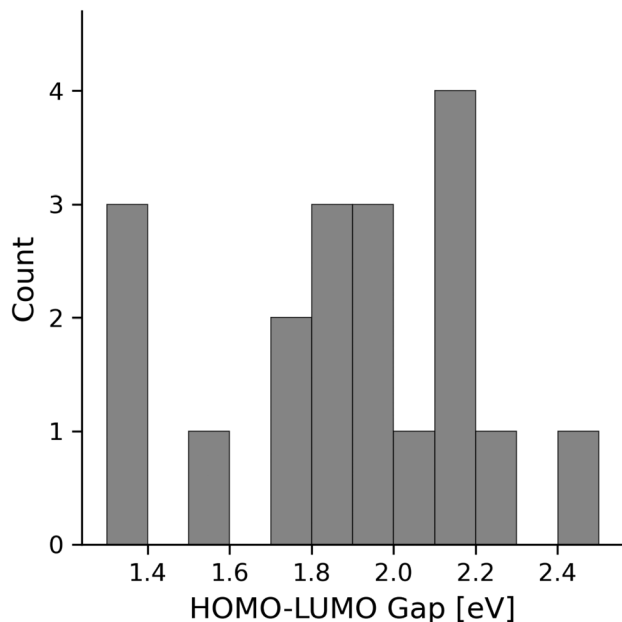


Fig. 5 Distribution of HOMO–LUMO gap energies obtained from B97-3c single-point calculations performed on MOPs containing the $[\text{Zr}_3\text{O}(\text{OH})_3(\text{C}_5\text{H}_5)_3]$ metal building unit and following geometry optimisation using the MACE-MATPES-r2SCAN-0 + D3 model.

and to scale to large datasets, we next applied these methods to investigate how the dataset can be exploited to identify promising materials for host–guest applications, and as an example target guest, we selected urea, which is relevant to biology and environmental monitoring. To begin, we extracted optimised MOPs with inner sphere radii that are 1.0 to 1.5 times the bounding sphere radii of urea (3.24 Å) and an average window diameter greater than 1.0 Å, yielding 52 MOPs. We then placed the urea molecule into each cage, varying the orientation to identify the most stable configuration using MACE-MATPES-r2SCAN-0 + D3 as the energy model. The geometries of the resulting configurations were then optimised to produce the

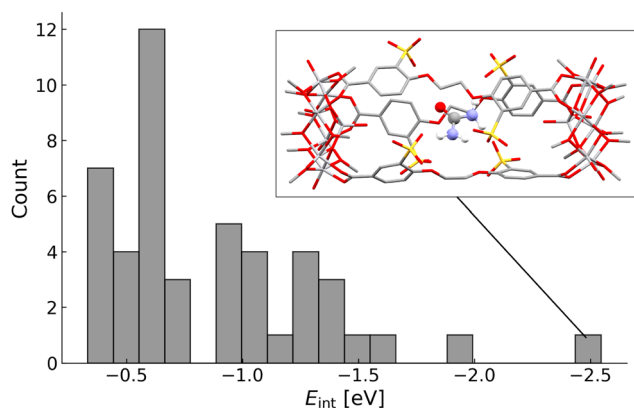


Fig. 6 Distribution of interaction energies for urea with the compatible MOPs calculated using MACE-MATPES-r2SCAN + D3 following geometry optimisation of the host–guest complex. The inset shows the most stable complex, which involves a (3-pyramidal) $_2$ (2-linear) $_3$ assembly model combining a vanadium metal CBU and a sulfate functionalised organic CBU.

energy minimised host–guest system and the final interaction energy was calculated. The distribution of interaction energies (E_{int}) is shown in Fig. 6. From this, we find a number of MOPs with highly stable interaction energies with the most stable MOP host featuring sulfate groups that form potentially three hydrogen bonds to the urea guest ($\text{N-H}\cdots\text{O} = 154.2, 165.7, 169.5^\circ$). The second most stable MOP host features the same assembly model and organic CBU but with a different metal CBU, thereby strongly suggesting that this combination of complementary cavity and hydrogen bonding sites is a promising design. It should be noted that the reported interaction energies are an approximation and neglect important temperature and solvent effects that are influential in host–guest behaviour. Nevertheless, the results demonstrate clearly how the integration of the knowledge graph and modelling results enables the dataset to be rapidly screened to find MOPs with potential for binding target substrates. Subsequent simulations,

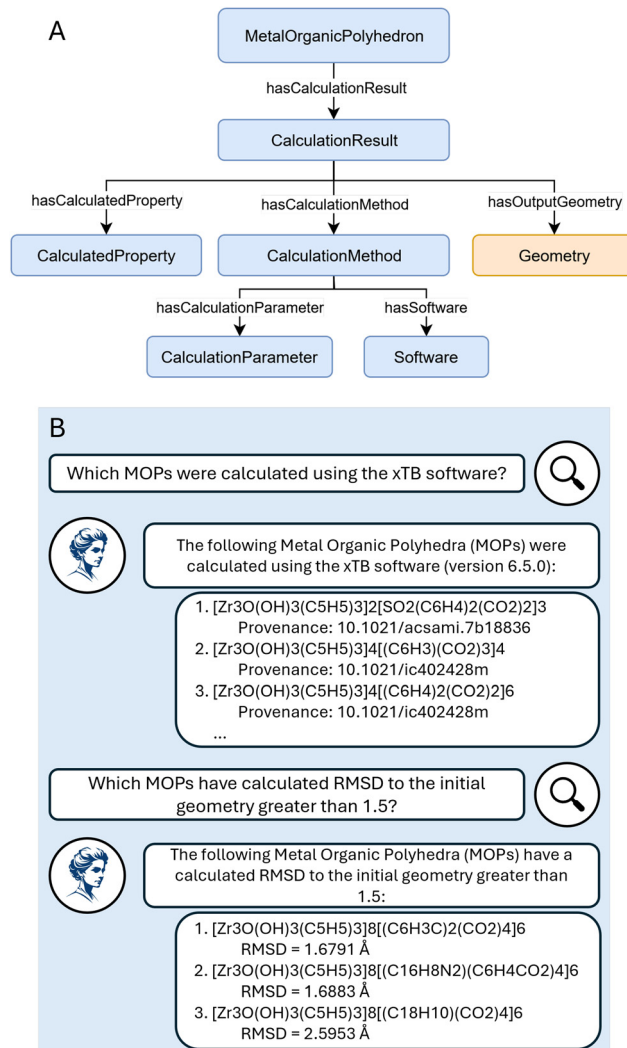


Fig. 7 The relevant concepts to incorporate calculation results into the OntoMOPs ontology (A) and examples of querying the MOP calculation results using Marie (B). In the ontology, blue concepts are part of OntoMOPs and orange concepts are part of OntoSpecies.⁴⁵



such as molecular dynamics, could be applied to more accurately estimate binding energies and rankings and characterise the response of the MOP to the guest. Considering the computational cost of modelling the large MOP structures, hierarchical workflows to efficiently narrow down the search space provide significant benefits in screening for applications.

3.5 Knowledge graph and natural language querying

To make the previous results more accessible, we extended the previous OntoMOPs ontology to allow for capturing calculation results. As shown in Fig. 7A, the extensions to the ontology consisted of adding a primary calculation result class along with classes for calculation methods, calculated properties, and calculation parameters. These classes are connected to the MOP concept by the `hasCalculationResult` predicate. These simple and general concepts were chosen to enable efficient uploading and querying of the data from the various computational methods used. Additionally, we have connected this data to Marie, the natural language agent of The World Avatar, allowing the calculation data to be queried using semantic reasoning without requiring programming knowledge. Example queries are shown in Fig. 7B. The data can further be accessed through our recently developed object-graph mapper Python package²⁵ interfaced with The World Avatar knowledge graphs. An example of this is provided as a Python notebook in the SI.

4 Conclusions

In this work, we have evaluated various computational workflows for refining MOP geometries assembled *in silico* and estimating their properties for downstream applications. Through benchmarking against experimental crystallographic data, we quantified the accuracy and reliability of various computational methods, with the results showing that universal MLIPs, such as the MACE and UMA models, offer an efficient route to achieve structural refinement. We further demonstrated the scalability of the post-assembly geometry optimisations by implementing a MLIP-based workflow and applying it to a dataset of 1435 hypothetical MOPs. Importantly, by fully ontologising these results within The World Avatar, we transformed the static calculation data into a dynamic, interoperable resource, which can be readily integrated into further simulations and screening tasks, as illustrated by the successful identification of candidate MOP hosts for urea. This work establishes a scalable, extensible workflow for reliably modelling hypothetical MOPs and, moreover, provides a clear example of how knowledge graphs and ontologies can be integrated with computational modelling to address issues with traceability and reproducibility in computational workflows.

Author contributions

P. W. V. B., A. S., S. D. R., A. F. F., T. S. C., A. K., and M. K. conceptualised the research and developed the methodology.

P. W. V. B., A. S., and A. F. F. performed the calculations. P. W. V. B., A. S., S. D. R., A. F. F. and A. K. analysed the results. P. W. V. B. prepared the figures and wrote the initial draft of the manuscript. T. S. C., S. M., J. A., and M. K. supervised the research. All authors contributed to reviewing and editing the final manuscript and approved the final submitted version.

Conflicts of interest

There are no conflicts to declare.

Data availability

The supplementary information (SI) contains additional figures illustrating the calculation results. See DOI: <https://doi.org/10.1039/d5cp04871k>.

The MOP Assembler software used is available on GitHub under the MIT license: <https://github.com/TheWorldAvatar/MOPTools>. The data supporting the results presented are available at Zenodo via DOI: <https://doi.org/10.5281/zenodo.17790355> along with an example Python notebook for interacting with the MOPs knowledge graph. Data from the calculations are also made available via Marie: <https://theworldavatar.io/demos/marie>.

Acknowledgements

This research was supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme. This project has received funding from the European Union's Horizon Europe research and innovation programme under grants 101058732 (JIDEP), 101074004 (C2IMPRESS), and 101188248 (CLIMATE-ADAPT4EOSC). This work was performed using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service (<https://www.csd3.cam.ac.uk>), provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/T022159/1), and DiRAC funding from the Science and Technology Facilities Council (<https://www.dirac.ac.uk>). M. K. gratefully acknowledges the support of the Alexander von Humboldt Foundation and the Massachusetts Institute of Technology. S. D. R. acknowledges financial support from Fitzwilliam College, Cambridge, and the Cambridge Trust. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

References

- 1 S. Lee, H. Jeong, D. Nam, M. S. Lah and W. Choe, *Chem. Soc. Rev.*, 2021, **50**, 528–555.
- 2 H. Vardhan, M. Yusubov and F. Verpoort, *Coord. Chem. Rev.*, 2016, **306**, 171–194.



- 3 A. J. Gosselin, C. A. Rowland and E. D. Bloch, *Chem. Rev.*, 2020, **120**, 8987–9014.
- 4 S. P. Argent, I. da Silva, A. Greenaway, M. Savage, J. Humby, A. J. Davies, H. Nowell, W. Lewis, P. Manuel, C. C. Tang, A. J. Blake, M. W. George, A. V. Markevich, E. Besley, S. Yang, N. R. Champness and M. Schröder, *Inorg. Chem.*, 2020, **59**, 15646–15658.
- 5 Y. Fang, J. A. Powell, E. Li, Q. Wang, Z. Perry, A. Kirchon, X. Yang, Z. Xiao, C. Zhu, L. Zhang, F. Huang and H.-C. Zhou, *Chem. Soc. Rev.*, 2019, **48**, 4707–4730.
- 6 A. O. Adeola, J. O. Ighalo, P. I. Kyesmen and P. N. Nomngongo, *J. CO₂ Util.*, 2024, **80**, 102664.
- 7 O. Calvo-Lozano, L. Hernández-López, L. Gomez, A. Carné-Sánchez, C. von Baeckmann, L. M. Lechuga and D. MasPOCH, *ACS Appl. Mater. Interfaces*, 2023, **15**, 39523–39529.
- 8 D. He, H. Ji, T. Liu, M. Yang, R. Clowes, M. A. Little, M. Liu and A. I. Cooper, *J. Am. Chem. Soc.*, 2024, **146**, 17438–17445.
- 9 B. Le Ouay, R. Minami, P. K. Boruah, R. Kunitomo, Y. Ohtsubo, K. Torikai, R. Ohtani, C. Sicard and M. Ohba, *J. Am. Chem. Soc.*, 2023, **145**, 11997–12006.
- 10 A. Tarzia and K. E. Jelfs, *Chem. Commun.*, 2022, **58**, 3717–3730.
- 11 Y. J. Colón and R. Q. Snurr, *Chem. Soc. Rev.*, 2014, **43**, 5735–5749.
- 12 C. E. Wilmer, M. Leaf, C. Y. Lee, O. K. Farha, B. G. Hauser, J. T. Hupp and R. Q. Snurr, *Nat. Chem.*, 2012, **4**, 83–89.
- 13 Y. Lim, H. Park, A. Walsh and J. Kim, *Matter*, 2025, **8**, 102203.
- 14 M. Hodapp and A. Shapeev, *Phys. Rev. Mater.*, 2021, **5**, 113802.
- 15 K. Choudhary and B. G. Sumpter, *AIP Adv.*, 2023, **13**, 095109.
- 16 A. M. Elena, P. D. Kamath, T. JaffreLOT Inizan, A. S. Rosen, F. Zanca and K. A. Persson, *npj Comput. Mater.*, 2025, **11**, 125.
- 17 I. Batatia, P. Benner, Y. Chiang, A. M. Elena, D. P. Kovács, J. Riebesell, X. R. Advincula, M. Asta, M. Avaylon, W. J. Baldwin, F. Berger, N. Bernstein, A. Bhowmik, F. Bigi, S. M. Blau, V. Cărare, M. Ceriotti, S. Chong, J. P. Darby, S. De, F. Della Pia, V. L. Deringer, R. Elijošius, Z. El-Machachi, E. Fako, F. Falcioni, A. C. Ferrari, J. L. A. Gardner, M. J. Gawkowski, A. Genreith-Schriever, J. George, R. E. A. Goodall, J. Grandel, C. P. Grey, P. Grigorev, S. Han, W. Handley, H. H. Heenen, K. Hermansson, C. H. Ho, S. Hofmann, C. Holm, J. Jaafar, K. S. Jakob, H. Jung, V. Kapil, A. D. Kaplan, N. Karimitari, J. R. Kermode, P. Kourtis, N. Kroupa, J. Kullgren, M. C. Kuner, D. Kuryla, G. Liepuoniute, C. Lin, J. T. Margraf, I.-B. Magdău, A. Michaelides, J. H. Moore, A. A. Naik, S. P. Niblett, S. W. Norwood, N. O'Neill, C. Ortner, K. A. Persson, K. Reuter, A. S. Rosen, L. A. M. Rosset, L. L. Schaaf, C. Schran, B. X. Shi, E. Sivonxay, T. K. Stenczel, C. Sutton, V. Svahn, T. D. Swinburne, J. Tilly, C. van der Oord, S. Vargas, E. Varga-Umbrich, T. Vegge, M. Vondrák, Y. Wang, W. C. Witt, T. Wolf, F. Zills and G. Csányi, *J. Chem. Phys.*, 2025, **163**, 184110.
- 18 I. Batatia, D. P. Kovács, G. N. C. Simm, C. Ortner and G. Csányi, MACE: Higher Order Equivariant Message Passing Neural Networks for Fast and Accurate Force Fields, in *Advances in Neural Information Processing Systems*, ed. A. H. Oh, A. Alekh, B. Danielle and C. Kyunghyun, 2022, <https://openreview.net/forum?id=YpPngE-ZU>.
- 19 D. P. Kovacs, I. Batatia, E. S. Arany and G. Csanyi, *J. Chem. Phys.*, 2023, **159**, 044118.
- 20 M. Neumann, J. Gin, B. Rhodes, S. Bennett, Z. Li, H. Choubisa, A. Hussey and J. Godwin, Orb: A Fast, Scalable Neural Network Potential, 2024, DOI: [10.48550/arXiv.2410.22570](https://doi.org/10.48550/arXiv.2410.22570).
- 21 B. Rhodes, S. Vandenhaute, V. Šimkus, J. Gin, J. Godwin, T. Duignan and M. Neumann, Orb-v3: Atomistic Simulation at Scale, 2025, DOI: [10.48550/arXiv.2504.06231](https://doi.org/10.48550/arXiv.2504.06231).
- 22 B. M. Wood, M. Dzamba, X. Fu, M. Gao, M. Shuaibi, L. Barroso-Luque, K. Abdelmaqsoud, V. Gharakhanyan, J. R. Kitchin, D. S. Levine, K. Michel, A. Sriram, T. Cohen, A. Das, A. Rizvi, S. J. Sahoo, Z. W. Ulissi and C. L. Zitnick, UMA: A Family of Universal Models for Atoms, 2025, DOI: [10.48550/arXiv.2506.23971](https://doi.org/10.48550/arXiv.2506.23971).
- 23 X. Fu, B. M. Wood, L. Barroso-Luque, D. S. Levine, M. Gao, M. Dzamba and C. L. Zitnick, Learning Smooth and Expressive Interatomic Potentials for Physical Property Prediction, 2025, DOI: [10.48550/arXiv.2502.12147](https://doi.org/10.48550/arXiv.2502.12147).
- 24 A. Kondinski, A. Menon, D. Nurkowski, F. Farazi, S. Mosbach, J. Akroyd and M. Kraft, *J. Am. Chem. Soc.*, 2022, **144**, 11713–11728.
- 25 J. Bai, S. D. Rihm, A. Kondinski, F. Saluz, X. Deng, G. Brownbridge, S. Mosbach, J. Akroyd and M. Kraft, *Digital Discovery*, 2025, **4**, 2123–2135.
- 26 A. Kondinski, A. M. Oyarzún, S. D. Rihm, J. Bai, S. Mosbach, J. Akroyd and M. Kraft, *Eur. J. Inorg. Chem.*, 2025, e202500115.
- 27 M. Q. Lim, X. Wang, O. Inderwildi and M. Kraft, in *Intelligent Decarbonisation: Can Artificial Intelligence and Cyber-Physical Systems Help Achieve Climate Mitigation Targets?*, ed. O. Inderwildi and M. Kraft, Springer International Publishing, Cham, 2022, pp. 39–53.
- 28 S. D. Rihm, D. N. Tran, A. Kondinski, L. Pascazio, F. Saluz, X. Deng, S. Mosbach, J. Akroyd and M. Kraft, *Data-Centric Eng.*, 2025, **6**, e22.
- 29 S. D. Rihm, F. Saluz, A. Kondinski, J. Bai, P. W. V. Butler, S. Mosbach, J. Akroyd and M. Kraft, *Digital Discovery*, 2025, **4**, 2893–2909.
- 30 E. Bitzek, P. Koskinen, F. Gähler, M. Moseler and P. Gumbsch, *Phys. Rev. Lett.*, 2006, **97**, 170201.
- 31 A. Hjorth Larsen, J. Jørgen 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. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Villhelmsen, M. Walter, Z. Zeng and K. W. Jacobsen, *J. Phys.: Condens. Matter*, 2017, **29**, 273002.
- 32 L. Barroso-Luque, M. Shuaibi, X. Fu, B. M. Wood, M. Dzamba, M. Gao, A. Rizvi, C. L. Zitnick and Z. W. Ulissi,



- Open Materials 2024 (OMat24) Inorganic Materials Dataset and Models, 2024, DOI: [10.48550/arXiv.2410.12771](https://doi.org/10.48550/arXiv.2410.12771).
- 33 D. S. Levine, M. Shuaibi, E. W. C. Spotte-Smith, M. G. Taylor, M. R. Hasyim, K. Michel, I. Batatia, G. Csányi, M. Dzamba, P. Eastman, N. C. Frey, X. Fu, V. Gharakhanyan, A. S. Krishnapriyan, J. A. Rackers, S. Raja, A. Rizvi, A. S. Rosen, Z. Ulissi, S. Vargas, C. L. Zitnick, S. M. Blau and B. M. Wood, The Open Molecules 2025 (OMol25) Dataset, Evaluations, and Models, 2025, DOI: [10.48550/arXiv.2505.08762](https://doi.org/10.48550/arXiv.2505.08762).
- 34 A. Sriram, L. M. Brabson, X. Yu, S. Choi, K. Abdelmaqsoud, E. Moubarak, P. de Haan, S. Löwe, J. Brehmer, J. R. Kitchin, M. Welling, C. L. Zitnick, Z. Ulissi, A. J. Medford and D. S. Sholl, The Open DAC 2025 Dataset for Sorbent Discovery in Direct Air Capture, 2025, DOI: [10.48550/arXiv.2508.03162](https://doi.org/10.48550/arXiv.2508.03162).
- 35 E. R. Johnson and A. D. Becke, *J. Chem. Phys.*, 2005, **123**, 024101.
- 36 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.
- 37 S. Takamoto, C. Shinagawa, D. Motoki, K. Nakago, W. Li, I. Kurata, T. Watanabe, Y. Yayama, H. Iriguchi, Y. Asano, T. Onodera, T. Ishii, T. Kudo, H. Ono, R. Sawada, R. Ishitani, M. Ong, T. Yamaguchi, T. Kataoka, A. Hayashi, N. Charoenphakdee and T. Ibuka, *Nat. Commun.*, 2022, **13**, 2991.
- 38 C. Bannwarth, S. Ehlert and S. Grimme, *J. Chem. Theory Comput.*, 2019, **15**, 1652–1671.
- 39 F. Neese, *Wiley Interdiscip. Rev.:Comput. Mol. Sci.*, 2012, **2**, 73–78.
- 40 F. Neese, *Wiley Interdiscip. Rev.:Comput. Mol. Sci.*, 2022, **12**, e1606.
- 41 J. G. Brandenburg, C. Bannwarth, A. Hansen and S. Grimme, *J. Chem. Phys.*, 2018, **148**, 064104.
- 42 M. Miklitz and K. E. Jelfs, *J. Chem. Inf. Model.*, 2018, **58**, 2387–2391.
- 43 V. Santolini, G. A. Tribello and K. E. Jelfs, *Chem. Commun.*, 2015, **51**, 15542–15545.
- 44 M. O'Shaughnessy, J. Glover, R. Hafizi, M. Barhi, R. Clowes, S. Y. Chong, S. P. Argent, G. M. Day and A. I. Cooper, *Nature*, 2024, **630**, 102–108.
- 45 L. Pascazio, S. D. Rihm, A. Naseri, S. Mosbach, J. Akroyd and M. Kraft, *J. Chem. Inf. Model.*, 2023, **63**, 6569–6586.

