MOFBuilder: Automated end-to-end modeling of MOF dynamics for high-throughput screening

Chenxi Li[1] and Mårten S. G. Ahlquist[1*]

[1*]Division of Theoretical Chemistry and Biology, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm, 10044, Sweden.

*Corresponding author(s). E-mail(s): ahlqui@kth.se;

S1. Computational details

S1.1 siRNA in NU-1000 System Setup and MD Protocols

The simulation of siRNA transport within the NU-1000 framework began with the construction of a double-stranded siRNA molecule using the Maestro suite (Schrödinger, LLC), based on established sequences Teplensky et al. (2019). The siRNA topology was generated in GROMACS Abraham et al. (2015) utilizing the AMBER14SB-OL21 force field Zgarbová et al. (2011); Maier et al. (2015). For the MOF framework, the NU-1000 organic linker topology and RESP charges were parameterized using the GAFF2 force field Wang et al. (2004); de Gracia Triviño et al. (2025). The zirconium nodes were modeled with the cationic dummy atom model (CDAM) to ensure an accurate non-bonded representation of the metal-organic interface Su and Ahlquist (2021). Independent 50 ns simulations were first conducted for both the siRNA and the MOF framework. Following these initial runs, the siRNA geometry was inserted into a primary channel of the NU-1000 framework. The merged system was re-solvated with TIP3P water Mark and Nilsson (2001) to maintain a consistent saline solution and equilibrated for a further 100 ns in the NPT ensemble.

S1.2 Descriptor Extraction in High-Throughput Screening of UiO-66 Derivatives

The setup and simulation of the UiO-66 derivative library were executed through the MOFBuilder workflow. Organic linkers were parameterized with GAFF2 and zirconium nodes were represented by CDAM. Quantum chemical properties for each linker were calculated using VeloxChem de Gracia Triviño et al. (2025) after geometry optimization at the B3LYP-d4/def2-SVP level of theory. To ensure a robust electronic characterization, each of the 30 linkers was modeled under three chemically distinct representations: the fully protonated neutral acid (n), the doubly deprotonated dianion, and a truncated central p-phenylene core (c).Static geometric descriptors (D_i, D_f, D_{if}, ASA, void fraction, and linker dimensions) were derived from relaxed crystal structures. Dynamic descriptors were extracted from solvated MD simulations, including diffusion coefficients (D_e), hydrogen bond statistics ($HB1$ water to linker and $HB2$ linker to water), and the average coordination number of water molecules around linkers ($water\ CN$). Electronic descriptors, including extreme RESP partial charges ($maxchg$, $minchg$), $HBdonor$, μ, α, and $\log P$(hydrophobicity), were derived independently for all chemical representations.

S1.3 General Molecular Dynamics Protocols

All MD simulations were performed using GROMACS-2024.2 under 3D periodic conditions. For the CO2-related simulations (Fig 2), all gas molecules were initially placed at a distance from the MOF models to avoid bias. Simulations utilized a 1 fs timestep. The equilibration protocol consisted of 100 ps of NVT followed by 100 ps of NPT, leading to a 100 ns production run at 298.15 K. Non-bonded Lennard-Jones and Coulombic interactions were calculated using a 10 Å cutoff, with long-range electrostatics managed via the Particle Mesh Ewald (PME) method. The V-rescale thermostat was employed for temperature regulation. Trajectory visualization and analysis—including Root-Mean-Square Deviation (RMSD), average hydrogen bond analysis, Radial Distribution Function (RDF), and coordination number analysis, were conducted using mdanalsyis and VIAMD software Skånberg et al. (2023).

S2. Performance Benchmarking

Performance benchmarks for the building process, as presented in Fig 1d, were executed on an Apple M2 chip with 16GB of RAM to characterize the efficiency of the workflow on consumer-grade hardware.

S3. Cases analysis

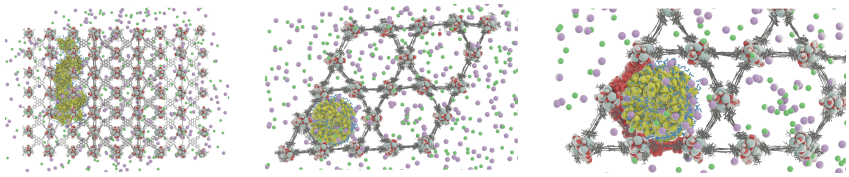

Fig. S1: Side view and top view of case1 system

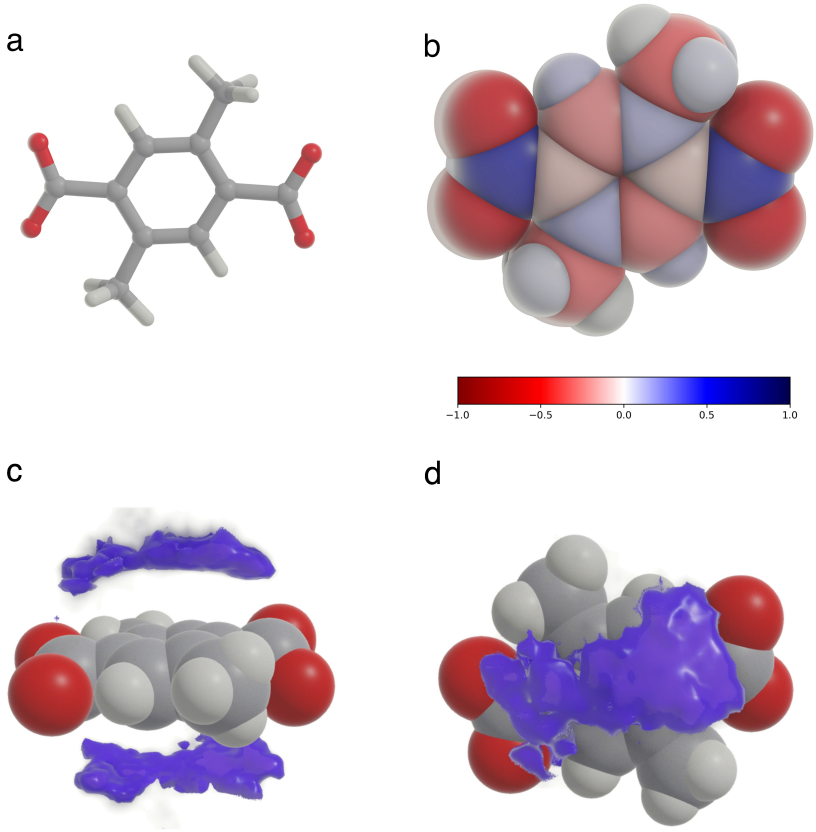

Fig. S2: (a) Visualization of linker 10#; (b) ESP charge analysis and (c-d) spatial distribution analysis of CO_2 around linker 10#

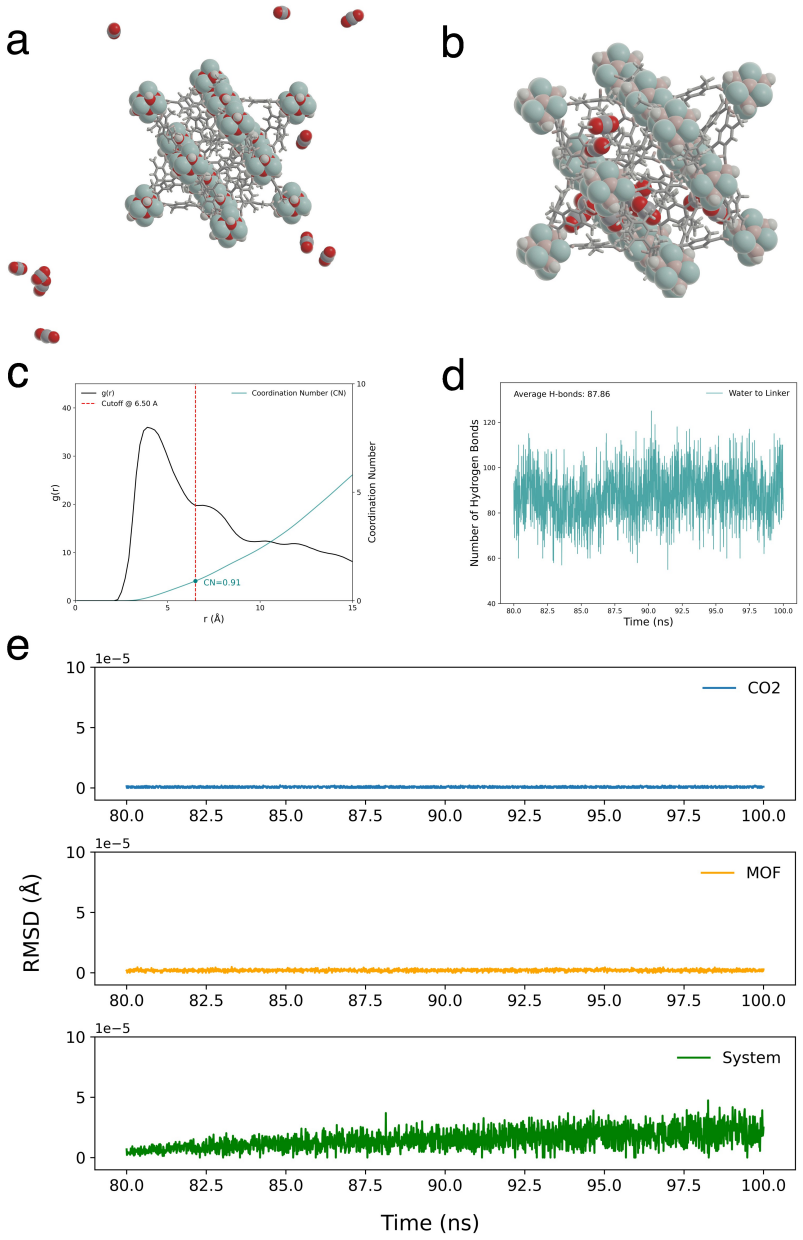

Fig. S3: (a) Snapshots of CO_2 at beginning of simulation; (b) Snapshots of CO_2 at end of simulation; (c) RDF analysis and coordination number extraction; (d) Hydrogen bond analysis from water to linker; (e) RMSD analysis of CO_2, MOF, and system during the last 20 ns production run

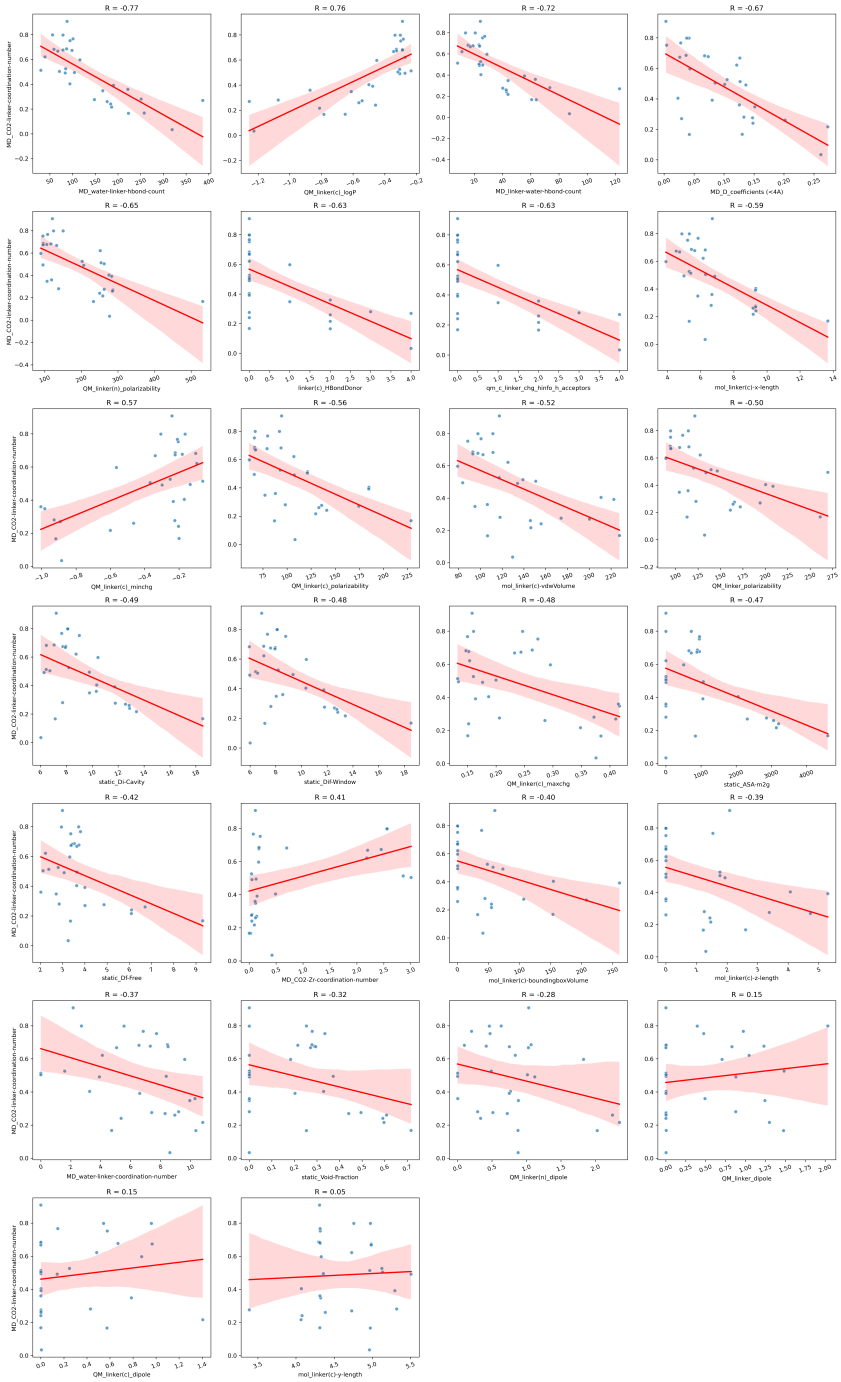

Fig. S4: Intercorrelations among descriptors

References

Abraham MJ, Murtola T, Schulz R, et al (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1-2:19–25. https://doi.org/10.1016/j.softx.2015.06.001, URL https://www.sciencedirect.com/science/article/pii/S2352711015000059

de Gracia Triviño JA, Brumboiu IE, Carrasco-Busturia D, et al (2025) VeloxChem Quantum–Classical Interoperability for Modeling of Complex Molecular Systems. The Journal of Physical Chemistry A 129(32):7575–7587. https://doi.org/10.1021/acs.jpca.5c03187, URL https://doi.org/10.1021/acs.jpca.5c03187, publisher: American Chemical Society

Maier JA, Martinez C, Kasavajhala K, et al (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. Journal of Chemical Theory and Computation 11(8):3696–3713. https://doi.org/10.1021/acs.jctc.5b00255, URL https://doi.org/10.1021/acs.jctc.5b00255, publisher: American Chemical Society

Mark P, Nilsson L (2001) Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. The Journal of Physical Chemistry A 105(43):9954–9960. https://doi.org/10.1021/jp003020w, URL https://doi.org/10.1021/jp003020w, publisher: American Chemical Society

Skånberg R, Hotz I, Ynnerman A, et al (2023) VIAMD: a Software for Visual Interactive Analysis of Molecular Dynamics. Journal of Chemical Information and Modeling 63(23):7382–7391. https://doi.org/10.1021/acs.jcim.3c01033, URL https://doi.org/10.1021/acs.jcim.3c01033, publisher: American Chemical Society

Su H, Ahlquist MSG (2021) Nonbonded Zr4+ and Hf4+ Models for Simulations of Condensed Phase Metal–Organic Frameworks. The Journal of Physical Chemistry C 125(11):6471–6478. https://doi.org/10.1021/acs.jpcc.1c00759, URL https://doi.org/10.1021/acs.jpcc.1c00759, publisher: American Chemical Society

Teplensky MH, Fantham M, Poudel C, et al (2019) A Highly Porous Metal-Organic Framework System to Deliver Payloads for Gene Knockdown. Chem 5(11):2926–2941. https://doi.org/10.1016/j.chempr.2019.08.015, URL https://linkinghub.elsevier.com/retrieve/pii/S2451929419303845

Wang J, Wolf RM, Caldwell JW, et al (2004) Development and testing of a general amber force field. Journal of Computational Chemistry 25(9):1157–1174. https://doi.org/10.1002/jcc.20035, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.20035, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jcc.20035

Zgarbová M, Otyepka M, Šponer J, et al (2011) Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. Journal of Chemical Theory and Computation

7(9):2886–2902. https://doi.org/10.1021/ct200162x, URL https://doi.org/10.1021/ct200162x, publisher: American Chemical Society