Freedom of design: towards in silico design of molecules with desired quantum-mechanical properties

Leonardo Medrano Sandonas[1], A. Tkatchenko[2], G. Cuniberti[1]

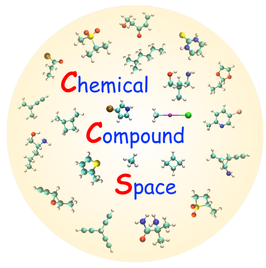
[1] Institute of Materials Science and Max Bergmann Center of Biomaterials, TU Dresden, Germany.

[2] Department of Physics and Materials Science, University of Luxembourg, Luxembourg.

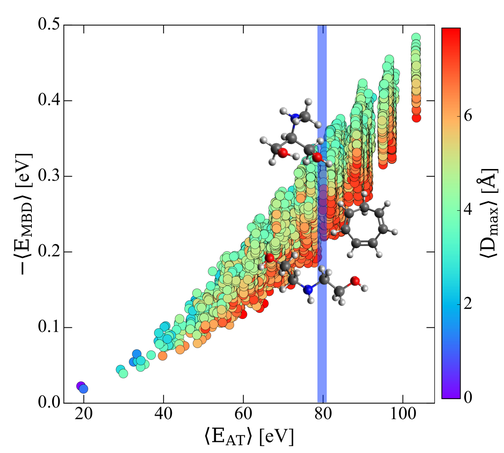
"Freedom of design" concept[1,2]

Flexibility when searching for a single molecule with a desired pair of quantum-mechanical (QM) properties or distinct molecules with a targeted set of QM properties.

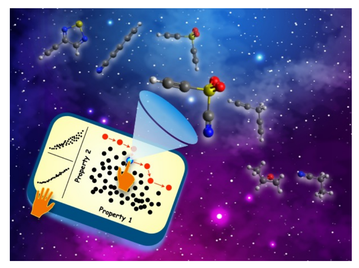
Property-to-property relationships



MBD energy vs Atomization energy



The challenge



Quantum-mechanical properties *(QM7-X and AQM datasets)* → **Inverse mapping** → Molecular structure

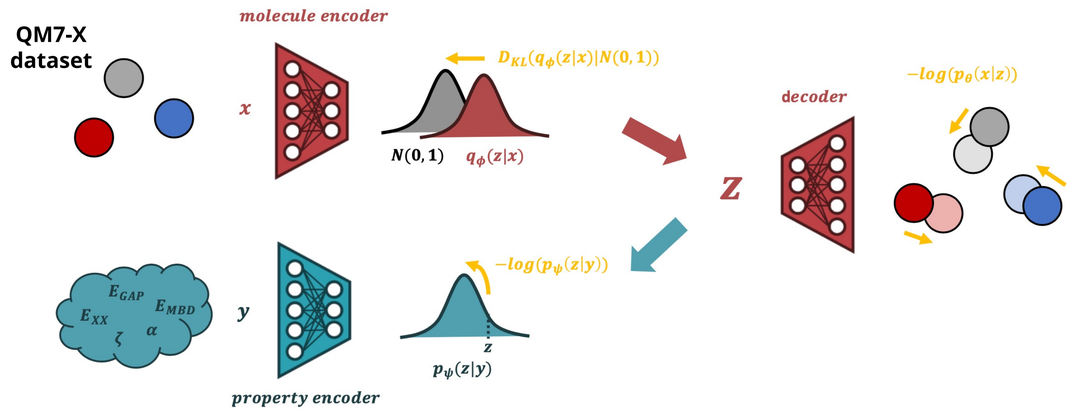
Conditional variational autoencoder

ML-augmented semi-empirical method

Conditional variational autoenconder[3]

Training the quantum inverse mapping (QIM) model

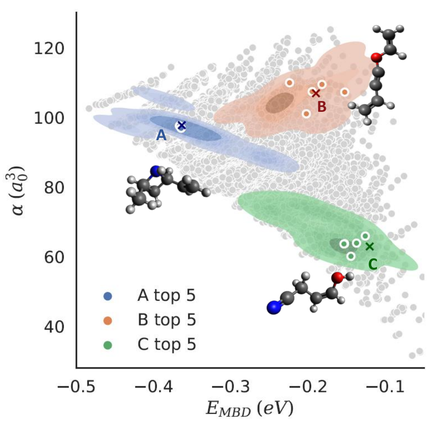
The QIM model was developed using a VAE architecture that employs Coulomb matrices as molecular representations along with the corresponding quantum-mechanical properties.



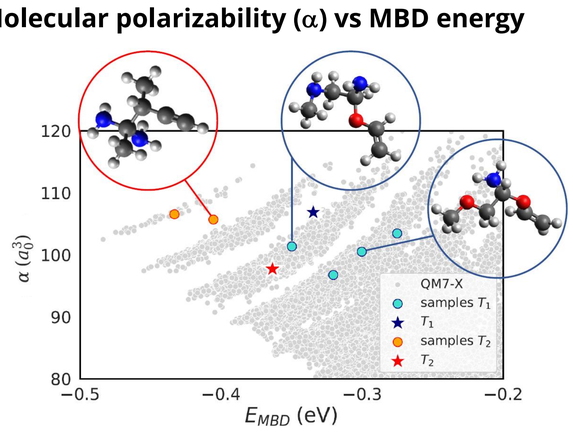
Structural data — Property data

$$\text{loss} = \boxed{D_{KL}[q_\phi(z|x)||p(z)] - \mathbb{E}_{q_\phi}[\log(p_\theta(x|z))]} - \boxed{\mathbb{E}_{q_\phi}[\log(p_\psi(z|y))]}$$

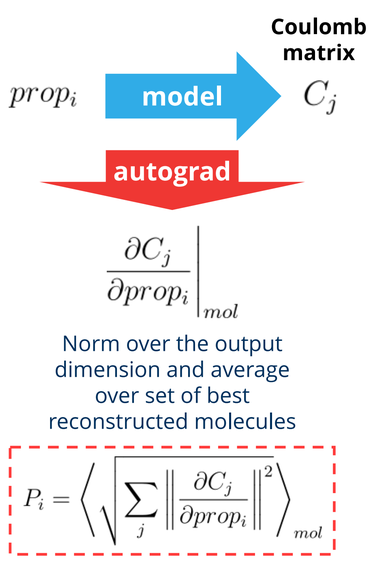
Multi-property design tasks

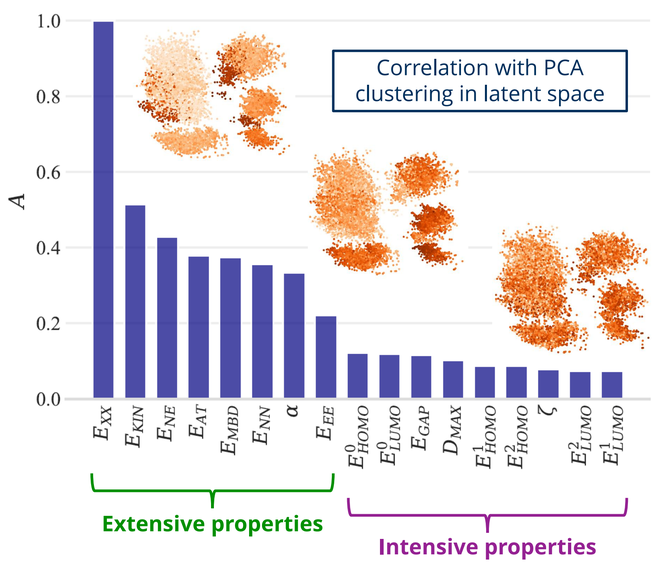


Molecular polarizability (α) vs MBD energy

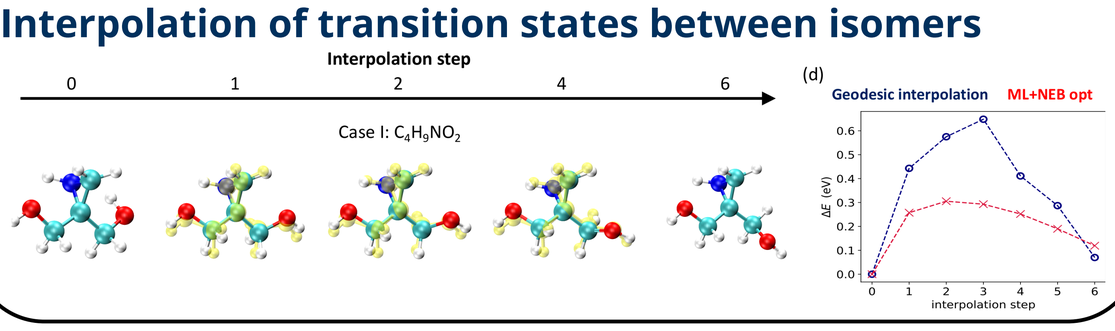


Interpretability of the QIM model





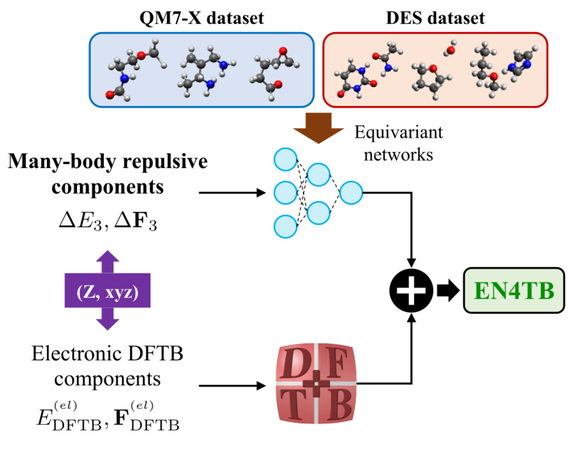
Interpolation of transition states between isomers

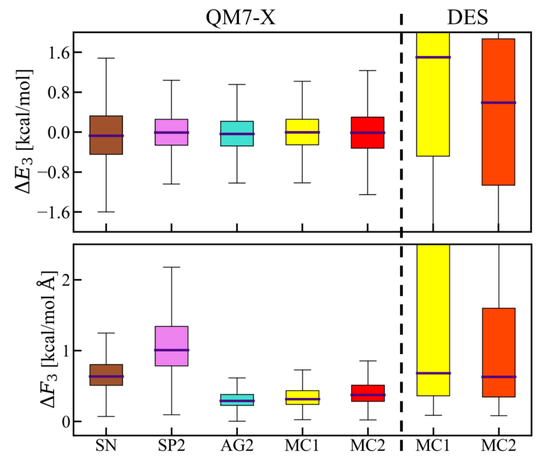


ML-augmented semi-empirical method[4,5]

Enhanced density functional tight-binding method

Equivariant Δ_3 many-body potentials have been trained to improve the performance of DFTB3 method, targeting PBE0 level of theory. We have considered state-of-the-art equivariant neural networks such as SpookyNet (SP), MACE (MC), and Allegro (AG).

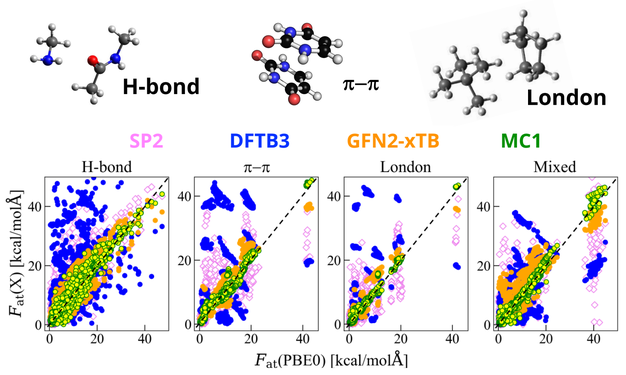


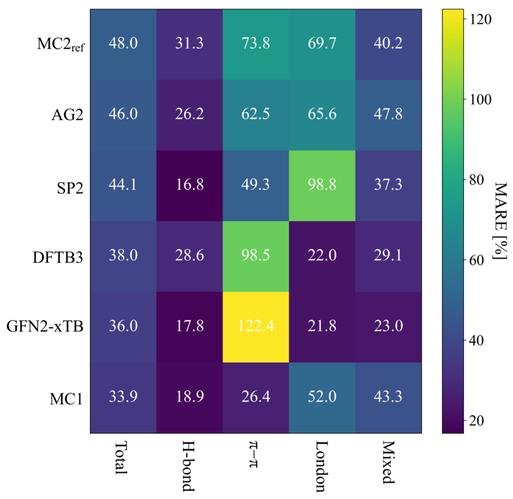


Predicting properties of non-covalent systems

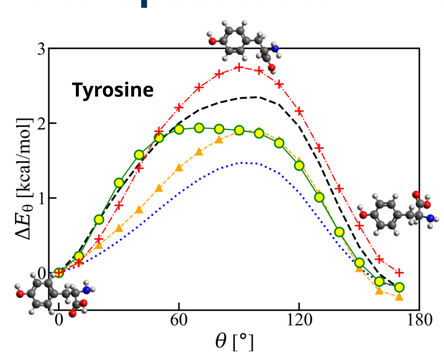
Performance of Δ_3 potentials in predicting the interaction energy E_{int} and atomic forces of molecular dimers in the S66x8 dataset.

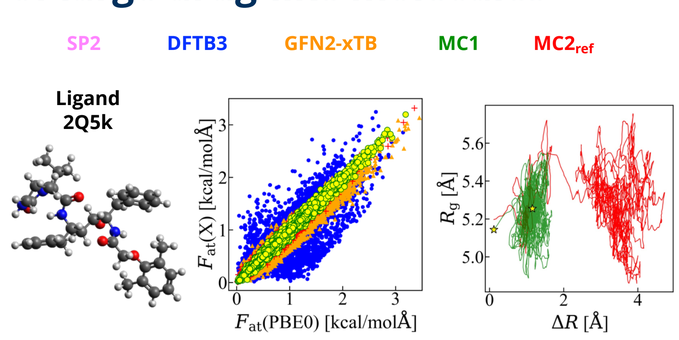
$$E_{int} = E_{dimer} - E_{mon1} - E_{mon2}$$



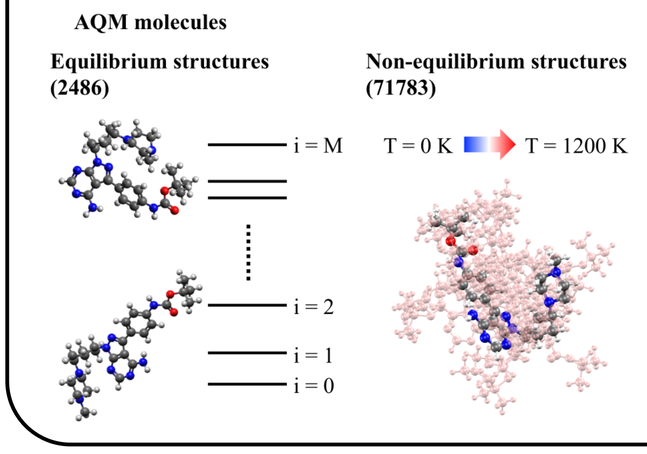


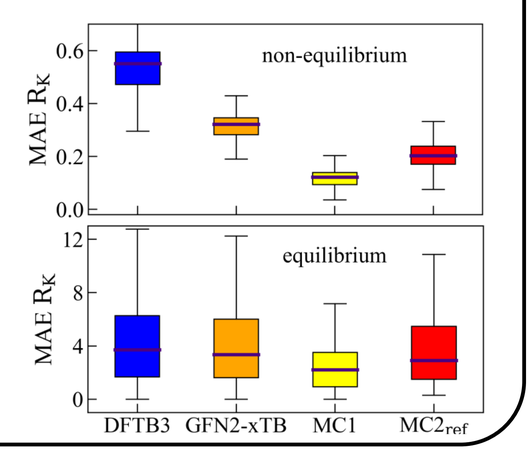
Non-equilibrium states of large drug-like molecules





Energetic ranking of molecular conformations





References

[1] **L. Medrano Sandonas** *et al.*, *Chem. Sci.* **14**, 39, (2023).
[2] J. Hoja, **L. Medrano Sandonas** *et al.*, *Sci. Data* **8**, 43, (2021).
[3] A. Fallani, **L. Medrano Sandonas**, and A. Tkatchenko. *Nat. Commun.* **15**, 6061, (2024).
[4] M. Stöhr, **L. Medrano Sandonas**, and A. Tkatchenko, *J. Phys. Chem. Lett.* **11**, 6835, (2020).
[5] **L. Medrano Sandonas** *et al.*, *Sci. Data* **11**, 742, (2024).

Contact:

Dr. Leonardo Medrano Sandonas

leonardo.medrano@tu-dresden.de
(+49) 351 46331424