ML4Chem

ML4CHEM 2026 · SUMMER SCHOOL

Modern ML in Chemistry

From (foundational) interatomic potentials to generative AI

Book of Abstracts

3–5 August 2026

Faculty of Chemistry, Leipzig University

Johannisallee 29, 04103 Leipzig, Germany

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

Funded by the European Union

Co-organized and funded by COST Action DAEMON (CA22154)

Contents

CONTRIBUTED TALKS

POSTERS

Contributed Talks

Aug 3, 9:00–09:15

ML-accelerated adsorption dynamics on semiconductor surfaces

Hendrik Bergner-Weiske[1], Patrick Mélix[1], Rhyan Barrett[1], Ralf Tonner-Zech[1], Julia Westermayr[1]

[1] Leipzig University

The functionalization of semiconductor surfaces is regarded as a cornerstone for tailoring electronic and chemical properties of materials. Ensemble-averaged snapshots of adsorption structures are provided by experimental techniques such as scanning tunnelling microscopy, yet the reaction pathways, transient intermediates, and thermally activated processes underlying these snapshots are not directly accessible. Conventional static DFT calculations are unable to capture ensembles, temperature effects, or dynamic reaction pathways, leaving a fundamental gap between theory and experiment.

This sampling bottleneck can be overcome by fine-tuning of a foundational MACE model on only a few thousand ab initio snapshots, achieving statistically converged adsorption dynamics at DFT accuracy for functionalized cyclooctyne on Si(001) [1]. The “molecular gun” workflow generates over 10,000 independent trajectories at a speedup exceeding three orders of magnitude, recovering rare intermediates and resolving the competition between adsorption motifs [1] (featured on the front cover of Digital Discovery). This work is extended to enable statistically meaningful, ab initio-quality dynamics on different semiconductor surfaces. A comparable fine-tuning strategy is applied across multiple material systems, revealing quantitative differences in adsorption statistics and substrate-dependent preferences. In addition, the impact of training-data composition is examined by comparing high-energy configurations from ab initio molecular dynamics with snapshots from static DFT calculations, systematically assessing the influence of data variety, quantity, and the sampling strategy on predictive accuracy [2]. This work establishes a general framework for probing adsorption dynamics on semiconductor surfaces at a fraction of the cost of DFT, bridging the gap between static theory and time-resolved experiment. With machine learning potentials reaching broad applicability across material classes, this approach is poised to become a standard tool in surface science and the computational design of organic/semiconductor interfaces.

We acknowledge funding from TU Dresden ZIH, PC² Paderborn, and SC Leipzig.

References:

[1] H. Weiske, R. Barrett, R. Tonner-Zech, P. Melix, J. Westermayr, Digital Discovery 2026, 5, 1501–1509 (Front Cover).

[2] H. Weiske, R. Tonner-Zech, in preparation.

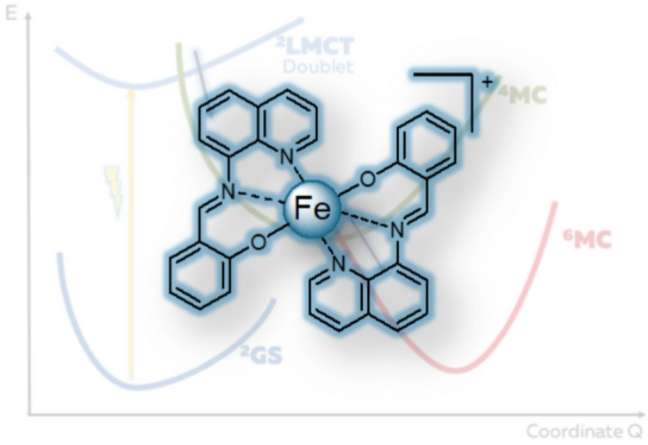
Aug 3, 14:00–14:15

Vibrational Fingerprints of Ultrafast Spin Crossover in an Fe(III) Complex

Dilara Farkhutdinova[1], Sebastian Mai[1], Leticia González[1]

[1] Institute of Theoretical Chemistry, University of Vienna, Währinger Str. 17, 1090 Vienna, Austria

Molecular bistability and light-switchable spin states in earth-abundant transition-metal complexes are of intense current interest for photomagnetic and molecular-electronic applications. Here we resolve the mechanism of light-induced spin crossover (SCO) in the iron(III) complex shown below and identify the specific nuclear motions that drive it. We investigate the mechanism with nonadiabatic dynamics simulations, using the trajectory surface hopping method SHARC (Surface Hopping including ARbitrary Couplings), on potential energy surfaces from a linear vibronic coupling (LVC) model. The LVC model is parametrised from multiconfigurational restricted active space self-consistent field calculations, capturing the complex character of the close-lying doublet, quartet, and sextet states. The simulations reveal a subpicosecond cascade: intersystem crossing from low-spin doublet ligand-to-metal charge transfer states to intermediate quartet states, followed by a second intersystem crossing to long-lived high-spin sextet states. Analysis of the trajectories exposes coherent normal modes exchanging energy through mode mixing, and pinpoints Fe–N and Fe–O stretching/bending together with quinoline-ring deformations as the modes that activate the transition, underscoring the intrinsically multidimensional nature of the non-adiabatic spin transition. Only a few well-defined modes govern the process, so the mechanism may carry over to chemically related systems. Testing this across the wider family of Fe(III) complexes is where machine-learned excited-state models and automated coordinate discovery offer a route to scale.



$[Fe(qsal)_2]^+$ compound and schematic picture of excited states processes

Aug 4, 9:00–09:15

Incorporating long-range interactions via the multipole expansion into ground and excited-state molecular simulations

Rhyan Barrett[1]

[1] Leipzig University, Faculty for Chemistry

Simulating long-range interactions remains a significant challenge for molecular machine learning (ML) potentials due to the need to accurately capture interactions over large spatial regions. In this work, we integrate the multipole expansion into equivariant ML potentials to model long-range interactions present in QM/MM simulations more accurately. By incorporating the multipole expansion, we are able to effectively capture environmental long-range effects in both ground and excited states. Benchmark evaluations demonstrate the superior performance of including higher-order features from atoms in the environment. To showcase the efficacy of our model, we accurately predict properties such as energies and forces for nickel complex systems and simulate the nonadiabatic excited-state dynamics of a ring-opening reaction in solution. Furthermore, we show that transfer learning from foundational models trained without any explicit environment enhances data efficiency, reducing the need to generate large QM/MM datasets before training. These examples demonstrate the versatility of our approach, paving the way for efficient, accurate, and scalable simulations of complex molecular systems and materials across electronic states.

Aug 4, 14:00–14:15

Generative Inverse Design of Tetrathiafulvalene-Based Molecular Materials for Organic Electronics

Patrícia Ramos[1], José Oliveira[2], Sandra Rabaça[3]

[1] Technical University of Porto, INESC TEC

[2] University of Porto, INESC TEC

[3] University of Lisbon

This work presents a generative framework based on stochastic-differential-equation diffusion models and graph-based molecular learning for the design of tetrathiafulvalene-containing (TTF-containing) materials for organic electronics. Starting from a curated dataset of over 5,500 crystallographic TTF-based structures, including neutral donors, radical-cation/charge-transfer salts, and coordination complexes, the workflow extracts relevant molecular components. It integrates graph reconstruction, joint 2D/3D diffusion generation, quantum-chemical screening, and machine-learning-guided property prediction. The model jointly represents molecular connectivity and three-dimensional geometry, enabling the exploration of diverse candidates while preserving the TTF donor motif. Generated molecules are evaluated with RDKit filters to remove invalid structures and assess valence consistency, connectivity, and geometrical plausibility. Valid candidates are ranked by novelty, TTF-core retention, and structural realism. To estimate electronic potential, selected molecules are screened using semiempirical tight-binding calculations, providing frontier orbital energies, HOMO-LUMO gaps, dipole moments, and molecular planarity. The most promising candidates are refined through density functional theory calculations to evaluate p-type organic electronic properties, including ionization potentials and hole reorganization energies. Finally, an active-learning loop uses xTB- and DFT-labelled molecules to train property predictors based on molecular fingerprints and graph-derived representations. These predictors guide subsequent generation toward chemically realistic and electronically promising regions of chemical space. By combining SDE-based molecular generation, quantum-chemical evaluation, and machine-learning prioritization, the framework supports efficient inverse design of TTF-based organic electronic materials while reducing expensive calculations for large-scale screening.

Posters

Poster 1

Nonadiabatic Excited-State Pathways from SOCT-ISC to Hydration-Driven Proton Transfer: A Trajectory-Based Perspective

Moumita Banerjee[1]

[1] Linköping University

Poster 2

ML vs. expert intuition for AD-HoC Nickel Cross-Coupling: Genetic algorithms coupled with regression algorithms to learn from heterogeneous experimental data

Daniel Bitterlich[1], Indrajit Ghosh[2,3], Burkhard König[2], Julia Westermayr[1,4]

[1] Faculty of Chemistry, Leipzig University, Leipzig, Germany

[2] Faculty of Chemistry and Pharmacy, University of Regensburg, Regensburg, Germany

[3] CEET, Technical University of Ostrava

[4] ScaDS.AI Dresden/Leipzig, Leipzig, Germany

Machine learning has emerged as a promising tool for advancing experiment optimization, but experimental datasets remain challenging: they are expensive to generate, often sparse and biased, and intrinsically noisy. Consequently, many machine learning studies rely either on nearly noise-free theoretical data or on highly curated databases from experiments that are often limited in chemical space and accuracy. Here, we investigate the potential of machine learning for yield prediction in AD-HoC nickel cross-couplings under photoredox conditions [1]. The reaction space is characterized by broad chemical diversity and high dimensionality with over $19 \cdot 10^6$ possible combinations of reactants. To navigate this space effectively and allow for yield prediction with machine learning, we develop and apply a genetic algorithm to identify an informative subset of reactions suitable for model training. The resulting models were then applied to virtual high-throughput screening of the whole chemical space to identify regions of high reaction yield, model uncertainty, and untested combinations. Comparison to expert predictions reveals insights into biases and limits of both machine learning algorithms and human intuition and shows how to overcome bottlenecks in real-world chemical optimization, enabling navigation of vast chemical spaces.

References:

[1] I. Ghosh et al., Nature, 2026, 619, 87-93, 10.1038/s41586-023-06087-4.

Poster 3

Gas-Phase Vibrational Spectroscopy of $PF_6^-(H_2O)_{0-18}$: From Isolated Clusters to the Condensed Phase

Hannah Buttkus[1]

[1] Leipzig University

Poster 4

Toward Closed-Loop Design of Enantioselective N-Heterocyclic Carbene Catalysts for the Benzoin Reaction

Shuai Chen[1], Robert Pollice[1]

[1] University of Groningen

Poster 5

A High-Throughput Generative Workflow for Data-Driven Reaction Environment Optimization

Robin Curth[1,2], Rhyan Barrett[1], Julia Westermayr[1,2]

[1] Wilhelm Ostwald Institute for Physical and Theoretical Chemistry, Leipzig University, Linnéstraße 2, Leipzig, 04103, Germany.

[2] Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Dresden/Leipzig, Germany.

The optimization of molecular environments to accelerate chemical reactions is a cornerstone of catalysis, drug development, and materials science. However, the enormous combinatorial complexity of possible environments makes conventional screening approaches inefficient and resource intensive. Here, we present a high-throughput automated workflow that combines predictive machine learning models, genetic algorithms, and generative deep learning to design chemical environments that minimize reaction activation barriers. Our approach iteratively refines the molecular environment for exploring new candidate configurations [1]. Applied to the Claisen rearrangement reaction of p-tolyl ether, our framework achieves a reduction in activation energy of 33 % (10 eV), which is confirmed by quantum chemical calculations. Beyond the performance enhancement, the method reveals important insights into favorable structural features to catalyze the reaction. Through enabling the inverse design of complex reaction environments, this platform represents a major step forward for data-driven research in sustainable chemistry.

References:

[1] Westermayr, J.; Gilkes, J.; Barrett, R.; Maurer, R. J.; Nat. Comput. Sci. 2023, 3 (2), 139–148.

Poster 6

Combining genetic algorithms with omits in kinematical refinements of cRED data: pathways towards higher accuracy

Alexander Feige[1], Paul Hager[2], Jungyoun Cho[2], Laura Pacoste[2], Xiaodong Zou[2], Julia Westermayr[1,3]

[1] Wilhelm-Ostwald-Institute for Physical and Theoretical Chemistry, Leipzig University

[2] Department of Chemistry, Stockholm University

[3] ScaDS.AI Leipzig (Center for Scalable Data Analytics and Artificial Intelligence)

Electron diffraction is an essential technique for determining crystal structures. Yet its accuracy frequently suffers due to dynamical scattering effects, leading to significant deviations in reflection intensities [1].

To address this challenge, we implemented a genetic optimization algorithm designed to selectively refine subsets of reflection data from continuous electron diffraction datasets [2], thereby enhancing the accuracy of structure refinement. This flexible algorithm accommodates multiple objectives, which can be either dependent or independent of refinement statistics. We tested our approach on datasets from a staircase-like crystal structure of the zeolite STW and multiple crystals of the inorganic salt KH_2PO_4. Our results demonstrate notable improvements in atomic displacement factors and facilitate precise localization of hydrogen atom positions, even under kinematical refinement conditions. These findings highlight the algorithm's efficacy as a practical method for identifying and excluding dataset regions strongly affected by dynamical scattering, being also applicable to other data collection strategies, such as serial electron diffraction.

References:

[1] T. Gruene, J. J. Holstein, G. H. Clever, B. Keppler, Nat. Rev. Chem. 2021, 5, 660.

[2] M. O. Cichocka, J. Ångström, B. Wang, X. Zou, S. Smeets, J. Appl. Crystallogr. 2018, 51, 1652.

Poster 7

Machine learning for smell: ordinal odor strength prediction of molecular perfumery components

Peter Fichtelmann[1]

[1] Wilhelm-Ostwald-Institute for Physical and Theoretical Chemistry, Leipzig University

Predicting olfactory perception from molecular structure is central to product design in perfumery, food and beverage, and health care. Among olfactory attributes, odor strength is key to perception, yet its modeling has been limited by scarce, fragmented intensity data. Here, we introduce an ordinal odor strength data set of over 2300 molecules, integrating two public sources and mapping structures to odorless, low, medium, and high categories. Comparing several molecular encodings and supervised learning algorithms, we identify effective prediction strategies. Dimensionality reduction and SHAP analysis reveal molecular shape, size, and polarity as primary drivers, consistent with mass-transport constraints on volatility, sorption, and receptor access. This scalable framework enables reliable odor-strength estimation for novel molecules, providing a foundation for in silico fragrance design.

Poster 8

First principles construction of Newns-Anderson Hamiltonians for hydrogen chemistry at surfaces

Nils Hertl[1], Zsuzsanna Koczor-Benda[1], Reinhard J. Maurer[1]

[1] University of Warwick

Poster 9

The Solvation Line: Mapping Close-Range Interactions in Reaction Mechanisms

F. Lauria[1], M. Blangetti[1], L. Degennaro[2], V. Pace[3], A. Maranzana[1]

[1] Department of Chemistry, University of Turin, P. Giuria street n. 7, I-10125, Turin, Italy

[2] Department of Pharmacy-Drug Science, University of Bari "Aldo Moro", E. Orabona street n. 4, I-70125, Bari, Italy

[3] Department of Chemistry, Sapienza University of Rome, Aldo Moro Square n. 5, I-00185, Rome, Italy

In the context of simulating reaction mechanisms in organic chemistry, there is a tendency to simplify computational models by omitting fundamental chemical species. These species, such as solvent molecules or counterions, may not be directly involved in the reaction but significantly influence the behaviour of the reactants and favour specific mechanistic pathways [1]. To simplify the system, these are often approximated using continuous medium models, such as the SMD model for solvents. Through Density Functional Theory (DFT) calculations, this study demonstrates how the omission of explicit chemical species can lead to results in disagreement with experimental data, thereby affecting our understanding of individual reaction steps. To facilitate this approach, we developed COMPLAX, a command-line utility tool to automate the preparation and ranking of microsolvated molecular systems (Figure 1). By leveraging xTB optimization algorithms [2], COMPLAX eliminates the need for manual manipulation via external visualization software. It provides a rapid and efficient method for generating high-quality, pre-optimized starting geometries, seamlessly bridging the gap between complex explicit solvation modeling and subsequent high-accuracy DFT calculations.

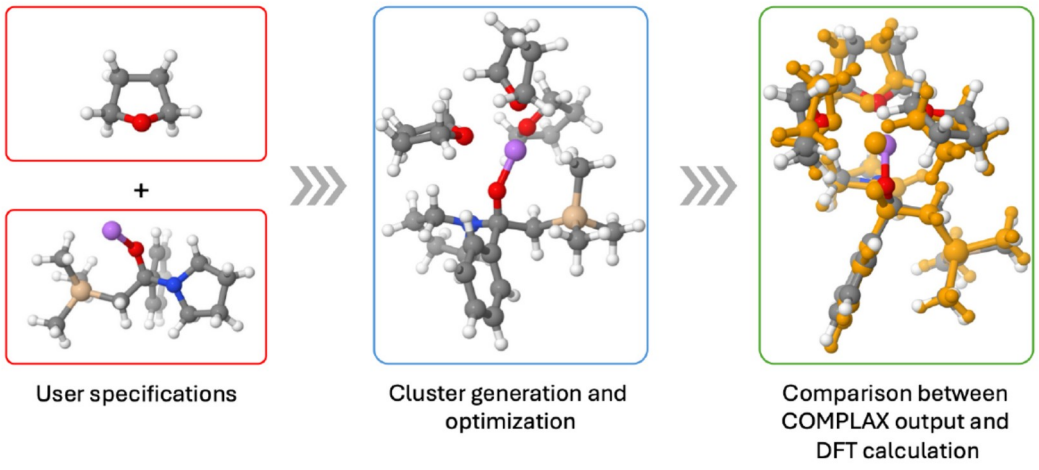


Figure 1. COMPLAX workflow for the automatic generation of microsolvated structures.

References:

[1] De Santis, G.; Maranzana, A.; Lauria, F.; Spennacchio, M.; Andresini, M.; Purgatorio, R.; Colella, M.; Luisi, R.; Degennaro, L., Advanced Synthesis & Catalysis, 2025, 367 (17). DOI: 10.1002/adsc.70083.

[2] Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S., WIREs Computational Molecular Science, 2020, 11 (2). DOI: 10.1002/wcms.1493.

Poster 10

Dataset Design for Buffer Region Neural Networks: from Sampling to Simulation

Bettina Lier[1], Chris Oostenbrink[1]

[1] Institute of Molecular Modeling and Simulation, BOKU University, Vienna, Muthgasse 18/6-54, 1190 Vienna, Austria

Buffer region neural network (BuRNN) simulations combine a hybrid QM/MM framework with a machine-learned interatomic potential for the quantum-mechanical region and buffer-region embedding at the QM/MM boundary. Their reliability depends on the accuracy and transferability of the machine-learned potential, which in turn rely on training data that capture the physically relevant regions of the potential-energy surface. Efficient model development therefore benefits from compact datasets with broad and representative structural coverage. Here, we present a workflow combining conformer generation, RMSD-based clustering, normal-mode sampling, neural-network training, and uncertainty-guided active learning to construct and iteratively refine such datasets. Model performance is assessed using both conventional prediction errors and structural observables from BuRNN simulations.

Poster 11

Mixed Metal Oxides in the Gas Phase: The Study of the FeAlO3+ and Al2O3+ Clusters

Lucas W. de Lima[1], Tatiana C. Penna[2,3], Vicente Zamudio-Bayer[4], J. Tobias Lau[4], Knut R. Asmis[2,3], Joachim Sauer[1], Michael Roemelt[1]

[1] Humboldt-Universität zu Berlin

[2] Wilhelm-Ostwald Institut für Physikalische und Theoretische Chemie, Universität Leipzig

[3] Fritz-Haber-Institut der Max-Planck-Gesellschaft

[4] Abteilung für Hochempfindliche Röntgenspektroskopie, Helmholtz-Zentrum Berlin für Materialien und Energie

Elucidating the atomistic structure of active sites in heterogeneous catalysis remains a fundamental challenge, owing to their intrinsic complexity and dynamic nature under operating conditions. The study of those systems involves techniques such as matrix isolation and gas-phase clusters, which enable detailed structural and electronic characterisation at the molecular level. In the realm of gas-phase clusters experiments, the investigation of mixed metal oxide catalysts uses state-of-the-art action spectroscopy approaches, where mass-selected ions (emulating the catalyst's active site) are characterised by techniques such as infrared photodissociation spectroscopy (IRPD) and X-ray absorption spectroscopy (XAS) [1]. Aluminium oxide (alumina) is a prototypical oxide material widely employed in heterogeneous catalysis, both as an active phase and as a support, due to its chemical and physical properties. The incorporation of transition metals into the alumina framework has emerged as an effective strategy to tailor catalytic performance, for instance, in isobutane synthesis [2]. In this regard, the study of the active site of iron-doped alumina is pivotal to understanding its catalytic activity. Herein, FeAlO3+ and Al2O3+ clusters were investigated as model systems for iron-doped alumina active sites by combining IRPD and XAS with electronic structure calculations. Structural assignment is achieved through comparison of experimental spectra with simulated harmonic infrared and X-ray absorption spectra at the oxygen K-edge. The results indicate that the measured cluster adopts a “key-like” structural motif exhibiting a C2v symmetry. Further insight into the electronic structure is obtained from X-ray absorption spectroscopy at both the oxygen K-edge and iron L-edge. The combined analysis reveals that FeAlO3+ is best described as an oxyl species with a general quintet spin state, in which the iron centre is in the +3 oxidation state. These findings afford a detailed atomistic picture of a prototypical iron-doped alumina motif, highlighting the power of combining action spectroscopy with simulations to resolve structure–property relationships in complex oxide systems. This work provides

fundamental insights into the nature of the active sites in mixed metal oxides and establishes a framework for the rational design of improved catalytic materials.

References:

[1] Schwarz, H., Asmis, K.R., et al. Chem Eur. J., 2019, 25, 2112.

[2] Venugopalan, A. T., Kandasamy, P., et al. Catal. Commun., 2021, 150, 106263.

Poster 12

E(3)-VERDE: A 3D-Equivariant Neural Network for the Complete Ground- and Excited-State Redox Landscape of Organic Photocatalysts

Andres Lopez[1], Leticia Gomes[1], Steven A. Lopez[1]

[1] Northeastern University

Poster 13

High-Precision Modeling of Nuclear Quantum Effects at Low Temperatures in CO/NaCl(100) Adsorption Using MACE Potentials and Path-Integral Methods

Mateo Pescador Arboleda[1], Shreya Sinha[1], Peter Saalfrank[1]

[1] Theoretical Chemistry, Institute of Chemistry, University of Potsdam, Potsdam

The adsorption of CO on NaCl(100) is a benchmark system for studying weak physisorption and nuclear quantum effects (NQEs). At cryogenic temperatures, the flat potential energy surface and subtle energetic differences dictate the stability and transitions between “C-bound” and metastable “O-bound” orientations. In this work, we employ the MACE architecture to develop a high-precision machine-learning interatomic potential for path-integral molecular dynamics simulations. First, we address a critical limitation of the initial model, where scaling to an extended cell revealed unphysical lateral interactions and system instabilities due to poorly captured multi-adsorbate effects. By incorporating extensive ab initio reference calculations in VASP, we successfully retrained the potential. Its precision is validated via classical molecular dynamics, where the Vibrational Density of States (VDOS) at 5 K and 30 K successfully resolves the Davydov splitting, a subtle spectroscopic signature of antiparallel CO orientations. Second, we employ Ring Polymer Instanton (RPI) theory to calculate thermal rate constants for CO rotation over a single sodium site. Utilizing our refined potential, we obtain semi-classical rates below the crossover temperature, where tunneling dominates. Above this threshold, we perform Ring Polymer Molecular Dynamics (RPMD) simulations to explore large-amplitude motions. We report the discovery of a novel isomerization mechanism involving a coupled rotation-translation toward neighboring sodium atoms. Crucially, this pathway is absent in classical molecular dynamics trajectories, providing direct evidence of heavy-atom tunneling at the interface. Our results demonstrate that high-fidelity machine-learned potentials enable the discovery of stationary points and quantum dynamical pathways often missed by conventional DFT, while drastically reducing the computational cost of path-integral simulations.

References:

[1] Sinha, S., Mladineo, B., Loncaric, I., & Saalfrank, P. (2024). The Journal of Physical Chemistry C, 128(49), 21117-21131.

[2] Choudhury, Arnab, et al. Natural Sciences 3.3 (2023).

[3] Sinha, S., Wodtke, A. M., & Saalfrank, P. (2025). Physical Chemistry Chemical Physics, 27(15), 7929-7942.

[4] Sinha, S., & Saalfrank, P. (2021). Physical Chemistry Chemical Physics, 23(13), 7860-7874.

[5] Sinha, S., & Saalfrank, P. (2026). Physical Chemistry Chemical Physics.

[6] Choudhury, A., DeVine, J. A., Sinha, S., Lau, J. A., Kandratsenka, A., Schwarzer, D., ... & Wodtke, A. M. (2022). Nature, 612(7941), 691-695.

Poster 14

Quantum nuclear and band-dispersion effects recover near-UV absorption in short-hydrogen-bonded organic crystals

Muhammad Nawaz Qaisrani[1]

[1] Technische Universität Ilmenau, Faculty of Mathematics and Natural Sciences, Institute of Physics, Department of Theoretical Solid State Physics

Near-UV optical absorption is increasingly reported in hydrogen-bonded organic and biomolecular materials lacking aromatic or extended π-conjugated chromophores. Yet, its microscopic origin remains elusive and electronic-structure calculations often overestimate experimental absorption onsets. Here, we combine machine-learned interatomic potentials for large-scale classical and quantum nuclear sampling with periodic excited-state calculations to address this discrepancy in L-pyroglutamine ammonium, a glutamine-derived molecular crystal with a crystallographically resolved short O–H···O hydrogen bond and non-aromatic near-UV optical response. We show that nuclear quantum effects stabilise proton-sharing configurations that are strongly suppressed classically, redshifting the lowest bright excitations by 0.5–0.8 eV and raising the fraction of configurations with bright excitations below 6 eV from approximately 3% to 30%. Explicit Brillouin-zone sampling provides a further, mechanistically distinct redshift of 0.5–1.1 eV, reflecting finite-k band-edge character rather than a purely Γ-point optical onset. Only when both effects are incorporated does the calculated onset recover the experimental 3.8–4.5 eV range. These results establish quantum proton fluctuations and finite-k band-edge effects as cooperative but physically distinct ingredients required for predictive optical spectroscopy of strongly hydrogen-bonded molecular materials.

References:

[1] A. D. Stephens, M. N. Qaisrani, et al., “Short hydrogen bonds enhance nonaromatic protein-related fluorescence,” *Proceedings of the National Academy of Sciences* **118**, e2020389118 (2021).

[2] M. N. Qaisrani, N. Kumar, C. Dreßler, R. Gebauer, and A. Hassanali, “Acid–Base Chemistry of Short Hydrogen Bonds: A Tale of Schrödinger’s Cat in Glutamine-Derived Crystals,” *Journal of Physical Chemistry Letters* **16**, 8588–8595 (2025).

[3] J. Hänseroth, M. Großmann, M. Grunert, A. Hassanali, E. Runge, C. Dreßler, and M. N. Qaisrani, “Quantum nuclear and band-dispersion effects recover near-UV absorption in short-hydrogen-bonded organic crystals,” arXiv:2606.24750 (2026).

Poster 15

Basin-Hopping Monte Carlo with Transferable Neural Network Potentials: Mapping Germanium and Aluminium Distributions in Zeolites

Indranil Saha[1,2], Andreas Erlebach[2], Oscar Bengtsson[3], Daniel Hewitt[3], Misbah Sarwar[4], Nicola Collis[4], Petr Nachtigall[2,‡], Christopher J. Heard[2], Ben Slater[3], Lukáš Grajciar[2]

[1] Multiscale Modelling of Heterogeneous Catalysis in Energy Systems, RWTH Aachen University, Schinkelstraße 8, 52062, Aachen, Germany

[2] Department of Physical and Macromolecular Chemistry, Faculty of Science, Charles University, Hlavova 2030/8, 128 43 Prague 2, Czech Republic

[3] Department of Chemistry, University College London, 20 Gordon Street, London WC1H 0AJ, UK

[4] Johnson Matthey Technology Centre, Blounts Ct Rd, Sonning Common, Reading RG4 9NH, UK

‡ Deceased.

Compositional disorder governs the catalytic activity, hydrolytic stability, and structural transformations of zeolites. Direct experimental characterisation of heteroatom distributions is often inconclusive, and brute-force ab initio screening of the configurational spaces involved (often >10^6 distinct configurations per unit cell) is intractable. We use basin-hopping Monte Carlo (BHMC) with transferable neural-network potentials (NNPs) as a general workflow to map low-energy heteroatom distributions. The same global-optimisation machinery accommodates NNPs trained for purely silicate frameworks and those trained for reactive aluminosilicate-water systems. For germanosilicates, a non-reactive NNP trained at PBE+D3(BJ) [1] was used to screen five D4R-containing topologies (UTL, BEC, UOV, IWW, *CTH) across Si/Ge from 2 to 19. Germanium clustering, not D4R preference alone, dominates the thermodynamics. Phase separation into Ge-rich and Ge-poor regions emerges at high loading regardless of structural unit type, and individual D4R units within the same unit cell show pronounced asymmetry in their Ge occupations. These trends help rationalise the mixed products observed in ADOR transformations. For sodiated aluminosilicates [2], BHMC was equipped with a custom sodium-positioning move-class on top of a reactive Al-Si NNP [3], screening CHA, MOR, and RHO at varying Si/Al. CHA and MOR favour Al-Si-Si-Al with strong avoidance of Al-O-Si-O-Al, while RHO inverts this preference at moderate Al content. Randomly generated and non-BHMC configurations sit far above the global minima, with direct consequences for catalysis modelling. BHMC paired with transferable NNPs enables broad mapping of compositional disorder across reactive and non-reactive zeolite chemistries at a fraction of the ab initio cost.

References:

[1] Saha, I., Erlebach, A., Nachtigall, P., Heard, C. J. & Grajciar, L. Germanium distributions in zeolites derived from neural network potentials. Catal. Sci. Technol. 2024, 14, 5838–5853. DOI: 10.1039/D4CY00763H

[2] Bengtsson, O., Hewitt, D., Saha, I., Sarwar, M., Collis, N., Erlebach, A., Grajciar, L., Heard, C. J. & Slater, B. Prediction of aluminium distributions in zeolites using basin-hopping and machine learning potentials. ChemRxiv 2026. DOI: 10.26434/chemrxiv.10001878/v2. Submitted to RSC Digital Discovery.

[3] Erlebach, A., Šípka, M., Saha, I., Nachtigall, P., Heard, C. J. & Grajciar, L. A reactive neural network framework for water-loaded acidic zeolites. Nat. Commun. 2024, 15, 4215. DOI: 10.1038/s41467-024-48609-2

[4] Erlebach, A., Nachtigall, P. & Grajciar, L. Accurate large-scale simulations of siliceous zeolites by neural network potentials. npj Comput. Mater. 2022, 8, 174. DOI: 10.1038/s41524-022-00865-w

Poster 16

Investigating ZnPd nanoparticles supported on ZnO using atomistic simulations with machine learning potentials

Roy Schumacher[1]

[1] Forschungszentrum Jülich

Supported metal nanoparticles play a central role in heterogeneous catalysis, yet their atomic-scale structure and its coupling to catalytic performance remain difficult to resolve using atomistic simulations. In our research we are investigating ZnPd nanoparticles supported on ZnO for methanol steam reforming, in close collaboration with transmission electron microscopy imaging and in-situ experiments. While ab initio simulation methods provide high accuracy, their computational cost limits their applicability to models much smaller than the few thousand atom models required to capture realistic particle shape and metal-support interactions. In addition to this the timescales accessible are typically only of the order of picoseconds. To overcome this limitation we are focusing on developing and applying machine learning potentials to our system, with the aim of reducing the gap between atomistic simulations and experimental work.

Poster 17

Restricted or unrestricted? The boundaries of universal MLIPs for organic photochemistry

Henrik Seng[1], Hanwen Zhang[1], Kjell Jorner[1]

[1] Institute for Chemical and Bioengineering, Department of Chemistry and Applied Biosciences, ETH Zürich

Universal machine-learning interatomic potentials (uMLIPs) approach DFT accuracy at a fraction of the cost, extending to triplet-driven photochemistry via charge- and spin-aware energetics. Their reliability beyond closed-shell molecules is not well established, and open-shell singlet (OSS) diradicals are the most challenging case, where even the DFT training data is itself unreliable. We benchmark five charge- and spin-aware uMLIPs (UMA-OMOL, MACE-OMOL, MACE-POLAR-1, OrbMol and AIMNet2-NSE) against DFT for the properties that drive triplet photochemistry: reaction energies and barriers, vertical and adiabatic singlet-triplet gaps, and full reaction profiles from closed to open shell. For closed-shell chemistry, including reaction energies and barriers, the best models reach DFT accuracy. For the open-shell singlet-triplet gaps the error grows several-fold, and its distribution stays heavy-tailed for every model, dominated by rare large-magnitude outliers rather than a uniform shift. Decomposing each uMLIP error separates a geometry contribution (misplaced structure) from an energy contribution (mis-ranked energy at that structure). The largest outliers are energy-dominated and concentrate on specific chemistry motifs. The open-shell error instead has a single dominant origin. Each model interpolates between the restricted (closed-shell) and unrestricted, broken-symmetry (open-shell) DFT solutions, and its position, not the diradical character, explains the error very accurately. uMLIPs that reproduce the spin polarization of OSS structures, such as MACE-POLAR-1, can flag these failures from their own output (best AUROC > 0.9), an intrinsic diagnostic that restricted models like AIMNet2-NSE lack.

Poster 18

Lighting up CO2 reduction: are organic photosensitizers paving the way to metal-free artificial photosynthesis?

Lucrezia Villa[1,*], Katharina Bica-Schröder[1]

[1] Institute of Applied Synthetic Chemistry, TU Wien, Getreidemarkt 9-163, 1060 Wien, Austria

Converting solar energy and CO2 into chemical fuels via artificial photosynthesis is a compelling strategy to address challenges in sustainable energy storage and carbon utilization [1]. Inspired by natural photosynthetic systems [2], many photocatalytic platforms have been developed, yet the most efficient still rely on noble-metal photosensitizers, whose scarcity, cost, and limited stability hinder scalability [3]. Fully organic photosensitizers based on thermally activated delayed fluorescence (TADF) have recently emerged as a promising alternative. These molecules are easily synthesized and combine long-lived excited states with highly tunable electronic structures, enabling efficient light harvesting and photoinduced electron transfer—essential for photoassisted CO2 reduction [4]. Their purely organic nature offers new opportunities for sustainable, metal-free photocatalytic systems. In this work, we systematically screened a series of organic TADF photosensitizers in homogeneous CO2 reduction systems, evaluating how variations in key reaction parameters influence catalytic performance. Optimization was guided by a detailed analysis of the absorption properties of individual components and solutions using UV–Vis spectroscopy. This approach, combined with steady-state and time-resolved photophysical measurements, aims to lay the groundwork for correlating photophysical properties with activity and stability, ultimately identifying the most efficient system [5] with the support of theoretical modeling to further advance this objective. Our results highlight TADF chromophores as efficient metal-free photosensitizers for photoassisted CO2 reduction, aiming to bridge photophysics and catalysis for next-generation artificial photosynthetic systems.

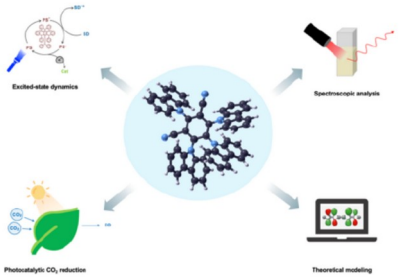


Figure 1. Overview of system testing and photophysical analysis to guide the design of efficient

References:

[1] IEA. Global energy review 2025. IEA, Paris, 2025.

[2] F. Droghetti, L. Villa, A. Sartorel, L. Dell’Amico, A. Ruggi, and M. Natali. ChemSusChem, 18:e202402627, 2025.

[3] L. Eisele, W. Chaikhan, S. Batool, A. Cherevan, D. Eder, and K. Bica-Schröder. ChemCatChem, 16:e202301454, 2024.

[4] A. Gualandi, M. Anselmi, F. Calogero, S. Potenti, E. Bassan, P. Ceroni, and P. G. Cozzi. Org. Biomol. Chem., 19:3527, 2021.

[5] F. Zhang, Y. Jiang, J. Liu, A. Jiang, Y. Cao, S. Yu, K. Zheng, and Y. Zhou. Fundamental Research, 5(6):2838–2849, 2025.

Poster 19

Exploring Excited-State Energy Landscapes Across Chemical Compound Space with Machine Learning: The Example of Green Fluorescent Protein Chromophores

Sophia Wesely[1], Rhyan Barrett[1], Julia Westermayr[1]

[1] Wilhelm-Ostwald-Institute for Physical and Theoretical Chemistry, Leipzig University

Understanding excited-state energy landscapes is central to designing photoreactive molecules for applications in bioimaging, optogenetics, or photopharmacology. The green fluorescent protein (GFP) chromophore, p-hydroxybenzylidene-2,3-dimethylimidazolinone anion ($HBDI^-$), is a model system for excited-state processes, but presents significant challenges for theory due to long-lived excited states, competing relaxation pathways, and the prohibitive cost of accurate quantum chemical simulations [1]. These factors have traditionally made systematic exploration of functional group modifications intractable. In this work, we overcome these limitations by combining our recently developed equivariant machine learning model for excited states, X-MACE [2], with nonadiabatic molecular dynamics. Trained on thousands of organic chromophores, X-MACE generates potential energy surfaces transferable across chemical space and electronic states with minimal additional data. This approach enables, for the first time, efficient screening of $HBDI^-$ derivatives to assess how structural changes modulate excited-state topologies and dynamics. Our results uncover design principles for tuning photophysical properties within chromophore compound space through targeted substitutions, demonstrating the potential of machine learning to unlock excited-state landscape exploration at scale.

References:

[1] N. H. List; C. M. Jones; T. Martínez, J. Commun. Chem. 2024 7, 25.

[2] R. Barrett; C. Ortner; J. Westermayr, arXiv prePrint 2025 arXiv:2502.12870.

Poster 20

From Analytical Microchips towards Self-Driving Chemistry – Experimental Challenges in the Age of Digitalisation

Hannes Westphal[1]

[1] Wilhelm-Ostwald-Institute for Physical and Theoretical Chemistry, Leipzig University

Chemistry is shifting from trial-and-error experimentation towards data-driven and increasingly autonomous discovery. Yet progress is often limited by the experimental bottleneck of producing consistent, traceable and machine-readable data at sufficient scale. Self-driving laboratories (SDL) tackle this challenge by closing the loop between automated experiments, integrated analytics, and data-driven decisions. Microfluidic lab-on-a-chip platforms offer ideal experimental units for such systems by enabling controlled reaction studies with subsequent analyses, while reducing reagent consumption, waste, and the risks associated with early-stage screening. The presented workflows demonstrate how automated and integrated miniaturised reaction platforms coupled to mass spectrometry (MS) can provide a practical experimental foundation for future ML-guided, closed-loop discovery systems.

Poster 21

Enhancing Product Specificity of Chalcone Synthase Using Natural Deep Eutectic Solvents

Mohammad Yousefi[1], Ludger Wessjohann[1], Martin Dippe[1*], Mehdi D. Davari[1*]

[1] Department of Bioorganic Chemistry, Leibniz Institute of Plant Biochemistry, 06120 Halle, Germany

Maintaining product specificity when enzymes are transferred from in vivo to in vitro conditions remains a central challenge in biocatalysis. Chalcone synthase (CHS) catalyzes the first committed step of plant flavonoid biosynthesis and serves as an attractive biocatalyst for producing flavonoid precursors and non-natural polyketide scaffolds, including chalcone analogues, benzophenone derivatives, and pyrone-type products. Under in vitro conditions, however, CHS produces substantial amounts of undesired products, notably p-coumaroyltriacetic acid lactone (CTAL). Here, we investigate natural deep eutectic solvents (NADESs) as green, attractive, non-conventional reaction media for improving CHS product specificity. Our preliminary results indicate that aqueous NADES substantially enhance the stability and specificity of CHS from Triticum aestivum (TaCHS) toward the desired chalcone pathway. Complementary quantum mechanical/molecular mechanical (QM/MM) simulations provide insight into how TaCHS control the product specificity. In addition, MD simulations revealed how interactions between NADES components and the enzyme alter enzyme dynamics and conformation, thereby enhancing product specificity and stability. Together, these findings suggest that media engineering offers a promising strategy for fine-tuning enzyme behavior and improving the selectivity of biocatalytic processes.

Poster 22

Another diamine cation challenges quantum chemistry methodology

Neda Zolghadri Jahromi[1], Marta Gałynska[1]

[1] Nicolaus Copernicus University in Toruń

Poster 23

Assessing the performance of quantum-mechanical descriptors in physicochemical and biological property prediction

Alejandra Irene Hinostroza Caldas[1]

[1] TU Dresden

Understanding how molecular structure determines physicochemical and biological properties is a central challenge in computer-aided drug design. A key obstacle for machine learning (ML) approaches is the development of molecular representations that capture both geometric and electronic characteristics. We present the QUantum Electronic Descriptor (QUED) framework [1], which combines geometric representations, including BOB and SOAP, with a quantum-mechanical representation encoding global and local electronic properties computed efficiently using the DFTB method. Applied to the prediction of biological properties, QUED shows strong predictive performance across several Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) endpoints, including acute toxicity (LD50) and lipophilicity (logD). These results highlight the importance of incorporating interpretable electronic-structure information alongside geometric features and demonstrate the value of hybrid representations for molecular property prediction.

References:

[1] Digital Discovery, 2026, 5, 803–818.