CHIMIA

CHIMIA 2024, Volume 78
ISSN 0009-4293
www.chimia.ch
Supplementa to Issue 7-8/2024

SCS
Swiss Chemical
Society

SCS Fall Meeting 2024
Lecture, Short Talk and Poster Abstracts

Computational Chemistry

September 5, 2024
University of Fribourg, Campus Pérolles
https://fm24.scg.ch

Swiss Chemical Society
Haus der Akademien
Postfach
3001 Bern
Switzerland
info@scg.ch
www.scg.ch

Simulating non-adiabatic dynamics in molecules: methodological aspects and attochemical applications

M. Vacher[1]

[1]Nantes Université, CNRS, CEISAM, UMR 6230, F-44000 Nantes, France - morgane.vacher@univ-nantes.fr

Non-adiabatic dynamics simulations are often key to the understanding of the mechanism, rate and yield of photochemical reactions. The most popular method to simulate photochemical processes in molecules is the mixed quantum-classical surface hopping method. In the first part, I will discuss different variants of surface hopping using different approximations to the non-adiabatic couplings. Our simulations identified two variants, the local diabatization scheme and a scheme based on biorthonormal wave function overlap, which reproduce at a much-reduced cost the reference dynamics. The other two schemes tested can give different results, and in some cases, even entirely incorrect dynamics [1,2,3].

Nonadiabatic dynamics simulations require underlying potential energy surfaces and couplings between them, calculated at a chosen level of theory. In the second part of my talk, I will focus on assessing the performances of various post-Hartree−Fock methods (namely, CIS, ADC(2), CC2, CASSCF and CASPT2) and exchange−correlation functionals (PBE, PBE0, and CAM-B3LYP) in a TD-DFT/TDA context, using the isomerization around a double bond as test case (Figure 1, left). Different relaxation pathways are identified, and the ability of the different methods to reproduce their relative importance and time scale is discussed [4]. In the last part, I will open the discussion to the simulation of attochemistry. Such molecular processes can, in principle, be simulated with various nonadiabatic dynamics methods, yet the impact of the approximations underlying the methods is rarely assessed. The performances of widely used mixed quantum-classical approaches, Tully surface hopping, and classical Ehrenfest methods are evaluated against the high-accuracy DD-vMCG quantum dynamics. This comparison is conducted for the valence ionization of fluorobenzene (Figure 1, right) [5].

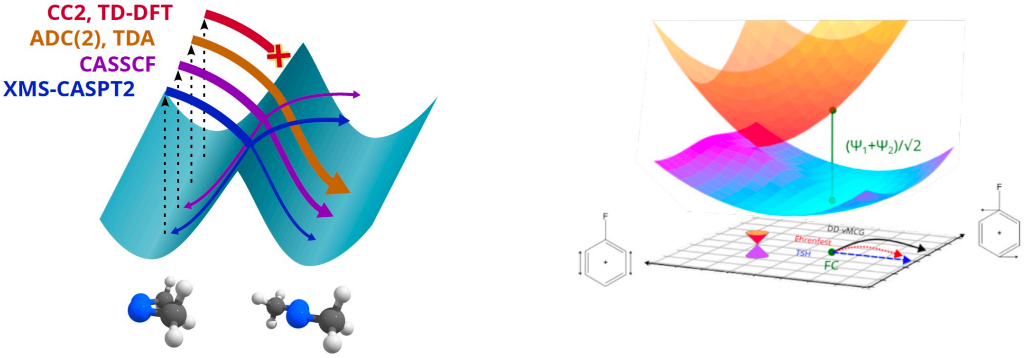


Figure 1: (left) impact of electronic structure on dynamics and (right) benchmark of dynamics methods for attochemistry.

[1] Merritt, Jacquemin and Vacher, *J. Chem. Theory Comput.*, **2023**, *19*, 1827-1842.
[2] Zhao, Merritt, Lei, Shu, Jacquemin, Zhang, Xu, Vacher and Truhlar, *J. Chem. Theory Comput.*, **2023**, *19*, 6577-6588.
[3] Li Manni et al, *J. Chem. Theory Comput.*, **2023**, *19*, 6933-6991.
[4] Papineau, Jacquemin and Vacher, **J. Phys. Chem. Lett.**, **2024**, *15*, 636-643.
[5] Tran, Fert. and Vacher, *J. Phys. Chem. Lett.*, **2024**, *15*, 3646-3652.

From capture to catalysis: Insights from atomistic simulations into MOF and COF host-guest interactions and guest dynamics

M. Ernst[1], J. Hutter[1]*

[1]University of Zurich, Department of Chemistry

A key feature of metal-organic frameworks (MOFs) and covalent organic frameworks (COFs) is their ability to capture, convert, and release guest molecules. This capability is governed by multiple factors such as pore volume and surface area, chemical environment, and the nature of the host-guest intermolecular interactions. X-ray diffraction (XRD) experiments are essential to resolve the atomic structures of the frameworks and, ideally, also the guests. However, while XRD provides irreplaceable structural information, it does not offer insights into the dynamic behaviour of guest molecules or the detailed electronic structure of the host-guest interactions. To bridge this gap, we employ various atomistic simulations: molecular docking to determine adsorption sites, DFT simulations to compute interaction energies, electron density and energy decomposition schemes to characterize the host-guest interactions, and ab initio molecular dynamics to simultaneously capture the precise electronic structure and the dynamics of the investigated processes.

First, we will discuss the correlation between simulated guest positions and those observed via XRD [1]. We will also show the characterization of host-guest interactions [2] and compare these interaction characteristics to experimental observations from drug release studies. This will highlight the promise of MOFs and COFs in drug delivery while also demonstrating the potential of simulations to predict effective drug delivery materials [3].

Following this, we will discuss COFs as catalysts. Catalysis within confined spaces often outperforms that in solution or on the surfaces of conventional heterogeneous catalysts, which can be attributed to the unique molecular interactions and reaction dynamics introduced by the spatial constraints. Results from AIMD calculations will be presented, illustrating the effect of confinement in a COF pore on a specific organic reaction, namely the Diels-Alder reaction (Figure).

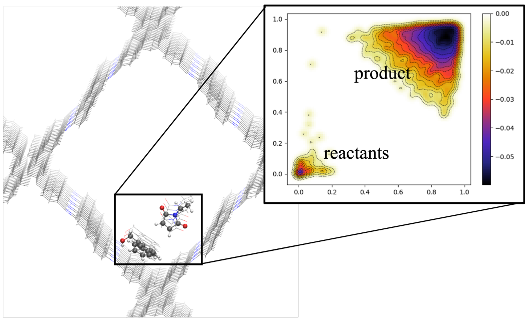


Figure Diels-Alder reactants in COF pore and corresponding simulated free energy surface (Hartree).

[1] M. Ernst, T. Por?ba, L. Gnägi, G. Gryn'ova, Journal of Physical Chemistry C, 2023 127, 523-531.

[2] M. Ernst, G. Gryn'ova, ChemPhysChem, 2022 23, e202200098.

[3] M. Ernst, G. Gryn'ova, Helvetica Chimica Acta, 2023, 106, e202300013.

Unravelling Tunnelling Effects in Photosensitization using Instanton Theory

M. Manae[1], J. Richardson[1]*

[1]Department of Chemistry and Applied Biosciences, ETH Zürich

Photosensitization of O_2 is widely made use of in organic synthesis, photovoltaic cells and is a key step in photodynamic therapy. Despite its significance, its comprehensive mechanism is still unknown.

$$^3PS + {}^3O_2 \rightarrow {}^1PS + {}^1O_2$$

A previous theoretical study calculated O_2-photosensitization rates by thiothymines (2-thiothymine, 4-thiothymine and 2,4-dithiothymine) and found a speed up by several orders of magnitude when rates were calculated from Fermi's golden rule (FGR) compared to Marcus theory, thus establishing the importance of quantum effects in driving photosensitization [1]. To understand more about the mechanism however, a quantum approach would be untenable. The gap between the need to account for quantum effects and computational cost can be bridged by semiclassical instanton theory (SCI), in particular, its extension to the inverted regime in the golden-rule limit [2]. SCI locates the optimal tunnelling pathway which provides mechanistic insights, and calculated rate constants are often in excellent agreement with experiment [3]. For O_2-photosensitization, we start by extending SCI to account for geometry-dependent coupling, which is a feature of photosensitization. This method is tested on a model system and we obtained perfect agreement with exact (FGR) rates [4]. Then, before applying the developed methodology to a real system, we take a closer look at the intersystem crossing (ISC) mechanism in three potential photosensitizers – the thiouracils (molecules closely related to thiothymines). This is because the mode which promotes photosensitization has been hypothesised to be the one along which ISC of the photosensitizer takes place [1]. Our calculations have shown evidence of room-temperature heavy-atom tunnelling in the ISC pathways of 4-thiouracil and 2,4-dithiouracil, while the tunneling effect in 2-thiouracil is negligible. Interestingly, we were able to explain why 2-thiouracil has a shorter triplet lifetime compared to 2-thiothymine, despite being very similar molecules. We then discuss extending the calculations to the O_2-photosensitization reaction by thiouracils.

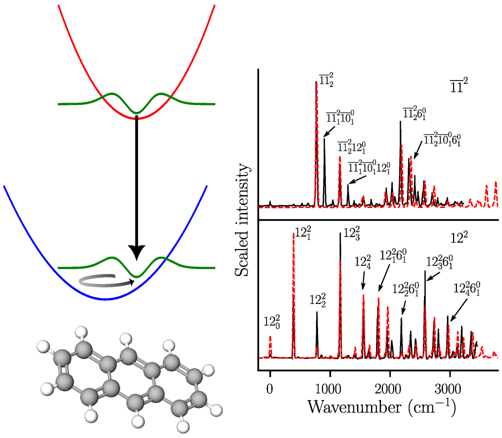
[1] Meghna A. Manae and Anirban Hazra, *Physical Chemistry Chemical Physics*, **2022**, *24*, 13266-13274.
[2] Eric R. Heller and Jeremy O. Richardson, *The Journal of Chemical Physics*, **2020**, *152*, 034106.
[3] Eric R. Heller and Jeremy O. Richardson, *Journal of the American Chemical Society*, **2021**, *143*, 20952-20961.
[4] Meghna A. Manae and Jeremy O. Richardson, *Chimia*, **2024**, *78*, 231-233.

Ab initio simulation of single vibronic level fluorescence spectra using Hagedorn wavepacket

Z. Zhang[1], J. Vanicek[1]*

[1]Laboratory of Theoretical Physical Chemistry, Institut des Sciences et Ingénierie Chimiques, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

Single vibronic level (SVL) fluorescence spectroscopy contributes to the understanding of molecular vibrational structures and relaxation processes. Here, we present a practical method for computing SVL fluorescence spectra of polyatomic molecules from arbitrary initial vibrational levels. Here, we use Hagedorn functions [1,2], which are products of a Gaussian and carefully generated polynomials, to represent SVL initial states. In systems where the potential is at most quadratic, the Hagedorn functions are exact solutions to the time-dependent Schrödinger equation and can be propagated with the same equations of motion as a simple Gaussian wavepacket. Having developed an efficient recursive algorithm to compute the overlaps between two Hagedorn wavepackets [3], we can now evaluate emission spectra from arbitrary vibronic levels using a single wavepacket trajectory [4]. We apply the method to compute SVL spectra of anthracene by performing wavepacket dynamics on a 66-dimensional harmonic potential energy surface constructed from density functional theory calculations [5]. With the Hagedorn approach, we not only reproduce the previously reported simulation results for singly excited 12^1 and 11^1 levels [6], but also are able to compute SVL spectra from multiply excited levels in good agreement with experiments.



[1] G. A. Hagedorn, *Annals of Physics*, **1998**, 269 (1), 77–104

[2] C. Lasser, C. Lubich, *Acta Numerica*, **2020**, 29, 229–401

[3] J. J. L. Vaníček, Z. T. Zhang, arXiv:2405.07880 [quant-ph], **2024**

[4] Z. T. Zhang, J. J. L. Vaníček, arXiv:2403.00577 [physics.chem-ph], **2024**

[5] Z. T. Zhang, J. J. L. Vaníček, arXiv:2403.00702 [physics.chem-ph], **2024**

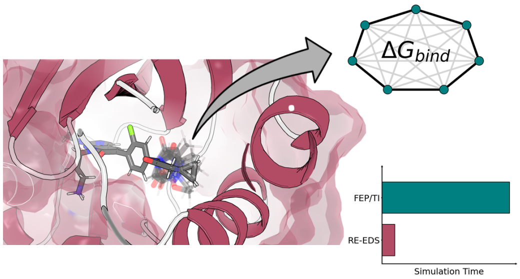
[6] E. Tapavicza, *Journal of Physical Chemistry Letters*, **2019**, 10, 20, 6003–6009

Accelerating Alchemical Free-Energy Calculation with the RE-EDS Multistate Method

C. Champion[1], R. Gall[1], B. Ries[1], S. Rieder[2], E. Pécora de Barros[1], P. Hünenberger[1]*, S. Riniker[1]*

[1]ETH Zurich, [2] ETH Zurich

Alchemical free-energy methods based on molecular dynamics (MD) simulations have become important tools to help prioritize compounds for synthesis during lead optimization. State-of-the-art methods (FEP/TI) evaluate pairwise free-energy differences, thus limiting their scalability when comparing N different ligands. To address this fundamental limitation, our group has developed the multistate method replica-exchange enveloping distribution sampling (RE-EDS), which enables the calculation of free-energy differences between multiple ligands (N > 2) from a single MD simulation. First, we demonstrate the general applicability of RE-EDS for drug-design campaigns by benchmarking the method on a set of four kinases with small molecule inhibitors, highlighting the decrease in computational cost when compared to pairwise methods. Second, we present an application of RE-EDS to molecules which exist in a mixture of protonation/tautomeric states under physiological conditions, underlining the importance of including both to predict the binding affinities of those compounds.



Automated Chemical Reaction Network Exploration of Oxidative Water Treatment

E. Petrus[1,2], U. von Gunten[1,3], M. Reiher[2], T. B. Hofstetter[1,4]*

[1]Swiss Federal Institute of Aquatic Science and Technology (Eawag), [2]Institute for Molecular Physical Science, ETH Zürich, Switzerland, [3]School of Architecture, Civil and Environmental Engineering (ENAC), EPF Lausanne, Switzerland, [4] Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, Switzerland

Chemical oxidants such as ozone, chlorine, and chlorine dioxide are key for eliminating waterborne diseases as well as for the abatement of micropollutants during (waste)water treatment [1]. However, chemical oxidants lead to the formation of transformation products (TPs) from micropollutants and undergo side reactions with water matrix constituents that can give rise to unknown and potentially toxic disinfection byproducts (DBP). Because a comprehensive experimental characterization of the complex reaction networks is extremely laborious and inherently incomplete, autonomous and unsupervised computational approaches are required to infer probable DBP formation pathways comprehensively. Here, we address the question of how to elucidate the oxidative water treatment chemistry and thus the intricate TPs/DBPs formation paths using autonomous quantum chemical reaction network (CRN) explorations. Because ab initio CRN explorations for the aqueous chemistry of chemical oxidants and their many reactive intermediates have not been attempted so far, this task involved (i) determining the quantum chemical methodology to balance computational efficiency and chemical accuracy, and (ii) identifying reliable benchmarks for the evaluation of computational simulations against experimental data. To that end, this study focused on the reactions of ozone with olefins given the paramount importance of ozone for water and wastewater treatment. Extensive previous electronic structure calculations and experimental kinetic data for the ozone reactions with olefins provide a reliable basis for the evaluation of the CRN exploration outcomes.

To identify the appropriate quantum chemical methodology and provide accurate structures and reaction energies for this exploration, we performed a benchmark study that focused on the 1,3-dipolar cycloaddition of ozone to ethene. Among the various semi-empirical and DFT methods evaluated, the LC-PBE functional provided the lowest RMSD compared to reference energy data published by Wheeler and coworkers [3]. This long-range corrected functional ensures an accurate description of zwitterionic compounds, which are frequent intermediates in reactions of ozone with olefins. However, no density functional considered provided reliable electronic energies; activation energy deviations were above 4 kcal·mol. By contrast, the (approximate) DLPNO-CCSD(T) formalism offered high accuracy, at comparatively low computational costs. With this methodological basis, we constructed a network of 1,244 compounds and 6,703 reactions for ethene and ozone in the gas phase using the Chemoton reaction exploration code [4]. To model the reactivity in the aqueous phase, we considered the implicit solvation model CPCM which was employed to reoptimize the molecular structures and recalculate the electronic energies of the entire reaction network. A recent implementation of a shortest-path algorithm [5] enabled the identification of the well-established Criegee mechanism as part of the computed reaction network. Microkinetic simulations in aqueous solution not only predicted the yield of the two main reaction products (formaldehyde, methyl-hydroxyperoxide) correctly, but also provided meaningful lifetimes of the ozonide intermediate [6]. This successful first autonomous CRN exploration of oxidative water treatment holds great promise for future applications that will target the identification of the currently unknown reaction paths to TPs and/or DBPs.

[1] U. von Gunten Environ. Sci. Technol. 2018, 52, 5062−5075.
[2] P. R. Tentscher, M. Lee, U. von Gunten Acc. Chem. Res.2019, 52, 3, 605–614
[3] S. E. Wheeler, D. H. Ess, K. N. Houk J. Phys. Chem. A 2008, 112, 8, 1798–1807
[4] J. P. Unsleber, S. A. Grimmel, M. Reiher J. Chem. Theory. Comput. 2022, 18, 5393-5409
[5] P. M. Türtscher, M. Reiher J. Chem. Inf. Model. 2023, 63, 1, 147-160
[6] P. Dowideit, C. von Sonntag Environ. Sci. Technol. 1998, 32, 1112-1119

Computational approaches in pharma on chemical reactivity and affinity optimization

H. Van Vlijmen[1]

[1]Johnson & Johnson Innovative Medicine R&D, Turnhoutseweg 30, 2340 Beerse, Belgium

Computational approaches are used for a wide variety of applications in the pharmaceutical industry. Here we will describe the use of quantum calculations to address the topics of chemical stability and optimization of synthesis in drug discovery. It will be shown that experiments and computations combined provide critical molecular level knowledge that cannot be obtained by either one alone.

A second topic will focus on the use of free energy perturbation calculations to estimate the affinity of compounds for a protein target by molecular dynamics simulations using a classical force field. The current level of accuracy of this method enables the selection of the most promising compound candidates for synthesis. Recent developments in compute power and methodological improvements have made these calculations a cornerstone of the early drug discovery process.

Applying a well-defined energy density for machine-learned density functionals

E. Polak[1], S. Vuckovic[1], H. Zhao[1]

[1]University of Fribourg

The recent integration of machine learning techniques in density functional theory (DFT) has established a powerful framework for developing next generation density functionals. While robust modelling of the exchange-correlation requires a well-defined energy density, conventional training sets usually rely on global quantities. We propose the application of the correlation energy density from second order perturbation theory that corresponds to the local slope in the non-interacting limit of the adiabatic connection approach in DFT [1]. The presentation will elucidate the methods for an efficient implementation of this quantity, with a focus on its spin-resolved components and its regularized version. Furthermore, we will highlight the potential of this strategy in paving the way for the next generation of machine-learned local dynamic hybrid functionals. Our results show a marked improvement in the prediction of observables while also maintaining computational efficiency.

[1] Stefan Vuckovic, Tom J. P. Irons, Andreas Savin, Andrew M. Teale, Paola Gori-Giorgi, *Journal of Chemical Theory and Computation,* **2016**, 12, 2598-2610.

Inverse Design of Singlet Fission Materials with Uncertainty-Controlled Genetic Optimization

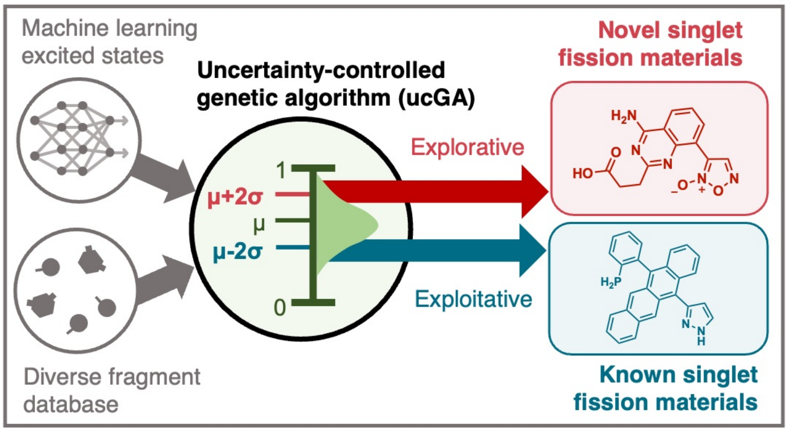
L. Schaufelberger[1], R. Laplaza[1], J. T. Blaskovits[1], K. Jorner[2]*, C. Corminboeuf[1]*

[1]EPFL Lausanne, Institute of Chemical Sciences and Engineering, [2]ETH Zürich, Department of Chemistry and Applied Bioscience

Singlet fission has shown potential for boosting the power conversion efficiency of solar cells, but the scarcity of suitable molecular materials hinders its implementation [1]. In this presentation, we introduce an uncertainty-controlled genetic algorithm (ucGA) built on ensemble machine learning (ML) predictions to discover singlet fission candidates. As conventional ML-driven setups can not control how far genetic optimization [2] moves away from previously known molecules (exploitation vs. exploration), our method builds on uncertainty-quantification to direct the genetic algorithm through chemical space. We first devise an interpretable excited state machine learning model incorporating diverse molecular representations, which was trained on 117k molecules characterized with time-dependent density functional theory [3]. Our multiobjective ucGA pipeline then concurrently optimizes predicted excited singlet and triplet energies, synthesizability, and singlet exciton size for the discovery of singlet fission materials.

Running the ucGA in an exploitative setup performs local optimization on variations of known singlet fission scaffolds, validating our approach. In an explorative mode, we generate new candidates displaying excellent quantum-chemical properties for singlet fission. We identify a class of heteroatom-rich mesoionic compounds as acceptors for charge-transfer mediated singlet fission, which show strong triplet localization and open-shell singlet characteristics appropriate for singlet fission. As the proposed candidates are composed of fragments from synthesized molecules, they are likely synthetically accessible.

The method presented in this talk is applicable to a variety of molecular design tasks, with our explorative methodology especially useful to avoid convergence towards already known and well-studied motifs, allowing the design of novel materials while addressing biases in datasets.



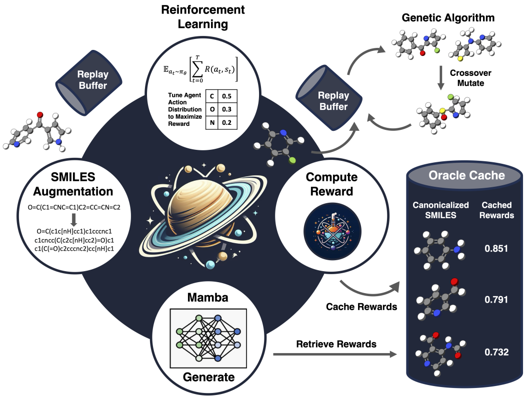
[1] Casillas, R. N.; Papadopoulos, I.; Ullrich, T.; Thiel, D.; Kunzmann, A.; Guldi, D. M. Energy Environ. Sci. 2020, 13, 2741.
[2] Laplaza, R.; Gallarati, S.; Corminboeuf, C. Genetic Optimization of Homogeneous Catalysts. Chemistry–Methods 2022, 2, e202100107.
[3] Blaskovits, J. T.; Laplaza, R.; Vela, S.; Corminboeuf, C. Data-Driven Discovery of Organic Electronic Materials Enabled by Hybrid Top-Down/Bottom-Up Design. Adv. Mater. 2024, 36, 2305602.

Saturn: Sample-efficient Generative Molecular Design using Memory Manipulation

J. Guo[1,2], P. Schwaller[1,2]*

[1]EPFL LIAC, Department of Chemistry and Chemical Engineering, [2]NCCR Catalysis

Generative molecular design for drug discovery has very recently achieved a wave of experimental validation, with language-based backbones being the most common architectures used. The most important factor for downstream success is whether an *in silico* oracle is well correlated with the desired end-point. To this end, current methods use cheaper proxy oracles with higher throughput before evaluating the most promising subset with high-fidelity oracles. The ability to *directly* optimize high-fidelity oracles would greatly enhance generative design and be expected to improve hit rates. However, current models are not efficient enough to consider such a prospect, epitomizing the sample efficiency problem. In this work, we introduce **Saturn**, which leverages the Augmented Memory algorithm and is the first application of the Mamba architecture for generative molecular design. We elucidate *why* Augmented Memory, which combines experience replay with data augmentation, improves sample efficiency and *how* Mamba synergistically exploits this mechanism. Saturn outperforms 22 models on multi-parameter optimization tasks relevant to drug discovery and may possess sufficient sample efficiency to consider the prospect of directly optimizing high-fidelity oracles.



[1] Guo, J., Schwaller, P. JACS Au, ASAP.

[2] Guo, J., Schwaller P. ICLR 2024.

A Metadynamics Study of Water Oxidation Reactions at (001)-WO_3/liquid-water Interface

R. Ketkaew[1], F. Creazzo[1], K. Sivula[2], S. Luber[1]

[1]Department of Chemistry, University of Zurich, Switzerland, [2]Institute of Chemical Sciences and Engineering (ISIC), EPFL, Switzerland

We computationally study the oxygen evolution reaction (OER) mechanism on the (001) WO_3 surface in the condensed phase (liquid water). We also investigate alternative mechanistic pathways to water oxidation, namely, hydroxyl radical ($OH^{\bullet}$) and hydrogen peroxide (H_2O_2) formations, which are of high interest since they are known to be favored over the OER on the (001) WO_3 surface according to experimental evidence. The metadynamics (MetaD) method within the framework of density functional theory-based molecular dynamics (DFT-MD) is used to calculate the free energy surfaces (FESs) of the OER. eXtended Social PeRmutation INvarianT (xSPRINT) coordinates and deep autoencoder neural network (DAENN) are used as collective variables (CVs) for MetaD to characterize the FES of the studied reactions. The MetaD simulation suggests that the oxygen-oxygen (O – O) bond formation induced by oxidative reactant species and oxygen atoms on the surface is the rate-determining step. The FES calculations using xSPRINT and DAENN CVs show that the formation of $OH^{\bullet}$ and H_2O_2, respectively, is more energetically (in terms of the free-energy barrier) favorable than the formation of oxygen gas at the WO_3 surface/liquid water interface. It is found that the O – O bond formation is sluggish because of the stable electronic state of the surface and that the formation of H_2O_2 occurs through an energy barrierless process.

[1] Rangsiman Ketkaew, Fabrizio Creazzo, Kevin Sivula, Sandra Luber, Chem Catalysis, 2024.

[2] Rangsiman Ketkaew, Fabrizio Creazzo, Sandra Luber, J. Phys. Chem. Lett., 2022, 13, 7, 1797–1805.

[3] Rangsiman Ketkaew, Sandra Luber, J. Chem. Inf. Model., 2022, 62, 24, 6352–6364.

Global and Local Electrophilicity as Measure of Endocrine Disruptor Genotoxicity Potential

S. Johnson[1], S. Guilbert[1], U. Rothlisberger[1]*

[1]École Polytechnique Fédérale de Lausanne

Exogenous endocrine disrupting chemicals (EDCs) bind to nuclear hormone receptors and are implicated in hormone-sensitive cancer development. Reactive metabolites of some EDCs form stable and depurinating adducts with DNA [1, 2] with the key step of adduct formation being a 1,4-Michael addition between a nucleophile (DNA) and electrophile (EDC) [3]. To evaluate potential genotoxicity of neutral EDC metabolites, we employ a conceptual DFT study (B3LYP-D3BJ/def2-TZVPD) to evaluate their global electrophilicity values via:

$$\omega \equiv \mu^2/2\eta \text{ where } \mu = (\partial E/\partial N)_{v(r)} \text{ and } \eta = (\partial^2 E/\partial N^2)_{v(r)}$$

We discover that many estrogenic EDC metabolites have comparable global electrophilicity values to endogenous, reactive estrogen quinone metabolites (see Fig. 1). We evaluate local electrophilicity using a Fukui function approach to identify potential covalent bond formation sites. Our classical and quantum mechanics/molecular molecular dynamics simulations with thermodynamic integration show an essentially barrierless approach of several metabolites into minor and major DNA grooves under biological conditions suggesting that, once present in the nucleus, such metabolites might easily locate at a favorable position along the DNA for covalent bonding.

Figure 1: Global electrophilicity values of selected quinone and non-quinone metabolites of known EDCs (several 2D metabolite structures shown on sides).

[1] E. L. Cavalieri, E.G. Rogan, Clinical and Translational Medicine 2016, 5, 12.

[2] N. Jalal, A. R. Surendranath, J. L. Pathak, S. Yu, C. Y. Chung, Toxicology Reports 2018, 5, 76-84.

[3] D. E. Stack, Metabolites 2015, 5, 475-488.

Heavy-atom tunnelling in singlet oxygen predicted with ab-initio instanton theory

I. Ansari[1], G. Trenins[1], J. O. Richardson[1], E. R. Heller[1]*

[1]Department of Chemistry and Applied Biosciences, ETH Zürich, 8093 Zürich, Switzerland

The reactive singlet state of the oxygen molecule decays nonradiatively in solution to its triplet state [1, 2]. The rate constant for this decay, which is a spin-crossover process (intersystem crossing) in the Marcus inverted regime, determines its lifetime and can in principle be calculated using Fermi's Golden Rule (FGR). However, it would require solutions to the nuclear Schrödinger equation and an exact treatment is thus computationally infeasible. Nonadiabatic instanton theory [3, 4] is a semiclassical approximation to FGR that captures multidimensional tunnelling and zeropoint energy effects by locating the optimal tunnelling path (reaction mechanism), called the instanton. This system is in the deep inverted regime and exhibits a completely new class of quantum tunnelling, which requires a new type of instanton theory. The tunnelling mechanism is found to comprise of an infinite ensemble of equally-likely paths [5]. Using this new theory combined with multireference electronic-structure calculations, we predict a rate constant with an enormous tunnelling effect— 27 orders of magnitude faster than the classical process—and which involves significant heavy-atom tunnelling contributions even at room temperature. The theory also correctly predicts the large experimentally-observed H2O/D2O kinetic isotope effect of approximately 20 and its temperature dependence.

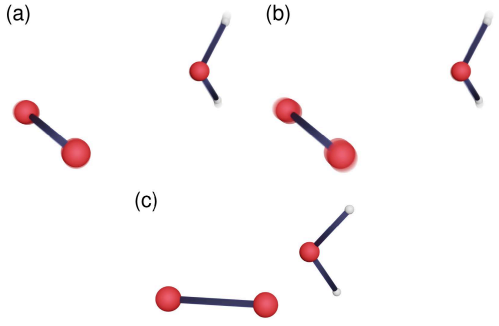


FIG. 1: The tunnelling mechanism for the decay of singlet oxygen in (a) water and (b) deuterated water, and (c) the geometry at the minimum-energy crossing point (MECP). The motion of the atoms involved in the tunnelling mechanism in (a) and (b) is depicted by a motion blur.

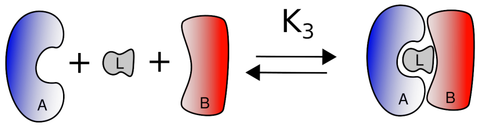
[1] C. Schweitzer and R. Schmidt, Physical mechanisms of generation and deactivation of singlet oxygen, *Chem. Rev.* **103**, 1685 (2003), https://doi.org/10.1021/cr010371d.
[2] M. Bregnhøj, M. Westberg, B. F. Minaev, and P. R. Ogilby, Singlet oxygen photophysics in liquid solvents: Converging on a unified picture, *Acc. Chem. Res.* 50, **1920** (2017), https://doi.org/10.1021/acs.accounts.7b00169.
[3] E. R. Heller and J. O. Richardson, Instanton formulation of Fermi's golden rule in the Marcus inverted regime, *J. Chem. Phys*. 152, 034106 (**2020**), arXiv:1911.06730 [physics.chem-ph].
[4] I. M. Ansari, E. R. Heller, G. Trenins, and J. O. Richardson, Instanton theory for Fermi's golden rule and beyond, *Phil. Trans. R. Soc. A*. 380, 20200378 (**2022**).
[5] I. M. Ansari, E. R. Heller, G. Trenins, and J. O. Richardson, Heavy-atom tunnelling in singlet oxygen deactivation predicted by instanton theory with branch-point singularities, *Nat Commun* 15, 4335, publisher: Nature Publishing Group.

Cooperative free energy: correlation, solvation, and conformation in protein-ligand-protein ternary complexation

S. Chen[1], R. Solazzo[1], M. Fouche[2], H. Roth[2], B. Dittrich[2], S. Riniker[1]*

[1]Department of Chemistry and Applied Biosciences, ETH Zurich, Vladimir-Prelog-Weg 2, Zurich, 8093, Switzerland, [2]
Novartis Institutes for BioMedical Research, Basel Switzerland

Cooperativity in the protein-ligand-protein ternary complex regulates the complexation efficiency and selectivity[1-2]. In terms of free energy, cooperativity can be decomposed into three three-body contributions. The gas-phase entropy part accounts for the correlation between the protein-ligand interactions, the solvation-free energy part accounts for the double-counted desolvation effect at the three-body interface, and the conformational free energy part accounts for the induced protein-protein interactions and the conformational penalty to bring the ligand, ligand-protein dimers, and protein-ligand-protein trimer to the stabilized conformational space. Under the theoretical framework, molecular dynamics simulations can be implemented to predict the cooperativity of the systems of interest. The study discusses the thermodynamics of the cooperative process, demonstrates *in silico* cooperative prediction, and is anticipated to facilitate the design of cooperative molecular glues.



$$-\beta\log(K_3) = \Delta G_{A\text{-}L}+\Delta G_{B\text{-}L}+\Delta G_{A\text{-}B}+\Delta G_{\alpha_0}$$

$$\Delta G_{\alpha_0}=\Delta G'_{\alpha_0,conf}+\Delta G'_{\alpha_0,solv}+\Delta G'_{\alpha_0,gas}$$

Induced Conformation

Solvation

Interaction Correlation

[1] Madita Wolter, Pim de Vink, João Filipe Neves, Sonja Srdanović, Yusuke Higuchi, Nobuo Kato,Andrew Wilson, Isabelle Landrieu, Luc Brunsveld, Christian Ottmann, J. Am. Chem. Soc. **2020**, 142, 27, 11772–11783

[2] Claire E. Grigglestone, Kap-Sun Yeung, ACS Med. Chem. Lett. **2021**, 12, 11, 1629–1632

Structure Determination of Single Crystal Quartz by Hirshfeld Atom Refinement

K. Chu[1], Y. Balmohammadi[1], S. Grabowsky[1]*

[1]DCBP, University of Bern, 3012 Bern, Switzerland

Hirshfeld atom refinement [1,2] is an advanced technique that determines structural parameters from single-crystal X-ray diffraction data and aspherical atomic form factors calculated through *ab initio* quantum chemistry calculation and partitioning of electron density. Despite its success in molecular crystal systems, applying it to network compounds such as quartz has been challenging [3], because their unit cells have no stable free-standing configuration.

In this presentation, we introduce the development and application of a new periodic Hirshfeld atom refinement using the quartz system as the main example. This was achieved by combining the open-source quantum crystallography program *Tonto* with the commercially available package for periodic calculation *Crystal23* through the interface software *lamaGOET* [4]. A self-consistent model structure of quartz was obtained through an iterative process alternating between *Crystal23* for crystal electron density calculations and *Tonto* for electron density partitioning and least-squares fitting. This approach opens new avenues for exploring a wider range of crystal systems from the structural point of view.

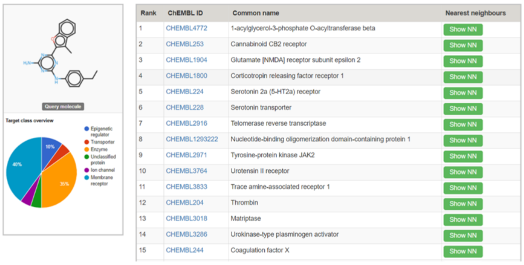
[1] Dylan Jayatilaka et al., *Acta Cryst. A*, **2008**, 64, 383-393.
[2] Silvia C. Capelli et al., *IUCrJ*, **2014**, 1, 361-379.
[3] Simon Grabowsky et al., Chem. Sci., **2017**, 8, 4159-4176.
[4] Lorraine A. Malaspina et al., J. Appl. Crystallogr., **2021**, 54, 987-995.

Predicting Off-targets from ChEMBL Data Using the Polypharmacology Browser

M. Darsaraee[1], J. Reymond[1]*

[1]Department of Chemistry, Biochemistry and Pharmacy, University of Berne, Freiestrasse 3, 3012 Berne, Switzerland

The public archive ChEMBL, which collects bioactive compounds and their associated targets from the literature, has been used by many groups to build models predicting the possible targets of small molecules to guide the experimental search for off-targets. In our group we have developed the polypharmacology browsers (PPB and PPB2), [1,2] which assign possible targets to a query molecule based on molecular fingerprint similarities to ChEMBL molecules, and provided critical insights in several practical case studies such as the identification of LPAATβ as the actual target of a putative kinase inhibitor (Figure). [3,4]



Rank	ChEMBL ID	Common name	Nearest neighbours
1	CHEMBL4772	1-acylglycerol-3-phosphate O-acyltransferase beta	Show NN
2	CHEMBL253	Cannabinoid CB2 receptor	Show NN
3	CHEMBL1904	Glutamate [NMDA] receptor subunit epsilon 2	Show NN
4	CHEMBL1800	Corticotropin releasing factor receptor 1	Show NN
5	CHEMBL224	Serotonin 2a (5-HT2a) receptor	Show NN
6	CHEMBL228	Serotonin transporter	Show NN
7	CHEMBL2916	Telomerase reverse transcriptase	Show NN
8	CHEMBL1293222	Nucleotide-binding oligomerization domain-containing protein 1	Show NN
9	CHEMBL2971	Tyrosine-protein kinase JAK2	Show NN
10	CHEMBL3764	Urotensin II receptor	Show NN
11	CHEMBL3833	Trace amine-associated receptor 1	Show NN
12	CHEMBL204	Thrombin	Show NN
13	CHEMBL3018	Matriptase	Show NN
14	CHEMBL3286	Urokinase-type plasminogen activator	Show NN
15	CHEMBL244	Coagulation factor X	Show NN

However, our PPB and PPB2 models associated only a single target per ChEMBL molecule. To better integrate the existing polypharmacology information available in ChEMBL, we are updating our PPB to handle multi-target information for ChEMBL molecules, using various machine learning models into account, and exploiting the latest version of the database featuring a total of 1.9 million molecule-target associations.

Keywords: Computer-aided drug design, Polypharmacology, Target prediction, Web-based tool, Cheminformatics

[1] M. Awale, J. L. Reymond, *J. Cheminf.* **2017**, *9*, 11, DOI: 10.1186/s13321-017-0199-x.

[2] M. Awale, J.-L. Reymond, *J. Chem. Inf. Model.* **2019**, *59*, 10, DOI: 10.1021/acs.jcim.8b00524.

[3] M. Poirier, M. Awale, M. A. Roelli, G. T. Giuffredi, L. Ruddigkeit, L. Evensen, A. Stooss, S. Calarco, J. B. Lorens, R.-P. Charles, J.-L. Reymond, *ChemMedChem* **2019**, *14*, 224, DOI: 10.1002/cmdc.201800554.

[4] M. R. Cunha, R. Bhardwaj, A. L. Carrel, S. Lindinger, C. Romanin, R. Parise-Filho, M. A. Hediger, J.-L. Reymond, *RSC Med. Chem.* **2020**, *11*, 1032, DOI: 10.1039/D0MD00145G.

Studying ion and water flow through aerolysin nanopores using molecular dynamic simulations

J. Duan[1], L. Perrin[1], C. Cao[1]

[1]Department of Inorganic and Analytical Chemistry, Chemistry and Biochemistry, University of Geneva; 1211 Geneva, Switzerland

Understanding the transport mechanisms of ions and water through nanopores is crucial for numerous applications in nanotechnology and biophysics. In this study, we utilize molecular dynamics (MD) simulations to investigate the flow of ions and water molecules through aerolysin nanopores. Aerolysin, a protein-based nanopore, along with its engineered mutant K238A, exhibits unique properties that make them promising candidates for biosensing and single-molecule sequencing technologies. Using MD simulations, we analyze the ion and water flow through aerolysin nanopores in different types of cation chloride solutions, including K^+, Li^+, and Gdm^+, and water molecules as they traverse the nanopore. We focus on key parameters such as ionic currents, conductivity, ion mobility, water flux, and possibility of interaction between cation with the pore, specifically their sticky characteristics. Our results reveal the influence of pore geometry, surface charge, and external electric fields on the flow behavior, providing insights into the fundamental mechanisms governing molecular transport in confined environments. This study enhances our understanding of nanopore-mediated transport processes and contributes to the development of advanced nanopore-based devices for diverse applications.

Accurate tunnelling splittings with second-order instantons

J. Dušek[1], J. E. Lawrence[2], J. O. Richardson[1]*

[1]Institute of Molecular Physical Science, ETH Zürich, [2]Simons Center for Computational Physical Chemistry, New York University

In Ring-polymer instanton theory (RPI) [1], we expand around a minimum-action trajectory called the instanton to approximately evaluate path integrals. This approach can be used to calculate tunnelling splittings and since we require information from just one trajectory, the method scales favourably for large systems (see e. g. [2]).

To improve the accuracy of RPI, we have previously introduced first-order perturbative corrections (PC1) which account for anharmonicity around the tunnelling trajectory [3]. With RPI+PC1, we still only need one trajectory, but additionally require third and fourth derivatives along it. This already dramatically improves the accuracy of the method, e. g. from −11% to 2% in the case of malonaldehyde proton transfer depicted in Fig. 1. Recently, we have derived second-order perturbative corrections (PC2), which improve the accuracy even further. Together with resummation techniques, we hope to create an affordable yet accurate technique for describing quantum tunnelling.

[1] Richardson, J. O., *Int. Rev. Phys. Chem*. 37:2, 171-216 (**2018**), DOI: 10.1080/0144235X.2018.1472353
[2] Nandi, A., et al., *J. Am. Chem. Soc*., 145, 17, 9655–9664 (**2023**), DOI: 10.1021/jacs.3c00769
[3] Lawrence, J. E., Dušek J. & Richardson, J. O., *J. Chem. Phys*. 159, 014111 (**2023**), DOI: 10.1063/5.0155579

Performance of the new non-decomposable approximant for the non-additive kinetic potential for embedded radical

T. Englert[1], T. A. Wesolowski[1]*

[1]University of Geneva, Department of Physical Chemistry, Quai Ernest-Ansermet 30, 1205 Genève

The bi-functional $v_t^{nad}[\rho_A,\rho_B]$ (non-additive kinetic potential) is the key ingredient of multi-level simulation methods based on Frozen-Density Embedding Theory (FDET) (see [1]). The commonly used approximations for $v_t^{nad}[\rho_A,\rho_B]$ are decomposable and semi-local. In some cases, these result in an artificial incorrect redistribution of electron density between the embedded species and the environment. The recently developed new approximant (NDCS) [2] is based on first principles and eliminates this flow as shown for model non-covalently bound dimers consisting of closed-shell molecules, atoms, or ions [2,3]. In this work we test the NDCS approximant on cases where the embedded species is a model radical ($OH^{\cdot}$) in a complex with alkaline cations and other closed shell species. Such systems are prone to artificial spin leak onto the cation resulting from the use of approximation for $v_t^{nad}[\rho_A,\rho_B]$. The comparisons with the corresponding reference results obtained for the complex, show that the NDCS outperforms significantly other semi-local non-decomposable approximations in case of Li^+ and K^+ reducing the relative error in the complexation induced shits of the isotropic hyperfine tensor component from 9.5% to 4% in case of Li^+ and from 4.8% to 0.4% in case of K^+. For Na^+ even using the simplest approximation to $v_t^{nad}[\rho_A,\rho_B]$ (decomposable functional derived from the Thomas-Fermi kinetic energy functional), yields leads to very small relative error of 0.9% and NDCS increases it marginally to 1.2%. The results obtained using basis sets of various qualities (always including atomic functions centred on BOTH the radical AND the cation) confirm the numerical soundness of the above observation [4]. These results are consistant with the ones previously observed in closed shell systems [2,3].

[1] Tomasz A. Wesolowski, Sapana Shedge, and Xiuwen Zhou, *Chemical Reviews*, **2015**, 115 (12), 5891-5928

[2] Elias Polak, Cristina E. González-Espinoza, Matin J. Gander, Tomasz A. Wesolowski, *Journal of Chemical Physics*, **2022**, 156, 044103.

[3] Elias Polak, Tanguy Englert, Martin J. Gander, Tomasz A. Wesolowski, *Journal of Chemical Physics*, **2023**, 158, 174107.

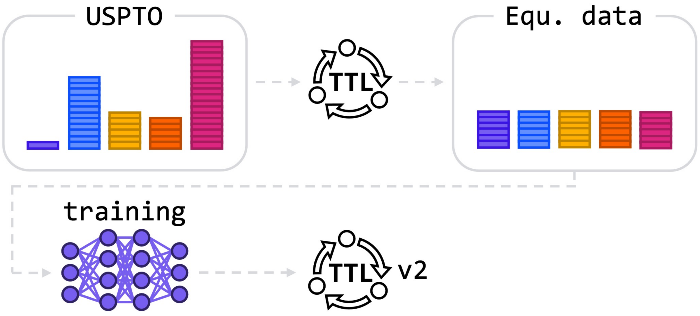
[4] Tanguy Englert, to be published.

Data augmentation leads to near quantitative single-step round-trip accuracy in transformer-based retrosynthesis models

Y. Grandjean[1], D. Kreutter[1], J. Reymond[1]*

[1]Department of Chemistry, Biochemistry and Pharmaceutical Sciences, University of Bern, Freiestrasse 3, 3012 Bern, Switzerland

In computer-aided synthesis planning (CASP), one trains machine learning models with organic reactions from literature, and challenges them to predict possible retrosyntheses of unseen molecules[1]. Many models perform well for single-step retrosynthesis, however only for reaction types that are well represented in the training data. Here we used our recently reported triple transformer loop (TTL) retrosynthesis model[2] to perform data augmentation and obtain a balanced dataset where each reaction type is represented by a similar number of examples. By retraining the TTL using the balanced dataset, we obtained a model showing near-quantitative performance in single-step round-trip accuracy prediction. Our data augmentation approach might be generally useful to broaden the scope of retrosynthesis models



[1] A. Thakkar, S. Johansson, K. Jorner, D. Buttar, J.-L. Reymond, O. Engkvist, *React. Chem. Eng.*, **2021**, 6, 27-51.
[2] D. Kreutter, J.-L. Reymond, *Chem. Sci.*, **2023**, 14, 9959-9969.

Structure-based machine learning models to predict environmental biodegradation half-lives and their uncertainty

J. Hafner[1,2], J. Cordero[1], A. Scheidegger[1], K. Fenner[1,2]*

[1]Swiss Federal Institute of Aquatic Science and Technology (Eawag), 8600 Dübendorf, Zürich, Switzerland, [2]University of Zürich, 8057 Zürich, Switzerland

Modelling the environmental persistence of chemicals is crucial to (i) develop new chemicals that are sustainable and safe by design, and (ii) to understand the environmental fate and hence exposure potential of existing micropollutants. Available biodegradation models fail to accurately predict the primary biodegradation half-lives (DT50s) of complex small molecules such as pharmaceuticals, pesticides, or industrial chemicals. This situation is mainly due to (i) the lack of abundant and standardized biodegradation data and (ii) to the high variability of experimental outcomes of biodegradation experiments.

In this work, we address these challenges by curating pesticide DT50s obtained from soil biodegradation experiments regarding their experimental variability, and by developing new Quantitative Structure-Biodegradation Relationships (QSBRs) to predict biodegradation half-lives for these structurally complex chemicals. The half-lives were obtained from the EAWAG-SOIL data package [1] on enviPath [2], resulting in a data set of 984 substances with at least one reported primary degradation half-life. For each substance, we applied a Bayesian inference technique to handle censored data points and to estimate the average DT50, its uncertainty, and its experimental variability across different studies [3]. QSBRs were constructed based on PaDEL descriptors, MACCS fingerprints, and enviPath's biotransformation rules, and machine learning models were trained using different regressors. While overall model performance was low (R2 < 0.4), Gaussian Process Regressors models not only predict a half-life value, but also assign a confidence metric to each prediction. We show that this confidence metric can be used to reliably distinguish accurate from inaccurate predictions, and to point towards classes of chemicals for which more experimental data is needed.

Our approach can be expanded to include more biodegradation data from other environmental compartments (e.g., activated sludge, water-sediment systems) in the future, with the objective to improve model performance, increase the coverage of the chemical space, and predict half-lives in other environments.

Acknowledgment: JH and KF are members of the Prioritisation and Risk Evaluation of Medicines In the EnviRonment (PREMIER). PREMIER has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 875508. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA (https://www.imi.europa.eu/).

Disclaimer: This abstract only reflects the author's view and the JU is not responsible for any use that may be made of the information it contains.

[1] D. A. R. S. Latino, J. Wicker, M. Gütlein, E. Schmid, S. Kramer, K. Fenner, Environmental Science: Processes & Impacts, 2017, 19, 449–464.

[2] J. Wicker, T. Lorsback, M. Gütlein, D. Latino, S. Kramer, K. Fenner, Nucleic Acids Research, 2016, 44, D502-D508.

[3] J. Hafner, K. Fenner, A. Scheidegger, Environmental Science & Technology Letters, 2023, 10, 859–864.

Spin-flip TDDFT within the Sternheimer formalism

L. I. Hernandez Segura[1], S. Luber[1]*

[1]University of Zurich, Zurich, Switzerland

In this work we present the state-of-the-art on low-scaling methods for spin-flip time dependent density functional theory (SF TDDFT) calculations. To this aim, the Sternheimer formalism[1] is employed to allow SF TDDFT calculations without the explicit need of unoccupied molecular orbitals, we name the resulting method SF TDDFPT. We implemented this method into the CP2K package (www.cp2k.org) to take advantage of its highly efficient methodologies, in particular, the Gaussian and plane-wave (GPW) method[2,3] and the auxiliary density matrix method (ADMM)[4]. The former is used to lower the computational scaling of the electron-electron repulsion and exchange-correlation contributions in the spin-flip kernel and the later is used for the calculation of the Fock exchange contribution with a reduced computational scaling. The resulting excitation energies are obtained within 0.3 eV with respect to highly accurate excitation energies[5], which is consistent with the reported accuracy of TDDFT[6,7].

[1] R. Sternheimer, *Physical Review,* **1951**, *84*, 244–253.

[2] G. Lippert, J. Hutter, M. Parrinello, *Molecular Physics,* **1997**, 92, 477-488.

[3] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, J. Hutter, *Computer Physics Communications,* **2005**, *167*, 103–128.

[4] M. Guidon, J. Hutter, J. VandeVondele, *Journal of Chemical Theory and Computation,* **2010**, *6*, 2348–2364.

[5] P.-F. Loos, A. Scemama, M. Boggio-Pasqua, D. Jacquemin, *Journal of Chemical Theory and Computation,* **2020**, *16*, 3720–3736.

[6] M. R. Silva-Junior, M. Schreiber, S. P. A. Sauer, W. Thiel, *Journal of Chemical Physics,* **2008**, *129*, 104103.

[7] T. M. Maier, H. Bahmann, A. V. Arbuznikov, M. Kaupp, *Journal Chemical Physics,* **2016**, *144,* 074106.

Reaction prediction in the low-data regime with transition state descriptors

L. Jacot-Descombes[1,2], J. Landis[1], K. Jorner[1,2]*

[1]Department of Chemistry and Applied Biosciences, ETH Zurich, Zurich, Switzerland, [2]NCCR Catalysis, ETH Zurich, Zurich, Switzerland

Abstract: Despite advancements in chemical reactivity prediction with machine learning (ML), the accuracy is limited by the small size of the datasets. Including knowledge from transition states (TS) to predict reactivity has been shown to lead to better results in the low data range [1–5]. To obtain the TS structures, density function theory (DFT) is the state-of-the-art method, but it presents a significant failure rate [1,3,6] and its high computational cost makes it challenging to integrate into ML pipelines. The SEAM method, derived from the empirical valence bond theory (EVB), offers a faster alternative for TS generation by approximating the TS as the intersection between the non-interacting force fields of the ground states of a reaction.

We developed the polanyi Python package to generate TS structures starting from the ground states coordinates. Polanyi implements the SEAM method based on the GFN-FF [7] force field via an interface to the xTB program [8]. We apply polanyi to the nucleophilic aromatic substitution (SNAr) [1] and the palladium-catalysed oxidative addition [9] reactions and compare it against higher-level computations and experimental rate constants. We then calculate electronic and steric descriptors on these generated TS geometries and use them to predict the rate and selectivity of the studied reactions.

[1] Kjell Jorner et al., *Chem. Sci.*, **2021**, 12, 1163–1175.

[2] Li-Cheng Xu et al., *Nat. Synth*, **2023**, 2, 321–330.

[3] Abbigayle E. Cuomo et al., *ACS Cent. Sci.*, **2023**, 9, 1768–1774.

[4] Eike Caldeweyher et al., *J. Am. Chem. Soc.*, **2023**, 145, 17367–17376.

[5] Manuel Orlandi et al., *Angew. Chem. Int. Ed.*, **2017**, 56, 14080–14084.

[6] Simone Gallarati et al., *Chem. Sci.*, **2021**, 12, 6879–6889.

[7] Sebastian Spicher, Stefan Grimme, *Angew. Chem. Int. Ed.*, **2020**, 59, 15665–15673.

[8] Christoph Bannwarth et al., *WIREs Comput. Mol. Sci.*, **2021**, 11, e1493.

[9] Jingru Lu et al., *Chem. Sci.*, **2022**, 13, 3477–3488.

Sampling strategies for expectation values within the Herman--Kluk approximation

F. Kröninger[1], C. Lasser[2]*, J. Vaníček[1]*

[1]Laboratory of Theoretical Physical Chemistry, Institut des Sciences et Ingénierie Chimiques, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland, [2]Department of Mathematics, Technische Universität München

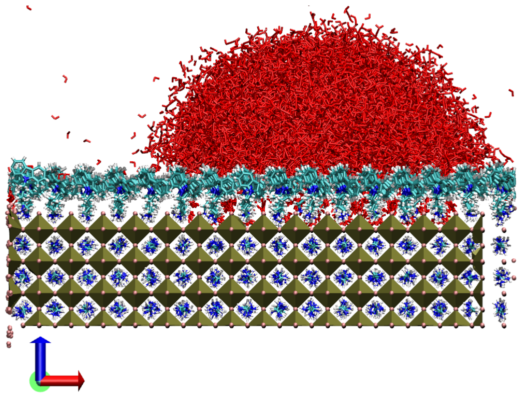
When computing quantum-mechanical observables, the curse of dimensionality limits the naive approach that uses the quantum-mechanical wavefunction. The semiclassical Herman--Kluk propagator provides a remarkable grid-free ansatz to evaluate the expectation values of these observables. In this paper, we investigate quadrature techniques for this high-dimensional and highly oscillatory propagator. In particular, we analyze Monte Carlo quadratures with three different initial sampling approaches. The first two, based on the Husimi density and its square root respectively, are independent of the observable while the third approach, which is new, incorporates the observable in the sampling to minimize the variance of the Monte Carlo integrand at initial time. We prove sufficient conditions for convergence of the Monte Carlo estimators and provide convergence error estimates. The analytical results are validated by numerical experiments in various dimensions on a harmonic oscillator and on a Henon-Heiles potential with an increasing degree of anharmonicity.

Multiscale simulations of hybrid halide-perovskite photovoltaic devices

N. Lempesis[1], V. Carnevali[1], L. Agosta[1], V. Sláma[1], A. Vezzosi[1], U. Röthlisberger[1]*

[1]Laboratory of Computational Chemistry and Biochemistry, Institute of Chemical Sciences and Engineering, École Polytechnique Fédérale de Lausanne

Hybrid halide perovskites rapidly captured the interest of the scientific community, in large part, due to their remarkable optical properties enabling high power conversion efficiencies their versatility and broad application spectrum including photovoltaics, light-emitting diodes, light detection, and energy storage. Most of this rapid progress is owed to changes in synthesis conditions, whose fundamental understanding at the atomic level is still elusive. In principle, computer simulation can elucidate the atomistic mechanisms underlying these phenomena. The efficient combination of multiscale simulation techniques spanning broad length/time scales can address formidable challenges related to having multiple atomic species, long relaxation times (high activation barriers), intricate system geometries, ir)reversible phase transitions, numerous metastable states and large system requirements to name a few approach to understanding structure-property relations entails the development, refinement and application of multiscale computational techniques acting synergistically to span electronic and atomistic levels of detail. Classical molecular dynamics simulations complemented by ab-initio molecular dynamics and density functional theory simulations are utilized to obtain an accurate description of the physics behind the mechanisms responsible for enhancing long-term stability and improving the efficiencies of PSCs, while also indicating possible new roadmaps for superior PSC devices.



[1] E. A. Alharbi, A. Krishna, N. Lempesis, M. Dankl, I. Mosquera-Lois, M. A. Hope, T. P. Baumeler, G. Kakavelakis, A.Mishra, F. T. Eickemeyer, O. Ouellette, T. Chawanpunyawat, A. Hagfeldt, S. M. Zakeeruddin, L. Emsley, L. Pfeifer, U.Roethlisberger, M. Grätzel, *Joule*, **2023**, 7, 183-200

[2] L. Li†, M. Wei†, V. Carnevali†, H. Zeng, R. Liu, N. Lempesis, F. T. Eickemeyer, M. Zeng, L. Luo, L. Agosta, M. Dankl, S. M. Zakeeruddin, U. Rothlisberger, M. Grätzel, Y. Rong, X. Li, *Advanced Materials*, **2024**, 36, 2303869

[3] D. Han, J. Wang, L. Agosta, Z. Zang, B. Zhao, L. Kong, H. Lu, I. Mosquera-Lois, V. Carnevali, J. Dong, J. Zhou, H. Ji, L. Pfeifer, S. M. Zakeeruddin, Y. Yang, B. Wu, U. Rothlisberger, X. Yang, M. Grätzel, N. Wang, *Nature*, **2023**, 622, 493-498

[4] S. M. Park†, M. Wei†, N. Lempesis†, W. Yu, T. Hossain, L. Agosta, V. Carnevali, H. R. Atapattu, P. Serles, F. T. Eicke- meyer, H. Shin, M. Vafaie, D. Choi, K. Darabi, E. D. Jung, Y. Yang, D. B. Kim, S. M. Zakeeruddin, B. Chen, A. Amassian, T. Filleter, M. G. Kanatzidis, K. R. Graham, L. Xiao, U. Rothlisberger, M. Grätzel, E. H. Sargent, *Nature*, **2023**, 1-3

[5] J. Jeong, T. Chawanpunyawat, M. Kim, V, Sláma, N. Lempesis, L. Agosta, V. Carnevali, Q. Zhang, F. T. Eickemeyer, L. Pfeifer, Y. J. Kim, J. W. Song, H. Lu, M. Amalki, S. Mo, S. M. Zakeerudin, U. Rothlisberger, D. S. Kim, P. J. Dysonand M. Graetzel, *Advanced Energy Materials*, **2024**, accepted

MiMiC Framework for Multiscale Simulations: Current Status and Recent Advancements

A. Levy[1], A. Antalík[1], U. Rothlisberger[1]*

[1]Laboratory of Computational Chemistry and Biochemistry, Institute of Chemical Sciences and Engineering, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

Multiscale simulations represent an indispensable part of computational chemistry, as they provide important insights into chemical and biological phenomena occurring at different space and time scales. The most prominent example of these is the hybrid quantum mechanics/molecular mechanics (QM/MM) method, where the chemically active part of the system is treated at the QM level, while the rest is treated at the MM level, with classical force fields. In this context, we developed the MiMiC [1,2], a framework for multiscale simulations, designed to be extremely flexible and, at the same time, efficient. At the moment, MiMiC offers an efficient implementation of electrostatic embedding QM/MM, interfacing GROMACS and CPMD as MM and QM clients, respectively. A large effort is being put into extending available clients and, at the same time, the capabilities of the framework, in terms of supporting new features and different multiscale methods. In this contribution, I will give an overview of the current status of the MiMiC framework and provide insights into some future advancements in the development of which I am personally involved. In particular, these include the interface with new client codes, such as OpenMM and CP2K, and the introduction of new features in MiMiC itself, such as a generalised version of the D-RESP approach [3]. These additions will further extend the flexibility of the MiMiC framework, expanding its capabilities in multiscale simulations

[1] A. Andrej, A. Levy, S. Kvedaravičiūtė, S. K. Johnson, D. Carrasco-Busturia, B. Raghavan, F. Mouvet, A. Acocella, S. Das, V. Gavini, D. Mandelli, E. Ippoliti, S. Meloni, P. Carloni, U. Rothlisberger, and J. M. Haugaard Olsen, *arXiv preprint*, **2024,** arXiv:2403.19035.
[2] J. M. Haugaard Olsen, V. Bolnykh, S. Meloni, E. Ippoliti, M. P. Bircher, P. Carloni, and U. Rothlisberger, *J. Chem. Theory Comput.*, **2019**, 15, 6, 3810–3823.
[3] L. Alessandro, J. VandeVondele, and U. Rothlisberger, *J. Phys. Chem. B*, **2002**, 106, 29, 7300-7307

Nonadiabatic dynamics simulations in periodic condensed phase systems and challenges with hybrid-based ΔSCF

M. Mališ[1], E. Vandaele[1], S. Luber[1]*

[1]Department of Chemistry, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

Almost all nonadiabatic (NA) processes encountered in nature happen in the condensed phase. City smog is a direct consequence of various nitrogen oxides and hydroxides generated by solar light photodecomposition of various manmade pollutants, such as nitrophenols, in the atmospheric water droplets. In technological applications like CO_2 photocatalytic reduction, photocatalyst attachment to a solid surface can enhance its absorbance, increase its photostability, and utilize the surface as a direct sacrificial electrode, to name a few benefits. In silico modelling of NA processes of photoexcited systems in condensed phase represents a significant theoretical and computational challenge. In addition to the system size increase due to additional environment molecules that must be accounted for the number of electronic states and the potential nonradiative deactivation (NRD) pathways all increase. Two such photoexcited systems in a condensed phase environment were studied through NA molecular dynamics (MD) simulations with all atoms at the same time-dependent density functional(TDDFT) level of theory, incorporating periodic boundary conditions for the condensed phase description. NRD mechanisms of ortho- and para-nitrophenol initiated from the brightest π to π^* singlet excited electronic state were investigated with trajectory surface hopping NA MD simulations on a manifold including several (up to twelve in total) singlet and triplet excited electronic states, mutually coupled by the spin-orbit coupling elements computed at the TDDFT level of theory. By careful analysis between such NRD pathways obtained on isolated and in water-solvated nitrophenols, the cage effect exhibited by the surrounding water molecules on electronically excited nitrophenol reduces the number of inter-system crossings and facilitates its NRD deactivation from the first singlet excited electronic states. When the water solvent was modelled at a force field level as usually performed when applying the computationally cheaper molecular mechanics/quantum mechanics (QM/MM) approach, the NRD mechanisms resemble more those of the isolated system than when the whole solvent was treated at the same level of theory as the chromophore.[1] An ultrafast excitation transfer from a localized Re-photoreduction catalyst to a TiO_2 anatase (101) surface was simulated with Ehrenfest NA MD utilizing the real-time TDDFT propagation of the initially ΔSCF generated excited electronic state density. In a subfemtosecond time scale, the metal-to-ligand charge transfer excitation localized on the Re-photoreduction catalyst transfers through the ligand frame and its phosphate groups, anchoring it to the anatase (101) surface, to the anatase (101) surface and suddenly polarizes all its layers.[2] Finally, we investigated the origin of the ΔSCF singlet excited energy shift when the electronic state is constructed using fractional occupation numbers with a hybrid density functional. Additional terms, some identical to the self-interaction error culprit terms, arise from the Coulomb and the exact Hartree-Fock exchange terms obtained from nonidempotent density matrices incorporating fractional occupation numbers. Ways to alleviate them are investigated, as this issue seems ubiquitous to any DFT calculation employing fractional occupation numbers with hybrid density functionals.[3]

[1] Eva Vandaele, Momir Mališ, Sandra Luber, *Journal of Chemical Theory and Computation*, **2024**, 20, 3258-3272.

[2] Eva Vandaele, Momir Mališ, Sandra Luber, *in revision*.

[3] Momir Mališ, Sandra Luber, *in preparation*.

Abraham solvation parameters model in the frame of graph convolutional neutral networks

M. Nedyalkova[1], A. Paluch [2], M. Lattuada [1]

[1]Department of Chemistry, Fribourg University, Chemin Du Musée 9, 1700 Fribourg, Switzerland, [2]Department of Chemical, Paper, and Biomedical Engineering, Miami University

Abraham model solute descriptors model have been leveraged for 7000 contaminates extracted from UFZ-LSER database [1]. This database contains experimentally determined Linear Solvation Energy Relationship (LSER) descriptors, E (excess molar refraction), S (polarizability/dipolarity), A (solute hydrogen bond acidity), B and/or B0 (solute hydrogen bond basicity), and L (logarithmic gas-hexadecane partition coefficient). The database also stores the McGowan's molar volume V calculated for each compound from its molecular structure. These solute descriptors can be used, in combination with LSER equations or other types of polyparameter linear free energy relationships (pp-LFERs), to calculate partition coefficients for various systems.

As a molecular graph representation, we used a graph convolutional neural network architecture (GCN). Our approach showed that several important chemical and thermodynamic properties could be predicted using the applied approach, including partition coefficients, molar solubility ratios, gas-liquid chromatography and HPLC retention data, infinite dilution activity coefficients, molar enthalpies of solvation, molar vaporization and sublimation at 298 K, vapor pressures, and fusion coefficients.

The GCN can predict more accurately than previous methods over broader data sets and is independent of the descriptors selection space.

[1] Ulrich, N.; Endo, S.; Brown, T.N.; Watanabe, N.; Bronner, G.; Abraham, M.H.; Goss, K.-U. UFZ-LSER Database v 3.2.1; Helmholtz Centre for Environmental Research-UFZ: Leipzig, Germany, 2017; Available online: http://www.ufz.de/lserd

Molecular dynamics simulation of aerolysin nanopore using high-resolution cryo-EM structure

L. W. Perrin[1], J. Anton[2], I. Iacovache[3], J. F. Bada Juarez[2], L. A. Abriata[2], M. J. Marcaida Lopez[2], M. Dal Peraro[2], B. Zuber[3], C. Cao[1]*

[1]Department of Inorganic and Analytical Chemistry, Chemistry and Biochemistry, University of Geneva; 1211 Geneva, Switzerland, [2]Institute of Bioengineering, School of Life Science, École Polytechnique Fédérale de Lausanne (EPFL); 1015 Lausanne, Switzerland, [3]Institute of Anatomy, University of Bern, 3012 Bern, Switzerland

Nanopores have recently emerged as a powerful single molecule sensing tool. A diverse array of pore-forming toxins proteins are now utilized as nanopore sensors to characterize biomolecules, advancing both life sciences and medical applications[1-4]. Among these, aerolysin stands out as an excellent candidate due to its narrow and long lumen, which confers high sensitivity[5]. Despite its promising applications, the current resolution of the available aerolysin structure is 7.9 Å[6], comparable to the radius of its barrel (10-17 Å). This limitation hinders our understanding of the sensing principles of the aerolysin nanopore. In this study, based on two newly obtained high-resolution structures of aerolysin via cryo-EM (2.2 Å) in both amphibol and lipid nanodiscs, we use molecular dynamics simulations to analyze their structure profiles and ionic current flows at the atomic-level. Our results reveal a previously unidentified constriction within the pore's lumen, which has been underestimated in previous studies. Furthermore, the ionic current calculated from molecular dynamics trajectories with these structures align more closely with experimental data. These findings enhance our understanding of aerolysin nanopores' sensing principle and suggest potential advancements in their applications as molecular sensors.

[1] Dorey, A. ; Howorka, S. Nanopore DNA Sequencing Technologies and Their Applications towards Single-Molecule Proteomics. Nat. Chem. 2024, 16 (3), 314-334.

[2] Hu, Z. ; Huo, M. ; Ying, Y. ; Long, Y. Biological Nanopore Approach for Single-Molecule Protein Sequencing. Angew. Chem. 2021, 133 (27), 14862-14873.

[3] Mohammadi, M. M.; Bavi, O. DNA Sequencing: An Overview of Solid-State and Biological Nanopore-Based Methods. Biophys. Rev. 2022, 14 (1), 99–110.

[4] Cao, C.; Magalhães, P.; Krapp, L. F.; Bada Juarez, J. F.; Mayer, S. F.; Rukes, V.; Chiki, A.; Lashuel, H. A.; Dal Peraro, M. Deep Learning-Assisted Single-Molecule Detection of Protein Post-Translational Modifications with a Biological Nanopore. ACS Nano 2024, 18 (2), 1504–1515.

[5] Cao, C.; Cirauqui, N.; Marcaida, M. J.; Buglakova, E.; Duperrex, A.; Radenovic, A.; Dal Peraro, M. Single-Molecule Sensing of Peptides and Nucleic Acids by Engineered Aerolysin Nanopores. Nat. Commun. 2019, 10 (1).

[6] Iacovache, I.; De Carlo, S.; Cirauqui, N.; Dal Peraro, M.; Van Der Goot, F. G.; Zuber, B. Cryo-EM Structure of Aerolysin Variants Reveals a Novel Protein Fold and the Pore-Formation Process. Nat. Commun. 2016, 7 (1), 12062.

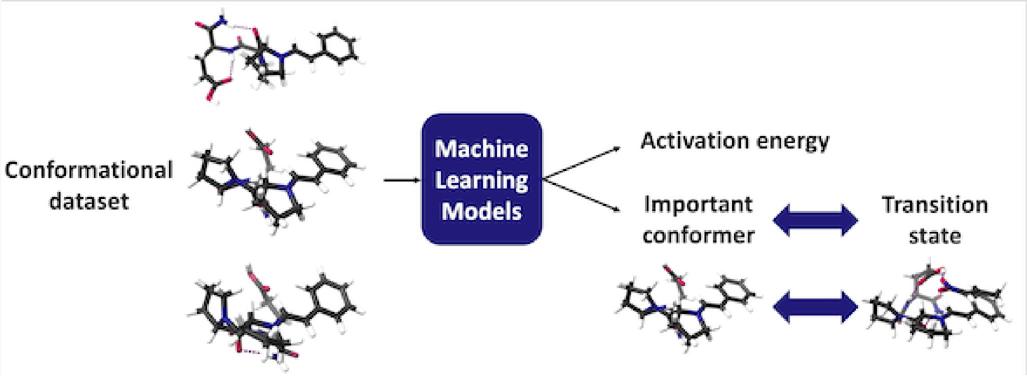
Reactivity prediction of highly flexible catalysts using conformationally enriched machine learning

S. P. Schmid[1], C. Ser[2,3], M. Skreta[3,4], A. Aspuru-Guzik[2,3]*, K. Jorner[1,5]*

[1]Department of Chemistry and Applied Biosciences, ETH Zurich, Zurich, Switzerland, [2]Department of Chemistry, University of Toronto, Toronto, ON, Canada, [3]Department of Computer Science, University of Toronto, Toronto, ON, Canada, [4]Vector Institute for Artificial Intelligence, Toronto, ON, Canada, [5]NCCR Catalysis, ETH Zurich, Zurich, Switzerland

Recent years have seen a rise in the development of machine learning (ML) methods applied to chemical reactions, ranging from ML-guided reaction optimization to catalyst design [1,2]. Most applications consider molecules as SMILES line-notation, or through representations derived from their 2D- or lowest-energy 3D-structure. Thus, they neglect the conformational freedom of molecules, which has been identified as a crucial factor in numerous catalytic reactions [3,4]. Recent work has benchmarked the benefits of including conformational information into ML models for moderately flexible BINOL-phosphoric acid catalysts [5,6]. However, conformational ML remains unexplored for more flexible catalysts, where such models could identify reactive conformers, providing insight into the reaction mechanism.

This work introduces conformer-enriched machine learning for reactivity prediction of highly flexible tripeptide catalysts. We design and create a dataset for activation energy prediction of a peptide-catalysed addition of aldehydes to nitroolefins, considering both energetically and structurally representative catalyst conformers. Semi-empirical methods [7] are employed for the ultra-fast generation of multiple reactant, transition state and product conformers of over 1’000 reactions, thus providing a reactivity dataset with energetically relevant and structurally diverse conformers of all reaction species for further model development. The dataset is used to create a conformer-augmented molecular representation, considering conformational information for reactivity prediction of highly flexible molecules. Important conformers for the prediction are further extracted from the ML model and compared to reaction transition states to assess the capability of the model to identify reactive conformers.



[1] Connor J. Taylor, et al. *Chem. Rev.* **2023**, *123, 6*, 3089-3126

[2] Julius Seumer, et al. *Angew. Chem. Int. Ed.* **2023**, *62*, e202218565

[3] Carla Rigling, et al. *J. Am. Chem. Soc.* **2018**, *140, 34*, 10829-10838

[4] Jennifer M. Crawford, et al. *Synthesis* **2019**, *51, 5*, 1021-1036

[5] Yanqiao Zhu, et al. *arXiv* **2023**, https://doi.org/10.48550/arXiv.2310.00115

[6] Dmitry Zankov, et al. *JCIM* **2023**, *63 (21), 6629-6641*

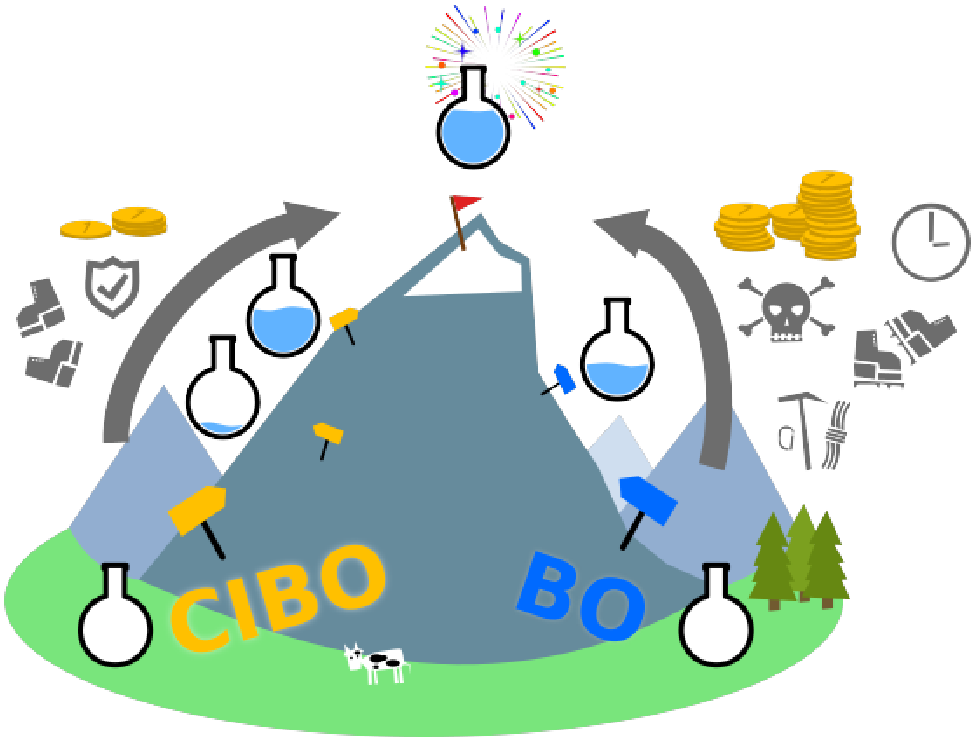
[7] Christoph Bannwarth, et al. *WIREs Comput. Mol. Sci.* **2020**, *11*, e01493

Cost-Informed Bayesian Reaction Optimization

A. A. Schoepfer[1,2], J. Weinreich[1,2], R. Laplaza[1,2], J. Waser[1,2]*, C. Corminboeuf[1,2]*

[1]Institute of Chemical Sciences and Engineering, École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland, [2]National Center for Competence in Research-Catalysis (NCCR-Catalysis), École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

Bayesian optimization (BO) is an efficient method for solving complex optimization problems, including those in chemical research, where it is consequently gaining significant popularity. Although effective in guiding experimental design, BO does not account for experimentation costs: testing readily available reagents under different conditions might be more cost and time-effective than synthesizing or buying additional ones. To address this issue, we present cost-informed BO (CIBO), an approach tailored for the rational planning of chemical experimentation that prioritizes the most cost-effective experiments. Reagents are used only when their anticipated improvement in reaction performance sufficiently outweighs their costs. Our algorithm tracks the available reagents, including recently acquired ones, and dynamically updates their cost during the optimization. Using literature data of Pd-catalyzed reactions, we show that CIBO reduces the cost of reaction optimization by up to 90% compared to standard BO. Our approach is compatible with any type of cost, *e.g.*, the cost of buying equipment or compounds, waiting time, and environmental or security concerns. We believe CIBO supersedes BO in chemistry and envision applications in both traditional and self-driving laboratories for experiment planning.



Assessing the performance and accuracy of Delta Self-Consistent Field (ΔSCF) method within Restricted-Open Kohn–Sham formalism

A. Sinyavskiy[1], M. Mališ[1], S. Luber[1]*

[1]Department of Chemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

The simulation of photochemical processes requires efficient electronic structure methods to calculate the properties of excited electronic states (such as vertical excitation energies, oscillator strengths, nonadiabatic and spin-orbit couplings, etc). For the last couple of decades the method of choice has been the Linear-Response Time-Dependent Density Functional Theory (TDDFT). It is fast and reliable, but fails to describe certain phenomena such as long-range charge-transfer and multi-electron excitations. A computationally affordable alternative is the Δ Self-Consistent Field (ΔSCF) method. In principle it's a generalization of the SCF procedure, but instead of occupying the lowest-lying molecular orbitals according to the Aufbau principle, an electron is promoted to a higher molecular orbital to construct a specific excited state of a given electronic density.

A thorough investigation of the effectiveness of the ΔSCF method within the Restricted-Open Kohn–Sham formalism has been carried out. A modification of the Maximum Overlap Method to improve the convergence of the ΔSCF method while preserving the order of the molecular orbitals is presented. For benchmarking purposes a large-scale comparison with TDDFT has been conducted: both single- and multi-reference singlet and triplet excited states of various molecules have been reproduced using ΔSCF and compared to the reference states provided by TDDFT. A decent correlation between constructed excited states electron densities and other properties has been observed. It has also been found that the energy gap alone is not a reliable criterion to compare the calculated excited states. The similarity between the electron densities and the transition dipole moments of the molecules should be used along with the energy gaps. [1]

[1] Andrey Sinyavskiy, Momir Mališ, Sandra Luber, in preparation.

Predicting Reaction Properties using SMILES/CGR

G. Sulpizio[1], K. Jorner[1]*

[1]Department of Chemistry and Applied Biosciences, ETH Zürich

Machine learning models have gained significant popularity for a wide variety of property predictions in chemistry, yet predicting reaction properties such as yield, activation energy, and kinetic rate constants remains a challenge. Currently, graph-based models seem to offer the highest accuracy and reliability, specifically models based on Condensed Graph of Reactions (CGR) [1]. Additionally, in recent years the field has seen the surge of large language models (LLMs) that have yielded alternative powerful methods for the task of predicting reaction properties mainly based on reaction SMILES [2]. In an effort to combine the advantages of the CGR representation with those of LLMs, we utilize the text representation of CGR (called SMILES/CGR [3]) to predict reaction properties using a BERT model. Preliminary results indicate that our model performs comparably to those based on reaction SMILES, although it is still outperformed by graph-based models.

[1] E. Heid, W. Green, J. Chem. Inf. Model., 2022, 62, 2101-2110
[2] P. Schwaller, A. Vaucher, T. Laino, J.-L. Reymond, Mach. Learn.: Sci. Technol., 2021, 2, 015016
[3] W. Bort et al., Sci. Rep., 2021, 11, 3178

Striking the right balance of encoding electron correlation in the Hamiltonian and wavefunction ansatz

K. Szenes[1], M. Reiher[1]

[1]ETH Zurich

Multi-configurational electronic structure theory delivers the most versatile approximations to many-electron wavefunctions, flexible enough to deal with all sorts of transformations, ranging from electronic excitations, to open-shell molecules and chemical reactions. Multi-configurational models are therefore essential to establish universally applicable, predictive ab initio methods for chemistry. Here, we present a discussion of explicit correlation approaches which address the nagging problem of dealing with static and dynamic electron correlation in multi-configurational approaches. We review the latest developments and then point to their key obstacles. Our discussion is supported by new data obtained with tensor network methods. We argue in favor of simple electrons-only correlator expressions that may allow one to define transcorrelated models, in which the correlator does not bear a dependence on molecular structure.

[1] K. Szenes, M. Moerchen, P. Fischill, M. Reiher, Faraday Discussions, 2024.

[2] A. Baiardi, M. Lesiuk, and M. Reiher, Journal of Chemical Theory and Computation, 2022, 18, 4203–1.

[3] A. Baiardi, M. Reiher, The Journal of Chemical Physics, 2020, 153, 164115.

Computational insights into organic halide perovskite solar devices

A. Vezzosi[1], N. Lempesis[1], V. Carnevali[1], V. Sláma[1], U. Röthlisberger[1]*

[1]Laboratory of Computational Chemistry and Biochemistry, École Polytechnique Fédérale de Lausanne (EPFL)

Hybrid organic/inorganic halide perovskites solar cells (PSCs) have become a promising class of materials for solar cells. Researchers have recently developed numerous methods to reduce defects, suppress non-radiative recombination, and enhance device efficiency and stability. Among these, additive and interface engineering methods have attracted extensive attention and are particularly popular for high-efficiency and stable PSCs. However, fundamental knowledge of the underlying atomic-level processes leading to improved performance and lifespan is often lacking. We employ DFT calculations, classical and ab-initio molecular dynamics (MD) simulations, with the help of state-of-the-art, polarizable classical force-fields and machine-learned interatomic potentials, trained on ab-initio MD trajectories, to unveil the underlying mechanisms connected to the facilitated interfacial charge transfer and improved stability of various PSCs devices. In particular, we apply our methodology to assess the stabilizing nature, as well as the optoelectronic properties of recently proposed hydrophobic organic cations in the context of layered two-dimensional perovskite materials and inverted PSCs.

[1] S. M. Park, M. Wei, N. Lempesis, W. Yu, T. Hossain, L. Agosta, V. Carnevali, H. R Atapattu, P. Serles, F. T Eickemeyer, H. Shin, M. Vafaie, D. Choi, K. Darabi, E. D. Jung, Y. Yang, D. B. Kim, S. M Zakeeruddin, B. Chen, A. Amassian, T. Filleter, M. G Kanatzidis, K. R Graham, L. Xiao, U. Röthlisberger, M. Grätzel, E. H Sargent, Nature, 2023, 624, 289-294.

[2] J. Jeong, T. Chawanpunyawat, M. Kim, V, Sláma, N. Lempesis, L. Agosta, V. Carnevali, Q. Zhang, F. T. Eickemeyer, L. Pfeifer, Y. J. Kim, J. W. Song, H. Lu, M. Amalki, S. Mo, S. M. Zakeerudin, U. Röthlisberger, D. S. Kim, P. J. Dyson and M. Graetzel, Advanced Energy Materials, 2024, accepted

[3] L. Li, M. Wei, V. Carnevali, H. Zeng, M. Zeng, R. Liu, N. Lempesis, F. T. Eickemeyer, L. Luo, L. Agosta, M.Dankl, S. M Zakeeruddin, U. Röthlisberger, M. Grätzel, Y. Rong, X. Li, Advanced Materials, 2024, 36, 2303869.

[4] C. Vona, M. Dankl, A. Boziki, M. P. Bircher, U. Röthlisberger, 2024, in preparation.

Ligand-based 3D Pharmacophore Search for Drug Analogues in Ultra-Large Combinatorial Libraries

M. von Korff[1], T. Sander[1]

[1]Schlossweg 63 / CH-4143 Dornach

In the field of drug discovery, "me-too" approaches are increasingly vital for identifying potential candidates for further development. Our study focuses on this aspect, evaluating both the performance and diversity of molecular analogues identified by an innovative pharmacophore search method in Ultra-Large Combinatorial Libraries (ULCLs). Our test set comprises eight recently approved first-in-class drugs.

We introduce the newly developed Pharos3D Method, which is a refinement of the Flexophore descriptor, a component of the open-source tools DataWarrior and OpenChemLib. Unlike classical ‘brute force’ approaches, Pharos3D enables the rapid screening of ULCLs within hours on computers with moderate hardware.

Here, we demonstrate that our new method identifies excellent fitting analogues for drugs. These analogues exhibit chemical diversity from the query molecule and can serve as a starting point for new medicinal chemistry programs aimed at addressing issues related to ADMET or intellectual property (IP).

3-Dimensional descriptors of molecular lipophilicity of macrocycles

F. Waibl[1], F. Dey[2], S. Riniker[1]*

[1]Institute of Molecular Physical Sciences, ETH Zürich, Zürich, Switzerland, [2]Roche Pharma Research and Early Development, Therapeutic Modalities, Roche Innovation Center Basel, F. Hoffmann-La Roche, Basel, Switzerland

In the development of novel therapeutic compounds, understanding polarity and lipophilicity is crucial to predict biophysical properties such as solubility or membrane permeability. However, these properties depend on the molecular conformation, which is inherently dynamic and responds to the molecular environment. Especially for larger drug compounds such as macrocycles, strong adaptations to the polarity of the environment help increase the membrane permeability, an effect that has been termed chameleonicity. Thus, conformation-dependent polarity is important in the computational modeling of chameleonicity and related biophysical quantities. Molecular dynamics (MD) is a powerful method to compute conformational ensembles and allows for detailed analysis of chameleonicity.

Here, we present a 3D-based polarity descriptor, based on an ML-base implicit solvent model previously developed in our group as well as the solvent excluded surface area, and combine it with MD simulations to study lipophilicity and chameleonicity of small molecules and macrocycles. Our results demonstrate that MD-based ensembles can improve in-silico predictions of lipophilicity and give novel insights into chameleonicity.

Stochastic Sampling around Ring-Polymer Instanton

Y. Wang[1], J. Richardson[1]

[1]Department of Chemistry and Applied Biosciences, ETH Zürich, 8093 Zürich, Switzerland

Quantum tunneling plays a crucial role in various chemical processes. By identifying dominant tunneling paths without prior knowledge of the tunneling path, the semiclassical instanton theory with the ring-polymer implementation[1] represents a valuable theoretical approach to describe tunneling phenomena. However, due to its reliance on the steepest descent approximation, it is unable to capture anharmonic effects perpendicular to the tunneling path if higher-order corrections are not further considered[2].

To overcome this limitation, here we propose a novel numerically exact approach called the instanton Monte Carlo method. The basic idea lies in conducting Monte Carlo random sampling around the instanton path. Specifically, utilizing a ring-polymer implementation, we transform the real coordinate space around the instanton solution into a normal-mode coordinate space, rescaled by the eigenfrequencies, while maintaining affine isomorphism with the original coordinates. Within this new space, we conduct random sampling weighted by the imaginary-time action. Applying this approach to compute the tunneling splitting in the one- and two-dimensional double-well models, the sampling procedure exhibits remarkably high efficiency, requiring only a few configuration updates to eliminate the autocorrelation between adjacent samples.

[1] J. O. Richardson, International Reviews in Physical Chemistry, 2018, 37, 171.

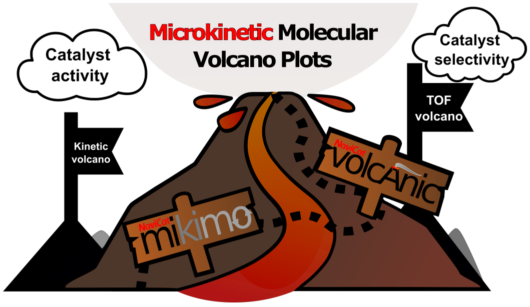
[2] J. E. Lawrence, J. Dušek, J. O. Richardson, The Journal of Chemical Physics, 2023, 159, 014111.

Microkinetic Molecular Volcano Plots for Enhanced Catalyst Selectivity and Activity Predictions

T. Worakul[1], R. Laplaza[1,2], S. Das[1], M. Wodrich[1,2], C. Corminboeuf[1,2]*

[1]Laboratory for Computational Molecular Design, Institute of Chemical Sciences and Engineering, École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland, [2]National Center for Competence in Research-Catalysis (NCCR-Catalysis), École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

Volcano plots and 2D activity maps,[1] derived from linear free energy scaling relationships, are valuable tools for predicting and rationalizing catalytic properties. By exploiting Sabatier's principle, these tools enable rapid identification of optimal catalysts through visual representations that correlate catalytic performance with a key energetic descriptor representative of the catalytic cycle. This work combines the rapid screening ability of molecular volcano plots with microkinetic modeling to create the next-generation of *microkinetic volcano plots* that predict the concentration of the product(s) at a given reaction time. We also introduce *mikimo*,[3] a lightweight Python program that is seamlessly integrated with our previously released volcano plot/activity map builder, *volcanic*[2] which automates the construction of microkinetic molecular volcano plots and activity/selectivity maps.



[1] Wodrich, M. D.; Sawatlon, B.; Busch, M.; Corminboeuf, C., The Genesis of Molecular Volcano Plots. *Accounts of Chemical Research* **2021,** *54* (5), 1107-1117.

[2] Laplaza, R.; Das, S.; Wodrich, M. D.; Corminboeuf, C., Constructing and interpreting volcano plots and activity maps to navigate homogeneous catalyst landscapes. *Nature Protocols* **2022,** *17* (11), 2550-2569.

[3] Worakul T, Laplaza R, Das S, Wodrich M, Corminboeuf C. Microkinetic Molecular Volcano Plots for Enhanced Catalyst Selectivity and Activity Predictions. *ChemRxiv*. **2024**; doi:10.26434/chemrxiv-2024-j79s3.