## Deep learning interatomic potentials development for LPS electrolytes

## <u>A. Azzali</u><sup>\*1</sup>, E. Bodo<sup>1</sup>

1Department of chemistry, Sapienza, università di Roma, Rome

\*alessandro.azzali@uniroma1.it

### Abstract

All-solid state (ASSB) batteries are one of the most promising alternatives of Li-ion liquid technology, thay have several advantages<sup>1</sup>, such as greater stability, better safety, and higher energy density<sup>2</sup>, but also an intrinsic and still unsolved problem, the low ion mobility. Amorphous Lithium thiophosphates, LPS (xLi<sub>2</sub>S-(1-x)P<sub>2</sub>S<sub>5</sub>) and their oxidised form LPSO are among the most promising<sup>3</sup> electrolytes. Although optimizing their performances by tuning their composition in order to achieve at least 10<sup>-3</sup> S/cm is an extremely complicated task. Using molecular modelling we can help making this process much quicker. So we decided, starting from quantum mechanics calculations, to create a deep learning force filed (DLFF). These FF can be used to perform large scale simulations that better foresee the electrochemical properties within short times, classical molecular dynamics speed, but with an accuracy depending on the theory level used to create the model, in this case Ab-initio molecular dynamics.

## Method

All DLFFs require a specific iter to be modelled.

- Data set acquisition or generation
- Training the neural network
- Validation of the model

First a data set must be generated, and the theory level must be chosen. This because, as said before, the input accuracy is the same as the output. Since we need a DFT theory level this step is the slowest. To generate the DLFFs we used the DeepMD kit program<sup>4</sup>, which is a solid package and heavily optimized code, that can generate the neural network and training it with data from other calculations, in this case ab-initio molecular dynamics perform with the program VASP, at DFT theory level, with PBE as functional, performed from 700K to 1500K.

We started acquiring crystal structures for ten different LPSs from XRD and more than six hundred quantum mechanically optimized structures of glass phase LPSs<sup>5</sup>. With this extensive set we perform the first training of the deep neural network (DNN) with a configuration of five layers and 2 million of training steps. Using the amorphous LPS as training set and the crystalline as validation set, we demonstrate that these systems are not so different for the deep learning model generated, since the error of single point energy (SPE) calculations with it is less than 10<sup>-2</sup> eV when predicting non-glassy structure energies (compared to DFT).

Then we proceeded to generate another model. This time aiming, not only to predict the energies of these systems, but also the dynamics proprieties. Because the ion conductivity is related to diffusion coefficient, we needed to generate a different data set, not only with optimized structures in the minimums of their potential energy surfaces (PES), as the previous data set, but mapping the PES during the entire movement of Li<sup>+</sup> in the solid. Now, since a good level accuracy and very long-time dynamic simulations (due to slow Li<sup>+</sup> movement in solids) are required, we create a new time and power saving procedure. In order to be green and faster, we selected four LPS crystals () and performed AIMD simulations with the semiempirical GFN1-xTB, implemented in the program DFTB+. This way obtain two thousand point proprieties of this systems in al very small time. All the simulations were performed in an NVT periodic ensemble at 700K and 300K, 500000 steps with a 0.1 fs time-step. This functional is extremely faster than traditional functionals and can produce reliable geometries but with much less accurate energies and forces. To recover accuracy we extracted two thousand geometries from these trajectory, and perform high accurate SPE calculation on each of them with DFT, B3LYP functional with D3 dispersion correction and Becke-Johnson damping, implemented in the program CP2K. In this way we drastically reduced the time and resources needed treating quantum mechanically relevant structures only, sampling the semi-empirical trajectory obtained less frequently.

## Results and discussion

We used the same configuration of the DNN described before to train it with this new data set and obtained a DLFF. The resulting model testing returns an error of 10<sup>-3</sup> eV on energies and 10<sup>-3</sup> eV/A on forces. With this DLFF we performed a NVT periodic ensemble at 700K molecular dynamics simulation on a much larger scale crystal (more than 6000 atoms), which only took few minutes, and compared the diffusion coefficient of Li<sup>+</sup> this way obtained with the literature. The value obtained with our is compatible with other ab-initio level calculations<sup>5</sup> but obtained thousand time faster. Comparing it to experimental<sup>5</sup> ones it's still overestimated, but this is probably because the data set is still small, we used only two thousand points to train the DNN, to get more accurate results.

## References

- [1] Sun, Y.-K., ACS Energy Lett. 2020, 5, 3221–3223.
- [2] Zhang, Z. et al., Energy Environ. Sci. 2018, 11, 1945–1976
- [3] Grady, Z. A. et al., Front. Energy Res. 2020, 8.
- [4] Han Wang, et al., Computer Physics Communications 228, 2018, 178-184.
- [5] Guo, H., et al. Chemistry of Materials, 2022, 34, 6702-6712.

### Phase-Cycling Two-Dimensional Infrared Spectra from Anharmonic Calculations

A. M. van Bodegraven, Braunschweig, C. R. Jacob, Braunschweig

Anna Maria van Bodegraven, Technische Universität Braunschweig, Institute of Physical and Theoretical Chemistry, Gaußstraße 17, 38106 Braunschweig, Germany

Two-dimensional infrared (2D IR) spectroscopy is a versatile experimental method to probe molecular structures and dynamics on various time scales. 2D IR experiments use multiple laser pulses in an advanced pump-probe setup studying vibrational couplings, energy transfer and relaxation or dynamics like protein folding mechanisms or the formation of hydrogen bonds. The complexity of spectra necessitates computational methods to analyze the results.[1]

Recently published experiments use a phase-cycling scheme to measure 2D IR spectra of isolated molecular ions showing waiting time dependent intensity variations of coupling vibrational modes.[3] Here, we set out a computational approach to calculate collinear phase-cycling spectra similar to our lately developed simulation method VIBRATIONS 2D for standard non-collinear 2D IR spectroscopy [2], presenting first quantum chemical results.

### Literature

- [1] P. Hamm and M. Zanni, *Concepts and methods of 2d infrared spectroscopy* (Cambridge University Press, Feb. 24, 2011).
- [2] J. Brüggemann, M. Wolter, and C. R. Jacob, J. Chem. Phys. 157, 244107 (2022).
- [3] Z. Ma, L. Chen, C. Xu, and J. A. Fournier, J. Phys. Chem. Lett. 14, 9683–9689 (2023).
- [4] H.-S. Tan, J. Chem. Phys. **129**, 124501 (2008).

## Assessing Time-Dependent DFT for Non-Covalent Interactions in Excited Complexes

ARIEL JONES<sup>1</sup> and LARS GOERIGK<sup>2</sup>

<sup>1</sup>School of Chemistry, The University of Melbourne, Parkville, Vic, Australia <sup>2</sup>School of Chemistry, The University of Melbourne, Parkville, Vic, Australia jonesac@student.unimelb.edu.au

Linear-response time-dependent DFT (TD-DFT) in the adiabatic approximation has emerged as the method of choice for excited-state calculations. Within this approximation, the exchange-correlation kernel is known from ground-state DFT. As such, any ground-state Density Functional Approximation (DFA) may be applied to TD-DFT. A known limitation of most DFAs is their inability to accurately capture non-covalent interactions (NCIs), including London dispersion forces, in the ground state. [1] However, the performance of TD-DFT methods for excited-state NCIs remains undertested. Exciplexes (excited complexes) are complexes which are more strongly bound in their electronically excited-state compared to the ground-state. As such, exciplex binding provides a useful case study for excited-state NCIs. We investigate the performance of various DFAs to describe exciplex binding. Additionally, we test whether ground-state dispersion corrections correctly recover dispersion forces in exciplexes.



#### References

 Kristyán, S.; Pulay, P. Can (semi)local density functional theory account for the London dispersion forces? *Chem. Phys. Lett.* **1994**, *229* (3), 175-180

## Computational Investigation of the PH<sub>3</sub>-H<sub>2</sub> Interaction Potential

#### Abstract

In this study, we present a detailed computational investigation of the interaction potential between PH<sub>3</sub> and H<sub>2</sub> molecules using highlevel *ab initio* methods. The potential energy surface (PES) was computed at the CCSD(T)-F12 level of theory with the aug-cc-pVTZ basis set, utilizing the MOLPRO software for all calculations. The six-atom system is described by 219 primitive Gaussian functions contracted into 165 functions. The interaction was mapped as a function of the radial distance (R) and the angular coordinates ( $\theta$ ,  $\phi$ ), with a detailed grid of 6630 geometries spanning R values from 4 to 30 Bohr. Angular variations were explored with uniform steps of 15° for both  $\theta$  and  $\phi$ . The Basis Set Superposition Error (BSSE) was corrected at all geometries using the counterpoise method. This study provides a precise description of the PH<sub>3</sub>-H<sub>2</sub> interaction potential, offering an accurate dataset for further theoretical and computational exploration of intermolecular forces.

In total, 6630 geometries were calculated, and the potential energy was corrected for Basis Set Superposition Error (BSSE) using the counterpoise method. This correction ensures the accuracy of the computed interaction energies. A 2D cut of the PES for the PH<sub>3</sub>-H<sub>2</sub> system was generated, showing the variation of energy as a function of R and  $\theta$  at a fixed value of  $\phi = 0^{\circ}$ . These computations provide a valuable and accurate representation of the intermolecular forces between PH<sub>3</sub> and H<sub>2</sub>, forming a reliable basis for further theoretical and experimental studies in the field of molecular interactions.

## Studying Nuclear Quantum Effects in Sub- and Supercritical Water with the SL-PIHMC-MIX method

Bo Thomsen (Thomsen.bo@jaea.go.jp), Motoyuki Shiga

Japan Atomic Energy Agency, Center for Computational Science and E-systems

Nuclear quantum effects (NQEs) are essential in the accurate modelling of water. In previous works, we used *ab initio* path integral molecular dynamics (AI-PIMD) to study the water auto-ionization constant [1], suband supercritical water [2], and waters isotopologues at room temperature [3]. These studies demonstrated that NQEs were necessary for obtaining quantitatively correct results, and that even at sub- and supercritical temperatures, NQEs continue to influence hydrogen bond contacts. However, AI-PIMD comes with the significant cost of calculating energies and gradients for multiple structures at each timestep, making it impractical for long simulations and large systems when based on *ab initio* energies and gradients.

In this study, we use the self-learning path integral hybrid Monte Carlo with mixed *ab initio* and machine learning potentials (SL-PIHMC-MIX) [4], which builds on the earlier self-learning hybrid Monte Carlo (SL-HMC) method [5]. This method requires an order of magnitude fewer *ab initio* calculations while still reproducing the results of AI-PIMD. It is implemented in the locally developed PIMD program [6], interfaced with CP2K [7] for electronic energy and gradient calculations, and with AENET [8] for obtaining and training machine learning potential energies and gradients.

One example application is shown in Figure 1, which illustrates the structure of high-temperature water. Here, we compare the performance of DFT functionals RPBE-D3, SCAN, rev-vdW-DF2, and rev-PBE0-D3, spanning the GGA to hybrid levels of Jacob's ladder of accuracy. While RPBE-D3 provides the best description of the first hydration shell at room temperature, its agreement with experimental data declines at higher temperatures. SCAN and rev-vdW-DF2 over-structure water at room temperature and appear to overestimate density at elevated temperatures. Although rev-PBE0-D3 also over-structures at room temperature, its agreement at higher temperatures suggests it better describes non-hydrogen bond interactions between water molecules, which are important at elevated temperatures.



Figure 1: The O-O RDFs from experiment [9] (black) and PIHMC-MIX simulations using the RPBE-D3 (blue), SCAN (red), rev-vdW-DF2 (green) and rev-PBE0-D3 (orange) functionals. The title bar above each plot shows the conditions used in the NPT ensemble and the corresponding experimental density in g/ml.

- [1] B. Thomsen, and M. Shiga, J. Chem. Phys. 154, 084117 (2021).
- [2] B. Thomsen, and M. Shiga, J. Chem. Phys. 155, 194107 (2021).
- [3] B. Thomsen, and M. Shiga, Phys. Chem. Chem. Phys. 22, 10851 (2022).
- [4] B. Thomsen, Y. Nagai, K. Kobayashi, I. Hamada and M. Shiga, J. Chem. Phys. 161, 204109 (2024)
- [5] Y. Nagai, M. Okumura, K. Kobayashi, and M. Shiga, Phys. Rev. B (Rapid Commun.) 102, 041124 (2020).
- [6] M. Shiga, PIMD, https://ccse.jaea.go.jp/software/PIMD/index.en.html (2023).
- [7] T. D. Kühne et al, J. Chem. Phys. 152, 194103 (2020).
- [8] N. Artrith, and A. Urban, Comput. Mater. Sci. 114, 135-150 (2016).
- [9] A. K. Soper, Chem. phys. 258, 121-137 (2000).

## Analyzing Microsolvation Shell Buildup of Helium Around the Zundel Cation Using Highly Accurate Path Integral Simulations Based on Machine Learning Potentials

## Eduard Gil Olaria<sup>1</sup>, Harald Forbert<sup>2</sup>, and Dominik Marx<sup>1</sup>

<sup>1</sup> Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany
<sup>2</sup> Center for Solvation Science ZEMOS, Ruhr-Universität Bochum, 44780 Bochum, Germany

Abstract: Since its discovery in 1968, the Zundel cation  $(H_5O_2^+)$  has been extensively studied as a fundamental hydrogen-bonded system. Despite its simple molecular formula, its highly fluxional nature and shallow potential energy surface yield intricate observables, such as a complex infrared spectrum. Advanced experimental techniques, including tagging with helium atoms or embedding in helium nanodroplets, have enabled detailed infrared spectroscopy of this challenging species. Recent experiments examining a single Zundel cation tagged with helium reveal no spectral changes upon adding up to eight helium atoms in the 3500–3700 cm<sup>-1</sup> range. In this project, the Zundel cation surrounded by different helium environments at ultracold temperatures will be simulated using high dimensional neural network potentials trained at CCSD(T\*)-F12a/aug-cc-pVTZ level in combination with path integral simulations, thus leading a deeper understanding of the arising helium-Zundel interactions, the buildup of the helium microsolvation shell, and its effects on the properties of the Zundel cation. Our results align with experimental evidence showing minimal structural changes with the different helium microsolvation shells, though many aspects of the helium atoms' quantum dynamics remain unresolved.

## Modeling Local Aerosol Surface Environments: Hydration and Embedding of Organic Acids and Ions

Georg Baadsgaard Trolle, Jakub Kubečka, Jonas Elm

Department of Chemistry, Aarhus University, 8000 Aarhus C, Denmark

Keywords: Molecular modeling, quantum chemistry, micro-hydration, embedding

Presenting author email: georgbt@chem.au.dk

Aerosols are microscopic particles in the atmosphere comprised of organic and inorganic matter. It is well-documented that new aerosol particle formation proceeds through small molecular clusters and an unknown transition regime before assembling to larger particles. However, despite the major advances in aerosol research recently, the influence of small organic acids on particle formation and growth, chemical composition and physico-chemical properties have remained elusive.

One such small organic acid is pyruvic acid, omnipresent in nature and found both in the gas phase and in the particle phase of the atmosphere as well as in the aqueous phase of the oceans. Despite much previously reported work on the photochemistry and stability of pyruvic acid in different chemical environments, the simultaneous interactions between gas-phase pyruvic acid, or similar acids, and water and ions are not well-understood. Thus, employing quantum chemical methods we studied micro-hydrated molecular clusters containing pyruvic acid and its structurally analog carboxylic acids lactic acid, propionic acid, and 2,2-dihydroxypropanoic acid. We carried out a detailed configurational sampling for all the clusters to identify the representative, low free energy structures. Further, we examined the geometries, binding free energies, hydrate distributions, and IR absorption spectra to assess the molecular and thermodynamic stabilities of the clusters. By viewing such clusters as a representation of the local aerosol surface environment, the overall aim is to elucidate how such small organic acids interact with the surface and how they are ultimately embedded in the particle.

We find that cluster geometries and especially their water structure are very sensitive to the presence and nature of ions. We show that the hydration of the organic acids is thermodynamically unfavorable in the gas phase and ions are necessary for mediating interactions between organic acids and water thereby stabilizing the clusters. Finally, we find a clear correlation between decreasing pyruvic acid carboxylic O–H stretching frequencies and increasing intensity when adding more water to the clusters and we observe a trend in decreasing O–H frequencies when adding ions to the clusters. The conclusions of this study may help interpret experimental spectroscopic data and could represent a crucial initial step toward unraveling the mechanisms driving the solvation and embedding of organic acids in aerosol particles.

# Quantifying Conformational Space Coverage with Convex Hulls

Jaiming Chung, Bienfait Isamura, Paul Popelier

Conformational sampling is a key component of any machine learning force field workflow. Existing methods for determining conformational space has been covered can be ambiguous, reference-dependent and qualitative. Convex hulls are a widely used technique for characterising the shape of a point cloud. We propose a new method for quantifying conformational space coverage using convex hull volumes. These volumes behave according to chemical intuition, and are therefore useful measures that can be used to compare different sampling methods or datasets.

## Redesigning Nucleic Acid – Metal Ion Interaction Potentials for Molecular Dynamics Force Fields

### <u>Leo Christanell</u><sup>1</sup>, Karl-Jakob König<sup>1</sup>, Hongqing Zhao<sup>1</sup>, Petros Mavromatis<sup>1</sup>, Julian Holzinger<sup>1</sup>, Anne Schütz<sup>1</sup> and Benjamin P. Fingerhut<sup>1</sup>

<sup>1</sup>Department Chemie and Centre for NanoScience, Ludwig-Maximilians-Universität München, Germany

Electrostatic interactions of nucleic acids and metal ions critically affect their structure and functionality in solution<sup>1–3</sup>. It has therefore been a constant effort of molecular dynamics (MD) simulations of aqueous nucleic acid systems to improve the parameterization of non-bonded interactions between metal ions and the biomolecular surface, in particular for doubly charged ions like Mg<sup>2+</sup>. Several force fields have been proposed to tackle the problems of missing polarization effects and mismatched surface-ion interactions, such as the 12-6-4 Lennard-Jones model, which introduces an additional induced dipole term into the 12-6 Lennard-Jones interaction model. An alternative approach relies on the screening of charges in condensed liquid phase, which can be effectively accounted for in a mean field way via a rescaling of charges<sup>4</sup>. However, recent investigations have shown that simulated RNA/DNA-ion interactions are too strong, resulting in inaccurate ion-water exchange rates<sup>4</sup>. In this work we unify the preceding approaches and parameterize a novel force field by fitting to *ab initio* data from a dimethyl phosphate (DMP)-metal-ion model system. Non-integer ion charges and either a 12-6 or 12-6-4 Lennard-Jones potentials are developed. The final nucleic acid-ion parameter sets are then extensively benchmarked in MD simulations. Preliminary results show an improved 12-6-4 parameter set with scaled charges. Further, using experimental NMR data we validate and refine our understanding of ion-phosphate interactions via the dependency of NMR shifts on the Mg<sup>2+</sup> coordination in the DMP-Mg<sup>2+</sup> model system. The novel parametrization approach promises to be a straightforward method to yield metal ion parameter sets for subsequent refinement steps that has broad applicability for atomistic force field simulations of nucleic acids in solution.

### References:

<sup>1</sup> A. Pyle, "Metal ions in the structure and function of RNA," JBIC J. Biol. Inorg. Chem. **7**(7–8), 679–690 (2002).

<sup>2</sup> J.K. Frederiksen, N.-S. Li, R. Das, D. Herschlag, and J.A. Piccirilli, "Metal-ion rescue revisited: Biochemical detection of site-bound metal ions important for RNA folding," RNA **18**(6), 1123–1141 (2012).

<sup>3</sup> B.P. Fingerhut, "The mutual interactions of RNA, counterions and water – quantifying the electrostatics at the phosphate–water interface," Chem. Commun. **57**(96), 12880–12897 (2021).

<sup>4</sup> I. Leontyev, and A. Stuchebrukhov, "Accounting for electronic polarization in non-polarizable force fields," Phys. Chem. Chem. Phys. **13**(7), 2613 (2011).

<sup>5</sup> K.K. Grotz, S. Cruz-León, and N. Schwierz, "Optimized Magnesium Force Field Parameters for Biomolecular Simulations with Accurate Solvation, Ion-Binding, and Water-Exchange Properties," J. Chem. Theory Comput. **17**(4), 2530–2540 (2021).

# Benchmarking machine learning potentials for aqueous electrolyte solutions

Lisa Hetzel<sup>a</sup>, Christopher J. Stein<sup>a</sup>

#### <sup>a</sup>TUM School of Natural Sciences, Technical University Munich Lichtenbergstr. 4, D-85748 Garching

In this work, we evaluate surrogate potentials for aqueous electrolyte solutions, a notoriously challenging task due to the long-range interactions of ions, which are often truncated by the cutoff radius used in most machine learning potentials to maintain scalability.<sup>[1]</sup> In particular, we select the MACE<sup>[2]</sup> potential, which, due to its message-passing architecture, spans a longer range by construction. We investigated two systems: First, aqueous alkali halide solutions as a toy model, and second, a charged graphene-water interface. Both systems are of high importance in the context of electrocatalysis.

In particular, we investigate the influence of model hyperparameters, such as the level of equivariance or number of uncoupled channels, on the MACE model's performance on both static and dynamic properties. Besides, we discuss the accompanying computational cost with an increase in the model's complexity.

Machine learning potentials for the aqueous potassium chloride system featured sufficient stability during molecular dynamics simulations without using an active learning scheme. At the same time, sodium cations proved to be more challenging to fit, most likely due to their smaller ion radius and stronger electric field, resulting in less structural diversity and higher force sensitivity, which makes the data set generation more challenging. Additionally, we benchmarked the computational cost for training the models and performing simulations with the obtained models using differing architectures. Nonetheless, the obtained computed physical quantities exhibited good agreement with experimental data, though limited by the chosen reference method. Furthermore, no significant differences between properties computed from a simulation employing the MACE potentials with architectures of different complexity and levels of equivariance were obtained.

The challenge for training a MACE potential for a charged system is that the isotropy of the system and, thus, the rotational symmetry of each atom's environment is violated for a directed electric field. While using a model architecture with sufficient range, we find that MACE can predict the energies and forces without significant outliers. The results are promising for the application of MACE potentials in the context of electrocatalysis.

[1] J. Behler, Chem. Rev. 2021, 121, 10037–10072.

[2] I. Batatia, D. P. Kovács, G. N. C. Simm, C. Ortner, G. Csányi, 2022.

## Environmental Shifts of Hydrogen Bond Vibrations in Ne Cryomatrices and He Nanodroplets – Building up an Experimental Database

<u>M. Bödecker, Göttingen/GER 37077</u>, M. A. Suhm, Institut für Physikalische Chemie, Universität Göttingen, Tammannstraße 6, 37077 Göttingen/GER

Cold rare gas embeddings such as neon matrices or helium nanodroplets are a common technique in vibrational spectroscopy used to immobilize clusters, stabilize high energy conformations and simplify infrared spectra by reducing rotational structure and suppressing hot vibrational transitions. [1] These techniques offer a significant advantage compared to measurements in supersonic jet expansions, being more sensitive and better able to trap metastable conformers. However, embedding molecules and molecular clusters in cold matrices leads to subtle shifts in vibrational frequencies due to environmental effects, which makes this experiment less suitable for direct comparison to gas phase theory.

If reliable correlations between vibrational wavenumbers in the cold gas phase and in matrix isolation experiments can be established, this could provide a powerful tool for isomer assignments in IR spectra or for predictions of cold gas phase wavenumbers from He nanodroplet or Ne matrix values, respectively.

The aim of this work is to build up an experimental database of neon matrix shifts [2] and helium nanodroplet shifts [3] of hydrogen-bonded O–H stretching fundamentals that can serve as a benchmark for theoretical simulations of the matrix effect.

Literature: [1] J. P. Toennies; A. F. Vilesov, Angew. Chem. Int. Ed., 2004, 43, 2622–2648. [2] M. Bödecker, D. Mihrin, M. A. Suhm, R. Wugt Larsen, J. Phys. Chem. A, 2024, 128, 7124–7136. [3] M. Y. Choi, G. E. Douberly, T. M. Falconer, W. K. Lewis, C. M. Lindsay, J. M. Merritt, P. L. Stiles, R. E. Miller, Int. Rev. Phys. Chem., 2006, 25:1-2, 15–75.

# Oriented stagnation graphs of the current-density field of molecules

M. Dimitrova<sup>a</sup>, D. Sundholm<sup>a</sup>, and R. J. F. Berger<sup>b</sup>

*a* Department of Chemistry, University of Helsinki *b* Department of Chemistry, University of Salzburg e-mail: maria.dimitrova@helsinki.fi

The Ampère-Maxwell law defines the classical relationship between a magnetic field and an electric current in a circuit. It turns out that the equation is applicable to quantum chemistry calculations of the magnetically induced current density (MICD) from the induced magnetic field in a molecule placed in an external magnetic field. Typically, MICD is calculated using the nuclear magnetic shielding tensor. In our recent studies, we have shown that it is possible to calculate the shielding tensor components on a loop to obtain the total MICD of a molecule or to accurately quantify the strength of the flux of a current-density vortex.<sup>1-3</sup>

The total MICD can be obtained by defining an integration plane originating at the centre of the molecule and extending to infinity above, below and away from the centre.1 Traditionally, the net current strength is calculated by integrating the MICD on points lying on the plane. Instead, we have calculated the nuclear magnetic shielding tensor on the closed loop defined by the perimeter of the integration plane.

The current-density field usually consists of multiple vortices. The zero points in the field define the stagnation graph (SG) of the MICD. We have performed thorough analyses of the SG in a series of molecules, and we have calculated the nuclear magnetic shielding tensor on the stagnation lines.<sup>2,3</sup> This approach allowed us to accurately assign the strength of each vortex, which is in contrast with the traditional approximate method of integrating the MICD on a series of 2D grids.



Figure 1: The two vortices in the magnetically induced current-density field of LiH.

- [1] R. F. J. Berger, M. Dimitrova, R. T. Nasibullin, R. R. Valiev, D. Sundholm. PCCP 2021, 24, 624.
- [2] R. F. J. Berger, M. Dimitrova. PCCP 2022, 24, 23089.
- [3] R. F. J. Berger, M. Dimitrova. (in preparation) 2024.

### **Research Data Management in Theoretical Chemistry**

Mario Wolter<sup>1</sup>, Christoph R. Jacob<sup>1</sup>

<sup>1</sup>Institute of Physical and Theoretical Chemistry, TU Braunschweig, Gaußstraße 17, 38106 Braunschweig, Germany m.wolter@tu-braunschweig.de, c.jacob@tu-braunschweig.de

**Collaboration Partners**: Vera Krewald (TU Darmstadt), Christian Mück-Lichtenfeld (Universität Münster), Maren Podewitz (TU Wien), Ralf Tonner-Zech (Universität Leipzig)

Effective data management has emerged as a critical component for scientific advancement and collaboration. While a detailed manual record-keeping of all parameters seems to be the standard in experimental disciplines, researchers in theoretical chemistry often simply rely on their folder structures and the saved input and output files. Arguably, this can lead to massive problems in reproduction of simulation results when things out of the scope of these files change.

In the first funding phase NFDI4Chem developed an Electronic Lab Notebook (ELN), extended data and meta-data formats, and ontologies to digitalize experimental data. In our ongoing work, we took the fist steps of making the Chemotion ELN[1] fit for use in theoretical chemistry. In order to do this, new functions had to be implemented in the ELN (e.g. reading general data formats like json and an API to interact with the ELN).

To make this integration of quantum-chemical data even more seamless an ontology for theoretical chemistry is absolutely essential. From this, the universal data (and metadata) format and even the structure of quantum-chemical code can be derived. The barriers of integration into already existing solutions (e.g. NOMAD[2]) and even on-site solutions are reduced to an absolute minimum. An adoption of these data-formats as standard output formats of quantum-chemical software will improve data handling not only in conjunction with ELNs. In the mean time, already established frameworks for quantum-chemical workflows (e.g. pyADF[3]) are able to fill this gap.

Once established and sufficently automated, this framework will enable real-time data capture, organization, and annotation. Researchers can effortlessly navigate, interpret, and reproduce computational results within the context of their electronic lab journals. By enhancing Findability, Accessibility, Interoperability, and Reusability[4], it not only facilitates effective collaboration but also strengthens the integrity and longevity of quantum-chemical research data.

- [1] Tremouilhac P, Nguyen A, Huang YC, et al., *J Cheminform.*, 9(1), 54 (2017) https://chemotion.net/
- [2] Scheffler, M., Aeschlimann, M., Albrecht, M., et al., *Nature* 604, 635–642 (2022) https://nomad-lab.eu, https://nomad-lab.eu/nomad-lab/nomad-oasis.html
- [3] Jacob, C.R., Beyhan, S.M., Bulo, R.E., Gomes, A.S.P., Götz, A.W., Kiewisch, K., Sikkema, J. and Visscher, L., *J. Comput. Chem.*, 32: 2328-2338 (2011)
- [4] Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J, et al., *Scientific data*, 3, 160018 (2016)

## Developing and Applying a Theoretical Receptor Approach to Novel Boroncontaining Autotaxin Inhibitors

Mauno J. Lius<sup>a</sup>, Andrew G. Leach<sup>a</sup>, Cristina Trujillo del valle<sup>b</sup>, Jonathan M. Skelton<sup>b</sup>

<sup>a</sup> Division of Pharmacy and Optometry, University of Manchester, Oxford Rd, Manchester M13 9PL

<sup>b</sup> Department of Chemistry, University of Manchester, Manchester M13 9PL



The use of boron-containing compounds (BCCs) in the pharmaceutical industry has been steadily increasing in the past decade. The interest arises because of boron's Lewis acidity and tendency to form reversible covalent interactions with proteins. However, computational methods to further investigate the energetic consequences of such binding is scarce and made difficult by the lack of efficient methods to model drug-protein complexes at the quantum-mechanical level. The current work has employed the use of a 'theoretical receptor' method<sup>1</sup> that aims to provide the best compromise between speed and accuracy in determining the binding geometry and affinity of drug-like compounds in a target protein by systematically extracting a sub-region of the protein (typically an active site) whose size is more amenable to traditional quantum-mechanical methods. In doing so, more accurate computational methods can be employed to best model the interaction of BCCs and their protein targets.

The method has been applied to the autotaxin (ATX) enzyme which is implicated in numerous diseases including cancers, pulmonary fibrosis, and neuropathy. This target was chosen because a range of boronic acid inhibitors have been identified in the literature making it an ideal protein to test our approach. Using GFN2-xTB<sup>2</sup>, accurate binding geometries and associated binding energies ( $\Delta E$ ) have been obtained for a range of inhibitors including many BCCs. The theoretical receptor method reveals that covalently bound boronic acids have a better binding energy than their non-covalent equivalents. Our multilinear regression model (including AlogP) suggests that we can predict pIC<sub>50</sub> values (N=37, R<sup>2</sup>=0.612, RMSE=0.825) with good overall accuracy, especially for boronic acids.

Future work on the ATX theoretical receptor will include conformational analysis of the active site to gauge whether it may improve  $pIC_{50}$  predictions for outlier ligands in the dataset.

[1] I. Lukac, H. Abdelhakim, A. R. Ward, S. A. St-Gallay, J. C. Madden, A. G. Leach, A. Chem. Sci, (2019) 10, 2218-2227.

[2] C. Bannwarth, S. Ehlert and S. Grimme, J. Chem. Theory Comput., (2019) 15, 1652–1671.

## DL\_FFLUX: A Robust, Accurate, and Cost-Effective Machine Learning Potential

Mohamadhosein Nosratjoo and Paul L.A. Popelier

Department of Chemistry, The University of Manchester, Oxford Road, Manchester M13 9PL, Great Britain

DL\_FFLUX is a state-of-the-art force field that has been developed to make the idea of a **robust** universal force field a reality [1]. By its novel construct, DL\_FFLUX "sees" the atoms and their electrons by utilising the quantum chemical topology (QCT) method. This results in quantum mechanically accurate **atomic** energies and multipole moments (MMs) in every step of an MD simulation. DL\_FFLUX uses machine-learned models to predict these energies and flexible MMs [2]. This approach's unique feature is that the machine learning does not carry out the atomic partitioning. Instead, QCT provides the atomic properties that are trained on, thereby guaranteeing their physical integrity. Our in-house FEREBUS [3] program is used to make models using the Gaussian Process Regression (GPR) method.





In this talk, the following questions will be answered:

- 1. Is it possible to have sub kcal/mol accuracy at speeds close to classical MD?
- 2. How to avoid *extrapolation*?
- 3. Can atoms be partitioned before machine learning?
- 4. Is it possible to have a non-bonded and non-parametrised force field?

### References

[1] Symons, B.C., Bane, M.K. and Popelier, P.L.A, 2021. DL\_FFLUX: a parallel, quantum chemical topology force field. *Journal of Chemical Theory and Computation*, *17*(11), pp.7043-7055.

[2] Symons, B.C. and Popelier, P.LA., 2022. Flexible multipole moments in smooth particle mesh Ewald. *The Journal of chemical physics*, *156*(24).

[3] Burn, M. J.; Popelier, P. L. A., FEREBUS: a High-performance Modern Gaussian Process Regression Engine. Digital Discovery 2023, 2, 152-164.

## Understanding vibrational temperature effects on Mössbauer quadrupole splitting using a sampling approach

Niklas von Rhein<sup>1,\*</sup>, Vera Krewald<sup>1</sup>

<sup>1</sup> Department of Chemistry, Eduard-Zintl-Institut, Technische Universität Darmstadt, 64287 Darmstadt, Germany \* Niklas.von rhein@tu-darmstadt.de

<sup>57</sup>Fe-Mössbauer spectroscopy is a powerful tool for investigating iron-containing species<sup>[1]</sup>. As an isotope specific method, it is especially useful for amorphous systems, such as FeNC catalysts<sup>[2]</sup>. While it yields information on spin state, oxidation state, and some structural parameters, it lacks the ability to elucidate the exact structure on its own. This is overcome by using reference data from model complexes, which are often investigated theoretically with quantum chemistry methods<sup>[2]</sup>.

In comparing experiment and theory, however, a problem arises: while experimental Mössbauer spectroscopy is temperature dependent<sup>[1]</sup>, quantum chemical calculations assume resting nuclei at absolute zero. This necessitates the use of low temperature Mössbauer experiments, usually below 80 K, to enable comparison with quantum chemical calculations<sup>[2]</sup>. Such low-temperature measurements can be experimentally highly demanding and cannot probe reactive conditions, thus limiting the use of theoretical model complexes to non-reactive Mössbauer experiments.

We have developed a theoretical method that helps understanding the temperature dependence of the quadrupole splitting due to vibrational deformation, which is probed via an ensemble of structures created by Wigner sampling<sup>[3]</sup>. Using these large ensembles we also gained a deeper insight into the connection between Mössbauer and structural or electronic parameters.



*Figure 1.* Schematic representation of our method for an exemplary FeN<sub>4</sub> center.

### References

[1] *P. Gütlich, E. Bill, A. X. Trautwein*, Mössbauer Spectroscopy and Transition Metal Chemistry, Springer Berlin Heidelberg, Berlin, Heidelberg, **2011**.

[2] C. Gallenkamp, U. I. Kramm, J. Proppe, V. Krewald, Calibration of computational Mössbauer spectroscopy to unravel active sites in FeNC catalysts for the oxygen reduction reaction, *Int. J. of Quantum Chemistry* 121 (**2021**) e26394

[3] *N. von Rhein, V. Krewald*, The temperature dependence of Mössbauer quadrupole splitting values: a quantum chemical analysis, *ChemRxiv* (**2024**); This content is a preprint and has not been peer-reviewed

#### Efficient and Accurate Local Equivariant Deep Neural Network Interatomic Potential for Large-Scale Porous Liquids Simulations

#### Ouail Zakary<sup>a</sup>, Perttu Lantto<sup>a</sup>

<sup>a</sup>NMR Research Unit, University of Oulu, P.O. Box 3000, FI-90014, Finland.

Porous materials serve diverse applications like molecular separations and catalysis [1], offering an energy-efficient method for capturing greenhouse gases (CO<sub>2</sub>, and CH<sub>4</sub>) and valuable noble gases (Xe, Ar, and Kr). Xenon — vital in optics, medicine, and nuclear fission processes [2] — poses extraction challenges due to its low atmospheric abundance (0.087 ppm by volume) and inertness, driving high commercial costs. Porous materials must exhibit precise size selectivity [3] and high adsorption capacity for effective xenon isolation in commercial applications. The recent development of porous liquids (PLs) with cavities formed by porous organic cages (POCs) has shown promise in addressing these challenges [4, 5]. Understanding the binding, occupancies, dynamics, and equilibrium between the host (PL and POC) and the guest (Xe) is vital in engineering new POCs with targeted functionalities. Molecular dynamics (MD) simulations have proven essential for understanding and exploring the physicochemical processes governing these systems. In MD simulations, atom movements are governed by forces derived as gradients of the potential energy surface (PES) of the system. Typically, the PES is accurately obtained by calculating the electronic structure using methods like density functional theory (DFT). While this combination of MD with DFT offers precise interatomic forces, their computational scaling limits simulations to tens of picoseconds and a few hundred atoms, which is far from capturing realistic timescales and size of these porous systems. Classical models for computing PES, although faster, lack accuracy. In recent years, machine learning (ML), particularly neural networks (NNs), has emerged as a promising solution to these limitations [6-8] by learning accurate interatomic potentials from a set of high-fidelity ab initio reference calculations while maintaining computational efficiency.

Here, we present a machine learning interatomic potential (MLIP) model constructed using the local equivariant deep NN architecture, Allegro [9]. This model was trained, validated, and tested on energies, forces, and virials — computed at the DFT-D4 level with PBE functional and periodic boundary conditions — of structures issued from semi-empirical MD simulations performed at RT and 600K, from a dataset comprising over 1600 structures consisting of 600 to 1170 atoms (H, C, N, O, F, Cl, and Xe). These structures comprise one to two xenon atoms in TBA-type, HAP-type, and DCT-type PLs, as well as CC3 POC where Xe atoms are positioned inside or near the cavities of the POC. The dataset encompasses 1.8 million atoms, with 12.5 million data points used in the process.

The MLIP will be applied to provide microscopic interpretation of experimental <sup>129</sup>Xe NMR observations. In this step, we will construct an ML model for magnetic shielding tensors,  $\sigma$ , obtained from the dataset. This ML model will later be used to predict  $\sigma$  in MD simulations to explain experimental, both static spectral and dynamic relaxation, data.

#### **References:**

- [1] A. G. Slater, A. I. Cooper, Science 348 (2015) 988.
- [2] W. M. Stacey, Nuclear Reactor Physics, Wiley-VCH, Weinheim, 2007.
- [3] T. Mitra, K. E. Jelfs, M. Schmidtmann, A. Ahmed, S. Y. Chong, D. J. Adams, A. I. Cooper, Nat. Chem. 5 (2013) 276–281.
- [4] T. Tozawa, J. T. A. Jones, S. I. Swamy, S. Jiang, D. J. Adams, S. Shakespeare, R. Clowes, D. Bradshaw, T. Hasell, S. Y. Chong, C. Tang, S. Thompson, J. Parker, A. Trewin, J. Bacsa, A. M. Z. Slawin, A. Steiner, A. I. Cooper, Nat. Mater. 8 (2009) 973–978.
- [5] L. Chen, P. S. Reiss, S. Y. Chong, D. Holden, K. E. Jelfs, T. Hasell, M. A. Little, A. Kewley, M. E. Briggs, A. Stephenson, K. M. Thomas, J. A.
- Armstrong, J. Bell, J. Busto, R. Noel, J. Liu, D. M. Strachan, P. K. Thallapally, A. I. Cooper, Nat. Mater. 13 (2014) 954–960.
- [6] S. Chmiela, H. E. Sauceda, K. R. Müller, A. Tkatchenko, Nat. Commun. 9 (2018) 3887.
- [7] S. Batzner, A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, B. Kozinsky, Nat. Commun. 13 (2022) 2453.
- [8] J. Vandermause, Y. Xie, J. S. Lim, C. Owen, B. Kozinsky, Nat. Commun. 15 (2021) 5183.
- [9] A. Musaelian, S. Batzner, A. Johansson, L. Sun, C. J. Owen, M. Kornbluth, B. Kozinsky, Nat. Commun. 14 (2023), 579.

## Static Embedding with Pair Coupled Cluster Doubles-Based Methods in DFT

Rahul Chakraborty, <sup>1</sup> Katharina Boguslawski, <sup>1</sup> Paweł Tecmer <sup>1</sup>

<sup>1</sup> Institute of Physics, Faculty of Physics, Astronomy and Informatics Nicolaus Copernicus University in Toruń, Poland

Reliable quantum-chemical modeling of large molecular systems still remains an elusive question. Wave function theory (WFT) based methods, although highly accurate, are limited by high computational expenses. Quantum embedding methods have shown promising results to address this challenge. [1, 2] In this approach, the molecular structure is partitioned into a system part studied by more reliable WFT methods and the environment part modeled by low-level methods. We propose a new embedding scheme, combining pair-coupled cluster doubles (pCCD)-based methods [3] as the WFT-based component for the system fragment with density functional theory (DFT) approximations for the environment. The pCCD method produces reliable results for strongly correlated systems with mean-field computational scaling. We also use *a posteriori* linearized coupled cluster (LCC) corrections on pCCD wave-function to account for the large extent of dynamic correlation, missing in pCCD ansatz and its extension to excited states. [4] All codes have been implemented in the developer version of the PyBEST software package.

We have tested our embedding methods for the challenging uranyl halides  $(UO_2X_4^{2-}, X=F, CI, Br)$ . The accuracy of the proposed model is assessed against the vertical excitation energies and orbital correlation analysis. [5] The embedding models reliably capture changes in the excitation energies going from bare uranyl fragments to supramolecular uranyl halide structures and represent a promising computational approach for excited states in large molecular systems. We also evaluate performance of our method for property calculations of some weakly bonded complexes. [6] Our work serves as a starting point for further development of pCCD-based embedding schemes.

<sup>[1]</sup> T. A. Wesolowski, A. Warshel, J. Phys. Chem. 97, 8050-8053 (1993)

<sup>[2]</sup> P. Huang, E. A. Carter, J. Chem. Phys. 125, 084102 (2006).

<sup>[3]</sup> P. Tecmer, K. Boguslawski, Phys. Chem. Chem. Phys. 24, 23026 (2022).

<sup>[4]</sup> K. Boguslawski, J. Chem. Theory Comput. 15, 18-24 (2019).

<sup>[5]</sup> R. Chakraborty, K. Boguslawski, P. Tecmer, Phys. Chem. Chem. Phys. 25, 25377–25388 (2023).

<sup>[6]</sup> R. Chakraborty, M. M. F. de Moraes, K. Boguslawski, A. Nowak, J. Świerczyński, P. Tecmer J. *Chem. Theory Comput.* **20**, 4689-4702 (2024).

## Predicting the Binding Affinity of Host-guest Complexes

The development of supramolecular macrocyclic chemistry has promoted the prosperous growth of host-guest chemistry.<sup>1</sup> The host-guest chemistry, in turn, is used in a range of applications such as chemical sensors,<sup>2</sup> catalysts,<sup>3</sup> stabilizers,<sup>4</sup> hazardous waste traps,<sup>5</sup> drug delivery vehicles,<sup>6</sup> and enantiomer separation<sup>7</sup>. The continual improvement of existing systems and the development of novel systems would benefit from understanding and quantifying the strength of the interactions between the host and the guest.

The calculation of binding affinity for the host-guest complexes poses two problems: the available methods are not accurate and fast enough to screen hundreds of structures within a reasonable time frame. Such problems are most commonly solved using Density Functional Theory (DFT); however, in many cases, it is not accurate enough to describe the weak interactions that govern the binding process. Often, the coupled cluster methods, e.g., the CCSD(T), are recommended and used for benchmarking to accurately describe the mentioned interactions, but the size of the host-guest complexes renders the use of these methods impractical.<sup>8, 9</sup>

Here, we present an approach that combines the existing experimental data with descriptors obtained from semi-empirical calculations using machine learning. The data for the training and testing process is curated from public sources that provide detailed information on the experimentally observed binding affinity. The experimental data is supplemented with descriptors from GFN2-xTB calculations. The resulting models predict the binding affinity between host-guest complexes in kJ per mol with an R<sup>2</sup> of 0.85 and RMSE of 6.2 kJ per mol for the test set.

<sup>&</sup>lt;sup>1</sup> Yu, J., Qi, D. & Li, J. Design, synthesis and applications of responsive macrocycles. *Communications Chemistry* vol. 3, 1-14(2020).

<sup>&</sup>lt;sup>2</sup> Zhu, Congzhi; Fang, Lei. Mingling Electronic Chemical Sensors with Supramolecular Host-Guest Chemistry. Bentham Science Publishers (2014)

<sup>&</sup>lt;sup>3</sup> Pattabiraman, M., Natarajan, A., Kaliappan, R., Mague, J. T. & Ramamurthy, V. Template directed photodimerization of trans-1,2-bis(n-pyridyl)-ethylenes and stilbazoles in water. *Chemical Communications*, 4542–4544 (2005).

<sup>&</sup>lt;sup>4</sup> Loftsson, T. & Duchêne, D. Cyclodextrins and their pharmaceutical applications. *International Journal of Pharmaceutics* vol. 329, 1–11 (2007).

<sup>&</sup>lt;sup>5</sup> Erdemir, S., Bahadir, M. & Yilmaz, M. Extraction of carcinogenic aromatic amines from aqueous solution using calix[n]arene derivatives as carrier. *J Hazard Mater* 168, 1170–1176 (2009).

<sup>&</sup>lt;sup>6</sup> Walker, S., Oun, R., McInnes, F. J. & Wheate, N. J. The potential of cucurbit[n]urils in drug delivery. *Israel Journal of Chemistry* vol. 51, 616–624 (2011)

<sup>&</sup>lt;sup>7</sup> Szente, L. & Szemán, J. *Cyclodextrins in Analytical Chemistry: Utility of Host-Guest Type Molecular Recognition*. Anal. Chem. 85(17), 8024-8030 (2013).

<sup>&</sup>lt;sup>8</sup> Grimme, S. Density functional theory with London dispersion corrections. *Wiley Interdiscip Rev Comput Mol Sci* 1, 211–228 (2011).

<sup>&</sup>lt;sup>9</sup> Hohenstein, E. G. & Sherrill, C. D. Wavefunction methods for noncovalent interactions. *Wiley Interdisciplinary Reviews: Computational Molecular Science* vol. 2, 304–326 (2011)

<u>Luka Simsive</u><sup>1</sup>, Oleksii Zdorevskyi <sup>1</sup>, Vivek Sharma<sup>1,2</sup> luka.simsive@helsinki.fi

<sup>1</sup>Department of Physics, University of Helsinki, 00014 Helsinki, Finland <sup>2</sup>HiLIFE Institute of Biotechnology, University of Helsinki, 00014 Helsinki, Finland

Mitochondrial Respiratory Complex I is the largest enzyme in the electron transport chain, which plays a pivotal role in aerobic cellular respiration. It couples NADH oxidation to the quinone reduction and pumping of four protons across the inner mitochondrial membrane. The generated charge gradient (proton motive force) is further employed by the ATP synthase for synthesis of ATP – energy currency of living cells. Despite the large amount of high-resolution structural data, the atomistic details of how redox chemistry is coupled to the proton pumping in respiratory complex I still remain unresolved. In our work, we advocate the hypothesis that this coupling may occur with the long-range transfer of protons from the quinone binding tunnel towards the proton pumping site(s). One of the least studied regions along this pathway is the E-channel – protein cavity connecting the quinone chamber with the central hydrophilic axis. In this work, we employ hybrid quantum mechanics/molecular mechanics (QM/MM) molecular dynamics simulations combined with free energy calculations to model the proton transfer within the whole breadth of the E-channel, also considering connectivity to the extended membrane domain. Our results demonstrate that protons can travel across the entire span of the E-channel via the Grotthuss hole mechanism, with low energy barriers and favourable thermodynamics. These results shed light on highly efficient coupling mechanism of respiratory complex I.



## Pyroelectric Effect in Ferroelectric Perovskites Studied with Density Functional Theory

#### Kim Eklund<sup>1</sup> and Antti J. Karttunen<sup>1</sup>

#### <sup>1</sup>Department of Chemistry and Materials Science, Aalto University

Ferroelectric perovskites such as BaTiO<sub>3</sub>, KNbO<sub>3</sub>, and PbTiO<sub>3</sub> and various solid solutions based on them are important functional materials in electrical engineering and related applications. In the pyroelectric effect, exhibited by all ferroelectrics, the spontaneous polarization of the polar material changes upon temperature fluctuation. This electric polarization is extractable as a current. Pyroelectrics are used in various sensor applications and other electronics, but also show potential for waste heat harvesting and catalysis applications. Measurement of the pyroelectric effect is rather complicated experimentally, and computational tools could aid in the design of novel pyroelectrics by providing an accurate route to modelling and screening of the pyroelectric properties.

Study of the finite-temperature phonon properties and finite-temperature atomic displacements is enabled by the self-consistent phonon calculation (SCPH) formalism implemented in the ALAMODE package developed by Tadano *et al.* [1,2]. We have developed a computational methodology, based on density functional theory (DFT), self-consistent phonon calculations (SCPH), and Berry phase theory of polarization, to evaluate the primary pyroelectric coefficient of tetragonal-structured perovskite ferroelectrics at finite temperatures [3,4]. The total pyroelectric coefficient consists of two parts, primary and secondary, where the secondary effect is a piezoelectric contribution. This too, can be evaluated with quasi-harmonic approximation (QHA) combined with the calculation of elastic and piezoelectric constants. Results obtained for BaTiO<sub>3</sub> show reasonable agreement with experimentally measured values, indicating that the supercell-based methodology can further be used for known and predicted perovskite solid solutions.



Figure 1: Pyroelectric coefficient of BaTiO<sub>3</sub> at finite temperatures calculated with DFT [3].

- [1] T. Tadano, Y. Gohda, and S. Tsuneyuki, J. Phys.: Condens. Matter 2014, 26, 225402.
- [2] T. Tadano and S. Tsuneyuki, Phys. Rev. B 2015, 92, 054301.
- [3] K. Eklund and A.J. Karttunen, J. Phys. Chem. C 2023, 127, 21806.
- [4] K. Eklund and A.J. Karttunen, J. Phys. Chem. C 2024, 128, 16199.

# Importance of a quinol anion in the catalytic cycle of respiratory complex I

**O. Zdorevskyi**<sup>1</sup>, J. Laukkanen<sup>1</sup>, V. Sharma<sup>1,2</sup> <sup>1</sup>Department of Physics, University of Helsinki, Helsinki, Finland <sup>2</sup> HiLIFE Institute of Biotechnology, University of Helsinki, Helsinki, Finland

Molecular function of respiratory complex I (NADH:ubiquinone oxidoreductase) has remained one of the most controversial problems in bioenergetics [1]. This large molecular machine (up to 1 MDa in certain species) utilises the energy from NADH oxidation for the ubiquinone reduction and pumps 4 protons across the inner mitochondrial membrane. Being spatially separated by ca. 200 Å, it is not completely understood how ubiquinone reduction and proton pumping processes are coupled together. Particularly, it remains unknown whether ubiquinone (UQ) is reduced directly to ubiquinol (UQH<sub>2</sub>) in the proximity of the N2 iron-sulfur cluster, or it first forms an intermediate species ubisemiquinone  $(UQ^{*})$  before being fully reduced to  $UQH_{2}$ . In our work [2], we address these questions by employing microsecond-long classical molecular dynamics (MD) in combination with state-of-the-art hybrid quantum-mechanical/molecularmechanical (QM/MM) free energy calculation methods. As model systems, we use high-resolution structures of respiratory complex I from Yarrowia lipolytica [3] and *Sus scrofa* [4]. Upon the first reduction, we see the formation of stable anionic ubisemiquinone which remains trapped at the redox active site for the second reduction step. After getting the second electron, a metastable anionic ubiquinol forms (UQH), which is then repelled back towards the exit of the cavity. These findings corroborate the latest site-directed mutagenesis data [5], and complement the proton-injection coupling model of respiratory complex I [6].

#### References

[1] V. Kravchuk et al. "A universal coupling mechanism of respiratory complex I" *Nature* 2022, 609.7928, 808-814.

[2] Zdorevskyi, Oleksii, Johannes Laukkanen, and Vivek Sharma. "Catalytic relevance of quinol anion in biological energy conversion by respiratory complex I." *bioRxiv*, **2024**, 2024-09.

[3] K. Parey et al. "High-resolution structure and dynamics of mitochondrial complex I - insights into the proton pumping mechanism" *Sci. Adv.* **2021**, 7.46, eabj3221.

[4] J. Gu et al. "The coupling mechanism of mammalian mitochondrial complex I" *Nat. Struct. Mol. Biol.* **2022**, 29.2: 172-182.

[5] C. Harter et al. "Quinone chemistry in respiratory complex I involves protonation of a conserved aspartic acid residue." *FEBS lett.* **2024**.

[6] A. Djurabekova et al. "Long-range electron proton coupling in respiratory complex I—insights from molecular simulations of the quinone chamber and antiporter-like subunits." *Biochem. J.*, **2024**, 481.7, 499-514.

## Characterization of surfactant behavior in aqueous solution with NMR techniques and MD simulations

<u>Pau Mayorga Delgado</u><sup>a</sup>, Ritu Ghanghas<sup>b</sup>, Anne Selent<sup>a</sup>, Sanna Komulainen<sup>a</sup>, Ilari Ainasoja<sup>a</sup>, Kalle Koitila<sup>a</sup>, Pär Håkansson<sup>c</sup>, Nønne Prisle<sup>d</sup>, Perttu Lantto<sup>a</sup>, Jiří Mareš<sup>e</sup>, and Ville-Veikko Telkki<sup>a</sup>

<sup>a</sup>NMR Research Unit, University of Oulu, Finland.

<sup>b</sup>Institute of Chemistry, University of Leiden, Netherlands

<sup>c</sup>Dispelix Oy, Oulu, Finland.

<sup>d</sup>Center for Atmospheric Research (ATMOS), University of Oulu, Finalnd.

<sup>e</sup>Department of Pharmaceutical Sciences, University of Vienna, Austria

Surfactants are long carbon chain molecules with a hydrophilic and hydrophobic group. At low concentration in water, they tend to distribute mainly in the droplet interface, lowering the surface tension of the solution. As concentration increases, the surfactants also distribute in the bulk and starts clustering. At certain concentration threshold, the so called critical micelle concentration (CMC), bigger aggregates are formed, called micelles. The distribution of surfactants in the bulk affects the water activity. This is an important property to understand how the atmospheric water droplets are formed, which would allow us to gain a deeper understanding on cloud formation and the impact of the surfactants on the climate. This study is focused on understanding the structures that surfactants form in water and their dynamics for the case of two mixtures: sodium decanoate with hexanol and sodium hexanoate with SDS. These are examples of surfactants present in the atmosphere. Our results are compared with NMR experimental data.

This study is focused on understanding the structures that surfactants form in water and their dynamics for the case of two mixtures: sodium decanoate with hexanol and sodium hexanoate with SDS. These are examples of surfactants present in the atmosphere. Our results are compared with NMR experimental data.

Molecular dynamics simulations are used to characterize these mixed systems at several concentrations above and below their CMC. The Stokes-Einstein model has been used to compute the diffusion of the surfactants, along with state-of-the-art NMR relaxation modelling, to gain insight on their dynamics.

Our computational results are the following. For the first mixture, we obtained similar diffusion coefficients for both surfactants. The relaxation rates are also similar, agreeing with the experimental trend. In contrast, the diffusion of the hexanoate and SDS from the second mixture differed by an order of magnitude and relaxation pointed towards slower dynamics.

The surfactants from the first mixture tended to self-assemble forming common aggregates. In the second mixture, below the CMC of the hexanoate, only the SDS formed micelles, while above the CMC both surfactants form a variety of aggregate sizes. This work provides a detailed understanding of their size distribution.

Bayesian Optimization Reduced Human Bias in the Adsorption Structure Search of Estradiol on Graphene

Poster abstract for the Helsinki Winter School of Theoretical Chemistry 2024

Saara Sippola / Department of Mechanical and Materials Engineering, University of Turku

17-Beta-estradiol (E2), which is a steroid hormone synthesized from cholesterol, has various impacts on health and the environment. Currently, the gold standard for its measurement in the body is a conventional blood test, but carbon-based electrochemical sensors have been proposed as an alternative because of their advantages, such as rapid analysis time and sensitivity.

To improve the atomic-level understanding of the interactions at the substrate surface, we performed density functional theory (DFT) simulations to study the nature of the adsorption of E2 on pristine graphene. Bayesian Optimization Structure Search (BOSS) was employed to reduce human bias in the determination of the most favorable adsorption configurations. Two stable adsorption minimum configurations were found. Analysis of their electronic properties indicates that E2 physisorbs on graphene.

Moving into more complex carbon-based materials, the importance of finding all possible minimum candidates with automated structure search tools is highlighted. Computational investigations facilitate tailoring substrate materials with outstanding performance and applications in neuroscientific research, fertility monitoring, and clinical trials. Combining them with experimental research carries significant potential to advance sensor design beyond the current state-of-the-art.



Reference to the published research article: Sippola, S., Todorović, M., & Peltola, E. (2024). First-Principles Structure Search Study of 17- $\beta$ -Estradiol Adsorption on Graphene. *ACS omega*, *9*(32), 34684-34691.