Battery Interface Genome – Materials Acceleration Platform

D2.5 – Initial protocols for experimental spectra prediction

VERSION

VERSION	DATE
1.0	

PROJECT INFORMATION

GRANT AGREEMENT NUMBER	957189
PROJECT FULL TITLE	Battery Interface Genome - Materials Acceleration Platform
PROJECT ACRONYM	BIG-MAP
START DATE OF THE PROJECT	1/9-2020
DURATION	3 years
CALL IDENTIFIER	H2020-LC-BAT-2020-3
PROJECT WEBSITE	big-map.eu

DELIVERABLE INFORMATION

WP NO.	WP2
WP LEADER	UU
CONTRIBUTING PARTNERS	3DS, CNR, CTH, UU
NATURE	Report
AUTHORS	D. Prezzi, L. Agosta, S. Rao, J. M. Carlsson, P. Johansson, K. Hermansson, E. Molinari
CONTRIBUTORS	M. Bonacci, F. Hanke, H. Hajiyani, M. Hernandez, P.N.O. Gillespie
CONTRACTUAL DEADLINE	30/04/2022
DELIVERY DATE TO EC	29/04-2022
DISSEMINATION LEVEL (PU/CO)	PU

ACKNOWLEDGMENT

SHORT SUMMARY

The scope of the WP2 work package is to advance the theoretical understanding of complex battery materials and interfaces at the atomistic scale. In this context, "*Task 2.5: Modelling operando spectroscopies and atomistic-level characterization*" aims at devising and applying protocols for calculating electronic properties and predicting experimental spectra (and other characterization techniques) at different levels of accuracy. This deliverable reports on the initial achievements obtained in terms of both protocol development and calculation of spectral features, focusing on different types of relevant systems, i.e. molecules in electrolytes, hybrid interfaces and solid-state interphases. A variety of key characterization techniques routinely employed for ex situ and in situ characterization of battery materials are targeted, such as valence-band photoelectron and UV-Vis spectroscopies, core-level spectroscopies, vibrational spectroscopies. Notably, quantum mechanics and classical molecular dynamics scales are bridged by leveraging artificial intelligence (AI) assisted clustering and analysis. AI techniques are also exploited in order to ease and standardize the comparison between calculated and experimental spectral features. Given the cross-cutting nature of T2.5 activities, they are conducted in close collaboration with multi-scale modelling (WP3), characterization (WP5-6) and infrastructures/interoperability (WP9) partners.

TABLE OF CONTENTS

1. Theory meets experiments: context of the deliverable

WP2 focuses on the generation of data and knowledge from simulations at the electronic and atomistic levels. Atomic-scale modelling can provide data of unrivalled detail and can assist experiments both by exploring compositions, structures and mechanisms in regions that experiments cannot (yet) reach and by providing an atomistic interpretation of complex spectral features, otherwise not easy to relate to the microscopic structure of the probed samples. While this is a fundamental step for understanding the intricate chemical processes that make lithium-ion batteries work – or their possible failures – , it is extremely challenging to realise because of the enormous complexity of battery materials, the lack of detailed structural information available from experiments, the kinetics-thermodynamics interplay, and the long timescales (from a modelling perspective) of many critical processes.

WP2 tackles this challenge by developing new modelling solutions, such as machine-learning (ML) approaches (see D2.2, M18) that can combine the efficiency and cost of classical molecular dynamics (MD) but retaining quantum-mechanical (QM) accuracy. The structural information contained in these atomistic simulations represents an indispensable input for the microscopic and mechanistic interpretation of experimental data (WP5, WP6), which can in turn provide targets for the validation of novel atomistic models (Figure 1). Within this context, WP2 builds on MD and QM geometries and focuses on the definition and application of protocols for calculating electronic properties and modelling experimental spectroscopies, as well as other characterization techniques, in order to directly compare with experimental findings. This deliverable provides the first BIG-MAP report on different protocols aiming at the *in silico* (spectroscopic) characterization of battery materials.
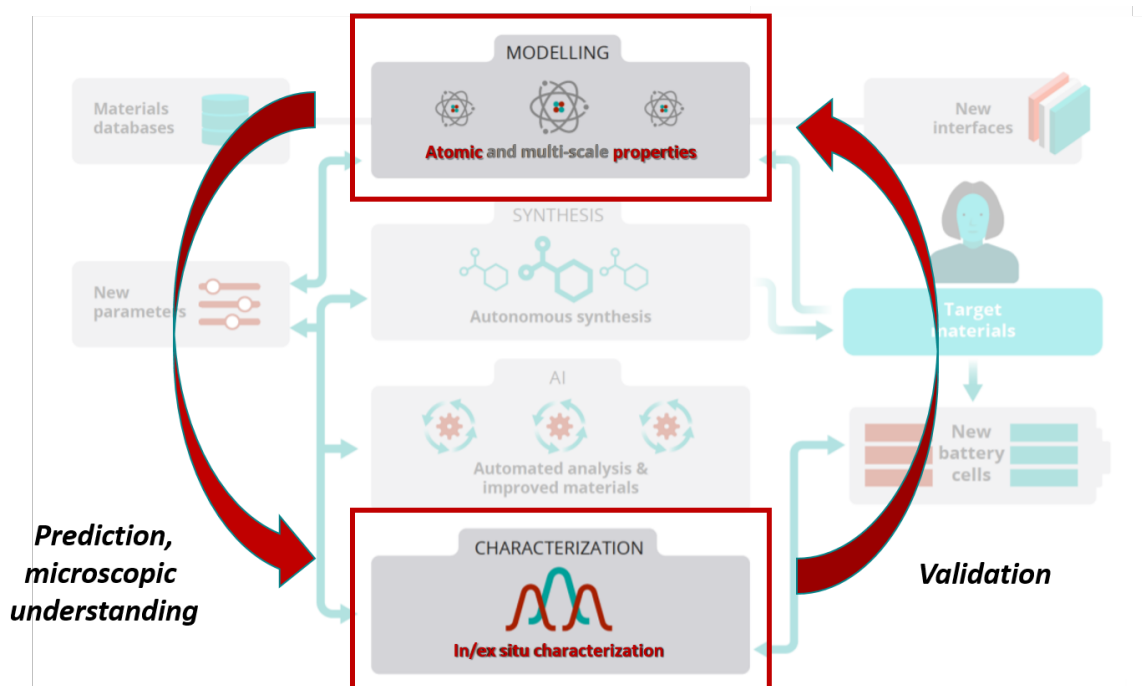

Figure 1. Schematic of the BIG-MAP project. Highlighted in red are the activities conceptually connected within D2.5.

2. Electronic properties, photoelectron and UV-Vis spectroscopies

Calculations of electronic properties – computed at different levels of accuracy depending on the purpose – are key for a number of activities that range from high-throughput screening of new materials and phases to comparison with experimental spectroscopic signatures, e.g. valence band photoelectron spectroscopy or UV-Vis optical spectroscopies.

However, especially for the most accurate techniques, even performing a single calculation might be a non-trivial task. This is the case of many-body perturbation theory (MBPT) approaches, which are the golden standard for electronic structure calculations in solids. These calculations are becoming more feasible, also in a high-throughput perspective, thanks to the (pre-)exascale transition of HPC resources and concurrent efforts on software development (see MaX CoE, http://www.max-centre.eu/). Nonetheless, the many connected steps needed to perform a single calculation, the several interdependent parameters to optimise, as well as the usage of distinct software packages and post-processing applications to obtain a single spectrum – most of them intimately linked to the specific implementation of this complex theory – are to date major obstacles for a widespread usage of these approaches.

Overcoming such difficulties is the key to better support of the experimental community and validation of new atomistic models/approaches – both targets of WP2. To this end, CNR has designed a general protocol that – starting from atomic structure information and with minimal input from the user – provides the requested properties at the requested level of accuracy, as outlined in the schematic of Figure 2. This protocol, currently under implementation, relies on existing and novel AiiDA work chains to orchestrate Quantum ESPRESSO and YAMBO packages for density-functional theory (DFT) and MBPT automated simulations, respectively. CNR has recently deployed a purpose-specific work chain (YamboConvergence work chain, indicated with a gear in Figure 2) that provides an automated search of the optimal converged set of parameters for a MBPT calculation: a key prerequisite for running automated or high-throughput calculations (see D9.3 and https://github.com/yambo-code/yambo-aiida for details). A further work chain for the band structure interpolation using Wannier functions is under development (CNR, EPFL). The full implementation of this protocol will allow us to compute, compare and collect relevant electronic and optical properties (at different levels of accuracy, see violet frame in Figure 2) for comparison with experimental data and model validation.
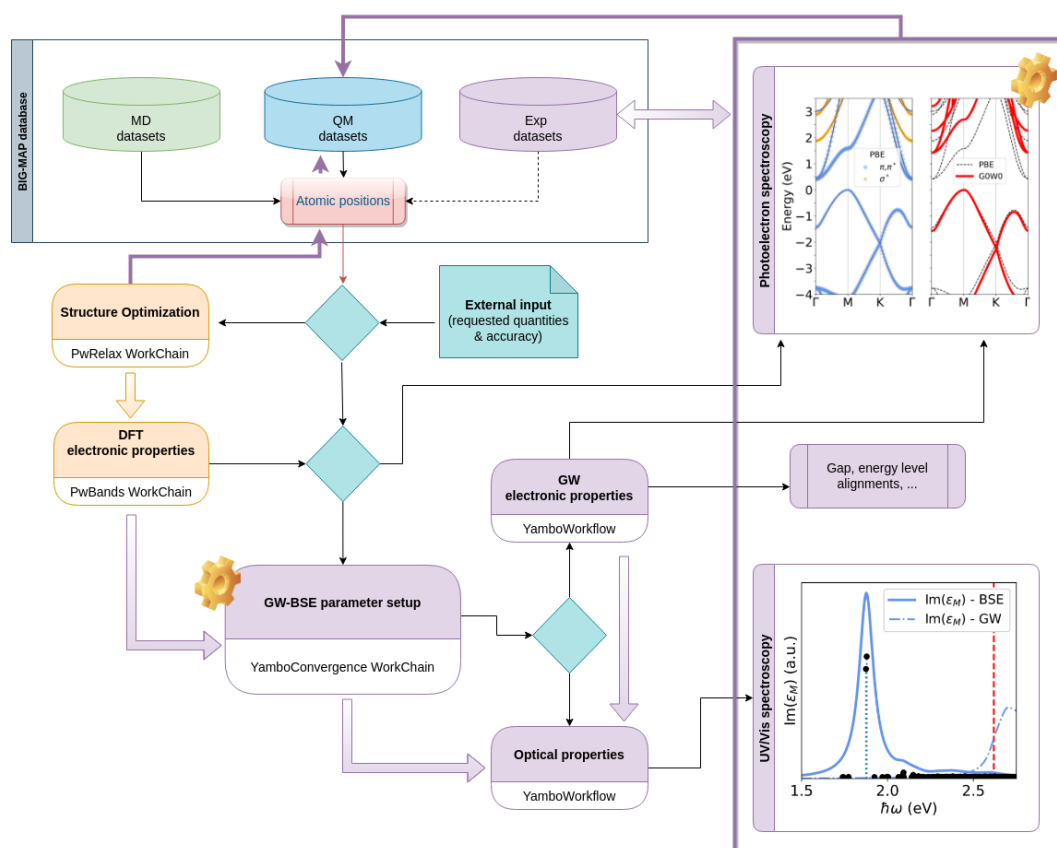
Figure 2. Pictorial representation of the protocol for electronic properties, photoelectron and UV-Vis spectroscopies. The two gears indicate specific work chains, developed and under development within BIG-MAP that, together with existing work chains and protocols, lay the foundation for fully automated calculations. The user is instead expected to provide minimal input (i.e., requested quantities and accuracy). The relevant simulated quantities produced within WP2 are depicted in the violet frame on the right.

3. Core-level spectroscopies

Core-level spectroscopies are fundamental techniques enabling the investigation of the chemical state and bonding as well as the local geometric structure around a target atomic site, by exploiting emission, absorption or scattering processes involving core-level states. X-ray Photoelectron Spectroscopy (XPS), X-ray Absorption Spectroscopy (XAS) and X-ray Raman scattering (XRS) are all different tools successfully employed for ex situ and operando investigation of materials changes upon cycling.

3.1 Electrodes and interfaces

Within WP3, different efforts and multiscale models are directed to the investigation of the morphological changes of materials upon cycling, including the formation of inorganic phases in the solid electrolyte interphase (SEI), as described elsewhere (see D3.1, Sec. 2.2.2 and 2.5 for SEI

builder and lithiated silicon nanowires, respectively). In order to characterize the predicted structures and directly compare with experiments, CNR (in collaboration with PDT and KIT) is developing a protocol for the ab initio simulation of XPS/XAS spectra starting from large MD datasets. Indeed, simulating from first principles the XPS/XAS spectra of several MD snapshots would require running, for each snapshot, about N_{atom} calculations for N_{atom}-large systems, where N_{atom} is of the order of several thousands of atoms. Such a direct approach would be totally out of reach for ab initio simulations. Considering that the XPS/XAS spectrum usually depends on the local environment of the absorbing atom only, CNR has designed a protocol that considerably reduces the computational costs, based on the following steps:

(i) Clustering each atom according to its environment (PDT, KIT, see e.g. D3.3, Sec. 2.4);
(ii) Identifying the minimal cluster size for reproducing the XPS/XAS spectrum of the absorbing atom;
(iii) Computing the XPS/XAS for a selection of clusters (size ~100 atoms) to sample all the identified environments;
(iv) Reproducing the final XPS/XAS signal as weighted average of the computed contributions.
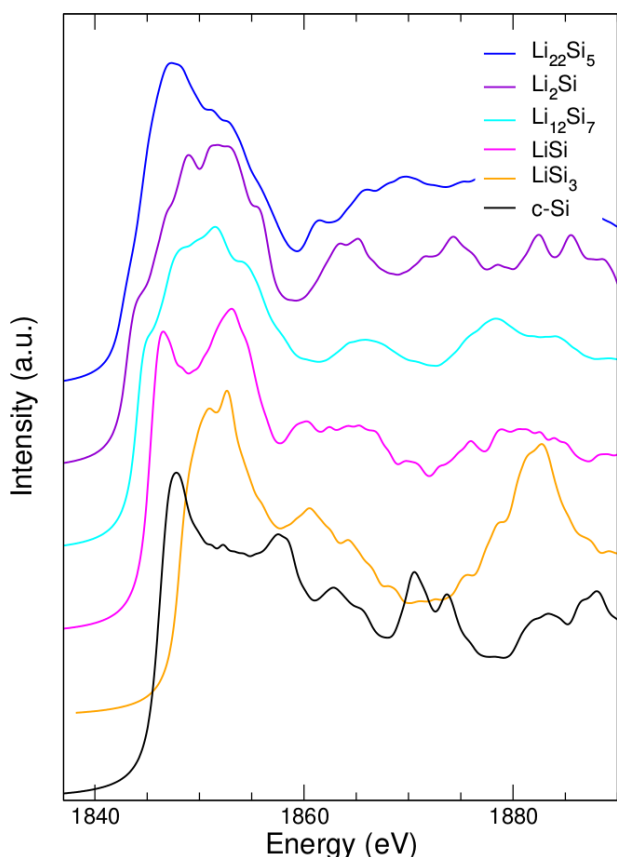
In order to automate the procedure and enable the investigation of large datasets, CNR has implemented specific AiiDA workflows for the ab initio simulation of XPS/XAS spectra, which are based on the Quantum ESPRESSO package (xspectra.x and pw.x codes), as detailed in D9.3. This has allowed the characterization of a number of crystalline reference systems, such as lithium silicides (stable structures taken from the Materials Project, see Figure 3), LiF, Li_2CO_3, $LiPF_6$, and is being used for the characterization of carbon-based structures. The objective is to exploit the computed spectral features to interpret experimental XPS/XAS measurements thus bringing in novel microscopic insights.

Figure 3. XAS spectra at the Si K-edge for some lithium silicides compounds reported in the Materials Project database, as computed using the xspectra-aiida plugin. Intensities are normalize to I_{max} for comparison.

3.2   Electrolyte and solid-electrolyte interphase

Within Task 2.4, 3DS has explored the growth of solid-electrolyte interphase (SEI) structures from mixed molecular dynamics and kinetic Monte Carlo (MC) simulations, namely the degradation mechanisms of electrolytes and salts (collaboration with FZJ). In order to validate the predicted structures from MC/MD simulations and use them to further interpret experimental data, 3DS has devised a complementary protocol for calculating XPS and XAS spectra for molecules in the electrolyte, able to directly process data from coupled MC/MD calculations (see Figure 4). XPS/XAS workflows have been implemented as simulation protocols in the Pipeline Pilot workflow engine (further details on the implementation are reported in D9.3), which uses the DFT code CASTEP available in Materials Studio. To obtain the XAS/XPS spectrum of particular target molecules in the calculated SEI structures (i.e. the result of MC/MD simulations, usually containing more than 10K atoms), this protocol selects a series of molecular clusters from the large MD structure, calculates the contribution from the individual atoms in the target molecule, and sums up the contributions from each atom to obtain the total XPS or XAS spectrum of the molecule. At the next step, different statistical methods, such as kernel density estimation (KDE), are used to estimate the probability density function of the XPS spectra. The simulations demonstrate the capability of predicting shifts in the spectra due to the environment, such that the shifts may reveal if the target molecules are located in the bulk electrolyte or in the SEI region. The theoretical spectra will later be compared with experimental XPS characterization.
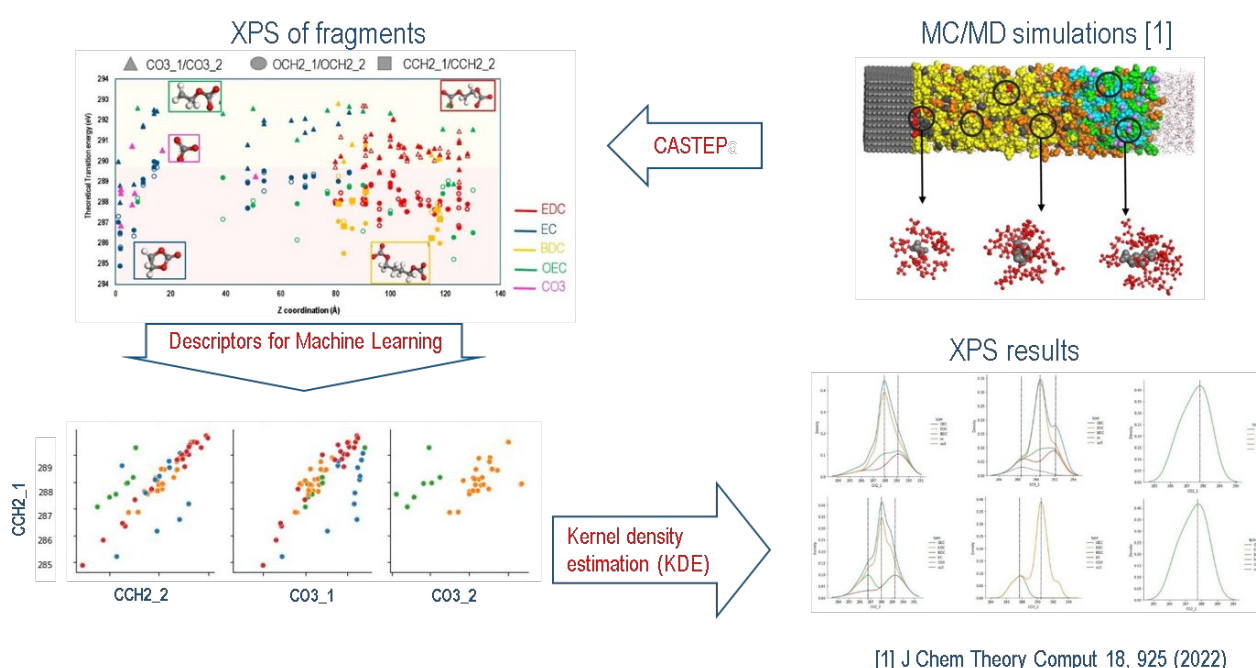

Figure 4. Schematic of the protocol for XPS simulations for electrolyte molecules in the SEI environment, as described in the text.

7

4. Vibrational spectroscopies

Vibrational spectroscopies, such as Raman and infrared (IR) spectroscopies are powerful non-destructive tools that allow retrieving specific structural fingerprint of a materials on the basis of its characteristic vibrational or phonon frequencies. Together with core-level spectroscopies, vibrational spectroscopies are among the most used characterization tools for the investigation of new materials as well as the identification of unknown samples and phases, of crucial importance for the study of battery materials and interphases upon cycling.

4.1   Raman and IR spectra of electrolytes

CTH has focused on collecting and benchmarking existing data on structure, IR, and Raman spectra on electrolytes. We first probe cation-solvent and solvent-solvent interactions under both vacuum and electrolyte conditions in detail through density functional theory calculations using M06-2X/6–311+G(d,p) by systematically computing and comparing binding/formation energies, solvation structures, IR and Raman spectra. We do this for $Li^+(S)_n$ solvation complexes, where solvents/additives 'S' is EC, PC, VC, VEC and FEC and the coordination number, n, varies from n=1 to 5. The solvation effects and Li coordination, along with the calculated spectra, were reported earlier in the mid-term report (Sec. 1.2.2.). IR and Raman spectra are also computed for $LiPF_6$ solvates in carbonate electrolytes. The computed IR and Raman data are compared with the experimental Raman measurements in collaboration with WP5. The data generated will also contribute to the WP9-KD8 joint effort towards an *Automated Workflow Demonstrator* for the integration of simulations and experiments. The goal of this spectroscopic demonstrator is to build the tools and capabilities that will facilitate the flow of data between experimental and calculated spectra. Central to this goal is the problem of data representation: simulated spectra from QM zero-Kelvin calculations are histograms (i.e., have zero broadening), while experimental spectra are distributions broadened by fundamental, instrumental, and environmental factors (Figure 5a). Representing both datasets in the same way is not only required for establishing similarity metrics to carry out clustering tasks, but also would enable combining experimental and simulated datasets to build the large training sets required for classification and regression tasks. This will be achieved by training convolutional neural networks to correctly predict the presence of functional groups (Figure 5b) from the spectral data.
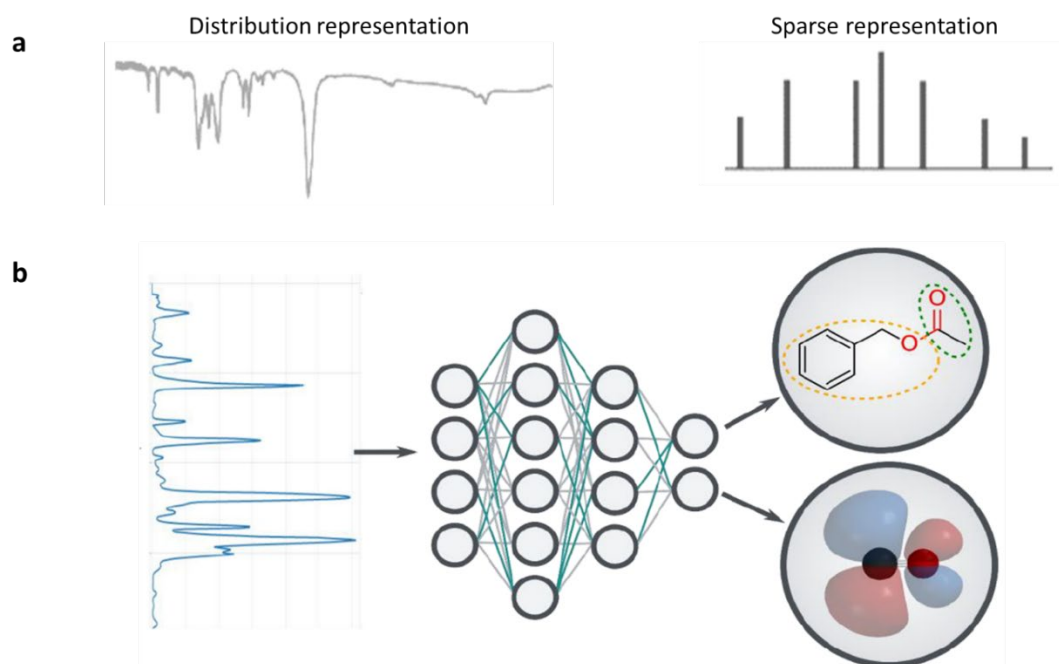
Figure 5. (a) Possible way to represent spectra: as a continuum distribution of overlapping bands (left), and as a sparse vector with non-zero elements in the positions where a band occurs (right). (b) Using convolutional neural networks to identify functional groups and atomic environments from spectra.

4.2   Infrared spectra of molecules at surfaces/interfaces

UU has focussed on methods for computing accurate IR spectra from ab-initio molecular dynamics (AIMD) simulations for an interface. Adsorbates at surfaces/interfaces represent particularly challenging cases because of the symmetry breaking (not to mention the frequent lack of detailed structural information). UU has explored the capabilities of a range of computational workflows against experimental frequency and intensity signatures.

Given the challenge of vibrational calculations of molecules at battery-relevant interfaces, it was decided to initially generate protocols for reliable workflows to generate molecular vibrational properties of a model interface system. The choice fell on the water/MgO(100) interface as this system is often taken as an archetypical reference system in surface science because of its stability and ostensibly simple structure, making experiments and their interpretation in principle relatively more accessible. Indeed, for this system several experimental studies (XRD and neutron spectroscopy) have been reported, confirming the formation of essentially ordered molecular patterns on the surface but revealing little further detail about the structural motifs. Infrared Reflection Absorption Spectroscopy (IRRAS) spectra can help to elucidate the local structural motifs as the local molecular group vibrations are sensitive to the environment around the oscillator, and several experimental vibrational spectroscopy studies have been published; they agree with each other in that at least three peaks are visible in the intramolecular stretching frequency domain.

Our computational workflow in quest of a protocol that can mimic and predict experiment consists of three stages, each designed to reach a particular goal.

- (Goal 1) Develop a modelling protocol that yields accurate intramolecular vibrational frequencies based on AIMD simulations (thus at non-zero temperature) for a given overall structure, and where any remaining sources of systematic frequency errors are known.
- (Goal 2) Screen plausible interface structure(s) using the protocol in (1) to find a fingerprint in satisfactory agreement with experimental vibrational frequencies.
- (Goal 3) Determine/develop a modelling protocol that manages to yield satisfactory agreement with experimental IRRAS intensities.

Goal (1) has been achieved for our model system after extensive testing of system sizes, vibrational anharmonicity effects and technical details for the AIMD simulations and their post-processing. Goal (2) has also been achieved after having explored 10 possible surface structures with the AIMD protocol from (1). Among them, only one gave close to perfect agreement with the experimental intramolecular stretching frequencies. That very configuration was tried only after revisiting the published experimental STM images in more detail. Our results highlight the strong link between structural features and vibrational spectra, which is both an opportunity and a challenge.
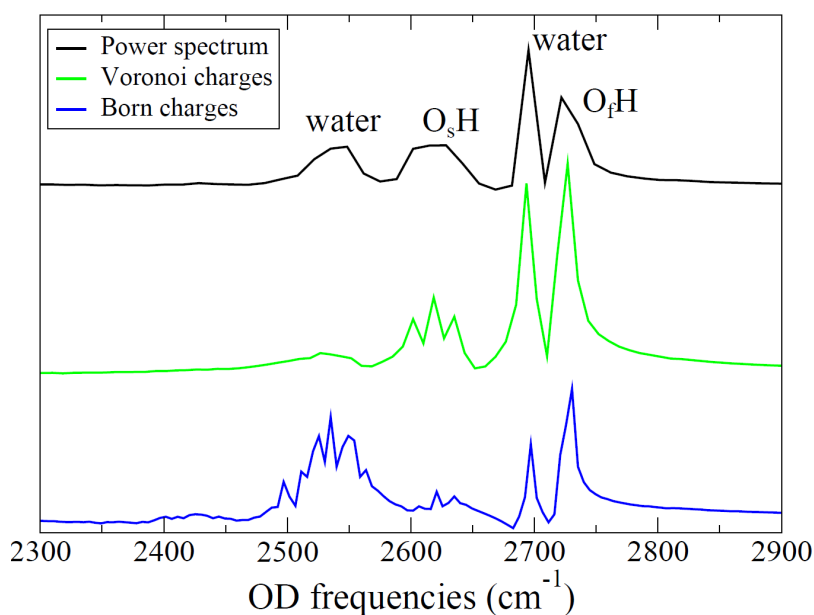

Figure 6. Calculated IRRAS spectra for an archetypical solid-liquid interface model system, namely liquid water (D₂O) on a MgO(100) surface. These three spectra were calculated based on the same AIMD simulation trajectory at the PBE-D3 level but different post-processing procedures were used. The labels O_sH and OH_f in the upper graph are generic and refer to the deuterated species.

The third important aspect, Goal (3), is the intensity match. IR intensities are calculated from time correlation of atomic dipoles. Atomic dipoles can enhance or suppress the presence of peaks in the power spectrum, thus yielding structural information. From ab-initio MD simulations, atomic dipoles and their variations can be calculated in different ways. On surfaces, the symmetry breaking complicates these calculations. Moreover, IRRAS experimental data collect part of the

polarized incident light, meaning that the excited dipoles registered only possess partial spatial components. These aspects must be considered in the intensities calculation. We have made use of several different – and to a large extent established – approaches for IR intensity calculations (with the extension of geometrical provisions to take the interface scenario into account) and find that the methods yield significantly different results (Figure 6), which can have both fundamental and (semi-spurious) technical reasons. Currently, yet another approach is being explored for validation against experiments both for the current system and for a second experimental system where very high-quality IRRAS interface data exist. These investigations will pave the way towards vibrational calculations of more complex, battery-related interfaces.

5. SUMMARY AND OUTLOOK

Thanks to the collaboration among theoretical teams using complementary strategies and approximations, we are able to deliver first results for computed spectra of a broad set of prototype systems: from molecules in electrolytes, to hybrid interfaces and solid-state interphases. These are based on a set of protocols that are currently further developed and standardized, enabling prediction of the most used and relevant experimental characterization techniques, from valence band and core photoelectron spectroscopies, to Raman and IR vibrational spectroscopies.

The next steps in the protocol development will focus on (i) systematic bridging of quantum mechanics and classical molecular dynamics scales; (ii) standardizing and automating the comparison and reciprocal feedback between calculated and experimental spectral features; (iii) porting the automated workflows to the forthcoming pre-exascale European supercomputers in view of massive high-throughput calculations.

Notably, the current work on steps (i) and (ii) incorporate machine learning techniques, which are integrated in (i) the automated clustering and analysis of classical MD structures, in order to extract the key reference motifs for the full quantum mechanical description through and beyond DFT, as well as (ii) the automated comparison of computed and experimental spectra. Continuing and strengthening the collaboration with WP3 for (i) and WP5 for (ii) will be essential for the successful advancement of the project.