CDCS
CENTER FOR DATA AND COMPUTING
IN NATURAL SCIENCES

OPENING SYMPOSIUM 2022

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

DESY.

TUHH

Contribution ID: **117**   Type: **Poster**

Descriptors based on the fragment molecular orbital method for machine learning prediction of X-ray absorption in proteins

Dynamic proton migration along the protein undergoes conformation structural changes being able to promote a folding/unfolding process. Those migration processes have been investigated by challenging near-edge X-ray absorption mass spectrometry (NEXAMS) experiments and computationally expensive calculations at high *ab initio* theory levels. Therefore, to obtain a solid understanding of protonation dynamics, less expensive quantum mechanical (QM) calculations and data analysis are essential for assessing and interpreting the underlying electronic structures changes in proteins when studied by X-ray spectroscopy. Fragment molecular orbital (FMO) calculations have been performed to obtain interaction energies among the amino-acid residues (fragments) providing new *ab-initio* chemical insights into protein inner structure and preserving chemical-physical properties as in conventional *in-silico* calculations. Structural conformations studied by FMO have the key advantage of quantitatively assessing the inter-fragment interactions energies (IFIE) and, thus, separating the energy of each pair of fragments into several energy contributions standing for the so-called pair interaction energy decomposition analysis (PIEDA). To this end, we made a deeper data exploration using PIEDA to prepare QM accurate descriptors for machine learning (ML). Protein atomistic representation were transformed obtaining new smaller structural regions according to the energy decomposition analysis from the IFIE. This new atomistic mapping was created as a function of the properties quantitatively measured by PIEDA. In the X-ray spectra data, NEXAMS exploits the fact that core-electron excitation processes probe the structure locally facilitating to get the complete molecule X-ray absorption as a sum of their fragmented parts (by mass spectrometry). In this connection, PIEDA energy contributions, such as charge transfer, enable to map structural regions of spectroscopic interest and build new smaller sets of molecules from the original one. Thus, leading potentially the X-ray spectra calculation as a sum of smaller molecule sets X-ray spectrum afterwards, which is computationally more efficient for proteins.

Primary author: Mr ORTIZ-MAHECHA, Carlos Andrés (Technische Universität Hamburg (TUHH))

Co-authors: Prof. BARI, Sadia (FS (Forschung mit Synchrotronstrahlung)); Prof. MEISSNER, Robert H. (Hamburg University of Technology)

Presenter: Mr ORTIZ-MAHECHA, Carlos Andrés (Technische Universität Hamburg (TUHH))

Session Classification: Poster session with buffet

Track Classification: CDL2 (Photon Science)