PhD position: Developing Machine Learning Methods for the Prediction of Stable Molecular Crystal Polymorphs

A PhD position is available in the Theoretical Chemical Physics (TCP) group, led by Prof. Alexandre Tkatchenko in the Physics and Materials Science Department (DPhyMS) at the University of Luxembourg. This PhD position belongs to the PHYMOL: A Marie Skłodowska–Curie Actions Doctoral Network (MSCA DN) on Intermolecular Interactions. As such, the PhD candidate will enjoy a broad collaboration with top-notech research groups, especially Avant-garde Materials Simulation Deutschland GmbH, in Freiburg, Germany, the world leading company in polymorph drug prediction.

Background: A method that can accurately and efficiently predict the stable phases of molecular crystals is highly desirable for use in the development of new pharmaceutical drugs [1]. If a drug is manufactured in a metastable phase that later converts to a more stable phase, it can render the drug insoluble and ineffective. Several pharmaceuticals, such as ritovanir, an HIV treatment, have had to be recalled from the market due to not discovering the most stable polymorph in time. One of the main reasons crystal structure prediction is so challenging is that one of the dominant interactions holding molecular crystals together is van der Waals interactions, an inherently quantum mechanical phenomenon. Additionally, while van der Waals interaction between two individual molecules is weak, many-body effects cause the collective interaction between the molecules in a system, even ones that are hundreds of angstrom apart, to become appreciable.

The complex nature of van der Waals interactions and the size of pharmaceutically relevant molecular crystals has impeded the use of computational methods to predict crystal structure, even making density functional theory (DFT), the workhorse of modern computational chemistry, ineffective. Two developments are changing this state of affairs. One, it was recently demonstrated in our group that DFT can be supplemented with an approximate model of van der Waals interactions called the many-body dispersion (MBD) model and this can lead to quantitative accuracy in predicting molecular crystal structure [2]. Two, enormous progress is being made in the realm of machine learned force fields such that they are now able to produce results of comparable accuracy to the quantum mechanical methods they are trained on at orders of magnitude reduced cost. We intend to capitalize on these new developments to improve the accuracy and efficiency of crystal structure prediction methods with a direct view on industrial applications.

Goal: The PhD candidate will use DFT+MBD to develop datasets for the potential energy surface of a set of molecular crystals and use them to develop machine learned force field methodologies that can be integrated into crystal structure prediction workflows, with an eye on properly balancing the complex interplay between different intra- and intermolecular interactions.

Day-to-day: The research will involve working on a linux-based high-performance computing cluster using python to run and analyse

calculations and implement/extend methodology for different aspects of the crystal structure prediction workflow. The TCP group has several experienced postdocs that will be able to assist the doctoral candidate in mastering the various aspects of this research. Additionally, this work will be co-supervised by our industrial partner in order to direct the research towards real world applications.

Pre-requisites: Good mathematical and programming skills, a good understanding of basic quantum mechanics, thermodynamics, and physical and chemical intuition. Experience with machine-learning is especially favorable.

Stipend and benefits: The yearly gross salary for every PhD at the University of Luxembourg is EUR 38.028,96 (full time)

For more information and applications contact: Prof. Dr. Alexandre Tkatchenko (alexandre.tkatchenko@uni.lu)

References:

[1] What is Crystal Structure Prediction? And why is it so difficult? - The Cambridge Crystallographic Data Centre (CCDC)

[2] Johannes Hoja and Hsin-Yu Ko and Marcus A. Neumann and Roberto Car and Robert A. DiStasio and Alexandre Tkatchenko "Reliable and practical computational description of molecular crystal polymorphs" Sci. Adv., 5, eaau3338 (2019) provides an excellent overview of the general problem and the method, DFT+MBD, that the doctoral candidate will use.