

Rolf Tarrach Award 2022

Martin Stöhr: Abstract of PhD Research

“van der Waals Dispersion Interactions in Biomolecular Systems: Quantum-mechanical Insights and Methodological Advances”

Computational modeling has by now become an integral part of chemistry, biology, and materials science. It plays a key-role for pharmacological and technological developments from drug design to next-generation catalysts and energy materials. On a macroscopic scale, systems can be described by classical mechanics, but below that, on the level of atoms and molecules, one enters the peculiar realm of quantum phenomena. The behaviors and properties of matter here follow the laws of quantum mechanics. Quoting Paul Dirac, Nobel laureate and one of the founding researchers of quantum mechanics, *“The underlying physical laws (...) are completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble”*. Over the past century, various developments lead to approximate methods for solving the resulting equations. As a general rule of thumb, the more sophisticated and reliable a given method is, the more steeply its computational workload scales with the size of the system one would like to study. Despite the tremendous growth in the availability of computational resources and high-performance implementations, this still limits the most accurate theories to rather small systems. Studying phenomena at increasing length and time scales requires to resort to increasingly approximate approaches. This bears an inherent complication: Practically-relevant systems, such as proteins or realistic energy materials and catalyst systems, are naturally nano-scale and thus small enough to render classical physics insufficient, but large enough for complex quantum many-body effects to cause decisive deviations from the well-studied behaviors of few-particle quantum systems. Identifying such deviations, potential shortcomings of conventional approximations and developing efficient solutions to them is therefore paramount for a reliable description of the emergent phenomena in practically-relevant systems.

The overarching goal of my PhD work (and my research since then) is to gain such a fundamental understanding and accurate description of nano-scale systems in realistic settings. To this end, I developed new methodological approaches and analysis tools, which combined with high-performance implementations enable the description and understanding of large-scale systems in hitherto unmatched microscopic detail. In a series of research projects, this further allowed to highlight the importance of such a rigorous quantum-mechanical many-body treatment going beyond the common assumptions of conventional, more approximate methods.

Machine-assisted Density-Functional Tight Binding The current work-horse method for the quantum-mechanical description of molecular systems and materials is density-functional theory (DFT). Despite its tremendous success and development over the past decades, the associated computational costs still limit its application to systems considerably smaller than practically-relevant length- and time-scales. A quantum-mechanical treatment at these scales can be achieved using an approximate DFT description as provided by the so-called Density-Functional Tight-Binding (DFTB) method. However, as a result of the introduced approximations, the applicability of DFTB is limited by a lack of accuracy and transferability. In my work, I highlight the failure of one of the critical approximations in one of the

terms entering the DFTB formalism (the so-called “repulsive energy”), which is essential for a correct description. In order to overcome the limitations posed by this approximation, I then proposed to combine DFTB with an artificial neural network model, where the latter is designed to describe this repulsive energy without resorting to the common assumptions entering this term. The resulting method then provides an accurate description of organic molecules with improved transferability. In essence, this allows for a *state-of-the-art* DFT treatment at orders of magnitude less computational costs and paves the way towards an accurate, fully quantum-mechanical description of realistic and practically-relevant systems. [M. Stöhr, L. Medrano Sandomas, A. Tkatchenko: *The Journal of Physical Chemistry Letters* **11**, 6835–6843 (2020)]

Non-local Electronic Polarizability in Liquids and Solvated Systems Another important aspect, in particular for biomolecules, is to account for the system’s environment. In nature, proteins are embedded in a complex matrix of water (and ions). Current methods meant to efficiently capture the effects of such an aqueous environment are based on a local description, assuming the so-called effective medium model. In this model, one assumes that the response and effect of the environment is homogeneous and uniquely defined. As part of my PhD work, I studied this response within liquids and for a protein embedded in water on a microscopic level and with explicit account for the inherent many-body character of the problem (*i.e.*, that the response between two atoms is affected by the presence of all other atoms). Based on this much more complete physical description, I found that the common assumption of an homogeneous medium is generally not valid at the nanoscale. Due to many-body effects, the response and effect of the environment at a given distance from the embedded system can vary strongly and does not have a single, unique value as assumed in the effective medium approach and derived models. Furthermore, the general trends can be strongly dependent on the shape of the interface between the system and the environment as well as on the level at which one represents the system (*i.e.*, whether we are interested in the effect of individual atoms in a protein or the effect of the full protein). The results obtained in this project, ultimately, provide the basis for the construction of more reliable models, that effectively account for the complex character of the effect a liquid environment has on an embedded system. [M. Stöhr, A. Tkatchenko: to be published soon]

Many-body van der Waals Interactions in Biomolecular Systems When it comes to describing the interactions relevant in (bio)molecular systems, one typically conceptually distinguishes so-called covalent interactions (everything that is concerned with the formation and stability of a chemical bond between close atoms) and non-covalent interactions (everything else). Covalent interactions are thereby naturally limited by the finite number of neighboring atoms and their limited spatial range. Non-covalent interactions, on the other side, involve every atom in the system, not only neighboring atoms, and thus play a particularly important role for large-scale systems. Moreover, a great deal of biomolecular function is dominated by global structure and dynamics rather than by the local arrangement of atoms. As such, non-covalent interactions are paramount for the stability and dynamics of biomolecular systems governing structure formation, regulation, and functionality. van der Waals (vdW) dispersion thereby contributes a crucial part to those interactions. As part of the long-range electron correlation energy, vdW interactions are inherently quantum-mechanical and many-body in nature. Common approaches to describe biomolecular systems, as used in *state-of-the-art* drug design or biophysical modeling, fail

to capture this complexity by adapting a purely phenomenological, pairwise formalism. Developing a new high-performance implementation of an accurate many-body vdW method and associated analysis tools during my PhD allowed me to study the characteristics of vdW forces in complex systems such as solvated proteins during folding. The results highlight the importance of many-body vdW forces for understanding the driving forces of protein folding and indicate a remarkable persistence of these interactions through biomolecules and aqueous environments. This forms the basis of a potential long-range interaction mechanism assisting in biomolecular regulation and allostery. With this, my thesis opens up the path to unravel the fundamental mechanisms in the biomolecular machinery in much greater detail. This includes the understanding and exploitation of non-local phenomena like allosteric regulation or the effect of mutations. [M. Stöhr, A. Tkatchenko: [Science Advances](#) **5**, eaax0024 (2019)]

Higher-order Non-covalent Interactions in Complex Systems Non-covalent interactions as introduced above are typically classified into electrostatics, polarization, exchange-repulsion, and vdW dispersion. These, however, only represent the leading-order contributions, beyond which there exist several more contributions. For instance, vdW interactions can change the way that electrons are distributed in the system. This phenomenon, called dispersion-polarization, introduces further, more complex interactions. Based on observations for small molecules, such effects are typically considered to be negligible. As a final part of my PhD research, I co-developed a practical formalism to accurately describe such higher-order terms. While indeed being negligible for small molecules, these contributions can be shown to become relevant with increasing size and complexity of the system. Within nano-structured environments, as relevant to biomolecular systems as well as nano-technological applications, such higher-order effects can even lead to qualitative changes in the interaction among molecules. [M. Stöhr, *et al.*: [Nature Communications](#) **12**, 137 (2021)]

Further details of my PhD research can be found in the following peer-reviewed publications: [Physical Review Letters](#) **121**, 183401 (2018); [Chemical Society Reviews](#) **48**, 4118 (2019); [Science Advances](#) **5**, eaax0024 (2019); [The Journal of Chemical Physics](#) **152**, 124101 (2020); [The Journal of Physical Chemistry Letters](#) **11**, 6835 (2020); and [Nature Communications](#) **12**, 137 (2021). The approaches developed during my PhD are incorporated in the free and open-source software packages [libmbd](#) and [DFTB+](#), making the devised methods and tools available to the wider scientific community.

The usual scheme in developing methods for atomistic modeling is to test approximations on small systems and then take their validity for granted in arbitrarily complex systems. On the contrary, my PhD work highlights how the quantum-mechanics of large-scale systems can in many ways be qualitatively different. Common wisdom and approximations validated for few atoms can fundamentally fail once considering thousands of atoms as in realistic practically-relevant systems. This motivates a drastic shift of paradigm in our understanding of nano-scale systems like biomolecules and their function, but at the same time suggests to reconsider our approach to method development in atomistic modeling. The results, methods, tools, and software provided through my PhD work thereby create part of the basis for such a transformation.