Doctoral Dissertation
Doctoral Program in Material Science and Technology (36^{th}cycle)

Computational and Statistical Approaches to Study Complex Molecular Systems and Their Emergent Properties

By

Martina Crippa

******

Supervisor(s):
Prof. Giovanni M. Pavan, Supervisor

Doctoral Examination Committee:
Asst. Prof. Davide Bochicchio, Referee, Università degli Studi di Genova
Prof. Pratyush Tiwary, Referee, University of Maryland
Prof. Eliodoro Chiavazzo, Politecnico di Torino
Prof. Marta Corno, Università degli Studi di Torino
Prof. Alessandro Laio, Scuola Internazionale Superiore di Studi Avanzati

Politecnico di Torino
2024

Declaration

I hereby declare that, the contents and organization of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

Martina Crippa
2024

* This dissertation is presented in partial fulfillment of the requirements for **Ph.D. degree** in the Graduate School of Politecnico di Torino (ScuDo).

Computational and Statistical Approaches to Study Complex Molecular Systems and Their Emergent Properties

Martina Crippa

Self-assembling molecular systems are typically difficult to characterize due to their complex internal structural dynamics. However, such intrinsic complexity plays a key role in determining emergent collective behaviors that have a strong impact on the physical properties of the entire molecular system.

This thesis proposes diverse methodologies that allow shedding light on the dynamic behavior of complex molecular systems. Employing a combination of multiscale molecular models, molecular simulations, and innovative statistical analysis, the work presented herein deepens the understanding of molecular and supramolecular systems, and the properties that emerge within them. In the first part of this thesis, I describe the use of minimalistic Coarse-Grained (mCG) models to study different classes of self-assembled supramolecular polymers. Such models allow investigating how these polymers dynamically communicate with each other in different systems and diverse external conditions, controlling the collective properties of the supramolecular material. Moving forward, the study expands to include another class of self-assembling supramolecular systems, *i.e* bi-component dynamical micelles. This part of the research incorporates the combination of high-dimensional structural descriptors - *i.e.* the Smooth Overlap of Atomic Position (SOAP) - and of advanced Machine Learning approaches to identify relationships between the chemical species that constitute such soft systems and their structural and dynamical features. The thesis then introduces a new descriptor, named Local Environment and Neighbor Shuffling (LENS), which tracks the changes in the identity-based neighborhood around each particle. This tool, able to efficiently detect local fluctuations in time, demonstrates its versatility across different systems, illustrating how changes in the local neighborhoods relate to the collective dynamics of the systems. Afterward, employing the comprehensive statistical approach proposed in the first part of this thesis, complemented with a LENS-based analysis, we investigate the stimuli-responsiveness of two classes of complex supramolecular polymers having different cooperativity degrees. We thus characterize how cooperativity determines

the different responses of the systems' equilibrium picture and communication network, specifically when perturbed by an external chemical stimulus. Furthermore, by integrating the dynamic characterization provided by LENS with a structural SOAP-based analysis, we demonstrate how it is possible to obtain unique microscopic-level structure-dynamic relationships relating the structural environment that populate a variety of molecular systems to the dynamical events that originate from them. Finally, I propose two exploratory works which contain ongoing results. In the first work, combining mCG models and SOAP-based characterization, we link the topology of the self-assembling building blocks to the equilibrium supramolecular structure that they generate by a high throughput exploration based on a Genetic Algorithm. This provides insights into the rational design of self-assembling molecules, revealing how the topology of an individual component may influence the structural properties of the whole assemblies. The second work provides an in-depth comparison of the structural characterization given by the SOAP and the Atomic Cluster Expansion (ACE) descriptors, understanding in what regimes they give equivalent results.

Overall, this thesis proposes a novel perspective and advanced analytical tools for studying molecular systems and their internal complexity. Integrating computational modeling, simulations, and advanced analysis methods, we provide a new approach to studying the underlying principles of such complex systems. This not only enriches our fundamental understanding of molecular and supramolecular systems' behavior but also opens new venues for the design of complex materials with controlled emergent properties.