Data-driven approaches to investigate intermolecular interactions and self-assembly in complex fluids: from data fusion to 2D correlation analysis

Duccio Tatini [1,2], Mert Acar [2,3], Michele Baglioni [1,2], Gabriella Tamasi [1,2], Claudio Rossi [1,2], Pierandrea Lo Nostro [2,3]

[1] Department of Biotechnology Chemistry and Pharmacy, University of Siena, 53100 Siena, Italy
[2] Centre for Colloid and Surface Science (CSGI), University of Florence, 50019 Sesto Fiorentino, Italy
[3] Department of Chemistry "Ugo Schiff", University of Florence, 50019 Sesto Fiorentino, Firenze, Italy

duccio.tatini@unisi.it

Understanding intermolecular interactions and self-assembly processes in complex fluids remains a central challenge in colloid science. In systems such as polysaccharides and wormlike micelles the presence of ions, cosolutes or polymers with different charge and nature can induce significant changes in the interfacial structuring, hydration properties and rheological response [1]. However, their interpretation is often hindered by the complexity of the underlying interactions. In this work, we present a novel data-driven framework that combines chemometric tools, machine learning, data fusion strategies, and two-dimensional correlation analysis (2D-COS) to investigate interaction mechanisms in soft matter systems by extracting and integrating relevant information from multiple techniques. First, polysaccharide dispersions are used as model systems to investigate Hofmeister effects on their hydration and thermal behavior. ATR-FTIR and thermogravimetric (TGA/DTG) data are combined through a data fusion approach, enabling sample differentiation according to the kosmotropic/chaotropic nature of the cosolutes (Figure 1a). This combined analysis reveals ion-specific trends not evident from single techniques, providing a comprehensive description of cosolute-dependent structuring. Second, 2D correlation spectroscopy is applied to ATR-FTIR data to elucidate the interaction pathways in polyelectrolyte–wormlike micellar systems. The analysis identifies distinct sequences of molecular rearrangements depending on polymer charge, separating direct electrostatic binding from mechanisms mediated by hydration and hydrogen bonding (Figure 1b). These results establish a direct link between molecular-level interactions and the resulting modifications of the micellar network, which determine the macroscopic rheological properties [2]. Finally, an extension of 2D-COS to non-spectroscopic datasets is proposed, combining rheological and calorimetric (DSC) data to investigate cosolute effects on wormlike micellar systems. The correlation of thermodynamic features with viscoelastic responses provides a comprehensive description of how cosolutes modulate molecular interactions, self-assembly, and bulk dynamics. Overall, this work demonstrates the potential of data fusion and data-driven strategies for the investigation of complex interaction phenomena in soft matter, offering a framework that can be extended to different macromolecular systems, cosolutes, and analytical techniques to advance the understanding of structure–property relationships in colloidal systems.

Keywords:
Machine learning, chemometrics, multiblock analysis, Hofmeister, IR, calorimetry, rheology

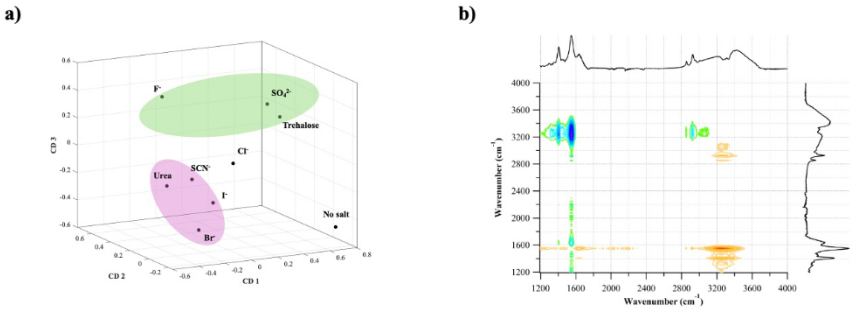

Figure 1. ComDim score plot obtained from the ATR-FTIR and TGA data (a) and asynchronous 2D-COS (b) spectra obtained from the ATR-FTIR data.

References:

[1] D. Tatini et al. J. Colloid Interface Sci. 617 (2022) 399–408
[2] D. Tatini et al. Colloids Surf. A Physicochem. Eng. Asp. 736 (2026) 139621