2.5.7 Research Area "Molecular Interactions in Organic and Biological Systems. Applications and Methodological Implementations" (E. Sánchez-García)

Involved: K. Bravo-Rodriguez, S. Mittal, P. Sokkar, J.M. Ramirez-Anguita

Objectives: Molecular interactions (MIs) play a key role in chemical and biological processes. However, despite the many scientific advances of the last decades, understanding and predicting the effect of MIs in biological systems is still challenging. In our group we develop computational models to help regulate strategically important molecular interactions in biological systems via the rational manipulation of three factors: *1. the introduction of ligands able to alter protein function, 2. the interactions of proteins with their biological environment,* and *3. the mutagenesis of key residues*. In addition, we also investigate molecular interactions in small model systems like reactive intermediates and closed-shell molecules as a basis to understand more complex systems (Figure 11). One important goal of our research is to predict improved therapeutic agents and approaches, for example in the context of neurodegenerative diseases and the development of new antibiotics.

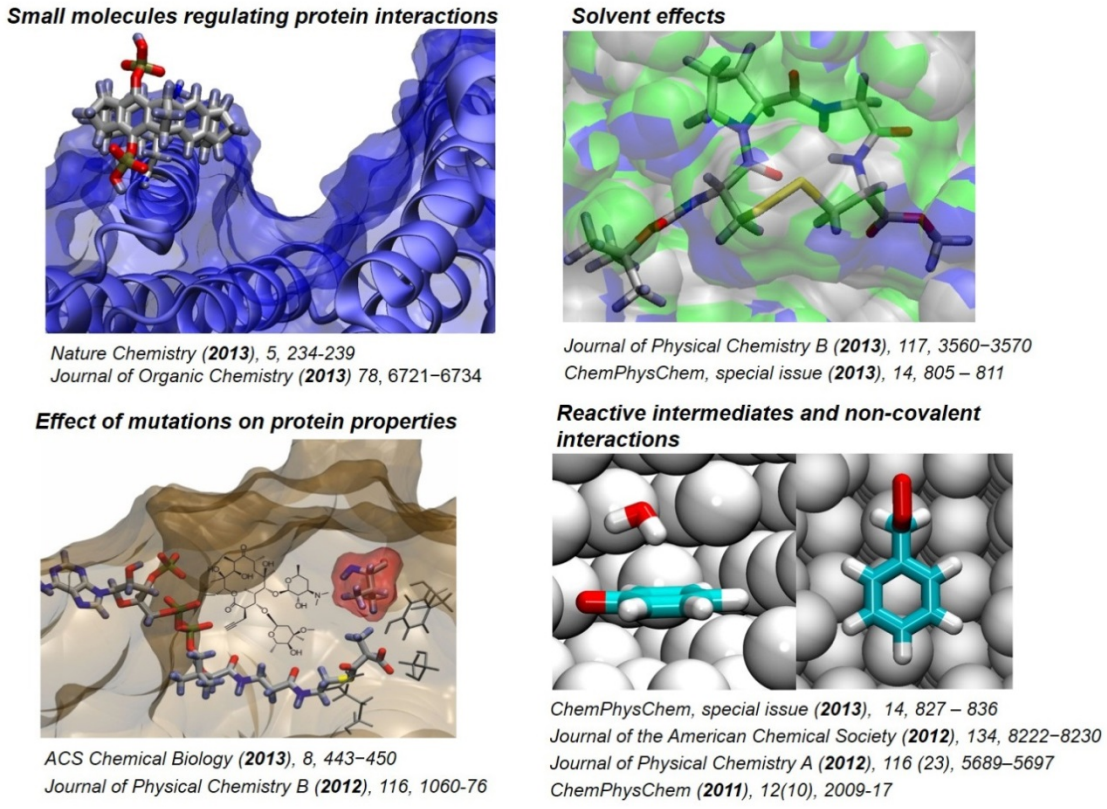

Figure 11. Main research lines and selected representative publications of our group in the last two years.

On the methodological side, we work on generalizing the combination of quantum mechanics (QM) and atomistic molecular mechanics (MM) into a triple-layer QM/MM/CG modeling approach by implementing coarse grained (CG) force fields in the ChemShell code. The aim is to extend the applicability of QM/MM computations to enable an efficient and realistic theoretical treatment of complex chemical, photochemical, and biochemical systems that are challenging for the existing computational methods.

Results:

Molecular interactions in organic and biological systems. Applications

a) Ligands for the regulation of protein properties

Ligands can be used to regulate protein aggregation, enzymatic activity, lipid attachment, and protein–protein interactions, among other important biological properties. In my group we investigate the effect of highly selective ligands such as molecular tweezers (MT) that bind specifically to lysine and arginine, as well as the effect of less specific molecules like aromatic heterocyclic derivatives.

Experimental studies have shown that MTs are able to inhibit the aggregation of Aβ (amyloidogenic peptide related to Alzheimer disease) and IAPP (islet amyloid polypeptide, related to type II diabetes mellitus) without toxic effects, which makes them promising candidates for drug development. In this context, we recently studied the binding mode of four water-soluble tweezers bearing phosphate, phosphonate, sulfate, or O-methylenecarboxylate groups at the central benzene bridge to amino acids and peptides containing lysine or arginine residues. The comparison between experimental and theoretical data provided clear evidence for the unique threading mechanism and the modulating effect of each anion on the interaction of the MTs with positively charged amino acids and peptides. These findings were explained on the basis of the host-guest complex structures obtained from molecular dynamics (MD) simulations and QM/MM methods [120].

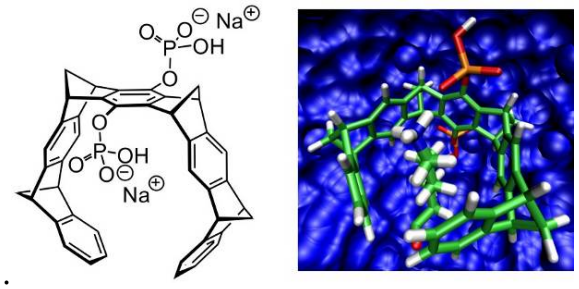

Figure 12. Bishydrogenphosphate tweezer CLR01 (left), CLR01 interacting with lysine, structure optimized in explicit water environment at the QM/MM level (right).

To answer the question of how the tweezers behave in the presence of complex systems with multiple lysine residues, we also studied the effect of molecular tweezers on proteins. We established that molecular tweezers are a powerful tool for regulating 14-3-3 protein-protein interactions for the 14-3-3 σ isoform via an entirely new interference mechanism. Furthermore, the combination of structural analysis and computer modeling allowed us to establish *general rules* for predicting the relative strength and type of interaction of the tweezers with lysine residues in 14-3-3 proteins [118]. We are currently investigating the effect of molecular tweezers on other protein hosts.

Importantly, the study of the interaction motifs and mechanisms of small molecules with several peptide and protein hosts allows us not only to answer important questions related to these systems, but also to guide the design of improved ligands able to reach specific protein regions of biological relevance.

b) Reactive intermediates and synthetic cyclic peptides. Solvent effects

Given the importance of radicals and radical reactions in such diverse aqueous environments as biological systems or tropospheric clouds, it is highly desirable to understand the influence of water and other solvents on their chemical and physical properties. Since we are interested in using molecular interactions with the environment to tune the properties of chemical and biological systems, we study the effect of solvents on reactive intermediates (e.g., radicals) using QM, MD, and QM/MM techniques.

We investigated the complexes of phenoxyl, anilinyl, and benzyl radicals with water using QM and QM/MM MD approaches [53,116]. The influence of molecular interactions on the properties of other model systems like furan complexes [9,115] was also addressed.

Cyclic polypeptides containing a photocleavable disulfide (S-S) bridge are useful models for the study of dynamical conformational changes that convert a polypeptide chain into a three-dimensional protein structure. The conformers of a series of cyclic peptides Cys-Pro-X-Cys (X: Gly, D-Leu, L-Leu, D-Val, L-Val) were studied in the gas phase and in the solvents water and acetonitrile, using both classical MD simulations and QM/MM approaches. We elucidated the effect of the stereochemistry of just a single residue on the conformation of the peptide and on solvent–peptide interactions, and found that the solvent plays a key role for the structural and spectroscopic properties of such systems [119,120].

c) Designed mutagenesis for the regulation of enzymatic activity

Quite often, mutations have deleterious effects and dramatically alter the properties of biological systems. On the other hand, mutations are essential for evolution; and controlled mutagenesis is an important tool to selectively manipulate protein properties. Hence, mutations can be used to regulate photochemical processes, protein folding, protein-protein interactions, and enzymatic activity, among others. Computer simulations play a key role for the molecular understanding and prediction of the effect of mutations.

In our work on *computational mutagenesis*, we aim at providing a molecular understanding of the effect of mutations in biological systems and at delivering computational predictions of targeted mutations able to modify protein properties. Hence, we have investigated disease-associated mutations in rhodopsin [25,121] and the effect of mutations on structural properties of other complex systems. Of especial interest to us are Polyketide Synthases (PKS), which are multienzyme complexes able to direct the formation of a diverse array of functional groups and stereocenters. They catalyze a cascade of Claisen condensations between the enzyme acyl thioester and malonic acid thioesters, which serve as extender units of a nascent polyketide chain. The complex process of extender unit building block selection and incorporation in PKS is mainly controlled by acyltransferase (AT) domains, which catalyze the malonylation of acyl carrier proteins (ACP). A promising option to modify building block specificity in PKS is the direct engineering of innate AT domains to alter their substrate scope by replacing selected amino acids instead of whole catalytic domains.

My group developed computational models for the last acyltransferase domain (AT6) of 6-Desoxyerythronolide B Synthase (DEBS), the key multienzyme complex in the biosynthesis of the antibiotic Erythromycin, to provide a molecular understanding of the effect of directed mutations and the role of specific residues on the enzymatic activity. We were able not only to *rationalize* the effect of point mutations on enzymatic activity, but also to successfully *predict* which single mutation would result in the incorporation of non-native building blocks [117]. The combination of experimental sequence-function correlations (cooperation with F. Schulz) and computational modeling revealed the origins of substrate recognition in PKS domains and enabled their targeted mutagenesis. We are currently extending this research to other systems like monensin PKS (cooperation with F. Schulz).

Implementation of coarse-grained methods in ChemShell

Hybrid QM/MM methods have widespread applications to model biomolecular, organic, inorganic, and organometallic systems with explicit solvation. ChemShell is a QM/MM modular package that allows the combination of several QM and MM methods. Currently, there is considerable interest in the development of coarse-grained (CG) force fields to improve the performance and sampling in MD simulations and geometry optimizations. The CG methodology has been successfully applied to very large molecular systems (such as virus capsids), which are cumbersome to handle at the atomistic level. However, CG approaches do not allow the study of fine structural details (such as fluctuations in secondary structures) due to their approximate representation. To circumvent this problem, hybrid coarse-grained / fine-grained (CG/FG) simulation protocols have been developed.

We work on the implementation of CG force fields in the ChemShell package (cooperation with W. Thiel). By developing novel triple-layer QM/MM/CG modeling approaches, we want to extend the applicability of QM/MM computations. This scheme allows us to treat the chemically important region at the QM level, the area around the QM region at the atomistic MM level, and the rest of the outer region (far from the QM site) at the CG level. In this approach, the presence of the CG region partially eliminates artifacts that may arise from the truncation of the MM region in very large systems. For instance, to study a membrane protein, the lipid bilayer and bulk water may be represented at the CG level while the more important regions are treated at higher resolution. Introducing a CG layer will also reduce the costs of MD simulations. This is especially helpful for QM/MM MD simulations that are currently computationally very expensive, even for smaller systems. Thus, the multi-scale QM/MM/CG approach can also be conveniently used to investigate chemical reactions in explicit solvent, with the bulk solvent being treated at the CG level.

At present, we have already developed a module for the ChemShell code that performs hybrid QM/MM/CG calculations with Martini CG water for the CG layer; we continue this work to include other systems and force fields in our implementation.