Juan Cortés

A data-driven approach for modeling highly-flexible proteins and regions

Protein structure prediction and protein design problems have been mostly formulated assuming that proteins fold into a well-defined three-dimensional form. Nevertheless, this is not always the case. Many proteins involve flexible regions, such as loops or linkers, which can play important functional roles in molecular recognition, enzyme catalysis, allosteric regulation or signaling. Moreover, the so-called Intrinsically Disordered Proteins (IDPs) lack of a permanent secondary or tertiary structure, and exploit their plasticity to perform highly specialized tasks that are complementary to those of their globular counterparts. The structural investigation of such highly-flexible proteins/regions is essential for the understanding of numerous functional or pathological processes. This is a challenging topic, requiring a tight coupling of experimental and computational methods. In this talk, I will present our approach to generate conformational ensemble models, which exploits information extracted from coil regions in high-resolution experimentally-determined protein structures, and shows the importance of local sequence-dependent structural propensities. Future advances in this field require the application of more sophisticated statistical models and methods.