Finding functional regions within intrinsically disordered proteins

Zsuzsanna Dosztányi
MTA-ELTE 'Momentum' Bioinformatics Group
Department of Biochemistry, Eötvös Loránd University, Budapest

Recent observations have shown that a large portion of the human genome encodes intrinsically unstructured/disordered proteins and protein regions (IDPs/IDRs) that do not adopt a stable structure in isolation. Many of these proteins function via binding to a structured partner and can undergo a disorder-to-order transition. In their bound state, some IDPs form a partially extended structure, burying large surface areas, while others bind through localized motifs with relatively small binding interfaces. These functional sites are key components of many cell signalling networks and have been associated with various diseases, especially with cancer. However, the protein-protein interactome mediated by binding regions within IDPs and its role in diseases, is still largely unexplored.

Bioinformatic methods can help to close the gap in the functional annotation of the structured versus the disordered proteome and to discover unknown disordered binding regions with functional importance. In recent years several approaches have been developed for the computational identification of these sites, including various machine learning methods. Using a different type of approach, ANCHOR predicts disordered binding regions by capturing their basic biophysical properties: their disordered status in isolation, and the ability to form favorable interactions with globular proteins that drive the binding. A conceptually different approach is based on short linear motifs (SLIMs) that correspond to short consensus sequence patterns distilled from proteins with a common interaction partner, such as PDZ, SH2 or SH3 domains. Known motifs can be used to investigate further candidate functional sites, however, the predicted instances of consensus motifs need to be filtered by additional criteria.

In this talk, I will give an overview of existing approaches to identify functional sites within disordered proteins and how these methods can be used to explore the IDP mediated interactome. Prediction of functional regions located within disordered proteins can also be used to gain insights into their role in different types of cancer. I will also present several examples where protein regions contain an increased number of mutations within their disordered binding regions.