Virtual BSC RS/Life Session: Protein structure informs functional mechanisms and the risk of disease

Objectives

Title: Protein structure informs functional mechanisms and the risk of disease

Abstract: Classifying proteins into evolutionary families is important for identifying conserved sequence and structure features that are key to the functional mechanisms of these proteins. Our in-house CATH classification currently classifies ~450,000 protein structures and nearly 150 million protein domain sequences into ~5500 evolutionary families. The recent success in protein structure prediction by DeepMind's AlphaFold2 (AF2) method and the expected release of hundreds of thousands of AF2 models, will change the scientific landscape by massively extending the structural data available for these protein evolutionary families. We have developed a strategy to bring this extensive new 3D data into CATH families and are examining how this data will expand our understanding of structure – function relationships and our ability to detect functional sites. Functional site predictions can be enhanced by combining structural features and evolutionary conservation patterns and some examples will be given of the application of CATH functional site data to understand protein splice events, the risk of Covid infection and the development of more aggressive lung cancer following whole genome duplication.

Speakers

Host: Alfonso Valencia, BSC Life Sciences department director

Speaker: Prof. Christine Orengo, Group Leader University College London (UCL) & President of the International Society of Computational Biology (ISCB)

Christine Orengo is a computational biologist, whose core research has been the development of robust algorithms to capture relationships between protein structures, sequences and functions. She has built one of the most comprehensive protein classifications, CATH, used worldwide by tens of thousands of biologists, and central to many pioneering structural and evolutionary studies.

CATH structural and functional data for hundreds of millions of proteins has enabled studies that revealed essential universal proteins and their biological roles, and extended characterisation of biological systems implicated in disease e.g. in cell division, cancer and ageing. CATH functional sites have revealed protein residues implicated in enzyme efficiency and bacterial antibiotic resistance. This data also identified genetic variations likely to be driving human diseases and the drugs that can be repurposed to offset the pathogenic effects.

Christine is President of the International Society of Computational Biology (ISCB). She is a Fellow of the Royal Society of Biology and Elected member of EMBO since 2014, a Fellow of ISCB since 2016 and a Fellow of the Royal Society since 2019. She is a co-founder of the ELIXIR 3DBioInfo Community in Structural Bioinformatics.

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

---