EMBO Practical Course

Integrative Modelling of Biomolecular Interactions

2-6 July 2018

Barcelona Supercomputing Center

Table of Contents

Protein interact to form large interconnected networks that are central to any cellular process. The atomic structures of these protein interactions are an invaluable source of information, paving the route towards a mechanistic understanding of biological function and providing a critical first step in the development of new therapeutic drugs. Because of the challenges associated with experimental structure determination, structural biology is relying more and more on computational methods to interpret and complement experimental data.

This EMBO Practical Course presents computational methods to predict how proteins interact with other biological macromolecules and therapeutic small molecules. Lectures will provide the theoretical background on state-of-the-art algorithms for generating and evaluating the quality of three-dimensional models of protein interactions, describe the use of low- and high-resolution information to drive or bias the modelling, and explore sequence conservation and co-evolution as predictors of interfaces. Further, protein-peptide and protein-ligand interactions, the use of coarse-grained models, and the young field of genome structure determination will also be covered to broaden the scope of the course.

We welcome you all in Barcelona for this exciting course!

The organizers

Alexandre Bonvin

Juan Fernandez-Recio

João Rodrigues

List of Sponsors

Programme

Day 1 (July 2nd)

09:00 - 09:30	Registration & Welcome address (Àgora Room) **Alfonso Valencia (BSC Life Director) & Alexandre Bonvin**
09:30 - 11:00	General Introduction to Protein-Protein Interactions (Àgora Room) **Shoshana Wodak & Marc F Lensink**
11:00 - 11:30	Coffee break (Àgora Room)
11:30 - 13:00	Principles of Docking I – Sampling (Àgora Room) **Paul Bates**
13:00 - 14:30	Lunch break (Vèrtex Cafeteria)
14:30 - 16:30	Practical: UDOCK - Interactive docking system (A5S102 room) **Matthieu Montes**
16:30 - 17:00	Coffee Break (FIB cafeteria)
17:00 - 19:00	Participants flash presentations I (Àgora Room)
19:00 -	Welcome drinks and dinner (Àgora Room)

Day 2 (July 3rd)

09:20 - 09:30	Day briefing (Àgora Room)
09:30 - 11:00	Principles of Docking II – Scoring (Àgora Room) **Juan Fernandez-Recio**
11:00 - 11:30	Coffee break (Àgora Room)
11:30 - 13:00	Information-driven Docking (Àgora Room) **Alexandre Bonvin**
13:00 - 14:30	Lunch break (Vèrtex Cafeteria)
14:30 - 16:30	Practical: Integrative docking with HADDOCK (A5S102 room) **Alexandre Bonvin, João Rodrigues**
16:30 - 17:00	Coffee Break (FIB cafeteria)
17:00 - 18:30	Practical: pyDock (A5S102 room) **Juan Fernandez-Recio & Brian Jimenez & Lucia Diaz**
18:30 - 19:00	Visit to Mare Nostrum supercomputer cluster

Day 3 (July 4th)

09:20 - 09:30	Day briefing (Àgora Room)
09:30 - 11:00	Bioinformatics & Interface Prediction (Àgora Room) **Raphaël Guerois**
11:00 - 11:30	Coffee break (Àgora Room)
11:30 - 13:00	Protein-ligand Interactions and Design (Àgora Room) **Birte Höcker**
13:00 - 14:30	Lunch break (Vèrtex Cafeteria)
14:30 - 17:00	Practical: Protein-ligand Interaction Design (A5S102 room) **Birte Höcker & Horst Lechner**
17:00 - 17:30	Coffee Break (FIB cafeteria)
17:30 - 19:00	Practical: Scoring docking models using co-evolution (A5S102 room) **Jessica Andreani**
19:00 - 21:00	Poster Session & Buffet Dinner (Sala Polivalent)

Day 4 (July 5th)

09:20 - 09:30	Day briefing (Àgora Room)
09:30 - 11:00	Integration of Cryo-EM data in Docking (Àgora Room) **Matteo Dal Peraro**
11:00 - 11:30	Coffee break (Àgora Room)
11:30 - 13:00	Integrative Structure Determination of Macromolecular Assemblies (Àgora Room) **Maya Topf**
13:00 - 14:30	Lunch break (Vèrtex Cafeteria)
14:30 – 16:30	Practical: Integrative Structure Determination of Macromolecular Assemblies (A5S102 room) **Maya Topf & Sony Malhotra**
16:30 - 17:00	Coffee Break (FIB cafeteria)
17:00 - 18:30	Bring your own problem & CASP-CAPRI Problem Set (A5S102 room)
18:30 - 19:30	Interactomes & Disease (Àgora Room) **Patrick Aloy**
21:00 -	Course dinner (Ca la Nuri Restaurant)

Day 5 (July 6th)

09:20 - 09:30	Day briefing (Àgora Room)
09:30 - 11:00	Structure Determination of Genomes (Àgora Room) **Marc Marti-Renom**
11:00 - 11:30	Coffee break (Àgora Room)
11:30 - 13:00	Elastic Network Modelling in the Study of Macromolecular Assemblies (Àgora Room) **Turkan Haliloğlu**
13:00 - 14:30	Lunch break (Vèrtex Cafeteria)
14:30 - 16:30	Practical: Binding Affinity Prediction **Iain Moal**
16:30 - 17:00	Coffee break (FIB cafeteria)
17:00 - 18:00	Q&A/Round Table: Challenges in the docking field (Àgora Room)
18:00 -	Departure

General Introduction to Protein-Protein Interactions

Shoshana Wodak* & Marc F Lensink*

VIB Structural Biology Research Center, VUB, Brussels, Belgium

I'll review some of the very early protein-ligand and protein-protein docking methods and their application to the elucidation of the quaternary structure change in hemoglobin, and how they gave rise to a thriving area of research. This will be followed by the description of a recent computational analysis of the mechanism and free energy profile of three-dimensional domain swapping reactions in protein homo-oligomers. In this analysis we propose that a large category of domain swapped systems form by a process that does not involve chain unfolding but proceeds by gradual exchange between intra- and intermolecular interactions between the subunits, which described a minimum energy path for the reaction. The computational analysis models both the conformational and dynamic aspects of the process at the atomic scale.

Principles of Docking (I): Sampling

Paul Bates*

The Francis Crick Institute, London, United Kingdom

Two major factors to consider for the successful computational docking of two or more unbound proteins into a macromolecular complex are the employment of efficient conformational space search engines and the use of appropriate energy functions. Proteins are not likely to find each other's binding sites in a pinpoint fashion. They are more likely to exhibit a swarm-like behaviour over each other's binding sites before descending their joint binding funnel. The docking process can therefore be split into two distinct processes, first, how the proteins find their approximate binding regions, from a sea of energetic decoys, and second, how the native binding funnel is descended to create highly specific, and general stabilising, atom-atom interactions.

We will present how our laboratory tackles these two aspects. Typically, a search is performed for likely binding regions using a hierarchical conformational space search engine based upon the principles of Particle Swarm Optimization (PSO). The benefit here is that in principle, although to a limited depth, the complete conformational search space can be sampled. For example, the complete surface of a receptor/ligand pair can be sampled, and within a computational time frame that is acceptable when performing multiple docking experiments. However, for a deeper conformational search, once the most likely true positive binding funnel has been identified, we employ a classical Molecular Dynamics (MD) protocol to refine the interacting proteins in explicit solvent. Here the MD simulations are restrained by local contacts from the top of the putative binding funnel, contacts identified from the coarse-grained docking methodology described above. Finally, we will present a discussion on how similar energetic MD profiles, for a true positive compared to false positive funnels, may be distinguished by an analysis integrating machine learning principles where scoring features are selected from knowledge-based as well as empirical potentials.

Principles of Docking (II): Scoring

Juan Fernandez-Recio*

IBMB-CSIC / BSC, Barcelona, Spain

The structural characterization of protein-protein interactions at molecular level will play an important role in analyzing the humongous amount of data being generated by genomic sequencing projects. In this context, protein-protein docking remains an important challenge in structural bioinformatics and many methods are being developed to complement experimental efforts in the structural modeling protein-protein interactions. From a technical point of view, the docking problem presents two main challenges: the efficient sampling of the conformational and orientation space in search of near-native structures (sampling), and the identification of such near-native structures among the many models generated (scoring). In many cases, the applicability of a given scoring function is strongly dependent on the sampling approaches used. Thus, the development of new functions that can be independently applied to different sets of docking models generated by a variety of docking methods is an active area of research. Here we will discuss future applications, report recent developments and identify areas requiring further investigation. Many functions have been developed to quantify the structural and energetic properties of interacting proteins. Current challenges include loop modeling, side-chain refinement, docking, multimer assembly, affinity prediction, affinity change upon mutation, hotspots location and interface design.

Information-driven Docking

Alexandre Bonvin*

Utrecht University, Faculty of Science - Chemistry, Utrecht, Netherlands

With the presently available amount of genetic information, a lot of attention focuses on systems biology and in particular on biomolecular interactions. Considering the huge number of such interactions, and their often weak and transient nature, conventional experimental methods such as X-ray crystallography and NMR spectroscopy will not be sufficient to gain structural insight into those. A wealth of biochemical and/or biophysical data can however easily be obtained for biomolecular complexes. Combining these data with docking, the process of modeling the 3D structure of a complex from its known constituents, should provide valuable structural information and complement the classical structural methods.

In this lecture, I will first describe the various potential sources of information for the study of biomolecular complexes and then discuss general aspects of docking. I will then describe the data-driven docking approach that we have developed for this purpose called HADDOCK (http://bonvinlab.org/software/HADDOCK2.2). HADDOCK distinguishes itself from ab-initio docking methods in the fact that it encodes information from identified or predicted protein interfaces in ambiguous interaction restraints (AIRs) to drive the docking process. It can integrate information derived from biochemical, biophysical or bioinformatics methods to enhance sampling, scoring, or both. The information that can be integrated is quite diverse: interface restraints from NMR, mutagenesis experiments, or bioinformatics predictions; shape data from small-angle X-ray scattering and, recently, cryo-electron microscopy experiments. In my talk I will illustrate HADDOCK's capabilities with various examples.

Bioinformatics & Interface Prediction

Raphael Guerois*, Jessica Andreani, Jinchao YU

Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ Paris-Sud, Université Paris-Saclay, 91198, Gif-sur-Yvette cedex, France, Gif-sur-Yvette, France

Protein-protein interactions are of fundamental importance in virtually all cellular processes. Many of these interactions have been conserved during the course of evolution. Structural prediction of protein interactions can be greatly improved by the analysis of evolutionary information [1]. In this lecture and in the accompanying practical, we will explore the extent to which evolutionary information can help recognize binding interfaces at the surface of proteins and predict binding modes/assemblies. Our understanding of protein complex co-evolution has been deepened by a systematic study of over 1,000 couples of homologous interfaces, in which we uncovered astonishing plasticity in the way interface structure evolves [2]. We also identified some rather invariant features which provide tracks for extracting meaningful information from multiple sequence alignments of binding partners. On that basis, we developed a method to score docking models using co-evolution which improves the prediction of protein interfaces [3,4] and was used in a number of protein docking cases [5], a selection of which will be discussed. Understanding how interfaces co-evolve also opens new perspectives for the design of inhibitors of protein-protein interactions, a topic we will also introduce.

[1] Andreani & Guerois (2014) Arch Biochem Biophys. 554:65-75.

[2] Andreani et al (2012) PLoS Comput Biol. 8(8):e1002677

[3] Andreani et al (2013) Bioinformatics. 29(14):1742-9.

[4] Yu el al (2016) Nucleic Acids Res. Apr 29. In press

[5] Lisboa et al (2014) Nucleic Acids Res. 42(11):7395-408.

Protein-ligand Interaction Design

Birte Höcker*

Institut für Biochemie, Bayreuth, Germany & MPI für Entwicklungsbiologie, Tübingen, Germany

Engineering specific interactions between proteins and small molecules is extremely useful for biological studies as these interactions are essential for molecular recognition. Many biotechnological applications are made possible by such an engineering approach, ranging from biosensors to the design of custom enzyme catalysts.

PocketOptimizer is a recently developed computational method to design protein binding pockets. Starting from a protein structure an existing small molecule binding pocket is optimized for the recognition of a new ligand. The modular program predicts mutations that will improve the affinity of a target small molecule to the protein of interest using a receptor-ligand scoring function to estimate the binding free energy. PocketOptimizer has been tested in a comprehensive benchmark and predicted mutations have also been used in experimental tests. I will provide an overview over this approach and general recommendations for usage.

Integration of Cryo-EM data in Docking

Matteo Dal Peraro*

Ecole polytechnique fédérale de Lausanne - EPFL

Revealing the atomistic architecture of macromolecular complexes is a fundamental step toward a deeper understanding of cellular functioning. This formidable task is facilitated by the advances in structural biology and proteomics, along with the emergence of integrative modeling approaches able to combine broad array of experimental data. My laboratory is currently active in the development of a new generation of such computational methods that are based on state-of-the-art evolutionary algorithms combined with molecular simulation. Within this framework we are thus able to take into account the dynamic rearrangements of the individual subunits upon assembly and can efficiently treat protein spatial restraints coming from heterogeneous experimental sources at different resolution. Both these aspects are in fact key for providing a more detailed understanding of the molecular determinants at the basis of macromolecular assembly and function.

Integrative Structure Determination of Macromolecular Assemblies

Maya Topf*

ISMB, Birkbeck/UCL University of London, London, United Kingdom

The integration of data derived from a variety of biophysical techniques at multiple levels of resolution is becoming increasingly common in structural determination of large macromolecular assemblies, with 3D electron microscopy (3D-EM) being one of the key players. Despite recent advances in the field the majority of the structures obtained by this technique is still not at atomic or near-atomic resolution. Fitting of atomic structures into 3D-EM density maps is often essential to gain further insights into the macromolecular assemblies they represent. However, density fitting is often very challenging, and its success depends on the resolution, size and shape of the complex, the number of components, and conformational differences between the fitted components and their corresponding density. Additionally, assessing the quality of the resulting model is not trivial. Here, I will introduce the various approaches to density fitting. Using examples with maps at different resolution and size, I will show how different methods could be used in different scenarios, including rigid fitting, flexible fitting and assembly fitting. I will discuss recent advances in fit/model assessment.

Interactomes & Disease

Patrick Aloy*

Institute for Research in Biomedicine (IRB Barcelona), Barcelona, Spain & Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Structural systems pharmacology offers a novel way of approaching drug discovery by considering the global physiological environment of protein targets, and the effects derived of tinkering with them, without losing the key molecular details. In this talk I will highlight some recent advances in the structural annotation of cell networks and discuss their potential impact on some of the hottest areas of drug development. In particular, I will analyse recent structure-based strategies to target networks, protein interaction interfaces and allosteric sites, and how they will help in the development of more potent and specific treatments. Finally, I will show how the mapping genetic variations onto protein networks, beyond the pharmacological targets, can rationalize inter-individual variability in drug response, giving valuable hints to advance towards personalized medicine.

Structure Determination of Genomes

Marc Marti-Renom*

Gene Regulation. Stem Cells and Cancer Program. Center for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Dr. Aiguader 88, Barcelona 08003, Spain., Barcelona, Spain & CNAG-CRG, Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Baldiri Reixac 4, Barcelona 08028, Spain., Barcelona, Spain

The three-dimensional (3D) architecture of a genome determines the spatial localization of regulatory elements and the genes they regulate. Thus, elucidating the 3D structure of a genome may result in significant insights about how genes are regulated. The current state of the art in experimental methods, including light microscopy and cell/molecular biology, are now able to provide detailed information on the position of genes and their interacting partners. However, such methods by themselves are not able to determine the high-resolution 3D structure of genomes or genomic domains. We have developed a computational module of the Integrative Modeling Platform (IMP, http://www.integrativemodeling.org) that uses chromosome conformation capture data to determine the 3D architecture of genomic domains and entire genomes. This approach, through the visualization of looping interactions between distal regulatory elements, allows characterizing global chromatin features and their relation to gene expression.

Elastic Network Modelling in the Study of Macromolecular Assemblies

Turkan Haliloğlu*

Chemical Engineering Department and Polymer Research Center, Boğaziçi University, 34342 Bebek, İstanbul

Proteins are complex machines that display a wide range of functions in the cell. Within the sequence, structure and function paradigm, the intrinsic dynamics is essential for function and underlies evolution. Elastic Network Models (ENMs) have proved to be an important means in the prediction of local to global fluctuations. Here, low resolution ENMs, Gaussian Network Model (GNM) and Anisotropic Network Model (ANM), and a hybrid method that combines ANM with Langevin Dynamics (ANM-LD), will be visited with the case studies. It will be demonstrated how the latter methods plausibly disclose dynamic infrastructure and functional motion and provide insights into mechanisms of protein-protein/ligand interactions, allosteric regulation and molecular assembly. Understanding and exploiting intrinsic dynamics of proteins are thus likely to open new ways for protein design and discovery.

Poster Presentations

P1	3D modeling of BmpA, BmpB, BmpC and BmpD from Borrelia burgdorferi
P2	High resolution structural insights into Species D Adenovirus receptor usage
P3	Introducing knowledge-based terms in a consensus approach for the scoring of protein-protein docking models
P4	Combining multi-scale modelling methods to decipher molecular motions of a Branching Sucrase from Glycoside-Hydrolase Family 70
P5	Probing Structural Flexibility of the Neuronal Calcium-Sensor Synaptotagmin-1 by Cross-Linking Mass Spectrometry and Molecular Dynamics
P6	High-Order Protein-Protein Interactions: a Benchmark for the Docking, Integrative, and Homology Modelling Community
P7	Advancing multi-scale simulation methods for biological membrane systems with applications to confined environments and membrane fusion
P8	Identification of binding sites for selective inhibitors of inducible nitric oxide synthase based on the interaction map of known structures
P9	A Complete Pipeline for Enabling Efficient and Timely NMR Structural Biology on Challenging Pharmaceutical Targets
P10	Structural studies of mTORC interactions and regulation.
P11	Discovering the interaction models of CPT1-C
P12	Experimental data-driven modeling of RNA and RNA-protein complexes
P13	Molecular specificity of TAM family receptors towards different ligands
P14	Protein-protein docking by MREMD simulations with the coarse-grained UNRES force field.
P15	Molecular Architecture and Spacer Acquisition in Type I-F Cas1:Cas2-3 CRISPR Complex by Structural Mass Spectrometry
P16	Membrane protein docking with HADDOCK
P17	Distinct motifs differentiate MYB-bHLH transcription factor interactions
P18	Implementation of a biocomputing platform to settle a new drug discovery pipeline towards post-synaptic receptors.
P19	A drug from a bug
P20	Small molecule based stabilization of kidney tight junction proteins: A novel nephro-protective strategy
P21	Interaction interfaces in toxin-antitoxin systems: Probing relationships among paralogues to evaluate potential for cross-reactivity and understand the mode of antitoxin action
P22	Inhibition of LRRK2 by interfering with its processive cycle
P23	Structural and functional characterization of RNA-binding proteins involved in the Amyotrophic Lateral Sclerosis
P24	In Silico Study on the Structure and Dynamics of the Ryanodine Receptor

P1: 3D modeling of BmpA, BmpB, BmpC and BmpD from Borrelia burgdorferi

Mia Åstrand[a], Julia Cuellar[b], Jukka Hytönen[b], Tiina A Salminen[a]

[a] Structural Bioinformatics Laboratory, Åbo Akademi University, Turku, Finland
[b] Department of Medical Biochemistry and Genetics, University of Turku, Finland

B. burgdorferi is one of the main Borrelia species causing Lyme disease in humans. The pathogens are transmitted by the Ixodes ticks, and there are 60 000 – 200 000 Lyme disease infections in Europe annually. The BmpA, BmpB, BmpC and BmpD proteins are expressed by B. burgdorferi in infected patients, but the exact role of the proteins is still unknown. The Bmp proteins are reported to be homologous to T. pallidum PnrA (Purine nucleoside receptor A), which has been characterized as a substrate-binding lipoprotein of the ATP-binding cassette (ABC) transporter family, preferentially binding purine nucleosides. Based on our 3D homology models, the Bmp proteins share the typical fold of the substrate-binding protein family. Moreover, the residues involved in binding the ribose moiety of the nucleoside are highly conserved in the Bmp models, whereas the residues in the purine binding site are less conserved. In particular, the BmpC model has differences in the residues binding the base moiety of the nucleoside. In conclusion, the revealed differences indicate that the Bmp proteins could prefer different nucleosides and, thus, might have distinct biological functions.

P2: High resolution structural insights into Species D Adenovirus receptor usage

Alexander Baker[a], Alexander Greenshields-Watson[b], Hanni Uusi-Kerttula[a], Lynda Coughlanc, David Cole[d], Pierre Rizkallah[b], Alan Parker[a]

[a] Division of Cancer and Genetics, Cardiff University, Cardiff, Wales, United Kingdom
[b] Division of Infection and Immunity, Cardiff University, Cardiff, Wales, United Kingdom
[c] Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York City, New York, USA
[d] Immunocore Ltd, Abingdon, Oxon, United Kingdom

Vectors based on adenovirus are of increasing clinical significance due to their ease of manipulation, ability to express large transgenes, and scale up potential. Furthermore, their highly immunogenic nature makes them promising candidates for vaccine-based applications, and recently vaccines based on the low seroprevalance, species D adenoviral vectors, Ad48 and Ad26 have shown responses and progressed to phase I and III clinical trials for the treatment of HIV and Ebola, respectively. However, this clinical development has proceeded despite limited knowledge surrounding the mechanisms by which these viruses infect cells.

To investigate the potential means of cellular attachment of Ad26 and Ad48 we generated high resolution X-ray crystallographic structures from recombinant fiber-knob proteins. We utilised these structures in silico to create energy minimised models of the fiber-knob in complex with the major adenovirus receptors hCAR and CD46.

Our finding in silico were confirmed by in vitro competition inhibition assays in Adenovirus receptor expressing cell lines. Our findings highlight that Ad26 and Ad48 fiber-knob domain is incapable of forming a stable complex with CD46. Models of the fiber-knob in complex with hCAR demonstrate that while Ad26 and Ad48 possess the necessary residues to form an interface, a previously undescribed mechanism of steric inhibition is likely to diminish this affinity.

These findings confirm that CD46 or hCAR are unlikely to function as the primary cellular receptor for Ad26 or Ad48, despite previous reports to the contrary. Additional studies are required to fully characterise the unknown receptor for these clinically significant viruses.

P3: Introducing knowledge-based terms in a consensus approach for the scoring of protein-protein docking models

Didier Barradas-Bautista[a], Anna Vangone[a], Zhen Cao[a], Luigi Cavallo[a], Romina Oliva[b]

[a] Kaust Catalysis Center, King Abdullah University of Science and Technology, Saudi Arabia
[b] Department of Sciences and Technologies, University "Parthenope" of Naples, Italy

Protein-protein complexes play critical roles in various biological processes, such as antibody recognition, enzyme catalysis, etc. Three-dimensional structures provide atomic details of such complexes allowing the development of binding molecules to regulate them. However, the discovery rate of protein-protein interactions surpasses the speed of their structural determination. Therefore, in silico modeling of by molecular docking can help to fill this increasing gap.

Protein-protein docking is a widely used modeling approach and is often capable of generating a pool of models within which a near-native structure can be found. These models have thus to be analyzed and scored in the attempt to single out the correct ones.Still, to date, no program can provide a single docking solution with a high enough confidence to be correct making scoring a critical step during the calculation.

Traditionally, scoring functions for protein-protein docking poses are energy-based, relying on combinations of physical or empirical energy terms, or "knowledge-based", using properties derived from experimental structures of protein-protein complexes.

CONSRANK is the first pure consensus method, which ranks models based on their ability to match the most conserved contacts in the ensemble they belong to.

We tested CONSRANK on CAPRI rounds, where it featured as one of the methods of scoring with the highest number of correct models submitted, having on average 9 over 10 correct models for the successful targets.

Clust-CONSRANK is a modified version that includes a contact-based clustering of the models as a preliminary step of the scoring process enriching the top ten ranking with near-native solutions.

Currently, we are developing an implementation of the CONSRANK scoring function, by introducing knowledge-based weights between residues at the complex interface based on their physicochemical nature. We'll present preliminary results of the scoring function(s) tested in a set of the protein-protein benchmark 5.

P4: Combining multi-scale modelling methods to decipher molecular motions of a Branching Sucrase from Glycoside-Hydrolase Family 70

Akli Ben Imeddourene[a], Jérémy Esque[a], Magali Remaud-Siméon[a], Isabelle André[a]

[a] Université de Toulouse; Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés, INSA, CNRS, INRA 135 Avenue de Rangueil, F-31077 Toulouse, France

Glucansucrases from Glycoside-Hydrolase Family 70 are valued tools in chemistry to generate glycodiversification. Of great biotechnological interest, these enzymes use sucrose, a very cheap and abundant agro-resource, to catalyze the synthesis of a range of carbohydrates. Here, our study was focused on the first engineered enzyme, called ΔN123-GB-CD2, which is specific of the α-(1→2) branching of dextran, a rare and non-digestible linkage. Although its original U-shape three-dimensional organization has been recently established, the detailed investigation and description of the structural organization and the functional role of macromolecular motions of these multi-domain enzymes are still missing to provide a comprehensive understanding of the enzyme reaction.

By combining long molecular dynamics simulation (1μs) and multiple analyses (NMA, PCA, Morelet Continuous Wavelet Transform and Cross Correlations Dynamics), we investigated here the dynamics of ΔN123-GB-CD2 alone and in interaction with sucrose substrate. Overall, our results provide the detailed picture at atomic level of the hierarchy of motions occurring along different timescales and how they are correlated, in agreement with experimental structural data. The findings of our in silico study might offer a novel insight on key regions and amino acid residues that could be targeted to design enzyme variants endowed with improved properties for biotechnological applications.

Keywords ΔN123-GB-CD2, glucansucrase, MD, NMA, carbohydrates, multi-scale modelling, enzyme sequence-structure-dynamics relationships

P5: Probing Structural Flexibility of the Neuronal Calcium-Sensor Synaptotagmin-1 by Cross-Linking Mass Spectrometry and Molecular Dynamics

Julian Bender[a], Caroline Haupt[a], Matteo T Degiacomi[b], Carla Schmidt[a]

[a] Interdisciplinary research center HALOmem, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany
[b] Department of Chemistry, Durham University, UK

A key step in neurotransmission is the fusion of synaptic vesicles with the presynaptic membrane. It is dependent on the intracellular level of Ca2+ and is tightly regulated by the calcium-sensor Synaptotagmin-1 (Syt-1). Syt-1 consists of a short transmembrane domain that links the protein to the membrane of synaptic vesicles followed by two calcium-binding domains connected through an intrinsically disordered linker region (IDR). This structural flexibility of Syt-1 may play a major role in the dynamics of neurotransmission.

To study the structural flexibility of Syt-1 under defined conditions, we recombinantly expressed Syt-1 in E. coli and purified the protein. Conformational space of Syt-1 in the absence of interaction partners was probed by chemical cross-linking using the lysine-reactive cross-linker BS3 followed by tryptic digestion and peptide analysis by LC/MS-MS (XL-MS). Cross-links were identified by database searching and manually validated by their spectral quality. In a second set of experiments the effect of Ca2+ on the structure was investigated by adding Ca2+ followed by XL-MS. As Syt-1 has been described to interact with membrane lipids, we investigated their effect on its dynamic behaviour. Therefore, Syt-1 was incorporated into liposomes of different composition and analyzed by XL-MS. For more detailed studies, we also generated and expressed truncated constructs of Syt-1 composed of the calcium-binding domains with or without the IDR .

To interpret the obtained distance restraints, we computed molecular dynamics simulations starting with two high-resolution structures of an open and a closed conformation of Syt-1. Using DynamXL, a software tool that incorporates side-chain flexibility into the calculation of inter-amino acid distances, we integrated the simulated structures with the experimental data. For this, we compared theoretical cross-linking distances of all structures determined in silico with those obtained experimentally. Our study indicates that Syt-1 features a high flexibility that is modified by specific interaction partners.

P6: High-Order Protein-Protein Interactions: a Benchmark for the Docking, Integrative, and Homology Modelling Community

Martino Bertoni[a], Patrick Aloy[a, b]

[a] Joint IRB-BSC-CRG Program in Computational Biology, Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology, Barcelona, Catalonia, Spain
[b] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Catalonia, Spain

Structural biology is projected towards determining the structure of molecular machines. These are big, multi-components protein complexes taking part to critical cellular processes. As highlighted by recent CASP-CAPRI experiments, modeling such high-order protein-protein interactions is an open challenge for the community. We extracted from the archive of known structures, the Protein Data Bank (PDB), a benchmark set of 498 non-redundant heteromeric protein complexes composed by more than three unique components. Each of these complexes can serve as "target" for modelling approaches and is complemented with available homologous subcomplexes or individual components (templates) at different range of sequence identities. 19% of the targets can be entirely modelled by homology docking, 32% has partial information on subunits' interactions but the full set of components is available for docking. The remaining 49% of targets has partial to none structural coverage and hence requires ab-initio modelling and docking. All datasets can be downloaded or visually inspected through an ad-hoc website.

P7: Advancing multi-scale simulation methods for biological membrane systems with applications to confined environments and membrane fusion

Astrid Brandner[a], Fabio Sterpone[a], Philippe Derreumaux[a], Marc Baaden[a]

[a] Laboratoire de Biochimie Théorique - UPR 9080 Institut de Biologie Physico-Chimique

Mitochondrial dynamics is based on a delicate balance between fusion and fission that allows a normal cell metabolism. To approach the study of mitochondral fusion, our current work focuses on the scaffolding and tethering of the mitochondrial outer membranes. Here we will use an homology model of Fzo1, an outer mitochondrial membrane protein from yeast, to study its contribution to membrane fusion and tethering[1]. For this purpose we have used both, atomistic and coarse-grained(CG) approaches to understand the conformational dynamics and assembly of this system at different oligomerisation states and levels of detail. By coarse-graining we reduce the degrees of freedom in a system in comparison to a traditional all-atom model. That gives the possibility to study bigger systems as well as extend the time-scale of our simulations with less computational cost. Besides from the study of Fzo1 tethering, we are interested in having a CG lipid model compatible with OPEP. OPEP is an implicit-solvent CG protein force field previously developed in the Lab that does not restraint the secondary structure of the proteins and has detailed representation of the backbone[2]. Here we present a first implementation of a CG lipid force field compatible with OPEP. This implementation is also available to study hydrodynamic effect[3], allowing an original multi-scale approach on simulations of complex biological systems where fluid dynamics is important.

[1] De Vecchis D et al. A membrane-inserted structural model of the yeast mitofusin Fzo1. Scientific Reports volume 7, Article number: 10217 (2017)

[2] Sterpone F et al. The OPEP protein model: from single molecules,amyloid formation, crowding and hydrodynamics to DNA/RNA systems. Chem. Soc. Rev., 2014,43, 4871

[3] Sterpone F et al. Protein Simulations in Fluids: Coupling the OPEP Coarse-Grained Force Field with Hydrodynamics. J Chem Theory Comput. 2015 Apr 14;11(4):1843-53

P8: Identification of binding sites for selective inhibitors of inducible nitric oxide synthase based on the interaction map of known structures

Maria Bzówka[a, b], Jan Vítečekc, d, Lukáš Kubalac, d, Jiří Damborskýc, e, Artur Góra[a]

[a] Tunneling Group, Biotechnology Centre, Silesian University of Technology, Krzywoustego 8, 44-100 Gliwice, Poland
[b] Faculty of Chemistry, Silesian University of Technology, ks. Marcina Strzody 9, 44-100 Gliwice, Poland
[c] International Clinical Research Center, Center of Biomolecular and Cell Engineering, St. Anne's University Hospital Brno, Pekařská 53, 656 91 Brno, Czech Republic
[d] Institute of Biophysics, AS CR, Královopolská 135, 612 65 Brno, Czech Republic
[e] Loschmidt Laboratories, Department of Experimental Biology and Research Centre for Toxic Compounds in the Environment RECETOX, Faculty of Science, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic

Nitric oxide synthases (NOSes) are enzymes catalysing the production of nitric oxide (NO), which is a crucial signalling molecule that regulates various physiological functions. NO acts in the cardiovascular system and it is also a neurotransmitter immune response modulator. Moreover, NO can be involved in numerous pathological processes like septic shock, hypertension, and atherosclerosis. The overproduction of NO contributes to the pathogenesis of acute and chronic inflammatory processes. In mammals, NO is being produced by one of the three NOS isoforms: neuronal (nNOS), endothelial (eNOS) or inducible (iNOS). A great effort has been made to design selective inhibitors of NOSes.

In the present study, we have analysed forty crystal structures of iNOS deposited in the Protein Data Base and summarized parts of the iNOS targeted by various inhibitors until today. The structures of iNOS-ligand complexes were analysed using Ligplot tool. Then, hierarchical clustering method was used to group amino acids based on the interaction similarity. Three main regions in the protein, where binding of ligands is possible, were identified: (i) the catalytic site located above the heme, (ii) the entrance channel to the active site and (iii) the part of the active centre with the tetrahydrobiopterin binding site. In total, 28 amino acid residues and the heme were identified as interacting with ligands, including 4 residues, which have not been identified by any previous research. These amino acids may become targets for designing new selective inhibitors for iNOS that will not compete with the natural substrate of the reaction and with other NOS isoforms.

Acknowledgement This work was supported by the National Science Centre, Poland (DEC-2015/18/M/NZ1/00427) and project no. LQ1605 from the National Program of Sustainability II (MEYS Czech Republic).

P9: A Complete Pipeline for Enabling Efficient and Timely NMR Structural Biology on Challenging Pharmaceutical Targets

Marta G. Carneiro[a], Stephan Theisgen[a], Eiso AB[a], Kavya Chepuri[a], Johan Hollander[a], Dipen Shah[a], Elseline van der Spek[a], Willem-Jan Waterreus[a], Gregg Siegal[a]

[a] ZoBio B.V.

By now the power of structure-based drug design (SBDD) is so widely recognized that it has become essentially the de facto approach for target based, small molecule campaigns. NMR structural biology is indispensable for cases where crystallization fails, or where crystal contacts create artefactual binding sites. NMR can provide structural information at different levels of resolution, with a trade-off between the amount of information/ambiguity versus throughput. Combining NMR data (e.g. chemical shift perturbations and intermolecular NOEs) with docking approaches, provides invaluable structural information for medicinal chemistry efforts. However, NMR structural biology efforts on pharmaceutical targets are often hindered by a variety of challenges including: poor yields of recombinant protein, limited solubility or instability of the protein and long experimental and analysis time needed for NMR resonance assignment and structural information. ZoBio has been implementing and developing comprehensive strategies to enable efficient and timely NMR structural biology and has routinely obtained protein-ligand co-structures by combining sparse NOE data with data-driven docking. The impact of these strategies on enabling NMR structural biology on difficult targets will be highlighted with various examples.

P10: Structural studies of mTORC interactions and regulation.

Yana Dem'yanenko[a], Nikolaus Dietz[a], Edward Stuttfeld[a], Stefan Imseng[a], Evelyn Sauer[a], Timm Maier[a]

[a] Biozentrum, University of Basel

Mammalian target of rapamycin complexes 1 and 2 (mTORC1,2) regulate cell growth and metabolism in response to metabolic and growth factor stimuli. Activation of otherwise cytosolic mTORC1 occurs at the lysosomal surface. Understanding the architecture of lysosomal mTORC1 super-complexes and the requirements for mTORC translocation to the lysosome is a prerequisite for deciphering mTORC signalling. In response to amino acid sensing, mTORC1 is recruited to the outer lysosomal membrane by an intricate interplay between complexes of lipid-anchored proteins, small GTPases, input sensing and adapter proteins. Additional protein complexes modulate the input sensing or act as guanine exchange factors or GTPase activating proteins. Here, I will discuss current strategies for studying mTORC1,2 activation via structural studies based on X-ray crystallography and cryo electron microscopy single particle analysis, as well as possible avenues for exploration using more high throughput techniques. Aberrant mTOR signalling has been linked to cancer and metabolic disorders. Our work will provide important insights for future design of pathway specific inhibitors.

P11: Discovering the interaction models of CPT1-C

Jose Luis Dominguez Meijide[a], Núria Casals Farré[a], et al.

[a] Department of Basic Sciences, International University of Catalunya

Carnitine Palmytoil Transferase 1 (CPT1) is a mitochondrial enzyme in charge of the acylation process in carnitines. It has three isoforms: CPT1-A, CPT1-B, and CPT1-C. CPT1-A (the liver isoform) is found in most body tissues, except skeletal muscle cells and brown adipose tissue (BAT). CPT1-B (the muscle isoform) is found in skeletal muscle cells and BAT. CPT1-C (the brain isoform) is present mainly in the neurons, and shows a series of key differences with the previous isoforms. Firstly, it is located mainly in the endoplasmic reticulum rather than mainly the mitochondria. Secondly, and perhaps more important, it doesn't perform a catalytic function, unlike its isoforms; rather, it seems to act as a malonyl-CoA sensor in the cell, triggering different metabolic responses based on the amount of malonyl-CoA present. However, the specific mechanisms by which CPT1-C interacts with other proteins when triggering these metabolic responses are not well known. Our project aims to discern some of the protein-protein mechanisms involving CPT1-C, using molecular dynamics simulations and docking simulations. Two of the target proteins studied are Glua1 (which is part of the AMPA receptors for glutamate) and Sac1 (involved in Golgi apparatus mobility); for their study we modeled several vectors of approach in a membrane surface between key areas of the two pairs of proteins (Glua1 and Sac1, each paired with a different part of CPT1C). In addition to the interaction trajectories, and in order to better understand the sensing mechanisms of CPT1-C, we're also performing a series of docking experiments using different acyl-CoA compounds and studying their interactions with the the different access areas to the analogue region of the catalytic area in CPT1-C.

P12: Experimental data-driven modeling of RNA and RNA-protein complexes

Pritha Ghosh[a], Janusz Bujnicki[a, b]

[a] Laboratory of Bioinformatics and Protein Engineering, International Institute of Molecular and Cell Biology in Warsaw, ul. Ks. Trojdena 4, PL-02-109 Warsaw, Poland
[b] Bioinformatics Laboratory, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, ul. Umultowska 89, PL-61-614 Poznan, Poland

The majority of known RNAs exert their function in complex with proteins to form ribonucleoproteins (RNPs), at one or more stages of their life cycle, which emphasizes the importance of studying RNA-protein interactions. During my PhD, I have developed and implemented a combination of structural and sequence-based methods to survey for RNA-binding proteins (RBPs) on a genome-wide scale. However, in order to understand the molecular function of these proteins, RNA-RBP complex structures are necessary. Unfortunately, due to the difficulties associated with the experimental determination of high resolution structures of RNA and RNA-protein complexes, computational methods have been developed for their three dimensional (3D) structure predictions. Work done in the Bujnicki laboratory, has led to the development of many structural bioinformatics algorithms for the prediction of 3D structures of RNAs, proteins, RNA-protein, as well as RNA-ligand complexes, for example, SimRNA, PyRy3D, NPDock, SimRNP etc. Over the years, work done by multiple research groups have shown the importance of integrating experimental data to refine and improve computationally predicted models. In the past few years, the Bujnicki group has also delved deeper into modeling of biomolecular structures by integrating data from various experimental sources, with particular emphasis on data derived from small-angle X-ray scattering (SAXS) and selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) experiments. As a part of the Bujnicki group, in the past few months, I have focussed on 3D structure modeling of RNAs and RNA-protein complexes. Here, I will present a case study for modeling the apo structure (aptamer and expression platform) of a riboswitch, with the aid of diffraction data from X-ray crystallography and structural restraints derived from SHAPE experiments, both of which have been acquired in our laboratory. We also plan to model the structural changes in the holo form of this riboswitch, upon binding of the ligand.

P13: Molecular specificity of TAM family receptors towards different ligands

Tülay Karakulak[a, b], Ezgi Karaca[a, b]

[a] Izmir Biomedicine and Genome Center
[b] Izmir International Biomedicine and Genome Institute

The transmembrane receptor tyrosine kinase TAM (Tyro3, Axl, Mer) family regulates an intriguing mix of cellular processes. TAM family members are activated when their extracellular domains are bound to one of their two paralogous ligands, i.e., Gas6 and Pros1. Although Gas6 and Pros1 are significantly similar (80% structure, 40% sequence similarity), Gas6 can bind to all TAM receptors with different affinities [Kd(Gas6-Axl) > Kd(Gas6-Tyro) > Kd(Gas6-Mer)] while Pros1 can interact only with Tyro3 and Mer [Kd(Pros1-Tyro) > Kd(Pros1-Mer)]. Expanding on this, we aim at characterizing TAM family's ligand selectivity through a structure-based approach. For this, we modeled the complex structures of all possible TAM-ligand interactions (Tyro-Gas6, Tyro-Pros, Mer-Gas6, Mer-Pros, Axl-Pros) by using the available Axl-Gas6 complex structure as a template. This process is followed by the refinement of each complex with HADDOCK. Strikingly, HADDOCK score of the complex structures, obtained at the end of refinement sessions, correlated very well with the observed experimental binding affinities. Detailed interaction analyses also revealed the grounds of TAM receptor selectivity towards its ligands: While Gas6 prefers electrostatic interactions, Pros1's binding is majorly guided through van der Waals contacts. To further validate our findings, we will integrate conservation and coevolution data into our analyses, which should lead us to the deduction of TAM-ligand interaction fingerprints. This generic strategy could in principle be applied to characterize the binding specificities of the other receptor tyrosine kinase families towards their ligands (e.g. epidermal growth factor and platelet-derived growth factor systems).

P14: Protein-protein docking by MREMD simulations with the coarse-grained UNRES force field.

Pawel Krupa[a, b], **Agnieszka Karczyńska[b]**, Adam Liwo[b], Cezary Czaplewski[b]

[a] Institute of Physics, Polish Academy of Sciences, Al. Lotnikow 32/46, PL-02-668 Warsaw, Poland.
[b] Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland.

Prediction of protein structures is one of the most important problems of current computational biology and bioinformatics. Many proteins in living organisms form multimeric structures or are involved in the complexes with other biomacromolecules, such as other proteins or small ligands.
Prediction of protein structure in the monomeric state is already very challenging and therefore the structures of their complexes are even more tricky to determine. Such complexes can be too large in a case for computational and time capabilities of present supercomputers, especially if every atom of every molecule in the complex is considered in the calculations. Therefore, utilizing a coarse-grained representation and some simplifications in calculations are recommended. The physics-based UNRES model developed in our lab assumes only two interaction sites per amino-acid residue: united peptide group and united side chain. Our protocol of molecular docking is consists of the generation of initial models of complex structures and performing multiplexed replica exchange molecular dynamics (MREMD) simulations with the coarse-grained UNRES force field.
The protocol was designed and tested recently for various protein-protein complexes with a chain length from 51 to 586 amino-acid residues (rigid-body docking; 20 complexes) and for various protein-ligand (small peptide) complexes (rigid protein structure and totally flexible ligand structure; 33 complexes). The results show that the UNRES force field is able to predict the structures of protein complexes with the good quality given sufficient sampling.

P15: Molecular Architecture and Spacer Acquisition in Type I-F Cas1:Cas2-3 CRISPR Complex by Structural Mass Spectrometry

Oleg Klykov[a, b], Fagerlund Robert D.[c], Wilkinson Max E.[c, d], Barendregt Arjan[a, b], Grant Pearce F.[e], Kieper Sebastian N.[c], Bostina Mihnea[c, f], Staals Raymond H. J.[c], Fineran Peter C.[c], Heck Albert J. R.[a, b], Scheltema Ricahrd A.[a, b]

[a] Biomolecular Mass Spectrometry and Proteomics, Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University, Universiteitsweg 99, 3584 CG Utrecht, The Netherlands
[b] Netherlands Proteomics Center, Padualaan 8, 3584 CH Utrecht, The Netherlands
[c] Department of Microbiology and Immunology, University of Otago, Dunedin 9054, New Zealand
[d] Department of Biochemistry, University of Otago, Dunedin 9054, New Zealand
[e] Biomolecular Interaction Centre and School of Biological Sciences, University of Canterbury, Christchurch 8020, New Zealand
[f] Otago Centre for Electron Microscopy, University of Otago, Dunedin 9054, New Zealand

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and their associated proteins (Cas) represent adaptive immune systems capture DNA fragments from invading bacteriophages and plasmids and integrate them as spacers into bacterial CRISPR arrays. Type I-F CRIPSR systems contain a unique fusion of Cas2, with the type I effector helicase and nuclease for invader destruction, Cas3, and employs "primed" mode of spacer acquisition. Primed acquisition is thought to have evolved to capture invaders that have escaped interference by mutating their protospacers or Protospacer Adjacent Motifs (PAMs). Altered PAM specificity is currently of high interest for genome editing, as the current strategies are constraint by one PAM making it difficult to e.g. target double-stranded breaks.

Integrative approach, comprising of crystallography, Electron Miscorscopy (EM) and Structural Mass Spectrometry (MS), leads to detailed characterization of the protein complex. Native MS and Shotgun Proteomics methods assist in defining of the complex stoichiometry. First crosslinking strategy uses the MS-cleavable disuccinimidyl sulfoxide (DSSO) linker together with a novel data analysis platform specifying initial structural models of similar complex components and providing insights into the Cas1:Cas2-3 complex organization. The second uses formaldehyde crosslinking to reveal the interaction site between the DNA and the CRISPR molecule and detected sites have been further validated through obtained EM-densities.

We present a structural model of the 400-kDa Cas14-Cas2-32 complex purified from Pectobacterium atrosepticum with bound protospacer substrate DNA. Two Cas1 dimers assemble on a Cas2 domain dimeric core, which is flanked by two Cas3 domains forming a groove where the protospacer binds to Cas1-Cas2. Postulated complex architecture is validated with the EM data and provides a solid basis for integration of this system in the CRISPR/Cas toolbox.

P16: Membrane protein docking with HADDOCK

Charlotte W. van Noort[a], **Panagiotis Koukos[a]**, Alexandre Bonvin[a]

[a] Bijvoet Center for Biomolecular Research, Faculty of Science - Chemistry, Utrecht University, Padualaan 8, Utrecht, 3584CH, the Netherlands

Membrane proteins play important roles in many aspects of cellular function such as cell communication, signal transduction and molecule transport. Due to these functions they account for about half of the known drug targets in humans. However their study by traditional structural characterisation methods, such as X-RAY crystallography, cryo-electron microscopy and tomography, and NMR is hampered by the fact that these proteins reside in the lipid bilayer making their expression, isolation and structural determination complicated. Computational methods are an attractive alternative for the study of membrane proteins.

HADDOCK is our information-driven biomolecular docking platform. It is one of the most popular docking servers with more than 10000 users worldwide. It has consistently been ranked as a top predictor in CAPRI - the blind docking competition. So far we have not implemented a membrane model in HADDOCK. In this work we present our ongoing efforts in this area. The implementation consists of a Z-restraining potential that limits the movement of transmembrane proteins along one axis. We have also implemented empirical desolvation energies optimised for membrane proteins. Our results indicate that we can accurately predict the conformations that proteins adopt in the membrane both in terms of tilt and residue coverage.

P17: Distinct motifs differentiate MYB-bHLH transcription factor interactions

Peter S. Millard[a, b], Birthe B. Kragelund[c], Meike Burow[a, b]

[a] Center of Excellence for Dynamic Molecular Interactions , Dept. of Plant and Environmental Sciences, University of Copenhagen, Denmark
[b] Copenhagen Plant Science Center, Dept. of Plant and Environmental Sciences, University of Copenhagen, Denmark
[c] Structural Biology and NMR Laboratory, Dept. of Biology, University of Copenhagen, Denmark

MYB and bHLH transcription factors (TFs) constitute the two largest TF families in plants, and play central roles in regulation of many plant-specific processes, including metabolism, environmental interactions and development (DOI: 10.1105/tpc.013839; DOI: 10.1016/j.tplants.2010.06.005). Several MYB and bHLH subgroups have co-evolved to regulate pathways, and proteins from the two TF families frequently interact physically with each other, both being required for specific promoter activation (DOI: 10.1111/j.1365-313X.2010.04459.x; DOI: 10.1016/j.molp.2014.11.022). In particular, MYB TFs regulating accumulation of anthocyanins interact with their bHLH interaction partners through a conserved motif in the N-terminal DNA-binding domain (DOI: 10.1111/j.1365-313X.2004.02183.x). We expected that such motifs derive from an ancestral interacting MYB-bHLH pair, and have since evolved to provide subgroup-specific interactions. However, we recently found that MYB TFs regulating accumulation of glucosinolates interact with their bHLH interaction partners through a different motif, located in the C-terminal transactivation domain. The finding that different MYB TFs utilize unrelated motifs to mediate interactions with bHLH TFs indicates that interaction between distinct MYB and bHLH subgroups may have evolved independently via convergent evolution. Our mutational analysis of the motif revealed that a few residues are absolutely critical for the interaction, while other residues could be important for modulating affinity or kinetics, and may contribute to subfunctionalisation among the MYB TFs carrying this motif (at least 8 in Arabidopsis thaliana). The motif was also found in MYB TFs from unrelated species that do not produce glucosinolates, indicating that the motif must be more ancient than this evolutionarily young pathway. The present work provides insight into the coevolution of MYB-bHLH complexes and the pathways they regulate.

P18: Implementation of a biocomputing platform to settle a new drug discovery pipeline towards post-synaptic receptors.

Rui Ribeiro[a], Alejandro Giorgetti[a, b]

[a] Department of Biotechnology, University of Verona, Ca' Vignal 1, Verona, Italy
[b] IAS-5 / INM-9: Computational Biomedicine - Institute for Advanced Simulation (IAS) / Institute of Neuroscience and Medicine (INM), Forschungszentrum Jülich, Jülich, Germany

GOMoDO is a GPCRs online modelling and docking web-server, developed in the Applied Bioinformatics Laboratory of the University of Verona. This biocomputing platform puts together state-of-the-art bioinformatics tools which allow users to effortlessly model GPCR structures and dock ligands to the model, and obtain biologically and pharmacologically relevant data. With a very easy user interface GOMoDo works through a consistent pipeline: protein sequence alignment, homology modelling and model quality assessment, and docking.

Although, GOMoDo was meant to be used by expert and non-expert users for GPCR-targeting drug design, we want to improve and extend it to other post-synaptic receptor families, specially to Cys-loop receptors. This superfamily of pentameric ion channels are widespread across the nervous system, being particularly involved in the learning and memory processes. They are also fundamental targets for clinically relevant drugs, such as neuromuscular blockers, barbiturates, benzodiazepines, and anaesthetics.

Today's cell and molecular biology no longer focus only on single macromolecules, but on the understanding of the chemical and biological functions of organisms. Therefore, we intend to implement in our biocomputing platform the ability to unveil the metabolic pathways and catalogue all the biological complexes and the relationships between them of the post-synaptic receptors.

Such complex biological systems rely on fundamental mathematical equations applied by computational methods. Hence, we will use computational frameworks that combine the strengths of rule-based and programmatic approaches (BioNetGen and Kappa) and Python numerical tools to assemble accurate, extensible, and reusable biological models.

Lastly, we also intend to provide the user the possibility of running in silico target fishing and virtual screening campaigns to find new targets against the post-synaptic receptors.

P19: A drug from a bug

Markus Roupé[a]

Hansa Medical

Donor-specific antibodies are a major obstacle to successful transplantation surgery. The current therapies to neutralize donor-specific antibodies are limited and ineffective in the most highly HLA-sensitized patients. IdeS is an IgG-degrading enzyme derived from Streptococcus pyogenes. It is an endopeptidase that very efficiently and specifically cleaves human IgG into F(ab')2 and Fc fragments. It thereby ameliorates both complement-dependent cytotoxicity and antibody-dependent cellular cytotoxicity in one single blow by cleaving human IgG in the hinge region. Hansa Medical is therefore trying to develop IdeS into a treatment that can be useful for desensitization of HLA incompatible patients before they receive a kidney transplant. In addition, the company seeks to broaden the scope of indications for the enzyme to also encompass treatment of autoimmune deceases. The major hurdle to overcome in order to make this a reality is the possible formation of anti-drug antibodies in patients. Since the enzyme is derived from a common human pathogen there are pre-existing memory B - and T-cells directed against IdeS and eventually high affinity antibodies are formed that start blocking the enzymatic activity. For long term treatment of autoimmune disease multiple dosings are necessary. The immunogenicity of IdeS therefore needs to be reduced significantly in order to reach this goal. Hansa Medical is therefore in the process of developing second and third generations of IdeS to meet these challenges. The B - and T-cell epitopes of these are being mapped in order to identify problematic amino acid residues. A better molecular model on the substrate-enzyme interactions is also being developed in order to help in the engineering process.

P20: Small molecule based stabilization of kidney tight junction proteins: A novel nephro-protective strategy

Amin Sagar[a, b], Ehtesham Arif[c], Ashish Solanki[c], FNU Ashish[d], Deepak Nihalani[c], et al.

[a] Indian Institute of Science Education and Research, Mohali, India.
[b] Centre national de la recherche scientifique, Montpellier, France.
[c] Division of Nephrology, Medical University of South Carolina, Charleston, SC, USA
[d] CSIR-Institute of Microbial Technology, Chandigarh, India

Junctional proteins Neph1 and ZO-1 and their interaction is required for maintaining the structural integrity of slit diaphragm, which is a critical component of kidney's filtration system. It has been previously shown that injury induces loss of this interaction. We hypothesized that strengthening this interaction may help in protecting kidney's filtration barrier and preserve kidney function. We derived a model of the cytoplasmic domain of Neph1 using Homology modelling and using Small Angle X-ray Scattering profile and Circular Dichroism as constraints. We then docked this structure to the crystal structure of ZO1 (cytoplasmic domain) and again used SAXS constraints along with mutational data to arrive at a model of Neph1-ZO1 complex. This model was screened for the presence of druggable pockets present at the interface of the two proteins. One such pocket was identified and a small molecule library was docked on to it. The six top scoring compounds were tested using biophysical/biochemical experiments with purified proteins and Isodesmosine (ISD), a rare naturally occurring amino acid was selected as the best candidate. Its ability to enhance Neph1-CD and ZO-1 binding was tested using various cell-based assays and animal models of kidney damage. Results from biochemical binding analysis showed that ISD enhanced Neph1 and ZO-1 interaction under in vitro and in vivo conditions. ISD treated podocytes were not only resistant to Puromycin aminonucleoside induced loss of transepithelial permeability, but ISD was also able to reverse the damage. Finally, mouse and zebrafish studies show that ISD protects from Nephrotoxic serum and Adriamycin induced renal damage. Our work demonstrates the power of integrative modelling in the field of drug discovery and presents the successful application of Protein-Protein interaction stabilizers for therapeutic intervention in kidney diseases which affect 10% of the world population.

P21: Interaction interfaces in toxin-antitoxin systems: Probing relationships among paralogues to evaluate potential for cross-reactivity and understand the mode of antitoxin action

Himani Tandon[a], Sankaran Sandhya[a], Raghavan Varadarajan[a], Narayanaswamy Srinivasan[a]

[a] Molecular Biophysics Unit, Indian Institute of Science, Bangalore

Type II toxin-antitoxin systems (TAs) are two-gene modules consisting of protein toxins and antitoxins. Under normal growth conditions, they occur as a complex but during adversities, lone toxins bind to cellular targets that are either RNA or enzymes like DNA gyrase, triggering survival responses in bacteria. We employed sequence and structure-based analyses to study 1) the evolutionary relationships within the 80 TA systems in M.tuberculosis and evaluate their potential for cross-reactivity 2) the mechanism by which an E.coli antitoxin reverts the activity of its cognate toxin. When clustered, TAs group into distinct sub-clusters based on their ability to relate to each other. Over 30 toxins, could relate to 40 other toxins, indicating that many are paralogous. An analysis of the alignments within each toxin sub-clusters showed similarities in their core domains and target-binding sites. Interestingly, evaluation of antitoxin alignments for each sub-cluster showed conservation in their DNA-binding sites. We extended our study to the interfaces between toxins and antitoxins to specifically identify a set of features common to each paralogous group, which can be used a) to design a peptide to modulate TA interaction b) to guide computational modelling of complexes for other TA systems. With this approach, TA complexes were built for 6 non-trivial cases in M.tuberculosis. The availability of these models will form the basis to support our ongoing studies of sequence and structural similarities at the predicted interface to explore cross-reactivity amongst the paralogous pairs. Moreover, our studies also identified and computationally characterized the interfaces of two hitherto unknown novel TA systems in M.tuberculosis which have also been validated by our bench-lab collaborators. Finally, we have employed normal-mode analysis on the E.coli CcdB-CcdA TA system to unveil the mechanism by which binding of CcdA at its interface on CcdB can have substantial effect on the gyrase binding interface.

P22: Inhibition of LRRK2 by interfering with its processive cycle

Dominika Wojewska[a], Ahmed Soliman[a], Arjan Kortholt[a]

[a] Department of Cell Biochemistry, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands

Mutations in leucine-rich repeat kinase 2 (LRRK2) are the most frequent genetic cause of Parkinson Disease (PD, OMIM #168600). LRRK2 belongs to the group of Roco proteins, which are characterized by the presence of Ras-like G-domain (Roc), C-terminal of Roc domain (COR), a kinase domain and several protein-protein interaction domains (InterPro ID: Q38SD2). LRRK2 has a complex activation mechanism, involving intra-molecular signalling and dimerization.

PD mutations in LRRK2 have been linked to increased kinase activity and decreased GTPase activity. Although LRRK2-specific, brain penetrant kinase inhibitors have been developed, most of them show side effects, and none have been approved for clinical use yet. As an alternative to direct kinase inhibition, GTPase inhibition or dimerization cycle disruption could be used [1].

LRRK2 exhibits a cross-talk between kinase and GTPase domains in its processive cycle. Studies have shown that dimeric form has both kinase and GTPase activity, while monomeric form is inactive [2, 3]. Therefore interfering with GTP hydrolysis and mono-dimerization cycle can inhibit the kinase domain indirectly. Therefore the goal of my project is obtaining a crystal structure of human RocCOR in GDP- and GTP-loaded states, to get detailed information on its dimerization interface and structural changes.

In order to disrupt dimerization, a series of conformationally restricted peptides was designed based on the crystal structure of a bacterial Roco homologue (PDB ID: 3DPU). Preliminary results show promise in this approach, and with the help of molecular docking and modelling a series of improved peptides, peptidomimetics or small compounds will be proposed.

[1] Rudenko et al., BMC Med. (2012), 10, 20
[2] Deyaert et al., Nat. Commun. (2017), 8, 1008
[3] Terheyden et al., Biochem. J. (2015), 465(1), 139–147

P23: Structural and functional characterization of RNA-binding proteins involved in the Amyotrophic Lateral Sclerosis

Juan Carlos Rengifo Gonzalez[a], Marie-Jeanne Clement Chai-Leong[a], Dominique Durand[b], David Pastre[a], Ahmed Bouhss[a].

[a] SABNP Laboratory, University of Evry Val d'Essonne, University Paris-Saclay, INSERM U1204, 91025 Evry, France
[b] Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, University Paris-Sud, University Paris- Saclay, Gif-sur-Yvette, France.

The Amyotrophic Lateral Sclerosis (ALS) is a neuronal disease caused by motor neuron degeneration in which several RNA-binding proteins (RBPs) aggregate in the cytoplasm forming toxic insoluble inclusions. The prion-like behavior of these proteins allows the spreading of the disease in the brain by a process still not well-understood. Among known RBPs, mutated TDP-43 has been found in these pathological cytoplasmic inclusions. TDP-43 contains two RRM (RNA-Recognition Motif) domains which are essential to RNA binding and a highly aggregation-prone domain involved in its dysfunction. The 3D structure of TDP-43/RNA complex is available. Our objective is to elucidate the main structural events responsible for the RNA recognition. We already obtained and purified several TDP-43 protein fragments labeled with 15N and 13C allowing us to analyze them by NMR. The comparison between NMR spectra obtained from the RNA-protein complexes allowed us to identify several residues which were mutated in order to determine their structural incidence on the RNA recognition mechanism. By combining NMR and Bioinformatic tools, an insight could be obtained about the induced effect on the TDP-43 structure upon RNA binding. Moreover, applying docking-based computational methods, we aim to screen new drugs which could interfere with the TDP-43 pathogenesis in affected cells. The in silico results will be validated by experimental data obtained by using NMR, biochemical and cellular approaches.

In Silico Study on the Structure and Dynamics of the Ryanodine Receptor

Anne-Elisabeth Molza[1*], Charles H. Robert[1],
Pierre Ducrot[2], and Marc Baaden[1]

[1] LBT-IBPC, 13 rue Pierre et Marie Curie, 75005 Paris
[2] Institut de Recherches Servier, 125 chemin de Ronde, 78290 Croissy sur Seine
* molza@ibpc.fr

Abstract

The Ryanodine Receptor (RyR) is a homotetrameric intracellular calcium channel. Each subunit contains more than 5000 residues (MW $\sim$ 2.2MDa). RyR1, the predominant isoform in skeletal muscle, plays a major role in the calcium release from the sarcoplasmic reticulum into the cytosolic compartment, being essential for excitation–contraction coupling [1,2].

Recent advances in the field of electron microscopy have made the elucidation of the 3D structure of membrane proteins possible, one example being RyR at a resolution of 3.8 Å. Based on this data, near-atomic cryo-EM-based models of both RyR1 and RyR2 in an open and a closed conformation could be obtained [3]. Nevertheless, these models are incomplete and do have many unknown and missing residues. Moreover, the allosteric mechanism of RyR is still unclear.

Here, we study the RyR1 structure and its conformational changes based on molecular modeling and normal mode analyses to suggest the transition pathway for the opening-closing mechanism. We obtain full-length models of RyR, analyze protein motions and suggest transition pathways between open and closed conformations.

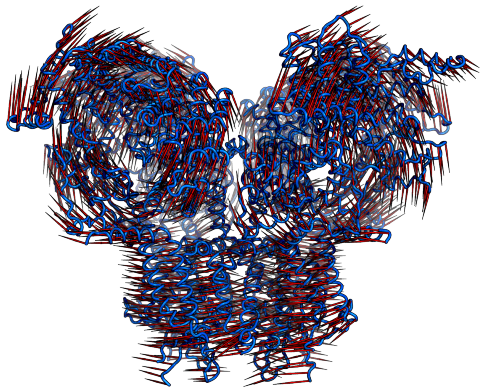


Figure 1: Prediction of the Cα atom displacements via low-frequency normal modes.

Keywords: Membrane protein, Allosteric motions, Molecular Modeling, Molecular dynamics, Integrative method

Acknowledgments
The Institut de Recherches Servier is acknowledged for funding and stimulating discussions.

References
[1] Nature. 2015 Jan 1; 517(7532): 44–49
[2] Protein Sci. 2017 Jan;26(1):52-68
[3] Cell. 2016 Sep 22;167(1):145-157.e17

List of Participants

Mia Åstrand
Åbo Akademi University
Tykistönkatu 6A
20660 Turku, Finland
mia.astrand@abo.fi

Alexander Baker
Cardiff University
Henry Welcome Building,
University Hospital of Wales, Academic Avenue
CF14 4XN Cardiff, United Kingdom
bakerat@cardiff.ac.uk

Dr. Didier Barradas-Bautista
KAUST
Beacon Av. House G3946
King Abdullah University of Science and Technology
23955-6900 Jeddah, Thuwal, Saudi Arabia
didier.barradasbautista@kaust.edu.sa

Akli Ben Imeddourene
Institut National des Sciences Appliquées
135 Avenue de Rangueil
31400 Toulouse, France
benimedd@insa-toulouse.fr

Julian Bender
Martin Luther University Halle-Wittenberg
Kurt-Mothes-Str. 3 a
06120, Halle (Saale), Germany
julian.bender@bct.uni-halle.de

Martino Bertoni
Institute for Research in Biomedicine
Parc Científic de Barcelona, C/ Baldiri Reixac 10
08028 Barcelona, Spain
martino.bertoni@gmail.com

Astrid Brandner
Laboratoire de Biochimie Théorique - UPR 9080
Institut de Biologie Physico-Chimique
13 rue Pierre et Marie Curie
75005 Paris, France
brandner@ibpc.fr

Maria Bzówka
Silesian University of Technology
Akademicka 2A
44-100 Gliwice, Poland
m.bzowka@tunnelinggroup.pl

Dr. Marta Carneiro
ZoBio B.V.
BioPartner 2 building
J.H. Oortweg 19
2333 CH Leiden, The Netherlands
mcarneiro@zobio.com

Yana Dem'yanenko
Biozentrum, University of Basel
Klingelbergstrasse 70
4001 Basel, Basel-Stadt, Switzerland
yana.demyanenko@unibas.ch

Dr. Pritha Ghosh
Laboratory of Bioinformatics and Protein Engineering
International Institute of Molecular and Cell Biology in Warsaw
ul. Ks. Trojdena 4
PL-02-109 Warsaw, Poland
pghosh@genesilico.pl

Tülay Karakulak
Izmir Biomedicine and Genome Center
Dokuz Eylul University Health Campus
Mithatpasa St.58/5
35340 Izmir, Turkey
tulaykarakulak@gmail.com

Agnieszka Karczynska
University of Gdansk
Jana Bażyńskiego 8
80-309 Gdańsk, Poland
agnieszka.karczynska@phdstud.ug.edu.pl

Oleg Klykov
Biomolecular Mass Spectrometry and Proteomics Group, Utrecht University
Padualaan 8
3583 CH Utrecht, The Netherlands
nazza39@gmail.com

Panagiotis Koukos
Utrecht University
Heidelberglaan 8
3584 CS Utrecht, The Netherlands
p.koukos@uu.nl

Dr. Jose Luis Dominguez Meijide
Departament of Basic Sciences
International University of Catalunya
Carrer de Bartrina 36
08191 Rubí, Barcelona, Spain
jldominguez@uic.es

Peter Sørensen Millard
University of Copenhagen
Thorvaldsensvej 40, T229
1871 Frederiksberg C, Denmark
pvm@plen.ku.dk

Anne-Elisabeth Molza
Institut de Biologie Physico-Chimique- Laboratoire de Biologie Théorique (IBPC-LBT)
13 rue Pierre et Marie Curie, 2nd floor, room 242
75005 Paris, France
molza@ibpc.fr

Juan Carlos Rengifo Gonzalez
University of Evry / University of Paris-Saclay
SABNP Laboratory
Rue du Père Jarlan, Bâtiment Maupertuis
91025 Evry, France
juancarlos.rengifogonzalez@univ-evry.fr

Rui Ribeiro
University of Verona
Laboratorio di Bioinformatica Applicata, DIP. Biotecnologie, Strada le Grazie 15 Ca'Vignal 1
37134 Verona, Italy
rui.ribeiro@univr.it

Dr. Markus Roupé
Hansa Medical
Scheelevägen 22
223 63 Lund, Sweden
markus.roupe@gmail.com

Dr. Amin Sagar
Indian Institute of Science Education and Research
Knowledge City, Sector 81
140306 Mohali, Punjab, India
aamin.sagar@gmail.com

Himani Tandon
Indian Institute of Science, Bangalore
CV Raman Rd, Bengaluru, Karnataka
560012 Bengaluru, Karnataka, India
himani@iisc.ac.in

Dominika Wojewska
University of Gronigen
Nijenborgh 7
9747 AG Groningen, Netherlands
d.n.wojewska@rug.nl

Grant Awardees

The following participants applied for and were awarded sponsoring, in the form of a travel grant or reduced registration fees:

Mia Åstrand (sponsored by Bruker)

Alexander Baker

Julian Bender

Astrid Brandner

Maria Bzówka (sponsored by Bruker)

Pritha Ghosh

Tülay Karakulak (sponsored by Bruker)

Agnieszka Karczynska

Rui Ribeiro

Amin Sagar

Himani Tandon

List of Lecturers

Dr. Patrick Aloy
Institute for Research in Biomedicine
c/ Baldiri i Reixac, 10-12
08028 Barcelona, Spain
patrick.aloy@irbbarcelona.org

Dr. Jessica Andreani
CEA I2BC (Institute for Integrative Biology of the Cell)
Institut des Sciences du Vivant Frédéric Joliot
Bâtiment 144 - Point courrier 22
91191 Gif sur Yvette, France
jessica.andreani@cea.fr

Dr. Paul Bates
The Francis Crick Institute
44 Lincoln's Inn Fields
WC2A 3LY London, United Kingdom
paul.bates@crick.ac.uk

Prof. Alexandre M.J.J. Bonvin
Utrecht University, Faculty of Science – Chemistry
Padualaan 8
3584 CH Utrecht, The Netherlands
a.m.j.j.bonvin@uu.nl

Dr. Lucia Diaz
Barcelona Supercomputing Center
Jordi Girona 29
08034 Barcelona, Spain
lucia.diaz@bsc.es

Prof. Juan Fernandez-Recio
IBMB-CSIC / BSC
Jordi Girona 29
08034 Barcelona, Spain
juanf@bsc.es

Dr. Raphaël Guerois
CEA I2BC (Institute for Integrative Biology of the Cell)
Institut des Sciences du Vivant Frédéric Joliot
Bâtiment 144 - Point courrier 22
91191 Gif sur Yvette, France
guerois@cea.fr

Prof. Turkan Haliloğlu
Boğaziçi University
Bebek
34342 Istanbul, Turkey
halilogt@boun.edu.tr

Prof. Birte Höcker
Universität Bayreuth
Universitätsstr. 30 / NWIII
95447 Bayreuth, Germany
birte.hoecker@uni-bayreuth.de

Dr. Brian Jiménez-García
Barcelona Supercomputing Center & Utrecht University
Jordi Girona 29
08034 Barcelona, Spain
br.jimenezgarcia@gmail.com

Dr. Horst Lechner
Universität Bayreuth
Universitätsstraße 30
Bayreuth, Germany
horst.lechner@uni-bayreuth.de

Dr. Marc Lensink
CNRS UGSF University of Lille
Parc de la Haute Borne
50 Avenue de Halley
59658 Villeneuve d'Ascq, France
marc.lensink@univ-lille.fr

Dr. Sony Malhotra
ISMB, Birkbeck/UCL University of London
University of London
London, United Kingdom
s.malhotra@mail.cryst.bbk.ac.uk

Prof. Marc A. Marti-Renom
Structural Genomics Group, CNAG-CRG
c/ Baldiri Reixac, 4. PCB - Tower I, 10th floor
08028 Barcelona, Spain
martirenom@cnag.crg.eu

Dr. Iain Moal
EMBL European Bioinformatics Institute
Wellcome Genome Campus Hinxton
CB10 1SD Cambridge, United Kingdom
moal@ebi.ac.uk

Prof. Matthieu Montes
Conservatoire National des Arts et Métiers
Laboratoire GBA, CNAM, 292 rue saint Martin
75003 Paris, France
matthieu.montes@cnam.fr

Prof. Matteo Dal Peraro
EPFL, Station 19
Lausanne, Switzerland
matteo.dalperaro@epfl.ch

Dr. João Rodrigues
Department of Structural Biology, Stanford University
318 Campus Drive, S210
94305 Stanford, California, United States of America
joaor@stanford.edu

Prof. Maya Topf
Birkbeck College, Malet Street
WC1E 7HX London, United Kingdom
m.topf@mail.cryst.bbk.ac.uk

Prof. Shoshana Wodak
VIB-VUB Center for Structural Biology
Pleinlaan 2
1050 Brussels, Belgium
shoshana.wodak@gmail.com

Venue and Housing Information

Venue

Barcelona Supercomputing Center
c/ Jordi Girona, 29 08034 Barcelona, Spain.
Phone: (+34) 93 413 77 16 (Switchboard)
Fax: (+34) 93 413 77 21
http://www.bsc.es

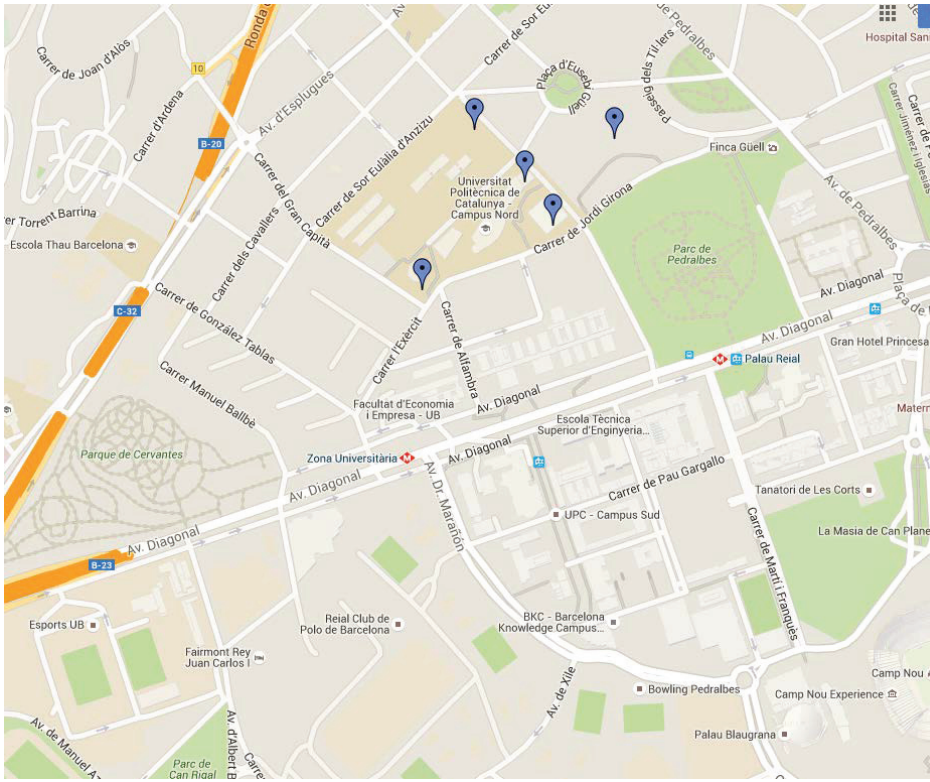
Accommodation

Students and instructors will be housed at Residencia Universitaria Torre Girona, a mere 5-minute walk from the site of the course. The Barcelona Supercomputing Center and MareNostrum supercomputer are located in a nice and quiet neighbourhood of the city, close to the University and the Metro station. The rooms have free internet connection, TV, telephone, en-suite bathroom, air-conditioning and a fully-equipped kitchen. Course participants will, in principle, share a room (2 persons per room).

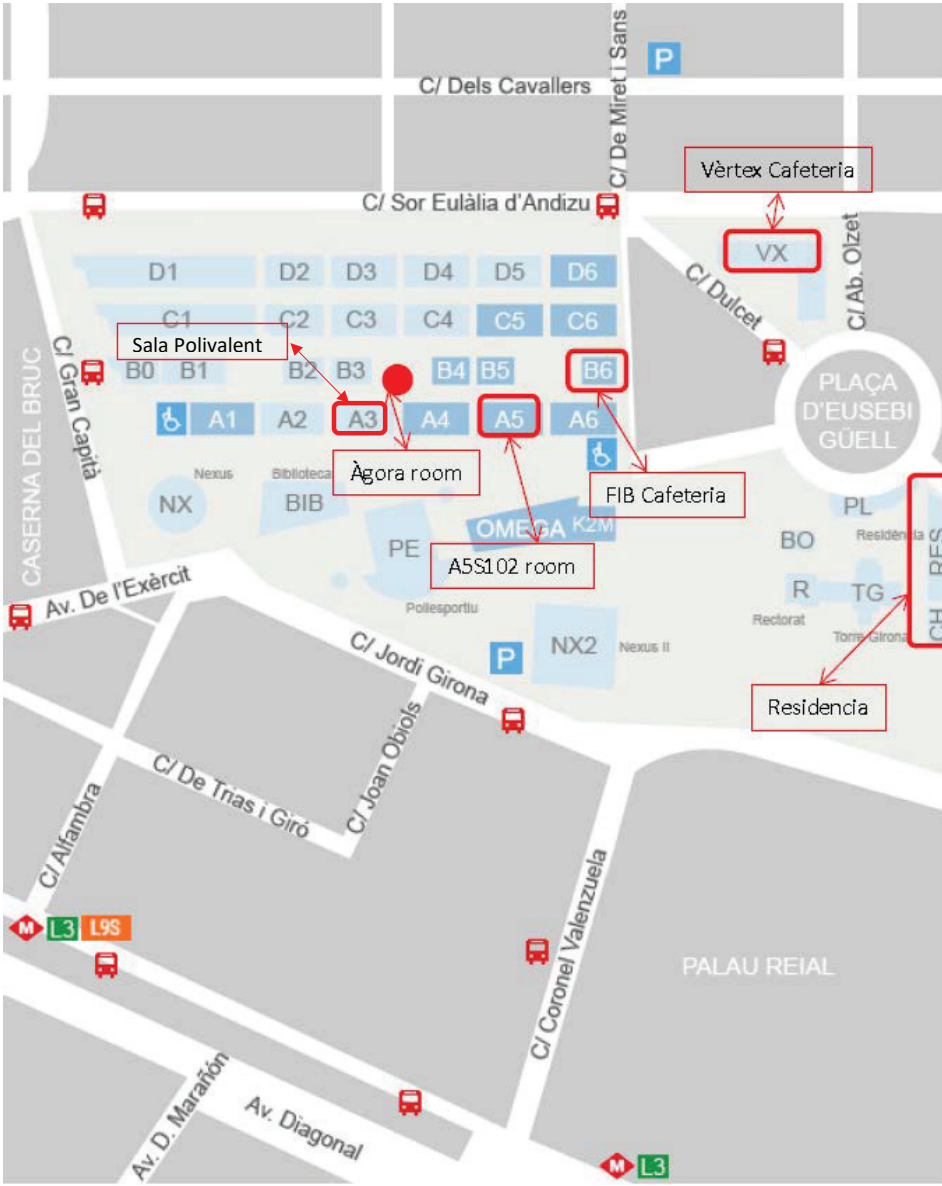
Torre Girona Residence Hall
Passeig dels Tillers, 19
08034 Barcelona
Telephone: (+34) 93 390 43 00
Fax: (+34) 93 205 69 10
E-mail: torregirona@resa.es
http://www.resa.es/eng/Residences/Torre-Girona

Transport

Details of how best to reach Barcelona Supercomputing Centre can be found here: http://www.bsc.es/about-bsc/contact-us/how-to-find-us



Map of UPC Campus Nord, showing lecture halls, cafeteria, and student residence



Location of Platja Ca la Nuri (Course Dinner on Thursday, July 5th 2018)

Address: Passeig Marítim de la Barceloneta, 55 08003 Barcelona
Phone: +34 932 21 37 75