UNIVERSITÀ DEGLI STUDI
DI MODENA E REGGIO EMILIA

BIGCHEM second Autumn School: Computer-Aided Drug Discovery

Wednesday October 25th

9:00 Alexandre Varnek, University of Strasbourg

Generative Topographic Mapping: an universal tool for chemical data visualization and modelling

In this talk I'll give some information about Generative Topographic Mapping and its application to Big Data visualization, library comparison, virtual screening and de novo design.

9:45 Martin Vogt, University of Bonn

Machine learning concepts in chemoinformatics

Applying machine learning methods for chemoinformatic problems requires the consideration of a number of different aspects ranging from data selection, preparation, and representation over method and parameter selection/optimization to the choice of metrics for performance evaluation. The lecture will discuss these aspects and how they are applied in practice.

10:30 Coffee break

11:00 Chris De Graaf, VU University Amsterdam

From structural protein-ligand interaction maps to medicinal chemical modulators

In the lecture I will discuss how the integration of structural interaction fingerprints in chemogenomics workflows can be used for the identification of function and protein-specific interaction hotspots to guide structure-based virtual screening and the construction of structure-based medicinal chemistry toolkits for ligand design. Computer-aided drug discovery workflows will be presented that combine structural cheminformatics tools and databases (https://3d-e-chem.github.io). I will show how breakthroughs in structural biology can be complemented by computational and experimental studies for a more accurate description and prediction of structural determinants of protein-protein interactions and ligand binding kinetics.

11:45 Christof Schwab (Molecular Networks GmbH)

Molecular structure representation in chemoinformatics applications

The power of any in silico tool for chemistry, biology or toxicity prediction depends on the accuracy of the representation of molecular structures, which are closely related to their chemical, physical and biological properties. This requires a robust, deterministic and reliable chemoinformatics platform to process, manipulate and store 2D and 3D chemical structures to be correctly interpreted and handled for computing molecular descriptors used in the predictions. Some of the basic concepts in the CORINA/MOSES chemoinformatics platform will be described including examples.

Thursday October 26th

9:00 Esther Kellenberger, University of Strasbourg

New insights on the conservation of fragment binding mode

In a short period of time, fragment-based approach has encountered plenty of successes and has become a new standard in drug discovery. We have questioned common beliefs on fragment binding by exploring the crystal structures of protein-fragment complexes in the Protein DataBank:

Do fragments preserve their binding mode during growing? We showed that the binding modes of fragments and their drug-like superstructures binding to the same protein are mostly conserved. Moreover, rigid fragments binding to a rigid protein pocket generally show a unique binding mode.

Do fragments have preferred sites in druggable pockets? We demonstrated that provided enough fragments have been co-crystallized with a protein pocket, good interaction coverage is achieved. Fragments do not cluster in specific spots.

Do promiscuous fragments have a conserved binding mode? We find that the majority of multi-target fragments show non-conserved binding modes, even if they bind in a similar conformation or to similar protein targets. Fragment chemical properties alone are not able to predict whether a fragment will exhibit a versatile or conserved binding mode.

9:45 Andrew Pannifer (University of Dundee)

The European Lead Factory: Collaborative Drug Discovery and Big Data.

The European Lead Factory (ELF) is a public-private consortium giving academic organisations and SMEs across Europe access to full industry-standard High Throughput Screening and hit follow-up capability. The screening library was created with diverse subsets contributed by each of the seven pharmaceutical companies in the consortium and is continually expanded through in-house synthesis. The consortium has developed a specialised triage workflow and developed software to balance the IP and scientific requirements of the diverse consortium members. The workflow and supporting triage tools will be described together with challenges encountered during the project and the ongoing retrospective data analysis.

10:30 Coffee break

11:00 Maxim Fedorov, Skolkovo Institute of Science and Technology

Computational methods for discovery and design of new bioactive molecules: at the interface between molecular theories and machine learning

The talk will overview recent developments and applications of new computational methods for discovery and design of new bioactive molecules (drugs, agrochemical, fragrancies etc). Modern automated systems for experimental screening and synthesis of new biomolecules provide new opportunities for development of compounds for medical, agricultural or food applications. However, while a vast number of bioactive molecules have already been found, these findings often result from trial and error rather than from a well-controlled design methodology. Another problem is the astronomically large number of potential candidates (estimates for the total number of molecules with potential bioactivity in 'chemical space' vary between 10^30 to 10^60). The use of high-performance computing and state-of-the-art computational methods for *rational* design of bioactive molecules targeted for a particular application can help to speed up the process of development of new compounds for practical applications.

The talk will discuss promising new hybrid approaches that combine molecular theory and molecular modelling methods with machine learning. Mathematical and numerical aspects of the new computational tools that would allow to accurately describe many molecular properties of biomolecules will be discussed. Development of such tools will lead to the predictive design of bioactive molecule with tailored functionality.

11:45 Ivan Kondratov, Enamine

Compound collections and their role in modern drug discovery

Most drug discovery processes start from screening of commercially available compounds. Further rapid and effective hit-to-lead optimization processes require easy availability of diverse building-blocks/reagents/capping agents from reliable suppliers.

In order to be competitive, suppliers have to learn and invent diverse advanced methods/approaches/techniques of compound synthesis in order to enrich their collections with novel and attractive compounds.

The presentation will demonstrate such an evolution of compound collections (mostly on the example of Enamine collections of Screening compound and Building Blocks) as well as further diverse perspectives of their application in drug discovery.

Friday October 26th

9:00 Oliver Koch, Technische Universität Dortmund

A medicinal chemistry guide to binding site comparison

The growing number of published protein structures and the vast amount of software to detect binding site similarities allow a fast and user-friendly comparison of thousands of protein pockets. The identification of binding site similarities can have an impact on binding site identification, function annotation, polypharmacology, off-target prediction, and drug repurposing. Nevertheless, the choice of a suitable tool, for example, to compare a binding site of interest to various known binding sites is still a difficult task. Based on an exhaustive evaluation it becomes obvious that the choice highly depends on the aim of the study. This presentation will focus on different applications and present a detailed guide to binding site comparison.

9:45 Andrea Cavalli, University of Bologna and Italian Istitute of Tecnology

Thermodynamics and kinetics of drug-target binding through molecular simulations

Drug-target binding represents the first event at the basis of the therapeutic action of drugs. This complex phenomenon needs to be properly described at an atomistic level to identify the major determinants of drug potency and in vivo drug efficacy. Molecular dynamics (MD) is emerging as a powerful tool for investigating protein-ligand binding, and is getting increasing consensus from the drug discovery community. While extensive MD simulations in the microsecond to the millisecond timescale are nowadays able to simulate protein-ligand binding "spontaneously", enhanced sampling methods, including metadynamics, steered-MD, umbrella sampling, etc., can improve the sampling of that part of free energy landscape that can be relevant for the biological process under investigation.

In this talk, I will be presenting the use of extensive MD simulations to investigate spontaneous protein-ligand binding. Then, I will show how free energy calculations allow the identification of the minimum free energy path from the bulk of the solvent into the protein-binding pocket, as well as the determination of thermodynamic and kinetic parameters associated to drug-target recognition and binding. The presentation will finally be focused on applications of enhanced sampling methods to accelerate ligand binding and unbinding and to estimate kinetics (kon and koff) and thermodynamics, in simulation timescale more compatible with the requirements of speed and accuracy of the pharmaceutical research. All these simulations will be discussed in the framework of drug design and discovery, highlighting the role of these approaches in real-life drug discovery endeavors.

10:30 Coffee break

11:00 Corrado Priami, The Microsoft Research – University of Trento Centre for Computational and Systems Biology (COSBI)

From multi-level data to mechanistic details of biological systems in systems pharmacology

Modeling biological systems emerging from the observation of many different (omics) data types is a current challenge of systems approaches (biology, medicine, pharmacology,

nutrition) that aim at fostering personalized diagnoses and interventions. The constant improvement of high-throughput technology is making complexity of models and analysis explode. At COSBI, we developed some pieces of technology that can help addressing these challenges. This talk will show how multi-level data can be pre-processed to obtain robust biomarkers with a rank-based novel algorithm. How data can be turned into pathways ranked by their level of activity according to a specific experimental condition. Then, selected pathways (possibly biomarker- related) can be semi-automatically mapped into graphical models suitable for stochastic simulation and in-silico what-if experiments.