FAKULTA
CHEMICKÉ TECHNOLOGIE
VŠCHT PRAHA

Studentská Vědecká Konference

2025

20. 11. 2025

SBORNÍK ANOTACÍ

(143)

ÚSTAVNÍ KOORDINÁTOR

Ing. Petr Čech, Ph.D.

SEZNAM SEKCÍ

1. Bioinformatika a chemická informatika

Bioinformatika a chemická informatika

MÍSTO: B1322

KOMISE

prof. Mgr. Daniel Svozil, Ph.D. (předseda)

Ing. Petr Čech, Ph.D.

Ing. Ivan Čmelo, Ph.D.

Mgr. Jan Pačes, Ph.D.

Ing. Martin Šícho, Ph.D.

PROGRAM

13:00 **zahájení**

13:05 **Marek Cuker** (M1, doc. Ing. Filip Lankaš, Ph.D.)
CYkAS – a tool for effective calculation of nucleic acid stiffness parameters from molecular dynamics simulations

13:20 **Ema Fialová** (M2, Mgr. Jan Pačes, Ph.D.)
Repetitive Elements in Avian Genomes

13:35 **Maria Fleshko** (B3, MSc. Wim Dehaen, Ph.D.)
Toxins of the Destroying Angel

13:50 **Adam Hanzlík** (M2, Ing. Milan Voršilák, Ph.D.)
A Consolidated SMARTS Filter Library for Nuisance Compound Handling

14:05 **Ester Machová** (M1, Ing. Petr Čech, Ph.D.)
Karnaugh map generator

14:20 **Dominika Soukupová** (M1, Ing. Martin Šícho Ph.D.)
Machine learning for Classification of Antimicrobial Peptides

14:35 **Jáchym Urban** (M1, Ing. Ivan Čmelo, Ph.D.)
Virtual screening for potential O-methyltransferase inhibitors

vyhlášení výsledků

SPONZOŘI - BIOINFORMATIKA A CHEMICKÁ INFORMATIKA

CYkAS – a tool for effective calculation of nucleic acid stiffness parameters from molecular dynamics simulations

Marek Cuker (M1)

školitel: doc. Ing. Filip Lankaš, Ph.D.

The mechanical properties of nucleic acid duplexes are key determinants of many cellular processes, such as chromatin remodelling and the regulation of gene expression. One property is bending, which plays an essential role in DNA–protein interactions as well as in DNA and RNA nanotechnology.

In this presentation, CYkAS, a program for analysing mechanical properties of nucleic acid duplexes from molecular dynamics (MD) trajectories will be introduced. CYkAS enables the effective calculation of stiffness constants, even for complex structures such as double cross-over (DX) motifs, key building blocks of DNA nanostructures. It also provides a novel method for calculating global bending angles, allowing for the comprehensive determination of global bending stiffness in nucleic acid duplexes.

Several applications of this approach will be demonstrated, including the study of length dependence of stiffness parameters, the characterization of bending anisotropy, and an analysis of polypurine sequences, which display notably distinct mechanical properties compared to random sequences.

Repetitive Elements in Avian Genomes

Ema Fialová (M2)

školitel: Mgr. Jan Pačes, Ph.D.

Avian genomes exhibit distinct characteristics, shaped by extensive adaptations during their evolution within the dinosaur lineage. For a long time, a number of genes were believed to be evolutionarily lost in avian genomes. However, recent studies suggest that many of these genes are not truly absent but rather located in regions that are technically difficult to analyze—such as microchromosomes. These regions are characterized by high GC content and a high density of repetitive elements, including gene stuttering. These patterns may contribute to genomic instability and their further analysis can provide insights into mechanisms of evolutionary change and selection. This thesis aims to identify and analyze repetitive features, including stutter-like patterns, within avian genomes. In this context, a stutter describes the multiplication of a portion of an exon into its adjacent intronic regions.

Toxins of the Destroying Angel

Maria Fleshko (B3)

školitel: MSc. Wim Dehaen, Ph.D.

Amanita virosa (Destroying Angel) is a highly toxic mushroom responsible for fatal poisonings upon ingestion. It contains amatoxins, phallotoxins, and virotoxins, which belong to the class of small cyclic or bicyclic peptides. Despite their chemical similarity, these toxins have very different receptors: amatoxins target a component of the largest subunit of RNA polymerase II, RPB1, while phallotoxins and virotoxins bind to filamentous actin. The aim of this work is to use computational methods (ligand-receptor docking, molecular dynamics, etc.) to investigate the molecular basis of the binding specificity and stability of these peptides.

For the analysis of interactions, protein structures from the PDB database, visualized using the PyMol and ChimeraX tools, will be utilized. Binding interactions will be evaluated using structure-based methods: molecular docking with Autodock Suite software, ADCP 1.1, and Alpha Fold 3. The Alpha Fold 3 platform will be used specifically for protein-ligand co-folding, i.e., blind docking without a crystal structure. To understand the stability of the complexes, molecular dynamics simulations (OpenMM) will be performed. The RDKit cheminformatics package will be used to analyze the thermostability and conformational constraints of the toxins.

A Consolidated SMARTS Filter Library for Nuisance Compound Handling

Adam Hanzlík (M2)

školitel: Ing. Milan Voršilák, Ph.D.

SMARTS substructure filters are useful for flagging nuisance compounds thanks to their smooth PostgreSQL/RDKit integration and portability. However, they also come with unintuitive pitfalls, most notably in RDKit, where implicit vs. explicit hydrogens can cause silent mismatches that look like genuine clean results. These issues appear even in well-known public databases such as those from ChEMBL.

To address this, we combined about 2,800 filters into a non-redundant set of ~1,800 and documented the technical edge cases most likely to mislead users. We also built an experimental workflow that clusters filters by their matching profiles and standardizes their names. Naming is aided by few-shot prompting with LLMs, using small-MW PubChem examples and similarity to established filters.

The resulting library integrates directly into database pipelines and routine compound annotation. It also supports ML workflows, where filter hits act as lightweight structural fingerprints. Overall, it strengthens cheminformatics analyses and reduces the risk of silent SMARTS-matching failures in planned deployments at CZ-OPENSCREEN.

Karnaugh map generator

Ester Machová (M1)

školitel: Ing. Petr Čech, Ph.D.

The Karnaugh map is one of the methods that are used to minimize logic functions. It works by mapping an n-dimensional table of logical variables and states into a two-dimensional map from which the minimized logic function can be expressed. The method is particularly applicable in the design of more advanced logic circuits, where the number of logic operations needs to be determined so that the smallest possible number of logic operations are used. The aim of this work is to design and develop an application that will generate a Karnaugh map from a given logic function, express the minimized function and determine the number of logic operations required to solve it.

Machine learning for Classification of Antimicrobial Peptides

Dominika Soukupová (M1)

školitel: Ing. Martin Šícho, Ph.D.

As antimicrobial resistance continues to be a pressing problem, the search for new antimicrobial agents is needed. A possible tool to deal with antimicrobial resistance is antimicrobial peptides. The advantage of these substances is that they physically disrupt bacterial cell membranes. Due to this mechanism, it is harder to evolve resistance.

In this work, machine learning techniques are used for the prediction of antimicrobial peptides. From the trained models, three predictions can be made:

- whether the peptide is antimicrobial
- whether the antimicrobial peptide is stable
- whether the antimicrobial peptide is active against Staphylococcus aureus

These models were evaluated with accuracies of 0.89, 0.84, and 0.76, respectively. To find out which peptides are likely to be active against Staphylococcus aureus, the AMPSphere database screening was performed. The database contains over 860 thousand potential AMPs and it was annotated with the trained models, and for each sequence, a joint decision was made, whether the peptide has a high probability of being active and stable. The expected accuracy of this activity prediction is 0.84 and the recall 0.78.

Virtual screening for potential O-methyltransferase inhibitors

Jáchym Urban (M1)

školitel: Ing. Ivan Čmelo, Ph.D.

Virtual screening is a set of computational methodologies often used to discover novel drug candidates within the vast chemical space of all accessible chemical compounds. In practice, this usually involves screening extensive virtual libraries of possible compounds, from catalogues provided by vendors or computationally generated libraries, against a biological target of interest.

This work involves structure-based virtual screening of O-methyltransferase from the pathogenic bacterium Francisella tularensis, the cause of tularemia in humans and animals both. The target protein was isolated and structurally resolved by the research group led by Evžen Bouřa. Methyltransferases are a compelling pharmacological targets, as they play a pivotal enzymatic role in a number of biological processes, including epigenetic regulation and metabolic control. Affecting these would affect the bacterium as a whole.

The developed virtual screening pipeline hinges on molecular docking, a two-step process comprising pose prediction and scoring. In this study, three docking approaches are compared and combined: AutoDock Vina (an open-source standard), Gnina (Vina enhanced with deep convolutional neural network scoring), and Protenix-Dock (an approach based on a combination of empirical knowledge and force fields). The study employs a consensus docking strategy with the objective of optimizing the reliability of virtual screening outcomes. The most promising candidate structures will be manually evaluated for viability and novelty and potentially purchased for in vitro testing.