**MEGA BIOINFORMATICS INTERNSHIP’24**

**Bversity School of Biological Sciences**

**26 Nov 2024 – 30 Nov 2024**

**"Comprehensive 2D and 3D Structural Analysis of Anti-Protozoan Therapeutics Targeting African Trypanosomiasis: A Case Study on Pentamidine Using the KNIME Analytics Platform"**

**Introduction:**

Drug repurposing for pentamidine involves finding new therapeutic applications beyond its original indication, typically as an antimicrobial agent for diseases like leishmaniasis and Pneumocystis pneumonia. The rationale for repurposing pentamidine—and the importance of understanding the process—stems from several scientific, medical, and economic factors. A thorough understanding of pentamidine's molecular mechanism of action enables researchers to hypothesize new uses based on its biological activity. Computational techniques like molecular docking and network pharmacology help predict interactions with other disease targets.

Pentamidine interacts with multiple cellular targets, including nucleic acids, enzymes, and membranes. This opens avenues for it to be effective against conditions like cancer, diabetes, or neurodegenerative diseases. Its ability to modulate pathways like DNA repair and apoptosis (programmed cell death) is particularly relevant for cancer therapeutics. Traditional drug discovery is a lengthy and costly process, often taking 10–15 years. Repurposing pentamidine skips the early discovery phases (e.g., pharmacokinetics and toxicology studies) since these are already known. Pathogens are becoming increasingly resistant to available drugs. Repurposing pentamidine could offer alternative therapies to combat drug-resistant strains. Screening pentamidine against large panels of cell lines or disease models allows for identifying new therapeutic windows. Genomic, transcriptomic, or proteomic analyses can identify patient subgroups who might benefit most from pentamidine.

Repurposing pentamidine aligns with personalized medicine approaches, as its use can be tailored to individuals with specific genetic or molecular disease profiles. For instance, patients with cancers characterized by vulnerabilities in DNA repair pathways might benefit from pentamidine’s mechanism. Drug repurposing for pentamidine represents an innovative, resource-efficient way to extend its therapeutic impact while addressing emerging challenges in medicine. This project deals with analysing 2D and 3D structures of anti – protozoan therapeutics targeting African trypanosomiasis for repurposing of a drug, pentamidine.

**Workflow:**

**Disease & Drug selection:**

**African Trypanosomiasis**, also known as **sleeping sickness**, is a vector-borne parasitic disease caused by the Trypanosoma species, primarily Trypanosoma brucei. The disease is transmitted to humans through the bite of an infected tsetse fly (Glossina species). Drugs such as **pentamidine, suramin,** and **eflornithine** are used in the early stages, while more advanced stages may require drugs like **melarsoprol** or **nifurtimox-eflornithine combination therapy (NECT). Pentamidine** is effective against T. brucei gambiense in the early stages but has limited effectiveness against T. brucei rhodesiense. Existing drugs do not always work against all stages of the disease, especially the late stages of infection, or they are ineffective in certain geographic regions where the parasite may have developed resistance. Advances in computational drug repurposing, including high-throughput screening, artificial intelligence, and bioinformatics, allow for the identification of drugs that can target Trypanosoma parasites effectively. These technologies can mine existing drug libraries to find candidates that may have anti-parasitic activity. Repurposing existing drugs for African trypanosomiasis is essential to address the limitations of current treatments, including toxicity, resistance, and access barriers.

Pathogens are becoming increasingly resistant to available drugs. Repurposing pentamidine could offer alternative therapies to combat drug-resistant strains. Traditional drug discovery is a lengthy and costly process, often taking 10–15 years. Repurposing pentamidine skips the early discovery phases (e.g., pharmacokinetics and toxicology studies) since these are already known.

**Data collection:**

Initially, 2D and 3D structures of the chosen drug compound were downloaded from pubchem in the structural data file format. This structural file format is used because it is easier for creating model with the algorithm.

**A screenshot of a computer

Description automatically generated**

This pentamidine drug structure similarity is compared with library of anti – protozoal compounds available in database.

A screenshot of a computer

Description automatically generated

**2D structural analysis**

A screenshot of a computer

Description automatically generated

1. From node repository, there are certain nodes which plays vital role in 2D structure similarity analysis such as SDF reader, RDKit from molecule, RDKit Fingerprint, Fingerprint similarity and HeatMap(JFreeChart).
2. The downloaded 2D structure of the target compound and library compounds are uploaded in the SDF Reader node which reads the file in SDF format.
3. After execution of SDF Reader node, we have used RDKit From Molecule node which converts file into a readable format or in a fingerprint format.
4. After execution of RDKit From Molecule node, the data are inputed to RDKit Fingerprint node and executed.
5. The comparison is done with the help of fingerprint similarity node where the structures are compared with the help of tanimoto coefficient.
6. The results of structure similarity analysis is done and it is viewed with the creation of heat map created with the help of HeatMap(JFreeChart).

**3D structural analysis**

A diagram of a computer program

Description automatically generated

1. From node repository, there are certain nodes which plays vital role in 3D structure similarity analysis such as SDF reader, 3D Coordinates, RDKit Open 3D Alignment, and HeatMap(JFreeChart).
2. The downloaded 3D structure of the target compound and library compounds are uploaded in the SDF Reader node which reads the file in SDF format.
3. After execution of SDF Reader node, we have used 3D Coordinates node for structure optimization.
4. After execution of 3D Coordinates node, the data are inputed into RDKit Open 3D Alignment node and executed.
5. The comparison is done with the help of RDKit Open 3D Alignment node where the structure similarity are compared with the help of root mean square deviation and score value.
6. The results of structure similarity analysis is done and it is viewed with the creation of heat map created with the help of HeatMap(JFreeChart).

**Results and Discussion:**

A heatmap of Tanimoto coefficients for 2D structure analysis typically visualize the pairwise similarity between chemical structures based on their molecular fingerprints. The Tanimoto coefficient ranges from **0** to **1**. **0**: No similarity between two structures. **1**: Perfect similarity (identical structures). Values closer to 1 indicate higher similarity, while values near 0 suggest low or no similarity.

A screenshot of a computer

Description automatically generated

RMSD quantifies the atomic positional differences between two 3D structures (e.g., ligand poses, protein conformations). Lower RMSD indicates higher structural similarity (closer alignment to a reference structure). Often used to assess pose stability or deviation during molecular dynamics or docking. Score values (e.g., docking scores, binding energies) quantify the stability or favorability of a molecular pose, typically representing: binding affinity and interaction energy between molecules (e.g., ligand and protein). Lower score values (more negative) indicate better binding or interaction stability.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**Conclusion:**

KNIME structural analysis platform provides boarder and deeper understanding of molecular docking by creating machine learning algorithms. This study provides a comprehensive analysis of anti-protozoan therapeutics targeting African trypanosomiasis, with a specific focus on Pentamidine. Using the KNIME Analytics Platform, both 2D and 3D structural characteristics were evaluated to assess molecular similarities, docking efficiencies, and potential binding conformations. The 2D structural analysis, based on Tanimoto similarity coefficients, revealed key structural motifs shared among anti-protozoan drugs, highlighting Pentamidine's distinct pharmacophore features within the therapeutic class. This insight supports its continued relevance as a cornerstone treatment for African trypanosomiasis while suggesting avenues for structural optimization in related analogs.

The 3D structural analysis, including RMSD and docking score evaluations, demonstrated Pentamidine's strong binding affinity and stable conformational alignment with target proteins associated with the disease. Low RMSD values coupled with favourable docking scores underscored its optimal interaction potential and validated its therapeutic efficacy. The integration of 2D and 3D analyses underscores the value of combining molecular similarity assessments with spatial and energetic evaluations to gain a holistic understanding of drug behavior. The use of KNIME streamlined the analysis pipeline, showcasing its utility for large-scale, reproducible structural studies.