REPORT: DRUG SIMILARITY ANALYSIS FOR METHICILLIN-RESISTANT STAPHYLOCOCCUS AUREUS

PIYUSH MESHRAM, NOVEMBER 30 2024

Problem Statement : We have a company that has to produce a new drug with similar set of action to an existing set of drugs in market for a particular disease. This new drug would be cost efficient than the pre-existing drug.

Disease : MRSA stands for Methicillin-resistant Staphylococcus aureus. It's a type of bacteria that's resistant to many antibiotics, making it difficult to treat. MRSA infections can range from mild skin infections to serious, life-threatening illnesses.

Here are some key points about MRSA:

Causes: MRSA is a type of staph bacteria that has developed resistance to many antibiotics.

Types: There are two main types of MRSA infections:

1. **Healthcare-associated MRSA (HA-MRSA):** This type is more common in healthcare settings like hospitals and nursing homes.
2. **Community-associated MRSA (CA-MRSA):** This type can occur in healthy people who haven't been in a healthcare setting.

Symptoms: Symptoms of MRSA infection can vary depending on the type of infection. Common symptoms include:

Skin infections: Red, swollen, painful bumps or boils

Serious infections: Fever, chills, fatigue, rapid heart rate

Here are some of the **drugs that are commonly used to treat MRSA infections:**

Vancomycin: This is one of the most commonly used antibiotics for MRSA infections. It is given intravenously and can be effective against many different types of MRSA. However, it can have some serious side effects, such as kidney damage and hearing loss.

Linezolid: This is another antibiotic that is effective against MRSA. It is given orally or intravenously and can be used to treat a variety of MRSA infections. However, it can also have some serious side effects, such as nerve damage and low blood cell counts.

Ceftaroline: This is a newer antibiotic that is effective against MRSA. It is given intravenously and can be used to treat a variety of MRSA infections. However, it can also have some serious side effects, such as seizures and allergic reactions.

Daptomycin: This is a newer antibiotic that is effective against MRSA. It is given intravenously and can be used to treat a variety of MRSA infections. However, it can also have some serious side effects, such as muscle pain and weakness.

The Drug I chose to work upon is **Linezolid**

Linezolid is a synthetic antibiotic that belongs to the **oxazolidinone class**. It works by inhibiting bacterial protein synthesis. Specifically, it binds to the 23S rRNA of the 50S subunit of the bacterial

ribosome, preventing the formation of the 70S initiation complex. This action effectively halts protein production and bacterial growth. This unique mechanism of action, coupled with its ability to penetrate tissues effectively, makes linezolid a potent choice for treating various MRSA infections.

- **KNIME – Workflow and Results :**

Input data:

- PubChem database integration to retrieve 2D and 3D conformers of Linezolid and library compounds [**Protein Synthesis inhibitors].**
- Linezolid selected as the reference molecule.

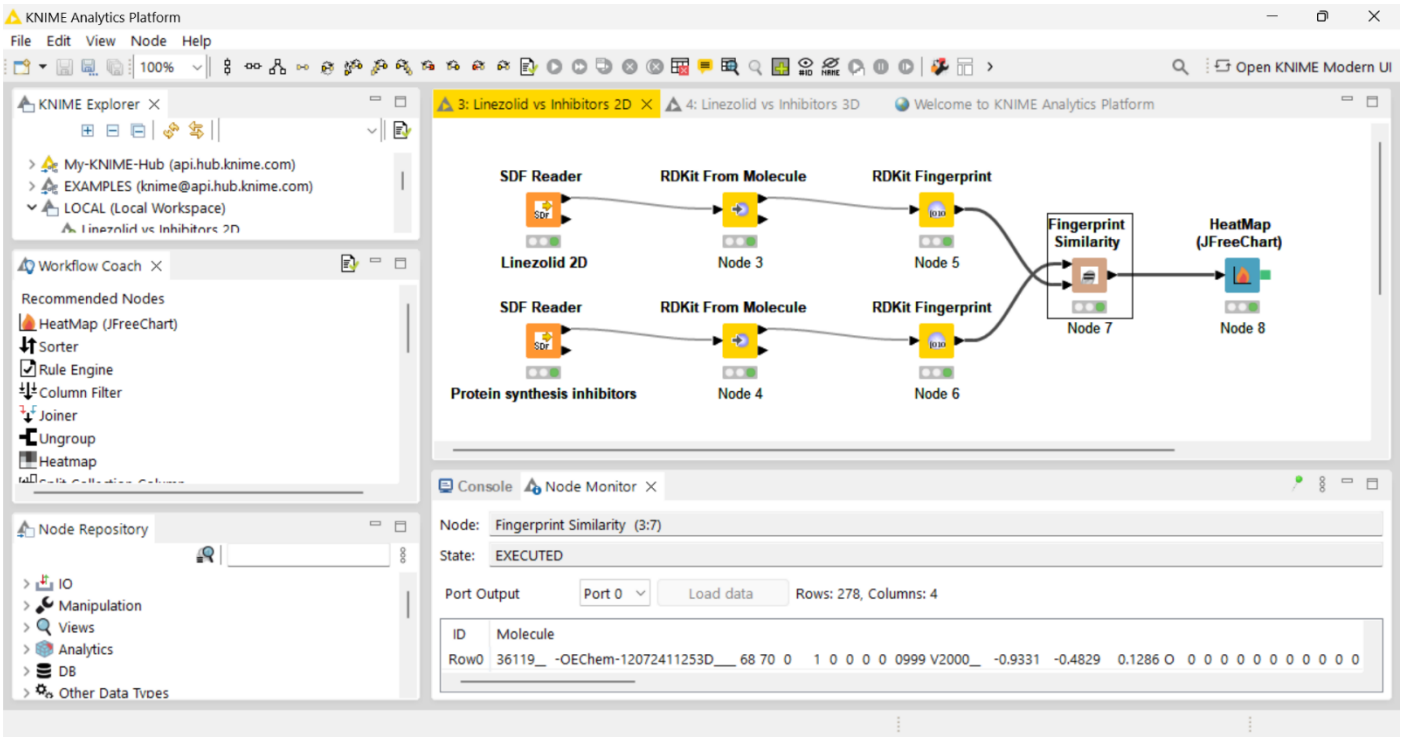
Data pre-processing:

- Molecules were converted to standard formats (e.g., SMILES or SDF) for compatibility.
- Structural normalization was applied via raw data

Feature Extraction:

2D and 3D molecular descriptors were computed for all compounds in the library using KNIME nodes.

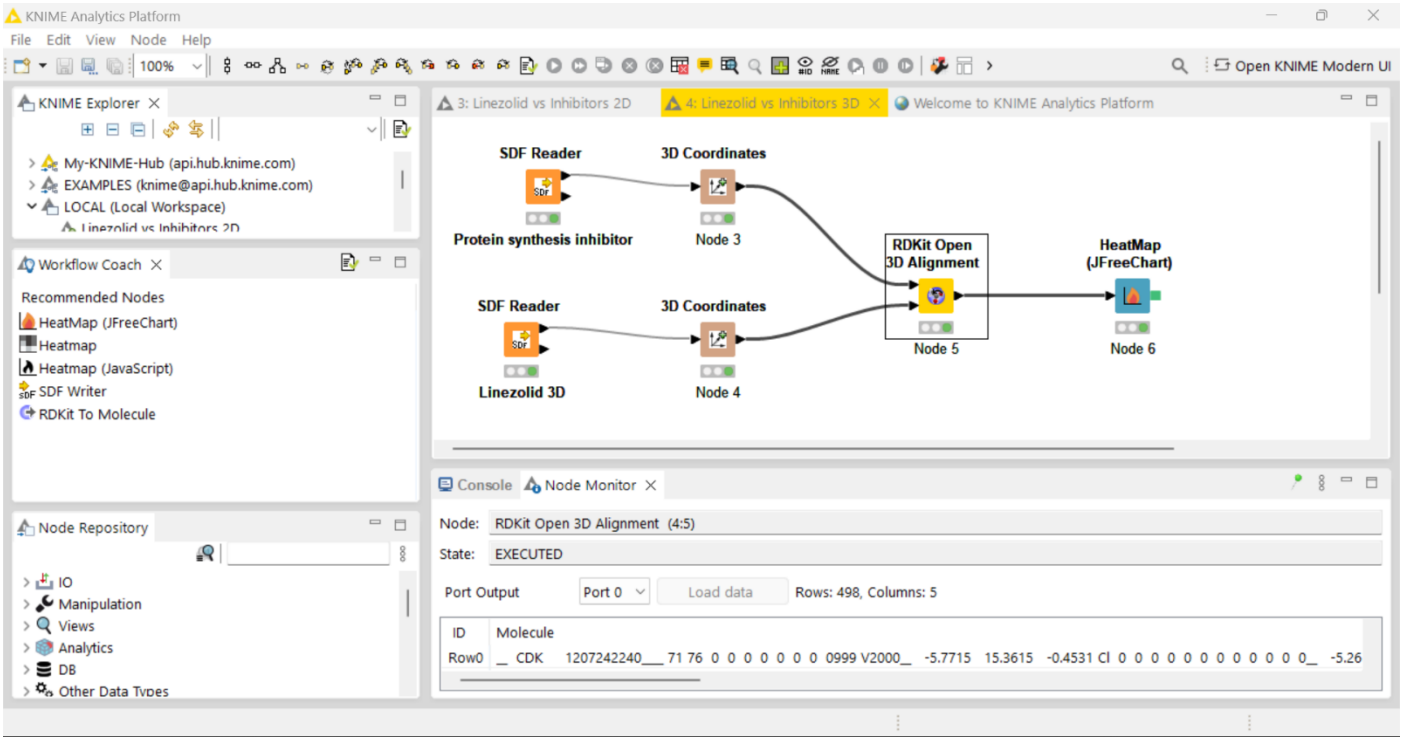
Workflow : Linezolid vs Protein synthesis inhibitors in 2D :



Nodes Used for 2D Workflow :

1. **SDF Readers**
2. **RDKit from Molecule**
3. **RDKit Fingerprint**
4. **Fingerprint Similarity**
5. **Heatmap (JFreeChart)**

Workflow : Linezolid vs Protein synthesis inhibitors in 3D :



Nodes Used for 3D Workflow :

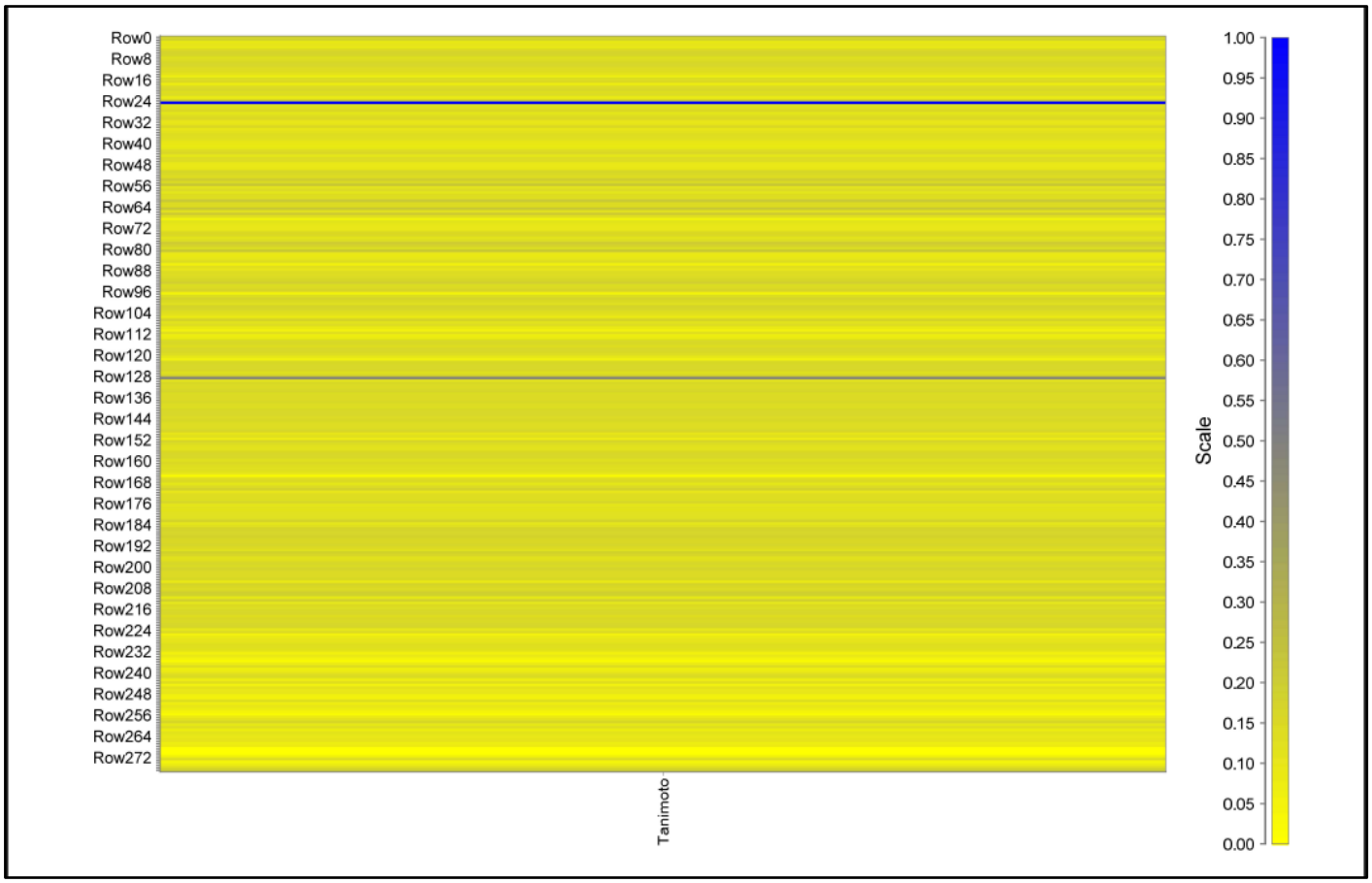
1. **SDF Readers**
2. **3D Coordinates**
3. **RDKit Open 3D Alignment**
4. **Heatmap (JFreeChart)**

Similarity Analysis:

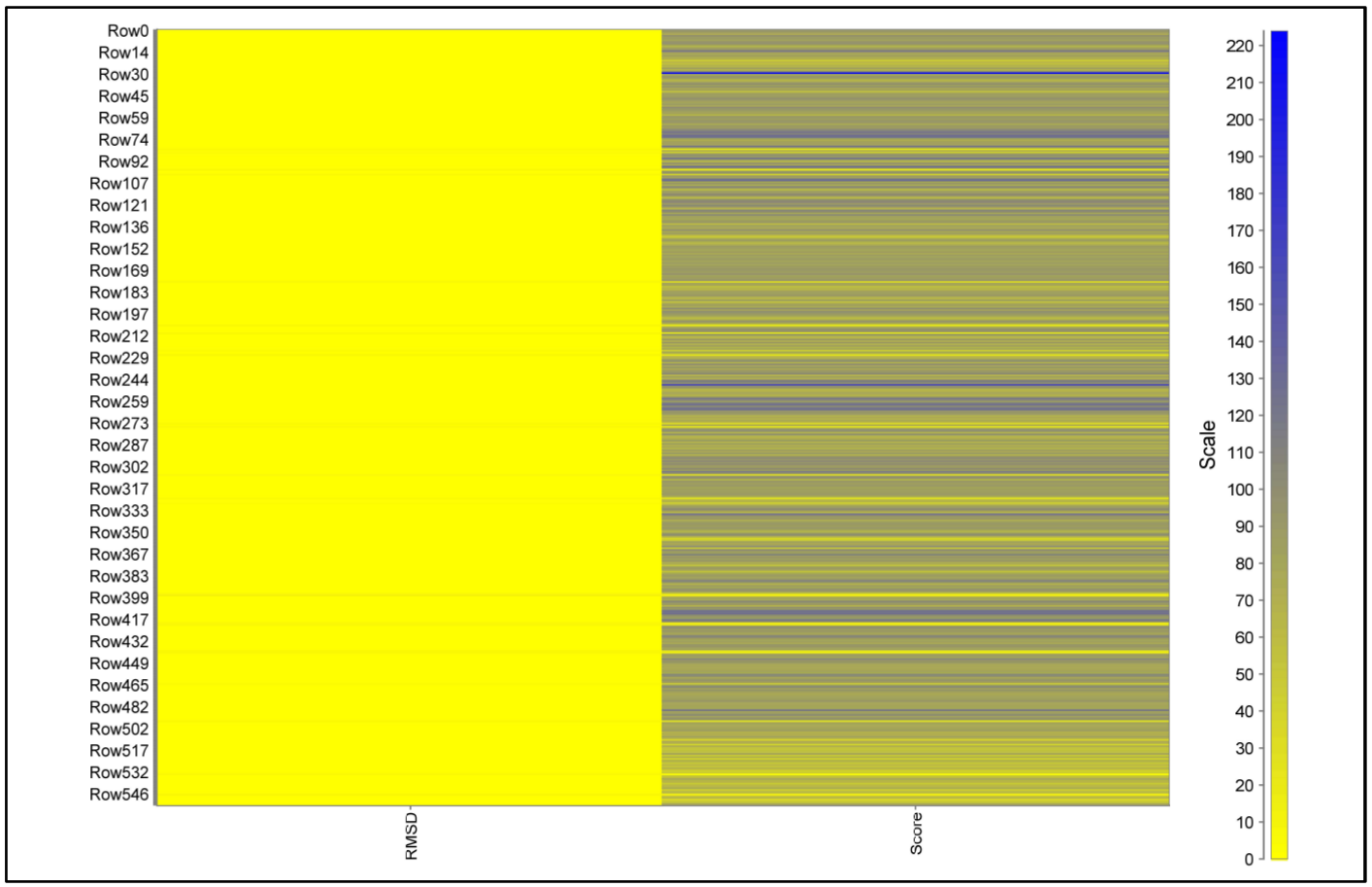
Structural and chemical similarity scores were calculated using **Tanimoto Coefficients for 2D similarity and RMSD for 3D alignment.**

Visualization: Results were visualized using scatter plots and heatmaps for better interpretability of similarity relationships.

2D HeatMap :



3D HeatMap :



In the heatmap, yellow represents lower values and blue indicates higher values, as indicated by the scale bar on the right-hand side: **Yellow: low values. Blue: High values.**

Rows: Represent individual molecules from the dataset. Each row corresponds to one **Protein Synthesis inhibitors** compound.

Column:

RMSD: shows the Root Mean Square Deviation (RMSD) values for the alignment of each **Protein Synthesis inhibitor** to Linezolid.

Score: Represents the alignment or similarity scores

- **Yellow:** Indicates lower scores or RMSD values.
- **Blue:** Indicates higher scores or RMSD values.

Filtering:

Compounds with similarity scores above a defined threshold (e.g., **Tanimoto – 0.92**) were shortlisted as like Linezolid.

▪ Conclusion :

2D workflow: Focuses on chemical similarity based on molecular feature (fingerprints). It is faster and can quickly suggest molecules with similar chemical structures but doesn't consider the 3D orientation.

3D workflow: Focuses on structural similarity based on how molecule align in 3D space. It helps you assess whether molecules might have a similar binding conformation or spatial arrangement to linezolid.

So, Molecules with high similarity in both 2D and 3D workflows are t most promising candidates for further research, as they are both chemically and structurally similar to Linezolid.

With all due regards -

Thanking you

Yours sincerely

Piyush Meshram

@Bversity School of Bioscience

@Bversity+ Community