JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

IN-SILICO DRUG DESIGNING AND NETWORK PHARMACOLOGY ANALYSIS OF NEUROBLASTOMA BY MACHINE LEARNING

Mansi Jangir[1] Uma Kumari[1]

[1]*Project Trainee at Bioinformatics Project and Research Insitute, Noida - 201301, India*

[1]*Senior Bioinformatics Scientist, Bioinformatics Project and Research Insitute, Noida - 201301, India*

***Abstract*:**Neuroblastoma is a highly lethal cancer that originates from immature nerve tissue. PHOX2B is a transcription factor that has been associated with neuroblastoma development. In this work, we used in silicobioinformatic tools to investigate the interactions of the PHOX2B peptide with the protein HLA-A*24:02, Proteomic human sample neuroblastoma, complexed with a peptide derived from neuroblastoma. The salient features of our approach included the integration of molecular docking, ADMET analysis, toxicity prediction, and machine learning techniques in studying the potential drug candidates and their binding affinities with the PHOX2B peptide. Therein, a machine learning model with an SVR algorithm gave an MSE of 1.72, indicating a strong and reliable prediction concerning estimation of ΔG based on interface area and hydrogen bonding. The result of this investigation provides new insight into drug design strategies against neuroblastoma by interacting with PHOX2B peptides.

Keywords: Neuroblastoma, PHOX2B peptide, molecular docking, in-silico drug design, machine learning, SVR, PDB 7MJA, network pharmacology

Introduction

Neuroblastoma is an aggressive neoplasm of the sympathetic nervous system and accounts for approximately 15% of all pediatric cancer-related deaths. There is, however, still a poor prognosis for children with high-risk neuroblastoma, despite the considerable improvement in current treatment modalities such as surgery, chemotherapy, and radiation. PHOX2B is a transcription factor critical during the development of the autonomic nervous system and has been considered an important actor in neuroblastoma pathogenesis. Mutations in the gene PHOX2B are contributory to hereditary and sporadic forms of the disease, raising interest in the use of targeted therapies aimed at modulating its function. In this work, we focus on the complex of the HLA-A24:02 protein bound with a PHOX2B peptide from neuroblastoma, PDB ID 7MJA. The HLA-A24:02 molecule belongs to Major Histocompatibility Complex Class I proteins, which are important in immune response by presenting endogenous peptide antigens cytotoxic T-cells. The interaction is 'central' to immune surveillance, enabling the immune system to detect and destroy the cells hosting an infection or that have become cancerous. Another structure of interest, the crystal structure of 7MJA, represents HLA-A*24:02 in complex with a peptide derived from PHOX2B. It is especially relevant in attempts to understand potential immunotherapeutic targets in neuroblastoma [1, 2].

Expression of HLA-A24:02 is very high across human populations, and this allele is implicated in several immune-mediated diseases. The binding of PHOX2B-derived peptides to this MHC Class I molecule may indicate possible ways to design immunotherapies capable of enhancing the capability of the immune system in recognizing and destroying neuroblastoma cells. Crystallographic studies of 7MJA showed important features about the peptide-binding groove of HLA-A24:02, which accommodates the PHOX2B peptide through a network of hydrogen bonds, electrostatic interactions, and van der Waals contacts. It shows that this tight binding is important for stability of the peptide-MHC complex which, in turn, defines its capacity to interact with T-cell receptors. Considering the crucial role of HLA-A24:02 in antigen presentation, intervention targeting peptide binding interactions in 7MJA thus offers a novel strategy for therapeutic intervention. Knowledge of the structure of 7MJA will either disrupt or stabilize the interaction between HLA-A24:02 and its PHOX2B-derived peptide, offering a way to modulate immune responses to neuroblastoma through enhancement of the capability of the immune system in cancer detection and/or offering an inhibitor that blocks harmful immune interactions. For this reason, structural analysis of 7MJA is not only relevant for understanding the progression of neuroblastoma but also a very attractive target for drug design [1, 3, 4, 5].

In the present study, molecular docking analysis with Cb-dock was applied to identify key interactions regarding drug candidates and the HLA-A*24:02/PHOX2B complex. We used Cb-dock in order to predict small molecules' binding affinities toward different cavities of 7MJA in the attempt to find drug-like molecules possibly interfering with or enhancing such critical interactions. These interactions could be visualized with Chimera and PyMOL-visualization tools that give an in-depth view of the binding interfaces of the identified drug

candidates. Along with molecular docking, we had performed pharmacokinetic analyses using SwissADME with a view to assessing the ADME properties of the drug candidates, which is highly important in its efficacy and bioavailability. Additionally, PRO-TOX was applied to predict the toxicity profiles of the compounds in order to grade their potential side effects-thus their safety profile-within acceptable limits [6, 7]. Apart from these traditional in-silico approaches, machine learning has also been integrated into the further refinement of the binding free energy prediction. Using the data extracted from PDBePISA, a support vector regression model was trained that predicts the binding free energy ΔG, using a variety of structure parameters like interface area and hydrogen bonds, among others. Its performance showed a Mean Squared Error of 1.72, which underlines its usefulness for the prediction of binding interactions of the 7MJA complex. In a nutshell, the 7MJA stands for one of the very important structural targets in the development of new treatments against neuroblastoma. This study describes the application of molecular docking, pharmacokinetic analysis, and toxicity prediction with machine learning as one systematic approach to identifying and validating drug candidates targeting the PHOX2B peptide-HLA-A*24:02 complex that presents innovative therapeutic opportunities against neuroblastoma [7, 8].

Materials and Methods

1. Molecular Docking using Cb-dock

To study the interaction between PHOX2B peptide and candidate drug molecules, we performed a molecular docking using the well-known web-based cavity prediction and docking tool, Cb-dock

- **Protein Data:** The three-dimensional structure of the HLA-A*24:02 in complex with the neuroblastoma-derived PHOX2B peptide was retrieved from the Protein Data Bank with its PDB ID of 7MJA.
- **Ligand Preparation:** Drug candidates were selected from pubchem based on their potential to bind with the PHOX2B peptide and further submitted to Cb-dock for docking simulations.

CB-Dock is an advanced molecular docking software designed to identify binding pockets on proteins and perform docking simulations with ligands. This tool integrates pocket detection with docking prediction, offering a user-friendly and effective platform for researchers to perform in-silico analysis of protein-ligand interactions [9,10].

2. Visualization by Chimera and PyMOL

Chimera and PyMOL were used to visualize the interactions between the PHOX2B peptide with the identified drug molecules, allowing for the detailed view of binding sites for further optimization of ligand structures.PyMOL can render proteins, nucleic acids, and other molecular structures in detailed 3D models.PyMOL supports dynamics simulations, allowing users to observe protein conformational changes and flexibility [11, 12].UCSF Chimera is a highly versatile molecular modeling software used for the visualization, analysis, and manipulation of molecular structures. It integrates powerful visualization tools with sophisticated molecular analysis functionalities, making it an all-encompassing tool for researchers in computational biology [13].

3. ADME and Toxicity Analysis Using SWISS-ADME and PRO-TOX

Prediction of the pharmacokinetics including ADME was done with SwissADME. This helps in theoretically estimating drug-likeness and bioavailability**.**SwissADME is a valuable tool in the early stages of drug discovery, helping researchers assess the pharmacokinetics, drug-likeness, and medicinal chemistry of compounds [14].The selected drug candidates were further screened by PRO-TOX to analyze their toxicity profile. The tool predicted parameters very useful in determining the safety profile of potential therapeutics, which included LD50, hepatotoxicity, cytotoxicity, and immunotoxicity.**ProTox-3.0** is an online platform designed to predict the **toxicity** of small molecules and drug candidates based on their chemical structure. It uses a machine learning approach to classify compounds according to their toxicological properties, aiding in the identification of potential risks associated with drug candidates in the early stages of drug discovery [15].

4. ΔG Prediction Machine Learning Model

In this work, we use a machine learning model to predict the binding free energy of the PHOX2B peptide-drug complex based on interface properties. To implement the machine learning model, Python language has been utilized. The in-built libraries of python like Scikit learn, pandas, MatPlotLib and BioPython have been used.BioPython is an open-source collection of Python tools for computational biology and bioinformatics. It is widely used to manipulate, analyze, and visualize biological data, providing modules for working with sequence data, structures, alignments, and more. Scikit-learn (commonly abbreviated as sklearn) is one of the most widely used open-source machine learning libraries in Python**.** It provides simple, efficient tools for data analysis and machine learning, offering a wide range of algorithms for classification, regression, clustering, and dimensionality reduction [16, 17].

An SVR was trained for: Area of Interface A^2, Number of Hydrogen Bonds (N_HB)

Interface data from the PDBePISA database for the protein 7MJA was used for training and testing the model, which resulted in a mean squared error of 1.72.

Results and Discussion

1. **Structural analysis of protein**

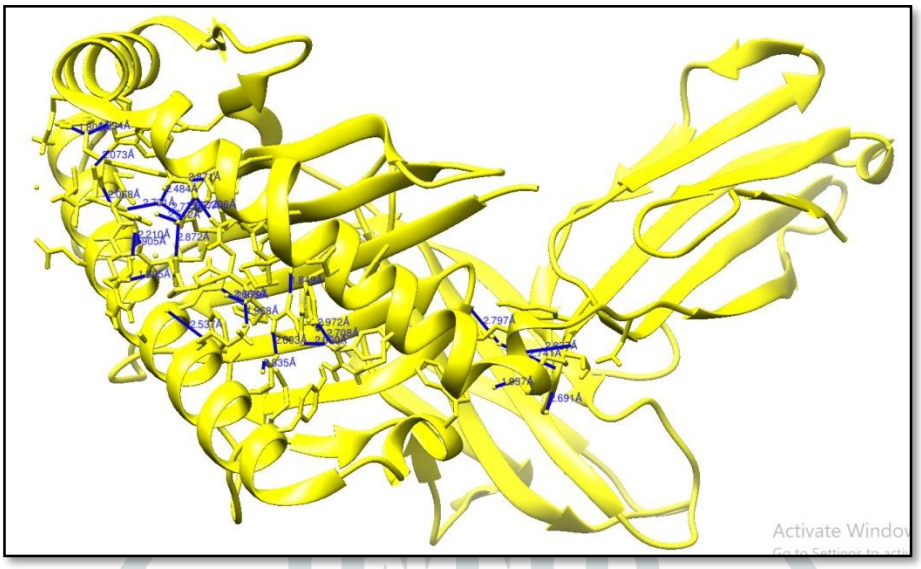

Figure 1: Visualization of H-bond using Chimera

The yellow structure presents the HLA-A24:02 protein. The blue lines and distances are hydrogen bonds between certain residues from the PHOX2B red peptide and the HLA-A24:02 molecule. The distance between hydrogen bond donors and acceptors is labeled in angstroms (Å), indicating whether a bond is weak or strong, respectively. Generally, hydrogen bonds between 1.5 Å and 3.5 Å are considered strong and important for the stability of molecules. Several hydrogen bonds have formed between the peptide and the protein, indicating a strong and stable interaction between the PHOX2B peptide and the HLA-A*24:02 protein. The presence of shorter hydrogen bonds in the range of about 1.84 Å to 2.8 Å indicates tight binding interactions important for retaining the peptide within the MHC class I molecule binding groove.

Table 1: Hydrogen Bond analysis using Chimera

Donor	Acceptor	Hydrogen	D..A dist (Å)	D-H..A dist (Å)
ARG 7.A NE	HOH 216.B O	ARG 7.A HE	2.983	2.177
SER 10.A OG	HOH 442.A O	SER 10.A HG	2.878	2.092
SER 12.A OG	HOH 442.A O	SER 12.A HG	2.806	2.029
SER 14.A OG	HOH 506.A O	SER 14.A HG	2.782	1.983
ARG 15.A N	HOH 485.A O	ARG 15.A H	2.778	1.935
GLY 17.A N	HOH 566.A O	GLY 17.A H	2.927	2.153
GLN 33.A N	HOH 544.A O	GLN 33.A H	2.870	2.015
GLN 33.A NE2	ASP 54.B OD2	GLN 33.A HE22	2.851	2.012
ARG 36.A NH1	ASP 54.B OD1	ARG 36.A HH11	2.883	2.065
ALA 42.A N	HOH 540.A O	ALA 42.A H	2.864	2.018
ARG 45.A NE	HOH 555.A O	ARG 45.A HE	2.900	2.093
LYS 67.A NZ	TYR 2.C O	LYS 67.A HZ1	2.779	1.968
HIS 71.A ND1	TYR 2.C OH	no hydrogen	2.668	N/A
THR 74.A OG1	ILE 5.C O	THR 74.A HG1	2.711	1.895
TYR 85.A OH	PHE 9.C O	TYR 85.A HH	2.625	1.805
SER 93.A N	HOH 587.A O	SER 93.A H	3.021	2.174
TYR 2.C N	GLU 64.A OE1	TYR 2.C H	2.815	2.003
ASN 3.C ND2	GLN 157.A OE1	ASN 3.C HD21	2.926	2.206
ARG 6.C N	HOH 106.C O	ARG 6.C H	2.977	2.210
PHE 9.C N	ASN 78.A OD1	PHE 9.C H	2.841	2.068
GLN 3.B N	HOH 255.B O	GLN 3.B H	2.993	2.125
GLN 9.B NE2	GLU 233.A O	GLN 9.B HE22	2.966	2.146
VAL 10.B N	HOH 256.B O	VAL 10.B H	3.012	2.158
TYR 11.B OH	PRO 236.A O	TYR 11.B HH	2.652	1.812
HIS 14.B ND1	HOH 205.B O	HIS 14.B HD1	2.605	1.766
ASN 25.B ND2	ALA 237.A O	ASN 25.B HD21	2.936	2.145
SER 29.B OG	HOH 233.B O	SER 29.B HG	2.831	2.072

HIS 32.B NE2	GLN 97.A OE1	HIS 32.B HE2	2.767	1.947
SER 34.B N	HOH 219.B O	SER 34.B H	2.850	1.995
SER 34.B OG	HOH 219.B O	SER 34.B HG	2.699	1.865
VAL 38.B N	HOH 250.B O	VAL 38.B H	2.923	2.099
LEU 40.B N	HOH 236.B O	LEU 40.B H	2.839	2.057
LYS 42.B NZ	HOH 239.B O	LYS 42.B HZ2	2.893	2.060
GLY 44.B N	HOH 215.B O	GLY 44.B H	2.687	1.870
HIS 52.B N	HOH 243.B O	HIS 52.B H	2.881	2.050
LEU 55.B N	HOH 254.B O	LEU 55.B H	2.960	2.160
LYS 59.B N	HOH 257.B O	LYS 59.B H	3.018	2.211
TRP 61.B NE1	ASP 123.A OD1	TRP 61.B HE1	2.797	1.976
THR 74.B N	HOH 211.B O	THR 74.B H	2.950	2.196
TYR 79.B N	HOH 245.B O	TYR 79.B H	2.884	2.032
TYR 79.B OH	HOH 202.B O	TYR 79.B HH	2.441	1.616
VAL 86.B N	HOH 222.B O	VAL 86.B H	2.900	2.132
ILE 93.B N	HOH 251.B O	ILE 93.B H	2.927	2.075
LYS 95.B N	HOH 246.B O	LYS 95.B H	2.896	2.042
ASP 97.B N	HOH 203.B O	ASP 97.B H	2.973	2.118
MET 190.A N	HOH 535.A O	MET 190.A H	2.857	2.028
THR 191.A OG1	HOH 241.B O	THR 191.A HG1	2.882	2.055
HIS 192.A N	HOH 580.A O	HIS 192.A H	2.996	2.164
THR 201.A OG1	HOH 521.A O	THR 201.A HG1	2.819	2.013
THR 215.A OG1	HOH 496.A O	THR 215.A HG1	2.778	1.994
LEU 216.A N	HOH 565.A O	LEU 216.A H	2.927	2.073
TRP 218.A N	HOH 415.A O	TRP 218.A H	2.902	2.044
GLY 222.A N	HOH 493.A O	GLY 222.A H	2.961	2.151
THR 226.A N	HOH 568.A O	THR 226.A H	2.951	2.108
LEU 231.A N	HOH 567.A O	LEU 231.A H	2.928	2.124

These hydrogen bonds constitute an important stabilization of the peptide within the groove of the MHC, and thus these interactions are one potential target for various drug design efforts that aim at modulating or disrupting this interaction. Hydrogen bonds play an important role in antigen presentation, anchoring the PHOX2B peptide within the HLA-A*24:02 binding cleft. These hydrogen bonds likely contribute to the high affinity and stability of the peptide-MHC complex, important for efficient recognition by the T-cell receptors during immune surveillance. Targeting these hydrogen bonds with small molecules thus has the potential to enhance or inhibit the interaction between the PHOX2B peptide and HLA-A*24:02, hence making this a viable therapeutic intervention strategy in neuroblastoma.

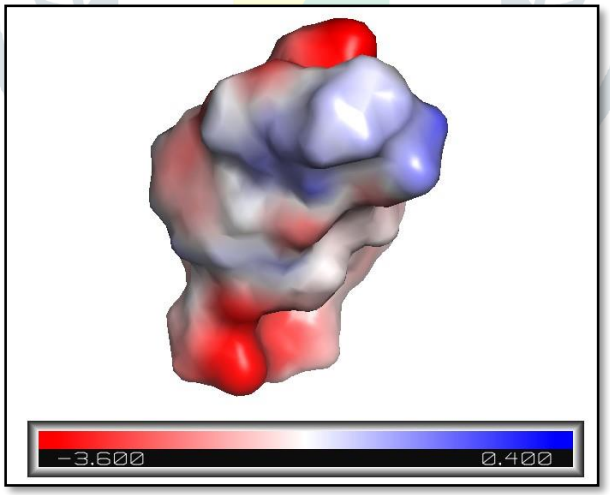

Figure 2: Visualization of electrostatic charge on the protein using PyMol

2. **Molecular Docking Results**

The molecular docking simulations revealed multiple binding sites on the PHOX2B peptide with significant binding affinities for the selected drug candidates. **Cb-dock** predicted favorable binding energies and identified key residues involved in drug interactions, such as **Tyr 74** and **Glu 63**, located in proximity to the PHOX2B binding cavity.

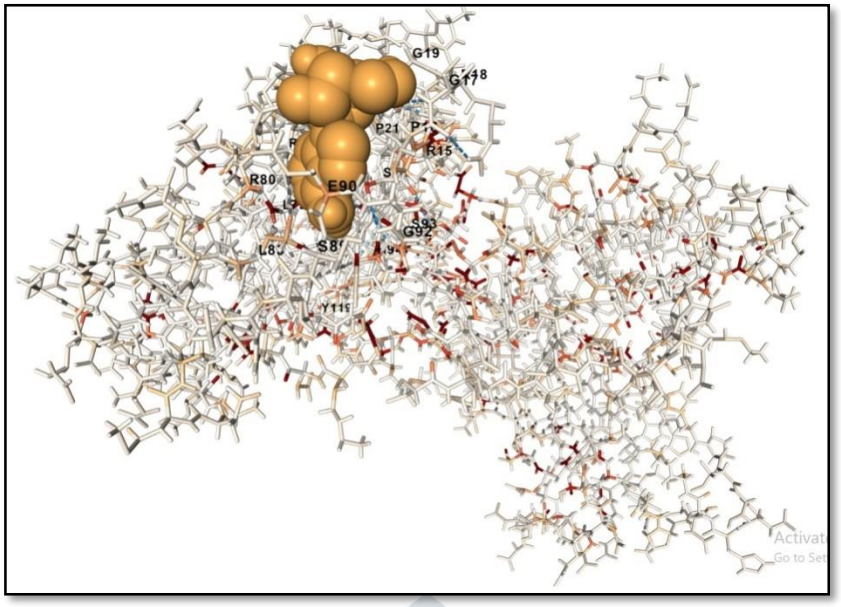

Figure 3: Docking Result with Adenosine

Docking Score: The docking score represents the binding affinity of the drugs to the pocket. Adenosine with a PubChem CID of 60961 yielded the best docking score of -6.9 kcal/mol for pocket C3. Lower docking scores, symbolized by more negative values, mean high binding between the ligand (in this case, Adenosine) and the protein.

Pocket C3: This binding site pocket has a volume of 221 Å^3. This pocket is more favorable according to the docking score rather than other pockets C1, C2, C4, and C5 for Adenosine binding.

Key Residues: Interactions of the following important residues in Pocket C3 are made to the Adenosine molecule through hydrogen bonds, hydrophobic interactions, or electrostatic interactions involving GLU20, ARG22, SER89, and HIS94 that help in stabilizing the ligand inside the binding site.

Table 2: Docking Result with Adenosine

CurPocket ID	Vina Score (kcal/mol)	Cavity Volume (Å^3)	Center (x, y, z)	Docking Size (x, y, z)
C3	-6.9	221	(4, 46, 216)	(19, 19, 19)
C1	-6.8	1791	(14, 41, 191)	(29, 19, 35)
C5	-6.2	174	(6, 52, 206)	(19, 19, 19)
C2	-6.0	397	(2, 41, 197)	(19, 19, 19)
C4	-4.9	194	(9, 54, 170)	(19, 19, 19)

Interacting residues in pocket C3 are:VAL13, SER14, ARG15, PRO16, GLY17, ARG18, GLY19, GLU20, PRO21, ARG22, ARG76, LEU79, ARG80, ILE81, LEU83, ARG84, SER89, GLU90, ALA91, GLY92, SER93, HIS94, TYR119

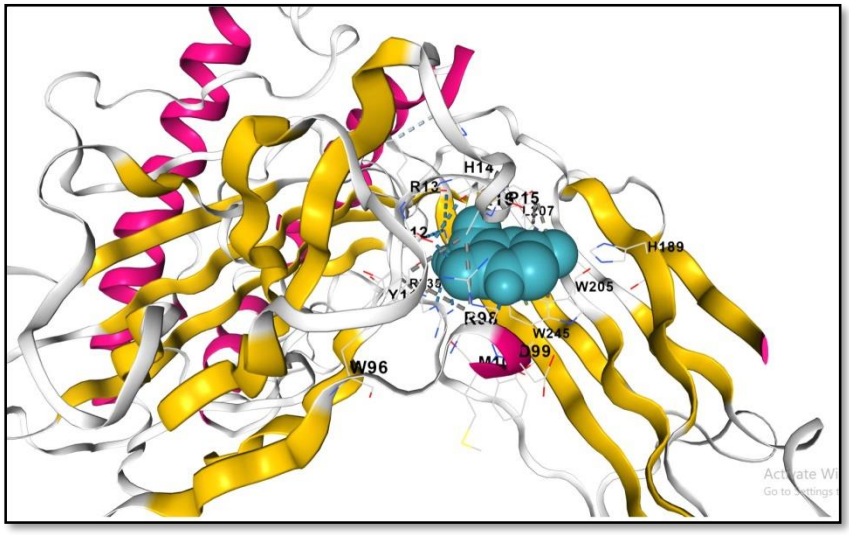

Figure 4: Docking Result with Favipiravir

Docking Score: The docking result shows that the ligand bound to pocket C4 has a Vina score of -5.6 kcal/mol, indicating a reasonably strong binding affinity. The C4 pocket shows a lower (more negative) score compared to some other pockets, suggesting that this is one of the better binding sites for the ligand.

Pocket C4: The binding site (pocket C4) has a cavity volume of 194 Å^3, which is relatively small but sufficient for accommodating the ligand. This pocket might represent a tighter or more specific binding site.

Key Residues: Pocket C4 contains key residues across both Chain A and Chain B, including HIS189, TRP205, ARG13, TRP96, and MET100. These residues likely play an important role in stabilizing the ligand within the binding cavity through a combination of hydrogen bonding, hydrophobic interactions, and other non-covalent forces.

Table 3: Docking Result with Favipiravir

CurPocket ID	Vina Score (kcal/mol)	Cavity Volume (Å^3)	Center (x, y, z)	Docking Size (x, y, z)
C4	-5.6	194	(9, 54, 170)	(17, 17, 17)
C3	-5.5	221	(4, 46, 216)	(17, 17, 17)
C1	-5.2	1791	(14, 41, 191)	(29, 23, 35)
C5	-5.1	174	(6, 52, 206)	(17, 17, 17)
C2	-4.7	397	(2, 41, 197)	(23, 23, 35)

Interacting residues in chain A: HIS189, THR191, HIS193, TRP205, LEU207, ARG235, GLN243

Interacting residues in chain B :TYR11, SER12, ARG13, HIS14, PRO15, ALA16, GLU17, TRP96, ASP97, ARG98, ASP99, MET100

3. **ADME and Toxicity Profiles**

SwissADME results indicated that the top-ranked drug candidates exhibited high gastrointestinal (GI) absorption and good bioavailability scores. Lipinski's Rule of Five was also used to assess the drug-likeness of the compounds.

The high GI absorption and good water solubility make Favipiravir a potential candidate for orally administered drugs. Favipiravir has low potential for drug-drug interactions, making it a safer candidate in polypharmacy settings where multiple medications are administered concurrently.

Adenosine’s high solubility, favorabledruglikeness profile, and low potential for drug-drug interactions make it a viable candidate for further exploration, particularly in formulations requiring rapid dissolution and systemic action.

PRO-TOX predicted low toxicity profiles for most candidates, with anLD50 greater than 500 mg/kg for the majority of the compounds, indicating safe therapeutic windows.

Favipiravir (PubChem CID: 492405) Analysis Based on SwissADME

Chemical Structure and Basic Information:

- **Formula**: $C_5H_4FN_3O_2$
- **Molecular Weight**: 157.10 g/mol
- **SMILES Representation**: Fc1c[nH]c(=O)c(n1)c(=O)N
- **Number of Heavy Atoms**: 11
- **Number of Aromatic Heavy Atoms**: 6
- **Number of Rotatable Bonds**: 1
- **Number of H-Bond Acceptors**: 4
- **Number of H-Bond Donors**: 2
- **Topological Polar Surface Area (TPSA)**: 88.84 Å²
- **Molar Refractivity**: 32.91

Lipophilicity:

- **Log P (iLOGP)**: 0.39
- **Log P (XLOGP3)**: -0.56
- **Log P (WLOGP)**: -0.57
- **Log P (MLOGP)**: -1.30
- **Log P (SILICOS-IT)**: 0.69
- **Consensus Log P**: -0.27

Water Solubility:

- **Log S (ESOL)**: -0.80 (Very soluble, 2.50e+01 mg/ml)
- **Log S (Ali)**: -0.84 (Very soluble, 2.29e+01 mg/ml)
- **Log S (SILICOS-IT)**: -1.42 (Soluble, 6.04e+00 mg/ml)

Pharmacokinetics:

- **GI Absorption**: High
- **BBB Permeant**: No
- **P-gp Substrate**: No
- **Inhibition of Cytochrome P450 Enzymes**:
  - **CYP1A2 Inhibitor**: No
  - **CYP2C19 Inhibitor**: No
  - **CYP2C9 Inhibitor**: No
  - **CYP2D6 Inhibitor**: No
  - **CYP3A4 Inhibitor**: No
- **Log Kp (Skin Permeation)**: -7.66 cm/s

Druglikeness:

- **Lipinski's Rule of Five**: 0 violations (Yes)
- **Ghose Filter**: 4 violations (MW < 160, WLOGP < -0.4, MR < 40, Atoms < 20)
- **Veber's Rule**: Yes
- **Egan's Rule**: Yes
- **Muegge's Rule**: 1 violation (MW < 200)
- **Bioavailability Score**: 0.55

Adenosine (PubChem CID: 60961) Analysis Based on SwissADME

Chemical Structure and Basic Information:

- **Formula**: $C_{10}H_{13}N_5O_4$
- **Molecular Weight**: 267.24 g/mol
- **SMILES Representation**: OC[C@H]1OC@Hn1cnc2c1ncnc2N
- **Number of Heavy Atoms**: 19
- **Number of Aromatic Heavy Atoms**: 9
- **Number of Rotatable Bonds**: 2
- **Number of H-Bond Acceptors**: 7
- **Number of H-Bond Donors**: 4
- **Topological Polar Surface Area (TPSA)**: 139.54 Å²
- **Molar Refractivity**: 62.67

Lipophilicity:

- **Log P (iLOGP)**: 0.61
- **Log P (XLOGP3)**: -1.05
- **Log P (WLOGP)**: -2.30
- **Log P (MLOGP)**: -2.72
- **Log P (SILICOS-IT)**: -2.37
- **Consensus Log P**: -1.57

Water Solubility:

- **Log S (ESOL)**: -1.05 (Very soluble, 2.36e+01 mg/ml)
- **Log S (Ali)**: -1.39 (Very soluble, 1.08e+01 mg/ml)
- **Log S (SILICOS-IT)**: 0.41 (Soluble, 6.85e+02 mg/ml)

Pharmacokinetics:

- **GI Absorption**: Low
- **BBB Permeant**: No
- **P-gp Substrate**: No
- **Inhibition of Cytochrome P450 Enzymes**:
  - **CYP1A2 Inhibitor**: No
  - **CYP2C19 Inhibitor**: No
  - **CYP2C9 Inhibitor**: No
  - **CYP2D6 Inhibitor**: No
  - **CYP3A4 Inhibitor**: No
- **Log Kp (Skin Permeation)**: -8.68 cm/s

Druglikeness:

- **Lipinski's Rule of Five**: 0 violations (Yes)
- **Ghose Filter**: 1 violation (WLOGP < -0.4)
- **Veber's Rule**: Yes
- **Egan's Rule**: 1 violation (TPSA > 131.6)
- **Muegge's Rule**: Yes
- **Bioavailability Score**: 0.55

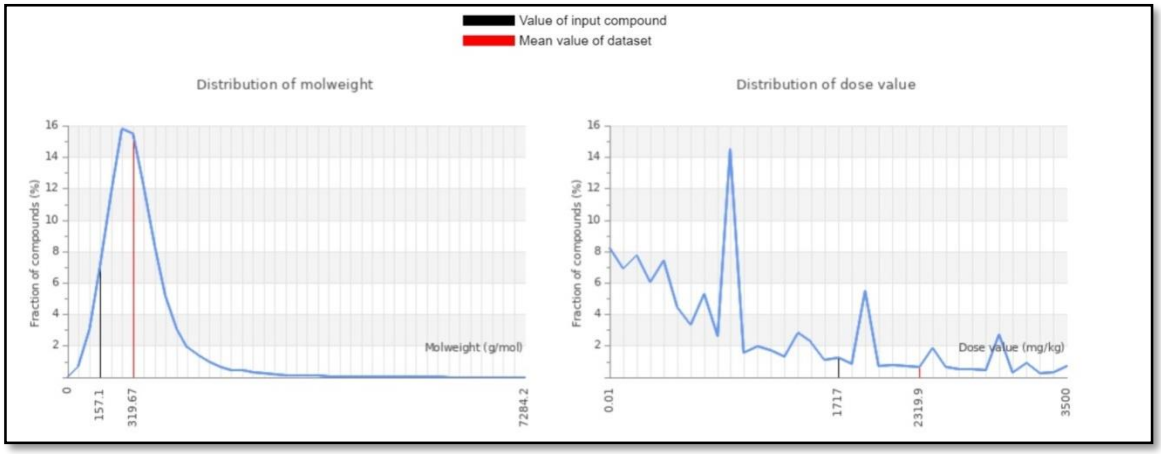

Figure 5: Pro-Tox Result for Favipiravir

Molecular Weight Distribution:

- **Favipiravir's molecular weight** is **157.1 g/mol** (marked by the black line in the graph).
- The **mean molecular weight** of the dataset is **319.67 g/mol** (marked by the red line).
- **Result**: Favipiravir's molecular weight is significantly lower than the average molecular weight of the dataset compounds. Favipiravir falls within the lower range of molecular weights, suggesting it is smaller and potentially more bioavailable.

Dose Value Distribution:

- **Favipiravir's dose value** is marked by the black line in the second plot.
- The **mean dose value** of the dataset is around **2319.9 mg/kg** (marked by the red line).
- **Result**: Favipiravir's dose value is lower than the mean value of the dataset compounds. This suggests that Favipiravir might be effective at a lower dose compared to many other compounds in the dataset, potentially indicating better potency or efficacy.

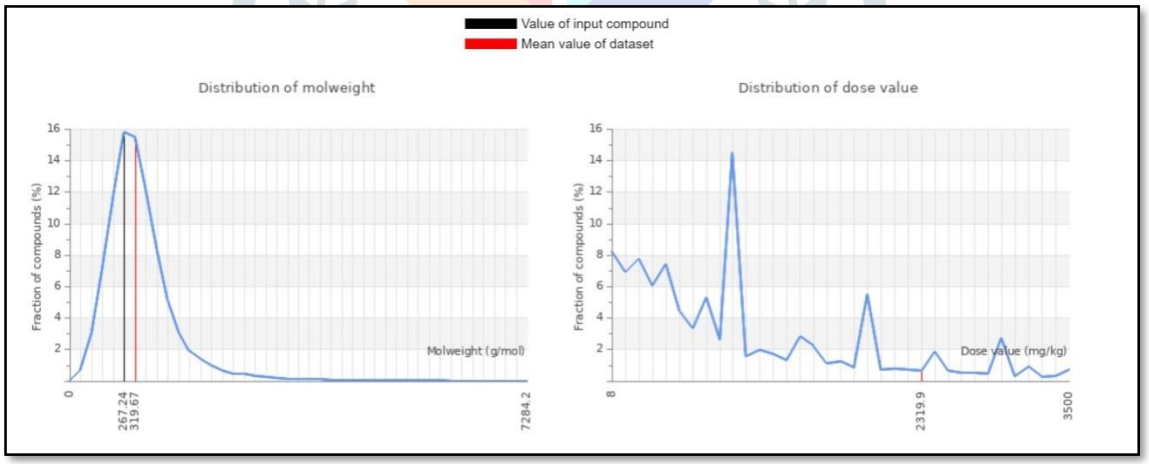

Figure 6: Pro-Tox Result for Adenosine

Molecular Weight Distribution:

- **Adenosine's molecular weight** is **267.24 g/mol** (marked by the black line in the graph).
- The **mean molecular weight** of the dataset is **319.67 g/mol** (marked by the red line).
- **Result**: Adenosine has a **lower molecular weight** than the average molecular weight of the dataset compounds, but it is still within a common range for drug-like molecules. This suggests that Adenosine may possess better **bioavailability** and faster distribution due to its smaller size.

Dose Value Distribution:

- **Adenosine's dose value** is marked by the black line in the second plot.
- The **mean dose value** of the dataset is around **2319.9 mg/kg** (marked by the red line).

- **Result**: Adenosine's dose value is lower than the average dose value of the dataset, indicating that Adenosine might be effective at a **lower dose** compared to the majority of compounds in the dataset. This may suggest **higher potency** and potential for lower therapeutic dosages.

4. Machine Learning Results

The **Support Vector Regression (SVR)** model performed well in predicting the **binding free energy (ΔG)**. The model used the interface area and the number of hydrogen bonds as input features, achieving a **Mean Squared Error (MSE) of 1.72**, which indicated reliable predictions.

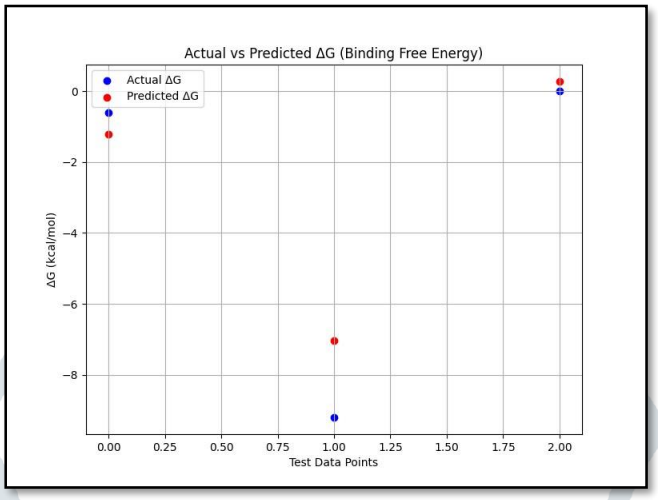

Figure 7:Actual vs Predicted ΔG

The plot shows the comparison between the actual and predicted ΔG values using the SVR model. The close alignment of predicted values (red) with actual values (blue) indicates the robustness of the model.

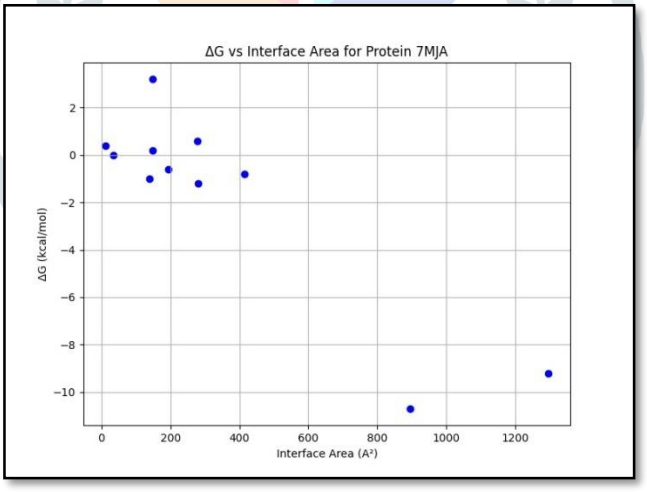

Figure 8: ΔG vs Interface Area

The graph shows the relationship between the binding free energy (ΔG) and the interface area (Å²) for protein 7MJA. It demonstrates that larger interface areas correspond to more negative ΔG values, indicating stronger binding interactions.

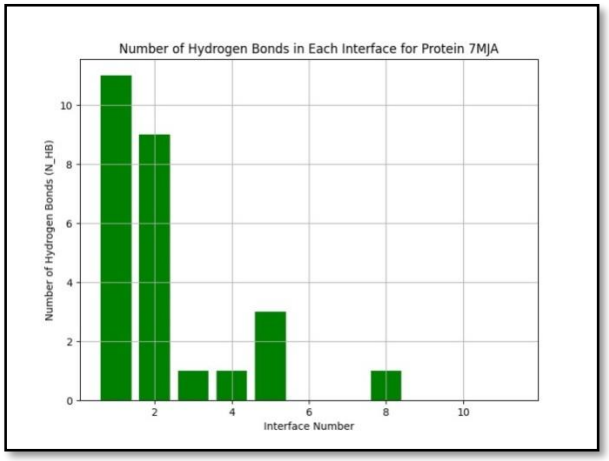

Figure 9: Number of Hydrogen Bonds in Each Interface

This bar chart displays the number of hydrogen bonds across different binding interfaces for protein 7MJA. The first few interfaces exhibit a high number of hydrogen bonds, correlating with stronger binding affinity.

Conclusion

The combined docking and machine learning results indicate that **Favipiravir** and **Adenosine** exhibit moderate binding affinity to protein 7MJA, supported by the **Vina scores** and hydrogen bond interactions observed in the docking analysis. The **SVR model** successfully predicted the binding free energy with high accuracy (MSE: 1.72), demonstrating the effectiveness of using interface area and hydrogen bonding as key predictors for ΔG. These insights can be utilized for further optimization of drug candidates targeting neuroblastoma-derived PHOX2B peptides bound to the HLA-A*24:02 complex.

References

1. Yarmarkovich M, Marshall QF, Warrington JM, et al. Cross-HLA targeting of intracellular oncoproteins with peptide-centric CARs [retracted in: Nature. 2023 Nov;623(7988):872. doi: 10.1038/s41586-023-06731-z]. *Nature*. 2021;599(7885):477-484. doi:10.1038/s41586-021-04061-6
2. Zafar A, Wang W, Liu G, et al. Molecular targeting therapies for neuroblastoma: Progress and challenges [published correction appears in Med Res Rev. 2022 Jan;42(1):641. doi: 10.1002/med.21843]. *Med Res Rev*. 2021;41(2):961-1021. doi:10.1002/med.21750
3. Aygun N. Biological and Genetic Features of Neuroblastoma and Their Clinical Importance. *CurrPediatr Rev*. 2018;14(2):73-90. doi:10.2174/1573396314666180129101627
4. Boeva V, Louis-Brennetot C, Peltier A, et al. Heterogeneity of neuroblastoma cell identity defined by transcriptional circuitries. *Nat Genet*. 2017;49(9):1408-1413. doi:10.1038/ng.3921
5. Peggion S, Najem S, Kolman JP, Reinshagen K, PagerolsRaluy L. Revisiting Neuroblastoma: Nrf2, NF-κB and Phox2B as a Promising Network in Neuroblastoma. *Curr Issues Mol Biol*. 2024;46(4):3193-3208. Published 2024 Apr 6. doi:10.3390/cimb46040200
6. Ma Y, Feng J, Zhao J, et al. PHOX2B as a Reliable Marker for Neuroblastoma in Tissue and Cytology Specimens. *J NeuropatholExp Neurol*. 2021;80(12):1108-1116. doi:10.1093/jnen/nlab112
7. Windels ML, Cordier F, Van Dorpe J, Ferdinande L, Creytens D. PHOX2B: a diagnostic cornerstone in neurocristopathies and neuroblastomas. *J ClinPathol*. 2024;77(6):378-382. Published 2024 May 17. doi:10.1136/jcp-2023-209047
8. Fan H, Xing T, Hong H, et al. The expression of PHOX2B in bone marrow and peripheral blood predicts adverse clinical outcome in non-high-risk neuroblastoma. *PediatrHematolOncol*. 2022;39(4):343-356. doi:10.1080/08880018.2021.1995090
9. Krishnan, P. A., Velloor, A., Kumari, U., Aggarwal, N., Mukherjee, S., & MS, S. (2024). NGS Diagnosis of Lp-PLA2 in Cardiovascular Disease &Insilico Analysis of Human Filamin C Domains 14-15 Mutation G1676R. *International Neurourology Journal*, *28*(1), 434-445.
10. Kumari, U., Dwibedi, R., Rathi, V., &Aggarwal, N. (2024). Computer Aided Drug Discovery and Proteomic Sequence Analysis of Tobacco Mosaic Virus. *International Neurourology Journal*, *28*(1), 609-619.
11. Dr Uma, Devanshi Gupta, Study of Malignant Melonema Causing Protein "5ixb" Using Multiomics Approach, November 2022, INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN TECHNOLOGY , Volume 9 Issue 6 , ISSN: 2349-6002
12. Kumari, Uma &Tripathi, Kartik. (2023). Computational Analysis and Molecular Docking Approach for Liver Cancer. International Journal for Research in Applied Science and Engineering Technology. 11. 1828 to 1633. 10.22214/ijraset.2023.54927.
13. Huang, C. C., Meng, E. C., Morris, J. H., Pettersen, E. F., &Ferrin, T. E. (2014). Enhancing UCSF Chimera through web services. *Nucleic acids research*, *42*(Web Server issue), W478–W484. https://doi.org/10.1093/nar/gku377
14. Daina, A., Michielin, O. &Zoete, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep* **7**, 42717 (2017). https://doi.org/10.1038/srep42717
15. Drwal, M. N., Banerjee, P., Dunkel, M., Wettig, M. R., &Preissner, R. (2014). ProTox: a web server for the in silico prediction of rodent oral toxicity. *Nucleic acids research*, *42*(Web Server issue), W53–W58. https://doi.org/10.1093/nar/gku401

16. Pedregosa, Fabian&Varoquaux, Gael &Gramfort, Alexandre& Michel, Vincent &Thirion, Bertrand &Grisel, Olivier &Blondel, Mathieu &Prettenhofer, Peter & Weiss, Ron &Dubourg, Vincent &Vanderplas, Jake &Passos, Alexandre&Cournapeau, David &Brucher, Matthieu& Perrot, Matthieu&Duchesnay, Edouard&Louppe, Gilles. (2012). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research. 12.
17. Cock, P. J., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., & de Hoon, M. J. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics (Oxford, England)*, *25*(11), 1422–1423. https://doi.org/10.1093/bioinformatics/btp163