IOSR Journal Of Pharmacy And Biological Sciences (IOSR-JPBS)
e-ISSN:2278-3008, p-ISSN:2319-7676. Volume 20, Issue 2 Ser. 2 (Mar. – Apr. 2025), PP 63-75
www.Iosrjournals.Org

Comprehensive Structure-Based Drug Designing And Toxicity Identification Of The Human Papillomavirus Type 18 E2 DNA-Binding Domain Bound To Its DNA Target

Juri Saikia, Tanmay Bandbe, Uma Kumari
(Project Trainee At Bioinformatics Project And Research Institute, Noida-201301, India.)
(Project Trainee At Bioinformatics Project And Research Institute, Noida-201301, India.)
(Senior Bioinformatics Scientist, Bioinformatics Project And Research Institute, Noida-201301, India.)

Abstract:
The Human papillomavirus (HPV) type 18 is a high-risk HPV type and a significant contributor to the development of cervical cancer. The E2 protein, a key regulator of HPV replication and transcription, binds to specific DNA sequences within the viral genome, making it a prime target for therapeutic intervention. This study presents an extensive, combined approach of structure-based drug design (SBDD) and computer-aided drug design (CADD) analysis to enhance the drug discovery process for this viral target and to identify potential inhibitors. The goal of this study is to identify promising drug candidates that can inhibit the E2 DNA-binding domain, potentially disrupting the HPV life cycle and reducing the risk of cervical cancer development. Homology modeling was employed to construct a 3D model of the HPV18 E2 DNA-binding domain. To ensure the integrity and quality of protein models, ERRAT (SAVES Server) was utilized to evaluate non-bonded atomic interactions, and PROCHECK analysis provided detailed stereochemical insights through the Ramachandran plot, confirming the structural quality of models. Subsequently, molecular docking and virtual screening were conducted using CB-Dock to facilitate the identification of promising inhibitors, which were then evaluated for binding affinity and specificity through molecular dynamics simulations and free energy calculations. Likewise, toxicity prediction was performed using ProTox 3.0 to ensure the safety and efficacy of the identified compounds. This integrative methodology enhances our understanding of the molecular interactions and stability of the E2-DNA complex, paving the way for the development of targeted antiviral therapies. These research findings highlight the potential drug candidates for therapeutic intervention, underscore the efficacy of CADD in accelerating the drug discovery process, and provide valuable insights into the molecular interactions involved in the pathogenesis, ultimately contributing to the development of targeted treatments for HPV-induced cervical cancer.
Background*: Human papillomavirus type 18 (HPV18), a high-risk oncogenic virus, is strongly associated with cervical cancer. The E2 protein, a key regulatory factor in the HPV lifecycle, exerts its function by binding to specific DNA sequences within the viral genome. Disrupting this interaction represents a promising therapeutic strategy against HPV-related diseases. Structure-based drug design leverages the three-dimensional structure of the E2-DNA complex to identify potential inhibitors. Furthermore, predicting the toxicity of these potential drug candidates is crucial for their development into safe and effective antivirals.*
Materials and Methods*: This study utilized computational methods to investigate the HPV-18 E2 DNA-binding domain. We retrieved protein structures from the PDB and visualized them using RasMol to analyze interactions. PyMOL was used to calculate RMSD for comparing structural similarities between protein conformations. The ERRAT and PROCHECK servers assessed the protein structures' quality and stereochemistry. CB-Dock2 was employed for molecular docking of anticancer drugs to predict binding modes and affinities, and ProTox 3.0 evaluated the safety and efficacy of selected ligands.*
Results*: The RasMol and PyMOL provided detailed 3D structures and comparative analysis of proteins 1JJ4 and 1F9F, revealing structural similarities (RMSD 0.556 Å), implying shared functions. The ERRAT analysis indicated high structural reliability for both proteins (scores > 93), with 1JJ4 slightly higher. PROCHECK, particularly the Ramachandran plot, confirmed the stereochemical quality of 1JJ4, with most residues in favoured regions. While the docking studies using CB-DOCK2 identified Irinotecan as having the highest binding affinity for 1JJ4 (score -9.5). Also, the ProTox 3.0 predicted Irinotecan as moderately harmful with lower acute toxicity compared to other tested ligands.*
Conclusion: *This research employs computational methods (SBDD and CADD) to find new drugs/inhibitors that target the HPV18 E2 protein's DNA-binding region, aiming to speed up drug discovery for this viral protein. The study identified key structural features of the protein-DNA that could disrupt their interaction.*

Irinotecan was identified as the most promising inhibitor based on docking simulations and toxicity profile. These findings provide a framework for designing targeted therapies against HPV18, potentially leading to more effective treatments for cervical cancer with fewer side effects.

***Key Word**: Human papillomavirus type 18 (HPV-18); in-silico drug design; Molecular Docking; Toxicity Prediction; HPV Oncogenesis; Ligand-Protein interaction*

Date of Submission: 18-04-2025 Date of Acceptance: 28-04-2025

I. Introduction

Cancer is one of the leading causes of death, making it a critical public health concern. Cervical cancer is one of the most preventable types of cancer; it remains the second most common cause of cancer deaths in women under 50[1, 2]. This disease is increasing in countries due to a lack of preventive measures. Cervical cancer is due to many environmental and genetic factors that lead to changes in epithelial cells and malignancy. It is a major health problem for women in the world. Mortality is high due to late and misdiagnosis. These tumors' initial formation and growth depend on two main oncogenes, E6 and E7, whose expression leads to cancer. This disease is a malignant tumor that develops in or on the lower part of the uterus and is divided into two categories depending on the type of cervical cells affected: cervical squamous cell which affects the squamous cells of the cervix, and cervical adenocarcinoma, which targets the mucus-producing cells of the cervix[3]. Most cases of cervical cancer follow infection with high-risk types of genital human papillomavirus (HPV)[4, 5].

HPV is an important human virus with some high-risk types, such as HPV-18, associated with cancer and other malignancies. Viruses act by expressing proteins, especially the E2 protein, which plays an important role in controlling viral infection and mutation. DNA binding domain (DBD) binds to specific sequences in viral proteins to regulate this process. The study of the structure of this domain is important to understand how the virus hijacks cellular machinery, providing an opportunity for drug development. It inhibits viral replication by identifying small molecules or inhibitors that can bind to this domain and prevent it from interacting with viral DNA. Risk factors for cervical cancer include: smoking tobacco, a n increasing number of sexual partners, early sexual activity, other sexually transmitted infections, a weakened immune system, and exposure to miscarriage prevention medicine. In recent years, the integration of bioinformatics, molecular modeling, and machine learning into the drug development process has achieved significant results in terms of accelerating the procedure and decreasing the costs associated with the development of new drug candidates[6]. This approach uses molecular docking, computational modeling, and simulation techniques to develop drug candidates.

Computer-aided drug design (CADD) enhances structure-based drug design (SBDD) by using computational tools to streamline and amplify the drug discovery system. Techniques, together with molecular docking, molecular dynamics (MD) simulations, and virtual screening, allow for the rapid evaluation of capable inhibitors.

For HPV-18 E2 DBD, CADD affords an efficient approach to exploring the chemical area, predicting binding affinities, and understanding the dynamic interactions between the protein and its DNA target[7]. The integration of SBDD and CADD processes to perceive and optimize small molecules capable of selectively targeting the HPV-18 E2 DBD. This dual technique not only speeds up the drug discovery technique but may enable the development of personalized treatment options for HPV-associated cancers and additionally increases the likelihood of identifying mighty and particular inhibitors with minimal off-target effects. This study aims to combine computational and experimental methods to develop a new strategy for HPV treatment[8, 9, 10, 11, 12, 13].

II. Material And Methods

Target Selection, Data Retrieval, and Protein Visualization using RasMol:

We retrieved the crystal structures of the HPV-18 E2 DNA-binding domain (PDB IDs: 1JJ4 and 1F9F) from the Protein Data Bank (PDB) (www.rcsb.org) in PDB format. To visualize and analyze these protein structures (1JJ4 and 1F9F), we used the RasMol molecular graphics program (version 2.7.5) (http://www.openrasmol.org/). RasMol offers a user-friendly interface to explore intricate molecular structures obtained from data or generated through computational modeling. We utilized commands within RasMol, such as wireframe, ribbons, and spacefill, to examine the structures in various representations. This allowed us to analyze features like hydrogen bonds, salt bridges, and other non-covalent interactions between the E2 protein and DNA. Additionally, we could identify alpha-helices and beta-sheets within the protein, which provided insights into how it interacts with DNA and other proteins[14].

RMSD Score Calculation using PyMOL:

We used PyMOL Molecular Graphics System (v3.10) (https://pymol.org/) to visualize and compare the conformations of proteins 1JJ4 and 1F9F from their PDB files. To assess structural differences, we superimposed the structures and calculated the Root Mean Square Deviation (RMSD). We focused our analysis on the protein and DNA by removing water molecules and other irrelevant heteroatoms (ligands, ions) before the RMSD calculation. A lower RMSD indicates higher structural similarity between the two proteins, while a higher value suggests significant conformational changes[15, 16].

Quality and Integrity Check of Protein using ERRAT:

We employed the ERRAT server online (https://saves.mbi.ucla.edu/) and uploaded PDB files (1JJ4 and 1F9F) as input. This server analyzed the statistics of non-bonded interactions between different atom types within the protein structures and compared these interactions to a database of well-refined, high-resolution structures. The output generated an overall quality factor that indicates the percentage of residues with acceptable non-bonded interactions and a plot of the error function versus residue position, highlighting regions with potentially problematic interactions[17].

Stereochemical Quality and Integrity Check of Protein using PROCHECK:

We accessed the PROCHECK tool online (https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/) to evaluate the stereochemical quality of our protein structure. We opted for the PDB file 1JJ4 as input over 1F9F due to its higher accuracy, which generated a validation report with default parameters. The report provided us insights into various aspects of protein geometry such as Ramachandran plots which displayed the distribution of φ (phi) and ψ (psi) dihedral angles in the protein structure as well as the percentage of residues in the most favored regions, additionally allowed/ disallowed regions, main-chain parameters such as bond lengths, bond angles, and torsion angles, and side-chain parameters to identify any unusual side-chain conformations[18].

Ligand Preparation and Molecular Docking using CB-Dock2:

We first obtained the 2D chemical structures (SDF format) of the ligands/ anticancer drugs (Irinotecan, Doxorubicin, Paclitaxel, Topotecan, and Topotecan Hydrochloride) specific to the E2 DNA-binding domain from PubChem. We then used AutoDock Vina to prepare each ligand for docking, which involved adding missing hydrogens and atoms, assigning atom charges, minimizing ligand energy, and generating potential ionization states. Later, the prepared protein structure and the library of prepared ligand/ drug structures were uploaded to the CB-Dock2 server .Before initiating the simulation, docking parameters on CB-Dock2, such as search space size, number of generated poses, and scoring function, were confirmed. Similarly, we performed docking simulations for each drug against the protein domain to predict the binding modes (how the drug interacts with the protein) and binding affinities (strength of the drug-protein interaction). Further, software like PyMOL or Chimera can be used to visualize the predicted binding modes for further analysis[19, 20].

Assessing Safety and Efficacy of Potential Ligand Molecules using ProTox 3.0:

We selected Irinotecan and Doxorubicin as potential ligands/ anticancer drugs for the HPV type 18 E2 DNA-binding domain based on their predicted binding affinity from CB-Dock2 analysis and given as input to ProTox to evaluate their safety and efficacy on humans. The profiles were compared to known toxicity data to validate the predictions. The output showed toxicity profiles that provide predictions for various toxicity endpoints such as acute toxicity (LD50 values), organ toxicity, carcinogenicity, and other toxicological endpoints[21].

III. Result

RasMol Analysis:

RasMol is a computational tool used for the visualization and exploration of molecular structures to understand the spatial arrangement and interactions within the molecule.

Figure 1: 1JJ4 RasMol Analysis

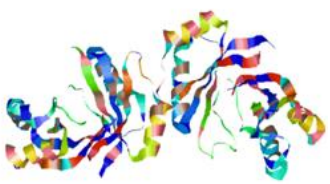


Figure 2: 1F9F RasMol Analysis

The structural analysis of proteins 1JJ4 and 1F9F using RasMol has provided valuable insights into their three-dimensional conformations and molecular interactions. With the high-resolution images generated by RasMol, we were able to examine the intricate architecture of these proteins, including their active sites and key residues involved in binding and functional processes.

Additionally, the visualization of spatial arrangements facilitated the identification of critical hydrogen bonds, hydrophobic interactions, and electrostatic forces. These findings demonstrate the utility of RasMol in structural biology, providing detailed visualizations that are essential for drug discovery research.

RMSD Score using PyMol:

Precise identification and analysis of protein active sites is crucial in structural biology and drug discovery. PyMOL facilitates the visualization and examination of these sites, contributing to a deeper understanding of protein function and the development of targeted drugs. The Root Mean Square Deviation (RMSD) score, calculated in PyMOL, measures the average deviation between corresponding atoms in two aligned protein structures.

According to RMSD scoring guidelines:

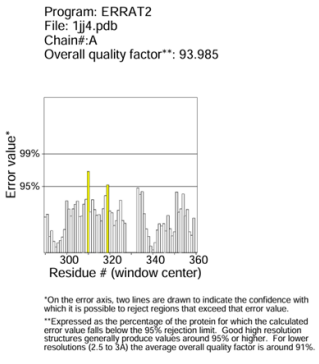
- 0-1 Å: Very close match; structures are nearly identical, suggesting high-resolution crystallographic data.
- 1-3 Å: Good match, with notable similarity.
- 3-5 Å: Moderate similarity.
- Above 5 Å: Significant structural differences.

A comparative analysis was performed using PyMOL on proteins 1JJ34 and 1F9F. The two structures were aligned over 1,156 atoms, resulting in a superposition score of 898. The RMSD score calculated between the structures 1JJ34 and 1F9F is 0.556 Å. This relatively low RMSD value indicates a reasonable degree of structural similarity between the two proteins. These structural similarities imply that these proteins share functional characteristics, which could be pivotal for understanding their biological roles and exploring potential therapeutic applications.

ERRAT Analysis:

ERRAT is an important tool in protein structure validation for evaluating the accuracy of protein models. ERRAT assesses the quality of a protein structure by comparing its non-bonded atom-atom interactions to a database of high-resolution structures, allowing researchers to confirm the accuracy and reliability of their protein models. The ERRAT plot illustrates protein structural regions based on their statistical reliability. High scores suggest the model is reliable and accurate, while low scores highlight areas that may need further refinement or validation. The 5% of high-quality protein structures may have flaws in the yellow-highlighted regions, which are those that can be rejected at the 95% confidence level. And, the red-colored regions indicate a 1% possibility of error and are rejected at the 99% confidence level.

Based on the ERRAT analysis, the overall quality factors of the protein structures 1JJ4 and 1F9F were determined to be 93.985 and 93.773, respectively. A score above 90 generally indicates a well-defined structure. Therefore, the ERRAT analysis suggests that both 1JJ44 and 1F9F are likely to have accurate and reliable protein structures. Although 1JJ4 exhibits a slightly higher overall quality factor compared to 1F9F. This suggests that 1JJ4 may have a more accurate and reliable structural model based on the criteria evaluated by the ERRAT. In structural-based drug design, precise models are essential. By providing a reliable structural framework, these models enable the testing of drug candidates against accurate targets, ultimately leading to more effective and focused treatments.



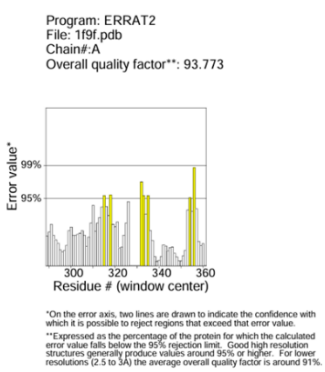


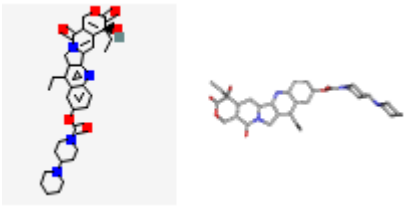
Figure 3: 1JJ4 ERRAT Analysis
Figure 4: 1F9F ERRAT Analysis

Hence, the substantial improvement in the 1JJ4 model enhances its reliability for identifying and developing targeted therapies against HPV-induced cancer.

PROCHECK Analysis:

PROCHECK analysis is a computational tool used to assess the stereochemical quality of protein structures. It performs a thorough analysis, including Ramachandran plots, to evaluate the overall quality and potential issues within a protein structure. PROCHECK assists researchers in identifying regions that may require additional refining or modeling by examining numerous factors like bond lengths, bond angles, and dihedral angles. A high percentage of residues in favoured regions and few outliers indicates a high-quality structure.

Of the total residues, 14 (10.6%) are located in additional, less optimal positions (yellow). While not ideal, these positions are still considered acceptable and do not significantly compromise the model's overall quality. Importantly, no residues were detected in the generously allowed (light yellow) or strictly disallowed (white) regions. The absence of residues in forbidden regions, which frequently signals structural inaccuracies in proteins, further reinforces the model's accuracy. The red-highlighted sections, representing the most preferred regions in this plot, encompass the majority (89.4%) of the examined amino acid residues (118 out of 182). This high percentage strongly indicates a stable and well-constructed protein model, as high-quality models typically contain over 90% of their residues in such regions. Furthermore, the plot provides detailed statistics on specific residue types.



Chain A: ASN297 Chain B: ARG296 ASN297 HIS321 LYS329

Figure 5: Structure of Irinotecan

It includes 6 glycine, 4 proline, and 132 other residues. Glycine residues, depicted as triangles, can occupy less populated regions due to their unique flexibility. Proline residues, with their restricted conformational space (φ-ψ angles), are positioned differently on the plot.

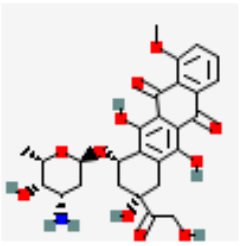
Overall, the Ramachandran plot analysis for the protein structure 1JJ4 reveals the quality of their conformations, with the majority of the residues falling within the favoured regions of the plot, indicating that the backbone dihedral angles are energetically favourable and consistent with known protein structures.

CB-Dock Analysis:

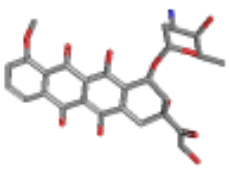
CB-Dock is a computational tool used in drug discovery and structural biology to predict the precise locations and orientations (binding poses) of small molecules (ligands) within the binding sites of proteins. It uses the AutoDock Vina program to generate potential binding poses and CurPocket to identify binding sites. The CB-Dock score, calculated based on various interactions (Van der Waals interactions, Electrostatic interactions, Hydrogen bonding), estimates the strength of ligand-protein binding. A lower CB-Dock score generally corresponds to a stronger binding affinity, implying a higher likelihood of the ligand binding to the protein at the predicted binding site.

Irinotecan

2D



3D



Chain B: ASP295 ARG296

ASN297 SER298 LYS300 LYS329

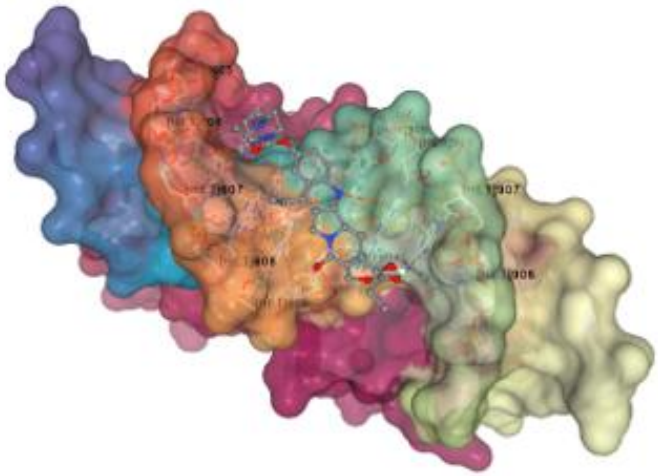
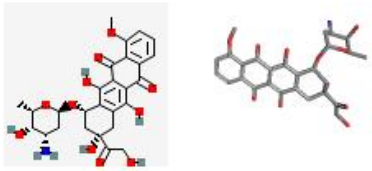


Figure 6: Protein-Ligand Docking (1JJ4-Irinotecan)

b. Doxorubicin

2D 3D



Chain B: ASP295 ARG296 ASN297 SER298 LYS300 LYS329

Figure 7: Structure of Doxorubicin

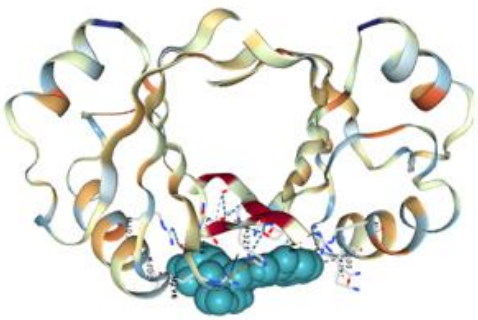
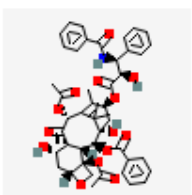


Figure 8: Protein-Ligand Docking (1JJ4-Doxorubicin)

c. Paclitaxel

2D



Chain B: ASP295 ARG296 ASN297 LYS300 LYS329

Figure 9: Structure of Paclitaxel

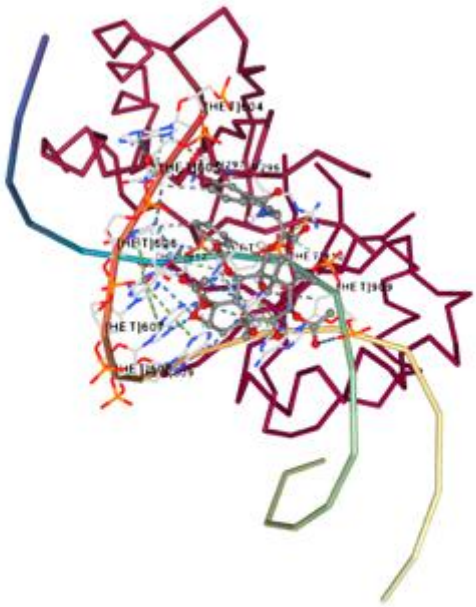
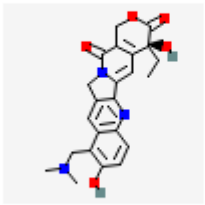
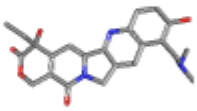


Figure 10: Protein-Ligand Docking (1JJ4-Paclitaxel)

d. Topotecan

2D 3D





Chain B: ASP295 ARG296 ASN297 SER298

HIS321 LYS329 THR330

Figure 11: Structure of Topotecan

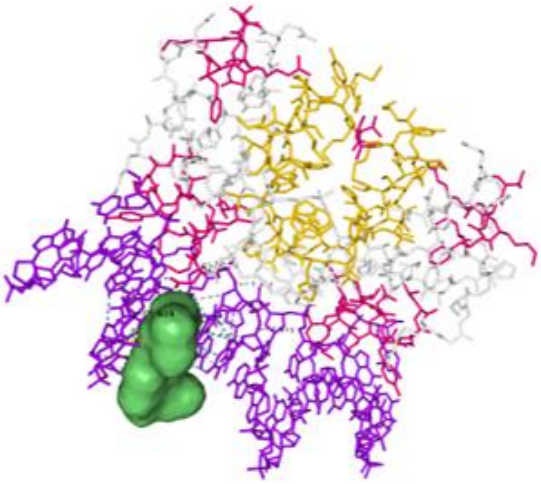
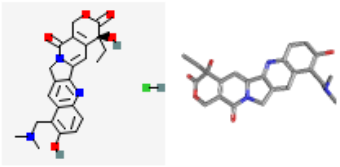


Figure 12: Protein-Ligand Docking (1JJ4-Topotecan)

e. Topotecan Hydrochloride

2D 3D



Chain B: ASP295 ARG296 ASN297 LYS300 LYS329

Chain B: ASP295 ARG296 ASN297 SER298 GLU328 LYS329 THR330

Figure 13: Structure of Topotecan Hydrochloride

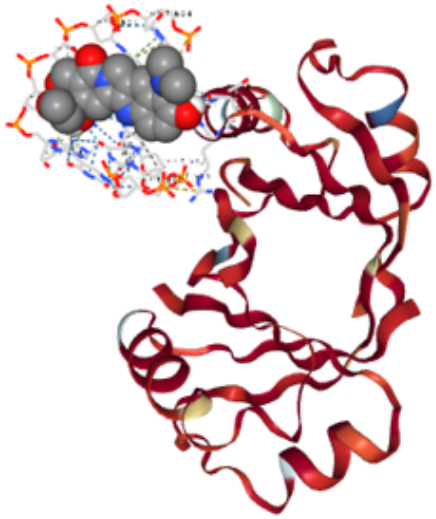


Figure 14: Protein-Ligand Docking (1JJ4-topotecan Hydrochloride)

Sr. No.	Ligand	PubChem ID	Docking Score	Center (x, y, z)	Remarks
1	Irinotecan	60838	-9.5	18, 22, 13	Very strong binding affinity
2	Doxorubicin	31703	-8.0	18, 23, 23	Strong binding affinity
3	Paclitaxel	46188928	-7.9	18, 23, 23	Strong binding affinity
4	Topotecan	25141092	-7.9	18, 23, 23	Strong binding affinity
5	Topotecan Hydrochloride	134436906	-7.6	18, 23, 23	Moderate binding affinity

Table no 1: CB-Dock Analysis of Protein-Ligand/ Drug Compound

This docking study aimed to evaluate the binding affinities of various ligands to protein 1JJ4 by using CB-DOCK2. Among the investigated ligands, Irinotecan exhibited the highest binding affinity with a docking score of -9.5, suggesting the strongest interaction with protein 1JJ4. Doxorubicin shows a strong binding affinity with a docking score of -8.0. Further, Paclitaxel and Topotecan displayed comparable strong binding affinities with docking scores of -7.9 each**.** Also, Topotecan Hydrochloride, a derivative of Topotecan, showed a slightly lower/ moderate binding affinity than the other inhibitors, with a score of -7.6. This suggests that while these ligands can bind effectively to protein 1JJ4, their selectivity might vary, impacting their therapeutic potential. Overall, Irinotecan emerged as a promising candidate based on its docking score. However, further experimental studies and investigations are necessary to elucidate the precise binding mode, specificity, and biological activity of these compounds.

ProTox 3.0 Analysis:

ProTox 3.0 is a powerful, revolutionary tool that can predict chemical toxicity by combining molecular similarity, fragment-based approaches, and machine learning. It classifies chemicals based on various toxicity endpoints, including acute toxicity, organ toxicity, and specific toxicological effects. By analyzing the chemical's structure and properties, this tool can identify potential risks to human health and the environment.

Sr. No.	Drug Name	Predicted LD50 (mg/kg)	Predicted Toxicity Class	Remarks	Prediction Accuracy
1	Irinotecan	765	4	moderate harmful (p53 Active)	100%
2	Doxorubicin	205	3	harmful (p53 Active)	100%

Table no 2: Toxicity Prediction of Drugs by ProTox 3.0

Irinotecan

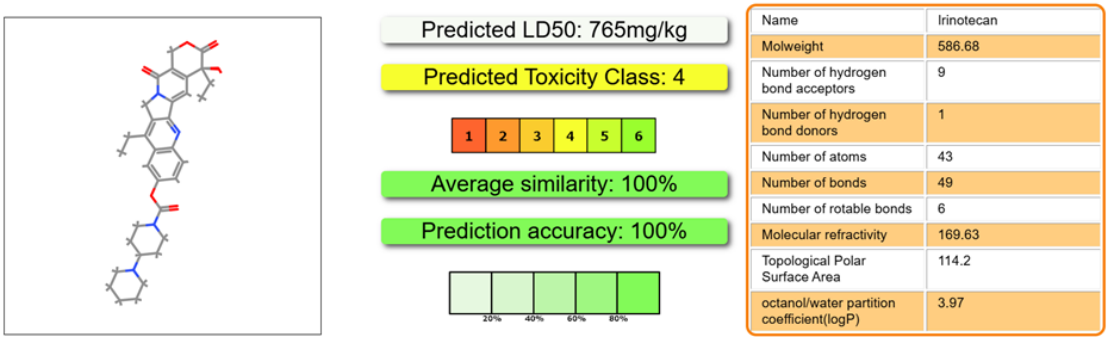


Figure 15: Oral Toxicity Prediction Results for Irinotecan

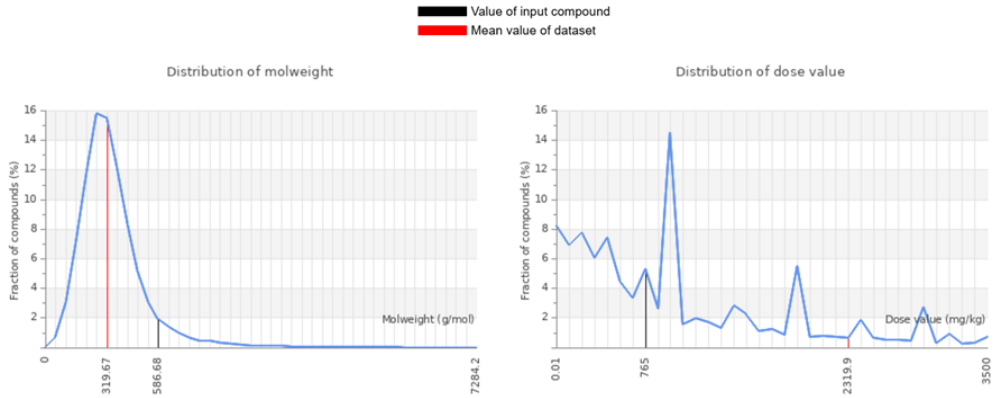


Figure 16: Comparison of Irinotecan with Dataset Compounds

Table no 3: Toxicity Model Report for Irinotecan

Classification	Target	Shorthand	Prediction	Probability
Organ toxicity	Cardiotoxicity	cardio	Inactive	0.80
Toxicity end points	Carcinogenicity	carcino	Inactive	0.61
Toxicity end points	Mutagenicity	mutagen	Inactive	0.67
Tox21-Nuclear receptor signalling pathways	Androgen Receptor (AR)	nr_ar	Not Calculated	Not Calculated
Tox21-Nuclear receptor signalling pathways	Androgen Receptor Ligand Binding Domain (AR-LBD)	nr_ar_lbd	Inactive	0.98
Tox21-Stress response pathways	Phosphoprotein (Tumor Supressor) p53	sr_p53	Active	0.60
Molecular Initiating Events	Thyroid hormone receptor alpha (THRα)	mie_thr_alpha	Inactive	0.90
Metabolism	Cytochrome CYP1A2	CYP1A2	Inactive	0.92

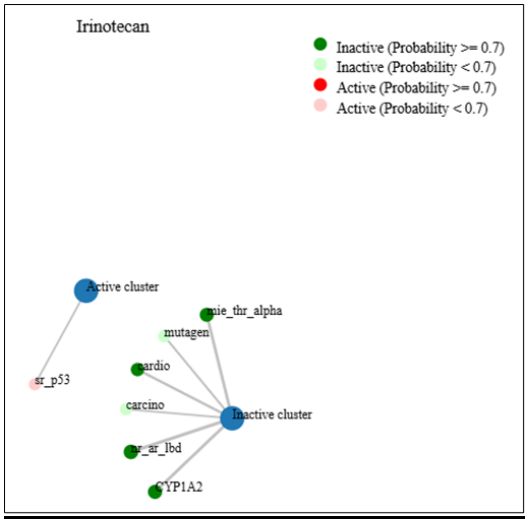


Figure 17: Network Chart for Irinotecan

Doxorubicin

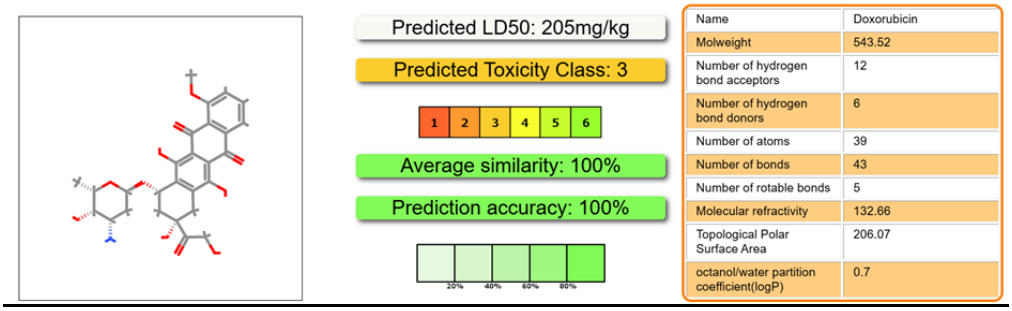


Figure 18: Oral Toxicity Prediction Results for Doxorubicin

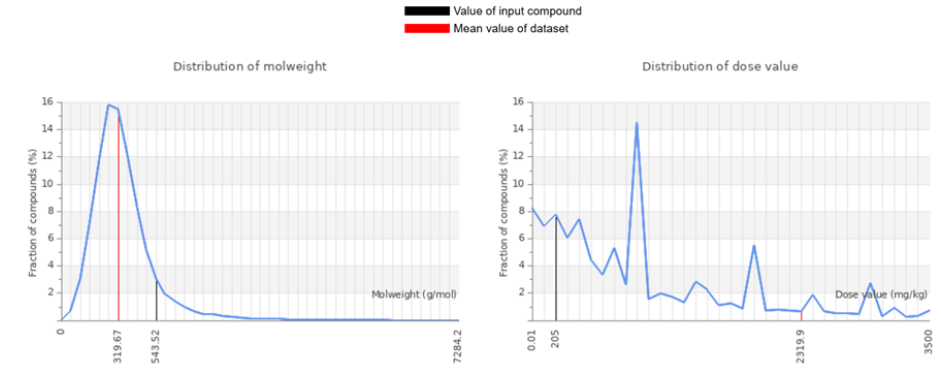


Figure 19: Comparison of Doxorubicin with Dataset Compounds

Table no 4: Toxicity Model Report for Doxorubicin

Classification	Target	Shorthand	Prediction	Probability
Organ toxicity	Cardiotoxicity	cardio	Active	0.64
Toxicity end points	Carcinogenicity	carcino	Inactive	0.90
Toxicity end points	Mutagenicity	mutagen	Active	0.98
Tox21-Nuclear receptor signalling pathways	Androgen Receptor (AR)	nr_ar	Not Calculated	Not Calculated
Tox21-Nuclear receptor signalling pathways	Androgen Receptor Ligand Binding Domain (AR-LBD)	nr_ar_lbd	Inactive	0.55
Tox21-Stress response pathways	Phosphoprotein (Tumor Supressor) p53	sr_p53	Active	0.52
Molecular Initiating Events	Thyroid hormone receptor alpha (THRα)	mie_thr_alpha	Inactive	0.90
Metabolism	Cytochrome CYP1A2	CYP1A2	Inactive	0.99

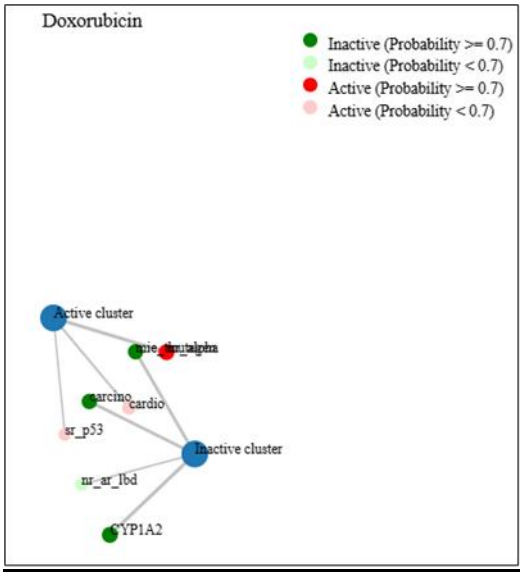


Figure 20: Network Chart for Doxorubicin

The Lethal Dose 50 (LD50), which quantifies the amount of a substance lethal to 50% of a population, inversely correlates with toxicity. Although classified as moderately harmful, Irinotecan exhibits a relatively low potential for acute toxicity due to its higher LD50 compared to the other drugs. Conversely, Doxorubicin, Paclitaxel, Topotecan, and Topotecan Hydrochloride are categorized as harmful with substantially lower LD50 values, suggesting a greater risk of toxicity, particularly at higher dosages. These predictions indicate the relative safety profiles of the drugs, with some showing activity in the p53 pathway, which is crucial for their therapeutic effects as the p53 gene plays a crucial role in cell cycle regulation and apoptosis. However, real-world testing is needed to confirm these in-silico findings.

IV. Discussion

The SBDD and CADD analyses offer a powerful approach to developing novel therapies for HPV-induced cancers. As this research successfully bridges the gap between computational analysis and practical drug discovery, several key areas warrant further investigation and development. To fully realize its potential, further validation of identified drug candidates is crucial, and in vitro and in vivo experiments will confirm their binding efficacy and biological activity, propelling them toward preclinical trials. CADD-identified drug candidates will undergo medicinal chemistry optimization to refine their chemical structures. This aims to enhance binding affinity, selectivity, metabolic stability, and pharmacokinetic profiles, ensuring they are suitable for therapeutic use. Also, broadening this research to encompass other high-risk HPV types, like HPV-16, could yield a comprehensive treatment for multiple HPV strains. Further, the synergy of AI, ML, and CADD can revolutionize drug discovery by boosting the predictive accuracy of molecular interactions, automating the drug design workflow, and expediting the identification of novel drug leads. To ensure the long-term effectiveness of HPV treatments, we must proactively study how the virus might evolve to evade new inhibitors. By understanding these mechanisms, we can design next-generation drugs that are one step ahead of resistance.

Likewise, investigating the potential synergistic benefits of pairing these identified inhibitors with other therapeutic modalities, including immune modulators or conventional chemotherapies, may lead to enhanced treatment outcomes. The findings of this research can contribute to the development of evidence-based public health policies and programs aimed at preventing and treating HPV-related diseases. Overall, through rigorous experimentation, optimization, and integration of cutting-edge technologies, this research has the potential to revolutionize the treatment of HPV-induced cervical cancer, ultimately improving patient prognosis, survival rate, and public health outcomes.

V. Conclusion

This research presents a comprehensive SBDD and CADD analysis to discover novel inhibitors targeting the HPV18 E2 DNA-binding domain. Through the application of advanced computational methodologies, we seek to expedite drug discovery efforts and elucidate the structural and functional aspects of this critical protein-DNA complex.

The analysis leverages a suite of computational tools to conduct the molecular interactions between the E2 protein and its cognate DNA sequence. The homology modeling and structural analysis revealed the

conserved structural features of the HPV18 E2 DNA-binding domain, including the α-helical and β-sheet motifs that are essential for DNA recognition and binding. Molecular docking simulations were performed to identify potential drug-like molecules that could disrupt the protein-DNA interaction. PDB, RasMol, and PyMol enable the visualization and exploration of the protein-DNA complex, facilitating a deep understanding of their structural and functional interplay. RMSD calculations quantified the conformational stability of the protein-DNA complex, while PubChem enabled the identification and evaluation of potential drug candidates. Furthermore, the study incorporated protein structure validation tools, including ERRAT and PROCHECK, to rigorously examine the protein structure for any potential errors or inconsistencies. This validation process ensures the reliability of the computational analysis and, consequently, the effectiveness of the subsequent drug design initiatives.

Through our analysis, we pinpointed critical residues that mediate DNA binding and maintain protein stability. These findings provide a robust framework for designing rational therapeutic interventions. The application of molecular docking techniques enabled the identification of potential drug candidates that could disrupt the protein-DNA interaction, thereby inhibiting viral replication, tumorigenesis, and potentially preventing the onset of cervical cancer. According to the docking results, Irinotecan appeared as the most promising candidate for further investigation as a potential inhibitor of the E2 protein. Using ProTox 3.0, we predicted the toxicity profiles of the identified compounds, ensuring their safety for therapeutic use. By providing rapid and accurate in silico predictions, ProTox 3.0 contributes to the reduction of animal testing and supports the development of safer chemicals and informed regulatory decision-making. However, further experimental validation, such as X-ray crystallography or biophysical assays, is required to determine their precise binding affinity, selectivity, and therapeutic potential.

In summary, this study provides a critical step forward in our understanding of the molecular mechanisms underlying HPV18-induced cervical cancer. By integrating computational and bioinformatics techniques, we have successfully identified potential drug candidates targeting this critical viral oncoprotein and outlined a path to accelerate the discovery of novel anti-HPV therapies. The identification of potent inhibitors through the proposed CADD pipeline represents a promising avenue for the development of targeted therapies that could offer improved efficacy and reduced side effects compared to conventional treatments. Future investigations will delve deeper into the interactions between the E2 protein and cellular factors, and explore the impact of small-molecule inhibitors on viral replication and oncogenesis. Ultimately, these efforts may lead to the development of novel anti-HPV therapies that can prevent cervical cancer and improve the lives of countless women worldwide.

References

[1]. Siegel, R. L., Miller, K. D., Fuchs, H. E., & Jemal, A. (2022). Cancer Statistics, 2022. CA: A Cancer Journal For Clinicians, 72(1), 7–33. Https://Doi.Org/10.3322/Caac.21708

[2]. Bandbe, T., Saikia, J., & Kumari, U. (2025). NGS Analysis Human Papillomavirus Type 18 E2 DNA-Binding Domain Bound To Its DNA Target With Biopython. South Eastern European Journal Of Public Health, XXVI 25 2025, 3781-3792. Https://Doi.Org/10.70135/Seejph.Vi.5856

[3]. Šarenac, T., & Mikov, M. (2019). Cervical Cancer, Different Treatments And Importance Of Bile Acids As Therapeutic Agents In This Disease. Frontiers In Pharmacology, 10, 484. Https://Doi.Org/10.3389/Fphar.2019.00484

[4]. Zur Hausen, H. (2002). Papillomaviruses And Cancer: From Basic Studies To Clinical Application. Nature Reviews Cancer, 2(5), 342–350. Https://Doi.Org/10.1038/Nrc798

[5]. Chung, S.-H., Franceschi, S., & Lambert, P. F. (2010). Estrogen And Erα: Culprits In Cervical Cancer? Trends In Endocrinology & Metabolism, 21(8), 504–511. Https://Doi.Org/10.1016/J.Tem.2010.03.005

[6]. Sun, D., Gao, W., Hu, H., & Zhou, S. (2022). Why 90% Of Clinical Drug Development Fails And How To Improve It? Acta Pharmaceutica Sinica B, 12(7), 3049–3062. Https://Doi.Org/10.1016/J.Apsb.2022.02.002

[7]. Niazi, S. K., & Mariam, Z. (2023). Computer-Aided Drug Design And Drug Discovery: A Prospective Analysis. Pharmaceuticals, 17(1), 22. Https://Doi.Org/10.3390/Ph17010022

[8]. Lionta, E., Spyrou, G., Vassilatis, D., & Cournia, Z. (2014). Structure-Based Virtual Screening For Drug Discovery: Principles, Applications And Recent Advances. Current Topics In Medicinal Chemistry, 14(16), 1923–1938. Https://Doi.Org/10.2174/1568026614666140929124445

[9]. Kumari, U., & Parashar, R. (2025). Structure Analysis And Molecular Docking Of Mesothelin-207 Fragment In Human Cancer. International Journal Of Emerging Technologies And Innovative Research (JETIR), ISSN:2349-5162, 12(3), G98-G106. Http://Doi.One/10.1729/Journal.44261

[10]. Kumari, U. & Agrawal, N. (2023). NGS And Mutational Profile Analysis Of Non-Small-Cell Lung Carcinoma (NSCLC). International Journal For Research In Applied Science And Engineering Technology (Ijraset), ISSN:2321-9653, 11(IV), 3090-3094. Https://Doi.Org/10.22214/Ijraset.2023.50880

[11]. Kumari, U., & Gupta, S. (2023). NGS And Sequence Analysis With Biopython For Prospective Brain Cancer Therapeutic Studies. International Journal For Research In Applied Science And Engineering Technology (Ijraset), ISSN: 2321-9653, 11(IV), 3318-3329. Https://Doi.Org/10.22214/Ijraset.2023.50885

[12]. Kumari, U., & Bajaj, T. (2025). Ngs Data Analysis And Active Site Identification Alpha-1-Acid Glycoprotein Bound To Potent Anti Tumor Compound UCN-01 In Malignant Brain Tumor. International Journal Of Emerging Technologies And Innovative Research (JETIR), ISSN:2349-5162, 12(3), G107-G116. Http://Doi.One/10.1729/Journal.44259

[13]. Kumari, U., & Mehrotra, K. (2025). NGS And Proteomic Gene Expression Analysis In Studies Of NUDT5 Silence Hormone Signaling In Breast Cancer. International Journal Of Emerging Technologies And Innovative Research (JETIR), ISSN:2349-5162, 12(2), C630-C638. Http://Doi.One/10.1729/Journal.43563

[14]. Sayle, R. (1995). RASMOL: Biomolecular Graphics For All. Trends In Biochemical Sciences, 20(9), 374–376. Https://Doi.Org/10.1016/S0968-0004(00)89080-5
[15]. Kumari, U., Bandbe, T., & Brahma, D. (2024, September 1). Structure-Based Drug Design And Molecular Target Prediction Of Transthyretin Variant A97S Complexed. International Journal Of Emerging Technologies And Innovative Research (JETIR), ISSN:2349-5162, 11(9), A316-A326. Http://Doi.One/10.1729/Journal.41352
[16]. Kumari, U., Johri, V., Dhopate, S., & Jha, T. (2024). Structure Based Drug Designing For The Prediction Of Epitope For Targeting Malignant Brain Tumor. International Journal Of Emerging Technologies And Innovative Research (JETIR), ISSN-2349-5162, 11(7), B689-B697. Http://Doi.One/10.1729/Journal.40433
[17]. Colovos, C., & Yeates, T. O. (1993). Verification Of Protein Structures: Patterns Of Nonbonded Atomic Interactions. Protein Science, 2(9), 1511–1519. Https://Doi.Org/10.1002/Pro.5560020916
[18]. Laskowski, R. A., Macarthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: A Program To Check The Stereochemical Quality Of Protein Structures. Journal Of Applied Crystallography, 26(2), 283–291. Https://Doi.Org/10.1107/S0021889892009944
[19]. Kumari, U., Verma, G., & Anbumegala M. (2025). IN SILICO DRUG DESIGN IN HUMAN APOPTOSIS INDUCING FACTOR (AIF) IN LUNG CANCER. International Journal Of Emerging Technologies And Innovative (JETIR), ISSN:2349-5162, 12(2), D139-D148. Http://Doi.One/10.1729/Journal.43606
[20]. Chaudhary, S. & Kumari, U. (2024). NGS, MOLECULAR DOCKING AND NETWORK PHARMACOLOGY REVEAL POTENT INHIBITOR FOR THE TREATMENT OF LUNG CANCER. International Journal Of Emerging Technologies And Innovative Research (JETIR), ISSN:2349-5162, 11(9), F116-F126. Https://Doi.One/10.1729/Journal.41696
[21]. Banerjee, P., Eckert, A. O., Schrey, A. K., & Preissner, R. (2018). Protox-II: A Webserver For The Prediction Of Toxicity Of Chemicals. Nucleic Acids Research, 46(W1), W257–W263. Https://Doi.Org/10.1093/Nar/Gky318