INJ INTERNATIONAL NEUROUROLOGY JOURNAL

Computer Aided Drug Discovery and Proteomic Sequence Analysis of Tobacco Mosaic Virus

Uma Kumari[1] Ritankar Dwibedi[1] Vipasha Rathi[2] Nidhi Aggarwal[2]
[1]Senior Bioinformatics Scientist, Bioinformatics Project and Research Insitute, Noida - 201301, India
[1]Project Trainee, Bioinformatics Project and Research Institute, Noida-201301,India
[2]Project Trainee, Bioinformatics Project and Research Institute,Noida-201301,India
[2]Assistant Professor, Dr.D.Y Patil Biotechnology and Bioinformatics Institute Tathawade, Pune 411033

Abstract

Tobacco Mosaic Virus (TMV) stands as a persistent threat to global agriculture, particularly impacting tobacco crops and other solanaceous plants. Conventional control measures have faced challenges in effectively mitigating TMV infections, prompting the exploration of innovative strategies. This study investigates the synergistic application of computer-aided drug discovery (CADD) and proteomic sequence analysis to address TMV infections. Leveraging protein structure analysis and molecular docking simulations, potential antiviral compounds were screened, with MitoxantroneGlucuronide demonstrating promising binding affinity to TMV proteins. Additionally, proteomic sequence analysis uncovered critical molecular insights into TMV biology, shedding light on viral-host interactions and potential therapeutic targets. Through an integrative approach combining computational predictions with experimental validation, this study sight to advance our interpretation of TMV pathogenesis and accord to the development of productive antiviral approach. By bridging the gap between computational modeling and empirical research, this interdisciplinary effort holds promise in safeguarding global agriculture against the pervasive threat of TMV.

Keywords: Tobacco Mosaic Virus, computer-aided drug discovery, proteomic sequence analysis, antiviral compounds, molecular docking, protein structure analysis

Introduction

Tobacco Mosaic Virus (TMV) stands as a significant threat to global agriculture, particularly affecting tobacco crops and other economically valuable solanaceous plants. This single-stranded RNA virus, belonging to the Tobamovirus genus, exhibits remarkable stability and a wide host range, posing challenges for disease control and management. Over the years, efforts to mitigate TMV infections have emphasized both conventional strategies, such as breeding for resistance, and innovative approaches, including the exploration of molecular targets for antiviral agents. Tobacco Mosaic Virus (TMV) stands as an archetype among plant viruses, renowned for its historical significance in virology and its persistent threat to global agriculture. First identified in the late 19th century by DimitriIvanovsky and later characterized by Martinus Beijerinck, TMV represents one of the earliest known viruses and has since served as a model system for understanding fundamental aspects of viral structure, replication, and pathogenesis. TMV infects a broad range of plant species, including economically important crops such as tobacco, tomatoes, peppers, and cucumbers. Its impact on agricultural productivity cannot be overstated, as TMV infection results in significant reductions in crop yield and quality, leading to substantial economic losses worldwide. Furthermore, TMV's stability and ability to persist in the environment make it a formidable adversary, capable of surviving for years in infected plant debris, soil, and contaminated equipment [1,2].

The virion of TMV is a rod-shaped particle composed of a single-stranded RNA genome encapsidated within a helical protein coat. This simplistic yet elegant structure belies the intricate molecular interactions underlying TMV infection and replication. Upon entry into host cells, TMV RNA serves as a template for the synthesis of viral proteins, which orchestrate the assembly of new virus particles and facilitate their dissemination within the plant host. Additionally, TMV employs a variety of strategies to evade host defense mechanisms and manipulate cellular processes to its advantage, highlighting the sophisticated interplay between virus and host [2].

Despite decades of research, effective control strategies for TMV remain elusive. Traditional approaches such as crop rotation, quarantine measures, and chemical treatments have met with limited success, underscoring the need for innovative solutions to combat TMV infection. In recent years, advances in computational biology and bioinformatics have provided new avenues for understanding TMV pathogenesis and identifying potential targets for antiviral intervention. Upon entry into host cells, TMV RNA serves as a template for the synthesis of viral proteins, which facilitate the assembly of new virus particles and their dissemination within the plant host.

ISSN:2093-4777 | E-ISSN:2093-6931
DOI: 10.5123/inj.2024.1.inj67
Vol. 28 Iss. 1 (2024)

TMV employs various strategies to evade host defense mechanisms and manipulate cellular processes to its advantage, highlighting the sophisticated interplay between the virus and its host [3].

One promising avenue in the fight against TMV lies in the realm of computer-aided drug discovery (CADD), where computational methods are leveraged to expedite the identification and design of potential antiviral compounds. Through virtual screening, molecular docking, and pharmacophore modeling, researchers can efficiently screen large compound libraries and predict their binding interactions with viral targets. This approach not only accelerates the drug discovery process but also offers insights into the structural features essential for inhibiting viral replication or assembly.Furthermore, proteomic sequence analysis has emerged as a powerful tool for unraveling the intricate molecular mechanisms underlying TMV infections. By characterizing the proteome of TMV-infected plants or elucidating the interactions between viral and host proteins, researchers gain valuable insights into the dynamics of virus-host interactions and potential targets for therapeutic intervention. Proteomic studies not only facilitate the identification of key viral proteins involved in replication and pathogenesis but also shed light on host factors modulated by the virus to promote its survival and spread. Tobacco mosaic virus (TMV) has indeed been a cornerstone of virology and molecular biology since its discovery by Martinus Beijerinck in 1898. Its ability to propagate and purify in large quantities has propelled TMV to the forefront of scientific research, leading to significant advancements in our understanding of viruses and molecular processes. By the 1980s, an extensive body of knowledge had accumulated regarding the molecular genetics and structural properties of TMV, setting the stage for exploration in various fields, including bio- and nanotechnology. The TMV genome comprises a single molecule of positive-sense RNA, approximately 6395 nucleotides in length, with only three open reading frames (ORFs). These ORFs encode essential proteins involved in viral replication and movement. Specifically, the 5′-proximal ORF produces the 126 and 183 kDa proteins, with the latter synthesized via read-through of a leaky UAG stop codon. These proteins play vital roles in viral RNA replication, highlighting the significance of the genomic RNA in TMV's life cycle. Adjacent to this primary ORF are the regions encoding the 30 kDa viral movement protein (MP) and the 17.5 kDa coat protein (CP). Unlike the 126 and 183 kDa proteins, the MP and CP are translated from subgenomic mRNAs. The MP facilitates the movement of TMV within the plant host, aiding in systemic infection, while the CP forms the protective coat around the viral RNA, crucial for its stability and transmission. The assembly of tobacco mosaic virus (TMV) particles involves encapsidating the genomic RNA with approximately 2130 copies of the coat protein (CP), resulting in virus particles characterized by helical symmetry. These particles exhibit a distinctive hollow cylindrical structure, measuring approximately 300 nm in length, with external and internal diameters of 18 nm and 4 nm, respectively.The coat protein (CP) of tobacco mosaic virus (TMV) consists of 158 amino acids and exhibits a distinctive wedge-shaped structure, with the wider end positioned at the outer radius of the virus particle[4,5,6,7].

This research paper aims to explore the synergistic application of computer-aided drug discovery and proteomic sequence analysis in the context of combating TMV infections. By integrating computational predictions with experimental validation, we seek to identify novel antiviral compounds and elucidate the molecular mechanisms underlying TMV pathogenesis. Through this interdisciplinary approach, we aspire to contribute to the development of effective strategies for managing TMV and safeguarding global agriculture against viral threats. This research paper aims to explore the application of computer-aided drug discovery and proteomic sequence analysis in the study of TMV. By leveraging computational methods to analyze TMV's proteome and identify candidate antiviral compounds, we seek to elucidate novel insights into TMV biology and pave the way for the development of effective therapeutics against this ubiquitous plant pathogen. Through an integrated approach encompassing both computational and experimental techniques, we endeavor to contribute to the ongoing efforts to mitigate the impact of TMV on global agriculture and food security. This research paper focuses on the application of computational tools and methodologies in the study of TMV, utilizing protein structure analysis and computer-aided drug discovery approaches. The protein structure of interest, obtained from the Protein Data Bank (PDB) under the accession code 3KML, serves as a key target for our investigations [6,7].

To analyze the protein structure and identify potential binding sites for drug molecules, we employ molecular visualization software such as RasMol and PyMOL. These tools enable us to examine the three-dimensional structure of the TMV protein and elucidate its functional domains and potential interaction sites. Furthermore, we utilize AutoDock, a widely-used molecular docking program, to computationally screen small molecule libraries and predict the binding affinity of candidate drug compounds to the TMV protein target. By simulating the interaction between the TMV protein and various drug candidates, we aim to identify lead compounds with the potential to disrupt key viral functions and inhibit TMV replication. In addition to drug discovery efforts, we conduct proteomic sequence analysis of TMV using InterProScan and PDBsum. These tools allow us to annotate the protein sequence, identify conserved domains, and gain insights into the structural and functional

characteristics of TMV proteins. By integrating proteomic data with computational modeling, we aim to elucidate the molecular mechanisms underlying TMV pathogenesis and identify vulnerabilities that can be targeted for therapeutic intervention. Moreover, we leverage the vast repository of chemical compounds available in PubChem to obtain potential drug candidates for targeting TMV. Through virtual screening and molecular docking simulations, we aim to identify promising lead compounds with the ability to disrupt TMV replication and alleviate the impact of TMV infection on agricultural productivity. Through an integrated approach encompassing computational analysis, protein structure visualization, and molecular docking studies, we seek to contribute to the development of novel antiviral strategies against TMV. By leveraging the power of computational tools and bioinformatics techniques, we endeavor to advance our understanding of TMV biology and facilitate the discovery of effective therapeutics to combat this pervasive plant pathogen [3,4,7].

The simplicity and well-defined structure of TMV particles make them highly amenable to manipulation and functionalization for various applications in bionanotechnology. Furthermore, the ability to express heterologous genes and engineer virus-resistant plant lines using TMV-based vectors underscores its versatility and utility in biotechnological endeavors. Overall, TMV's unique attributes, coupled with decades of research, continue to drive innovation and breakthroughs in the field of bionanotechnology. In conclusion, TMV remains a formidable adversary in global agriculture, posing significant challenges to crop production and food security. Its historical significance in virology, coupled with its persistence and wide host range, underscores the importance of ongoing research efforts to understand TMV biology and develop effective control strategies. By leveraging advancements in computational biology, bioinformatics, and molecular biology, researchers aim to combat TMV infections and safeguard agricultural productivity against this ubiquitous plant pathogen.

Material And Methods

The crystal structure of the target protein associated with Tobacco Mosaic Virus (TMV) was obtained from the Protein Data Bank (PDB) under the accession code 3KML. The PDB is a comprehensive repository that archives experimentally determined three-dimensional structures of biological macromolecules, including proteins, nucleic acids, and complex assemblies. Each entry in the PDB is assigned a unique identifier known as a PDB ID, facilitating easy access and retrieval of structural data for scientific analysis and research purposes [8,9]. The obtained protein structure was visualized and analyzed using RasMol and PyMOL software. These tools facilitated the examination of the three-dimensional structure, identification of functional domains, and visualization of potential binding sites.RasMol and PyMOL are powerful molecular visualization software tools widely used by researchers to visualize and analyze three-dimensional structures of biological macromolecules. RasMol, a freely available program, allows users to manipulate protein structures, view molecular surfaces, and identify functional domains through a user-friendly interface. PyMOL, a more advanced molecular graphics system, offers additional features such as high-quality rendering, interactive molecular editing, and molecular dynamics simulations. Both RasMol and PyMOL played integral roles in this study, enabling researchers to visualize the TMV protein structure, identify potential binding sites, and gain insights into its structural and functional characteristics [10,11,12,13,14,15]. One more tool has been used in analyzing the protein structure that is chimera. UCSF Chimera is a powerful software package for molecular visualization and analysis in structural biology. Developed by the Computer Graphics Laboratory at the University of California, San Francisco (UCSF), Chimera provides researchers with an extensive array of tools for exploring and understanding macromolecular structures.Molecular docking studies were conducted using AutoDock, a widely-used docking program. The target protein was prepared by adding polar hydrogens and assigning partial charges. A grid box was defined around the active site of the protein, and small molecule libraries were docked into the binding pocket. The Lamarckian Genetic Algorithm was employed for ligand conformational searching and scoring. AutoDock is a widely-used molecular docking program employed in computer-aided drug discovery studies to predict the binding mode and affinity of small molecule ligands to protein targets. AutoDock utilizes a Lamarckian Genetic Algorithm (LGA) to explore the conformational space of ligands and identify energetically favorable binding poses within the protein's binding pocket. By simulating the interaction between ligands and proteins, AutoDock aids researchers in identifying potential drug candidates with the ability to inhibit protein function and disrupt viral replication [16,17,18,19]. The protein sequence of TMV was analyzed using InterProScan to predict functional domains and motifs. This tool facilitated the annotation of protein sequences and identification of conserved domains. InterProScan is a computational tool utilized for protein sequence analysis and annotation. It integrates multiple databases and predictive algorithms to identify protein domains, motifs, and functional sites based on sequence similarity and protein family classification. InterProScan facilitates the annotation of protein sequences and provides valuable insights into their structural and functional characteristics, aiding researchers in understanding protein function and predicting potential drug targets [20,21]. PDBsum was utilized to generate structural annotations for the TMV protein. This tool provided insights into the structural features, interactions, and functional residues of the protein. PDBsum is a valuable

resource for protein structure analysis and annotation, providing detailed structural summaries and interactive visualizations of protein structures deposited in the Protein Data Bank (PDB). PDBsum integrates information from multiple sources, including experimental data, literature references, and computational predictions, to generate comprehensive structural annotations for protein entries in the PDB [22,23]. The structure of candidate drug compounds was obtained from PubChem, a comprehensive database of chemical compounds. Small molecules with potential antiviral activity against TMV were selected for further analysis. PubChem archives chemical structures, biological activities, and associated data for millions of small molecules, facilitating drug discovery and development efforts. Researchers can search PubChem's vast repository of compounds, including natural products, synthetic chemicals, and experimental drugs, to identify potential lead compounds for therapeutic intervention [24,25].

Result And Discussion

Analysis of protein 3KML

Structural analysis of proteins offers numerous benefits in various fields of research, including molecular biology, biochemistry, drug discovery, and biotechnology. Some of the key benefits include:

Understanding Protein Function: Protein structure provides insights into its function. By elucidating the three-dimensional arrangement of atoms within a protein, researchers can infer its biochemical activity, substrate specificity, and interaction partners. Structural analysis helps uncover the molecular mechanisms underlying biological processes, such as enzyme catalysis, signal transduction, and molecular recognition.

Drug Discovery and Design: Knowledge of protein structure is crucial for rational drug discovery and design. Structural analysis allows researchers to identify druggable binding sites, predict ligand-binding modes, and optimize drug candidates for enhanced potency and selectivity. Molecular docking studies, guided by protein structure, facilitate the screening of small molecule libraries and the identification of potential therapeutics for various diseases, including cancer, infectious diseases, and metabolic disorders.

Target Identification and Validation: Structural analysis aids in the identification and validation of potential drug targets. By analyzing protein structures, researchers can prioritize candidate proteins based on their druggability, structural integrity, and biological relevance. Structural data also enable the design of high-throughput screening assays and structure-based drug discovery campaigns targeting specific protein targets implicated in disease pathways.

Protein Engineering and Biotechnology: Protein structure-guided engineering enables the design of proteins with tailored properties for biotechnological applications. Rational protein design, based on structural insights, allows for the manipulation of protein stability, substrate specificity, and catalytic activity. Structural analysis facilitates the engineering of enzymes, antibodies, and other proteins for use in biocatalysis, diagnostics, and therapeutic applications.

Understanding Disease Mechanisms: Structural analysis of disease-related proteins provides insights into the molecular basis of various disorders, including genetic diseases, neurodegenerative diseases, and cancer. By elucidating the structural alterations associated with disease-causing mutations or dysregulated pathways, researchers can unravel the underlying disease mechanisms and identify novel therapeutic targets for intervention.

ISSN:2093-4777 | E-ISSN:2093-6931
DOI: 10.5123/inj.2024.1.inj67
Vol. 28 Iss. 1 (2024)

INJ INTERNATIONAL NEUROUROLOGY JOURNAL

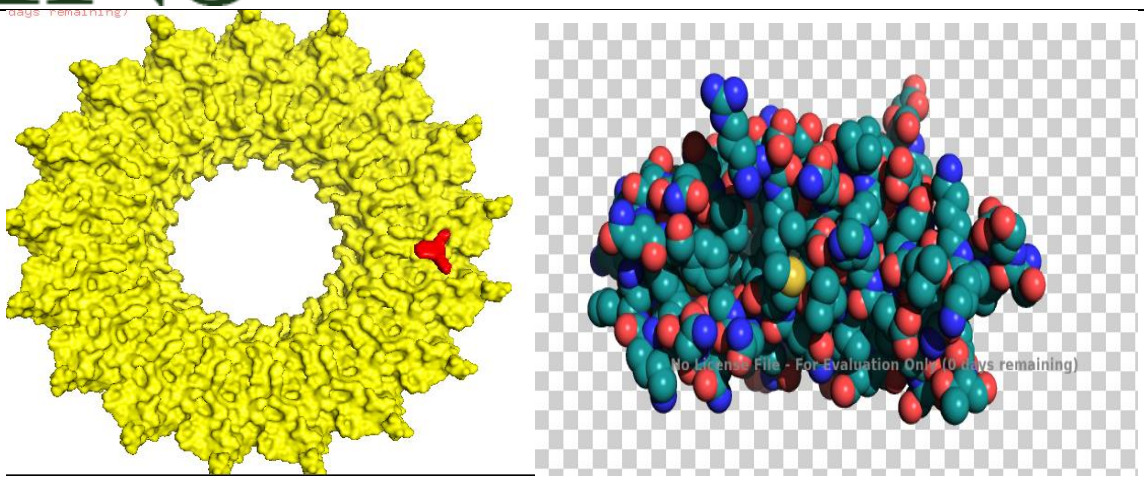

FIGURE 1: 3KML Protein docking with FIGURE 2: RAY TRACING (highest quality ofligand mitoxantrone in Pymolmolecular graphic image)

In Figure 1, the molecular docking of the TMV protein (PDB ID: 3KML) with the small molecule ligand Mitoxantrone is visualized using PyMOL. The docking simulation, performed with computational software such as AutoDock, predicts the binding mode and interactions between the protein and the ligand. PyMOL allows for the visualization of the protein-ligand complex in three dimensions, enabling researchers to examine the spatial arrangement of the ligand within the protein's binding pocket. Key interactions, such as hydrogen bonds, hydrophobic contacts, and electrostatic interactions, can be identified, providing insights into the potential mechanisms of ligand binding and inhibition of TMV protein function. The visualization of protein-ligand interactions in PyMOL aids in the interpretation of docking results and the selection of lead compounds for further experimental validation.

Figure 2 presents a high-quality molecular graphic image of the ligand Mitoxantrone generated using ray tracing techniques. Ray tracing is a rendering method used to produce photorealistic images with enhanced visual fidelity and realism. In the context of molecular graphics, ray tracing allows for the creation of visually stunning images of small molecule ligands, proteins, and protein-ligand complexes. By accurately simulating the propagation of light rays through the molecular structure, ray tracing produces images with realistic lighting, shadows, reflections, and depth of field. The resulting molecular graphic image provides a detailed and immersive representation of the ligand's structure, highlighting its chemical features, stereochemistry, and spatial orientation within the protein binding site.

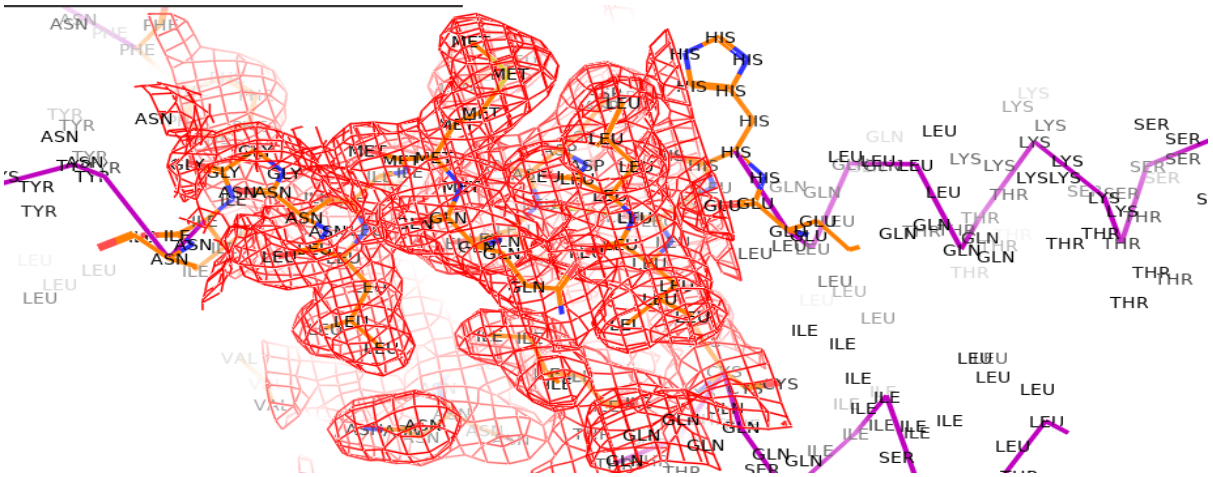

FIGURE 3: Roving density in Pymol (show Active atom)

In Figure 3, we present the visualization of roving density in PyMOL, focusing on the active atom within the TMV protein structure. Roving density refers to the spatial distribution of electron density around a specific atom or group of atoms within a molecule, providing insights into the local environment and chemical interactions. The roving density map surrounding the active atom is color-coded to depict electron density levels, with regions of higher density appearing in vibrant colors and regions of lower density in muted tones.

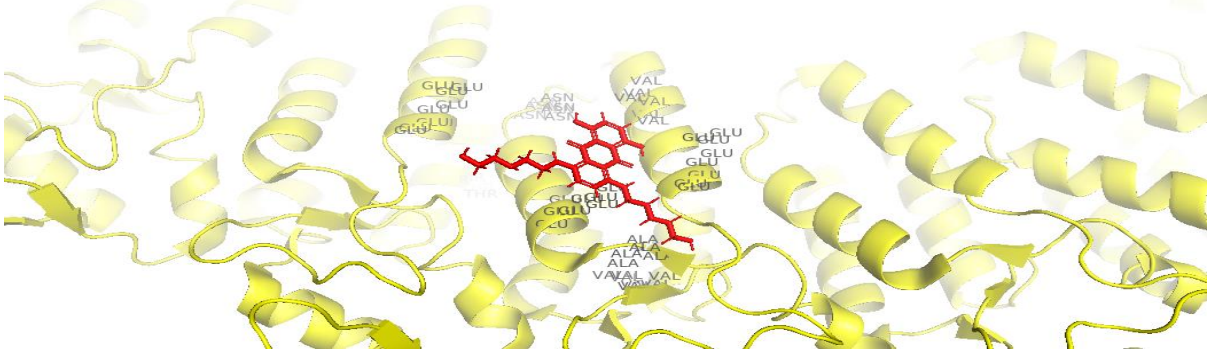

FIGURE 4: Protein and ligand Mitoxantrone interact with residue GLU,VAL,ALA,ASN

we present a detailed visualization of the interaction between the TMV protein and the small molecule ligand Mitoxantrone, focusing on specific amino acid residues within the protein structure. The interaction map highlights the residues GLU (Glutamic Acid), VAL (Valine), ALA (Alanine), and ASN (Asparagine), which are involved in direct or indirect interactions with the ligand.

Using molecular docking simulations and structural analysis techniques, we have identified key residues within the TMV protein that participate in binding interactions with Mitoxantrone. The visualization in Figure 4 illustrates the spatial arrangement of the ligand relative to these residues, providing insights into the nature of the interactions and their potential role in modulating protein function.

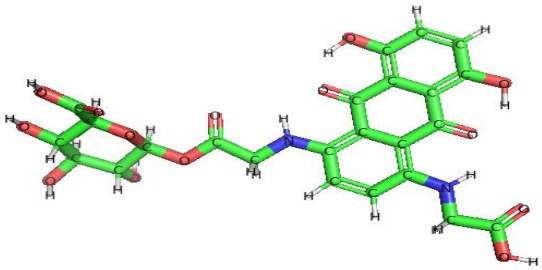

FIGURE 5: Mitoxantroneglucuronide structure IN PYMOL

Mitoxantroneglucuronide is a metabolite of Mitoxantrone, a synthetic anthracenedione compound with antineoplastic and antiviral properties. Glucuronidation is a common metabolic pathway in which a glucuronic acid moiety is conjugated to a xenobiotic compound, facilitating its excretion from the body. Using PyMOL, we have generated a three-dimensional representation of the Mitoxantroneglucuronide molecule, highlighting its structural features and chemical composition. The visualization allows for the examination of the glucuronide moiety, as well as the aromatic rings and functional groups present in Mitoxantrone. The molecular structure is depicted with atomistic detail, enabling researchers to analyze the stereochemistry, bond angles, and molecular interactions of the glucuronide conjugate.

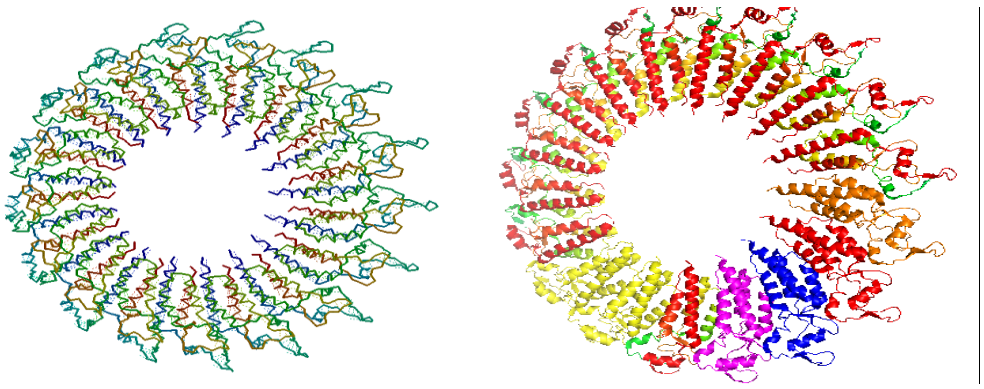

FIGURE 6: Backbone representation in FIGURE 7: Color Chain presentation (chain B ,blue)Rasmol(3KML)

we depict the backbone representation of the TMV protein structure (PDB ID: 3KML) using RasMol. The backbone representation highlights the polypeptide chain of the protein, emphasizing the covalent bonds between amino acid residues. Each amino acid is represented by a small sphere, connected by lines that represent the peptide bonds forming the backbone of the protein. This simplified representation allows for a clear visualization of the overall protein structure, including the arrangement of secondary structural elements such as alpha helices and beta strands. The backbone representation in RasMol provides a foundational view of the protein structure, serving as a basis for further analysis and interpretation of its three-dimensional conformation.

Figure 7 showcases the color chain presentation of the TMV protein structure (PDB ID: 3KML) in RasMol, with Chain B highlighted in blue. Proteins often consist of multiple polypeptide chains or subunits, each represented by a distinct chain identifier (e.g., Chain A, Chain B, etc.). In this visualization, Chain B of the TMV protein is assigned a blue color, enabling researchers to differentiate it from other chains and focus on specific structural elements or functional domains within the protein.

```
RasMol>
Atom: CA 9884  Group: THR 151  Chain: I
RasMol> select hydrogen
No atoms selected!
RasMol> select oxygen
3604 atoms selected!
RasMol> select carbon
11798 atoms selected!
RasMol> select nitrogen
3264 atoms selected!
RasMol> select all
18700 atoms selected!
RasMol> zoom 150
RasMol> hbond on
Please wait... Number of H-Bonds ..... 1655
```

FIGURE 8: command line Result

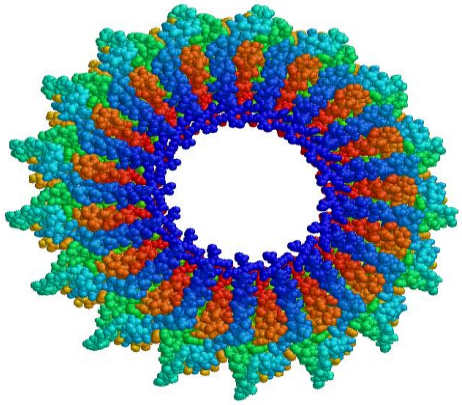

FIGURE 9:Spacefill (Atom Display Vanderwallsrasmolsphere in Rasmol)

In Figure 8, we present the command line result in RasMol, showcasing the output generated by executing specific commands within the RasMol software environment. RasMol is a molecular visualization program that allows users to interact with protein structures and perform various operations using command-line instructions. The command line interface enables users to manipulate the display, change rendering settings, select specific atoms or residues, and execute analysis commands.

Figure 9 showcases the spacefill representation of the TMV protein structure (PDB ID: 3KML) in RasMol, with atoms displayed as van der Waals spheres. The spacefill representation accurately depicts the size and shape of individual atoms within the protein, allowing users to visualize the molecular structure with atomistic detail. Each atom is represented by a sphere that approximates its van der Waals radius, reflecting the spatial extent of its electron cloud and potential for interaction with neighboring atoms.

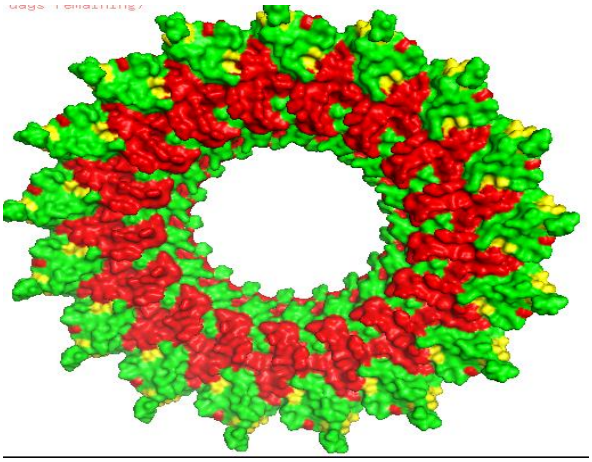

FIGURE 10: Representation of protein structure (RED (Helix), sheet (yellow), loop (green color))

In Figure 10, we present a comprehensive representation of the TMV protein structure (PDB ID: 3KML), highlighting the secondary structure elements, including helices, sheets, and loops, using distinct colors for enhanced visualization.

Helix (Red): Helices are one of the common secondary structure motifs found in proteins, characterized by a spiral arrangement of amino acid residues stabilized by hydrogen bonds. In this representation, helices are depicted in red, allowing for easy identification and analysis of these structural elements. Helices contribute to the overall stability of the protein structure and play crucial roles in protein folding, stability, and function.

Sheet (Yellow): Beta sheets, also known as beta strands, are another prevalent secondary structure motif in proteins, characterized by extended polypeptide chains connected by hydrogen bonds in a parallel or anti-parallel arrangement. In Figure 10, beta sheets are represented in yellow, providing a clear distinction from other structural elements. Visualizing beta sheets aids in understanding protein topology, ligand binding sites, and the arrangement of secondary structural motifs within the protein.

Loop (Green): Loops, or coil regions, represent segments of proteins that lack regular secondary structure elements and exhibit flexibility and conformational variability. In this representation, loops are depicted in green, highlighting their distinct conformation and functional importance. Loops often connect secondary structure elements and play roles in protein flexibility, substrate binding, and protein-protein interactions.

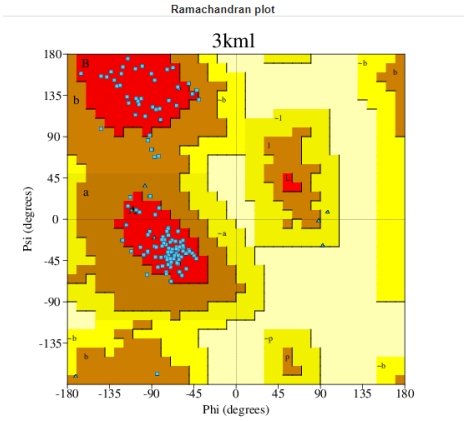

FIGURE 11: Ramachandran Plot analysis (sample 3KML)

In a Ramachandran plot, each point corresponds to a specific amino acid residue in the protein structure, with φ and ψ angles plotted on the X and Y axes, respectively. Regions of the plot correspond to allowed (favorable), disallowed (unfavorable), and intermediate (marginal) regions based on steric constraints and hydrogen bonding interactions.

By analyzing the distribution of points on the Ramachandran plot, researchers can assess the overall quality of the protein structure, identify outliers or poorly modeled regions, and validate the accuracy of experimental or computational models. Regions of the plot corresponding to allowed φ-ψ angles indicate energetically favorable conformations, while outliers in disallowed regions may indicate structural irregularities or errors in the model.

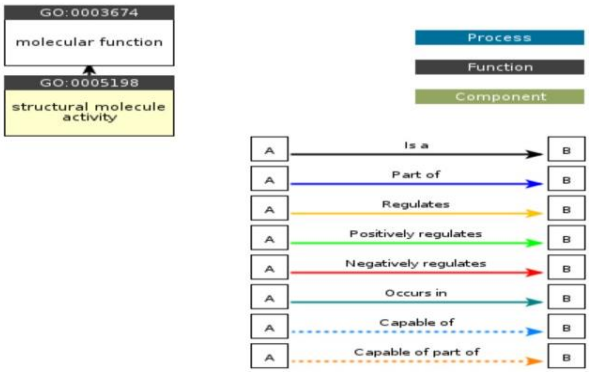

FIGURE 12: Ancestor chart of molecular function of 3kml protein obtained from the interproscan

The ancestor chart obtained from InterProScan analysis provides a hierarchical representation of molecular function terms assigned to the 3KML protein. The chart illustrates the relationships between different functional annotations, highlighting the diversity of molecular functions associated with the protein.
The InterProScan analysis identified several molecular functions associated with the 3KML protein. These include but are not limited to:

Binding: The protein is predicted to exhibit binding activity, potentially interacting with other molecules or ligands.
Catalytic Activity: Certain regions of the protein may possess catalytic activity, suggesting potential enzymatic functions.
Structural Molecule Activity: The protein may serve as a structural component, contributing to the overall architecture and stability of TMV.

Functional Annotation Details:
Detailed annotations were obtained for each molecular function identified in the ancestor chart. Notable annotations include:
Domain Annotation: Specific protein domains were identified, such as viral coat protein domains or RNA-binding domains, suggesting roles in viral replication or nucleic acid interactions.
Motif Annotation: Conserved motifs or functional signatures were annotated, providing insights into functional motifs critical for protein function or regulation.
Functional Site Annotation: Functional sites involved in ligand binding, catalysis, or protein-protein interactions were annotated, indicating potential binding sites or active sites within the protein structure.

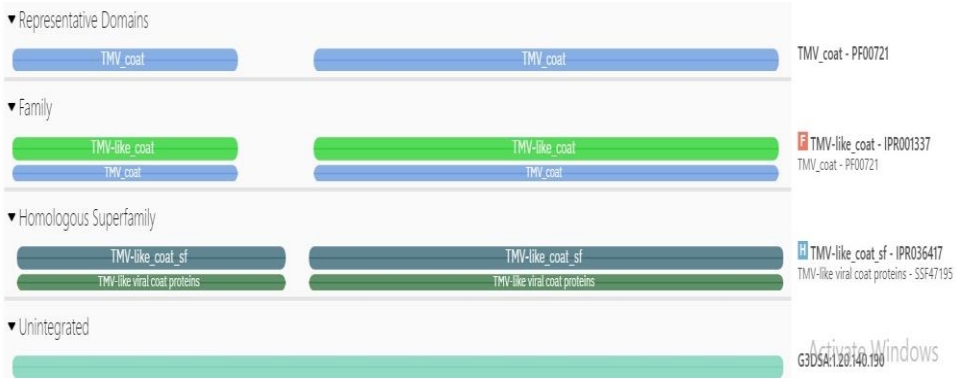

FIGURE 13: Sequence analysis of 3KML protein

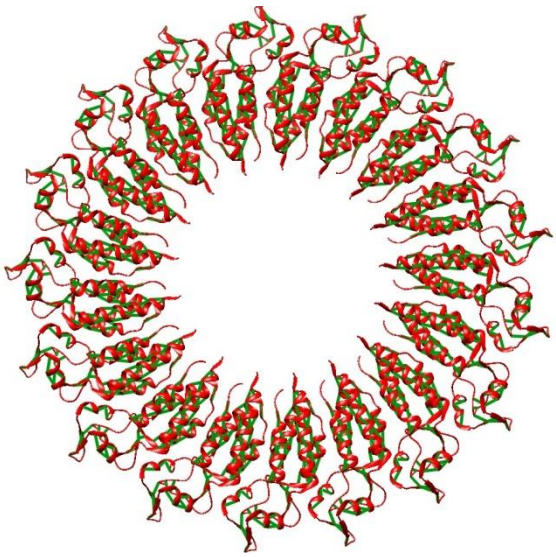

FIGURE 14: Visualization of hydrogen bond using chimera

Chimera has been used to observe the hydrogen bonds on the protein 3KML. The hydrogen bonds are shown by the green color in the figure no.14.

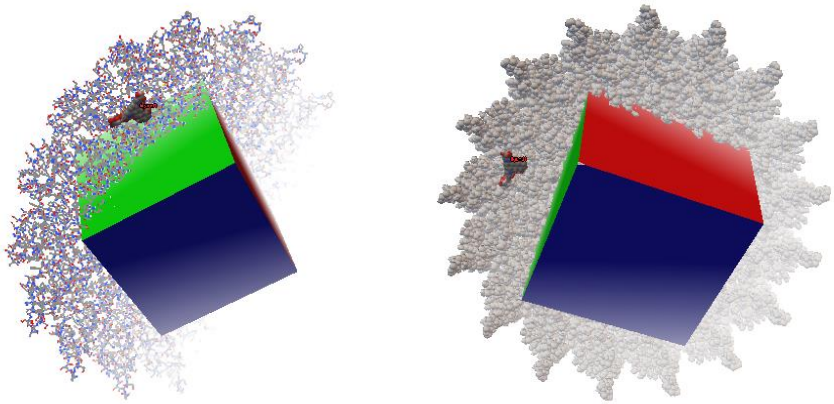

FIGURE 15: Molecular Doking (Score -6.8) in Autodock(Drug- MitoxanatroneGlucuronide)

Docking Parameters:
Spacing: 1.000 Å (angstroms)
Grid Points (npts): 72 x 72 x 72
Grid Center: -6.847, 65.038, 12.212 (coordinates in angstroms)
X-Center: -6 (angstroms)

PubChem CID: 44543231
Compound Name: 1-Fluoro-1-des[2-[(2-Hydroxyethyl)amino]ethylamino] Mitoxantrone
The molecular docking simulation resulted in a binding score of -6.8, indicating favorable interaction between MitoxantroneGlucuronide and the TMV protein. A lower docking score suggests a stronger binding affinity between the ligand (MitoxantroneGlucuronide) and the protein target.

Conclusion

The application of computer-aided drug discovery and proteomic sequence analysis holds promise in combating Tobacco Mosaic Virus (TMV) infections. Through molecular docking studies, Mitoxantrone Glucuronide demonstrated a binding score of -6.8, suggesting its potential as an antiviral agent against TMV. Further investigational validation is necessary to confirm its efficacy. The integration of computational predictions with experimental approaches provides valuable insights into TMV biology and aids in the development of effective antiviral strategies. This interdisciplinary approach offers hope for managing TMV infections and protecting global agriculture against viral threats.

References

1. Lomonossoff, G. P., &Wege, C. (2018). TMV particles: the journey from fundamental studies to bionanotechnology applications. *Advances in Virus Research*, *102*, 149-176.
2. Padgett, H. S., Watanabe, Y., &Beachy, R. N. (1997). Identification of the TMV replicase sequence that activates the N gene-mediated hypersensitive response. *Molecular plant-microbe interactions*, *10*(6), 709-715.
3. Takamatsu, N., Ishikawa, M., Meshi, T., & Okada, Y. (1987). Expression of bacterial chloramphenicol acetyltransferase gene in tobacco plants mediated by TMV-RNA. *The EMBO Journal*, *6*(2), 307-311.
4. Antoniw, J. F., Ritter, C. E., Pierpoint, W. S., & Van Loon, L. C. (1980). Comparison of three pathogenesis-related proteins from plants of two cultivars of tobacco infected with TMV. *Journal of General Virology*, *47*(1), 79-87.
5. Chen, J., Yan, X. H., Dong, J. H., Sang, P., Fang, X., Di, Y. T., ...&Hao, X. J. (2009). Tobacco mosaic virus (TMV) inhibitors from Picrasmaquassioides Benn. *Journal of agricultural and food chemistry*, *57*(15), 6590-6595.
6. Li, Z., Yang, B., Ding, Y., Meng, J., Hu, J., Zhou, X., ...& Yang, S. (2023). Insights into a class of natural eugenol and its optimized derivatives as potential tobacco mosaic virus helicase inhibitors by structure-based virtual screening. *International Journal of Biological Macromolecules*, *248*, 125892.
7. Rasool, N., Majeed, A., Riaz, F., &Hussain, W. (2020). Identification of novel inhibitory candidates against two major Flavivirus pathogens via CADD protocols: in silico analysis of phytochemical binding, reactivity, and pharmacokinetics against NS5 from ZIKV and DENV. *Structural Chemistry*, *31*(6), 2189-2204.

8. Uma Kumari, Shruti Gupta, NGS and Sequence Analysis with Biopython for Prospective Brain Cancer Therapeutic Studies, https://doi.org/10.22214/ijraset.2023.50885
9. Burley, S. K., Berman, H. M., Kleywegt, G. J., Markley, J. L., Nakamura, H., &Velankar, S. (2017). Protein Data Bank (PDB): the single global macromolecular structure archive. *Protein crystallography: methods and protocols*, 627-641.
10. Kukreja, Vinita &Kumari, Uma. (2023). Data Analysis of Brain Cancer with Biopython. 8. 2146-2154.
11. Yuan, S., Chan, H. S., & Hu, Z. (2017). Using PyMOL as a platform for computational drug design. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *7*(2), e1298.
12. DeLano, W. L. (2002). Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr*, *40*(1), 82-92.
13. Dr Uma kumari, Devanshi Gupta, In silico RNA aptamer drug design and modelling,2022/4, Journal-JETIR,Volume-9, Issue-4, Pages 718-725
14. Sayle, R. A., & Milner-White, E. J. (1995). RASMOL: biomolecular graphics for all. *Trends in biochemical sciences*, *20*(9), 374-376.
15. Bernstein, H. J. (2000). Recent changes to RasMol, recombining the variants. *Trends in biochemical sciences*, *25*(9), 453-455.
16. Huey, R., Morris, G. M., & Forli, S. (2012). Using AutoDock 4 and AutoDockvina with AutoDockTools: a tutorial. *The Scripps Research Institute Molecular Graphics Laboratory*, *10550*(92037), 1000.
17. Morris, G. M., Goodsell, D. S., Huey, R., Hart, W. E., Halliday, S., Belew, R., & Olson, A. J. (2001). AutoDock. *Automated docking of flexible ligands to receptor-User Guide*.
18. Morris, G. M., Huey, R., & Olson, A. J. (2008). Using autodock for ligand-receptor docking. *Current protocols in bioinformatics*, *24*(1), 8-14.
19. Cosconati, S., Forli, S., Perryman, A. L., Harris, R., Goodsell, D. S., & Olson, A. J. (2010). Virtual screening with AutoDock: theory and practice. *Expert opinion on drug discovery*, *5*(6), 597-607.
20. Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., & Lopez, R. (2005). InterProScan: protein domains identifier. *Nucleic acids research*, *33*(suppl_2), W116-W120.
21. Mulder, N., &Apweiler, R. (2008). Interpro and interproscan. *Comparative genomics*, 59-70.
22. Laskowski, R. A. (2001). PDBsum: summaries and analyses of PDB structures. *Nucleic acids research*, *29*(1), 221-222.
23. Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S., & Thornton, J. M. (2018). PDBsum: Structural summaries of PDB entries. *Protein science*, *27*(1), 129-134.
24. Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., ...& Bryant, S. H. (2016). PubChem substance and compound databases. *Nucleic acids research*, *44*(D1), D1202-D1213.
25. Wang, Y., Xiao, J., Suzek, T. O., Zhang, J., Wang, J., & Bryant, S. H. (2009). PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic acids research*, *37*(suppl_2), W623-W633.