International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Validation of NGS and Multi-omics Approach for Active Site Identification CDK2-Associated Protein 1: Oral Cancer

Navya Srivastava[1], Uma Kumari[2], Kartik Tripathi[3]

[1]Project Trainee, Bioinformatics Project and Research Institute, Noida - 201301, India
[1]Senior Bioinformatics Scientist, Bioinformatics Project and Research Institute, Noida - 201301, India
[2]Project Trainee at Bioinformatics Project and Research Institute, Noida - 201301, India
[1]*Corresponding Author Uma Kumari (uma27910@gmail.com)

ABSTRACT

This study investigates the role of Cyclin-Dependent Kinase 2-Associated Protein 1 (CDK2AP1) in oral cancer, utilizing a multidisciplinary approach encompassing structural analysis, domain identification, molecular docking, Next-generation Sequencing (NGS) such as CATH classification, and STRING network enrichment analysis. The Protein Data Bank (PDB) entry 2KW6 provides the three-dimensional structure of CDK2AP1, serving as the primary dataset. Structural analysis reveals intricate hydrogen bonding patterns and domain organization within CDK2AP1, indicative of its molecular architecture and stability. Domain identification identifies distinct functional units within CDK2AP1, highlighting its multifaceted role in oral cancer progression. Active site residue identification uncovers critical residues involved in protein-ligand interactions, presenting potential therapeutic targets. Molecular docking simulations predict putative binding interactions between CDK2AP1 and molecular interactions of three chemotherapy drugs, namely Hydroxyurea, Fluorouracil, and Cisplatin, using molecular docking simulations. Autodock scores were calculated to assess the binding affinity of these drugs to their respective targets. Hydroxyurea, with an Autodock score of -8.3, is a medication used in the treatment of various cancers by inhibiting ribonucleotide reductase, thereby halting DNA synthesis and cell proliferation. Fluorouracil, with an Autodock score of -1.3, functions as an antimetabolite by inhibiting thymidylate synthase, disrupting DNA synthesis, and cell division. Cisplatin, with an Autodock score of -7.1, forms cross-links in DNA, preventing cell division and inducing apoptosis in cancer cells, unveiling potential mechanisms of action and signaling pathways implicated in oral cancer development. Overall, these findings deepen our understanding of CDK2AP1's involvement in oral carcinogenesis and underscore its potential as a therapeutic target for improving patient outcomes.

Keywords: Oral Cancer, Multi-omics, Structural analysis, NGS, CDK2AP1, Therapeutic Targets, Molecular Interactions, Domain identification.

INTRODUCTION

Oral cancer also referred to as oral cavity cancer or oral squamous cell carcinoma (OSCC)(ShahafGivony, 2020), encompasses cancers that develop within the mouth, including areas like the lips, tongue, gums, inner cheeks, and the roof or floor of the mouth(Sachidananda *et al*.., 2023). Oral

cancer predominantly manifests as squamous cell carcinomas, comprising over 90% of cases(Iacopo *et al*., 2019). A significant portion of these occurrences, approximately two-thirds, are observed in developing countries, with South Asia being particularly affected by Cancers (Poonam Joshi*et al*., 2014). Notably, India alone reports around 100,000 new cases annually(Bhawna Gupta*et al*., 2013). While men typically face a higher risk, with rates often double those of women, exceptions exist, such as in Bangladesh, where the male-to-female ratio stands at 10:1(Akter, 2022). These risk factors include tobacco use (both smoking and chewing), heavy alcohol consumption, betel quid chewing, poor oral hygiene, human papillomavirus (HPV) infection(Saeeda Baig*et al*., 2015), sun exposure (particularly for lip cancer), and a diet lacking in fruits and vegetables(Petti, 2009). Genetic predisposition and a family history of oral cancer can also play a role(Joseph, 2002).In the context of oral cancer,CDK2-Associated Protein 1 (CDK2AP1), also known as DOC-1 (Deleted in Oral Cancer-1), is a gene associated with tumor suppression(Atul Katarkar*et al*., 2015). It is located on chromosome 12q24 and encodes a protein that interacts with cyclin-dependent kinase 2 (CDK2), regulating its activity(Juan Chai *et al*., 2016). Originally identified as a potential tumor suppressor gene, CDK2AP1 is frequently deleted or downregulated in oral cancer cell lines and primary tumors(Wong*et al*., 2011).Numerous studies have highlighted the significance of CDK2AP1 in suppressing tumor growth and metastasis, particularly in oral cancer(Mei Sun *et al*., 2013). By inhibiting the activity of CDK2, CDK2AP1 helps control cell cycle progression and prevent uncontrolled cell proliferation(Wong*et al*., 2011). Loss or downregulation of CDK2AP1 can lead to dysregulated cell cycle progression, promoting tumor growth and metastasis in oral cancer(Xiong& Zhong, 2015).In oral cancer, CDK2AP1 downregulation is associated with advanced tumor stage, lymph node metastasis, and poor prognosis. Various mechanisms contribute to the dysregulation of CDK2AP1, including genetic alterations (e.g., deletions, mutations, promoter methylation) and post-transcriptional regulation (e.g., microRNA-mediated suppression)(Ping Liu*et al*., 2017). Therefore, assessing CDK2AP1 expression may help predict patient outcomes and guide treatment decisions in oral cancer(Li-Ching *et al*., 2012).Restoration of CDK2AP1 expression or function represents a potential therapeutic strategy for oral cancer treatment.

NGS enables high-throughput sequencing of DNA and RNA from oral cancer samples, facilitating the identification of genetic mutations, copy number variations, and gene expression changes in oral Cancer (Gorenchtein& Mike, 2016). By analyzing large cohorts of patients, researchers can uncover recurrent genetic alterations in CDK2AP1 and related pathways, offering valuable insights into the molecular mechanisms driving oral cancer progression(Parag Gulati, 2017). Additionally, NGS-based transcriptomic analyses allow for the characterization of gene expression signatures associated with protein dysregulation, potentially revealing novel biomarkers for diagnosis, prognosis, and treatment response in oral cancer patients(Priya & Uma, 2024).Metabolomic analyses can uncover metabolic alterations associated with CDK2AP1 dysregulation, providing mechanistic insights into the metabolic reprogramming observed in oral cancer cells(Xiong& Zhong, 2015).

A multi-omics approach in oral cancer research represents a comprehensive strategy to elucidate the complex molecular mechanisms underlying this disease. By integrating data from various omics disciplines, researchers aim to gain a holistic understanding of the genetic, epigenetic, transcriptional, proteomic, metabolomic, and microbiome alterations associated with oral carcinogenesis(Sajib Chakrabort*et al*., 2018). Epigenomic studies uncover modifications to DNA and histones that influence gene regulation and chromatin structure, impacting gene expression in oral cancer cells(Jun Liu*et al*., 2020).Additionally, microbiome investigations shed light on the complex interplay between the oral

microbiome and the host environment, potentially influencing cancer development and treatment response(Sajib Chakrabortyet *al*., 2018).

Next-generation sequencing (NGS) and multi-omics approaches have emerged as powerful tools in elucidating the molecular landscape of oral cancer(Liang Zhonget *al*., 2018)(Sajib Chakrabortyet *al*., 2018), including the role of CDK2AP1 (CDK2-Associated Protein 1). These technologies offer comprehensive insights into the genetic, epigenetic, transcriptomic, proteomic, and metabolomic alterations associated with oral carcinogenesis, providing a holistic understanding of the disease(Kshreeraja *et al*., 2023).

The research aims to conduct a comprehensive investigation into the molecular mechanisms underlying oral cancer, with a particular focus on the role of CDK2AP1 (CDK2-Associated Protein 1). By employing Next-Generation Sequencing (NGS) and multi-omics approaches, the study seeks to elucidate the genetic, epigenetic, transcriptomic, proteomic, and metabolomic alterations associated with oral carcinogenesis.

MATERIALS & METHODS

The study begins by obtaining the three-dimensional structure of CDK2AP1 from the Protein Data Bank (PDB) entry 2KW6, alongside relevant protein sequences and structural data from reputable databases. Structural analysis of 2KW6 will be conducted using molecular visualization software such as PyMol and RasMol, focusing on identifying hydrogen bonding patterns to understand its stabilizing interactions and structural characteristics. Bioinformatics tools like InterPro Scan will aid in identifying functional domains within the protein sequence, facilitating the interpretation of its biological role and interactions. Active site residues crucial for enzymatic activity or ligand binding will be identified through computational methods and literature review, providing insights into 2KW6's functional significance in oral cancer. Molecular docking simulations will predict potential binding interactions with other biomolecules, focusing on active site residues to identify putative binding sites and interaction modes. Integration of multi-omics data from genomics, transcriptomics, and proteomics disciplines will complement structural analysis, offering a comprehensive understanding of CDK2AP1's role in oral cancer and its cellular interactions. Utilizing the CATH database, the protein structure will be classified into respective architecture classes, topologies, and homologous superfamilies, revealing evolutionary relationships and structural diversity. Network enrichment analysis via the STRING server will uncover functional associations and interactions of CDK2AP1 with other proteins, shedding light on the cellular pathways and biological processes it is involved in within the context of oral cancer.

RESULTS

Figure 1: H-bond on representation in RasMol.

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Chain A, Cyclin-dependent kinase 2-associated protein 1 [Homo sapiens]	Homo sapiens	129	129	100%	3e-37	100.00%	65	2KW6_A
cyclin-dependent kinase 2-associated protein 1 isoform X2 [Xyrauchen texanus]	Xyrauchen texanus	115	115	98%	2e-31	87.50%	87	XP_051973780.1
cyclin-dependent kinase 2-associated protein 1 isoform X2 [Ctenopharyngodon idella]	Ctenopharyngodon i...	115	115	98%	2e-31	87.50%	87	XP_051764649.1
cyclin-dependent kinase 2-associated protein 1-like isoform X2 [Cyprinus carpio]	Cyprinus carpio	114	114	98%	2e-31	87.50%	87	XP_042621502.1
cyclin-dependent kinase 2-associated protein 1 isoform X2 [Misgurnus anguillicaudatus]	Misgurnus anguillic...	114	114	98%	2e-31	87.50%	87	XP_055063612.1

Figure 2: BLAST Description Result.

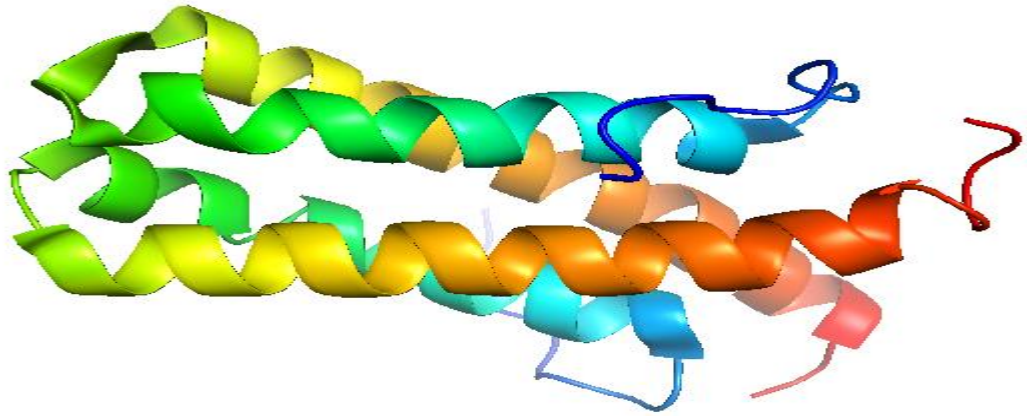

Figure 3: C-terminal RED N-Terminus Blue in PyMol.

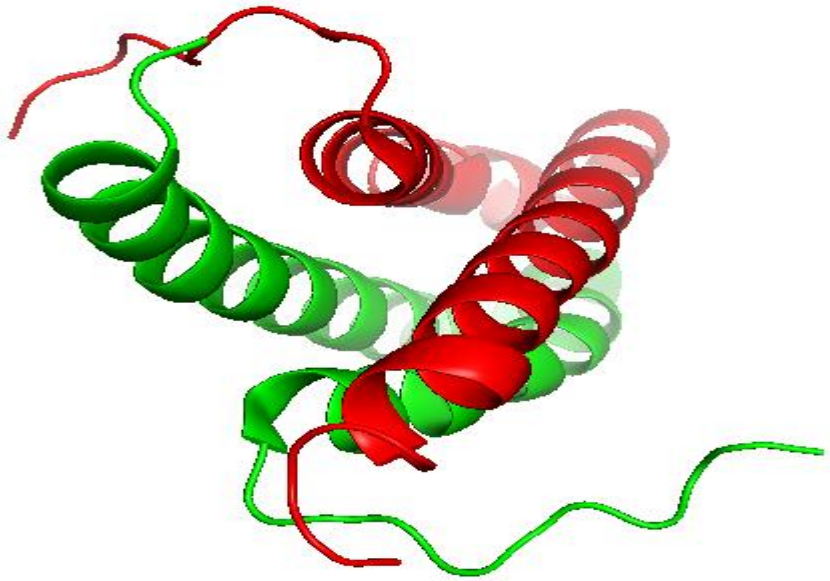

Figure 4: Chain-A redchain-B green (domain identification in PYMOL) Python command.

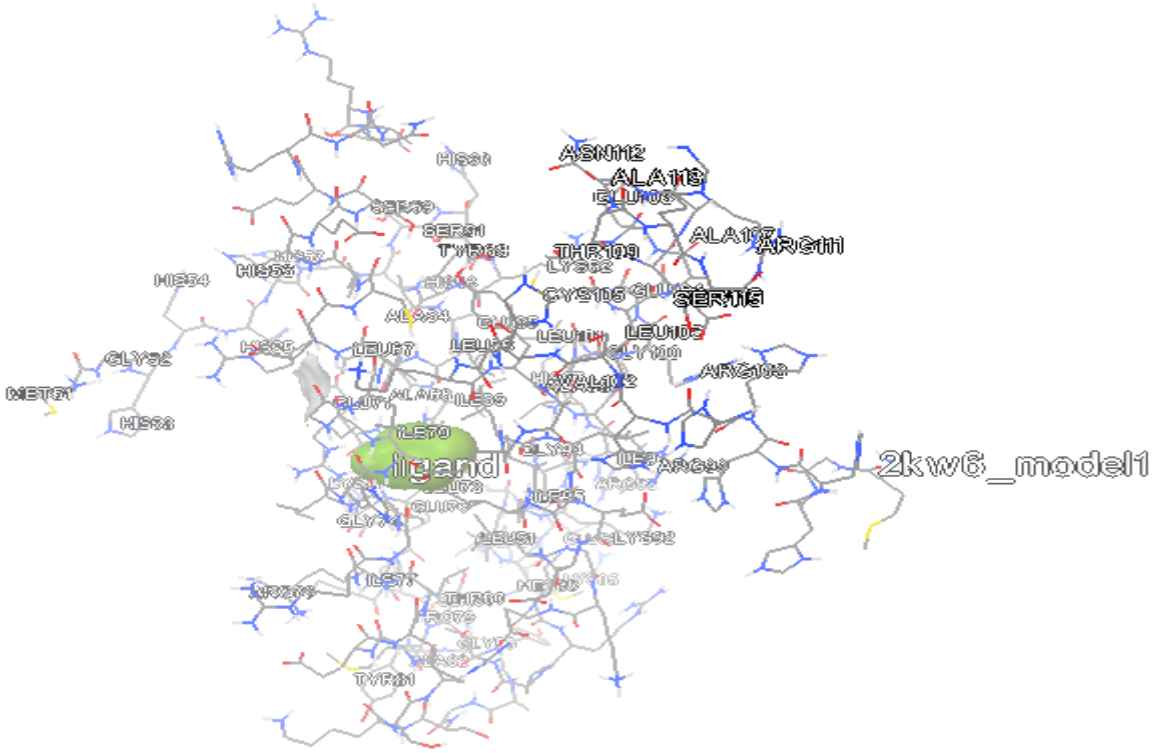

Figure 5: Active site Residue representation.

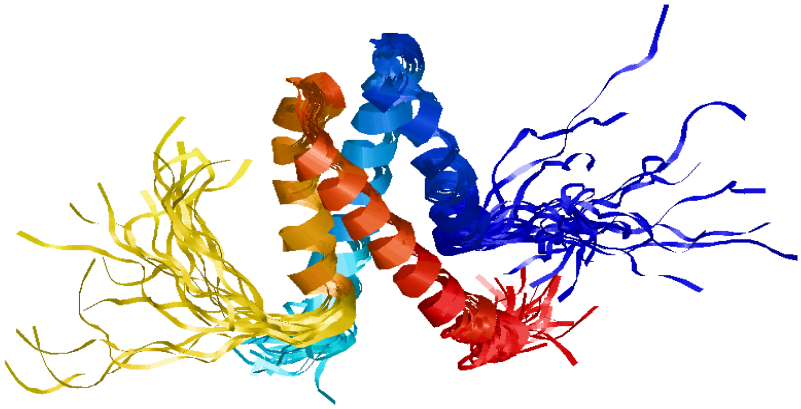

Figure 6: Spectrum define (N-terminalBlue, C-terminal Red).

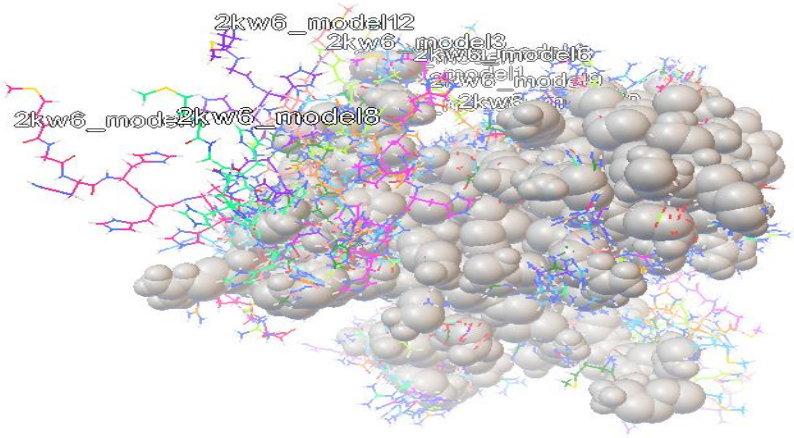

Figure 7: Prepared protein 2KW6 (model 1) in pdb.qt format.

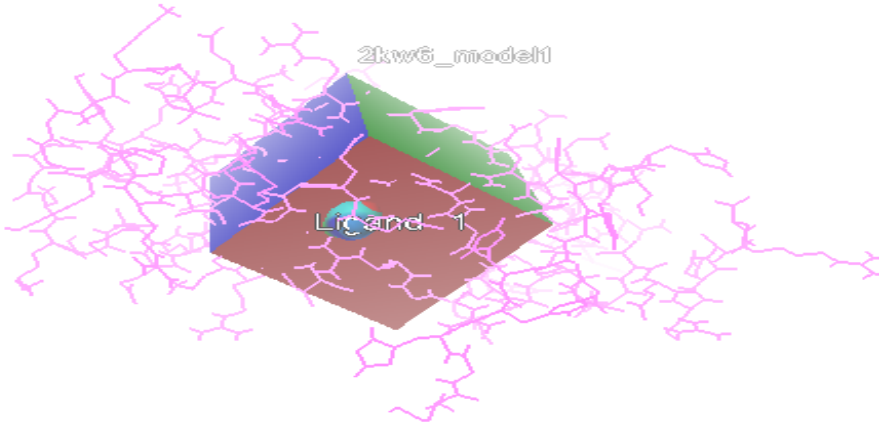

Figure 8: Molecular docking (Grid box)Hydroxyurea.

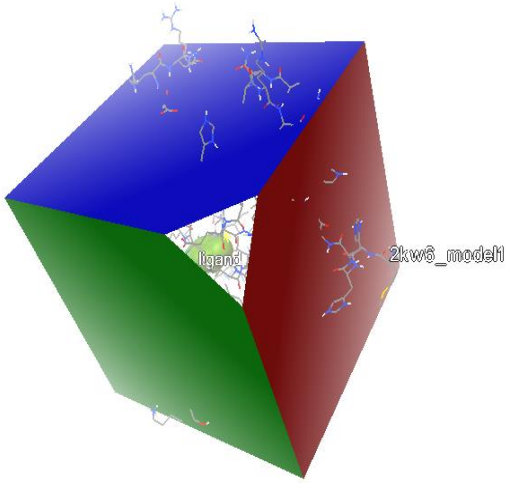

Figure 9: Molecular docking (Grid box)5-fluorouracil.

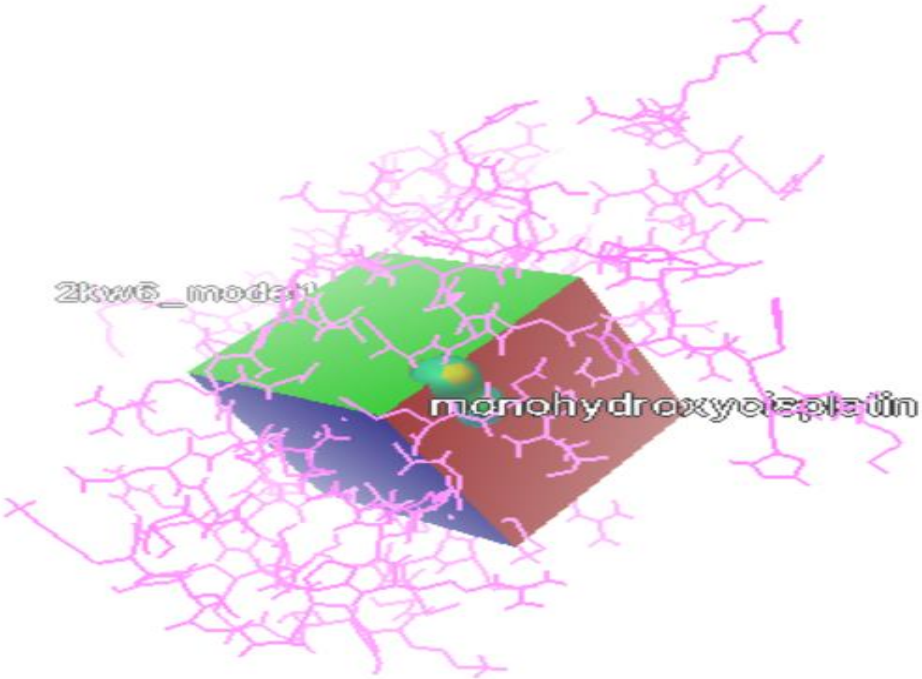

Figure 10: Molecular docking (Grid box) Cisplatin

Table 1: Showing the Grid score of the drug used for Docking.

S. No.	Drug Name	PDB ID	Autodock score	Center total grid			Center Grid Box		
				x	y	z	x	y	z
1.	Hydroxyurea	2KW6	-8.3	40	40	40	-8.3	-3.9	-2.8
2.	Fluorouracil		-1.3	40	40	40	-0.16	-0.17	-2.88
3.	Cisplatin		-7.1	40	40	40	-2.0	-3.0	-2.8

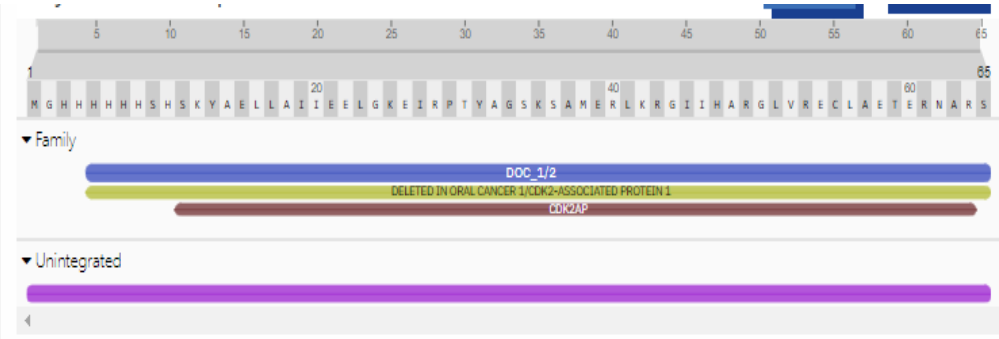

Figure 11: Interpro scan (Domain Analysis using Interpro scan)

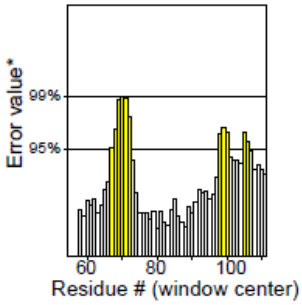

Figure 12: ERRAT Result with Overall quality factor 85.470.

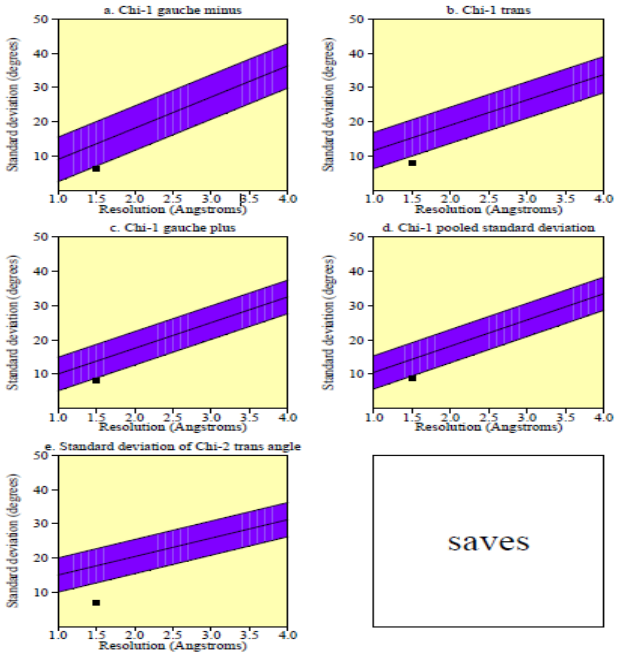

Figure 13: Side Chain Parameter Saves of the protein 2KW6.

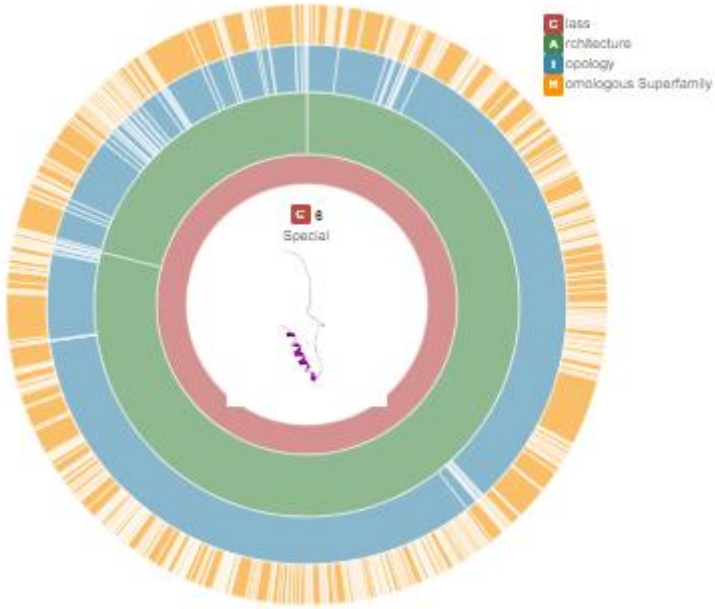

Figure 14 : (Representation of Class alpha helix representation)

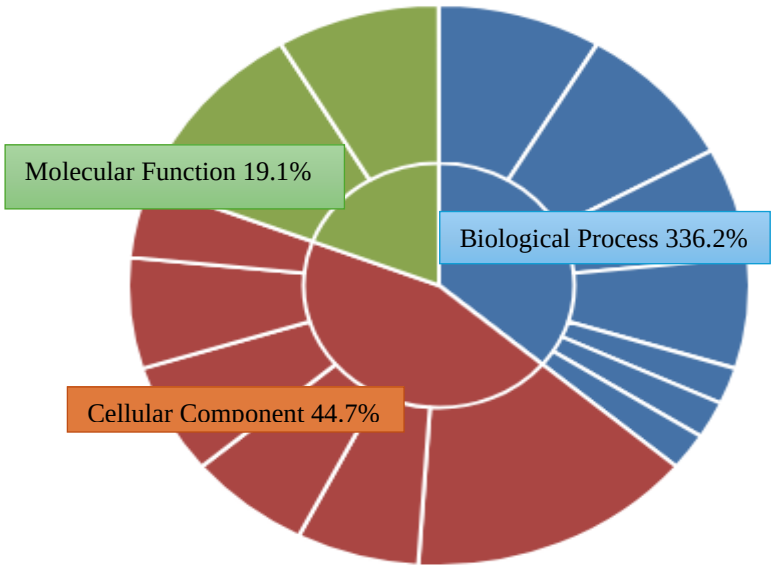

Figure 15: Gene Ontology (GO) Diversity distribution in Molecular, Biological, and Cellular components.

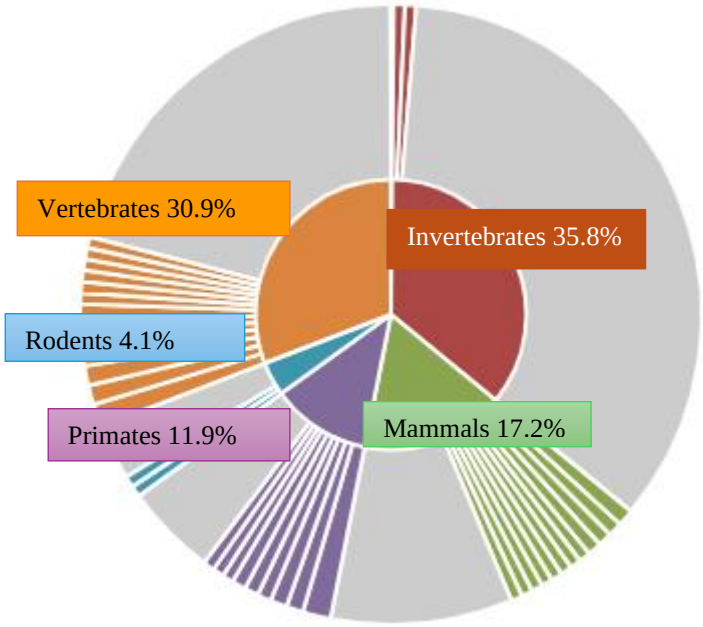

Figure 16: Species Diversity distribution

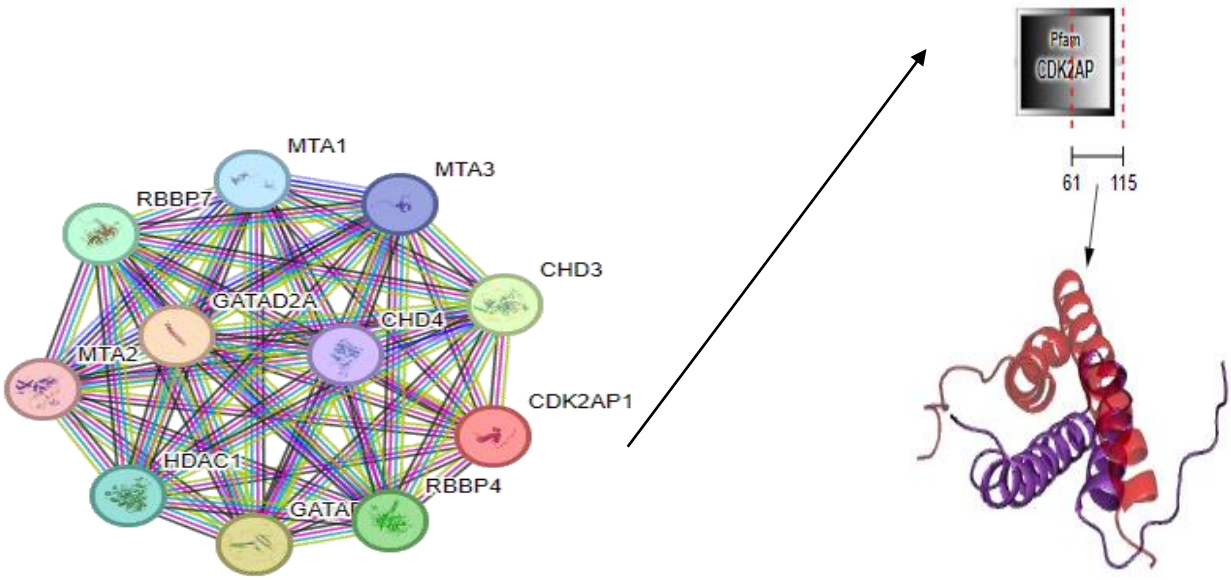

Figure 17: Network Enrichment analysis from string server showing the CDK2AP1.

CONCLUSION

In conclusion, this study has provided a comprehensive analysis of Cyclin-Dependent Kinase 2-Associated Protein 1 (CDK2AP1) in the context of oral cancer. Through various analytical techniques, several significant findings have emerged.The main finding of this research is the elucidation of CDK2AP1's structural features and functional implications in oral carcinogenesis. Structural analysis revealed intricate hydrogen bonding patterns and domain organization within CDK2AP1, shedding light

on its molecular architecture and stability. Domain identification uncovered distinct functional units within CDK2AP1, highlighting its multifaceted role in oral cancer progression.

Active site residue identification identified critical residues involved in protein-ligand interactions, presenting potential targets for therapeutic intervention. Molecular docking simulations predicted putative binding interactions between CDK2AP1 and other biomolecules, revealing potential mechanisms of action and signaling pathways implicated in oral cancer development.Furthermore, CATH classification categorized CDK2AP1 into specific protein architecture classes, providing insights into its evolutionary relationships and structural diversity. Network enrichment analysis using the STRING server unveiled functional associations of CDK2AP1 with other proteins, offering additional context for its role in oral cancer biology.

Overall, these findings contribute to a deeper understanding of CDK2AP1's involvement in oral carcinogenesis and highlight its potential as a therapeutic target. This research lays the foundation for future investigations aimed at developing innovative diagnostic and therapeutic strategies for oral cancer management, with the ultimate goal of improving patient outcomes and quality of life.

REFERENCES

1. Akter, S. (2022). Pattern Of Cancer And Its Risk Factor In Chattogram. *PhD diss., Chattogram Veterinary & Animal Science University, Khulshi*.
2. Atul Katarkar, L. P. (2015). Association of oral tumor suppressor gene deleted in oral cancer-1 (DOC-1) in progression of oral precancer to cancer. *Oral Science International, 12*(1), 15-21.
3. Bhawna Gupta, A. A. (2013). Oral cancer in India continues in epidemic proportions: evidence base and policy initiatives. *International Dental Journal, 63*(1), 12-25.
4. D.T.W. Wong, J. K.-H. (2011). Double Edge: CDK2AP1 in Cell-cycle Regulation and Epigenetic Regulation. *Journal of Dental Research, 91*(3), 235-241.
5. Givony, S. (2020). Oral squamous cell carcinoma (OSCC) an overview. *Medical Sciences, 8*(13), 67-74.
6. Gorenchtein, M. (2016). Microrna deregulation within optically altered oral cancer fields of non-smoker patients. *Open Collections*.
7. Gulati, P. (2017). DECIPHERING CELL SIGNALING PATHWAYS IN ORAL CANCER: A COMPREHENSIVE REVIEW. *International Journal of Medical Science and Dental Health, 3*(11), 1-4.
8. Iacopo Panarese, G. A. (2019). Oral and Oropharyngeal squamous cell carcinoma: prognostic and predictive parameters in the etiopathogenetic route. *Expert Review of Anticancer Therapy, 19*(2), 105-119.
9. Kshreeraja S. Satish, K. S. (2023). Leveraging technology-driven strategies to untangle omics big data: circumventing roadblocks in clinical facets of oral cancer. *Frontiers in Oncology, 13*, 1183766.
10. Liang Zhong, Y. L. (2018). Biomarkers: paving stones on the road towards the personalized precision medicine for oral squamous cell carcinoma. *BMC Cancer, 18*, 1-20.
11. Li-Ching Wu, Y.-L. C.-R.-F.-Y.-W.-L.-Y.-H.-P.-J.-L. (2012). Expression of cyclin-dependent kinase 2-associated protein 1 confers an independent prognosticator in nasopharyngeal carcinoma: a cohort study. *Journal of Clinical Pathology, 65*(9), 795-801.

12. Mei Sun Rui Jiang Guan Wang Chao Zhang Jindong Li Chengyan Jin, X. Z. (2013). Cyclin-dependent kinase 2-associated protein 1 suppresses growth and tumorigenesis of lung cancer. *International Journal of Oncology*, *42*(4), 1376-1382.
13. Petti, S. (2009). Lifestyle risk factors for oral cancer. *Oral Oncology*, *45*(4-5), 340-350.
14. Ping Liu, F. Y. (2017). Dysregulated expression of microRNAs and mRNAs in pulmonary artery remodeling in ascites syndrome in broiler chickens. *Oncotarget*, *8*(2), 1993–2007.
15. Poonam Joshi, S. D. (2014). Head and Neck Cancers in Developing Countries. *Rambam Maimonides Medical Journal*, *5*(2), 9.
16. Priya Arya, U. K. (2024). INTEGRATIVE ANALYSIS OF SERINE DEHYDRATASE-LIKE (SDSL) GENE IN LIVER CANCER CELLS THROUGH NGS WITH BIOPYTHON. *International Journal of Medicine and Pharmaceutical Sciences*, *14*(1), 11-22.
17. Sachidananda Mallya P., A. B. (2023). Recent advances in oral cancer - A review. *Biomedicine*, *43*(6), 1646-1652.
18. Saeeda Baig, Z. R. (2015). Chewable risk factors - Threatened oral cancer HPV's Looming Epidemic in Pakistan. *Euopean Journal of Biotechnology and Bioscience*, *3*(1), 39-45.
19. Sajib Chakraborty, M. I. (2018). Onco-Multi-OMICS Approach: A New Frontier in Cancer Research. *BioMed Research International*, 14.
20. Simone Marcone, A. B.-L. (2021). Proteomic signatures of radioresistance: Alteration of inflammation, angiogenesis and metabolism-related factors in radioresistant oesophageal adenocarcinoma. *Cancer Treatment and Research Communications*, *27*, 100376.
21. Uma Kumari; Prahalad Krishnakumar "ANALYSIS OF GASTRO-INTESTINAL STROMAL TUMOR BY COMPUTER AIDED DRUG DESIGNING AND BIOINFORMATICS TOOLS ", International Journal of Emerging Technologies and Innovative Research, ISSN:2349-5162, Vol.11, Issue 2, page no. ppc56-c65, February-2024, Available at : http://www.jetir.org/papers/JETIR2402210.pdf
22. Uma Kumari Priya Arya "MOLECULAR DOCKING APPROACH FOR HUMAN ONCOLOGY RESEARCH WITH CHEMOINFORMATIC", International Journal of Emerging Technologies and Innovative Research, ISSN:2349-5162, Vol.11, Issue 2, page no.a658-a666, February-2024, Available :http://www.jetir.org/papers/JETIR2402085.pdf
23. UMA KUMARI , Bhoomika A , Bhavana KS, "INSILICO COMPUTER-AIDED DRUG DESIGNING OF THE MUTATED EGFR KINASE DOMAIN OF NSCLC (NON-SMALL CELL LUNG CANCER)", International Journal of Emerging Technologies and Innovative Research , ISSN:2349-5162, Vol.10, Issue 12, page no. ppb226-b235, December-2023,http://doi.one/10.1729/Journal.37000
24. Xiangming He, H. X. (2014). CDK2-AP1 inhibits growth of breast cancer cells by regulating cell cycle and increasing docetaxel sensitivity in vivo and in vitro. *Cancer Cell International*, *14*(130), 1-10.
25. Xiong, G. Z. (2015). miR-205 promotes proliferation and invasion of laryngeal squamous cell carcinoma by suppressing CDK2AP1 expression. *Biological Research*, *48*(60), 337.