iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** IX **Month of publication:** September 2023

DOI: https://doi.org/10.22214/ijraset.2023.55614

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IX Sep 2023- Available at www.ijraset.com

Proteomics and NGS Analysis of Induced Myeloid Leukemia Cells

Uma Kumari[1], K. Sreenath Yadav[2]
[1]*Senior Bioinformatics Scientist, Bioinformatics Project and Research Institute, Noida - 201301, India*
[2]*Project Trainee, Bioinformatics Project and Research Institute, Noida-201301, India*

Abstract: ***Induced myeloid leukemia cells have emerged as critical models for unraveling intricate molecular mechanisms of leukemogenesis. Integration of proteomics and Next-Generation Sequencing (NGS) analyses has yielded profound insights. Mcl-1's significance in tumorigenesis and therapeutic resistance highlights Mcl-1 inhibitors' promise as anticancer agents. A selectively designed macrocyclic compound for Mcl-1 with high affinity, currently in clinical development, signifies a milestone in precision therapy. Activation of the Bak-dependent mitochondrial apoptotic pathway in acute myeloid leukemia offers novel insights into apoptotic regulation and potential interventions. Strikingly, complete tumor regression was achieved in diverse myeloma and acute myeloid leukemia models with a single tolerated dose of the compound, as monotherapy or combined with bortezomib or venetoclax, showcasing its therapeutic potential. Structural analyses through PyMOL and RasMol reveal intricate protein architecture, informing amino acid distributions, topology, and electrostatic profiles. These insights underpin understanding of protein functionality and intermolecular interactions.***

Simultaneously, multi-omics strategies, spanning genomics, epigenomics, transcriptomics, and proteomics, illuminate induced myeloid leukemia cells. These approaches uncover recurrent mutations, copy number variations, and fusion genes driving leukemogenesis. Epigenetic modifications, notably DNA methylation and chromatin remodeling, modulate gene expression in leukemic contexts.Together, these efforts underscore the significance of integrating diverse omics layers to decode leukemic transformation intricacies. The findings offer therapeutic prospects and insights into disease progression, fostering nuanced treatments across oncogenic landscapes beyond myeloid leukemia.

Keywords: ***Acute myeloid leukemia, proteomics, Structure analysis, leukemia, PyMol, homology modeling***

I. INTRODUCTION

Proteomics is the large-scale study of proteins expressed by a cell, tissue, or organism. In the context of induced myeloid leukemia cells, proteomics analysis can provide insights into the changes in protein expression, post-translational modifications, and protein-protein interactions that occur during leukemogenesis. Techniques used in proteomics analysis include mass spectrometry (MS) and liquid chromatography (LC), often coupled together as LC-MS/MS. Proteomics can help identify differentially expressed proteins between healthy and leukemia cells, providing a deeper understanding of the molecular mechanisms underlying the disease [1].

Next-Generation Sequencing (NGS) is a high-throughput technology that allows for rapid sequencing of DNA and RNA. In the context of induced myeloid leukemia cells, NGS can be used to study the genetic and epigenetic changes associated with leukemogenesis. Combining proteomics and NGS data can offer a comprehensive view of the molecular changes associated with induced myeloid leukemia cells. Integrative analyses can reveal how changes at the DNA, RNA, and protein levels collectively contribute to disease progression and identify potential therapeutic targets. Induced myeloid leukemia cells serve as invaluable models for deciphering the intricate mechanisms underlying leukemogenesis. This review article provides an extensive overview of recent advancements in understanding induced myeloid leukemia cells, with a focus on the integration of multi-omics approaches. We delve into the realms of genomics, epigenomics, transcriptomics, proteomics, and functional studies to unveil the complexity of induced myeloid leukemia cells. By amalgamating information from diverse omics layers, we aim to elucidate the molecular processes governing leukemic transformation and identify potential therapeutic avenues. Myeloid leukemia, characterized by the uncontrolled proliferation of myeloid lineage cells, represents a significant health challenge. Induced myeloid leukemia cells, cultivated through genetic and molecular manipulations, have emerged as powerful tools for dissecting the molecular basis of leukemia initiation and progression. Genomic aberrations play a pivotal role in the pathogenesis of myeloid leukemia. We discuss the utility of whole-genome sequencing (WGS) and whole-exome sequencing (WES) in identifying recurrent mutations, copy number alterations, and fusion genes associated with induced myeloid leukemia cells. By dissecting the mutational landscape, we shed light on key driver events and potential therapeutic targets [2,3,4,5].

Epigenetic modifications intricately regulate gene expression patterns in leukemia cells. We explore the application of ChIP-Seq and DNA methylation profiling in deciphering the epigenetic alterations that contribute to leukemic transformation. The interplay between chromatin remodeling, histone modifications, and DNA methylation offers insights into the epigenetic plasticity of induced myeloid leukemia cells [6,7].

In 1878, ERNST NEUMANN suggested for the first time that the bone marrow was the site of origin of leukemia and used the term myelogenus (myelogenus) leukemia. Myeloid is a term that relates to bone marrow or spinal cord. It also refers to a type of blood cell that arises from a common progenitor cell in the bone marrow. Myeloid cells include red blood cells granulocytes, monocytes & platelets. Myeloid leukemia is a type of leukemia affecting myeloid tissue. Myeloid tissue is the bone marrow arising tissue: bone marrow is the place where blood cells are synthesized. Bone marrow is two types of red marrow known as myeloid tissue & yellow marrow. Bone marrow contains two type's stem cells hematopoietic cells & stromal cells. Stem cells have the remarkable potential to renew themselves. They can develop into many different cell types in the body during early life and growth.There are several main categories: the "pluripotent" stem cells (embryonic stem cells and induced pluripotent stem cells) and non-embryonic or somatic stem cells (commonly called "adult" stem cells). Pluripotent stem cells have the ability to differentiate into all of the cells of the adult body. Adult stem cells are found in a tissue or organ and can differentiate to yield the specialized cell types of that tissue or organ.Early mammalian embryos at the blastocyst stage contain two types of cells – cells of the inner cell mass, and cells of the trophectoderm. The trophectodermal cells contribute to the placenta. The inner cell mass will ultimately develop into the specialized cell types, tissues, and organs of the entire body of the organism. Previous work with mouse embryos led to the development of a method in 1998 to derive stem cells from the inner cell mass of pre-implantation human embryos and to grow human embryonic stem cells (HESCs) in the laboratory. Blood carries oxygen and nutrients to living cells and takes away their waste products. Blood is deemed so precious that is also called "red gold". The straw-colored fluid that forms the top layer is called plasma. The middle white layer is composed of white blood cells (WBCs) and platelets, and the bottom red layer is the red blood cells (RBCs) [8,9,10,11,12,13].

II. MATERIALS AND METHODS

The PDB id of the protein used in this paper is 6FS0. The year was 1971 when the concept of the PDB was conceived, born out of the necessity for a centralized resource to house the growing number of experimentally determined protein structures. Recognizing the value of sharing these invaluable insights, researchers from around the world embarked on a journey of collaboration that would ultimately lead to the establishment of the PDB.

At the heart of the PDB are the intricate three-dimensional structures of proteins, meticulously deciphered through techniques such as X-ray crystallography, NMR spectroscopy, and cryo-electron microscopy. These structures offer a blueprint of a protein's architecture, revealing its folds, twists, and turns [14,15]. In the fascinating world of molecular biology and structural biology, visualizing complex molecular structures is a crucial task. Researchers and students alike rely on specialized software tools to gain insights into the 3D arrangements of molecules, which in turn helps in understanding their functions and interactions. Two widely used tools for this purpose are RasMol and PyMOL. RasMol, short for "Ras macromolecular visualization," can be considered a trailblazer in the field of molecular visualization software.

Developed by Roger Sayle in the early 1990s, RasMol was one of the first tools that made it possible for researchers to visualize and manipulate macromolecular structures interactively.

It offers various rendering options like wireframe, stick, and sphere representation, which allow users to visualize molecular structures at different levels of detail. Users can assign colors to different elements of a molecule and add labels to atoms, residues, or chains for better annotation. PyMOL is a more recent addition to the molecular visualization landscape, developed by Warren L. DeLano in the early 2000s.

It quickly gained popularity due to its advanced features and capabilities, becoming an essential tool for researchers and educators in the biological sciences. PyMOL offers high-quality rendering capabilities, including ray-tracing for photorealistic images, ambient occlusion, and advanced lighting effects. PyMOL allows users to create molecular animations and generate movies, which are invaluable for presentations and publications [16,17,18,19].PDBsum is a web-based resource that provides comprehensive and detailed structural information about proteins, as derived from their Protein Data Bank (PDB) entries.

It offers a wealth of visualizations, annotations, and analyses to help researchers understand the structure and function of proteins [20].

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IX Sep 2023- Available at www.ijraset.com

III. RESULT AND DISCUSSION

A. *3D Structural analysis using RasMol and PyMol*

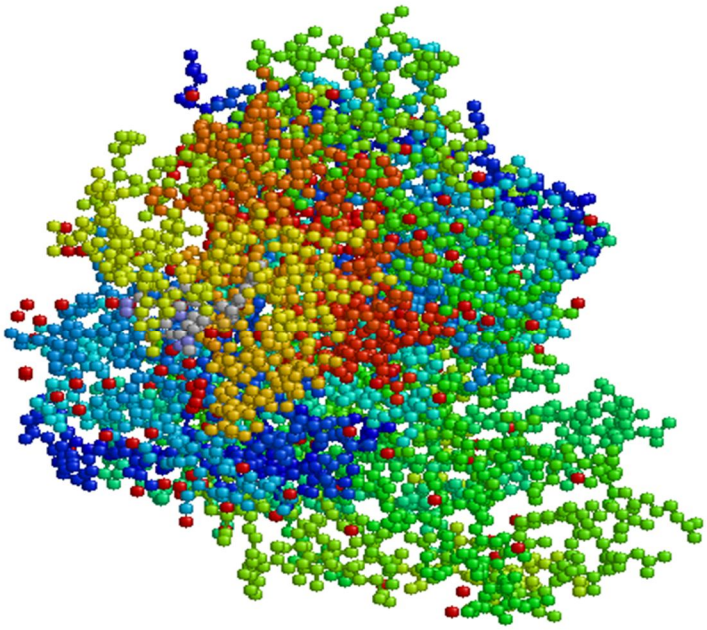


Figure1: Ball & stick in RASMOL display type. This structure defines Residue sphere chemical bonds and sticks. Each color defines different amino acid based upon color. It shows atoms as ball and chemical bonds as sticks

Table1: Results obtained from RasMol

COLOR REPRESENTATION	AMINO ACID
Green	Alanine
Olive green	Leucine
Medium orange	Serine
Dark orange	Glutamine
Dark blue	Arginine

Knowing the specific amino acids present in a protein is essential for a comprehensive understanding of its structure, function, and interactions. Here are some key benefits of knowing the amino acids present in a protein-

1) *Functional Insights:* Amino acids play a critical role in determining a protein's function. Understanding the sequence of amino acids can provide clues about the protein's enzymatic activity, binding partners, catalytic sites, and other functional features.
2) *Structure-Function Relationships:* The sequence of amino acids dictates the three-dimensional structure of the protein. By knowing the amino acid sequence, you can make predictions about the protein's secondary and tertiary structures, which are directly related to its function.
3) *Protein Interactions:* Amino acids that participate in protein-protein interactions or ligand binding sites are of particular interest. Understanding the residues involved in these interactions can shed light on how the protein interacts with other molecules in the cellular environment.
4) *Structural Analysis:* The specific amino acids present in a protein determine its folding patterns, secondary structures, and overall conformation. This information is crucial for structural biology studies and can aid in understanding how changes in the sequence might impact the protein's stability and structure.

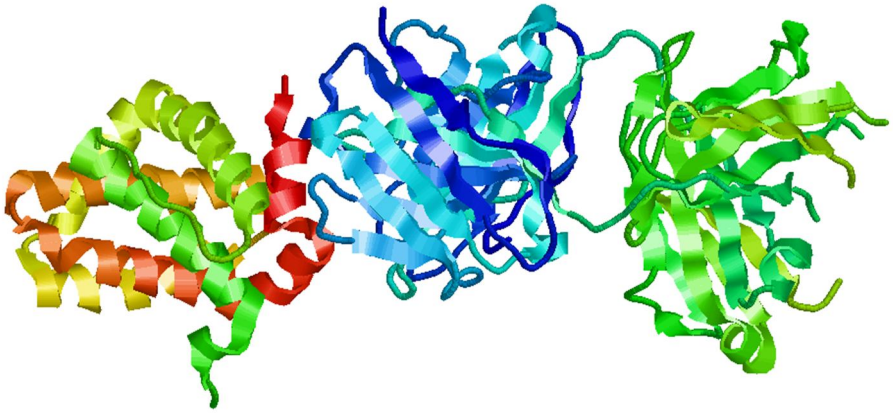


Figure 2: It represents N terminal and C terminal represents free amine end of the chain is called N terminus or Amino terminus. Free carboxylic acid end is called the C terminus or carboxyl terminus

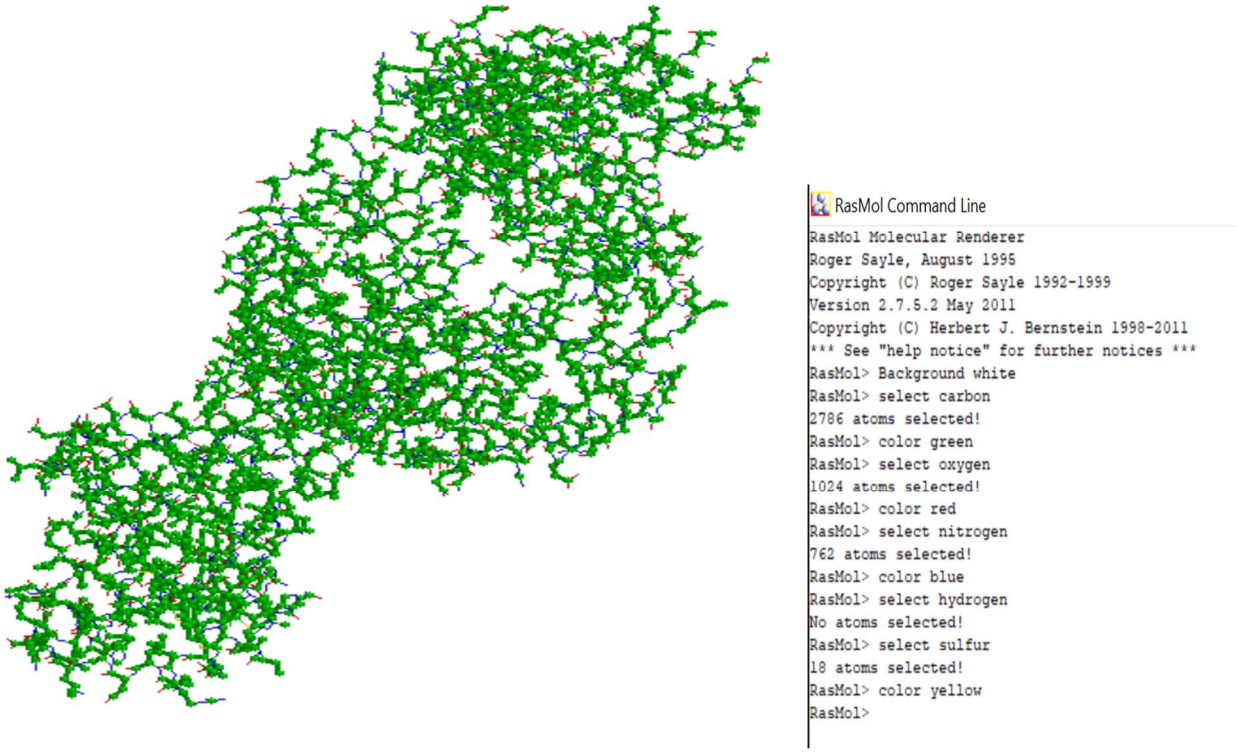


Figure 3: Visualizing protein atom structure through RasMol

Table2: Results obtained from RasMol

ATOM NAME	OBSERVED NO. OF ATOMS	ATOM SHOWN BY COLOR
carbon	2786	Green
Oxygen	1024	Red
Nitrogen	762	Blue
Hydrogen	No Atoms Present	NA
sulfur	18	Yellow

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue IX Sep 2023- Available at www.ijraset.com

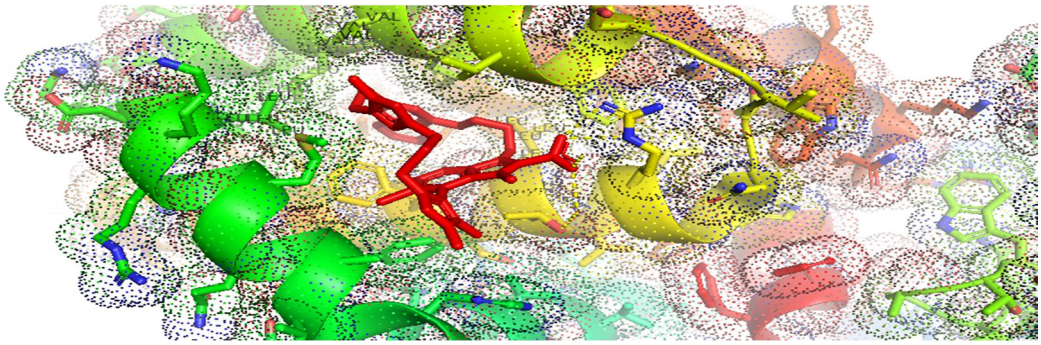


Figure 4:Licorice dot sticks are used in PyMOL to visualize molecular structures, such as protein or DNA. They represent atoms as dots and connect them with sticks to show the bond between them. It's a great tool for analyzing and studying complex molecules.

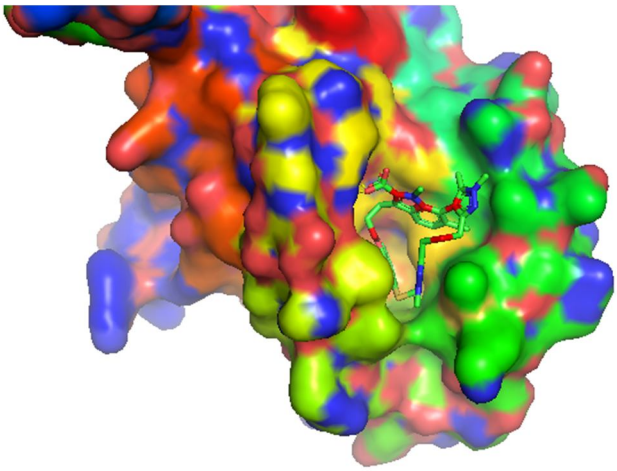


Figure 5: PYMOL result in surface form is visualize the accessible surface area of the molecule. This can be useful for studying protein-ligand interactions or identifying potential binding sites. The surface form provides a more detailed representation of the molecules shape and can help in understanding its structural properties.

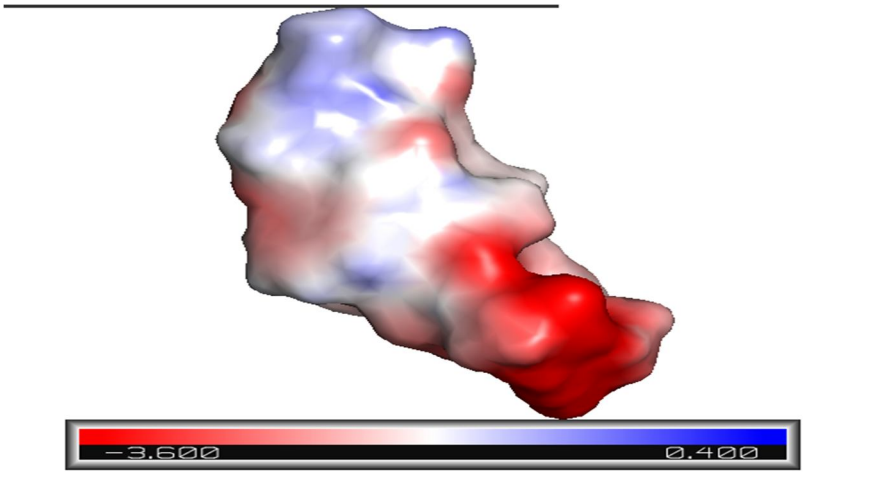


Figure 6: Protein is visualized through 'electrostatic' scheme in which positively charged region is shown by blue colour, negatively charged region is shown by red colour and neutral

Visualizing proteins using an "electrostatic" color scheme, where positively charged regions are shown in blue, negatively charged regions in red, and neutral regions in white, is a common technique to highlight the distribution of electrostatic charges on the protein's surface. This visualization can provide insights into the protein's potential interactions with other molecules, such as ligands, substrates, or other proteins.

Here's why this visualization is useful:

1) *Charge Distribution:* The distribution of charges on a protein's surface can play a crucial role in its interactions with other molecules. Positively charged regions might be involved in binding negatively charged ligands, and vice versa.
2) *Binding Sites:* Electrostatic interactions are often key determinants of molecular recognition and binding. Visualizing charged regions can help identify potential binding sites where complementary charges might lead to specific interactions.
3) *Enzyme Active Sites:* Enzymes often have catalytic residues with specific charges that help facilitate chemical reactions

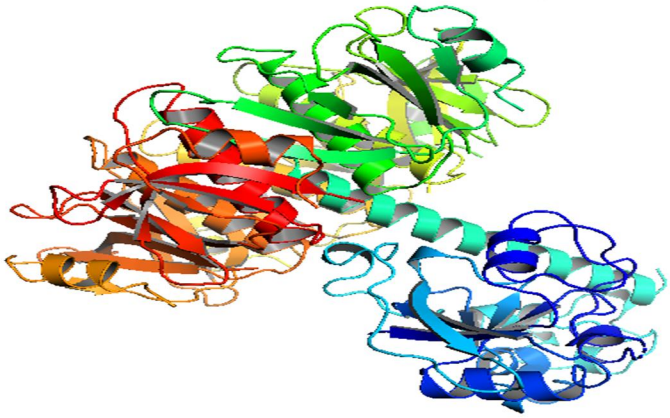


Figure 7: MIA is visualized through 'cartoon ribbon' scheme in which high density region is shown by red colour and low density region is shown by blue colour

B. 2D Structural analysis usingPDBsum

There are 3 chains in the protein structure. These are as follows-

- A chain (156 amino acids)
- H chain (208 amino acids)
- L chain (213 amino acids)

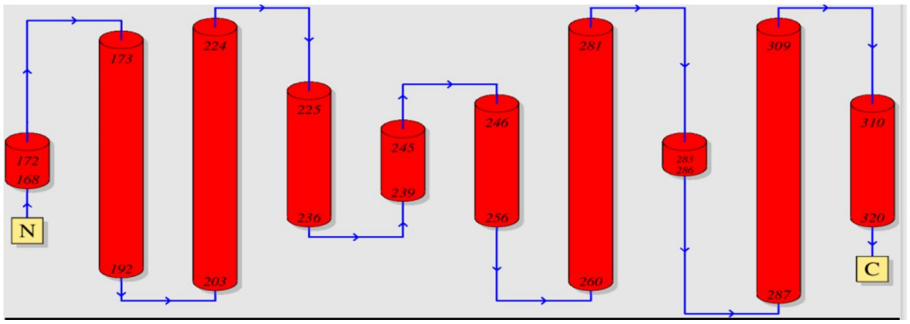


Figure 8: Topology visualization in PDBSum for A chain

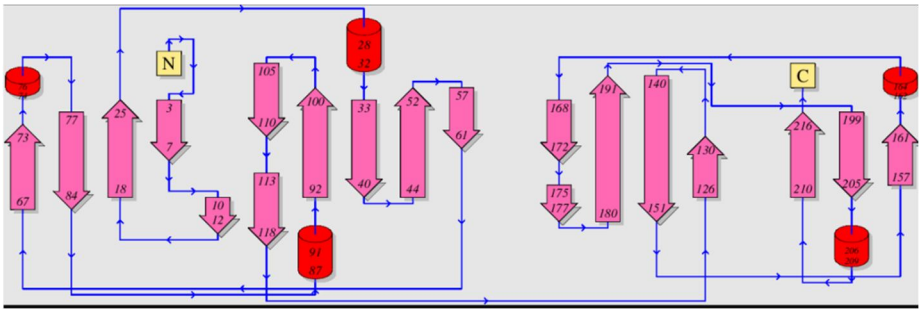


Figure 9: Topology visualization in PDBSum for H chain

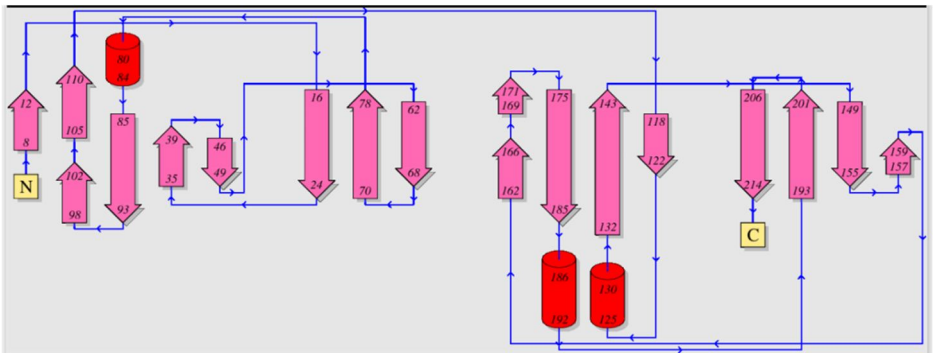


Figure 10: Topology visualization in PDBSum for L chain

Visualizing protein topology is essential for understanding its three-dimensional structure and how different parts of the protein are interconnected. Protein topology refers to the arrangement of secondary structural elements (such as alpha helices and beta sheets) and the connectivity between these elements.

Topology diagrams are schematic representations that emphasize the connectivity between secondary structure elements. These diagrams typically use lines to represent strands and cylinders to represent helices, showing how they are connected.

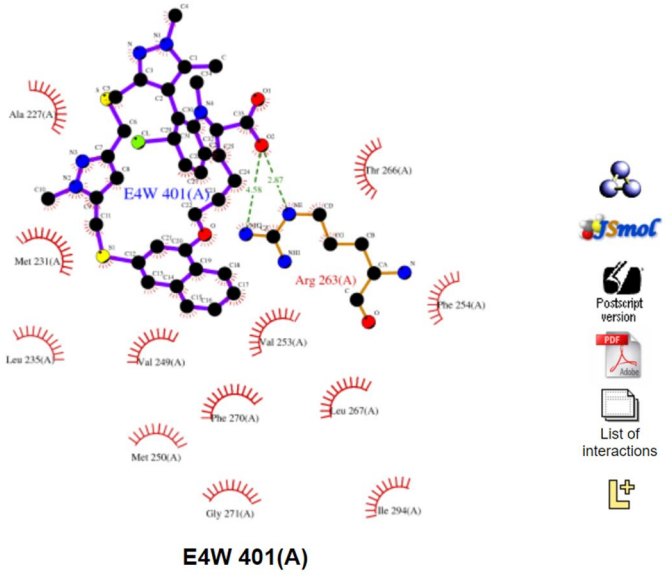


Figure 11:Ligand bonds are depicted by purple lines, non-ligand bonds by yellow lines, hydrogen bonds by green dashed lines, non-ligand residues in hydrophobic contact by red curve lines, and atoms in hydrophobic contact by black dots.

It provides detailed visualizations and analyses of protein-ligand complexes, highlighting key interactions and their contributions to binding. Here's a brief overview of LigPlot and its capabilities:

LigPlot Features:

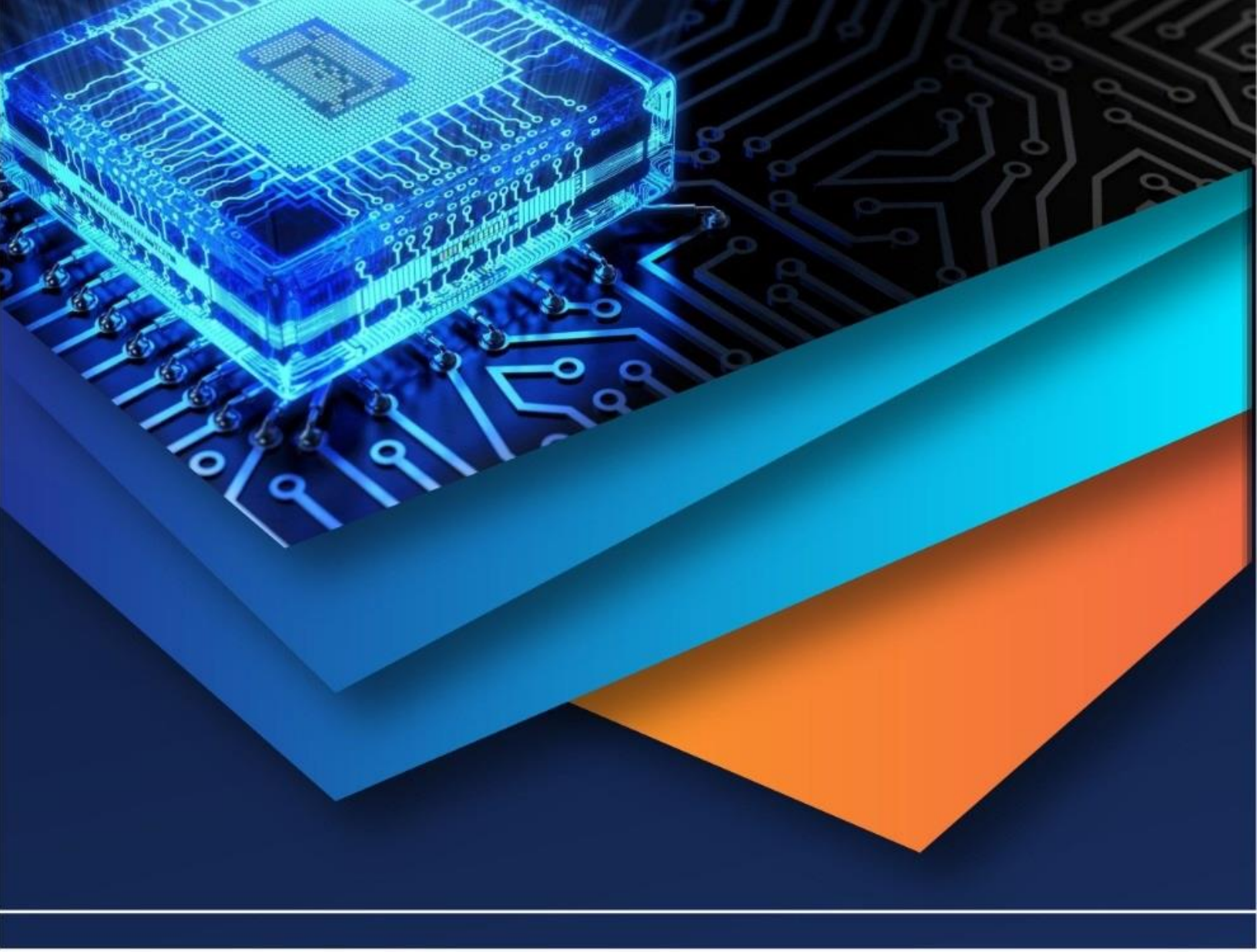
1) *Visualizing Interactions:* LigPlot generates schematic diagrams that display the interactions between a protein and its ligand(s). It highlights hydrogen bonds, hydrophobic interactions, metal coordination, and other non-covalent interactions.
2) *Protein-Ligand Contacts:* LigPlot identifies specific residues in the protein and atoms in the ligand that are involved in interactions. It shows the distances and angles between interacting atoms, helping to assess the strength of the interactions.
3) *Interaction Maps:* The software generates 2D interaction maps that display the location and type of interactions on a 2D plane, aiding in understanding the spatial distribution of interactions.
4) *Multiple Ligand Analysis:* LigPlot can handle complexes with multiple ligands, allowing you to analyze interactions in more complex scenarios.

IV. CONCLUSION

In conclusion, the integration of proteomics and Next-Generation Sequencing (NGS) analysis has provided a comprehensive understanding of induced myeloid leukemia cells and their underlying molecular mechanisms. The discovery of Mcl-1 as a key player in tumorigenesis and resistance to anticancer therapies highlights the potential of Mcl-1 inhibitors as promising anticancer drugs. The development of a rationally designed macrocyclic molecule with high selectivity and affinity for Mcl-1 showcases a significant advancement in targeted therapy. The use of structural analysis tools like PyMOL and RasMol has allowed for a detailed exploration of protein structures, offering insights into amino acid distribution, topology, and electrostatic potential. These analyses are crucial for understanding the structural basis of protein function and interactions. Collectively, these studies emphasize the importance of integrating various omics layers to decipher the molecular processes underlying leukemic transformation. The advancements in our understanding of induced myeloid leukemia cells offer potential therapeutic avenues and insights into disease progression. By bridging the gap between basic research and clinical applications, these findings pave the way for the development of more effective and targeted treatments for myeloid leukemia and potentially other cancers.

REFERENCES

[1] Al-Amrani, S., Al-Jabri, Z., Al-Zaabi, A., Alshekaili, J., & Al-Khabori, M. (2021). Proteomics: Concepts and applications in human medicine. World journal of biological chemistry, 12(5), 57–69. https://doi.org/10.4331/wjbc.v12.i5.57
[2] Besser, J., Carleton, H. A., Gerner-Smidt, P., Lindsey, R. L., & Trees, E. (2018). Next-generation sequencing technologies and their application to the study and control of bacterial infections. Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases, 24(4), 335–341. https://doi.org/10.1016/j.cmi.2017.10.013
[3] Sharma, T. R., Devanna, B. N., Kiran, K., Singh, P. K., Arora, K., Jain, P., Tiwari, I. M., Dubey, H., Saklani, B., Kumari, M., Singh, J., Jaswal, R., Kapoor, R., Pawar, D. V., Sinha, S., Bisht, D. S., Solanke, A. U., &Mondal, T. K. (2018). Status and Prospects of Next Generation Sequencing Technologies in Crop Plants. Current issues in molecular biology, 27, 1–36. https://doi.org/10.21775/cimb.027.001
[4] Koboldt, D. C., Steinberg, K. M., Larson, D. E., Wilson, R. K., &Mardis, E. R. (2013). The next-generation sequencing revolution and its impact on genomics. Cell, 155(1), 27–38. https://doi.org/10.1016/j.cell.2013.09.006
[5] Dorado, G., Gálvez, S., Rosales, T. E., Vásquez, V. F., & Hernández, P. (2021). Analyzing Modern Biomolecules: The Revolution of Nucleic-Acid Sequencing - Review. Biomolecules, 11(8), 1111. https://doi.org/10.3390/biom11081111
[6] Weinhold B. (2006). Epigenetics: the science of change. Environmental health perspectives, 114(3), A160–A167. https://doi.org/10.1289/ehp.114-a160
[7] Miller G. (2010). Epigenetics. The seductive allure of behavioral epigenetics. Science (New York, N.Y.), 329(5987), 24–27. https://doi.org/10.1126/science.329.5987.24
[8] Horton, S. J., &Huntly, B. J. (2012). Recent advances in acute myeloid leukemia stem cell biology. Haematologica, 97(7), 966–974. https://doi.org/10.3324/haematol.2011.054734
[9] Wang, X., Huang, S., & Chen, J. L. (2017). Understanding of leukemic stem cells and their clinical implications. Molecular cancer, 16(1), 2. https://doi.org/10.1186/s12943-016-0574-7
[10] Chan, W. I., &Huntly, B. J. (2008). Leukemia stem cells in acute myeloid leukemia. Seminars in oncology, 35(4), 326–335. https://doi.org/10.1053/j.seminoncol.2008.04.003
[11] Garg, S., Shanmukhaiah, C., Ghosh, K., &Madkaikar, M. (2016). Cell signaling in putative leukemic stem cells and blast population in acute myeloid leukemia. Leukemia & lymphoma, 57(9), 2195–2198. https://doi.org/10.3109/10428194.2015.1124996
[12] Xie, C., Zhou, H., Qin, D. et al. Bcl-2 inhibition combined with PPARα activation synergistically targets leukemic stem cell-like cells in acute myeloid leukemia. Cell Death Dis 14, 573 (2023). https://doi.org/10.1038/s41419-023-06075-6
[13] Jahn, E., Saadati, M., Fenaux, P. et al. Clinical impact of the genomic landscape and leukemogenic trajectories in non-intensively treated elderly acute myeloid leukemia patients. Leukemia (2023). https://doi.org/10.1038/s41375-023-01999-6
[14] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The Protein Data Bank. Nucleic acids research, 28(1), 235–242. https://doi.org/10.1093/nar/28.1.235
[15] Stephen K Burley and others, RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy, *Nucleic Acids Research*, Volume 47, Issue D1, 08 January 2019, Pages D464–D474, https://doi.org/10.1093/nar/gky1004
[16] Pikora, M., &Gieldon, A. (2015). RASMOL AB - new functionalities in the program for structure analysis. ActabiochimicaPolonica, 62(3), 629–631. https://doi.org/10.18388/abp.2015_972
[17] Amber D. Fraley, Katherine E. Odegaard,Victor H. Vilchiz,ChulHee Kang,Cathy Lee, The American Biology Teacher (2022) 84 (6): 369–373. https://doi.org/10.1525/abt.2022.84.6.369
[18] Seeliger, D., & de Groot, B. L. (2010). Ligand docking and binding site analysis with PyMOL and Autodock/Vina. Journal of computer-aided molecular design, 24(5), 417–422. https://doi.org/10.1007/s10822-010-9352-6
[19] Xue, Q., Liu, X., Russell, P., Li, J., Pan, W., Fu, J., & Zhang, A. (2022). Evaluation of the binding performance of flavonoids to estrogen receptor alpha by Autodock, AutodockVina and Surflex-Dock. Ecotoxicology and environmental safety, 233, 113323. https://doi.org/10.1016/j.ecoenv.2022.113323
[20] Laskowski R. A. (2001). PDBsum: summaries and analyses of PDB structures. Nucleic acids research, 29(1), 221–222. https://doi.org/10.1093/nar/29.1.221



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY