

NGS Diagnosis of Lp-PLA2 in Cardiovascular Disease & Insilico Analysis of Human Filamin C Domains 14-15 Mutation G1676R

Ponnambil Anantha Krishnan¹ Amrutha Velloor¹ Uma Kumari² Nidhi Aggarwal² Saptarshi Mukherjee³ Smitha MS³

¹Amity Institute of Microbial Technology ,Amity University, Noida, Uttar Pradesh 201313

¹Amity Institute of Microbial Technology ,Amity University, Noida, Uttar Pradesh 201313

²Senior Bioinformatics Scientist, Bioinformatics Project and Research Insitute, Noida - 201301, India

²Assistant Professor,Dr.D.Y Patil Biotechnology and Bioinformatics Institute Tathawade,Pune 411033

³Project Trainee at Bioinformatics Project and Research Insitute, Noida - 201301, India

³Assistant Professor Amity Institute of Microbial Technology ,Amity University, Noida, Uttar Pradesh 201313

Abstract

This study presents a comprehensive exploration of two key components in cardiovascular health: Human Lipoprotein-Associated Phospholipase A2 (Lp-PLA2), represented by the PDB code 5JAD, and filamin C (FLNC) domain 14-15 with mutation gene G1676R represented by the biological sample 7P0E, and By employing an integrated approach that combines structural, genetic, and functional analyses, this study aims to unravel the multifaceted roles of these enzymes in cardiovascular diseases (CVD) and their potential implications for therapeutic interventions. Initially, our investigation focuses on elucidating the structural and functional characteristics of Lp-PLA2, the enzyme represented by 7P0E, and Filamin C domain 14-15 with PDB Code 5JAD. Through rigorous analysis utilizing various bioinformatics tools and techniques, including Molecular Docking, structural validation, Next-generation Sequencing (NGS), Drug discovery, and functional domain analysis, we reveal the complex protein-ligand interactions of Lp-PLA2 and FLNC. Notably, our findings highlight potential therapeutic targets for both enzymes, with molecular docking analyses indicating promising interactions with various ligands, suggesting avenues for targeted interventions in CVD management. Concurrently, computational analyses and bioinformatics methodologies are employed to explore the genetic landscapes of Lp-PLA2 and FLNC, and their implications in cardiovascular health. Structural analyses using Command lines and Python-based software provide insights into the three-dimensional architectures of these enzymes, while BLAST assessments elucidate genetic variations. Furthermore, gene ontology and pathway analyses contextualize their roles within cardiovascular pathways, offering a holistic understanding of their contributions to disease pathogenesis. By integrating structural, genetic, and functional analyses of Lp-PLA2 and FLNC, this research advances our understanding of their roles in cardiovascular health. The findings not only provide valuable insights into potential therapeutic targets but also offer directions for future research and the development of targeted interventions aimed at combating the burden of cardiovascular diseases.

KEYWORDS: Cardiovascular Disease, Human Lipoprotein-Associated Phospholipase A2, Filamin C domain, Molecular Docking, therapeutic targets, structural variations, NGS, atherosclerosis.

Introduction

Cardiovascular disease (CVD) remains a pervasive global health challenge, contributing significantly to morbidity and mortality rates worldwide. With its multifaceted nature encompassing conditions such as coronary artery disease, heart failure, and hypertrophic cardiomyopathy, CVD poses complex diagnostic and therapeutic dilemmas (Miano, 2014). Despite advancements in understanding its etiology and risk factors, including genetic predispositions and lifestyle influences, there persists a critical need for innovative approaches to enhance risk prediction and treatment strategies (Hitesh *et al.*, 2021). The intricate interplay between genetic factors, environmental influences, and molecular mechanisms underscores the complexity of CVD pathogenesis. The intricate processes leading to cardiovascular diseases often involve the gradual accumulation of plaque within the arteries, a condition known as atherosclerosis (M. Zaromitidou *et al.*, 2016). This buildup narrows the vessels, restricting blood flow and potentially leading to complications such as heart attacks or strokes. The risk factors for CVD are multifaceted, involving a complex interplay of genetic predisposition, lifestyle choices, and environmental influences (Sekhara & Easterly, 2002). To address this complexity, researchers have increasingly turned to biomarkers and genetic mutations as potential indicators of cardiovascular risk and targets for therapeutic intervention (Hitesh *et al.*, 2021). In this context, lipoprotein-associated phospholipase A2 (Lp-PLA2) has emerged as a promising biomarker, offering insights into the inflammatory processes underlying atherosclerosis, a key contributor to many cardiovascular conditions. The enzyme is particularly relevant in the context of atherosclerosis, a condition characterized by the buildup of

plaque within arteries (Brian *et al.*, 2012). In atherosclerotic plaques, Lp-PLA2 activity contributes to the inflammatory processes, potentially leading to the destabilization of plaques and increasing the risk of cardiovascular events, including heart attacks and strokes (Devin Hasanally *et al.*, 2014). Measuring Lp-PLA2 levels in the bloodstream has been studied as a potential biomarker for assessing cardiovascular risk. Elevated levels of Lp-PLA2 have been associated with an increased likelihood of coronary heart disease and stroke (Arup *et al.*, 2019). Therefore, understanding the role of Lp-PLA2 in the inflammatory cascade within arterial walls is of significance in cardiovascular research and risk assessment. Concurrently, Mutations in the FLNC gene are associated with various muscle-related disorders, including cardiomyopathies, which are conditions affecting the heart muscle (Jakub *et al.*, 2018). Different mutations in Filamin C can lead to diverse clinical manifestations, including hypertrophic cardiomyopathy (HCM) and dilated cardiomyopathy (DCM) (Matthias & Nobert, 2021). Understanding the role of Filamin C in muscle function and its genetic variations is crucial for comprehending the molecular basis of these disorders and developing potential therapeutic interventions (Kavitha & Shankar, 2019). The Filamin C (FLNC) gene, particularly the G1676R mutation in domains 14-15, has garnered attention for its implications in hypertrophic cardiomyopathy and other cardiovascular disorders (Miyazaki, 2023). FLNC, a crucial component of muscle structure and function, plays diverse roles in sarcomere integrity and signaling pathways, making it a compelling target for genetic investigations in cardiovascular health (Matthias *et al.*, 2005). To comprehensively explore the molecular landscape of cardiovascular diseases, this research integrates insights from Lp-PLA2 biomarker analysis and FLNC G1676R mutation studies.

Leveraging next-generation sequencing (NGS) technologies and molecular docking, the study aims to unravel the genetic and structural underpinnings of CVD. By bridging the gap between genetic variation and clinical phenotypes, this multidisciplinary approach seeks to advance our understanding of CVD pathogenesis and pave the way for personalized therapeutic interventions targeting specific genetic mutations associated with cardiovascular risk (Andrew & Colin, 2005). Through this integrated analysis, the research endeavors to shed light on the intricate mechanisms governing cardiovascular health and contribute to the development of precision medicine strategies aimed at mitigating the impact of CVD on a molecular level (Priya & Uma, 2024). Both lipoprotein-associated phospholipase A2 (Lp-PLA2) and filamin C (FLNC) are significant players in cardiovascular health, each contributing unique insights into the pathogenesis and management of cardiovascular diseases. Lp-PLA2, residing primarily within circulating lipoproteins, particularly low-density lipoprotein (LDL) particles, plays a crucial role in catalyzing the hydrolysis of oxidized phospholipids (Roland, 2013), thus generating pro-inflammatory products that contribute to the progression of atherosclerosis. Elevated levels of Lp-PLA2 have been correlated with an increased risk of adverse cardiovascular events, making it a valuable biomarker for assessing cardiovascular risk and guiding treatment decisions. By measuring Lp-PLA2 levels, clinicians gain insights into the inflammatory processes underlying atherosclerosis, thereby facilitating risk stratification and personalized interventions (Adam *et al.*, 2017). In contrast, FLNC, a critical component of muscle structure and function, particularly in sarcomere integrity, is implicated in various cardiovascular conditions, including hypertrophic cardiomyopathy (HCM). Mutations in the FLNC gene, such as the G1676R mutation in domains 14-15, induce significant sarcomeric abnormalities, contributing to the pathogenesis of HCM and elevating the risk of sudden cardiac death. Beyond cardiomyopathy (Matthias & Nobert, 2021), FLNC also plays a role in muscle regeneration and may influence signal transduction and cellular migration during muscle repair processes (Matthias *et al.*, 2005). It's important to note that there are likely other mutations in the Filamin C gene that can also contribute to various health conditions. The specific impact of each mutation can vary, and ongoing research aims to elucidate the precise roles of Filamin C in cellular processes and the consequences of genetic variations in this gene (Shen Song *et al.*, 2021). Integrating insights from both Lp-PLA2 and FLNC provides a comprehensive understanding of the molecular landscape of cardiovascular diseases (Michael *et al.*, 2023). While Lp-PLA2 offers insights into inflammatory processes associated with atherosclerosis and aids in cardiovascular risk assessment, FLNC mutations provide valuable clues about structural abnormalities in cardiac muscle and their implications for various cardiovascular conditions. This type of mutation is referred to as a missense mutation, as it leads to the substitution of one amino acid for another in the resulting protein sequence (Gwen & Kylie, 2022). This integrated approach facilitates risk prediction, diagnosis, and the development of targeted therapeutic interventions tailored to individual patients, ultimately advancing the management of cardiovascular diseases on multiple fronts (Sean *et al.*, 2005).

Biopython, a specialized library in the Python programming language for bioinformatics, provides a versatile set of tools for tasks ranging from sequence analysis to structural bioinformatics (Vinita & Uma, 2023). In the realm of sequence analysis, Biopython facilitates the retrieval of genetic sequences associated with proteins from public databases or experimental data (Ki Chon *et al.*, 2022). Tools like BioPython's pairwise module enable sequence alignment, aiding in the comparison of protein sequences across individuals or populations to identify conserved regions. Moving into structural bioinformatics, Biopython interfaces with tools for the prediction of

the three-dimensional structure of protein. This can provide insights into its functional domains and potential binding sites (Moon *et al.*, 2021). Additionally, BioPython's Bio.PDB module facilitates the visualization of protein structures, aiding in the interpretation of protein's structural characteristics relevant to its function in cardiovascular disease (Priyash *et al.*, 2023).

Methodology

Developing a comprehensive methodology for investigating both Human Lipoprotein-Associated Phospholipase A2 (Lp-PLA2) with PDB code 5JAD and the Filamin C domain 14-15 mutation G1676R with PDB code 7P0E in cardiovascular disease involves several key steps. Initially, biological samples are retrieved and prepared using the Protein Data Bank (PDB), accessing the 3D coordinates of the proteins. BLAST searches are then conducted to analyze E-values and sequence similarities, facilitating the identification of homologous proteins and evolutionary relationships. Structural analysis is performed using RasMol and PyMOL to visualize the 3D structures, identify relevant domains or regions, and analyze active sites. RMSD calculations are conducted to compare structural deviations induced by mutations. Python scripts using Biopython are employed for command-based structural analyses, including distance and angle calculations. Conversion of the 3D structures to their corresponding amino acid sequences is done using the Molecular Modeling Database (MMDB), allowing for further analysis. Quality estimation of the protein models is assessed using ERRAT and PROCHECK. Potential binding sites are predicted using the CB Dock server for molecular docking studies. Structural classification is carried out using the CATH database, categorizing the proteins into structural classes and domains. Gene Ontology terms associated with Lp-PLA2 and the mutated Filamin C protein are retrieved and analyzed to understand their biological roles. Pathway analysis using KEGG pathway analysis tools explores interactions with other proteins and pathways related to cardiovascular disease. Ramachandran plot analysis, along with further protein structure and interaction analysis using the PDBsum server and functional analysis using the InterPro server, is conducted to identify conserved domains and functional sites affected by mutations. Additionally, BioPython is employed for additional structural analysis tasks and to streamline data processing. This comprehensive methodology integrates various computational tools and servers to elucidate the molecular consequences of mutations in Lp-PLA2 and Filamin C, providing insights into their potential implications for protein function and disease pathogenesis in cardiovascular disease research.

Results

Chain A, Platelet-activating factor acetylhydrolase

PDB: 5JAD_A

>pdb|5JAD|A Chain A, Platelet-activating factor acetylhydrolase

```
MAAASFGQTKIPRNGPYVSGCTDLMFDHTNKGTFRLRYPSQDNDRLDRTLWIPNKEYFWGLSKFLG
THLMGNILRLLFGSMTPANWNSPLRPGEKYPLVVFHSHGLGAFRTLYSAIGIDLASHGFVAAVEHRDR
SASATYYFKDQSAEIGKSWLYLRTLKQEEETHIRNEQVRQRAKECSQALSLILDIDHGKPKVKNALDLK
FDMEQLKDSIDREKIAVIGHSFGGAVIQTLSEDQRFRCGIALDAWMFPLGDEVYSRIPQPLFFINSEYFQY
PANIIMKMKCYSPDKERKMITIRGVSVHQNFADFTFATGKIIGHMLKLKGDIDSNVAIDLNSKASLAFLOK
HLGLHKDFDQWDCLIEGDENLIPGTNINTTNQHSHHHHH
```

Figure: PDB (Biological Sample)

Chain A, Isoform 2 of Filamin-C

PDB: 7P0E_A

>pdb|7P0E|A Chain A, Isoform 2 of Filamin-C

```
GPLPAHDASKVRASGPGPLNASGIPASLPVEFTIDARDAGEGLLTVQILDPEGKPKKANIRDNGDGTYYT
S
YLPDMSGRYTITIKYGGDEIPYSPFRIHALPTGDASKCLVTVSIGGHGLGACLPRIQIGQETVITVDAK
AAGERKVTVSTPDGAELDVDVVENHDGTFDIYYTAPPEPGKYVITIRFGGEHIPNSPFHVLATE
```

Figure: PDB (Biological Sample)

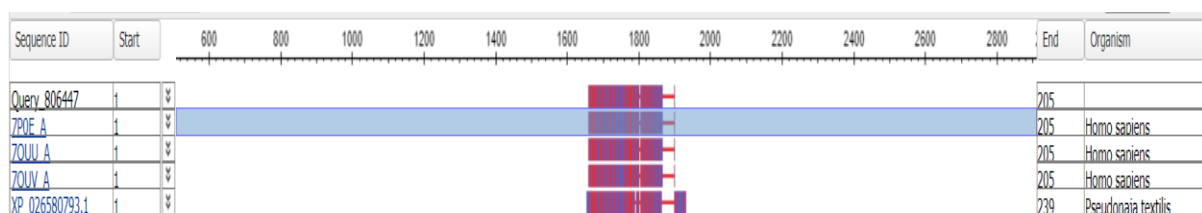


Figure: Representation of color from red to blue reflects the size of Amino acids (Red representing smaller side chain and Blue represents larger side chain).

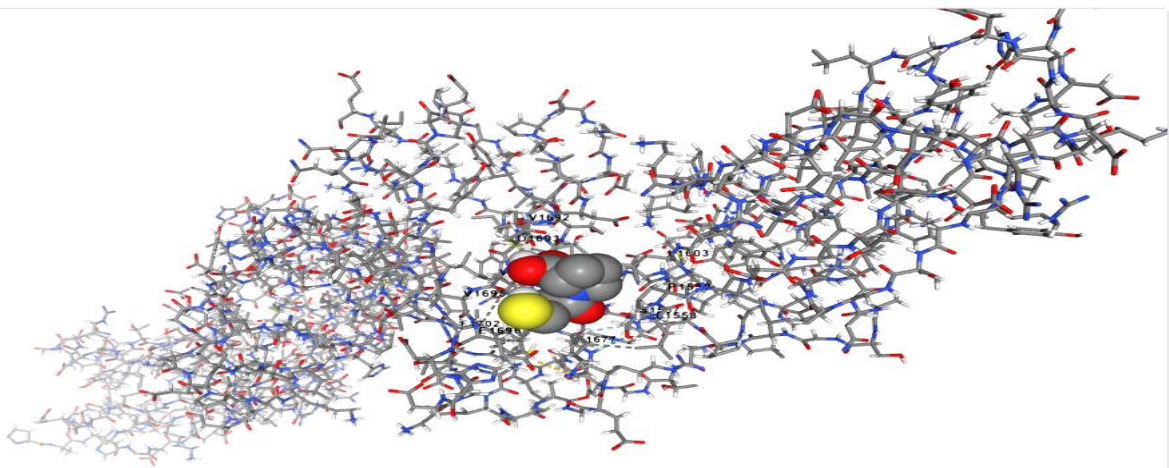


Figure: Protein-ligand interaction with score -4.6 with Captopril ligand with CBDock showing very good binding and considered as the cure drug for a human cardiovascular biological sample.

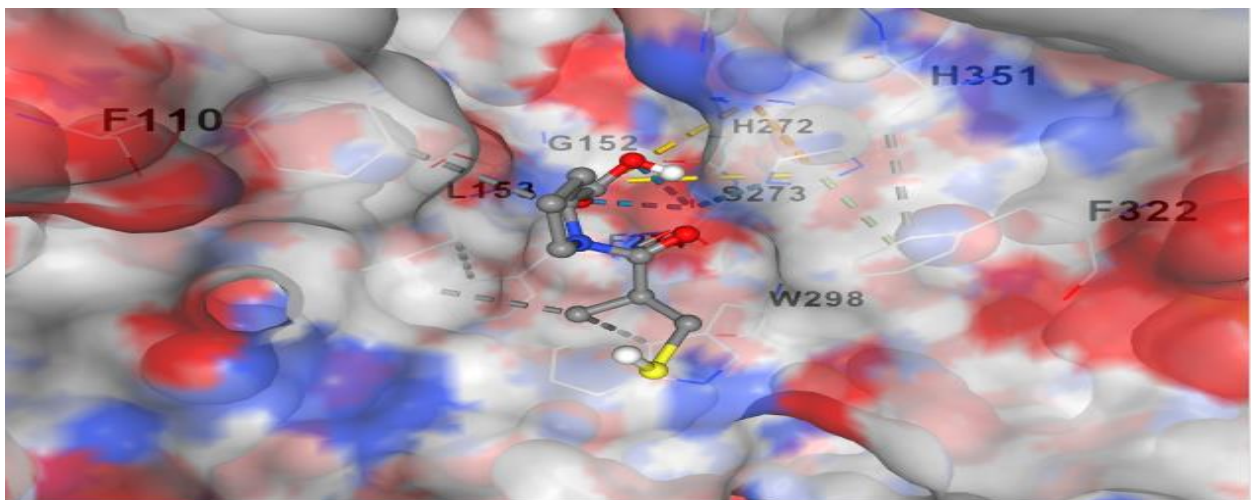


Figure:Captopril Docking result.

Table: Representing the Binding Score of the Effective Drugs.

PDB Sample code	Drug Name	Vina Score	Centre			Docking Size		
			x	y	z	x	y	z
5JAD	Aspirin	-5.2	19	6	-5	26	17	17
	Captopril	-5	19	6	-5	26	18	18
7P0E	Benazepril	-6.9	-12	-1	10	25	18	18
	Captopril	-4.6	-3	-15	-9	23	23	23
	Mexiletine	-5.1	-12	-1	10	25	18	18

[9] !drug-db show-drug Mexiletine

Scaling factors for drug Mexiletine and FPC 1

Name	Value
scale_drug_INa	0.9991
scale_drug_INaI	0.7491
scale_drug_Ito	1.0
scale_drug_ICaI	0.904
scale_drug_IKr	0.9938
scale_drug_IKs	1.0
scale_drug_IK1	1.0
scale_drug_IKb	1.0
scale_drug_INab	1.0
scale_drug_ICab	1.0
scale_drug_IpCa	1.0
scale_drug_Isacns	1.0
scale_drug_Isack	1.0

Figure: drug work using python representing the scaling factor

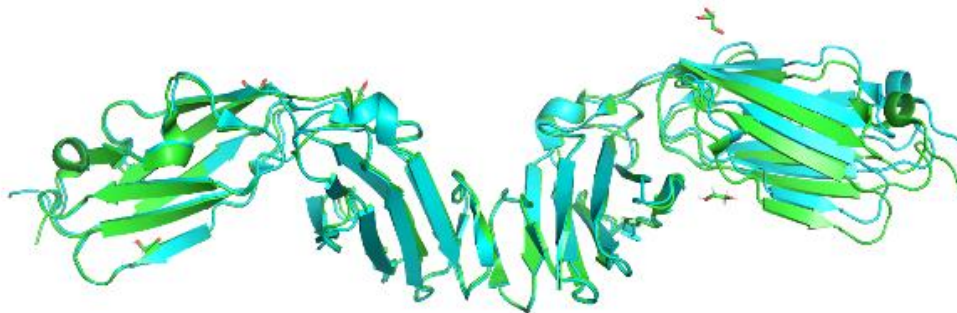


Figure: RMSD value score calculation(7P0E,7OUU) square root of the mean showing the distance between the matched atoms determining the RMSD values. RMSD = 0.711 (4521 to 4521 atoms)

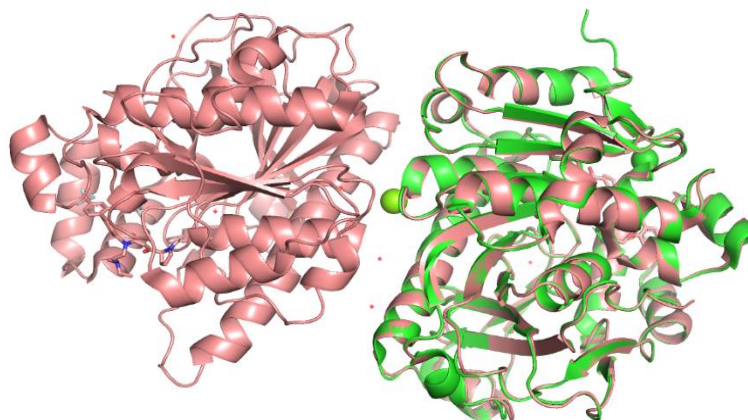


Figure: RMSD Value score calculation (6m07,5JAD) is 0.312 which is close to zero and represents the good binding of both the proteins.

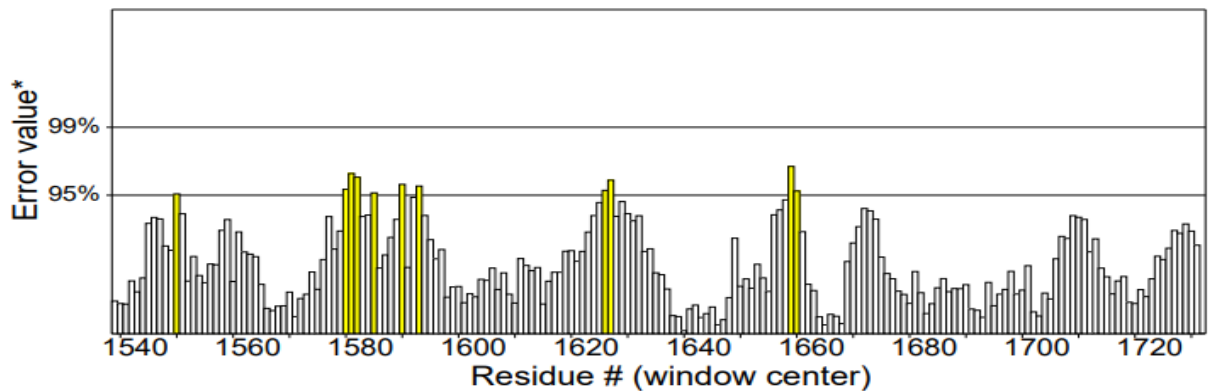


Figure: ERRATE analysis (structure validation 94.5) yellow showing the rejected protein part.

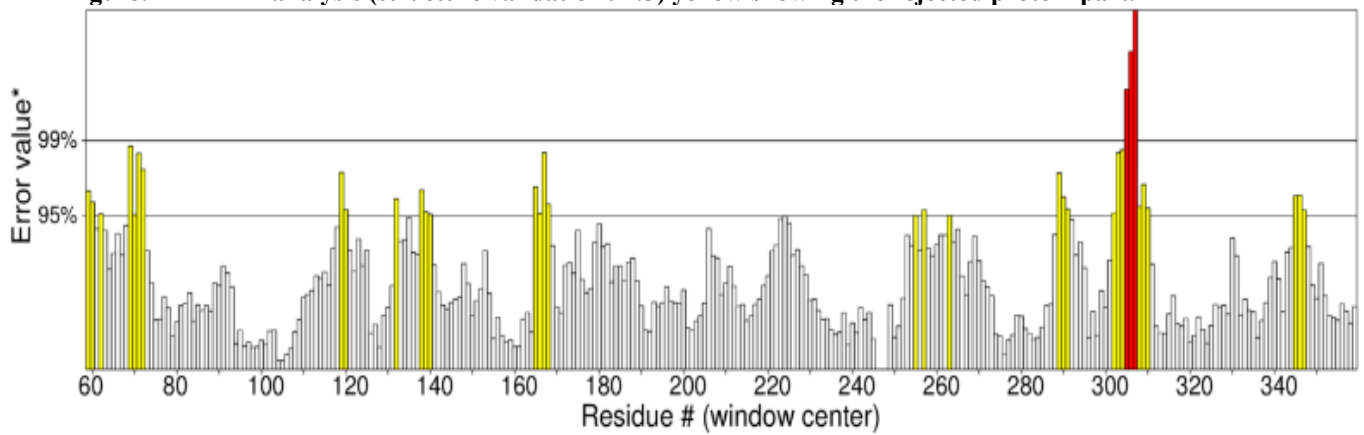


Figure: ERRAT result shows different color phases (yellow showing the rejected protein part and red color representing the not working protein part).

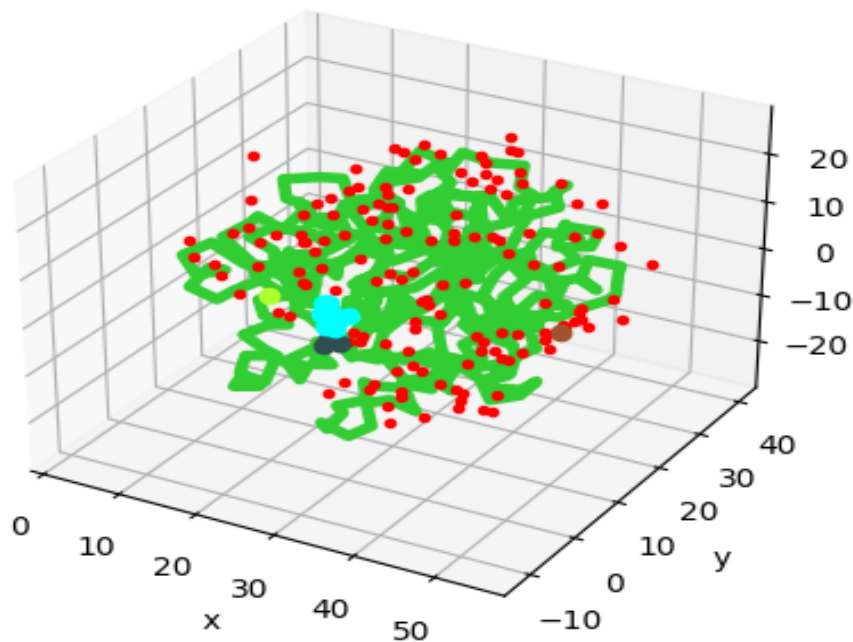


Figure: Structure Visualization of the protein (5JAD) with the help of Bio.PDB in BioPython.

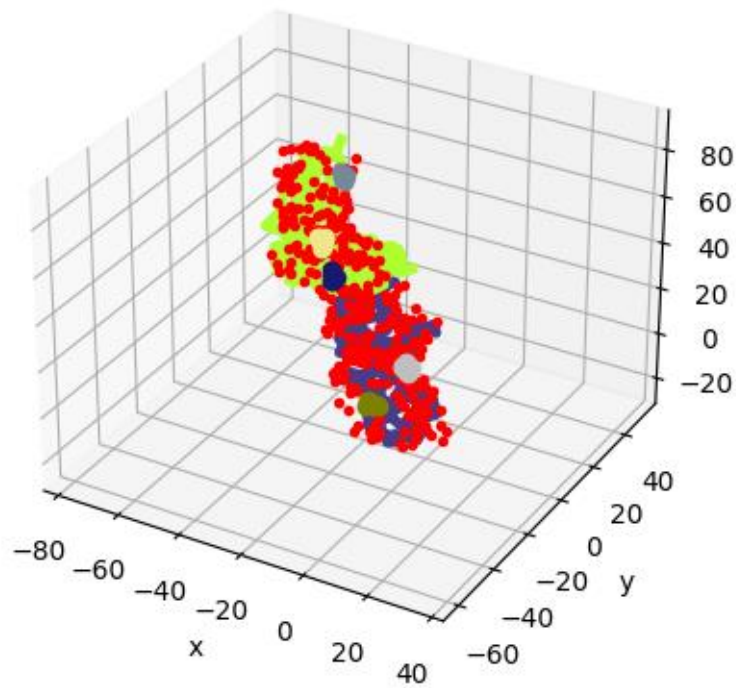


Figure: Representing the Structure Visualization of the protein (7P0E) with the help of Bio.PDB in BioPython.

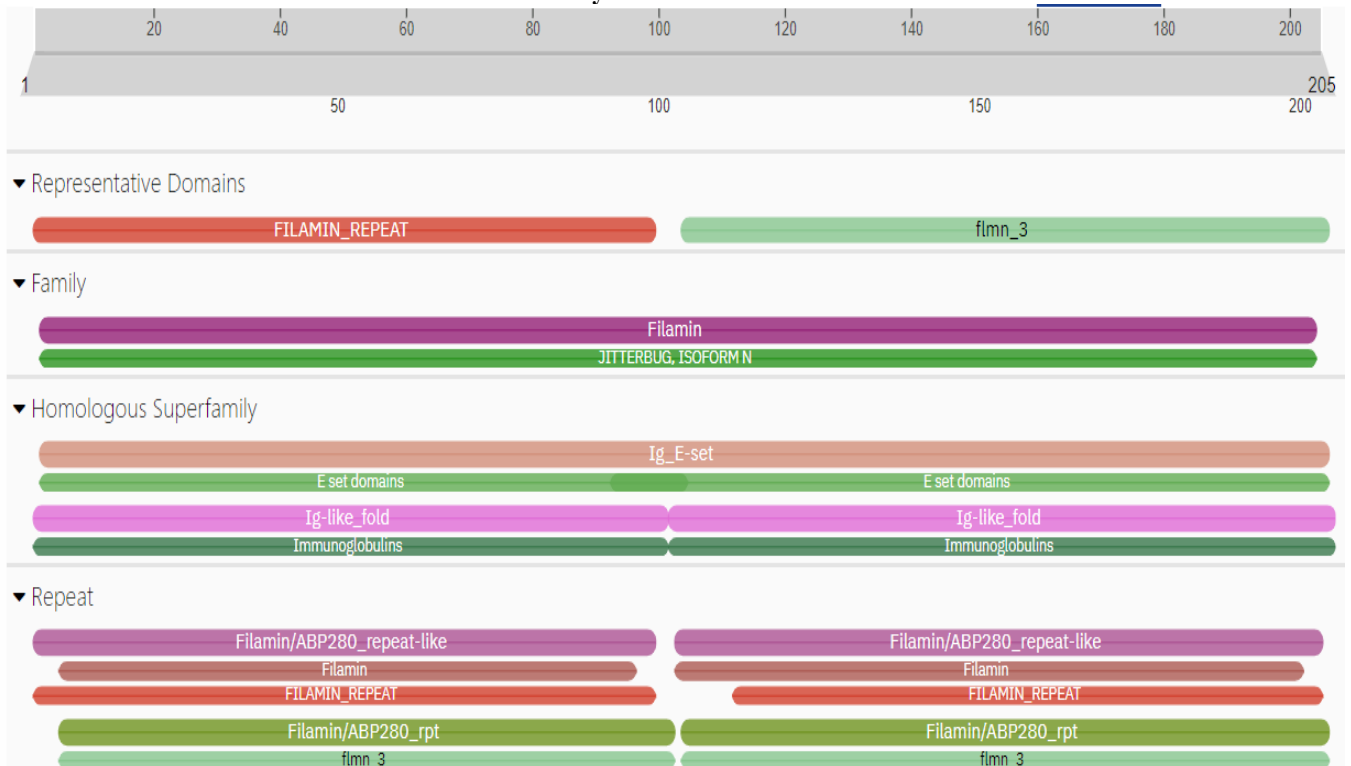


Figure: Functional Domain analysis (InterPro) representing the domain features.

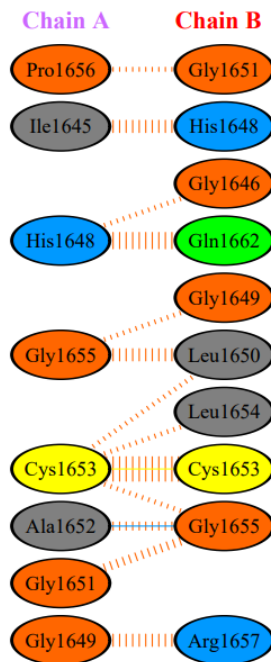


Figure: Protein-protein interface shows residue interactions across interfaces colored by residue type.

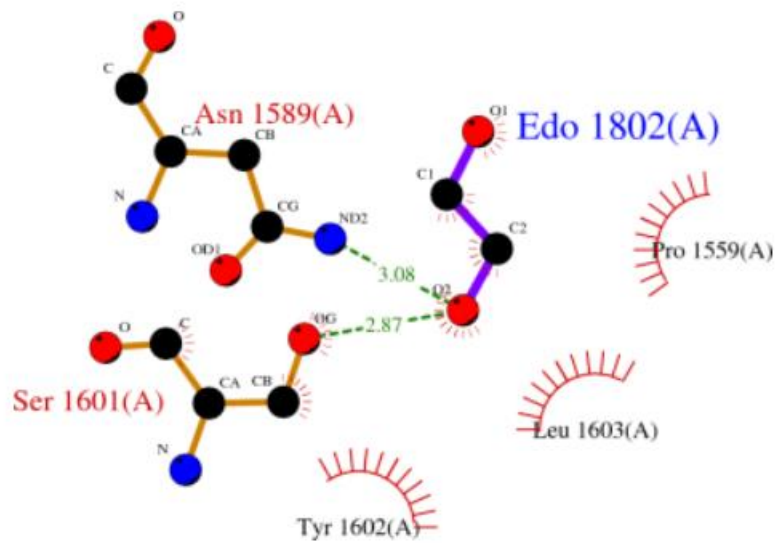


Figure: Ligplot of interactions involving ligand EDO of 7P0E.

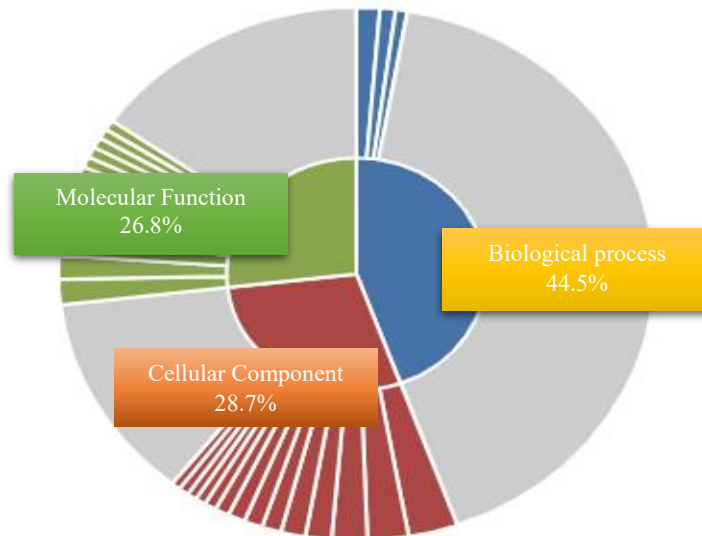


Figure: Gene ontology result showing the different components of the protein (5JAD)

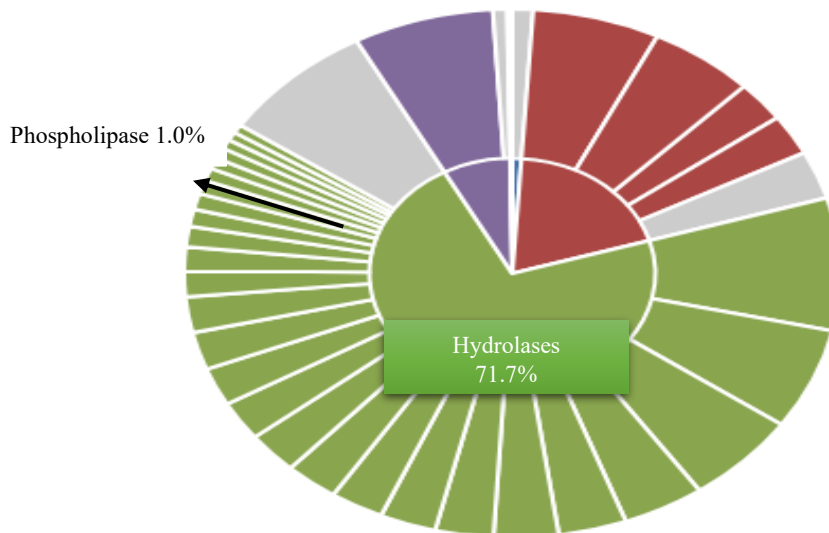
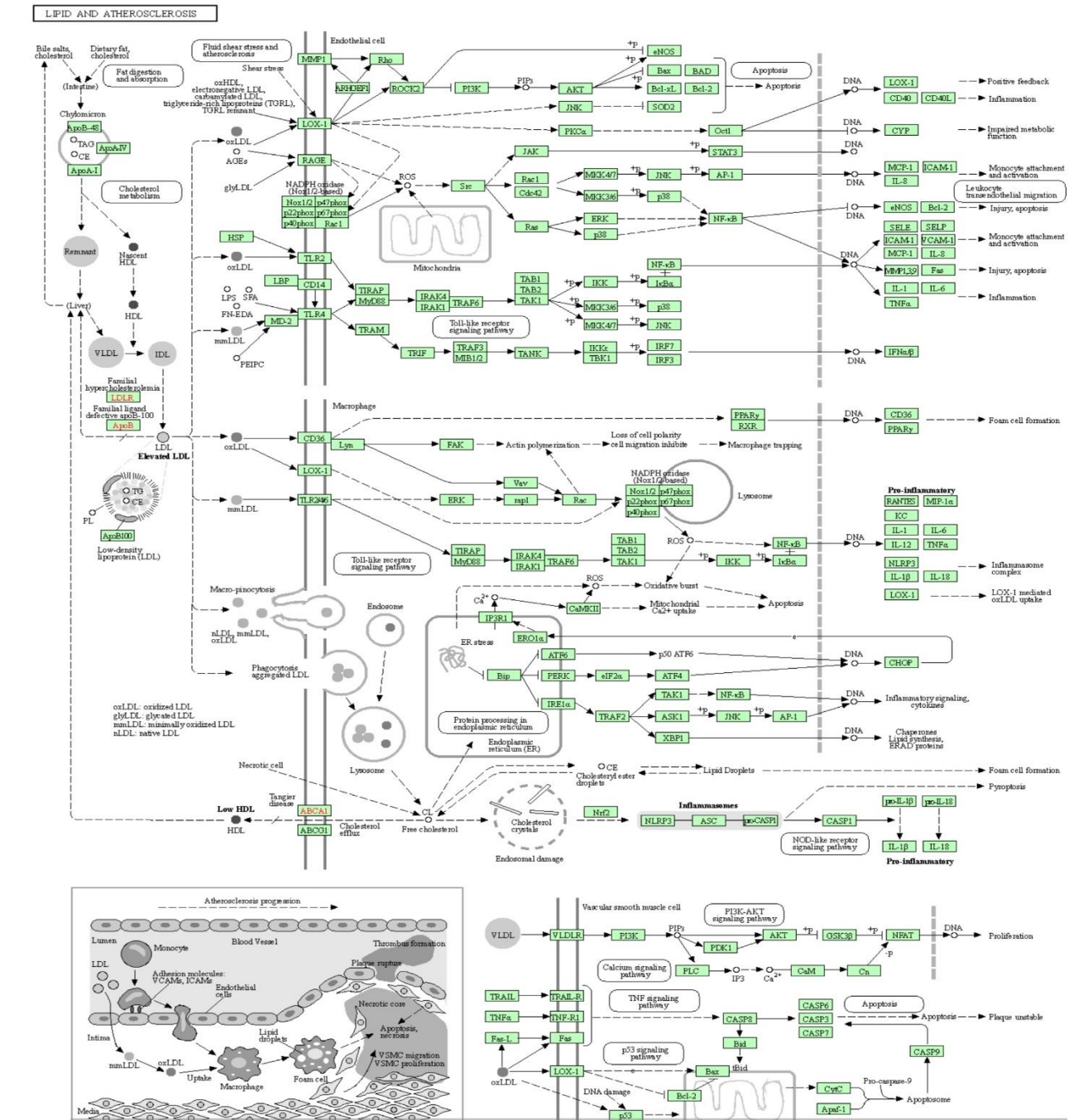


Figure: EC diversity of the 5JAD showing the majority part of the Hydrolases class of the enzyme.



Conclusion

In conclusion, our comprehensive methodology, underpinned by the intricate exploration of Human Lipoprotein-Associated Phospholipase A2 (Lp-PLA2) and the Filamin C domain 14-15 mutation G1676R within the context of cardiovascular disease, marks a significant advancement in our understanding of this pervasive health challenge. Through a synergistic integration of computational prowess, biochemical insights, and genetic exploration, we embarked on a profound journey into the heart of cardiovascular pathogenesis. From the initial stages of sample retrieval and preparation from the Protein Data Bank (PDB) to the nuanced analyses of genetic sequences and protein structures, our methodology meticulously unfolded the intricate landscape of cardiovascular disease etiology. Employing a sophisticated array of analytical tools including BLAST searches, RasMol and PyMOL visualization, RMSD calculations, and Python-driven analyses, we meticulously dissected the structural and functional nuances of Lp-PLA2 and the Filamin C mutation. Our methodology transcended mere structural elucidation, delving into the deeper realms of sequence analysis,

structural classification, and pathway elucidation. With the aid of invaluable resources such as the Molecular Modeling Database (MMDB), CATH database, and KEGG pathway analysis tools, we untangled the complex web of genetic variants, structural motifs, and molecular pathways implicated in cardiovascular disease. Moreover, through meticulous analyses including Ramachandran plots, PDBsum server insights, InterPro domain exploration, and BioPython scripting, we delved into the molecular intricacies of conserved domains, pivotal residues, and molecular interactions modulated by these mutations.

Importantly, our docking studies revealed promising interactions between the Lp-PLA2 enzyme and the drug Captopril, indicating effective binding that holds potential therapeutic implications. This finding underscores the significance of targeting Lp-PLA2 in cardiovascular disease management and suggests avenues for further exploration in drug development and personalized medicine strategies.

In summation, our integrated methodology stands as a testament to the collaborative efforts of multidisciplinary research, offering a panoramic view of the genetic, structural, and functional landscape of cardiovascular health. By unraveling the mysteries surrounding Lp-PLA2 and the Filamin C mutation, we pave the way for transformative advancements in diagnostics, therapeutics, and personalized medicine, ultimately ushering in a new era of cardiovascular disease management with enhanced precision and efficacy.

References

1. Adam Oesterle, U. L. (2017). Pleiotropic Effects of Statins on the Cardiovascular System. *Circ Res*, 120(1), 229-243.
2. Andrew Zalewski, C. M. (2005). Role of Lipoprotein-Associated Phospholipase A2 in Atherosclerosis. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 25, 923–931.
3. Arup Kr. Malakar, D. C. (2019). A review of coronary artery disease, its risk factors, and therapeutics. *Journal of Cellular Physiology*, 234(10), 16812-16823.
4. Brian W. Wong, n. M. (2012). The Biological Role of Inflammation in Atherosclerosis. *Canadian Journal of Cardiology*, 28(6), 631-641.
5. Devin Hasanally, R. C. (2014). Role of Phospholipases and Oxidized Phospholipids in Inflammation. *Phospholipases in Health and Disease*, 10, 55–72.
6. Fubao Huang, K. W. (2019). Lipoprotein-associated phospholipase A2: The story continues. *Medicinal Research Reviews*, 40(1), 79-134.
7. Fiskesund, R. (2013). *Anti-Phosphorylcholine Antibodies In Cardiovascular Disease*. Sweden: Stockholm.
8. Gwen R. Buel, K. J. (2022). Can AlphaFold2 predict the impact of missense mutations on structure? *Nature structural & molecular biology*, 29, 1-2.
9. Hitesh Shukla, J. L. (2021). Identifying genetic markers associated with susceptibility to cardiovascular diseases. *Advances in Medical Biochemistry, Genomics, Physiology, and Pathology*, 177-200.
10. Haider, R. (2023). Cardiovascular Risk Factors. *Toxicology and Applied Pharmacology Insights*, 6(1), 77-87.
11. Jakub Piotr Fichna, A. M. (2018). Myofibrillar myopathy in the genomic context. *Journal of Applied Genetics*, 59, 431–439.
12. Kavitha Mukund, S. S. (2019). Skeletal muscle: A review of molecular structure and function, in health and disease. *WIREs Systems Biology and Medicine*, 12(1), 1462.
13. Ki Chon, J. M.-Q. (2022). Cardiovascular Disease Prediction Via Literature Embedding Model Using Risk Factors, Symptoms, and Genotype Information. *ymptoms, and Genotype Information*, 21.
14. Matthias Eden, N. F. (2021). Cardiac Filaminopathies: Illuminating the Divergent Role of Filamin C Mutations in Human Cardiomyopathy. *Journal of Clinical Medicine*, 10(4), 577.
15. Matthias Vorgerd, P. F. (2005). A Mutation in the Dimerization Domain of Filamin C Causes a Novel Type of Autosomal Dominant Myofibrillar Myopathy. *The American Journal of Human Genetics*, 77(2), 297-304.
16. Miano, J. M. (2014). Myocardin in biology and disease. *The Journal of Biomedical Research*, 29(1), 3-19.
17. Michel M. Joosten, J. K. (2012). Associations Between Conventional Cardiovascular Risk Factors and Risk of Peripheral Artery Disease in Men. *Jama*, 308(16), 1660-1667.
18. Moon, J., Posada-Quintero, H. F., Kim, I., & Chon, K. H. (2021). Preliminary Analysis of the Risk Factor Identification Embedding Model for Cardiovascular Disease. *International Conference of the IEEE Engineering in Medicine & Biology Society*, 1946-1949.
19. M. Zaromitidou, G. S. (2016). Chapter 2 - Atherosclerosis and Coronary Artery Disease: From Basics to Genetics. *Cardiovascular Diseases*, 3-24.

20. Miyazaki, T. (2023). Calpain and Cardiometabolic Diseases. *International Journal of Molecular Sciences*, 24(23), 16782.
21. Priya Arya, U. K. (2024). Molecular Docking Approach For Human Oncology Research With Chemoinformatic. *Journal of Emerging Technologies and Innovative Research*, 11(2), 658-666.
22. Sekhara Rao Basavaraju, C. E. (2002). Pathophysiological effects of radiation on atherosclerosis development and progression, and the incidence of cardiovascular complications. *Medical Physics*, 29(10), 2391-2403.
23. Sean C. Goetsch, C. M. (2005). Myogenic Progenitor Cells Express Filamin C in Developing and Regenerating Skeletal Muscle. *Stem Cells and Development*, 14(2), 181-187.
24. Shen Song, A. S. (2021). Filamin C in cardiomyopathy: from physiological roles to DNA variants. *Heart Failure Reviews*, 27, 1373–1385.
25. Uma Kumari, K. T. (2023). Computational Analysis and Molecular Docking Approach for Liver Cancer. *International Journal for Research in Applied Science & Engineering Technology*, 11(7), 1628-1633.
26. Uma Kumari, S. G. (2023). NGS and Sequence Analysis with Biopython for Prospective Brain Cancer Therapeutic Studies. *International Journal for Research in Applied Science & Engineering Technology*, 11(4), 3318-3329.
27. Vinita Kukreja, U. K. (2023). Data Analysis of Brain Cancer with Biopython. *International Journal of Innovative Science and Research Technology*, 8(3), 2146-2154.
28. Zhenfeng Mao, F. N. (2020). Structure and Function of Filamin C in the Muscle Z-Disc. *International Journal of Molecular Sciences*, 21(18), 2696.