COMPUTATIONAL ENRICHMENT ANALYSIS OF SAP PROTEIN ON LYSOSOMAL STORAGE DISORDER – METACHROMATIC LEUKODYSTROPHY

FSP2999–Full Semester Project Report

a project report submitted by

ANTONITA RACHEAL (URK17BI013)

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

in

BIOTECHNOLOGY

Under the supervision of

DR. PREMNATH (Internal Supervisor)

Dr. SAHAYA PRAVIN (External Supervisor)

DEPARTMENT OF BIOTECHNOLOGY

KARUNYA INSTITUTE OF TECHNOLOGY AND SCIENCES

(Deemed-to-be-University)

Karunya Nagar, Coimbatore – 641 114, INDIA

APRIL 2021

BONAFIDE CERTIFICATE

Certified that this project report **COMPUTATIONAL ENRICHMENT ANALYSIS OF SAP PROTEIN ON LYSOSOMAL STORAGE DISORDER – METACHROMATIC LEUKODYSTROPHY** is the bonafide work of **"Antonita Racheal (Reg: No. URK17BI013)"** who carried out the project work under my supervision.

SIGNATURE

Dr. JIBU THOMAS
HEAD OF THE DEPARTMENT
Professor of Biotechnology
Department of Biotechnology

SIGNATURE

Dr. PREMNATH DHANARAJ
SUPERVISOR
Assistant Professor
Department of Biotechnology

Submitted for the Full Semester Viva Voce held on …….

Internal Examiner **External Examiner**

Annexure 1

الاختصاصي للحلول الطبية ذ.م.م.
SPECIALIZED MEDICAL SOLUTIONS W.L.L.

30 April 2021

Attendance Certificate

This is to certify that **Ms. Antonita Racheal** pursuing IV B. Tech of Karunya Institute of Technology and Sciences has carried out her Project work entitled "Computation Enrichment Analysis of SAP Protein on Lysosomal Storage Disorder – Metachromatic Leukodystrophy under the supervision of Dr. Sahaya Pravin, from 4th January 2021 to 30th April 2021 in the company Specialized Medical Solutions, Qatar.

30/April/21

Dr. Sahaya Pravin
Lab Manager

س.ت : ١٣٤٢٩٣ - ص.ب : ٢٠٣٨٧ - الدوحة - قطر - تليفون : ٤٤٦٢٠٤٩١ ٩٧٤+ - فاكس : ٤٤٦٢٦٧٦١ ٩٧٤+
C.R. No.: 134293 - P.O. BOX : 20387 - Doha Qatar - Tel. : +974 44620491 - Fax : +974 44626761
Email : info@smsqatar.org : البريد الكتروني Web : www.smsqatar.org

Annexure II

COMPUTATIONAL ENRICHMENT ANALYSIS OF SAP PROTEIN ON LYSOSOMAL STORAGE DISORDER – METACHROMATIC LEUKODYSTROPHY

ANTONITA RACHEAL (REG NO.URK17BI013)

Dr.Premanth Dhanaraj, Assistant Professor

Dr. Sahaya Pravin, Lab Manager
Address: Building No.11, Zone 43, Street 640, Al-Mamoura,
P.O. BOX: 20387 – Doha, Qatar
Phone no: +97455146963
Email ID: pravin@smsqatar.org

Abstract

A rare, yet fatal disease, known as Metachromatic leukodystrophy (MLD), is a neurodegenerative, hereditary malady, which falls under lysosomal storage disorder (LSD). The frequency rate for MLD was estimated to be about 1 per 40,000. However, the prognosis is severe in which more than 50 percent of the patients die at a late infantile stage, which is approximately a lifespan of 5 years. Lack of ARSA enzyme, paves way for the accumulation of sulfatides, in lysosomes. As a result of this sulfatide deposition, the protective fatty layer, myelin sheath, guarding nerves in both the central and peripheral nervous systems is damaged. It is to be noted that no other molecule can compensate for the ARSA enzyme.

No cure for MLD is available as of now. Treatment of the condition is aimed at the improvement and regulation of the symptoms. This study reports, the analysis done with the help of research methods, that includes quantitative and qualitative data, gained from machine learning algorithms and computational tools, to study particularly deleterious mutations that occur in the *ARSA* gene of MLD patients, which will foster future wet lab research and molecular diagnosis.

Keywords: Metachromatic leukodystrophy (MLD), Lysosomal Storage Disorder (LSD), ARSA gene, ARSA enzyme, Sulfatides.

ACKNOWLEDGEMENT

First and foremost, I praise and thank **ALMIGHTY GOD** for being with me through out the project.

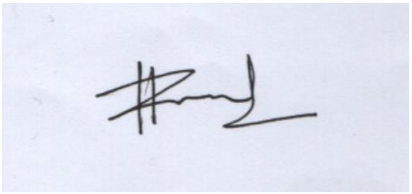
I am grateful to our beloved founders **Late Dr. D.G.S Dhinakaran, C.A.I.I.B., Ph.D** and **Dr. Paul Dhinakaran, M.B.A, Ph.D,** for their constant prayers and blessings.

I extend my thanks to our Vice-Chancellor, **Dr. P. Mannar Jawahar**, Pro Vice Chancellor (Research and Collaborations) **Dr.E.J.James,** Pro-Vice Chancellor (Quality Standards) **Dr. Ridling Margaret Waller and** Registrar, **Dr. R. Elijah Blessing**, for giving me this opportunity to do the project.

I express my sincere thanks to our respected Dean, **Dr. Sajan Kurien**, for motivating us to carry out the project successfully.

My heartfelt thanks to our respected Head of the Department, **Dr. Jibu Thomas** for expressing keen interest in our academic pursuits and facilitating us to carry out this project work successfully.

I Express my sincere gratitude to my guide **Dr.Premanth Dhanaraj** for supporting and guiding me throughout my project



Antonita Racheal

TABLE OF CONTENTS

	Page No.
BONAFIDE CERTIFICATE	ii
ABSTRACT	iv
TABLE OF CONTENTS	vi
LIST OF TABLES	vii
LIST OF FIGURES	viii
LIST OF SYMBOLS AND ABBREVIATIONS	ix
1. INTRODUCTION	1
1.1 Motivation and background	
1.2 Aim and objective of the research	
2. LITERATURE REVIEW	4
2.1 Metachromatic Leukodystrophy (MLD)	
2.1.0 Inheritance	
2.1.1 Types of MLD	
2.1.2 Signs and Symptoms	
2.1.3 Causes	
2.1.4 Frequency	
2.1.5 Epidemiology	
2.1.6 Diagnosis	
2.1.7 Treatment	
2.2 Conclusion obtained from literature review	
2.3 Research Gap	
2.4 Formulation of the problem	
2.5 Rationale of the study	
3. MATERIALS AND METHODS	10
3.1 Data Collection	
3.1.1 Identifying the gene target	
3.1.2 Retrieval of SAP mutations	
3.2 Data Prediction and screening	
3.3 Dataset of *in-silico* mutant modeling	
3.4 Energy Minimization	
3.5 Single Model Analysis	
3.5.1 Intramolecular Interaction	
3.5.2 Secondary Structure Analysis	
3.5.3 Fluctuation Analysis	
3.5.4 Contact Map Analysis	
3.5.4 Solvent Accessible Surface Areas (SASA)	
3.6 Docking	
4. RESULTS	36
5. DISCUSSION	38
6. CONCLUSION	39
7. REFERENCES	40

LIST OF TABLES

TABLE NO.	TITLE
TABLE 1	LIST OF SUPPORT VECTOR MACHINE BASED SCREENED DELETERIOUS MUTATIONS
TABLE 2	ENERGY MINIMIZATION SCORE FOR EACH MUTANT FROM YASARA PROGRAM.
TABLE 3	INTRAMOLECULAR INTERACTIONS - PART 1
TABLE 4	INTRAMOLECULAR INTERACTIONS - PART 2
TABLE 5	OVERALL PERCENTAGE OF SECONDARY STRUCTURE ANALYSIS
TABLE 6	OVERALL CONTACT MAP ANALYSIS
TABLE 7	SASA CALCULATIONS - PART 1
TABLE 8	SASA CALCULATIONS - PART 2
TABLE 9	DOCKING

LIST OF FIGURES

FIGURE NO.	TITLE
FIGURE 1	LYSOSOMAL STORAGE DISORDER
FIGURE 2	METACHROMATIC LEUKODYSTROPHY
FIGURE 3	T274M
FIGURE 4	R311Q
FIGURE 5	R370W
FIGURE 6	R390W
FIGURE 7	C489G
FIGURE 8	HYDROPHOBIC INTERACTIONS
FIGURE 9	SECONDARY STRUCTURE ANALYSIS - 1AUK
FIGURE 10	1ST LINE COMPARISON - SECONDARY STRUCTURE ANALYSIS
FIGURE 11	2ND LINE COMPARISON - SECONDARY STRUCTURE ANALYSIS
FIGURE 12	3RD LINE COMPARISON - SECONDARY STRUCTURE ANALYSIS
FIGURE 13	4TH LINE COMPARISON - SECONDARY STRUCTURE ANALYSIS
FIGURE 14	5TH LINE COMPARISON - SECONDARY STRUCTURE ANALYSIS
FIGURE 15	BAR GRAPH - REPRESENTING PERCENTAGE OF EACH STATE
FIGURE 16	T274M FLUCTUATION PLOT
FIGURE 17	R311Q FLUCTUATION PLOT
FIGURE 18	R370W FLUCTUATION PLOT
FIGURE 19	R390W FLUCTUATION PLOT
FIGURE 20	C489G FLUCTUATION PLOT
FIGURE 21	T274M CMA
FIGURE 22	R311Q CMA
FIGURE 23	R370W CMA
FIGURE 24	R390W CMA
FIGURE 25	C489G CMA
FIGURE 26	STEP 1
FIGURE 27	STEP 2
FIGURE 28	STEP 3
FIGURE 29	STEP 4
FIGURE 30	STEP 5
FIGURE 31	T274M
FIGURE 32	R311Q
FIGURE 33	R370W
FIGURE 34	R390W
FIGURE 35	C489G
FIGURE 36	CONTACT MAP OVERLAP
FIGURE 37	AFFINITY

LIST OF ABBREVATIONS

ABBREVATIONS	DETAIL
LSD	Lysosomal Storage Disorder
MLD	Metachromatic leukodystrophy
ARSA	Arylsulfatase A
SAP	Single Amino Acid Polymorphism
SVM	Support Vector Machines
MRI	Magnetic Resonance Imaging
CMV	Contact Map View
PDB	Protein Data Bank
SASA	Solvent Accessible Surface Areas

1. Introduction

1.1 Motivation and Background

Lysosome, also known as suicide bags of the cell, are responsible for disposing of unwanted and dead cell organelles, with the help of digestive enzymes. When the cell is damaged beyond repair, it proceeds to self-destruct, causing it to burst open and release the digestive enzyme, which digests its own cell, causing the cell to die. Macromolecules such as proteins, lipids, and carbohydrates, are degraded through this process (Futerman & van Meer, 2004).

Lysosomal Storage Disorder (LSD), is a group of rare genetic metabolic disorders, where the accumulation of toxic substances takes place in the cells of the body, particularly lysosomes. The deposition of these toxic substances is caused due to the deficiency of enzymes, which are essential to speed up the functional process of the human body. This build-up of harmful substances leads to damaged cells and organs in the body. More than 50 types of rare disorders are found under LSD (https://www.facebook.com/WebMD, 2016).

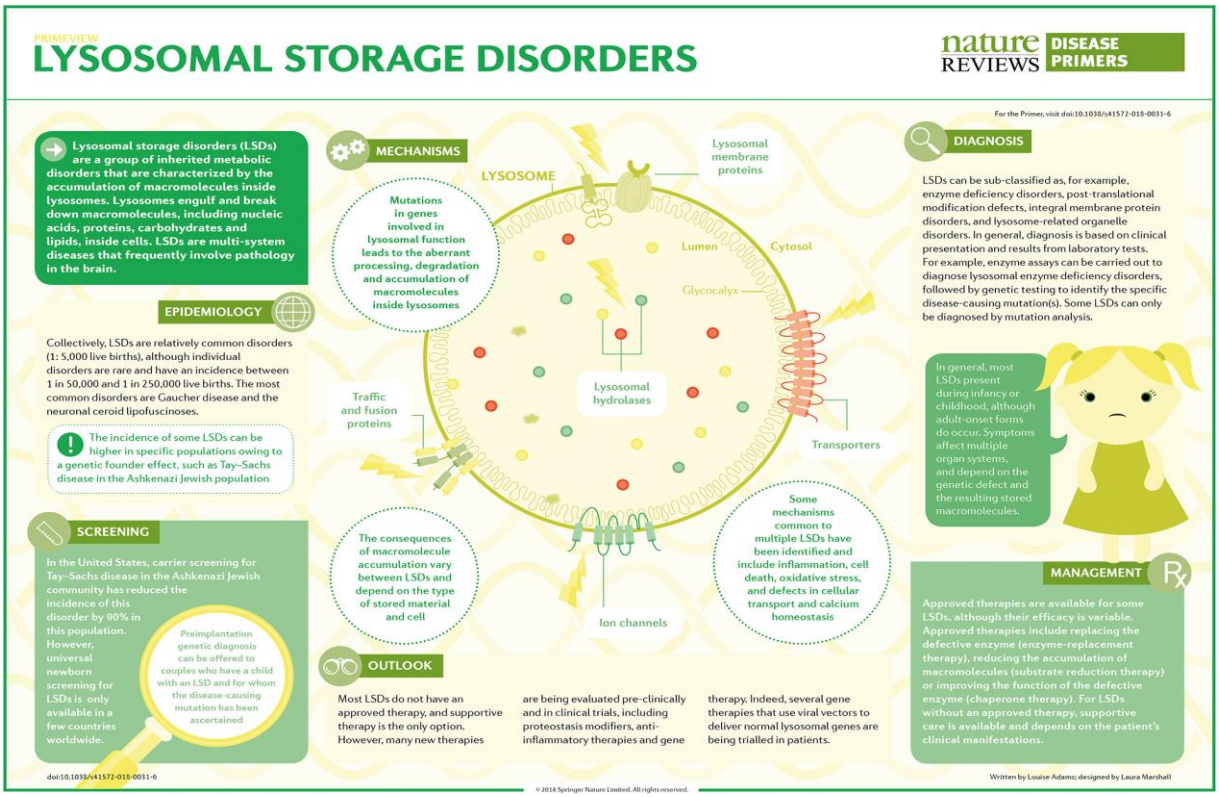


FIGURE 1 - LYSOSOMAL STORAGE DISORDER

Almost all of the LSD are autosomal recessive, except 3, which are x- linked disorders are Fabry disease X-linked disorders, Hunter syndrome, and Danon disease (Staretz-Chacham, Lang, LaMarca, Krasnewich, & Sidransky, 2009). LSD is classified into sphingolipidoses, oligosaccharides, mucolipidosis, mucopolysaccharidoses (MPSs), lipoprotein storage disorders, lysosomal transport defects, neuronal ceroid lipofuscinoses, and many others, it is based on the substance which is accumulated (Nelson, 2021).

Under LSD categorization, I have chosen to work with Metachromatic leukodystrophy (MLD) (OMIM 250100), which falls under Sphingolipidoses (Pastores, 2010). It is an autosomal recessive, rare disorder. Most of the symptoms and signs for MLD are visible during late infantile and juvenile MLD, which is around the age between 1 - 16 years. Multiple disabilities followed by an increase in mortality rate (Harrington et al., 2019), is why I choose to work with MLD. Furthermore, being a hereditary disease, it would require molecular diagnosis (Narayanan et al., 2019), as the changes begin at a genetic level. Thus, with the help of computational tools and software, this study will provide a comparison of the deleterious mutations, which are present in the gene of MLD.

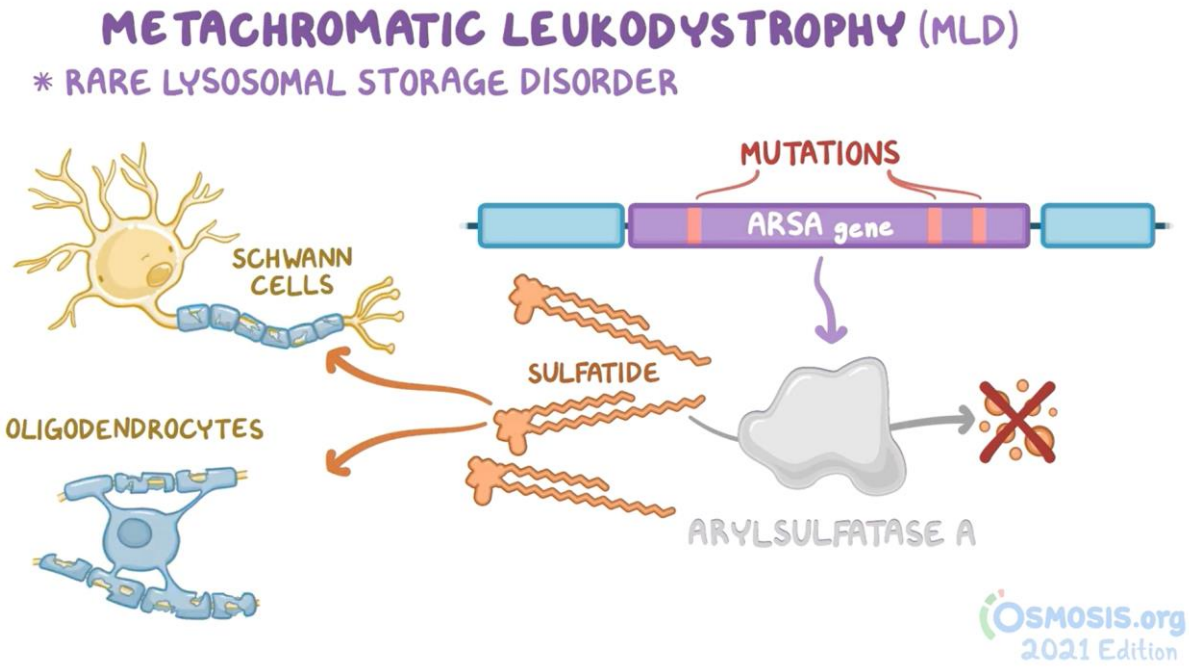


FIGURE 2 - METACHROMATIC LEUKODYSTROPHY

1.2 Aim and Objective

Aim:

To determine deleterious SAPs and perform comparative and structural analysis in the ARSA gene in patients with metachromatic leukodystrophy (MLD).

Objective:

- ➔ Identification of gene target and retrieval of SAP sequences from various data sources.
- ➔ The retrieved dataset is validated and screened for identifying deleterious SAP's using SNPs&GO based on Support Vector Machines (SVM).
- ➔ The Selected Protein is Modelled, Validated, and Mutated using various software's.
- ➔ Mutated compounds are validated, through an energy minimization with YASARA.
- ➔ Comparative Analysis is performed for both Native and Mutant proteins.

2. Literature Review

2.1 Metachromatic Leukodystrophy (MLD)

2.1.0 Inheritance

MLD is an autosomal recessive rare disorder, which means the parents of the patient, each carries a copy of the mutation. Yet there are no signs or symptoms of MLD in parents. ("Metachromatic leukodystrophy: MedlinePlus Genetics," 2020). Thus the probability of inheriting MLD, when both the parents carry the mutation, is 1 in 4 chance of expressing the abnormal gene, from both our parents. 1 in 2 chance, in expressing one abnormal gene, thus being a carrier of MLD. ("Autosomal recessive: MedlinePlus Medical Encyclopedia," 2016).

2.1.1 Types of MLD

There are mainly 3 types of MLD, late infantile (30 months), Juvenile MLD (2.5 - 16 Years), and Adult (Above 16 years), based on age-onset symptoms. In the end, all subtypes affect intellectual along with Physiological functions similarly (Narayanan et al., 2019).

Late infantile - approximately 50% of the cases occur at this stage. These deficits may be detected by neuropsychiatric examination. Juvenile MLD - it is further classified into early (4-6 years) and late (6-16 years) juvenile, which is characterized by declined motor skills and intellectual skills like seizures, memory loss, dementia, etc. Adult - It is often confused with bipolar disorder and dementia. The symptoms and signs are quite similar to the juvenile onset (Anish Lamichhane & Franklyn Rocha Cabrero, 2020).

2.1.2 Signs and Symptoms

MLD is a gradual illness. In other words, the effects get worse, over time. People with this condition eventually lose all Physiological and Psychological functions. Lifespan often relies on the time, it's first diagnosed in a patient (Burke, 2012).

By the age of five, the majority of children in the infantile form have died. The juvenile form's symptoms worsen over time, with death occurring 10 to 20 years after onset. Adults that are infected, typically die within 6 to 14 years after the symptoms ("Metachromatic Leukodystrophy Information Page | National Institute of Neurological Disorders and Stroke," 2021). The common symptoms are:

- abnormal muscle movement
- behaviour problems
- decreased mental function
- decreased muscle tone
- difficulty walking
- difficulty eating or feeding
- frequent falls
- incontinence
- irritability
- a loss of muscle control
- problems with nerve function
- seizures
- difficulty speaking
- difficulty swallowing

2.1.3 Causes

Metachromatic leukodystrophy (MLD), an autosomal recessive, rare neurodegenerative, hereditary disease, which falls under lysosomal storage disorder, is caused due by mutations in the Arylsulfatase A gene (ARSA), this leads to deficiency of the ARSA enzyme (Wu et al., 2020). No other molecule can compensate for the ARSA enzyme. Lack of ARSA enzyme, paves way for the accumulation of 3-O-sulfogalactosylceramide (sulfatide), in lysosomes of various cell types including oligodendrocytes, Schwann cells, microglia, and subpopulations of neurons (Doerr et al., 2015). As a result of this sulfatide deposition, the protective fatty layer, myelin sheath guarding nerves in both the central and peripheral nervous systems is damaged, along with the steady decline of white matter (leukodystrophy) in the entire nervous system.

White matter destruction causes disruptions in neurological and physiological functions. Symptoms such as loss of sensation, incontinence, seizures, paralysis, an inability to speak, blindness, and hearing loss. Eventually, they lose awareness of their surroundings and become unresponsive. It is also reported that sulfatide deposition occurs on other organs and tissues, mostly the gallbladder.

In certain rare cases, mutation can also occur in the Prosaposin PSAP gene, which is responsible for the production of a protein called saposin B. Saposin B is a sphingolipid activator, which is responsible for the activation of the ARSA gene. Thus this could also be a possible reason for MLD (“Metachromatic Leukodystrophy - NORD (National Organization for Rare Disorders),” 2021).

2.1.4 Frequency

The frequency rate for late infantile MLD was estimated to be about 1 per 40 000 (GUSTAVSON & HAGBERG, 1971). The prevalence rate is higher in certain isolated communities such as 1 in 75 in a small group of Jews who immigrated to Israel from southern Arabia, 1 in 2,500 in the western portion of the Navajo Nation, and 1 in 8,000 among Arab groups in Israel ("Metachromatic leukodystrophy: MedlinePlus Genetics," 2020).

2.1.5 Epidemiology

Even in developing countries, epidemiological data on leukodystrophies is limited (Amin, Elsayed, & Ahmed, 2017). The rates of morbidity and mortality differ from the influence of disorder. Young patients generally have the fastest-progressing illness, while patients with adult MLD have more of a chronic and gradual progression (Blenda, 2021). In the overall population, 1 in 40,000 equates to 1 in 100 carriers frequency. A record of 3,600 MLD births per year, with 1,900 alive in the US, 3,100 in Europe, and 49,000 alive worldwide. (Wikipedia Contributors, 2021)

2.1.6 Diagnosis

If the symptoms of a patient, show the signs of MLD, a doctor can order the following tests to confirm your diagnosis (Burke, 2012):

- Blood tests would be done to determine whether you have an enzyme deficiency.
- A urine test can be used to determine whether you have a sulfatide accumulation.
- A genetic test to see whether you have the gene that causes MLD.
- A nerve conduction analysis can determine how electrical impulses travel through the nerves and muscles and also assess MLD-related nerve damage.
- A magnetic resonance imaging (MRI) scan of your brain can be performed. Sulfatides build up in the brain, as well as affect the white matter of the brain, an MRI will reveal this.

2.1.7 Treatment

MLD has no known remedy. In some infantile-onset cases, bone marrow transplantation can help to slow the progression of the disease. Symptomatic and supportive care is the only other choice. In an experimental model of MLD and clinical trials, significant progress has been made with gene therapy. ("Metachromatic Leukodystrophy Information Page | National Institute of Neurological Disorders and Stroke," 2021) . Types of treatment are listed below(Blenda, 2021):

- Hematopoietic stem cell transplantation (HSCT) - Bone marrow or umbilical cord blood transplantation
- Gene Therapy - To correct the underlying genetic abnormality.
- Enzyme Replacement Therapy - Recombinant human ARSA (rhARSA) enzyme, Metazym (Shire HGT)

2.2 Conclusion obtained from literature review

Metachromatic leukodystrophy (MLD), an autosomal recessive, rare neurodegenerative, hereditary disease, which falls under lysosomal storage disorder, is caused due by mutations in the Arylsulfatase A gene (ARSA). The ARSA enzyme can't be compensated by any other molecule. The lack of the ARSA enzyme allows for the aggregation of sulfatide. Deterioration of motor and cognitive skills are the most prominent symptoms. Based on age at onset and symptoms, three separate clinically-based MLDs are diagnosed: late infantile (3 years), juvenile (3-16 years), and adult (after 16 years). There is no cure for MLD. The line of therapy is to strengthen and regulate your symptoms.

2.3 Research Gap

MLD, falling under a rare hereditary disorder, itself, makes it hard for wet lab, as well as dry lab, research, and analysis. The need for molecular diagnosis and genetic analysis could be aided, with extensive research on patients with MLD. Clinicians should be aware of the importance of diagnosing leukodystrophies which are time-sensitive therapies and of the initial diagnosis, suggested referrals, and services available to patients and families who have been diagnosed with suspected leukodystrophy (Ruzhnikov, Brimble, & Hickey, 2021).

2.4 Formulation of the problem

MLD, though a rare disorder, has a high fatality rate, almost 50 percent in the late infantile state (Anish Lamichhane & Franklyn Rocha Cabrero, 2020). Adding on, the prognosis of MLD is also poor, with most deaths occurring before the age of 5 ("Metachromatic Leukodystrophy Information Page | National Institute of Neurological Disorders and Stroke," 2021). A study emphasizes the importance of molecular diagnosis in all MLD patients so that MLD families can acquire carrier evaluation, genetic counselling, and prenatal testing (Narayanan et al., 2019).

2.5 Rationale of the study

This study, reports on comparative analysis of the deleterious mutation that takes place in ARSA genes, of MLD patients by doing so, this molecular investigation and statistical evaluation, could help aid the research community. In this study, we report, the molecular analysis was done with the help of multiple computational tools, to study deleterious mutations present in the ARSA gene, to compare and contrast, their statistical and genetic analysis, which would provide data for further molecular research and diagnosis.

3. Materials and Methods

3.1 Data Collection

3.1.1 Identifying the gene target

Metachromatic leukodystrophy (MLD), is caused due to the mutations in the Arylsulfatase A gene (ARSA), which leads to deficiency of the ARSA enzyme. Lack of ARSA enzyme paves way for the accumulation of 3-O-sulfogalactosylceramide (sulfatide) (Wu et al., 2020).

The structure of an Arylsulfatase A mutant P426L (PDB ID: 1E33) was retrieved from Protein Data Bank. Mutational and other data were referred from Uniprot database (Uniprot ID: P15289). For further structural and molecular analysis, I chose the structure 1AUK, as it had 3D conformations and a high resolution to work with.

3.1.2 Retrieval of SAP mutations from various data sources

Through literature review, collected data of MLD patients, to have known and novel mutations, particularly missense mutations, from various regions of the world. Below given points will have the collected SAP mutations and their source of data.

1) A Late-onset spectrum of ARSA variations in Asian Indian patients with Arylsulfatase A deficient metachromatic leukodystrophy (Narayanan et al., 2019).
   Novel Mutations

   p180q q139k r299w g34e g392r g127r f399s

   Known Mutations

   g156r r390w r311q t274m r299w t304m d430y p426l f247c g293s t272m r244h g309c r370w g245r y33s c489g

2) Three novel variants in the arylsulfatase A (ARSA) gene in patients with metachromatic leukodystrophy (MLD) (Hettiarachchi & Dissanayake, 2019).
   Novel Mutations

   n350s y201h d281n r372q p84k d283n r86w t393s r313q a344d

   Know Mutations

   R370Q

3) Clinical, Biochemical, and Molecular Characterization of Metachromatic Leukodystrophy Among Egyptian Pediatric Patients: Expansion of the ARSA Mutational Spectrum (Amr et al., 2020).
   Novel Mutations

   R60p v112d g124d p193s p276k g311r p150l t276k g295s w195c w320c n284s

4) Comprehensive clinical, biochemical, radiological, and genetic analysis of 28 Turkish cases with suspected metachromatic leukodystrophy and their relatives (Pekgül et al., 2020).
   Novel Mutations

   t276m a214p g311d w320c g298d t393g i181s r301l r349w g247r

5) Chinese Cases of Metachromatic Leukodystrophy with the Novel Missense Mutations in ARSA Gene (Wu et al., 2020).
   Novel Mutations

   r372q g298v g101v p194r l433v g449r p84q

6) Late-onset MLD with normal nerve conduction is associated with two novel missense mutations in the ASA gene (Gallo, 2004).
   Novel Mutations - G293D C489G

3.2 Data Prediction and Screening

With the dataset collected from various credible data sources, we are now going to screen for the most deleterious mutations with the help of SNPs&GO. It is a machine learning algorithm, which follows the Support Vector Machine Model (SVM), to predict the most deleterious SNPs.

The outcome is a table with the number of the mutated location in the protein sequence, the wild-type residue, the new residue, and whether the associated mutation is expected as disease-related or neutral polymorphism, as well as the RI value for that mutation (Reliability Index). Below I have attached a table with the most deleterious, which was screened based on the high-Reliability Index, and which is hand in hand with the 1AUK structure.

Mutation	Prediction	PhD-SNP	PANTHER	RI	Probability
T274M	Disease	F[T]=93% F[M]=0% Nali=995	F[T]=98% F[M]=0%	9	0.947
R311Q	Disease	F[R]=92% F[Q]=0% Nali=992	F[R]=94% F[Q]=0%	9	0.929
R370W	Disease	F[R]=55% F[W]=0% Nali=949	F[R]=62% F[W]=0%	9	0.945
R390W	Disease	F[R]=81% F[W]=0% Nali=884	F[R]=84% F[W]=0%	9	0.958
C489G	Disease	F[C]=66% F[G]=2% Nali=107	F[C]=69% F[G]=2%	9	0.946

TABLE 1 - SCREENED DELETERIOUS MUTATION

Mutation: WT+POS+NEW, WT: Residue in wild-type protein, POS: Residue position, NEW: New residue after mutation, RI: Reliability Index, Predictor of human Deleterious Single Nucleotide Polymorphisms, Probability: Disease probability (if >0.5 mutations are predicted Disease, F[X]: Frequency of residue X in the sequence profile
Nali: Number of aligned sequences in the mutated site

3.3 Dataset of *in-silico* mutant modelling

Specifically, 5 mutations have been detected, the next step is followed by mutating the native model with the help of Swiss PDB Viewer. In this process, we change the wild type of the amino acid into its mutant type, this is then saved in PDB format, for further analysis.

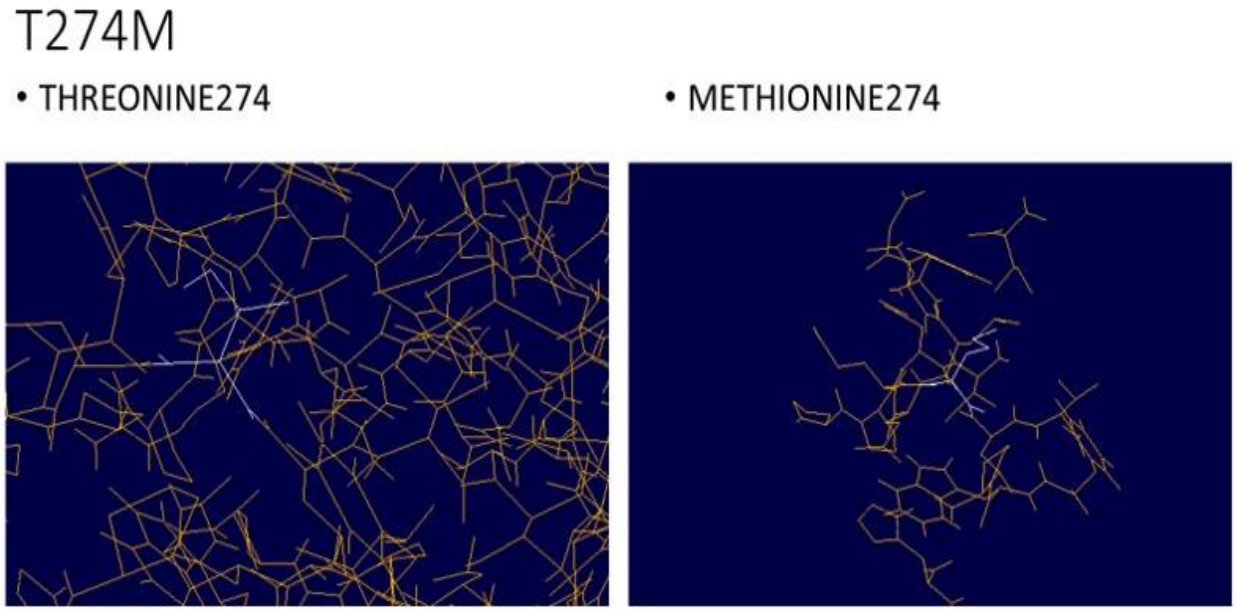


FIGURE 3 - T274M

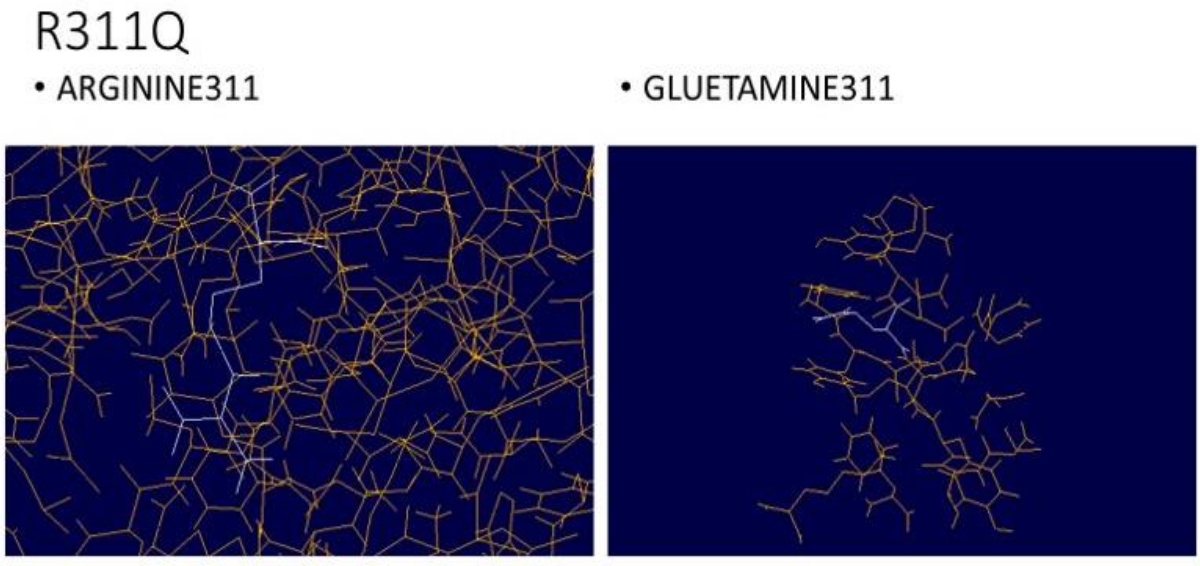


FIGURE 4 - R311Q

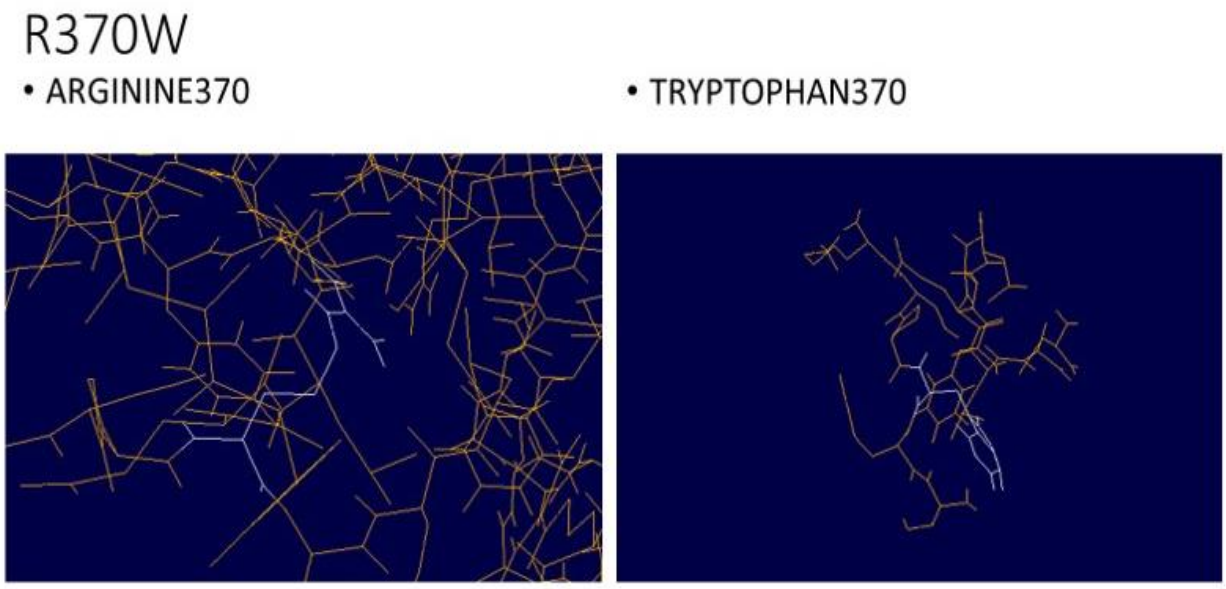


FIGURE 5 - R370W

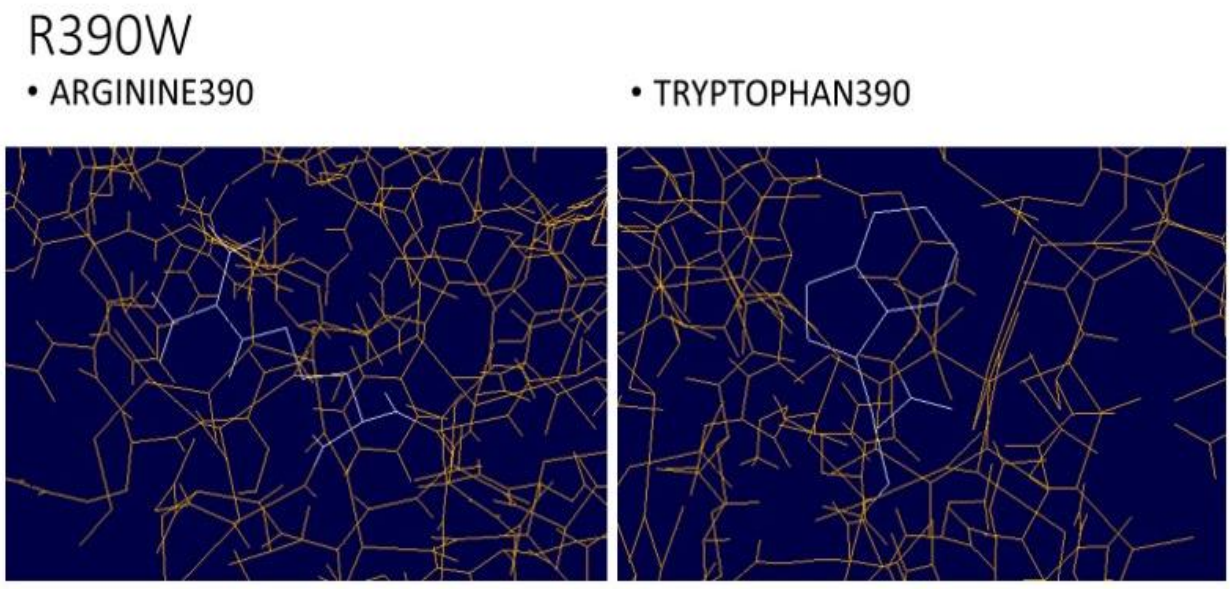


FIGURE 6 - R390W

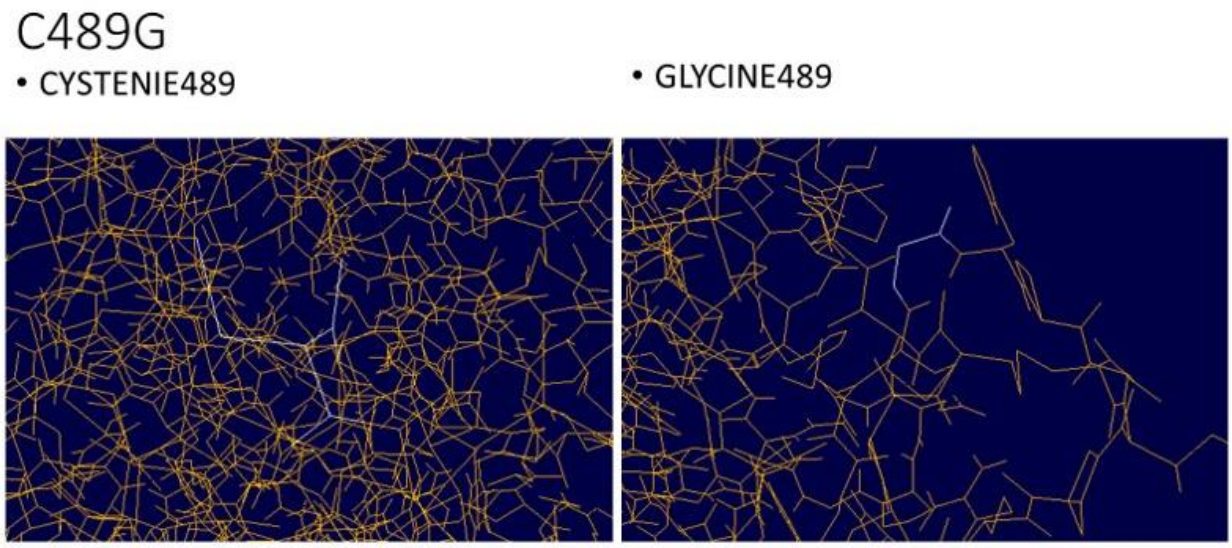


FIGURE 7 - C489G

3.4 Energy Minimization

These mutated models, which are saved as PDB files are now energy minimized with the help of YASARA program. This is a tool for molecular graphics, modelling, and simulation and an easy-to-use application. It uses PVL (Portable Vector Language) as the engine behind it, a modern development platform that outperforms conventional apps. Given below is a table of the energy minimization score.

Compounds	Energy Minimization Score
1AUK	-249060.3
T274M	-247354.6
R311Q	-246171.2
R370W	-248250.5
R390W	-247801.7
C489G	-247130.3

TABLE 2 - ENERGY MINIMIZATION SCORE FOR EACH MUTANT FROM YASARA PROGRAM.

3.5 Single Model Analysis

The single model analysis is performed through various software and analytical tools as listed below. This exploration is done for both native and mutant models.

3.5.1 Intramolecular Interactions

Understanding the molecular basis of protein stability and function requires taking into account interactions within a protein structure along with interactions between proteins in an array (Tina, Bhadra, & Srinivasan, 2007). Thus, Intramolecular interactions are exhibited, through a web server called Protein Interaction calculator (PIC).

Various kinds of interactions are noted such as disulphide bonds, hydrophobic interactions, ionic interactions, hydrogen bonds, aromatic- aromatic interactions, aromatic-sulphur interactions, and cation-cation interactions. Below (Table 3) is attached for intramolecular interactions for both native and mutant compounds.

Native and Mutants	**Intra Protein Hydrophobic Interactions: Hydrophobic Interactions within 5 Angstroms**	**Interprotein Disulphide Bridges Disulphide bridges: Between sulphur atoms of cysteines within 2.2 Angstroms**	**Interprotein Main Chain-Main Chain Hydrogen Bonds**	**Interprotein Main Chain-Side Chain Hydrogen Bonds**	**Intra Protein Side Chain-Side Chain Hydrogen Bonds**
Native - 1AUK	436	6	527	247	133
T274M	442	6	533	235	148
R311Q	430	6	519	254	141
R370W	448	6	519	247	169
R390W	444	6	520	238	139
C489G	434	5	518	244	146

TABLE 3 - INTRAMOLECULAR INTERACTIONS - PART 1

Native and Mutants	Interprotein Ionic Interactions: Ionic Interactions within 6 Angstroms	Interprotein Aromatic-Aromatic Interactions: Aromatic-Aromatic Interactions within 4.5 and 7 Angstroms	Interprotein Aromatic-Sulphur Interactions Aromatic-Sulphur Interactions within 5.3 Angstroms	Interprotein Cation-Pi Interactions
Native - 1AUK	48	11	8	11
T274M	45	13	7	13
R311Q	41	12	8	11
R370W	46	14	8	13
R390W	43	13	9	12
C489G	48	10	9	12

TABLE 4 - INTRAMOLECULAR INTERACTIONS - PART 2

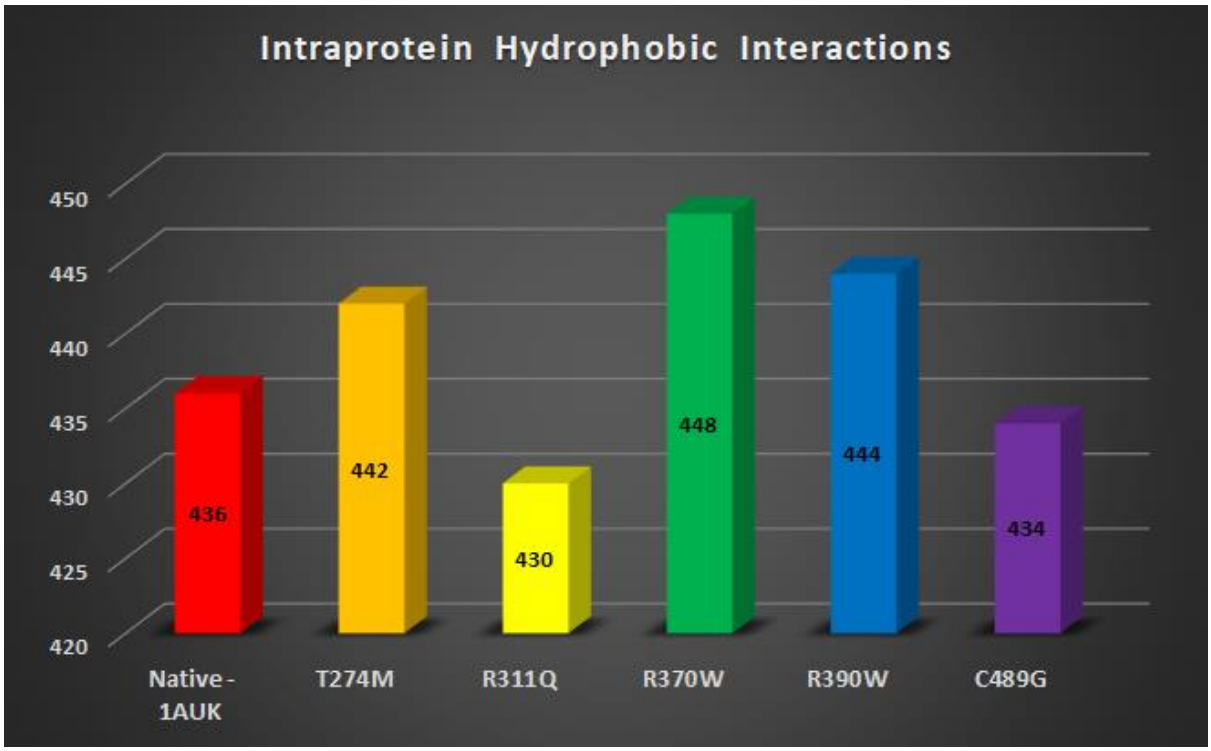


FIGURE 8 - HYDROPHOBIC INTERACTIONS

3.5.2 Secondary Structure Analysis

A protein's primary structure is simply the amino acid sequence in the polypeptide chain. Protein secondary structures are repeating standard conformations on the polypeptide chain. A protein may be folded into a stable three-dimensional structure from its secondary structures, which is classified as the tertiary structure of a protein (Zhang, Li, & Lü, 2018).

Here, we use a tool called POLYVIEW-MM, combining an interactive view and analysis of conformational changes with the development of customized 2D description plots that show the proteins' molecular motion. Below displayed (Figure 9) is an example of the output from polyview mm.

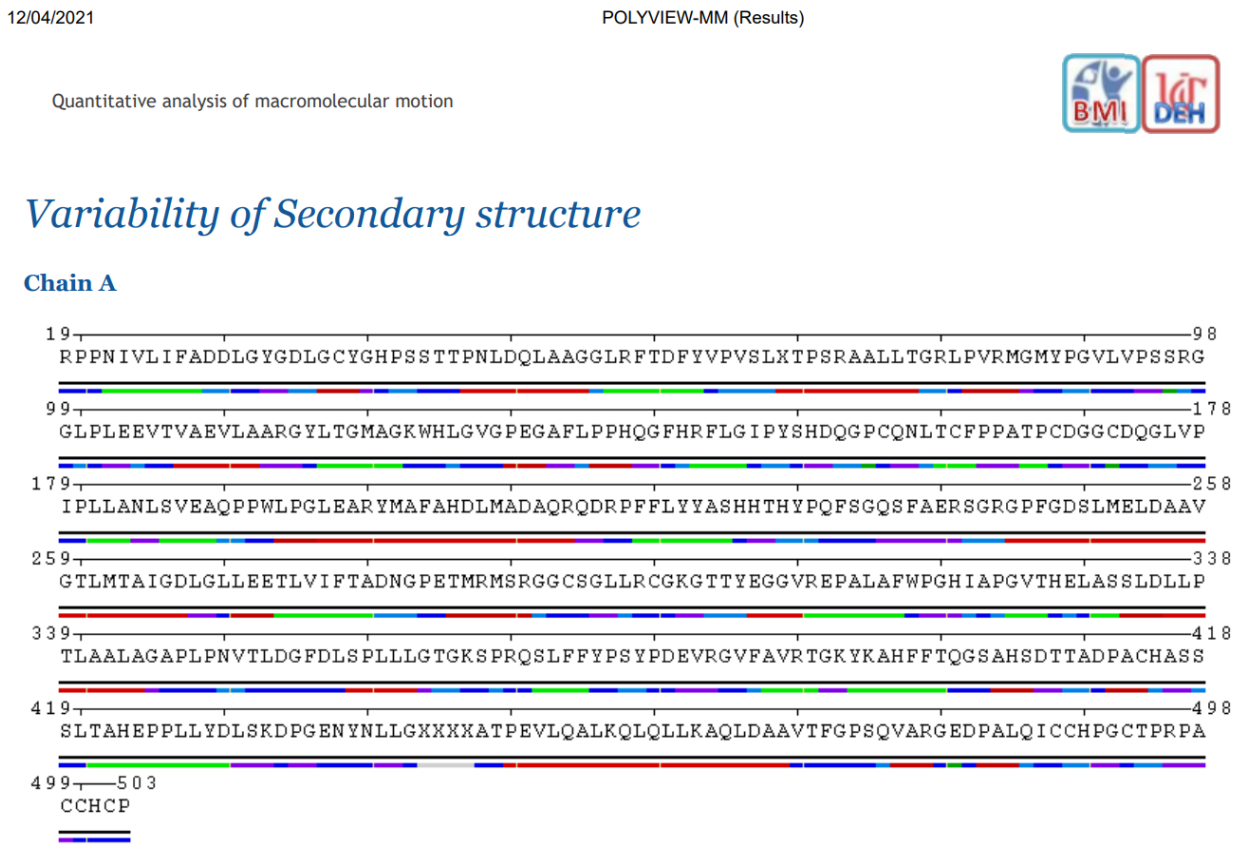


FIGURE 9 - SECONDARY STRUCTURE ANALYSIS - 1AUK

Furthermore, you will find the comparative analysis, between the native and the mutants, where visual changes between them are boxed for better understanding are the figures below:

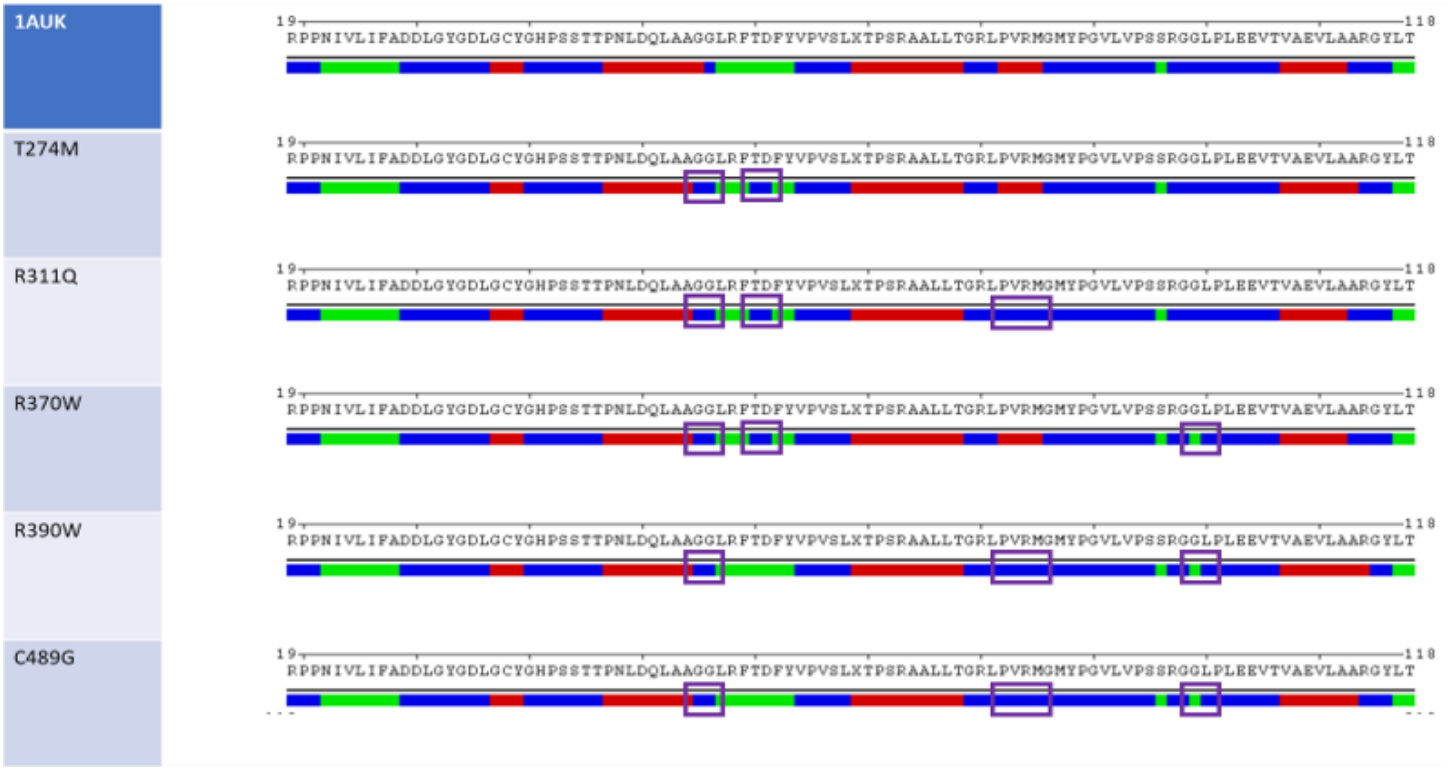


FIGURE 10 - 1ST LINE COMPARISON - SECONDARY STRUCTURE ANALYSIS

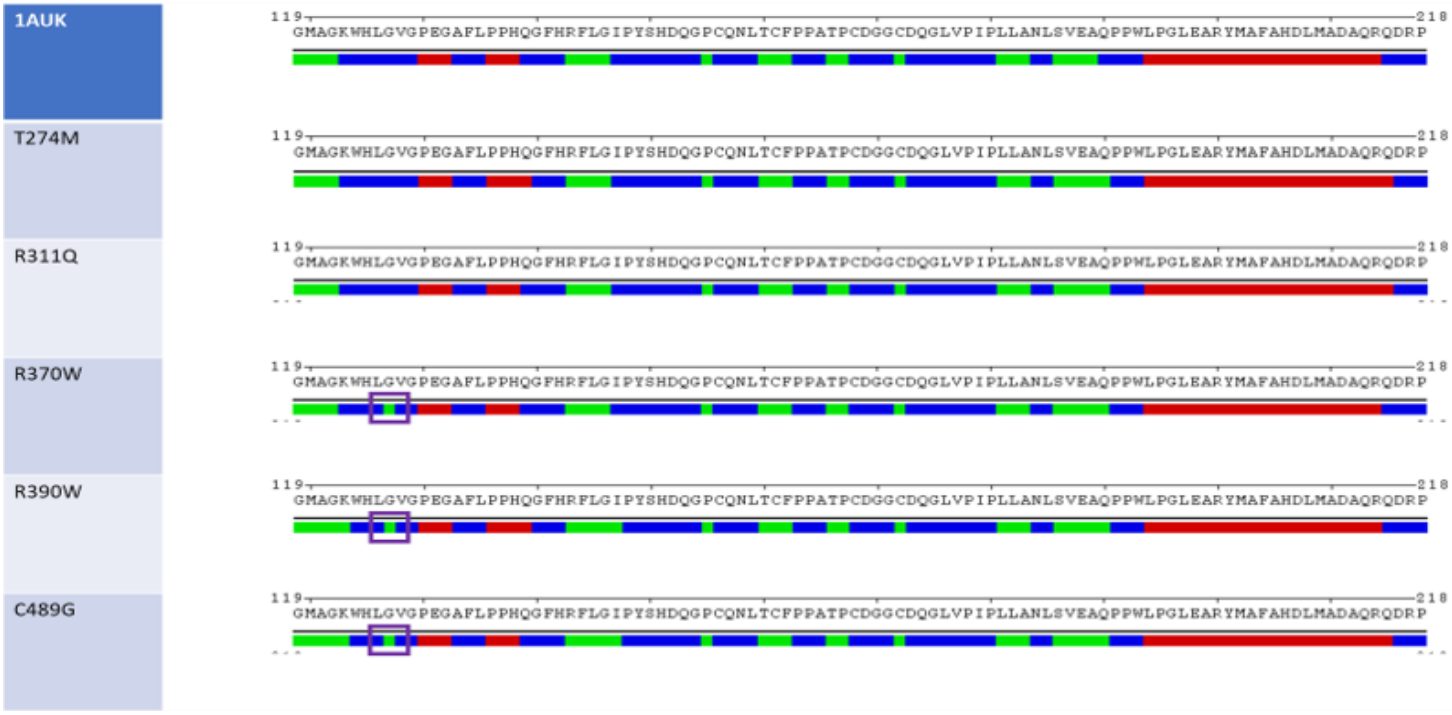


FIGURE 11 - 2ND LINE COMPARISON - SECONDARY STRUCTURE ANALYSIS

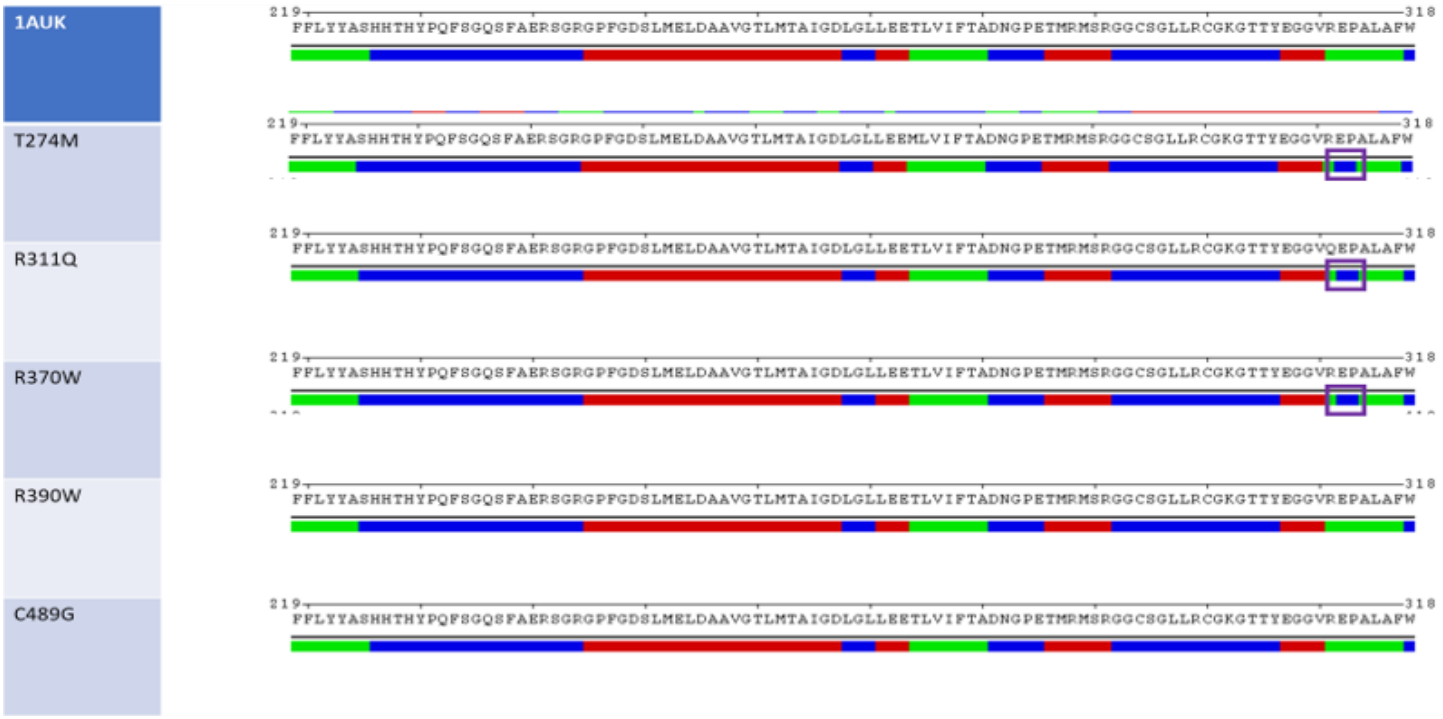


FIGURE 12 - 3RD LINE COMPARISON - SECONDARY STRUCTURE ANALYSIS

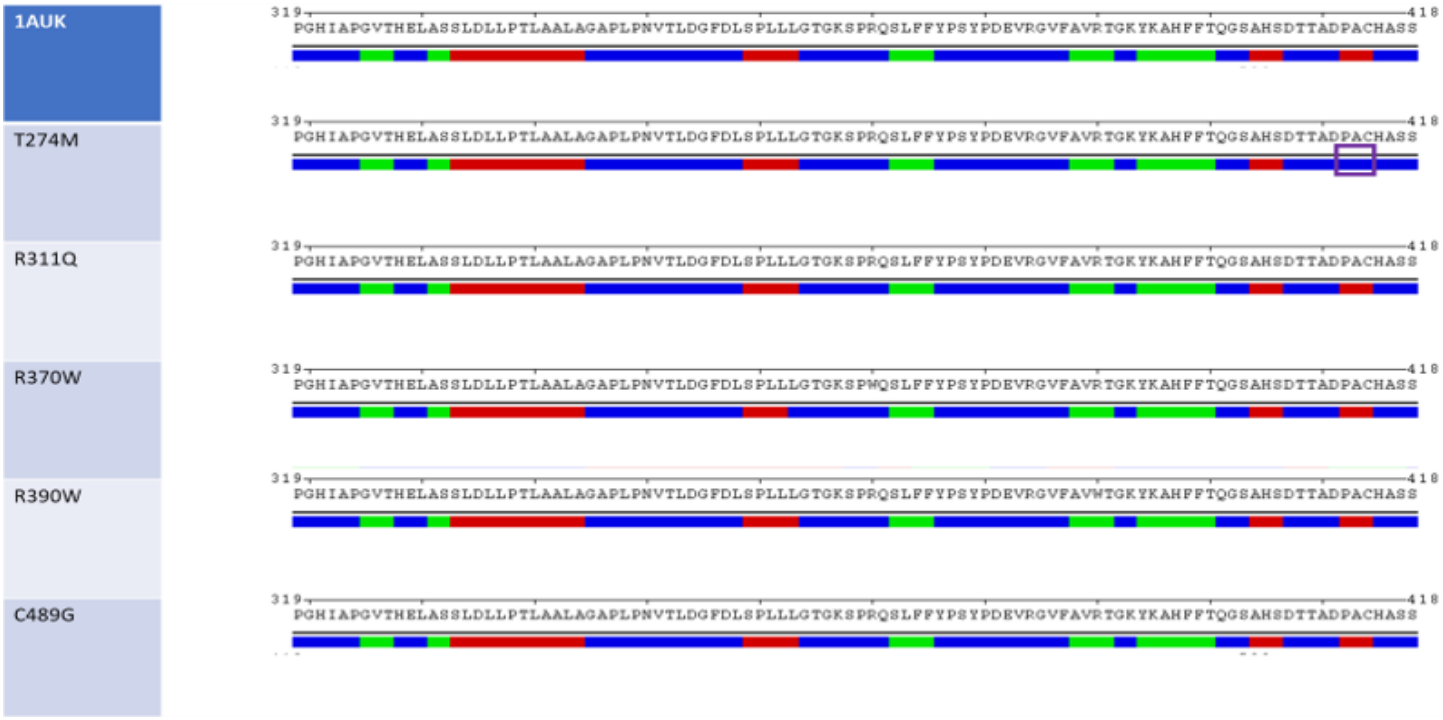


FIGURE 13 - 4TH LINE COMPARISON - SECONDARY STRUCTURE ANALYSIS

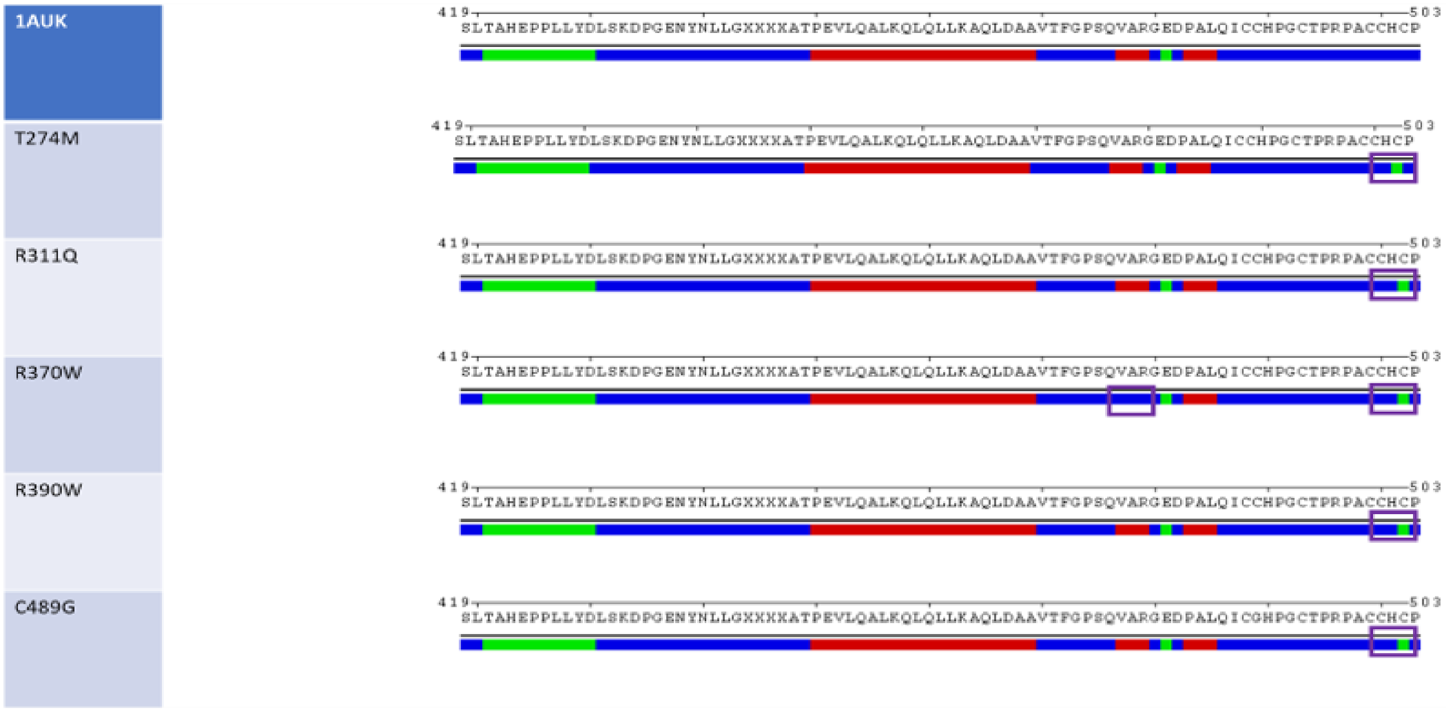


FIGURE 14 - 5TH LINE COMPARISON - SECONDARY STRUCTURE ANALYSIS

States	1AUK	T274M	R311Q	R370W	R390W	C489G
Unstructured Loop, Coil	52%	53%	53%	53%	51%	51%
Beta Bridges	18%	18%	18%	18%	19%	20%
Alpha Helix Turn	30%	29%	29%	29%	30%	29%

TABLE 5 - OVERALL PERCENTAGE OF SECONDARY STRUCTURE ANALYSIS

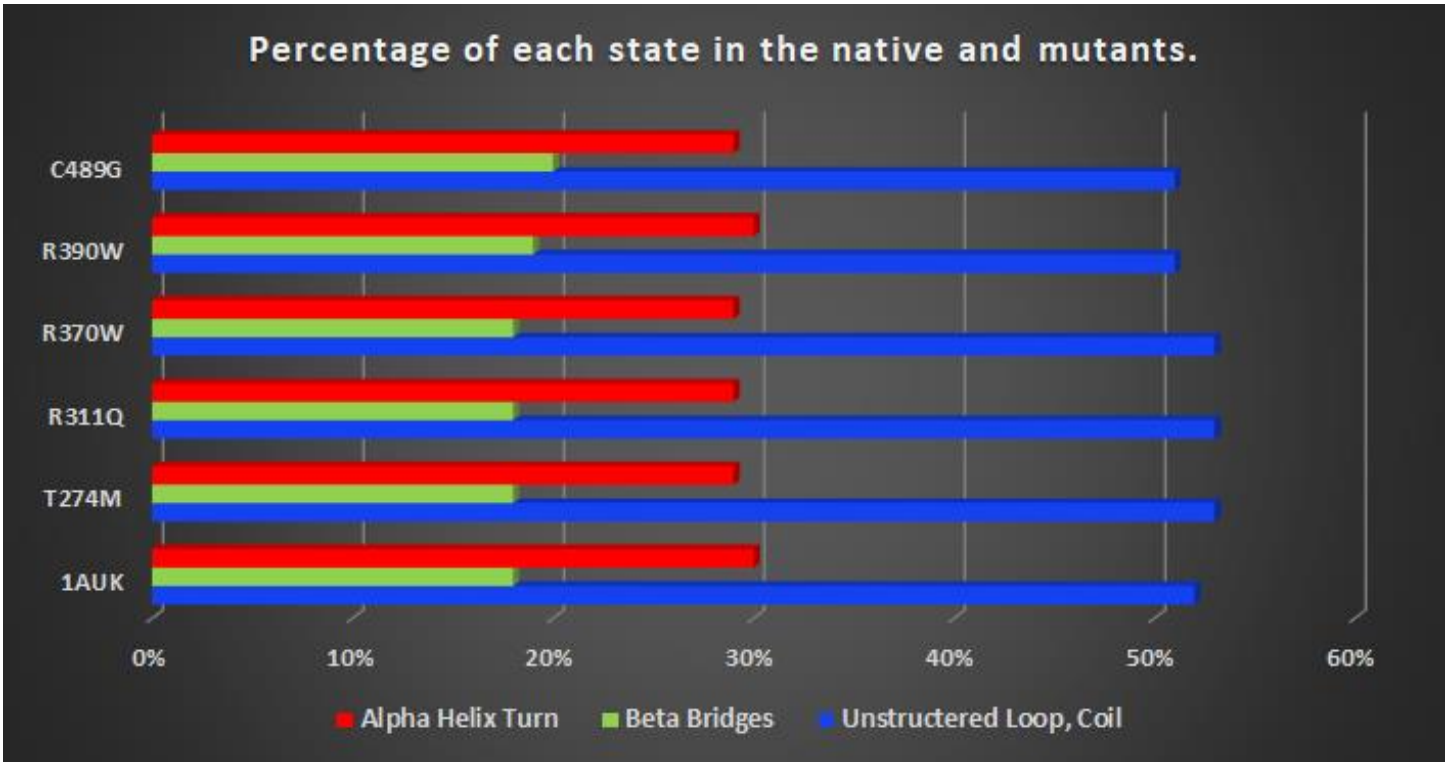


FIGURE 15 - BAR GRAPH - REPRESENTING PERCENTAGE OF EACH STATE

3.5.3 Fluctuation Analysis

Since its introduction by Luria and Delbrück in 1943, the estimation of mutation parameters in fluctuation analysis has been the subject of numerous studies (Luria & M. Delbrück, 1943). To our knowledge, fluctuation analysis is the only method to determine causality in systems where the perturbation is not possible (Welf & Danuser, 2014).

Fluctuation analysis allows researchers to investigate protein activity in a linear range. Natural variations in component behaviour offer an opportunity to research small changes in component activity, which we refer to as "linear biology."(Welf & Danuser, 2014) Similarly, this analysis is dedicated to the native and mutant of the ARSA gene.

CABS-flex is a quick modelling technique for protein structure flexibility simulations. It's focused on the CABS model, which is a well-known coarse-grained protein modelling method (Kmiecik et al., 2016).

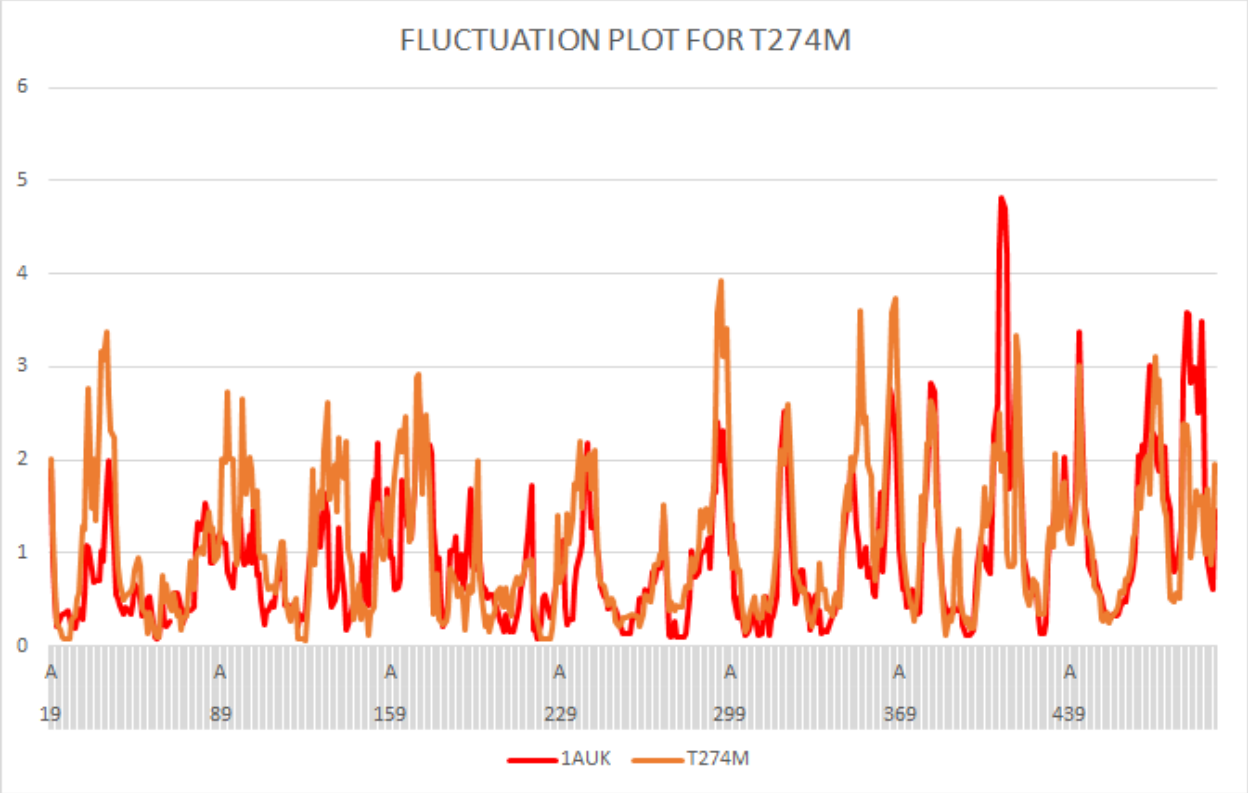


FIGURE 16 - T274M FLUCTUATION PLOT

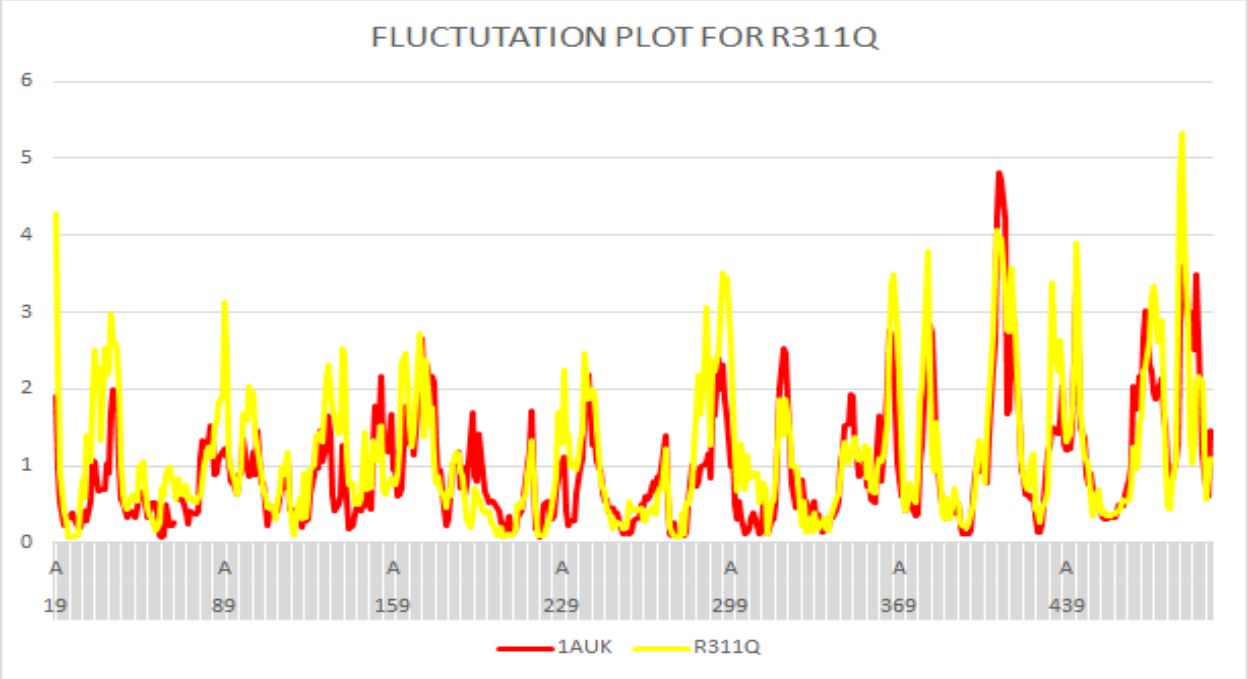


FIGURE 17 - R311Q FLUCTUATION PLOT

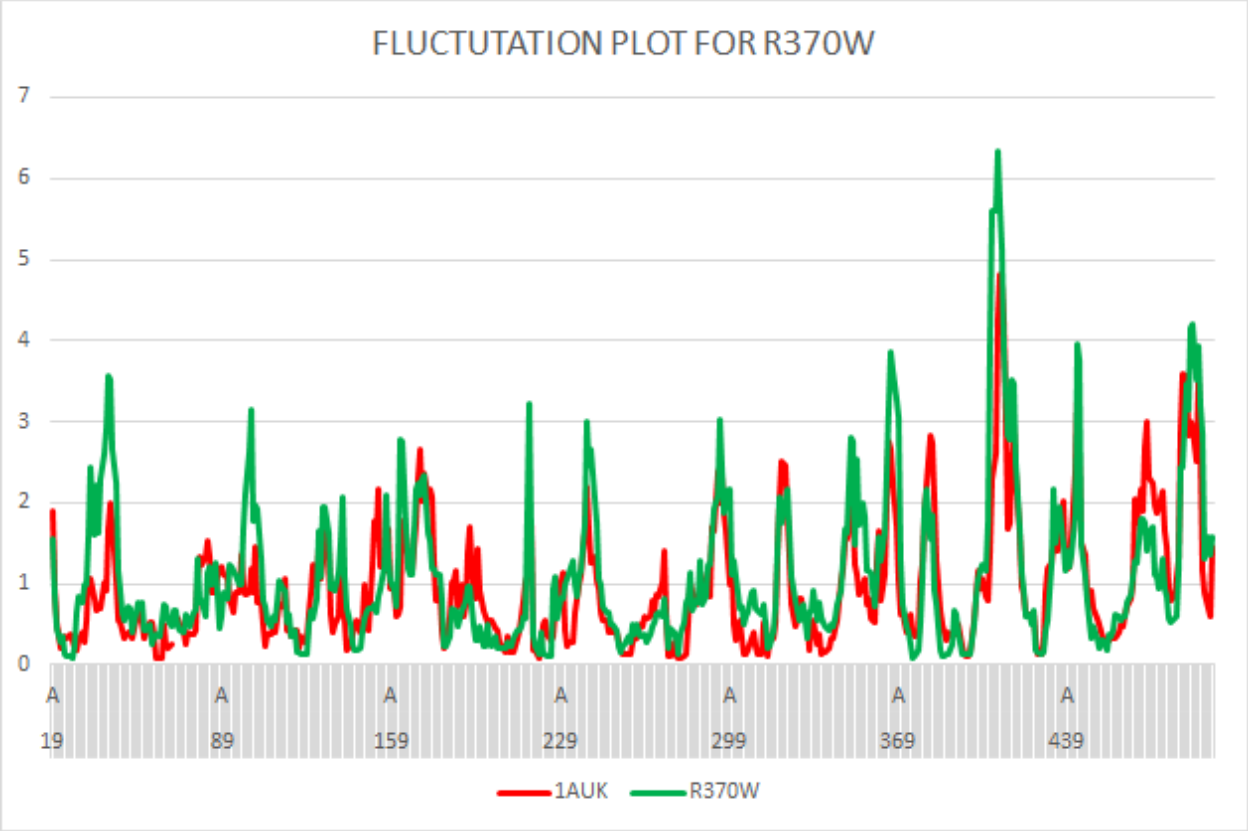


FIGURE 18 - R370W FLUCTUATION PLOT

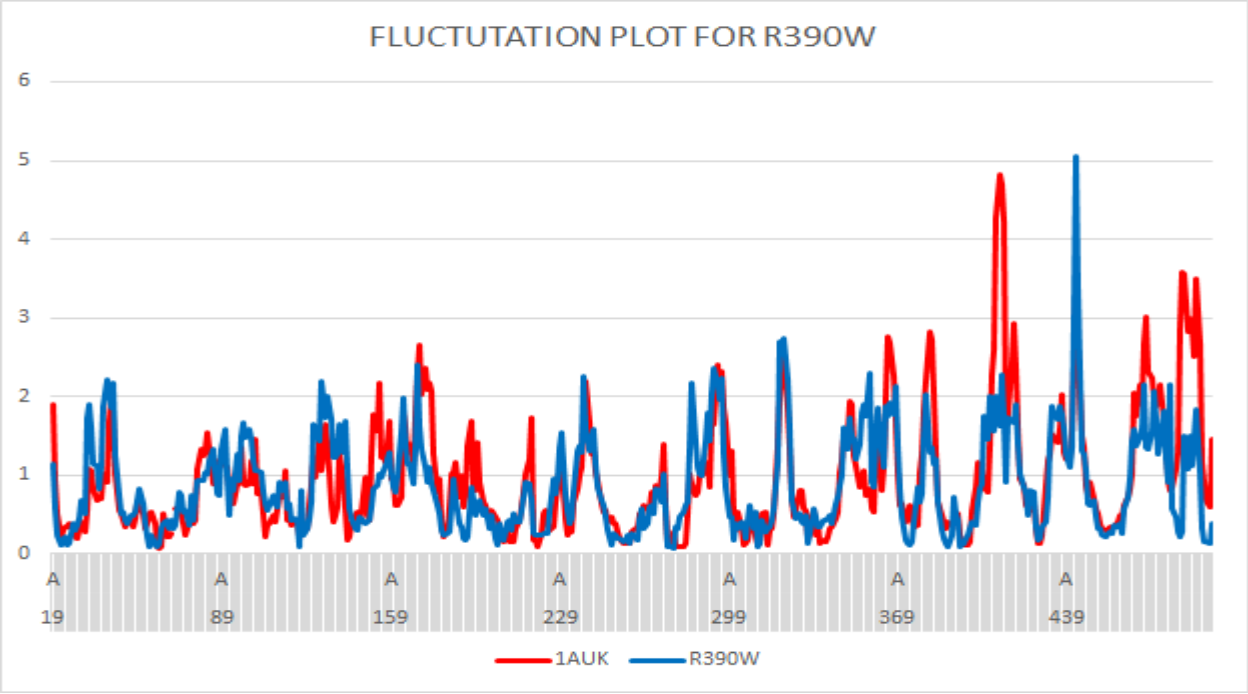


FIGURE 19 - R390W FLUCTUATION PLOT

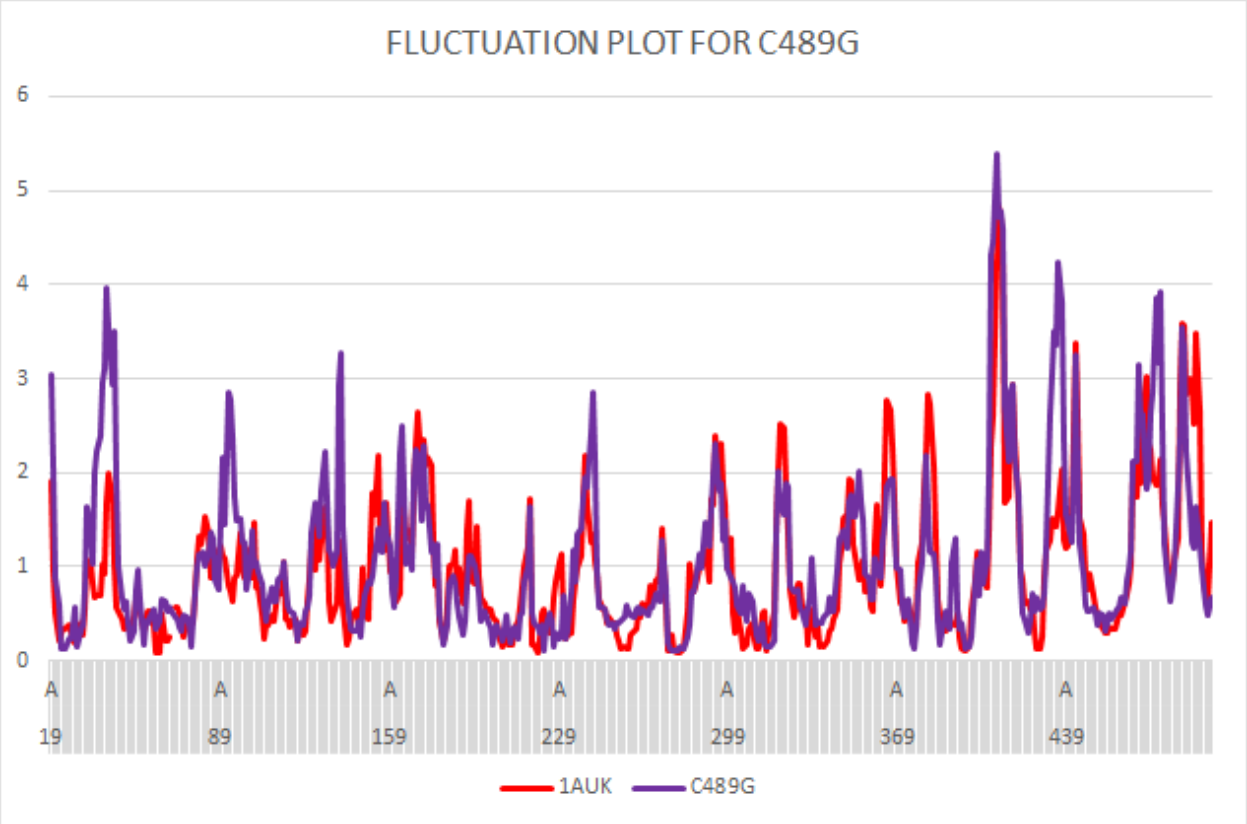


FIGURE 20 - C489G FLUCTUATION PLOT

3.5.4 Contact Map Analysis

In a three-dimensional protein structure, a protein contact map contains a binary two-dimensional matrix that describes the distance between all possible amino acid residue pairs. Helices, parallel and anti-parallel beta-sheets, loops can also be determined from the contact map graph (Wikipedia Contributors, 2020).

Here, we use a software called Contact Map View (CMView), to identify similarities and differences, in the structural conformation of both the native and the mutant. It shows the list of unique and similar contacts, with also the percentage of the contact map overlapping. This is also used in studying, patterns of interaction between secondary structural components can also be observed through contact map analysis. It follows Needleman Wunsch Analysis.

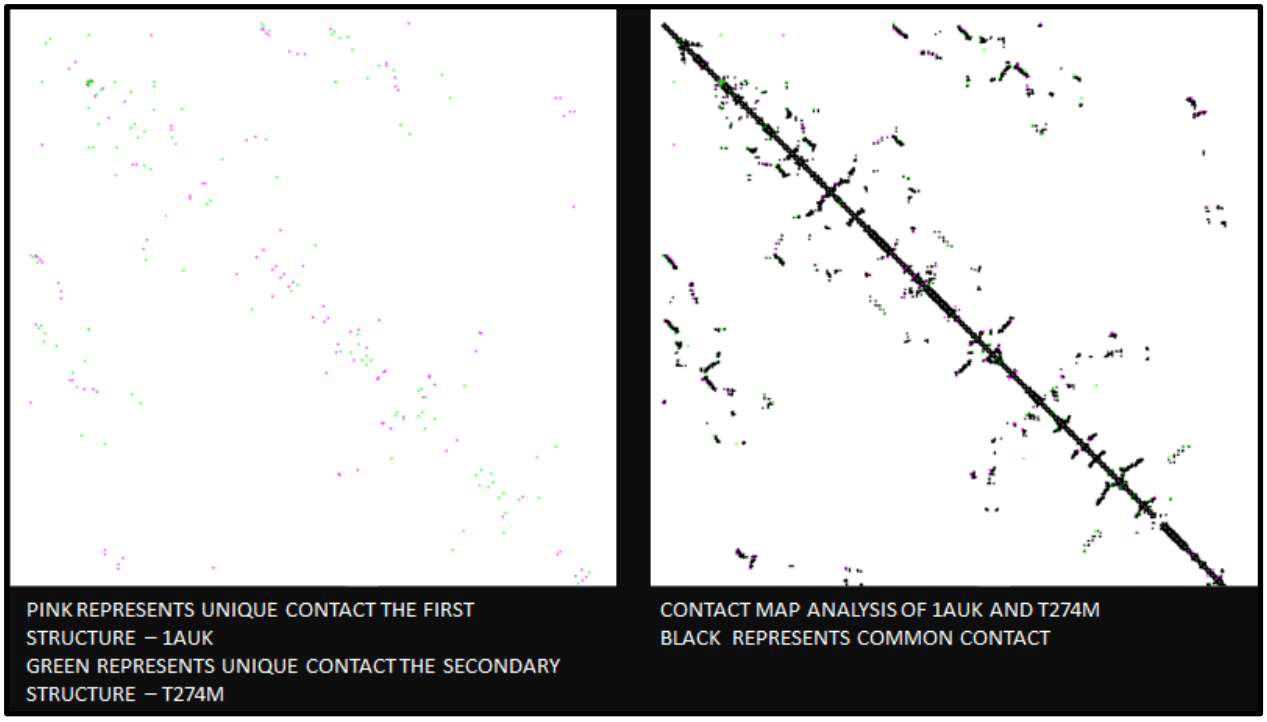


FIGURE 21 - T274M CMA

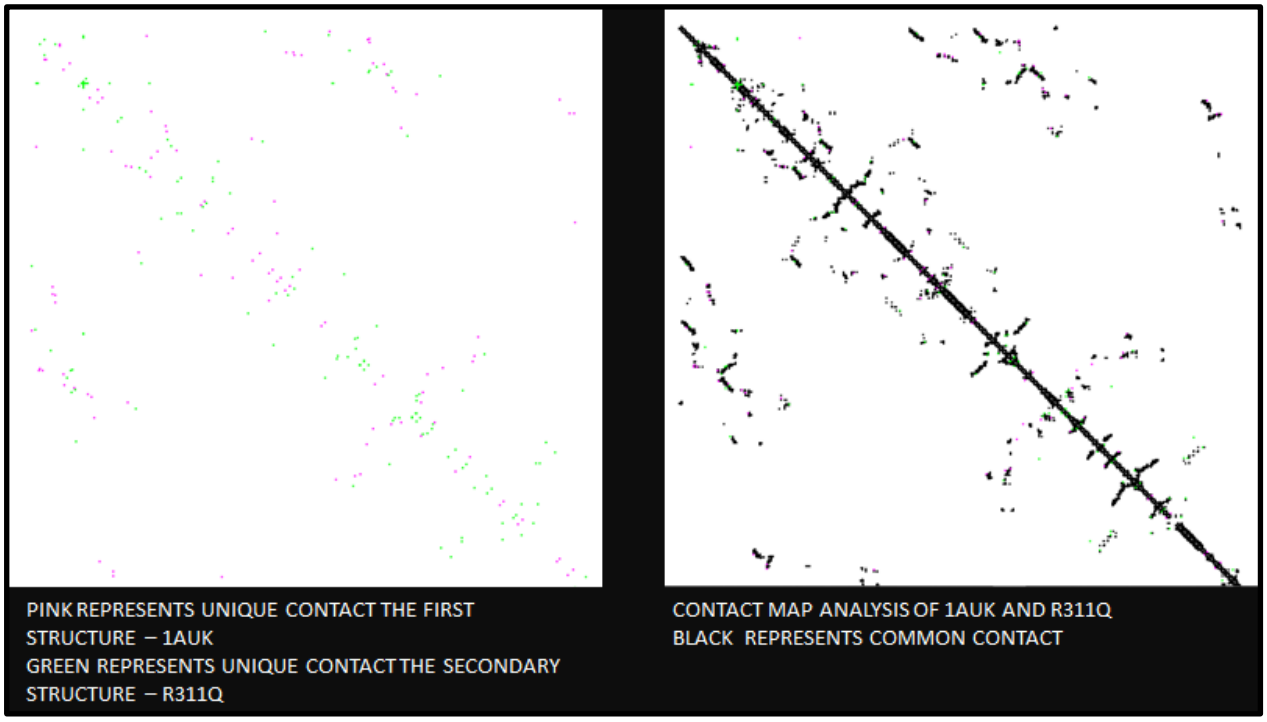


FIGURE 22 - R311Q CMA

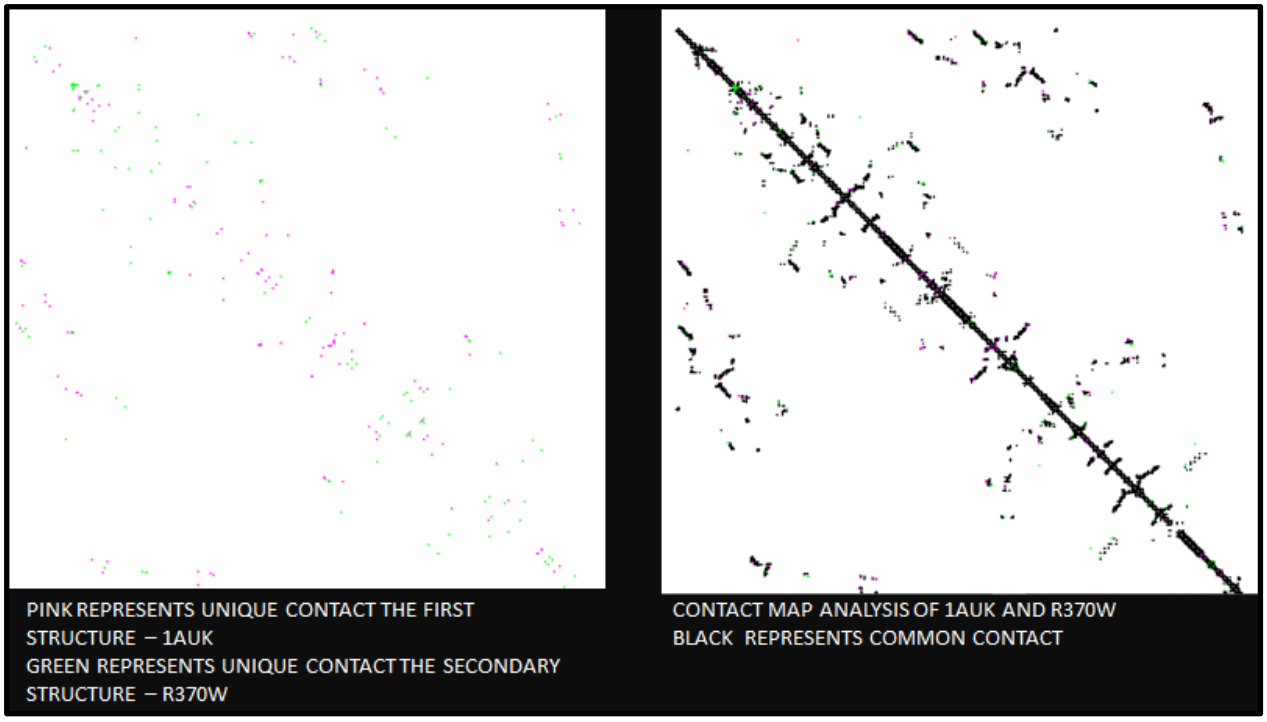


FIGURE 23 - R370W CMA

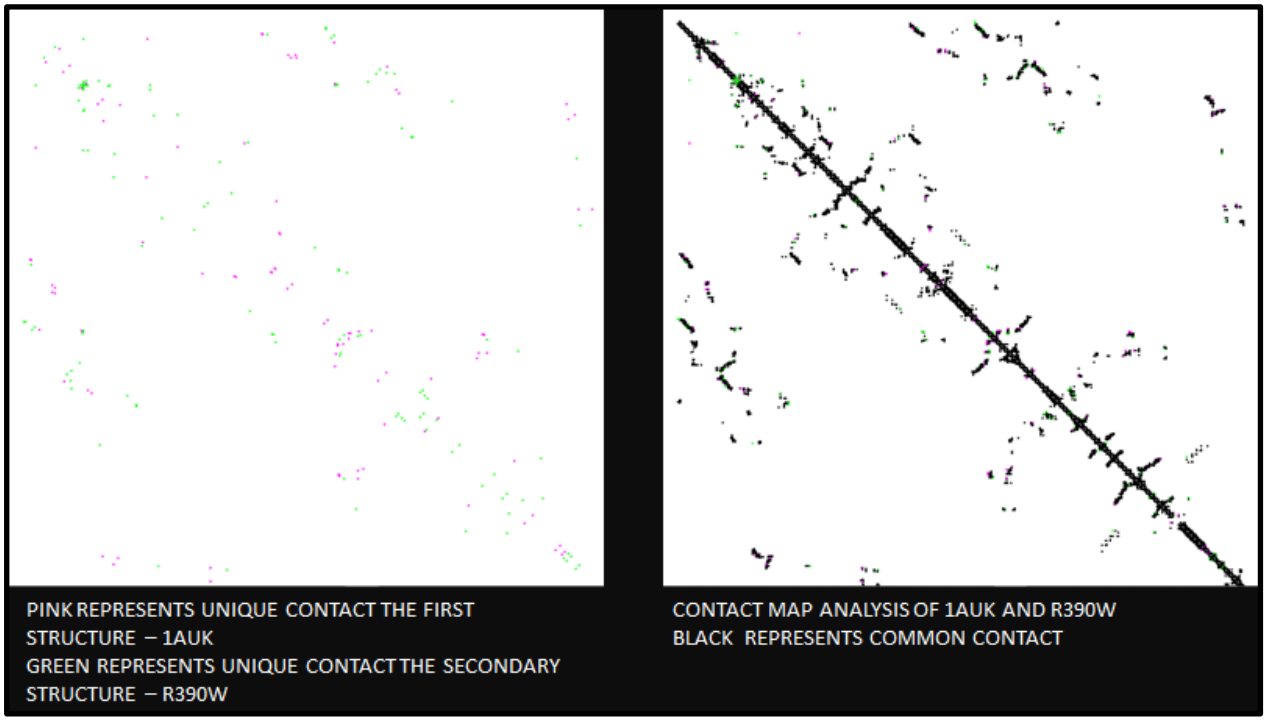


FIGURE 24 - R390W CMA

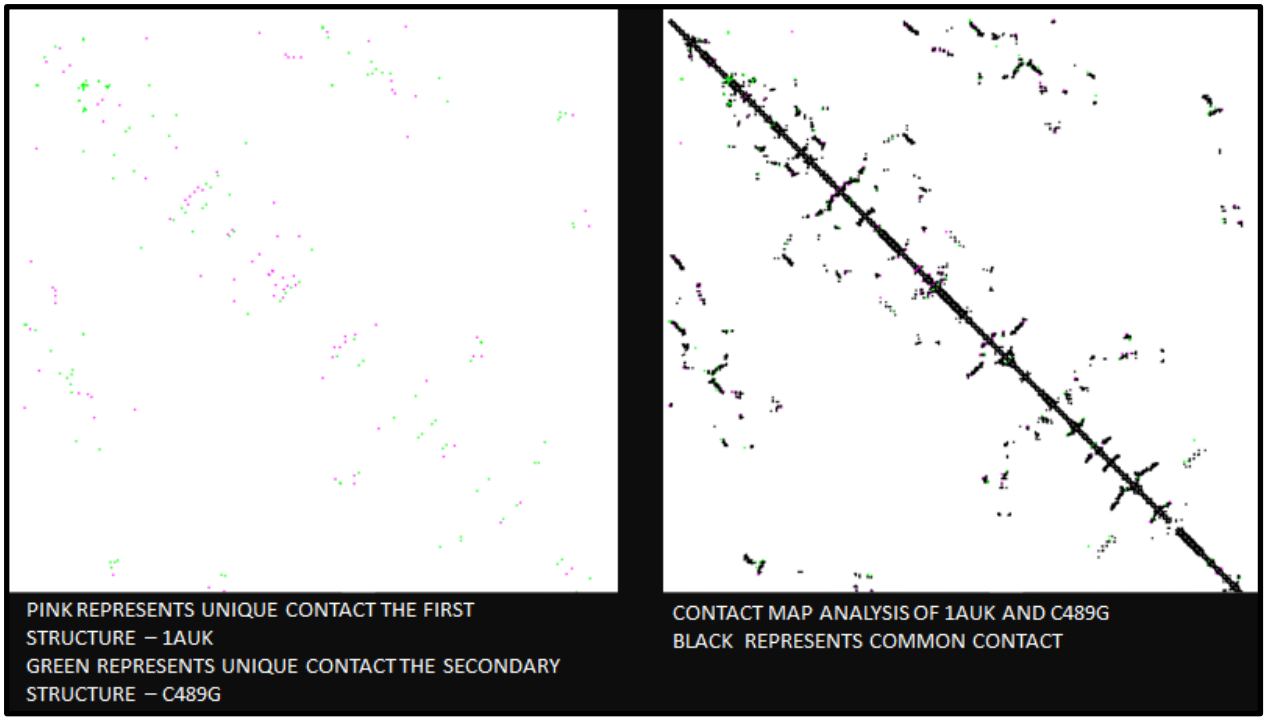


FIGURE 25 - C489G CMA

Native and Mutants	No. of Contacts	No. of Unique Contacts -NATIVE	No. of Unique Contacts - MUTANTS	No. of Common Contacts	Contact Map Overlap
Native - 1AUK	2611	-	-	-	-
T274M	2608	55	52	2556	96.00%
R311Q	2613	54	56	2557	95.90%
R370W	2602	69	60	2542	95.20%
R390W	2615	50	54	2561	96.10%
C489G	2616	48	53	2563	96.20%

TABLE 6 - OVERALL CONTACT MAP ANALYSIS

3.5.5 Solvent Accessible Surface Areas (SASA)

SASA which is also known as accessible surface area (ASA), is the facet of the entire compound, which would come in contact with the solvent. This would completely depend on whole structure of the compound. A great way to measure SASA, is through units (Durham, Dorr, Woetzel, Staritzbichler, & Meiler, 2009).

Native and Mutants	POLAR area/energy	APOLAR area/energy	UNKNOW area/energy	Total area/energy
Native - 1AUK	6131.66	11352.01	0	17483.67
T274M	6141.85	11342.84	0	17484.69
R311Q	6154.04	11359.83	0	17513.87
R370W	6124.73	11332.57	0	17457.3
R390W	6125.6	11345.01	0	17470.61
C489G	6131.61	11336.21	0	17467.82

TABLE 7 - SASA CALCULATIONS - PART 1

Native and Mutants	Number of surface atoms	Number of buried atoms	Number of atoms with ASP=0	Probe Radius
Native - 1AUK	1822	1768	914	1.4
T274M	1835	1756	913	1.4
R311Q	1841	1747	911	1.4
R370W	1830	1763	915	1.4
R390W	1804	1797	900	1.4
C489G	1799	1797	905	1.4

TABLE 8 - SASA CALCULATIONS - PART 2

3.6 Docking

Docking is a molecular simulation technique for predicting how small molecules like ligands interact with proteins like enzymes. The ability of a protein (enzyme) or nucleic acid to interact with small molecules to form a supramolecular complex has a big impact on the protein's dynamics, which could either help or hinder its biological function. (Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment, 2015)

3.6.1 Choice of Protein and Ligand

From the literature review, our protein is from the gene ARSA. Thus we are going to keep the 1AUK PDB structure as our protein, and 3-O-Sulfogalactosylceramide is our ligand because it falls under the category of sulfatides, which cause demyelination in MLD patients (Narayanan et al., 2019). Both the molecules are collected from the protein data bank (PDB).

3.6.2 Tools used for docking

A Web Server for Protein-Ligand Docking that is Free, known as DockThor was used for finding the binding sites, for the entire protein. Furthermore, these results were modelled with the help of software called Pymol. Pymol executed the part where, native needed to align with the mutant, in order to see where the ligands are placed, in the superimposed compound.

3.6.3 Structure Visualization

The following steps are been followed for superimposing the native to the mutant model, in the Pymol software. The protein and ligand have been prepped, with the help of DockThor. The bright green colour represent 1AUK and the light blue colour represent the ligand in all the mutants. Any colour present in the in the structure will be the mutant and its ligand. This is done in order to differentiate between the structures.

Step 1: Load the native protein (1AUK), Into the Pymol software

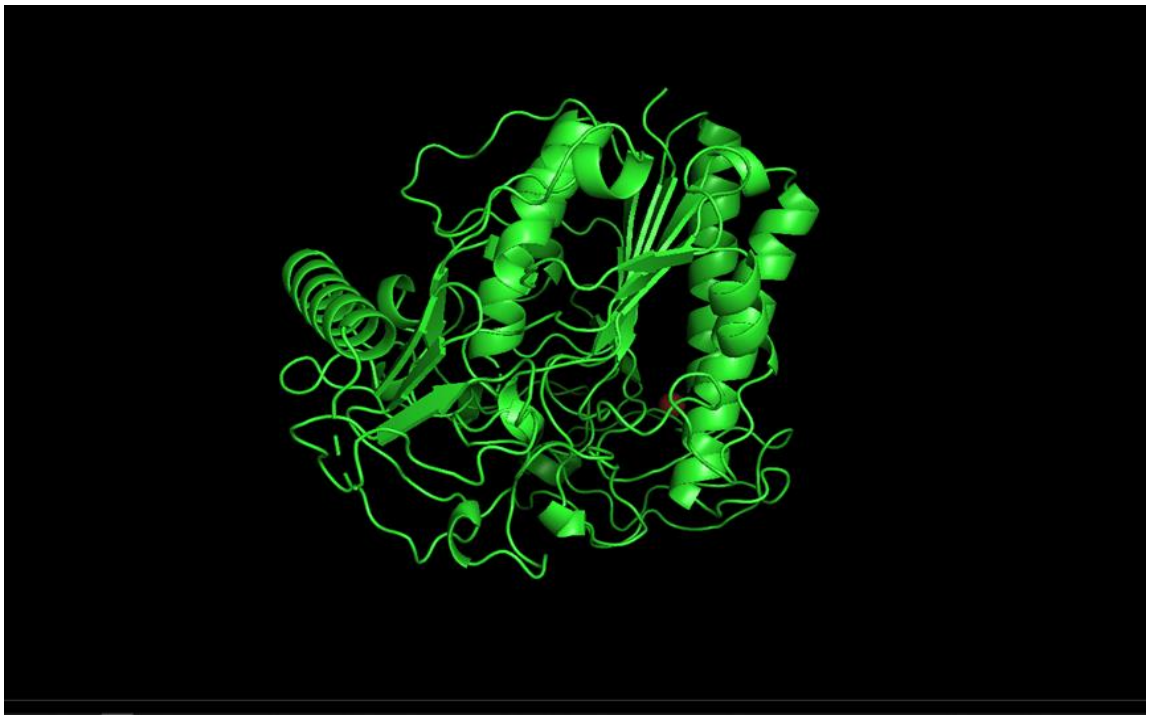


FIGURE 26 - STEP 1

Step 2: Load the ligand (3-O-Sulfogalactosylceramide).

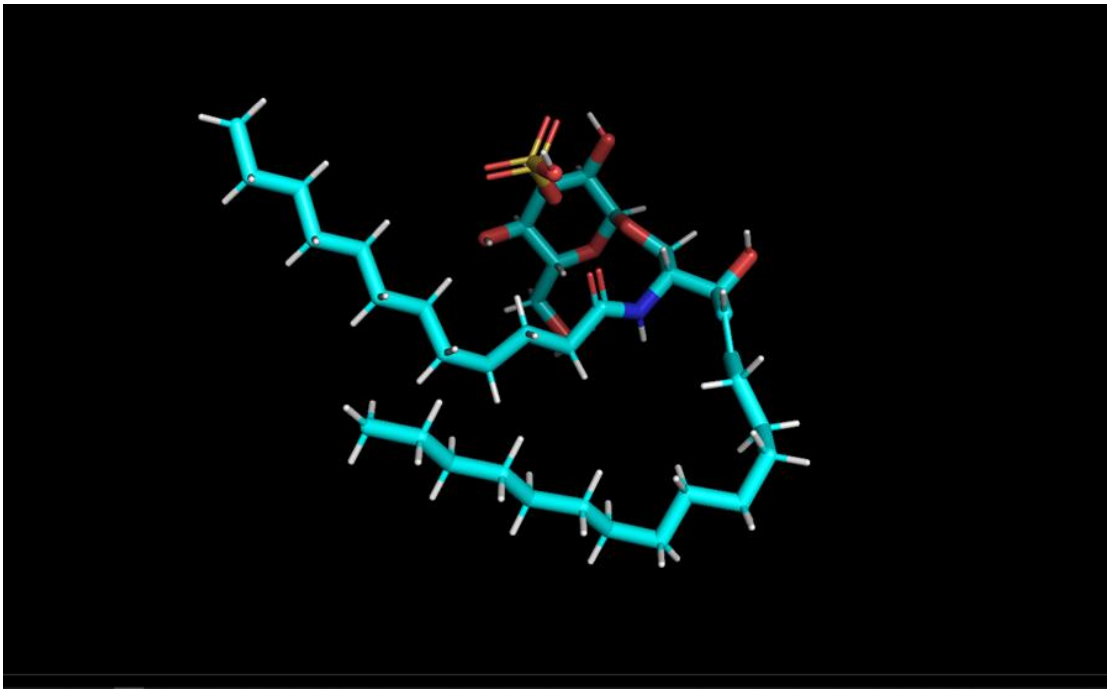


FIGURE 27 - STEP 2

Step 3: The enzyme has located the binding site of the enzyme.

FIGURE 28 - STEP 3

Step 4: Add mutant (T274M) protein to this native (1AUK) protein molecule.

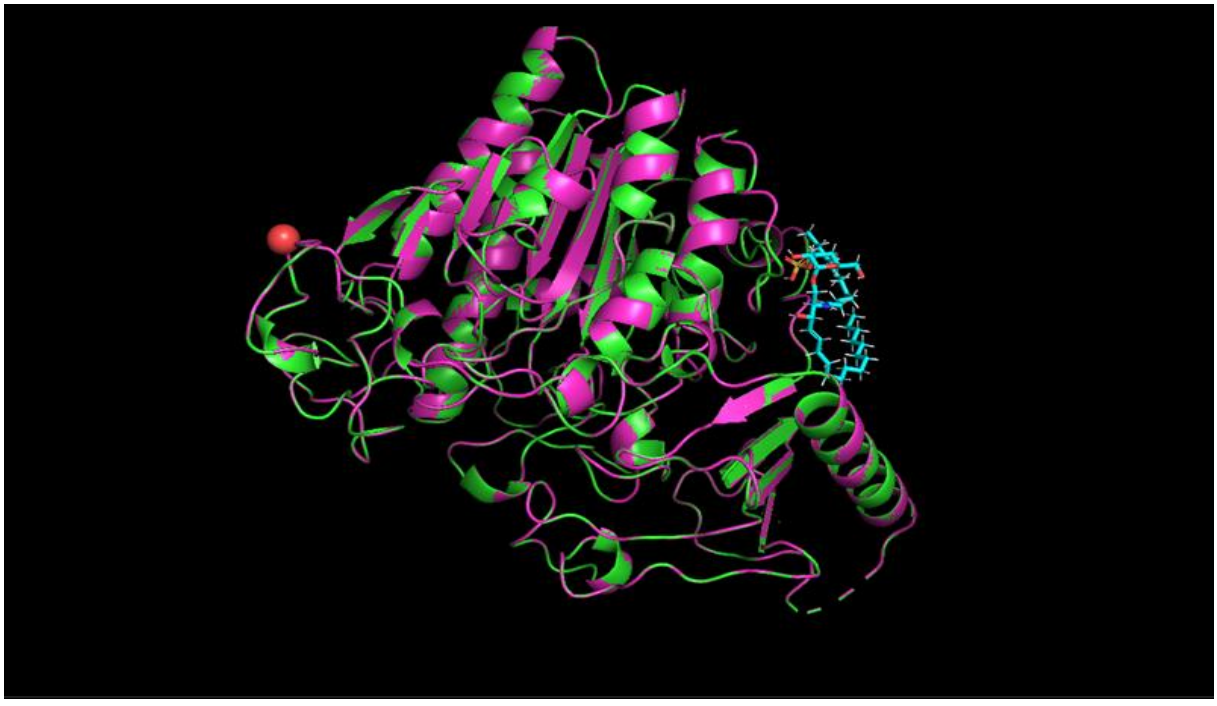


FIGURE 29 - STEP 4

Step 5: Add ligand, of the mutant (T274M), to the current structure. T274M DOCKED.

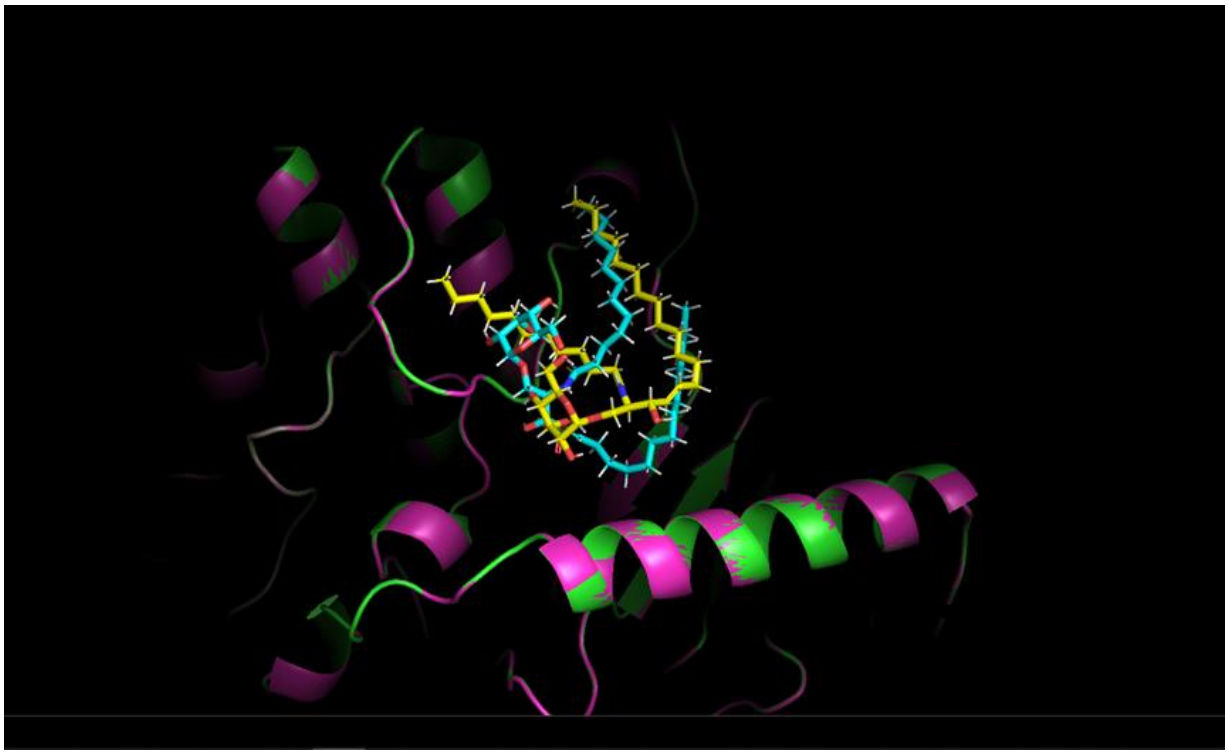


FIGURE 30 - STEP 5

Step 6: Final Model -T274M

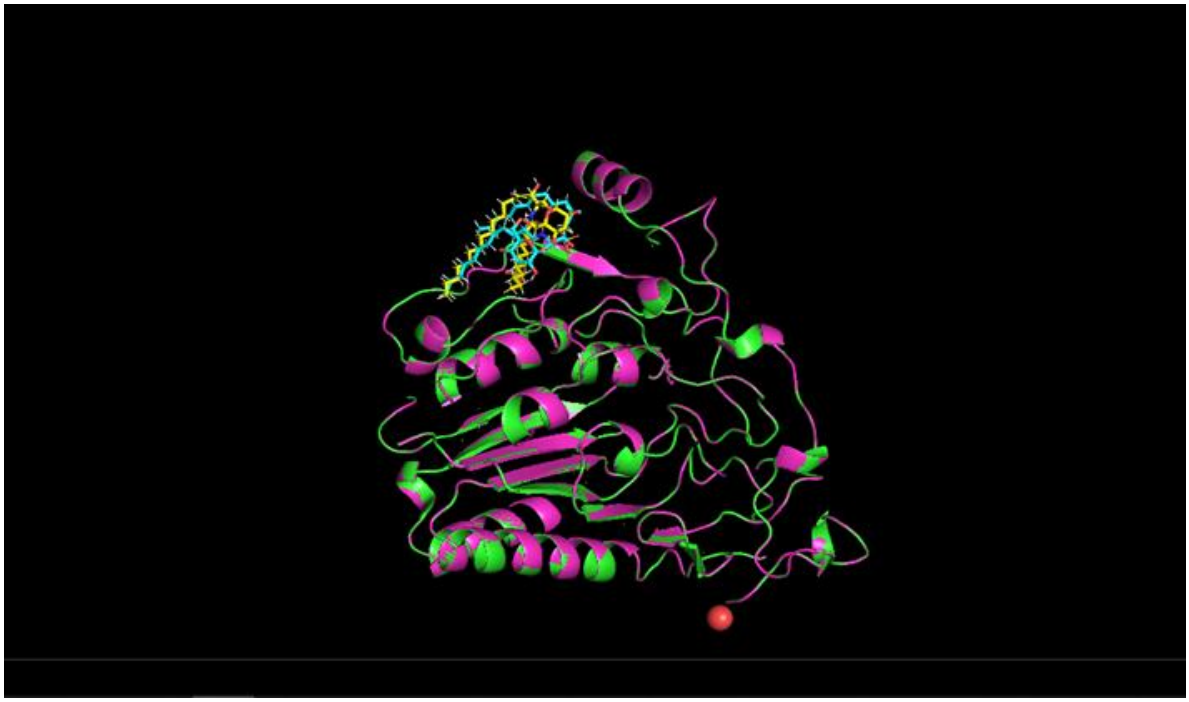


FIGURE 31 - T274M DOCKE

Final Model - R311Q

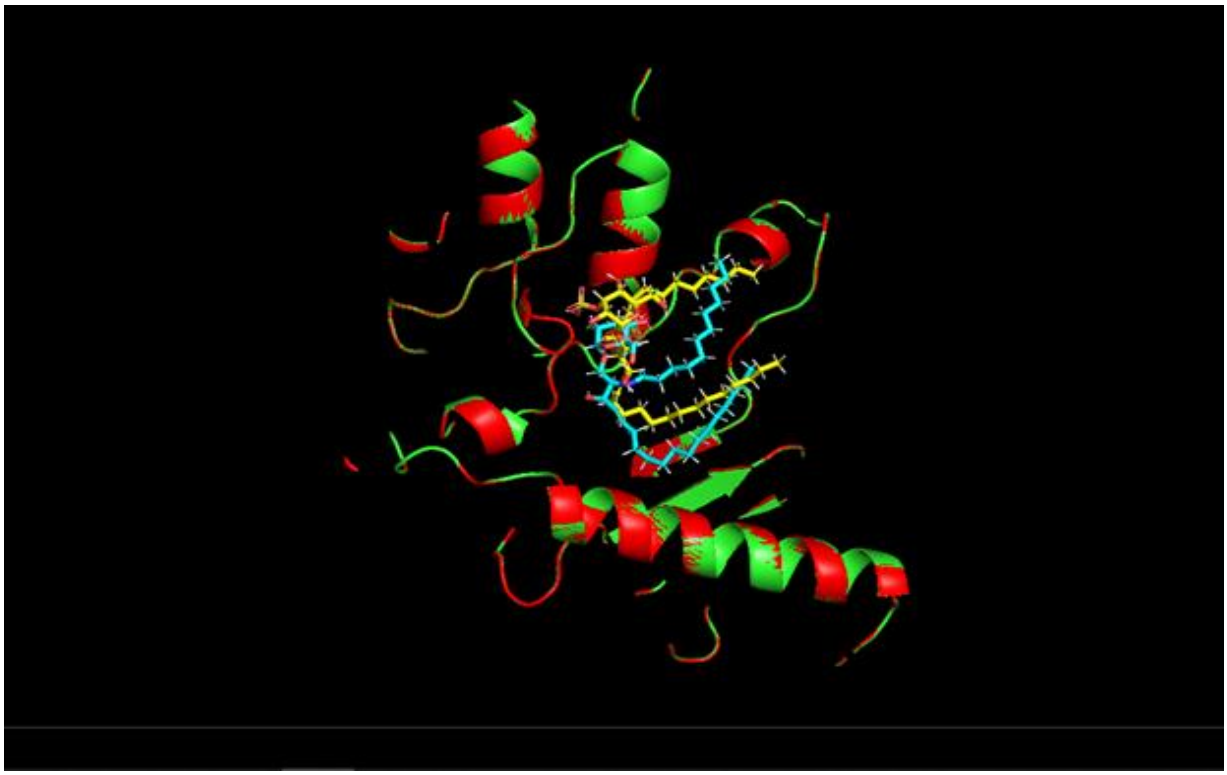


FIGURE 32 - R311Q DOCKED

Final model R370W

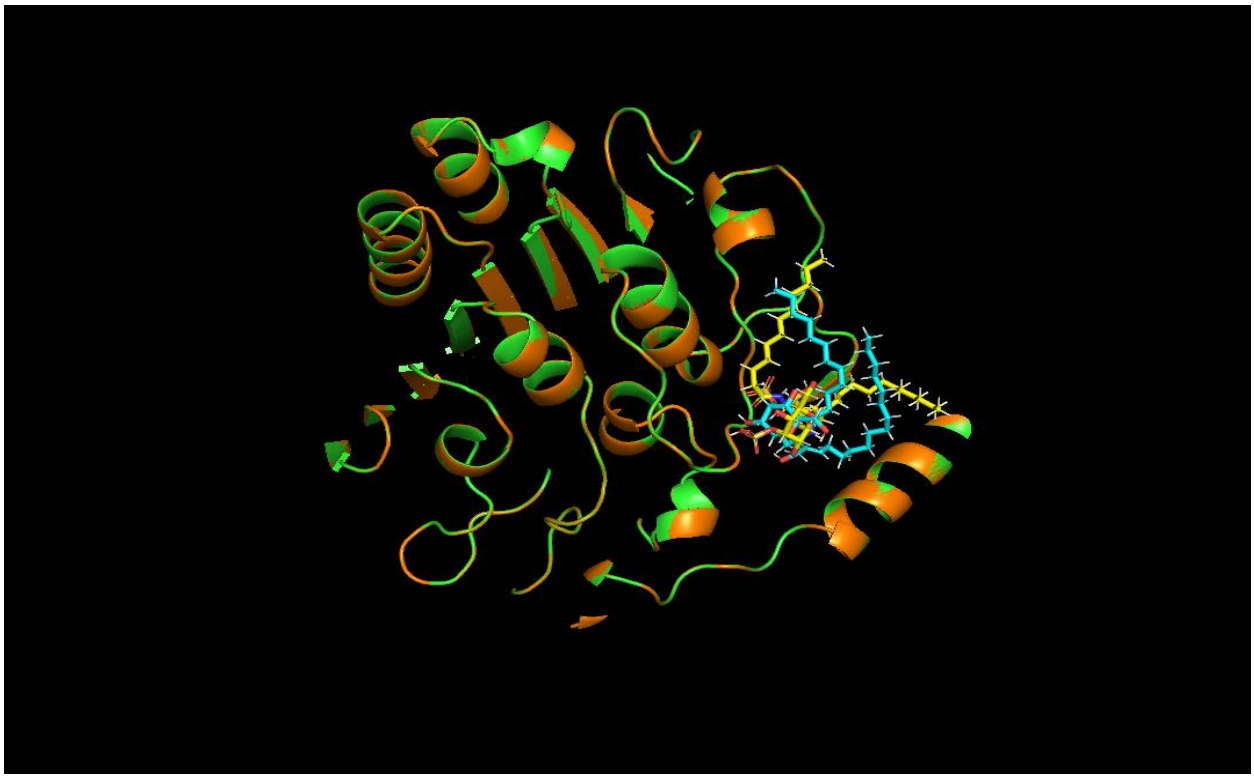


FIGURE 33 - R370W DOCKED

Final model R390W

FIGURE 34 - R390W DOCKED

Final model C489G

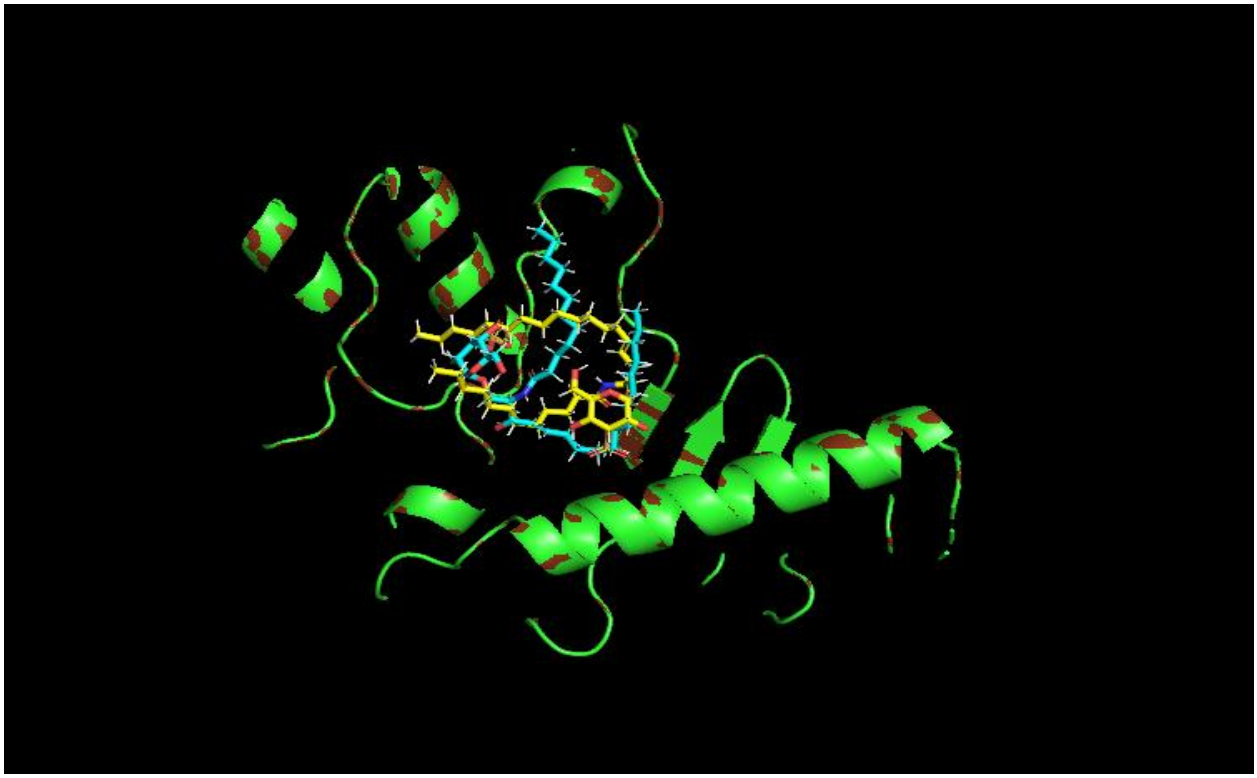


FIGURE 35 - C489G DOCKED

4. Results

The RCSB database was used to extract the PDB structure of 1AUK. Alongside that, the Uniprot ID: P15289 for the 1AUK structure was derived from a Uniprot database entry. The native and mutant structures of 1auk were used as inputs for a single model analysis and further simulation.

Under single model analysis, we have done intramolecular interactions, for native and mutant, a comparative study showing the hydrophobic interaction was done and presented graphically, refer to (Figure 8).Followed up but the secondary structure analysis, where the overall percentage of the alpha helix turn, beta bridges and unstructured loop, coil were calculated mentioned in the (Table 5). Further fluctuation analysis was obtained between the native and every mutant, the line graph (Figure 16 - 20), points out the differences in.SASA calculations were also noted and presented in the (Table 7 and 8) Contact map analysis was also done and the values are presented in (Table 6). Below given (Figure 36), is a bar graph with the comparison of the mutant and the native contact map overlap.

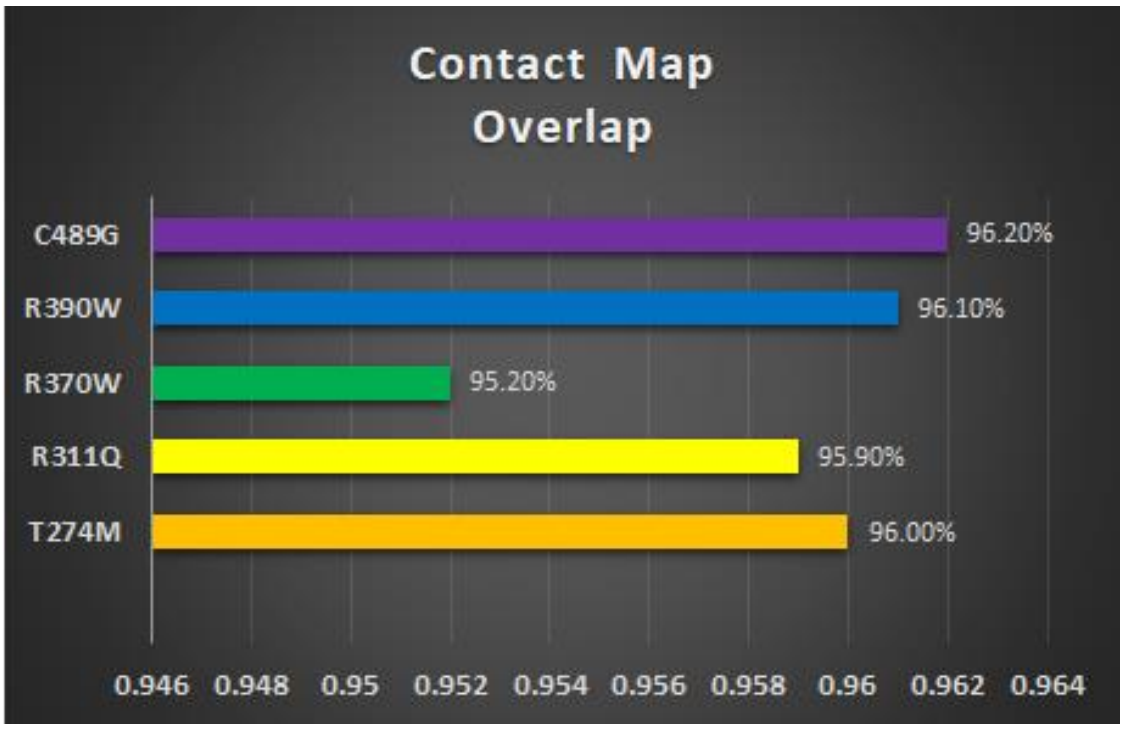


FIGURE 36 - CONTACT MAP OVERLAP

Finally, docking was done for the native, and for each mutant structure, the docked images can be found in every mutant can be found in (Figure 31 - Figure 35). Below given (Table 9), is the statistics of the docking compound.

Compounds	Affinity	Total Energy	VDW Energy	Elec. Energy
1AUK	-8.43	-3.146	-24.424	-15.503
T274M	-8.542	-7.314	-21.154	-15.007
R311Q	-8.089	-2.974	-25.659	-10.29
R37OW	-8.799	-3.816	-26.435	-9.939
R390W	-7.807	-3.484	-20.461	-10.274
C489G	-8.053	-8.19	-28.536	-6.473

TABLE 9 - DOCKING

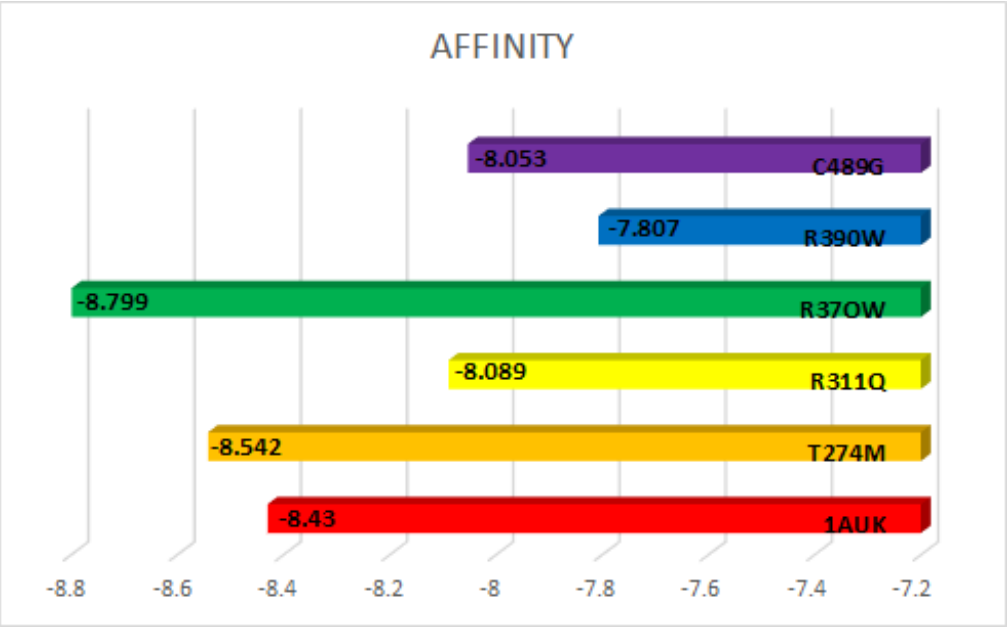


FIGURE 37 - AFFINITY

5. Discussion

In this study, we assess deleterious SAPs and conduct comparative and structural analysis in the native and mutant of the ARSA gene. Methods such as energy minimization and single model analysis were done to predict the quantitative and qualitative study of each mutant, as well as the difference between the native and mutants.

Energy minimization was done for the native and the mutant where the Score returned as negative values stating that all compounds are energetically stable. Further, in single model analysis, under molecular intermolecular interactions, we found that mutations R370W have the highest count of hydrophobic interactions, followed by R390W which means they are more rigid in format compared to the other mutants. This was followed by the secondary structural analysis where the overall percentage difference between the compounds varied from 1% to 3%, but their end result would be significant. This was backed up by fluctuations analysis, where the RMSF of residues, continuously varied in comparison with native and every mutant. Under contact map analysis, R370W has the highest number of unique contacts, which was followed up by R311Q, which leads to the least percentage of the contact map overlap.

The structural integrity, folding mechanisms, and functional properties of proteins are all known to be influenced by various interactions (Akke & Forsén, 1990). Under SASA analysis. It was found that the number of surface atoms was high, in the native and mutant, except for one mutant C489G, where the no. of buried atoms was the highest, this states that this mutant is more compact than the rest of the compounds. Further, docking was done for the native and mutant, with the ligand 3-o-sulfogalactosylceramide, where we can visualize how the native superimposed the mutant, and changes in the ligand pose due to these interactions.

Narayanan et al., 2019 states that their research emphasizes the importance of molecular diagnosis in all MLD patients. Thus, we have studied the deleterious mutations that occurred in the ARSA gene, and have produced a qualitative, and structural analysis of the mutant and native, for patients with MLD.

6. Conclusion

We have demonstrated a comparative molecular study of the deleterious mutants present in the ARSA gene, where various interactions and conformations changes were noted. These predictions are built on the exciting evidence of clinical trials done on MLD patients (Narayanan et al., 2019). Our study on these mutants, emphasizes the significance of genetic research, which is crucial for patients with MLD. In order to address the diagnostic challenges of MLD, future clinical trials and genetic research is recommended.

7. References

Amin, M., Elsayed, L., & Ahmed, A. E. (2017). Clinical and Genetic Characteristics of Leukodystrophies in Africa. *Journal of Neurosciences in Rural Practice*, *08*(S 01), S089–S093. https://doi.org/10.4103/jnrp.jnrp_511_16

Amr, K., Fateen, E., Mansour, L., Tosson, A. M., Zaki, M. S., Salam, G. MH. A., … El-Bassyouni, H. T. (2020). Clinical, Biochemical, and Molecular Characterization of Metachromatic Leukodystrophy Among Egyptian Pediatric Patients: Expansion of the ARSA Mutational Spectrum. *Journal of Molecular Neuroscience*. https://doi.org/10.1007/s12031-020-01734-1

Anish Lamichhane, & Franklyn Rocha Cabrero. (2020, October 23). Metachromatic Leukodystrophy. Retrieved April 28, 2021, from Nih.gov website: https://www.ncbi.nlm.nih.gov/books/NBK560744/#article-25045.s4

Beerepoot, S., van Dooren, S. J. M., Salomons, G. S., Boelens, J. J., Jacobs, E. H., van der Knaap, M. S., … Wolf, N. I. (2020). Metachromatic leukodystrophy genotypes in The Netherlands reveal novel pathogenic ARSA variants in non-Caucasian patients. *Neurogenetics*, *21*(4), 289–299. https://doi.org/10.1007/s10048-020-00621-6

Blenda, A. V. (2021, April 3). Metachromatic Leukodystrophy: Background, Pathophysiology, Epidemiology. Retrieved April 28, 2021, from Medscape.com website: https://emedicine.medscape.com/article/951840-overview#a6

Doerr, J., Böckenhoff, A., Ewald, B., Ladewig, J., Eckhardt, M., Gieselmann, V., … Koch, P. (2015). Arylsulfatase A Overexpressing Human iPSC-derived Neural Cells Reduce CNS Sulfatide Storage in a Mouse Model of Metachromatic Leukodystrophy. *Molecular Therapy*, *23*(9), 1519–1531. https://doi.org/10.1038/mt.2015.106

Durham, E., Dorr, B., Woetzel, N., Staritzbichler, R., & Meiler, J. (2009). Solvent accessible surface area approximations for rapid and accurate protein structure prediction. *Journal of Molecular Modeling*, *15*(9), 1093–1108. https://doi.org/10.1007/s00894-009-0454-9

Futerman, A. H., & van Meer, G. (2004). The cell biology of lysosomal storage disorders. *Nature Reviews Molecular Cell Biology*, *5*(7), 554–565. https://doi.org/10.1038/nrm1423

Gallo, S. (2004). Late onset MLD with normal nerve conduction associated with two novel missense mutations in the ASA gene. *Journal of Neurology, Neurosurgery & Psychiatry*, *75*(4), 655–657. https://doi.org/10.1136/jnnp.2003.017400

GUSTAVSON, K.-H. ., & HAGBERG, B. (1971). THE INCIDENCE AND GENETICS OF METACHROMATIC LEUCODYSTROPHY IN NORTHERN SWEDEN. *Acta Paediatrica*, *60*(5), 585–590. https://doi.org/10.1111/j.1651-2227.1971.tb06994.x

Harrington, M., Whalley, D., Twiss, J., Rushton, R., Martin, S., Huynh, L., & Yang, H. (2019). Insights into the natural history of metachromatic leukodystrophy from interviews with caregivers. *Orphanet Journal of Rare Diseases*, *14*(1). https://doi.org/10.1186/s13023-019-1060-2

Hettiarachchi, D., & Dissanayake, V. H. W. (2019). Three novel variants in the arylsulfatase A (ARSA) gene in patients with metachromatic leukodystrophy (MLD). *BMC Research Notes*, *12*(1). https://doi.org/10.1186/s13104-019-4773-3

Kmiecik, S., Gront, D., Kolinski, M., Wieteska, L., Dawid, A. E., & Kolinski, A. (2016). Coarse-Grained Protein Models and Their Applications. *Chemical Reviews*, *116*(14), 7898–7936. https://doi.org/10.1021/acs.chemrev.6b00163

Luria, S. E., & M. Delbrück. (1943). MUTATIONS OF BACTERIA FROM VIRUS SENSITIVITY TO VIRUS RESISTANCE. *Genetics*, *28*(6), 491–511. Retrieved from https://www.genetics.org/content/28/6/491

Narayanan, D. L., Matta, D., Gupta, N., Kabra, M., Ranganath, P., Aggarwal, S., … Dalal, A. (2019). Spectrum of ARSA variations in Asian Indian patients with Arylsulfatase A deficient metachromatic leukodystrophy. *Journal of Human Genetics*, *64*(4), 323–331. https://doi.org/10.1038/s10038-019-0560-1

Pastores, G. M. (2010). Review: Therapeutic approaches for lysosomal storage diseases. *Therapeutic Advances in Endocrinology and Metabolism*, *1*(4), 177–188. https://doi.org/10.1177/2042018810384429

Pekgül, F., Eroğlu-Ertuğrul, N. G., Bekircan-Kurt, C. E., Erdem-Ozdamar, S., Çetinkaya, A., Tan, E., … Özkara, H. A. (2020). Comprehensive clinical, biochemical, radiological and genetic analysis of 28 Turkish cases with suspected metachromatic leukodystrophy and their relatives. *Molecular Genetics and Metabolism Reports*, *25*, 100688. https://doi.org/10.1016/j.ymgmr.2020.100688

Porollo, A., & Meller, J. (2010). POLYVIEW-MM: web-based platform for animation and analysis of molecular simulations. *Nucleic Acids Research*, *38*(Web Server), W662–W666. https://doi.org/10.1093/nar/gkq445

Ruzhnikov, M. R. Z., Brimble, E., & Hickey, R. E. (2021). Early Signs and Symptoms of Leukodystrophies: A Case-Based Guide. *Pediatrics in Review*, *42*(3), 133–146. https://doi.org/10.1542/pir.2019-0184

Staretz-Chacham, O., Lang, T. C., LaMarca, M. E., Krasnewich, D., & Sidransky, E. (2009). Lysosomal Storage Disorders in the Newborn. *PEDIATRICS*, *123*(4), 1191–1207. https://doi.org/10.1542/peds.2008-0635

Tina, K. G., Bhadra, R., & Srinivasan, N. (2007a). PIC: Protein Interactions Calculator. *Nucleic Acids Research*, *35*(Web Server), W473–W476. https://doi.org/10.1093/nar/gkm423

Tina, K. G., Bhadra, R., & Srinivasan, N. (2007b). PIC: Protein Interactions Calculator. *Nucleic Acids Research*, *35*(Web Server), W473–W476. https://doi.org/10.1093/nar/gkm423

Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment. (2015). https://doi.org/10.1016/c2014-0-00286-9

Welf, Erik S., & Danuser, G. (2014). Using Fluctuation Analysis to Establish Causal Relations between Cellular Events without Experimental Perturbation. *Biophysical Journal*, *107*(11), 2492–2498. https://doi.org/10.1016/j.bpj.2014.10.032

Wu, S., Hou, M., Zhang, Y., Song, J., Guo, Y., Liu, P., … Chen, Z. (2020). Chinese Cases of Metachromatic Leukodystrophy with the Novel Missense Mutations in ARSA Gene. *Journal of Molecular Neuroscience*, *71*(2), 245–251. https://doi.org/10.1007/s12031-020-01643-3

Ycart, B. (2013). Fluctuation Analysis: Can Estimates Be Trusted? *PLoS ONE*, *8*(12), e80958. https://doi.org/10.1371/journal.pone.0080958

Zhang, B., Li, J., & Lü, Q. (2018). Prediction of 8-state protein secondary structures by a novel deep learning architecture. *BMC Bioinformatics*, *19*(1). https://doi.org/10.1186/s12859-018-2280-5

Autosomal recessive: MedlinePlus Medical Encyclopedia. (2016). Retrieved April 28, 2021, from Medlineplus.gov website: https://medlineplus.gov/ency/article/002052.htm#:~:text=Autosomal%20recessive%20is%20one%20of,disease%20or%20trait%20to%20develop.

Burke, D. (2012, August 20). Metachromatic Leukodystrophy. Retrieved April 27, 2021, from Healthline website: https://www.healthline.com/health/metachromatic-leukodystrophy#treatments

DockThor. (2021). Retrieved May 1, 2021, from Lncc.br website: https://dockthor.lncc.br/v2/

https://www.facebook.com/WebMD. (2016, December). What Are Lysosomal Storage Disorders? Retrieved April 27, 2021, from WebMD website: https://www.webmd.com/children/what-are-lysosomal-storage-disorders

Metachromatic Leukodystrophy - NORD (National Organization for Rare Disorders). (2021, January 25). Retrieved April 28, 2021, from NORD (National Organization for Rare Disorders) website: https://rarediseases.org/rare-diseases/metachromatic-leukodystrophy/

Metachromatic Leukodystrophy Information Page | National Institute of Neurological Disorders and Stroke. (2021). Retrieved April 28, 2021, from Nih.gov website: https://www.ninds.nih.gov/Disorders/All-Disorders/Metachromatic-Leukodystrophy-Information-Page

Metachromatic leukodystrophy: MedlinePlus Genetics. (2020). Retrieved April 28, 2021, from Medlineplus.gov website: https://medlineplus.gov/genetics/condition/metachromatic-leukodystrophy/#frequency

Nelson, S. L. (2021, February 16). Lysosomal Storage Disease: Overview, Classification of Lysosomal Storage Diseases, Glycogen Storage Disease Type II. Retrieved April 28,

2021, from Medscape.com website: https://emedicine.medscape.com/article/1182830-overview#a1

RCSB Protein Data Bank. (2011a). RCSB PDB - 1AUK: HUMAN ARYLSULFATASE A. Retrieved April 28, 2021, from Rcsb.org website: https://www.rcsb.org/structure/1AUK

RCSB Protein Data Bank. (2011b). RCSB PDB - 1E33: Crystal structure of an Arylsulfatase A mutant P426L. Retrieved April 28, 2021, from Rcsb.org website: https://www.rcsb.org/structure/1E33

SNPs&GO - Predicting disease associated variations using GO terms. (2021). Retrieved April 29, 2021, from Biofold.org website: https://snps.biofold.org/snps-and-go/snps-and-go.html

Wikipedia Contributors. (2020, December 14). Protein contact map. Retrieved April 29, 2021, from Wikipedia website: https://en.wikipedia.org/wiki/Protein_contact_map

Wikipedia Contributors. (2021, April 27). Metachromatic leukodystrophy. Retrieved April 28, 2021, from Wikipedia website: https://en.wikipedia.org/wiki/Metachromatic_leukodystrophy#cite_note-:0-17

YASARA - Yet Another Scientific Artificial Reality Application. (2013). Retrieved April 29, 2021, from Yasara.org website: http://www.yasara.org/