SRI RAMACHANDRA
INSTITUTE OF HIGHER EDUCATION AND RESEARCH
(Category-I Deemed to be University) Porur, Chennai, India
[Accredited by NAAC with "A++" Grade]

Investigating the Deleterious nsSNPs in causing Phenylketonuria by Molecular Docking Approach

Guide: **Dr. MAGESH R,** Department of Biotechnology,
SRIHER

By
Susitra G
Reg no: B0119017

IV year, B.Sc.(Hons) Biomedical Sciences - Elective Biotechnology

Introduction

- **Phenylketonuria** PKU) is an autosomal recessive inborn error of phenylalanine (Phe) metabolism (Williams et al. 2008).
- It results from a deficiency in phenylalanine hydroxylase (PAH), often caused by mutations in the PAH gene(Williams et al. 2008).
- The PAH gene spans about 171 kb and 13 exons and is found on chromosome 12q23.2, encoding the L-phenylalanine hydroxylase enzyme (Williams et al. 2008).
- About 30,000 people in the United States have phenylketonuria (Thompson 2008).

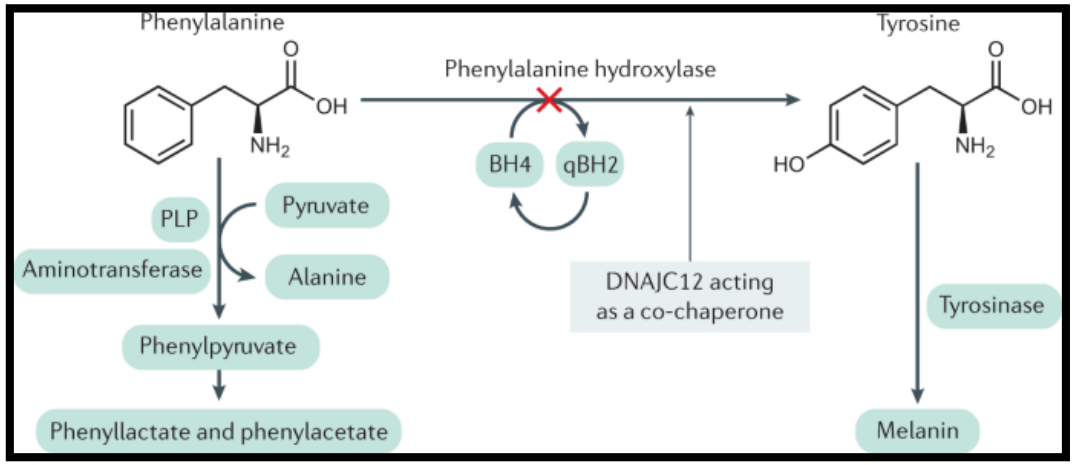


Figure 1: **Phenylalanine metabolism and PKU [(van Spronsen et al. 2021)]**

Review of Literature

- Most forms of PKU and hyperphenylalaninemia arise from mutations in the PAH gene. (Williams et al. 2008)
- The conversion of phenylalanine (Phe) to tyrosine (Tyr) involves a hydroxylating system comprising phenylalanine hydroxylase (PAH), the unconjugated pterin cofactor tetrahydrobiopterin (BH4), and enzymes responsible for BH4 regeneration, namely dihydropteridine reductase and 4α-carbinolamine dehydratase. (Williams et al. 2008)

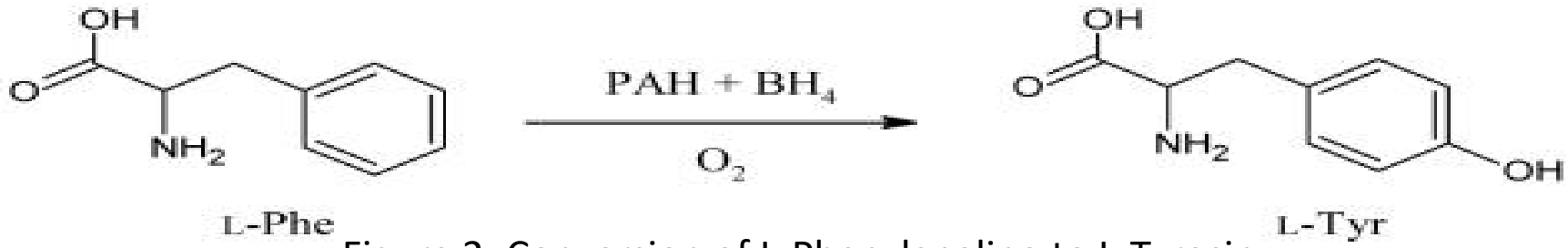


Figure 2: Conversion of L-Phenylanaline to L-Tyrosin

- Untreated phenylketonuria manifests with microcephaly, profound mental retardation, and, in certain instances, progressive supranuclear motor disturbances. (Pietz 1998)
- Sapropterin dihydrochloride, as per FDA-approved labelling, effectively lowers blood phenylalanine levels in tetrahydrobiopterin-responsive phenylketonuria. (Thompson 2008)

Aim

- To identify and analyze the response of the deleterious variants with Sapropterin dihydrochloride to reduce phenylketonuria conditions.

Objectives

- To retrieve the protein sequence and structure.
- To identify the deleterious amino acid variants using various computational tools.
- To perform docking analyses to check the response of the drug towards the wild and mutant protein structure

Materials and Methods

RETRIEVAL OF PROTEIN INFORMATION

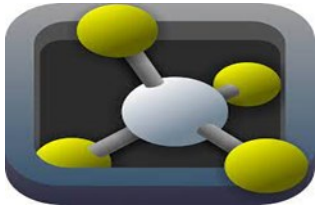
SNP ANALYSIS

PATHOGENICITY ANALYSIS

CONSERVATION ANALYSIS

BIOPHYSICAL PROPERTY ANALYSIS

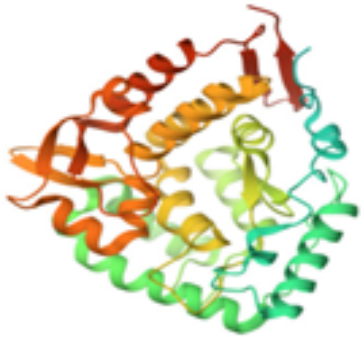
STABILITY ANALYSIS



STRUCTURAL ANALYSIS

DOCKING ANALYSIS

Results

Gene	Protein	Uniprot ID	Sequence	PDB	Structure
PAH	**Beta-hexosaminidase subunit alpha**	P00439 452 aa	>sp|P00439|PH4H_HUMAN Phenylalanine-4-hydroxylase OS=Homo sapiens OX=9606 GN=PAH PE=1 SV=1 MSTAVLENPG LGRKLSDFGQ ETSYIEDNCN QNGAISLIFS LKEEVGALAK VLRLFEENDV NLTHIESRPS RLKKDEYEFF THLDKRSLPA LTNIIKILRH DIGATVHELS RDKKKDTVPW FPRTIQELDR FANQILSYGA ELDADHPGFK DPVYRARRKQ FADIAYNYRH GQPIPRVEYM EEEKKTWGTV FKTLKSLYKT HACYEYNHIF PLLEKYCGFH EDNIPQLEDV SQFLQTCTGF RLRPVAGLLS SRDFLGGLAF RVFHCTQYIR HGSKPMYTPE PDICHELLGH VPLFSDRSFA QFSQEIGLAS LGAPDEYIEK LATIYWFTVE FGLCKQGDSI KAYGAGLLSS FGELQYCLSE KPKLLPLELE KTAIQNYTVT EFQPLYYVAE SFNDAKEKVR NFAATIPRPF SVRYDPYTQR IEVLDNTQQL KILADSINSE IGILCSALQK IK	6HPO	

SNP Retrieval

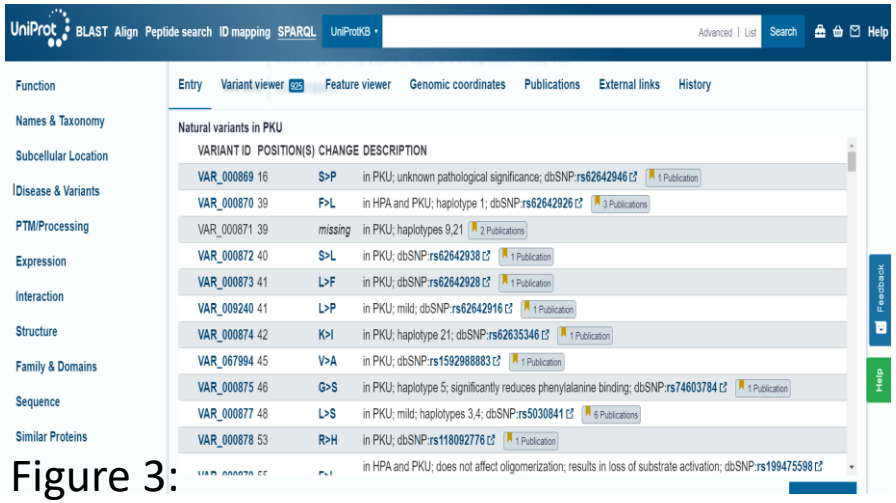
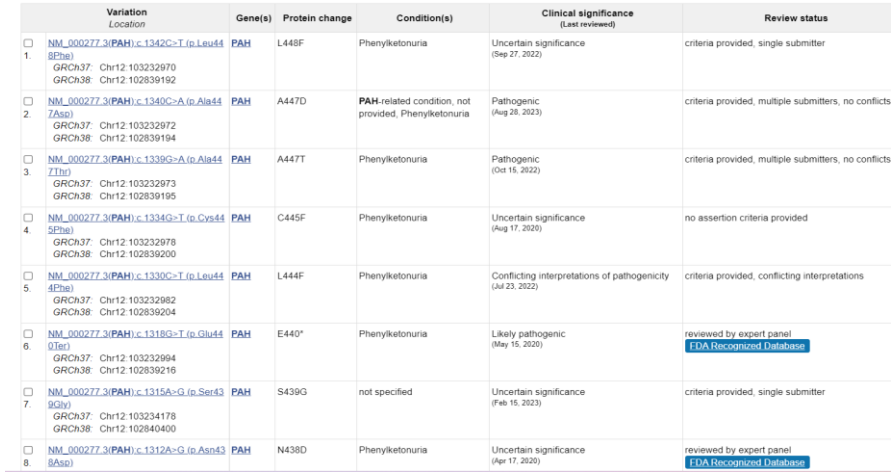


Figure 3:



	Variation Location	Gene(s)	Protein change	Condition(s)	Clinical significance (Last reviewed)	Review status
1.	NM_000277.3(PAH):c.1342C>T (p.Leu448Phe) GRCh37: Chr12:103232970 GRCh38: Chr12:102839192	PAH	L448F	Phenylketonuria	Uncertain significance (Sep 27, 2022)	criteria provided, single submitter
2.	NM_000277.3(PAH):c.1340C>A (p.Ala447Asp) GRCh37: Chr12:103232972 GRCh38: Chr12:102839194	PAH	A447D	PAH-related condition, not provided, Phenylketonuria	Pathogenic (Aug 28, 2023)	criteria provided, multiple submitters, no conflicts
3.	NM_000277.3(PAH):c.1339G>A (p.Ala447Thr) GRCh37: Chr12:103232973 GRCh38: Chr12:102839195	PAH	A447T	Phenylketonuria	Pathogenic (Oct 15, 2022)	criteria provided, multiple submitters, no conflicts
4.	NM_000277.3(PAH):c.1334G>T (p.Cys445Phe) GRCh37: Chr12:103232978 GRCh38: Chr12:102839200	PAH	C445F	Phenylketonuria	Uncertain significance (Aug 17, 2020)	no assertion criteria provided
5.	NM_000277.3(PAH):c.1330C>T (p.Leu444Phe) GRCh37: Chr12:103232982 GRCh38: Chr12:102839204	PAH	L444F	Phenylketonuria	Conflicting interpretations of pathogenicity (Jul 23, 2022)	criteria provided, conflicting interpretations
6.	NM_000277.3(PAH):c.1318G>T (p.Glu440Ter) GRCh37: Chr12:103232994 GRCh38: Chr12:102839216	PAH	E440*	Phenylketonuria	Likely pathogenic (May 15, 2020)	reviewed by expert panel FDA Recognized Database
7.	NM_000277.3(PAH):c.1315A>G (p.Ser439Gly) GRCh37: Chr12:103234178 GRCh38: Chr12:102840400	PAH	S439G	not specified	Uncertain significance (Feb 15, 2023)	criteria provided, single submitter
8.	NM_000277.3(PAH):c.1312A>G (p.Asn438Asp)	PAH	N438D	Phenylketonuria	Uncertain significance (Apr 17, 2020)	reviewed by expert panel FDA Recognized Database

Figure 4:

Accession Number	Amino acid change	Codon number	Phenotype	Reference
CM000541	M-R	1	Phenylketonuria	Hennermann (2000) Hum Mutat 15, 254 Additional report available to subscribers
CM920539	M-I	1	Phenylketonuria	Eiken (1992) Hum Mutat 1, 388 Additional report available to subscribers
CM890092	M-V	1	Phenylketonuria	John (1989) Am J Hum Genet 45, 905 Additional report available to subscribers
CM155973	M-L	1	Phenylketonuria	Vela-Amieva (2015) Clin Genet 88, 62 Additional report available to subscribers
CM010947	S-P	16	Phenylketonuria	Guldberg (1995) PAH UP, #229 Additional report available to subscribers
CM112733	S-Y	16	Phenylketonuria	Okano (2011) J Hum Genet 56, 306
CM1511421	Q-P	20	Phenylketonuria	Li (2015) Sci Rep 5, 15769
HM972014	T-K	22	Phenylketonuria	Froster (2001) Hum Genet 109, 127 Additional report available to subscribers
CM1314660	L-P	37	Phenylketonuria	Bik-Multanowski (2013) Acta Biochim Pol 60, 613 Additional report available to subscribers Additional phenotype report available to subscribers
				Forrest (1991) Am J Hum Genet 49, 175 Additional phenotype report available to subscribers Additional report available to subscribers Additional report available to subscribers Functional characterisation report available to subscribers Additional report available to subscribers Functional characterisation report available to subscribers Functional characterisation report available to subscribers Additional report available to subscribers Functional characterisation report available to subscribers

Figure 5:

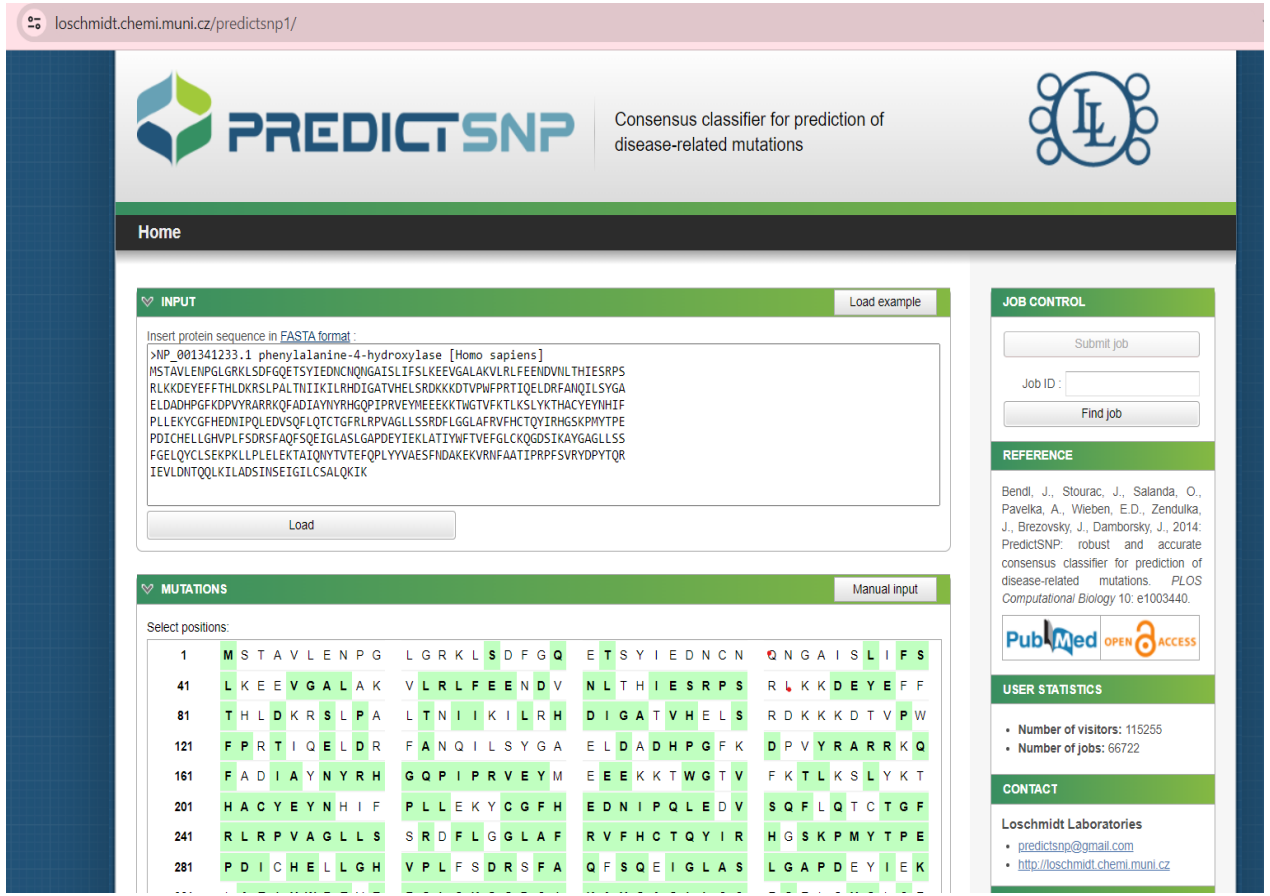
Gene	PAH
Uniprot	195
HGMD	493
ClinVar	157
Total with repeats	845
Total without repeats	514

UniProt: https://www.uniprot.org/uniprotkb/P06865/entry#disease_variants

ClinVar: https://www.ncbi.nlm.nih.gov/clinvar

HGMD: https://www.hgmd.cf.ac.uk/ac/all.php

Pathogenicity Analysis



Pathogenicity Tools	Variants
Predict SNP prediction	377
MAPP prediction	323
PhD-SNP prediction	427
PolyPhen-1 prediction	346
PolyPhen-2 prediction	397
SIFT prediction	428
SNAP prediction	277
PANTHER prediction	426

Total common variants 191

Predict SNP (Home page): https://loschmidt.chemi.muni.cz/predictsnp1/

Conservation Analysis

Figure 7.

Figure 7.1: Consurf

CONSERVATIONAL	SCORES
Score 9	122
Score 8	38
Score 7	30

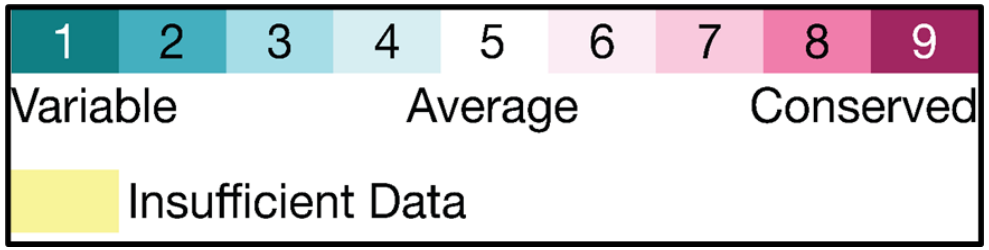


Figure 7.2: Scores

Biophysical Property Analysis

Substitution	GV	GD	Prediction
L48S	0.00	144.08	Class C65
E66K	0.00	56.87	Class C55
R68S	0.00	109.21	Class C65
R68G	0.00	125.13	Class C65
D129G	0.00	93.77	Class C65
D129Y	0.00	159.94	Class C65
H146Y	0.00	83.33	Class C65
P147L	0.00	97.78	Class C65
D151H	0.00	81.24	Class C65
D151E	0.00	44.60	Class C35
D151G	0.00	93.77	Class C65
Y154N	0.00	142.23	Class C65
Y154H	0.00	83.33	Class C65
Y154C	0.00	193.72	Class C65
Y154F	0.00	21.61	Class C15
R157S	0.00	109.21	Class C65
R157I	0.00	97.59	Class C65
R157N	0.00	85.11	Class C65

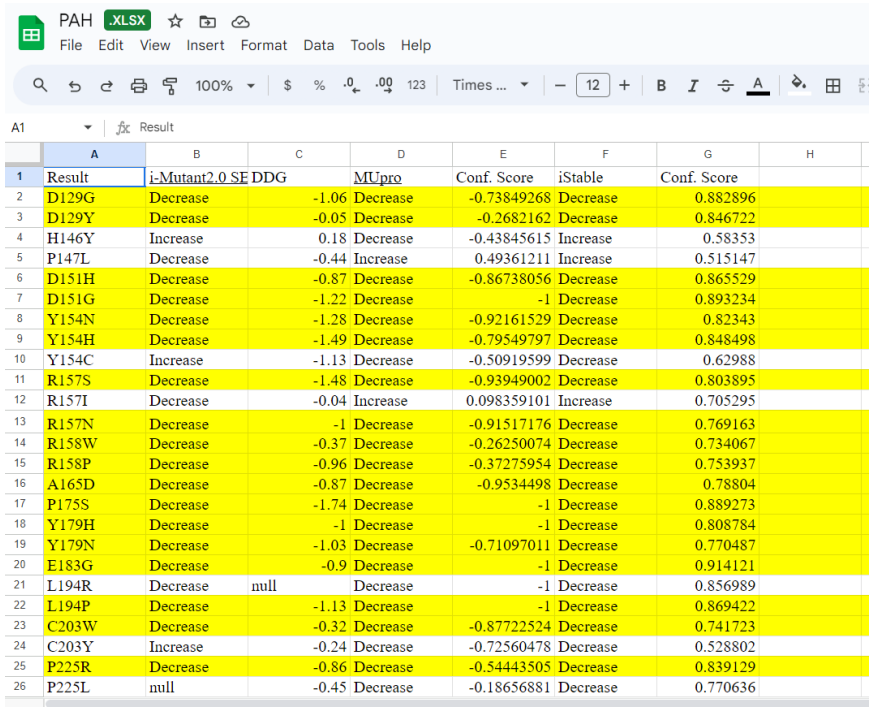
Figure 8.1 : Align GVGD

The following classifiers, ordered from most likely to interfere with function to least likely, were used:
GD>=65+Tan(10)x(GV^2.5) => Class C65 <=> most likely
GD>=55+Tan(10)x(GV^2.0) => Class C55
GD>=45+Tan(15)x(GV^1.7) => Class C45
GD>=35+Tan(50)x(GV^1.1) => Class C35
GD>=25+Tan(55)x(GV^0.95) => Class C25
GD>=15+Tan(75)x(GV^0.6) => Class C15
Else (GD<15+Tan(75)x(GV^0.6)) => Class C0 <=> less likely

Align GVGD: http://agvgd.hci.utah.edu/cgi-bin/agvgd_output.cgi

GENE	Biophysical Property Analysis
TOTAL	122
C65	81

Stability Analysis



PAH .XLSX

File Edit View Insert Format Data Tools Help

A1 Result

	A	B	C	D	E	F	G	H
1	Result	i-Mutant2.0 SE	DDG	MUpro	Conf. Score	iStable	Conf. Score	
2	D129G	Decrease	-1.06	Decrease	-0.73849268	Decrease	0.882896	
3	D129Y	Decrease	-0.05	Decrease	-0.2682162	Decrease	0.846722	
4	H146Y	Increase	0.18	Decrease	-0.43845615	Increase	0.58353	
5	P147L	Decrease	-0.44	Increase	0.49361211	Increase	0.515147	
6	D151H	Decrease	-0.87	Decrease	-0.86738056	Decrease	0.865529	
7	D151G	Decrease	-1.22	Decrease	-1	Decrease	0.893234	
8	Y154N	Decrease	-1.28	Decrease	-0.92161529	Decrease	0.82343	
9	Y154H	Decrease	-1.49	Decrease	-0.79549797	Decrease	0.848498	
10	Y154C	Increase	-1.13	Decrease	-0.50919599	Decrease	0.62988	
11	R157S	Decrease	-1.48	Decrease	-0.93949002	Decrease	0.803895	
12	R157I	Decrease	-0.04	Increase	0.098359101	Increase	0.705295	
13	R157N	Decrease	-1	Decrease	-0.91517176	Decrease	0.769163	
14	R158W	Decrease	-0.37	Decrease	-0.26250074	Decrease	0.734067	
15	R158P	Decrease	-0.96	Decrease	-0.37275954	Decrease	0.753937	
16	A165D	Decrease	-0.87	Decrease	-0.9534498	Decrease	0.78804	
17	P175S	Decrease	-1.74	Decrease	-1	Decrease	0.889273	
18	Y179H	Decrease	-1	Decrease	-1	Decrease	0.808784	
19	Y179N	Decrease	-1.03	Decrease	-0.71097011	Decrease	0.770487	
20	E183G	Decrease	-0.9	Decrease	-1	Decrease	0.914121	
21	L194R	Decrease	null	Decrease	-1	Decrease	0.856989	
22	L194P	Decrease	-1.13	Decrease	-1	Decrease	0.869422	
23	C203W	Decrease	-0.32	Decrease	-0.87722524	Decrease	0.741723	
24	C203Y	Increase	-0.24	Decrease	-0.72560478	Decrease	0.528802	
25	P225R	Decrease	-0.86	Decrease	-0.54443505	Decrease	0.839129	
26	P225L	null	-0.45	Decrease	-0.18656881	Decrease	0.770636	

Figure 9: iStable

Amino Acid Mutations

Amino acid	Overall Stability	Torsion*	Predicted ΔΔG (kcal/mol)
ALA	Destabilising	Unfavourable	-9.62
VAL	Destabilising	Unfavourable	-14.31
LEU	Destabilising	Unfavourable	-9.11
ILE	Destabilising	Unfavourable	-8.51
MET	Destabilising	Unfavourable	-6.01
PRO	Destabilising	Unfavourable	-16.2
TRP	Destabilising	Unfavourable	-5.78
SER	Destabilising	Unfavourable	-6.02
THR	Destabilising	Unfavourable	-8.47
PHE	Destabilising	Unfavourable	-7.34
GLN	Destabilising	Unfavourable	-8.75
LYS	Destabilising	Unfavourable	-14.65
TYR	Destabilising	Unfavourable	-9.35
ASN	Destabilising	Unfavourable	-6.65
CYS	Destabilising	Unfavourable	-6.21
GLU	Destabilising	Unfavourable	-8.54
ASP	Destabilising	Unfavourable	-7.93
ARG	Destabilising	Unfavourable	-10.12
HIS	Destabilising	Unfavourable	-14.25

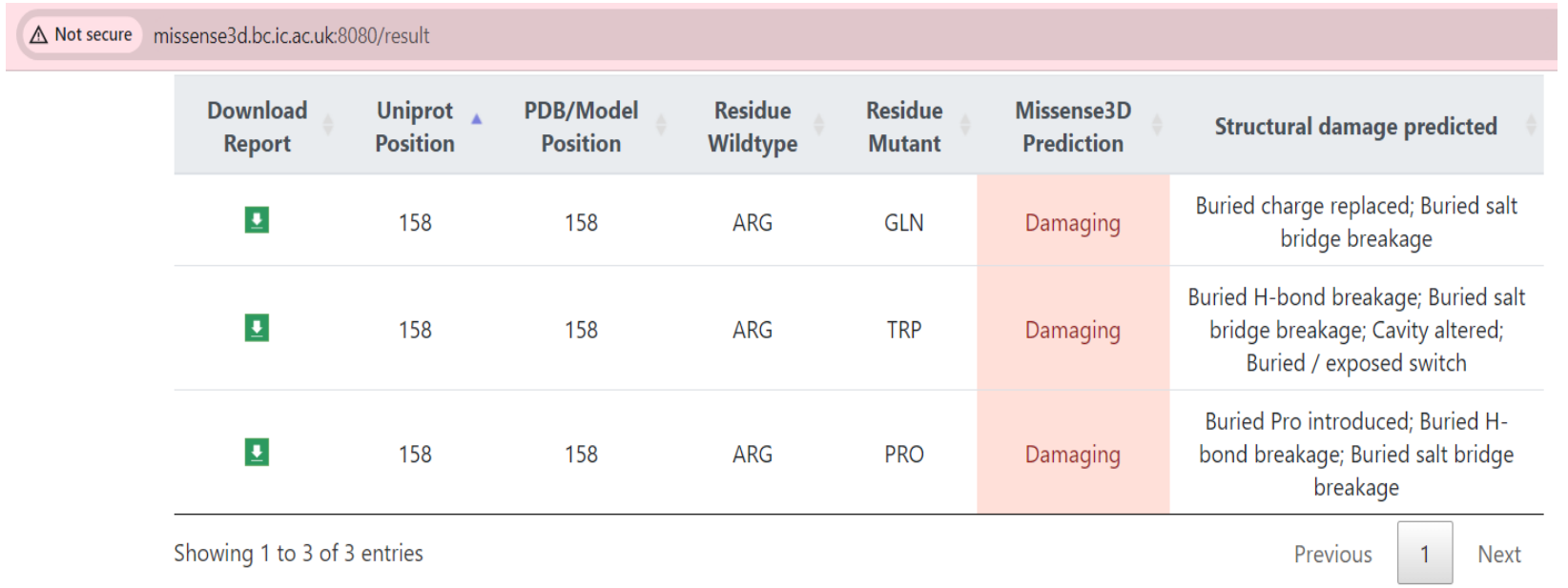
Figure 10: CUPSAT

Total	i Stable	CUPSAT
81	46	27

iStable : http://predictor.nchu.edu.tw/iStable/

CUPSAT : https://cupsat.brenda-enzymes.org/resultall.jsp

Biophysical Property Analysis



Download Report	Uniprot Position	PDB/Model Position	Residue Wildtype	Residue Mutant	Missense3D Prediction	Structural damage predicted
	158	158	ARG	GLN	Damaging	Buried charge replaced; Buried salt bridge breakage
	158	158	ARG	TRP	Damaging	Buried H-bond breakage; Buried salt bridge breakage; Cavity altered; Buried / exposed switch
	158	158	ARG	PRO	Damaging	Buried Pro introduced; Buried H-bond breakage; Buried salt bridge breakage

Showing 1 to 3 of 3 entries Previous 1 Next

Figure 11: Missense 3D

Missense 3D : http://missense3d.bc.ic.ac.uk:8080/result

GENE	Variants
PAH	12

Variant Effect Analysis

2.5 Conclusion

Finally we can conlude that:

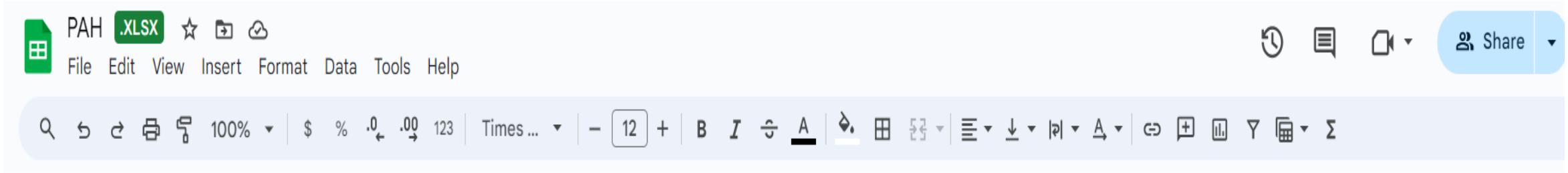
- Based on TANGO, the mutation decreases the aggregation tendency of your protein.
- Based on WALTZ, the mutation does not affect the amyloid propensity of your protein.
- Based on LIMBO, the mutation does not affect the chaperone binding tendency of your protein.
- Based on FoldX, the mutation severely reduces the protein stability.

Figure 12: SNP Effect 4.0

SNP Effect 4.0 (Home page): https://snpeffect.switchlab.org/

GENE	VARIANTS
PAH	3

SNP Effect 4.0



	A	B	C	D	E
1		TANGO	WALTZ	LIMBO	FoldX
2	Y154N	does not affect the aggregation tendency of your protein	does not affect the amyloid propensity of your protein	does not affect the chaperone binding tendency of your protein	reduces the protein stability
3	R158W	does not affect the aggregation tendency of your protein	does not affect the amyloid propensity of your protein	does not affect the chaperone binding tendency of your protein	reduces the protein stability
4	R158P	does not affect the aggregation tendency of your protein	does not affect the amyloid propensity of your protein	does not affect the chaperone binding tendency of your protein	severly reduces the protein stability
5	L194P	does not affect the aggregation tendency of your protein	does not affect the amyloid propensity of your protein	does not affect the chaperone binding tendency of your protein	severly reduces the protein stability
6	C203W	does not affect the aggregation tendency of your protein	decreases the amyloid propensity of your protein	does not affect the chaperone binding tendency of your protein	severly reduces the protein stability
7	P225R	does not affect the aggregation tendency of your protein	does not affect the amyloid propensity of your protein	does not affect the chaperone binding tendency of your protein	severly reduces the protein stability
8	L248R	does not affect the aggregation tendency of your protein	does not affect the amyloid propensity of your protein	does not affect the chaperone binding tendency of your protein	no effect on the protein stability
9	R261P	does not affect the aggregation tendency of your protein	does not affect the amyloid propensity of your protein	does not affect the chaperone binding tendency of your protein	reduces the protein stability
10	Y268H	does not affect the aggregation tendency of your protein	does not affect the amyloid propensity of your protein	does not affect the chaperone binding tendency of your protein	reduces the protein stability
11	P281L	does not affect the aggregation tendency of your protein	does not affect the amyloid propensity of your protein	does not affect the chaperone binding tendency of your protein	reduces the protein stability
12	G289R	does not affect the aggregation tendency of your protein	does not affect the amyloid propensity of your protein	decreases the chaperone binding tendency of your protein	severly reduces the protein stability
13	Y386D	decreases the aggregation tendency of your protein	does not affect the amyloid propensity of your protein	does not affect the chaperone binding tendency of your protein	severly reduces the protein stability
14					

Energy Minimization (swiss pdbv)

Variant	E-value(Kj/MOL)
Native	-17511.434
C203W	-17528.086
G289R	-17052.732
Y386D	-17883.717

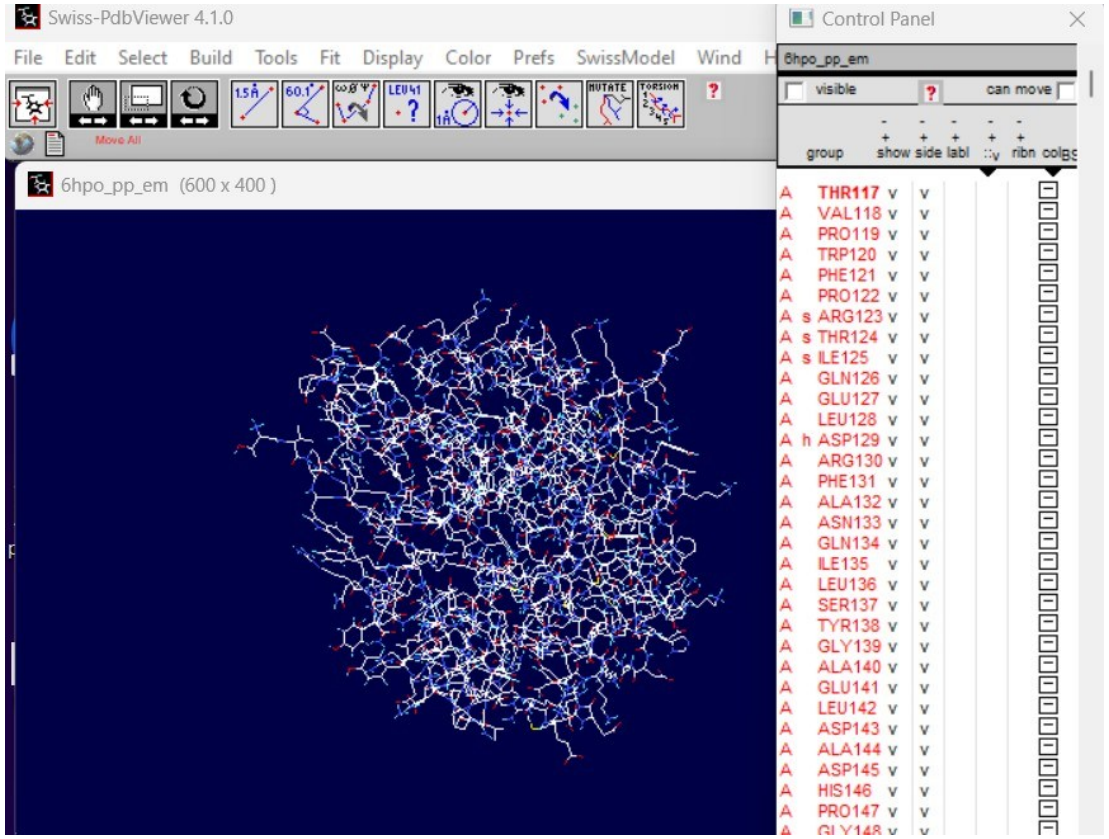


Figure13.1: Native

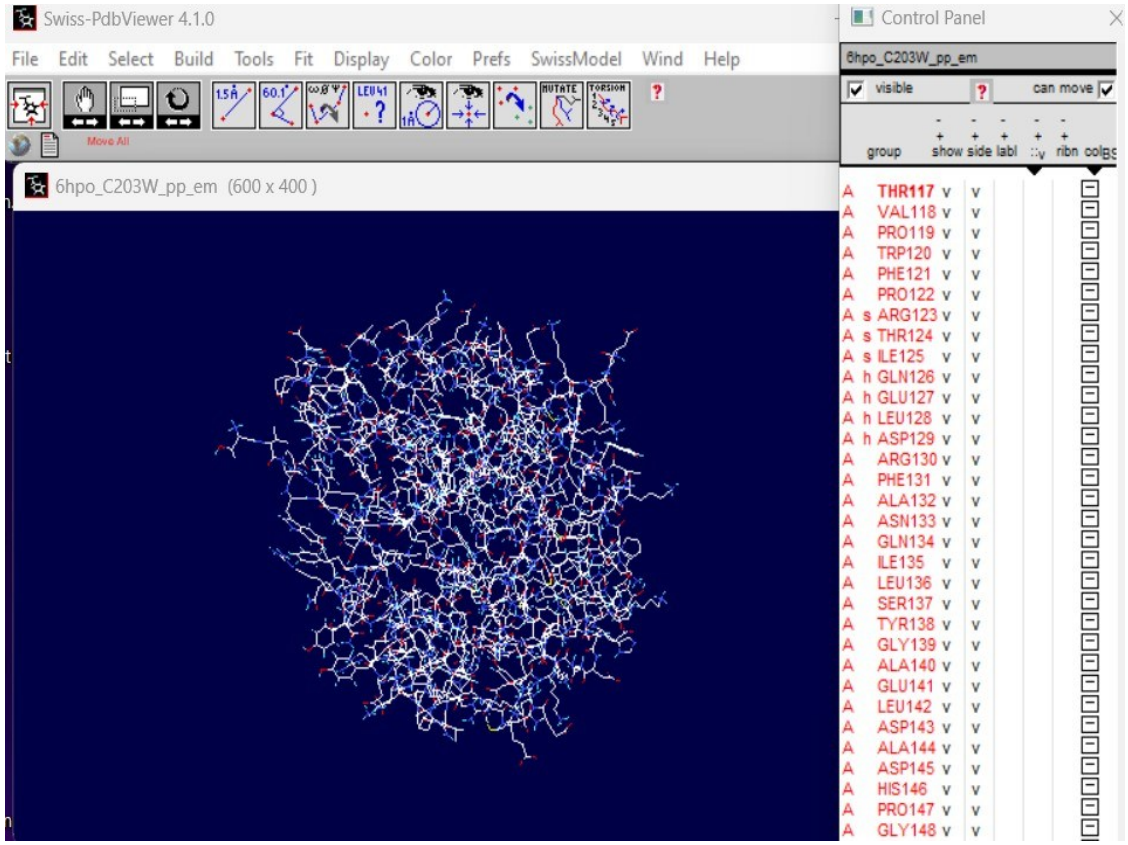


Figure 13.2: C203W

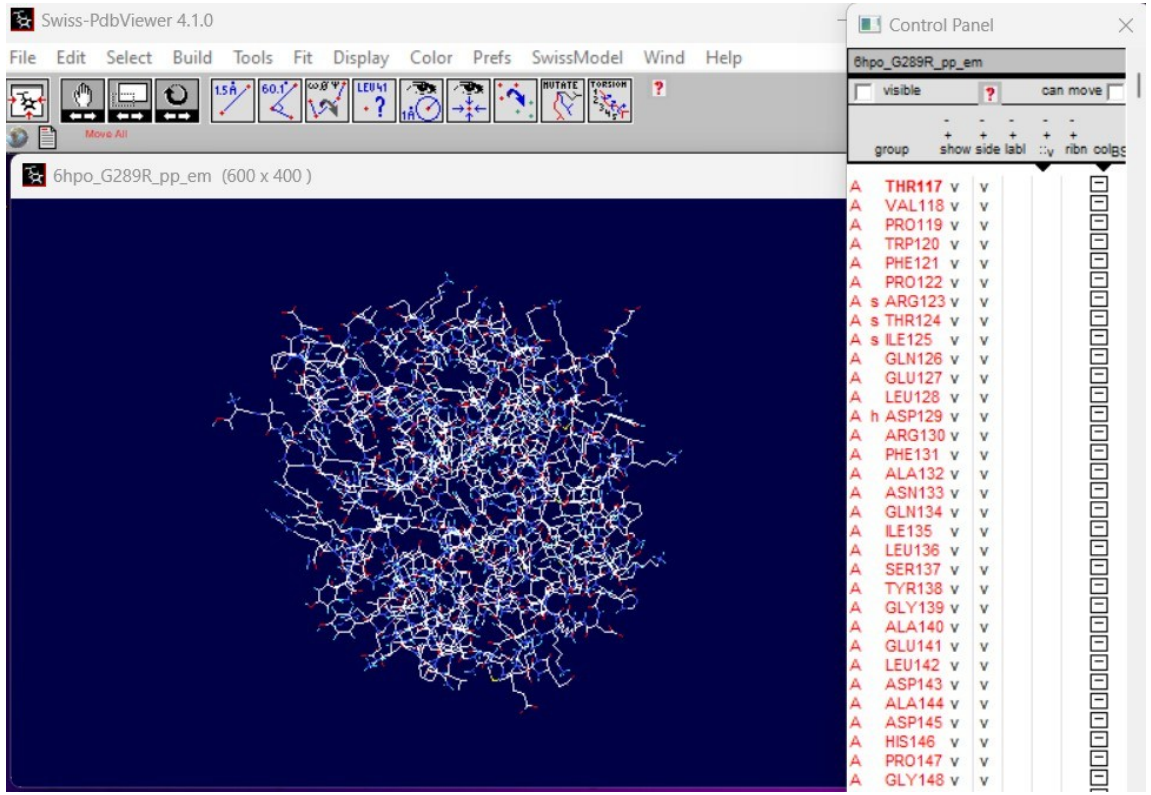


Figure 13.3: G289R

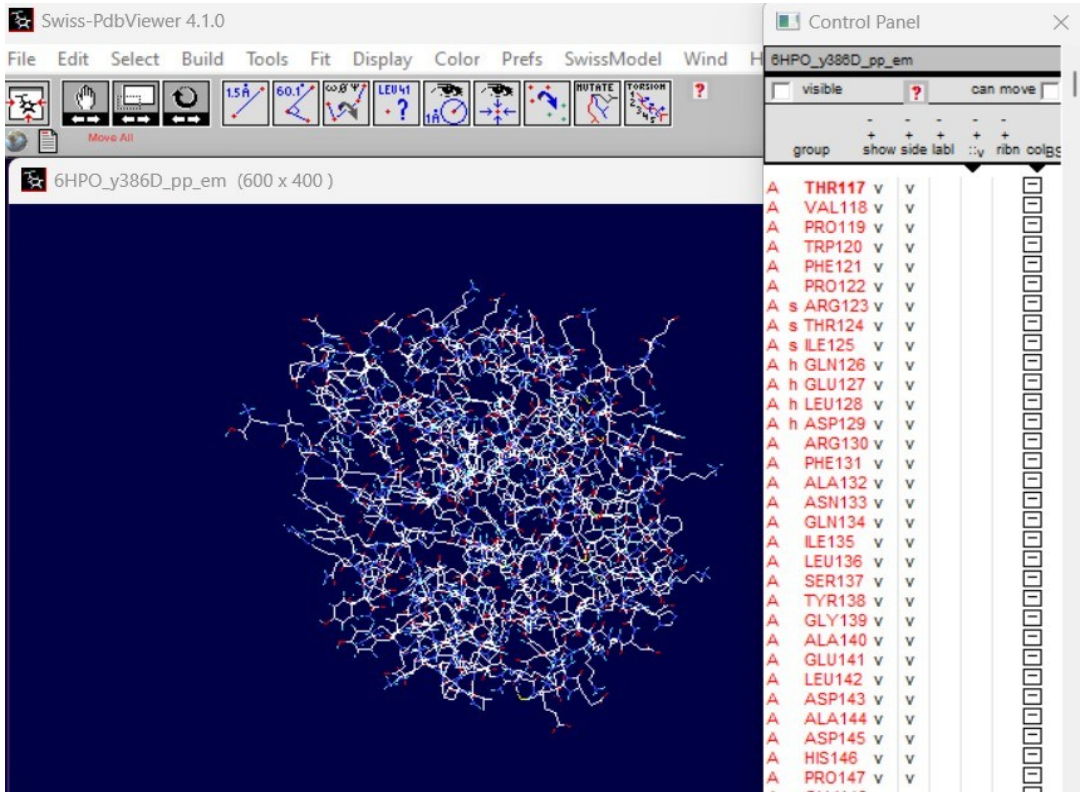


Figure 13.4: Y386D

Molecular Docking with Sapropterin:

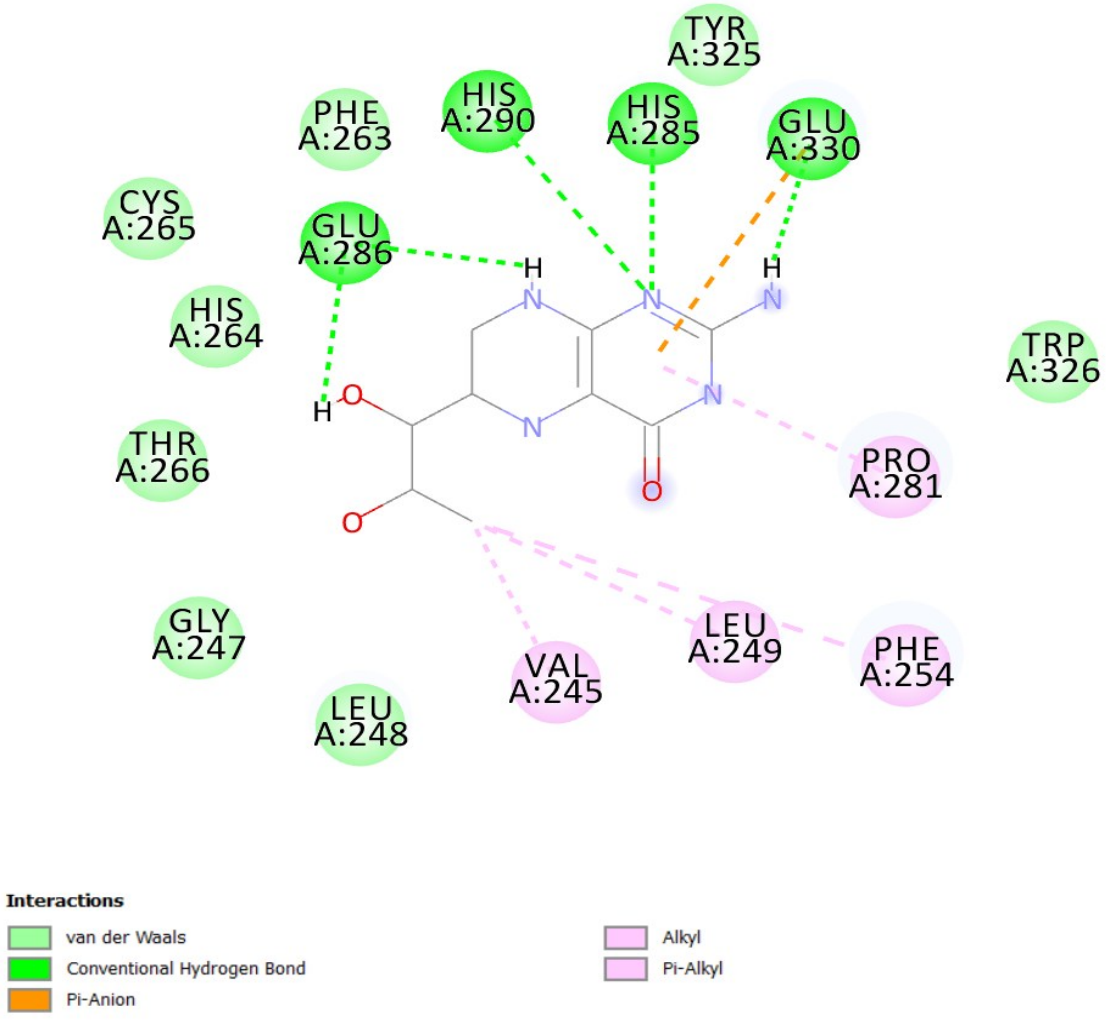


Figure14.1: **Native (-6.29kJ/mol)**

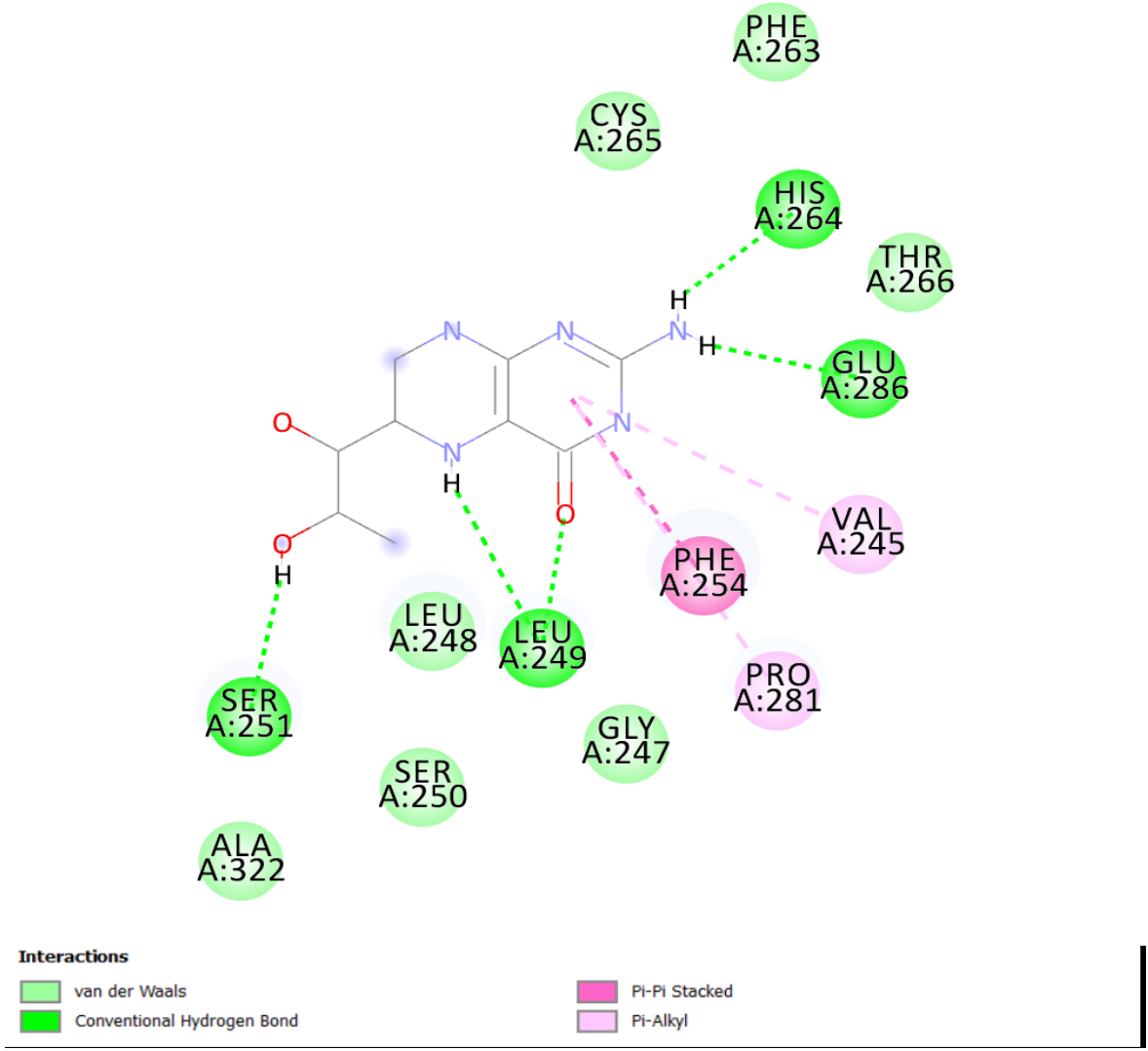


Figure14.2: **C203W (-6.31kJ/mol)**

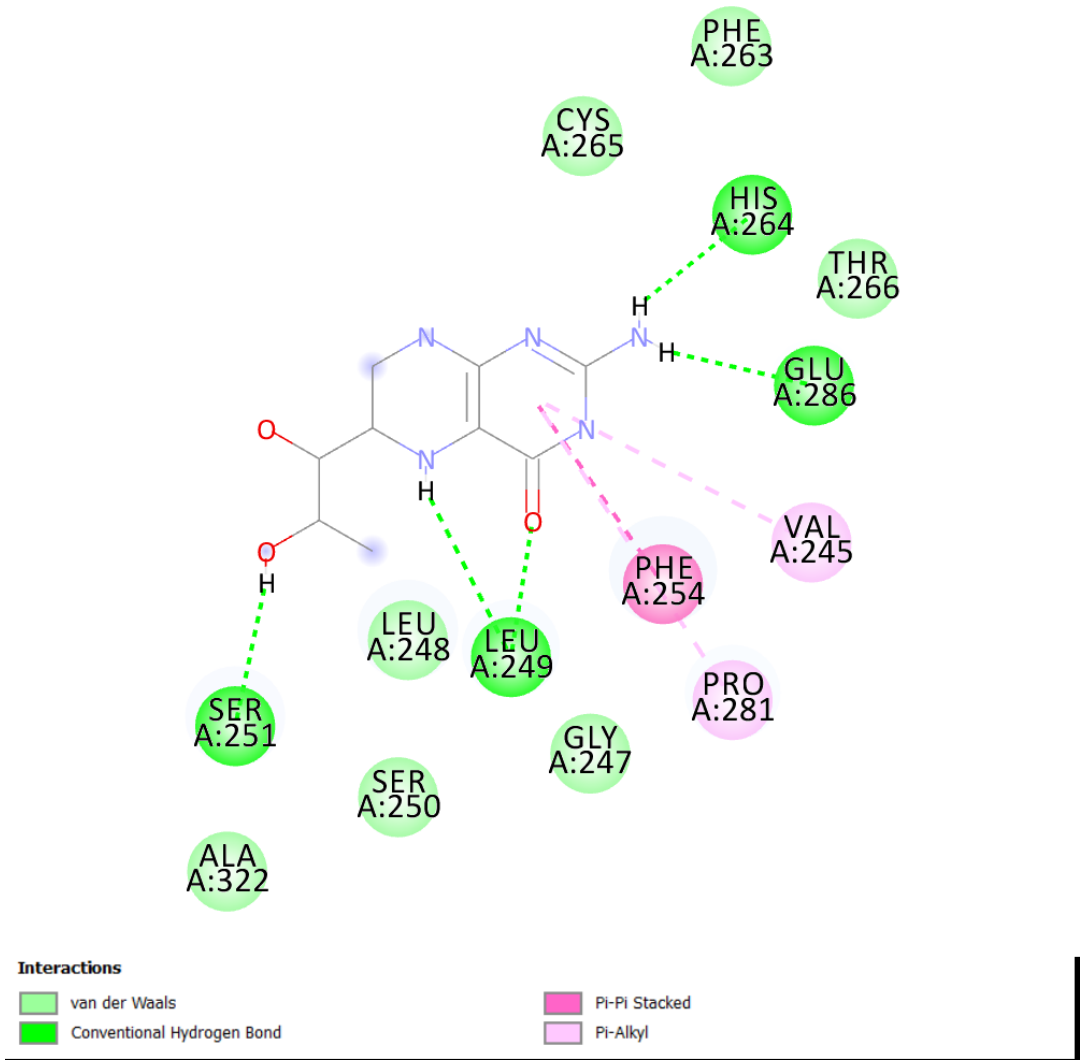


Figure14.3: **G289R (-6.72kJ/mol)**

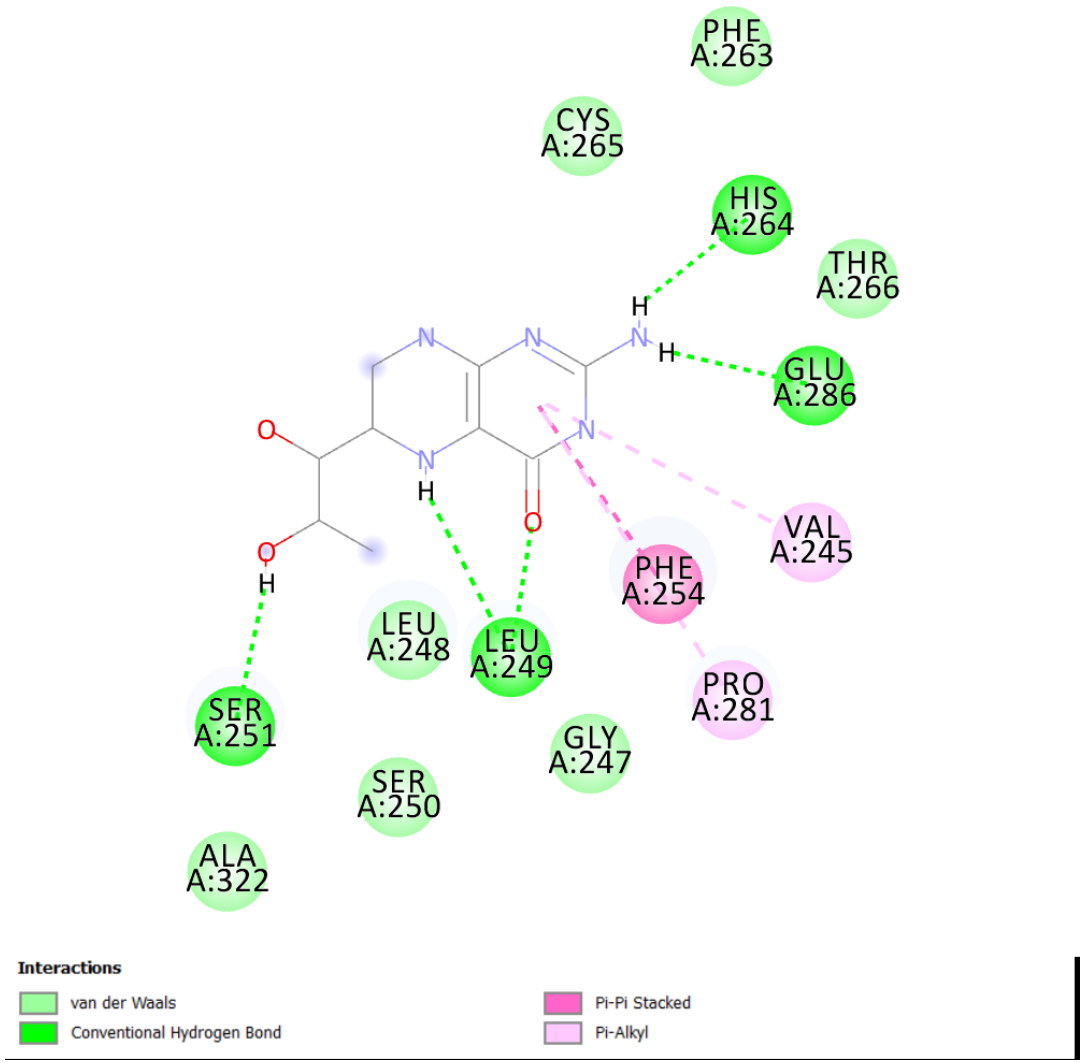


Figure14.4: **Y386D (-6.35kJ/mol)**

Conclusion

- Y386D, G289R, and C203W were determined to be the most harmful version by the use of various computational techniques.

- In order to comprehend how Sapropterin dihydrochloride interacts with both native and the highest and lowest scoring mutant kinds, Docking was performed.

- Research in the future can help produce customized treatment.

Outcome of this study:

- Identified harmful variations that may lead to Phenylketonuria.
- Examined drug (**sapropterin)** interaction with the protein.

Future Study:

- Dynamics can be carried out to understand the structural flexibility and entropic effects of the drug–target recognition and binding.

Acknowledgment

- First I extend my gratitude to the **Almighty**
- I thank SRIHER for giving an opportunity to present my work
- I greatly indebted to my respected guide, **Dr. R. Magesh**, Assistant Professor, Department of Biotechnology, for his exemplary guidance, vital encouragement, and motivation to come up in this project.
- I express my heartfelt thanks to **Dr. Sudha Warrier**, Principal of Biomedical Sciences and Head of Department of Biotechnology, SRIHER, for her valuable support for the completion of our project work.
- I express my heartfelt thanks to **Dr. Charles Emmanuel Jebaraj**, Professor, Department of Biotechnology, for his valuable support and providing all the necessary facilities and resources for the completion of our project work.
- I also would like to thank **Dr. Sumathy Joseph**, Professor & Head, Department of Biomedical Sciences, and all the faculty members of the department for their constant support and encouragement.
- It is my pleasure to express my profound thanks to Ms. Madhana Priya. N, Research scholar for her constant support and help in successful completion of our project work.

References

- Scriver, Charles R. 2007. “The PAH Gene, Phenylketonuria, and a Paradigm Shift.” Human Mutation 28 (9): 831–45.
- Dubois, Eline A., and Adam F. Cohen. 2010. “Sapropterin.” British Journal of Clinical Pharmacology 69 (6): 576–77.
- Flickr, Follow us on. n.d. “How Do Health Care Providers Diagnose Phenylketonuria (PKU)?” Https://www.nichd.nih.gov/. Accessed December 18, 2023. https://www.nichd.nih.gov/health/topics/pku/conditioninfo/diagnosed.
- Pietz, Joachim. 1998. “Neurological Aspects of Adult Phenylketonuria.” Current Opinion in Neurology 11 (6): 679.
- “Psychiatric Symptoms and Disorders in Phenylketonuria.” 2010. Molecular Genetics and Metabolism 99 (January): S59–63.
- Rocha, Júlio César, Kirsten K. Ahring, Heather Bausell, Deborah A. Bilder, Cary O. Harding, Anita Inwood, Nicola Longo, et al. 2023. “Expert Consensus on the Long-Term Effectiveness of Medical Nutrition Therapy and Its Impact on the Outcomes of Adults with Phenylketonuria.” Nutrients 15 (18). https://doi.org/10.3390/nu15183940.
- Vernon, H. J., C. B. Koerner, M. R. Johnson, A. Bergner, and A. Hamosh. 2010. “Introduction of Sapropterin Dihydrochloride as Standard of Care in Patients with Phenylketonuria.” *Molecular Genetics and Metabolism* 100 (3): 229–33.