Chapter – 4

RESULT

Chapter – 4

RESULT

4.1. Prevalence of Glucose-6-Phosphate Dehydrogenase deficiency

All together 2310 subjects belonging to Proto-Australoid and Mongoloid population were screened. A total of 1188 subjects from the Proto-Australoid population were screened, among which 99 (8.3%) were found to be G6PD deficient. Among the Mongoloid population, 1122 subjects were screened, out of which 45 (4.01%) were found to be G6PD deficient. All the deficient subjects identified in our study were asymptomatic. The district wise findings of G6PD screening has been presented in Table 15. Both Udalguri and Kokrajhar districts showed a G6PD deficiency prevalence of 6.6%, 6.03% in Baksa and 5.6% in Chirang. Among the total deficient subjects, 66 were severely deficient and 78 were intermediates. The overall prevalence of G6PD deficiency in the study area was found to be 6.2% as depicted in Figure 7. Higher prevalence of the deficiency was found in the Proto-Australoid population.

Table 15. Result of G6PD screening in four districts of BTR, Assam.

District	Race	Individuals screened	G6PD deficient	G6PD normal
Udalguri	Proto-Australoid	301	32	269
	Mongoloid	290	07	283
Baksa	Proto-Australoid	288	23	265
	Mongoloid	295	12	283
Chirang	Proto-Australoid	301	13	288
	Mongoloid	246	18	228
Kokrajhar	Proto-Australoid	298	31	267
	Mongoloid	291	08	283

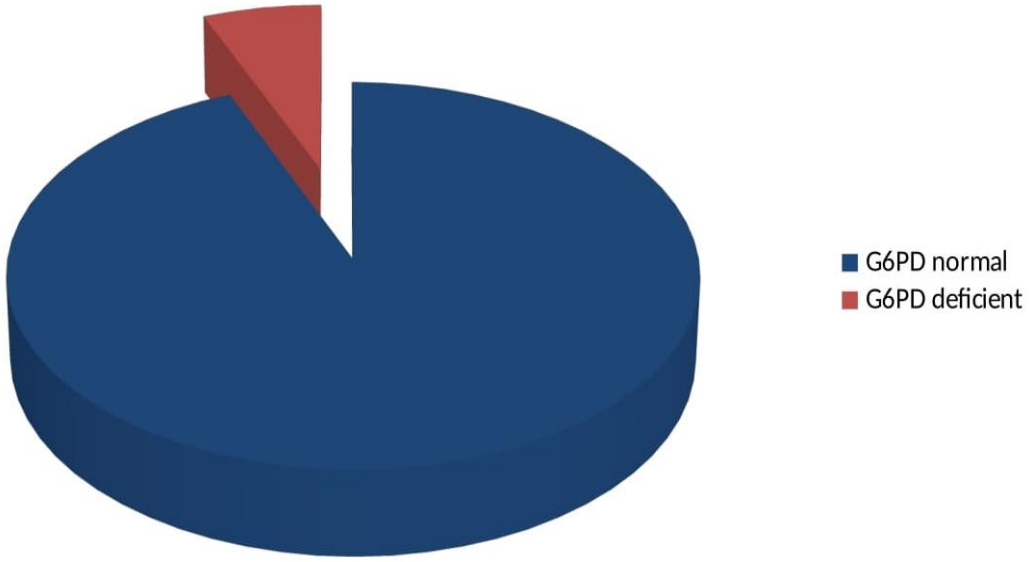


Figure 7. G6PD status of the study subjects.

4.2. Screening for abnormal haemoglobin variants by Complete Blood Count and Haemoglobin-typing

Based on the Hb% detected during screening, the subjects with low Hb% (anaemic) were selected for CBC (Appendix – III). Out of 2310 individuals, 1436 were found anaemic. All the anaemic samples were subjected to Hb-typing. An overall prevalence of 42.7% of haemoglobinopathies was found. The details of Hb-typing results are presented in Table 16. It was observed that 944 samples carried the normal Hb type. HbS variant was observed in 473 samples, constributing to about 32.9% of the total anaemic cases, out of which 417 were carriers and 56 were disease cases. All cases of HbS were from the tea tribe (Proto-Australoids). A total of 506 samples (35.2% of anaemic cases) were observed with presence of HbE, 342 were HbE trait and 164 were HbE disease. Among all HbE, four were from the tea tribe and others were from Mongoloids. β–thalassaemia trait was observed in three samples from the tea tribe population. Two samples from tea tribe population were identified to have compound heterozygous condition for HbS and HbE. Another one sample from the tea tribe was found to have HbE - β–thalassemia trait. One case was detected with both G6PD deficiency and HbE haemoglobinopathy. This case was from Mongoloid population. We also found 86 cases showing high HbF values (Hereditary

persistence of foetal Hb). The Hb-typing results of selected cases are depicted in Appendices IV – XI.

Table 16. Results of Hb-typing of anaemic subjects.

Hb – typing	1436
Normal	448
HbS trait	417
HbS disease	56
HbE trait	342
HbE disease	164
HbSE trait	02
β–thalassemia trait	06
HbE - β–thalassemia trait	01
Hereditary persistence of foetal Hb	86

4.3. Statistical analysis of results of Haemoglobin-typing

In a normal adult, blood contains 95-98% of adult haemoglobin (HbA_0) and 2-3% of a smaller component of adult haemoglobin (HbA_2). About 1-2% of foetal haemoglobin (HbF) is also present. The Mean ± SD of the Hb components in different categories of Hb disorders are presented in Table 17. In normal individuals, the mean ± SD value of HbA_0 was observed to be within the normal range (96.34 ± 7.2). In HbE trait and HbS trait lower values of HbA_0 were observed, 66.48 ± 9.5 and 60.81 ± 4.9, respectively. Very low values of HbA_0 were seen in HbE (5.19 ± 3.9) and HbS disease (0.46 ± 0.99) cases. Normal levels of HbA_2 were observed in all cases. High HbF values were observed in both HbE and HbS disease cases. In HbE disease, the Mean ± SD value of HbF was 4.06 ± 3.4, which was as high as 19.93 ± 6.6 in HbS disease cases.

Table 17. Mean ± SD of Hb types in different categories of haemoglobinopathies.

Category of Hb disorder	Parameters (Mean ± SD)				
	HbA_0	**HbA_2**	**HbE**	**HbS**	**HbF**
Normal (n=448)	96.34 ± 7.2	2.77 ± 0.6	0	0	0.19 ± 0.5
HbE trait (n=342)	66.48 ± 9.5	0.98 ± 1.6	31.19 ± 9.8	0	0.66 ± 1.2
HbE disease (n=164)	5.19 ± 3.9	1.16 ± 2.0	89.04 ± 4.2	0	4.06 ± 3.4
HbS trait (n=417)	60.81 ± 4.9	3.09 ± 0.54	0	34.8 ± 7.7	0.39 ± 0.9
HbS disease (n=56)	0.46 ± 0.99	2.5 ± 1.3	0	77.1 ± 6.3	19.93 ± 6.6

4.4. Clinical evaluation of Glucose-6-Phosphate Dehydrogenase deficient subjects

The RBC indices include Hb (g/dL), RBC count (Million/Cumm), MCV (fL), MCH (pg), MCHC (g/dL), PCV (%) and RDW (%). The WBC indices are TLC (Cells/Cumm), LYM (%), Monocytes (%), Eosinophils (%), Basophils (%). The platelet counts are represented in 10^3 µl. The statistical analyses of the RBC indices along with platelets are presented in Table 18. Slightly lower values of Hb, RBC, MCH, MCHC and PCV were seen in G6PD deficient subjects compared to the normal subjects. A slightly high RDW value (15.98 ± 3.07) was recorded in the G6PD deficient subjects, which was 15.56 ± 2.69 in the normal subjects. Platelet count was lower in the G6PD deficient subjects (199.27 ± 70.06) compared to the normal ones (205.82 ± 66.64). Statistically significant positive correlations of Hb, RBC and MCHC were seen with G6PD, with Pearson correlation coefficient of 0.081, 0.067 and 0.069 respectively. Other parameters such as MCH, PCV and platelets showed positive correlation with G6PD, however, the correlation was not

found to be statistically significant. Negative correlation was observed with MCV and RDW, which was also not found to be statistically significant.

Table 18. Mean ± SD values of RBC parameters and platelets, and Pearson correlation coefficient with G6PD along with p-values.

Parameter	**Normal Mean ± SD (n=2166)**	**G6PD deficient Mean ± SD (n=144)**	**Pearson correlation Coefficient**	**P-value**
Hb (g/dL)	10.81 ± 2.56	10.59 ± 3.69	0.081	<0.05
RBC (Million/Cumm)	4.27 ± 0.82	4.16 ± 1.04	0.067	<0.05
MCV (fL)	84.45 ± 12.07	84.44 ± 11.99	-0.010	>0.05
MCH (pg)	25.38 ± 3.96	25.09 ± 3.89	0.012	>0.05
MCHC (g/dL)	29.90 ± 1.41	29.62 ± 1.63	0.069	<0.05
PCV (%)	35.8 ± 4.73	35.43 ± 5.17	0.065	>0.05
RDW (%)	15.56 ± 2.69	15.98 ± 3.07	-0.046	>0.05
Platelets (10^3 µl)	205.82 ± 66.64	199.27 ± 70.06	0.059	>0.05

Table 19 represents the WBC indices of the G6PD deficient and normal subjects. The values of the WBC indices, viz., leucocytes, lymphocytes and eosinophils were slightly higher in the deficient subjects, and similar values of monocytes were observed. Lymphocytes, monocytes and eosinophils were found to be negatively correlated with G6PD (Pearson correlation coefficients -0.050, -0.007 and -0.012 respectively); however, the correlation was not statistically significant. Leucocytes showed a positive correlation coefficient of 0.047 with G6PD, which was also not statistically significant.

Table 19. Mean ± SD values of WBC parameters and Pearson correlation coefficient with G6PD along with p-values.

Parameter	**Normal Mean ± SD (n=2166)**	**G6PD deficient Mean ± SD (n=144)**	**Pearson correlation Coefficient**	**P-value**

Leucocyte (Cells/Cumm)	8143.81 ± 2450.507	8161.05 ± 2441.52	0.047	>0.05
Lymphocytes (%)	37.3 ± 9.09	38.68 ± 10.148	-0.050	>0.05
Monocytes (%)	6.31 ± 1.33	6.36 ± 1.46	-0.007	>0.05
Eosinophils (%)	5.70 ± 3.28	5.96 ± 3.94	-0.012	>0.05
Basophils (%)	0	0	----	----

4.5. Association between Glucose-6-Phosphate Dehydrogenase deficiency and gender

The Chi-square statistic for testing the association between G6PD deficiency and gender is presented in Table 20. The results showed that 31.3% of the deficient cases were females with intermediate form of deficiency. The remaining deficient cases (68.7%) were males with either severe or intermediate form of deficiency. The G6PD values in the males were between 1.5 U/g Hb and 5.9 U/g Hb, while in females the values were between 2.2 U/g Hb and 6.0 U/g Hb.

Table 20. Cross tabulation between G6PD deficiency and gender.

G6PD status	**Gender**		**Total**
	Male	**Female**	
Deficient	68.7%	31.3%	100%
Normal	31.3%	68.7%	100%

4.6. Molecular analysis of Glucose-6-Phosphate Dehydrogenase deficient samples

4.6.1. Agarose gel electrophoresis of extracted DNA

The genomic DNA extracted from the deficient blood samples were subjected to agarose gel electrophoresis (0.8%) to confirm and check the quality of the extracted DNA. The images of extracted DNA run on 0.8% agarose is shown in Figure 8. In the figure, lane

M represents 1kb DNA ladder and lanes 1-12 is the extracted DNA of the G6PD deficient samples.

4.6.2. Quantification of DNA

The concentration and purity of the extracted DNA of 144 G6PD deficient samples was determined using the UV-VIS Spectrophotometer. The quantity of extracted DNA ranged between 157-322 µg/ml. The A_{260}/A_{280} ratio was found to be ≥ 1.8 in 115 samples, indicating pure extracted DNA. In 29 samples, the A_{260}/A_{280} ratio was found to be either 1.6 or 1.7, which indicates the presence of small amount of impurities (may be protein and RNA) in the DNA.

4.6.3. Polymerase Chain Reaction for amplification of Glucose-6-Phosphate Dehydrogenase gene

Amplification of the exonic region of *g6pd* gene was done using different primer sets for exons 3-12. Since the start codon of *g6pd* gene lies in the base 115 of exon 2, and no mutations have been reported from India in exons 1, 2 and 8, so the amplication was done by using primer sets for exons 3-7 and 9-12. The amplified products were then run on 1.8-2% agarose gel electrophoresis. The exon-wise PCR conditions and images of the PCR products run on 1.8-2% agarose are described as follows:

4.6.3.1. Amplification of exon 3

The target product for exon 3 was obtained by running the amplication process as follows: Initial denaturation and denaturation at 95°C for 5 minutes and 30 seconds, respectively, annealing at 60°C, extension and final extension at 72°C for 45 seconds and 10 minutes respectively. The amplified products of exon 3 are shown in Figure 9. In the figure, lane M represents 100bp DNA ladder and lanes 1-12 is the amplified products of exon 3 with product size of 352bp.

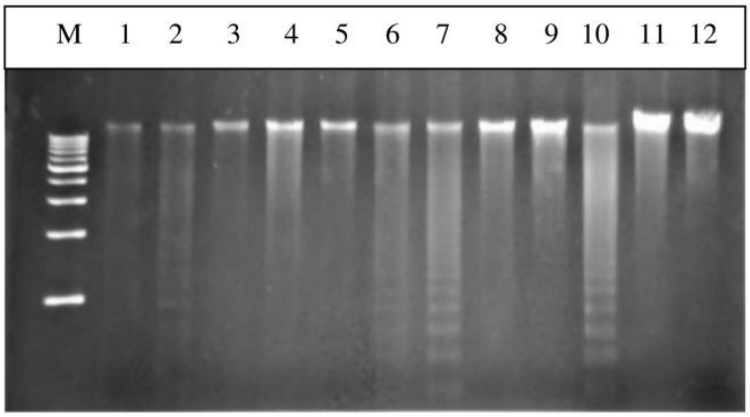


Figure 8. Images of genomic DNA run on 0.8% agarose gel electrophoresis. M → 1kb DNA ladder and lanes 1-12 → genomic DNA of the G6PD deficient samples.

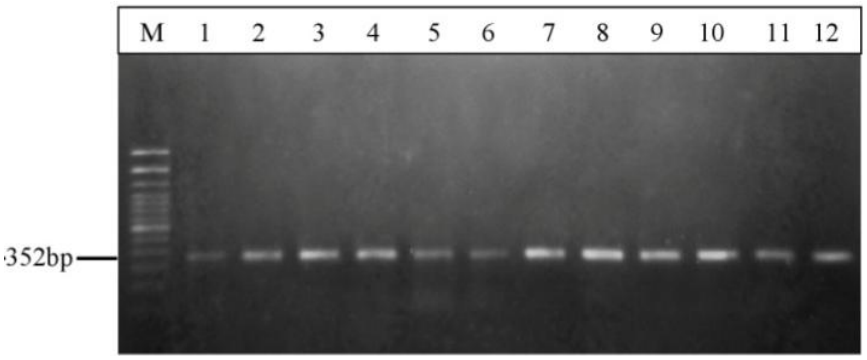


Figure 9. Agarose gel electrophoresis images of PCR products of exon 3. M → 100bp DNA ladder, lanes 1-12 → 352bp PCR products of exon 3.

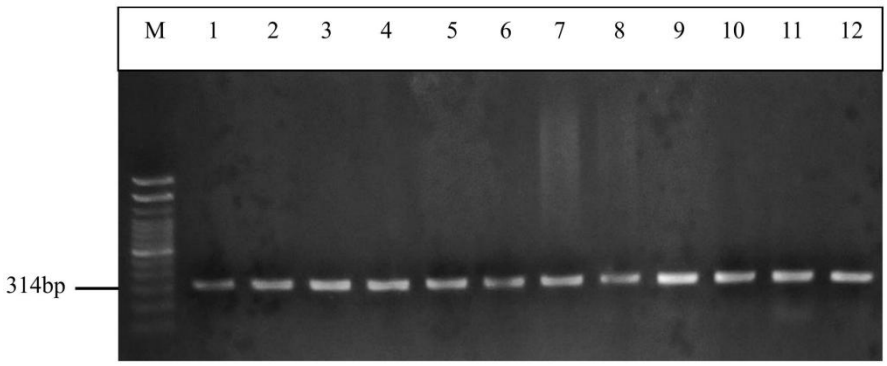


Figure 10. Agarose gel electrophoresis images of PCR products of exon 4. M → 100bp DNA ladder, lanes 1-12 → 314bp PCR products of exon 4.

4.6.3.2. Amplification of exon 4

Exon 4 was amplified by running the amplication process as follows: Initial denaturation and denaturation at 95°C for 5 minutes and 30 seconds, respectively, annealing at 64°C, extension and final extension at 72°C for 45 seconds and 10 minutes respectively. Figure 10. represents the agarose gel electrophoresis images of PCR products of exon 4. Lane M represents 100bp DNA ladder and lanes 1-12 are the amplified products of exon 4 having product size of 314bp.

4.6.3.3. Amplification of exons 4-5

Amplification of exons 4-5 having a product size of 701bp was obtained by performing PCR with the following conditions: Initial denaturation and denaturation at 95°C for 5 minutes and 30 seconds, respectively, annealing at 61°C, extension and final extension at 72°C for 45 seconds and 10 minutes respectively. The PCR products of exons 4-5 are presented in Figure 11. Lane M is the 100bp DNA ladder and lanes 1-12 represent the 701bp PCR products of exons 4-5.

4.6.3.4. Amplification of exons 6-7

Exons 6-7 was amplified by running the amplification process with the conditions as initial denaturation and denaturation at 95°C for 5 minutes and 30 seconds, respectively, annealing at 66.5°C, extension and final extension at 72°C for 45 seconds and 10 minutes respectively. The agarose gel electrophoresis images of PCR products of exons 6-7 are presented in Figure 12. Lane M is the 100bp DNA ladder and lanes 1-12 represent the 545bp PCR products of exons 6-7.

4.6.3.5. Amplification of exon 9

PCR was done for obtaining the amplified products of exon 9 by performing PCR with the following conditions: Initial denaturation and denaturation at 95°C for 5 minutes and 30 seconds, respectively, annealing at 61.6°C, extension and final extension at 72°C for

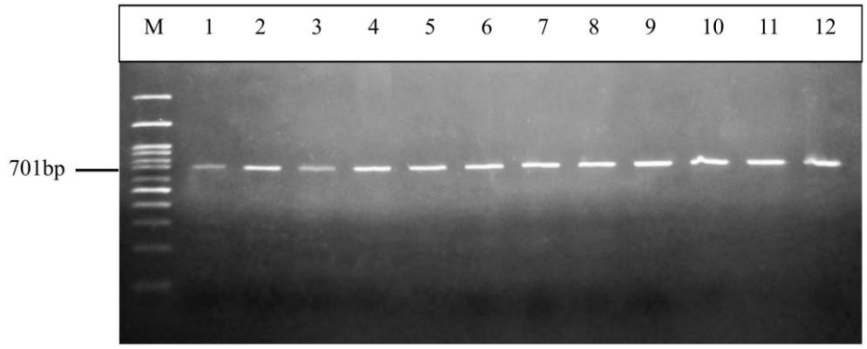


Figure 11. Agarose gel electrophoresis images of PCR products of exons 4-5. M → DNA 100bp ladder, lanes 1-12 → 701bp PCR products of exons 4-5.

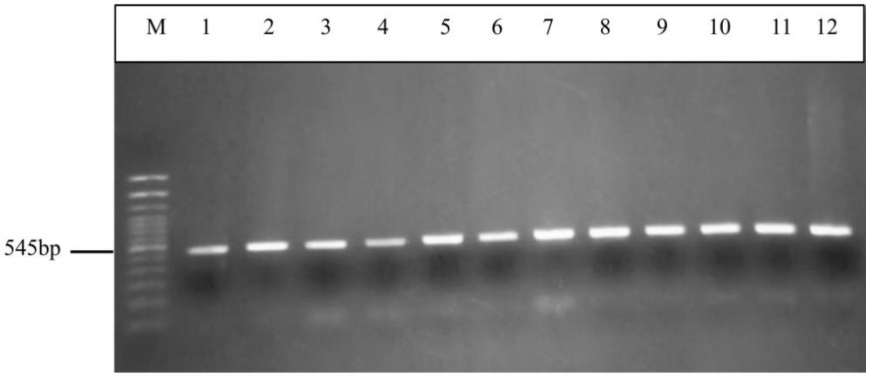


Figure 12. Agarose gel electrophoresis images of PCR products of exons 6-7. M → 100bp DNA ladder, lanes 1-12 → 545bp PCR products of exons 6-7.

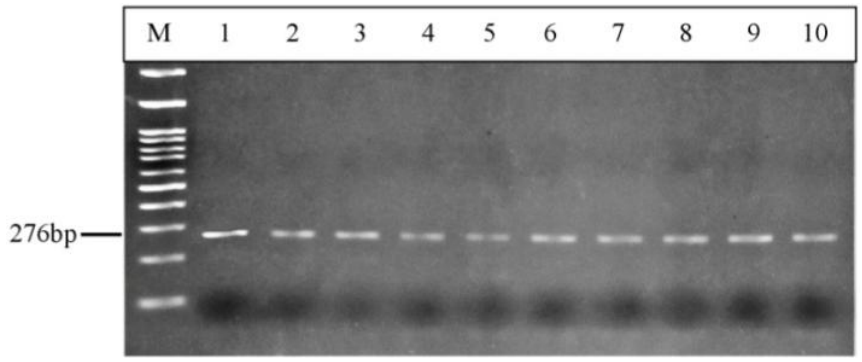


Figure 13. Agarose gel electrophoresis images of PCR products of exon 9. M → 100bp DNA ladder, lanes 1-12 → 276bp PCR products of exon 9.

45 seconds and 10 minutes respectively. Figure 13 shows the images of PCR products of exon 9 run in 2% agarose gel elctrophoresis. Lane M is the 100bp DNA ladder and lanes 1-12 represent the 276bp PCR products of exon 9.

4.6.3.6. Amplification of exon 10

The PCR conditions were set as initial denaturation and denaturation at 95°C for 5 minutes and 30 seconds, respectively, annealing at 63°C, extension and final extension at 72°C for 45 seconds and 10 minutes respectively, for obtaining the amplified products of exon 10. The agarose gel electrophoresis images of PCR products of exon 10 are presented in Figure 14. Lane M represents 100bp DNA ladder and lanes 1-12 represent the 342bp PCR products of exon 10.

4.6.3.7. Amplification of exon 11

The PCR conditions for amplification of exon 11 were set as follows: Initial denaturation and denaturation were set at 95°C for 5 minutes and 30 seconds, respectively, annealing was set at 68°C, extension and final extension was done at 72°C for 45 seconds and 10 minutes respectively. The 214bp PCR products of exon 11 are presented in Figure 15. In the figure, Lane M represents the 100bp DNA ladder and lanes 1-12 represent the 214bp PCR products of exon 11.

4.6.3.8. Amplification of exon 12

Amplified products of exon 12 were obtained by carrying out PCR with the following conditions: Initial denaturation and denaturation at 95°C for 5 minutes and 30 seconds, respectively, annealing at 62°C, extension and final extension at 72°C for 45 seconds and 10 minutes respectively. The PCR products of exon 12 are presented in Figure 16. Lane M is the 100bp DNA ladder and lanes 1-12 represent the 227bp PCR products of exon 12.

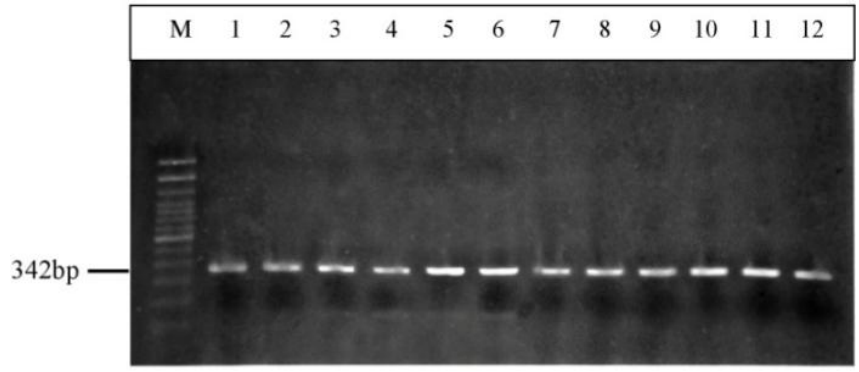


Figure 14. Agarose gel electrophoresis images of PCR products of exon 10. M → 100bp DNA ladder, lanes 1-12 → 342bp PCR products of exon 10.

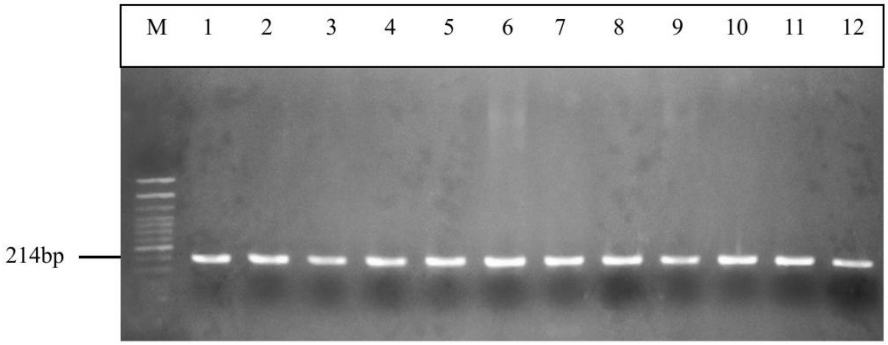


Figure 15. Agarose gel electrophoresis images of PCR products of exon 11. M → 100bp DNA ladder, lanes 1-12 → 214bp PCR products of exon 11.

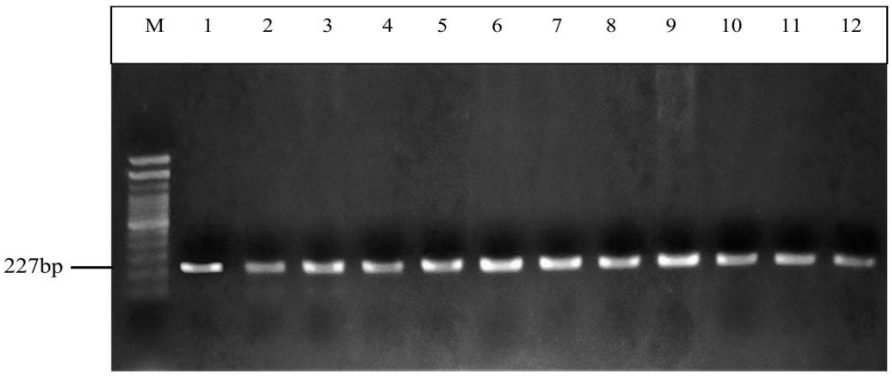


Figure 16. Agarose gel electrophoresis images of PCR products of exon 12. M → 100bp DNA ladder, lanes 1-12 → 227bp PCR products of exon 12.

4.6.4. Restriction Fragment Length Polymorphism

Following amplification of the selected exons, the PCR products were treated with specific restriction endonucleases for detection of the respective mutations via Restriction Fragment Length Polymorphism (RFLP) analysis. The detailed observations of RFLP are described as follows:

4.6.4.1. Detection of Orissa variant (131C>G) in exon 3 by HaeIII restriction digestion analysis

HaeIII recognizes and cleaves the DNA at 5′GG/CC3′. The mutation 131C>G eliminates a cut site for HaeIII in exon 3. For detection of this mutation, all 144 G6PD deficient samples were subjected to digestion of PCR product (exon 3) by restriction endonuclease HaeIII. The samples were digested with HaeIII at 37°C for 1-2 hours. A total of 92 samples were observed to have this mutation. Figure 17 represents the samples digested with HaeIII. Lane M represents the 100bp DNA ladder, lane 1 has the undigested 352bp PCR product of exon 3. From the figure, it is evident that in lanes 2, 3, 5, 6, 10, 11 and 12, a cut site is eliminated, indicating the presence of 131C>G mutation. In lanes 4, 7, 8 and 9, the mutation was absent.

4.6.4.2. Detection of Namoru variant (208T>C) in exon 4 by NlaIII restriction digestion analysis

Normal exon 4 does not have any cut site for NlaIII, which is 5′CATG/3′, but the mutation 208T>C generates a cut site for NlaIII in exon 4. All 144 G6PD deficient samples were examined for the presence of the variant G6PD Namoru (208T>C) in exon 4 by digesting with restriction endonuclease NlaIII at 37°C for 1 - 2 hours. Figure 18 shows the agarose gel electrophoresis images of PCR products (exon 4) treated with NlaIII. Lane M represents 100bp DNA ladder, lane 1 represents untreated PCR product of exon 4. No cuts were observed in any of the samples (lanes 2-12), and hence, the mutation 208T>C was absent in the samples.

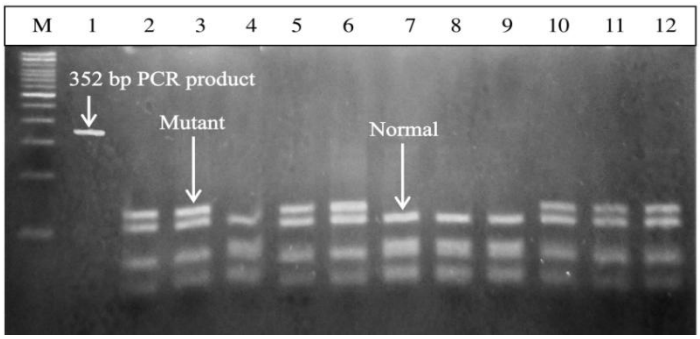


Figure 17. Agarose gel electrophoresis of PCR products digested with HaeIII (Orissa 131C>G). M → 100bp DNA ladder, lane 1 → undigested PCR product of exon 3, lanes 4, lanes 2, 3, 5, 6, 10-12 → 131C>G mutation is present, lanes 7-9 → 131C>G mutation is absent.

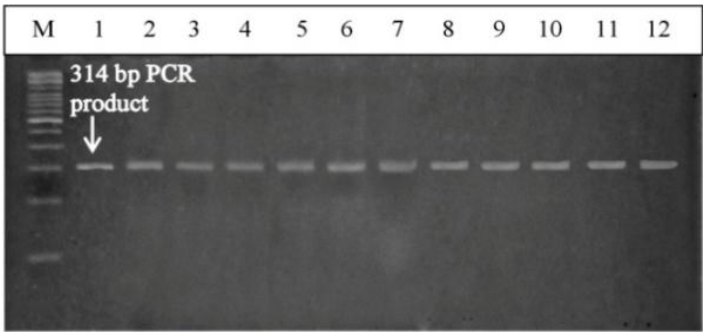


Figure 18. Agarose gel electrophoresis of PCR products digested with NlaIII (Namoru 208T>C). M → 100bp DNA ladder, lane 1 → undigested PCR product of exon 4, lanes 2-12 → mutation 208T>C was absent in all samples.

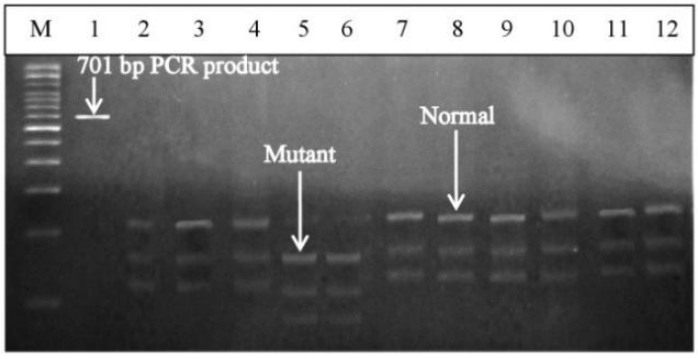


Figure 19. Agarose gel electrophoresis of PCR products digested with FokI (A^+ 376A>G). M → 100bp DNA ladder, lane 1 → undigested PCR product of exons 4-5, lanes 5, 6 → 376 A>G mutation is present, lanes 2-4, 7-12 → 376 A>G mutation is absent.

4.6.4.3. Detection of A^+ variant (376A>G) in exon 5 by FokI restriction digestion analysis

For detection of 376A>G mutation in exon 5, the PCR products of exon 5 of the 144 G6PD deficient samples were subjected to digestion by restriction endonuclease FokI at 37°C for 1-2 hours. FokI recognizes 5′-GGATG-3′ and creates two cuts, one in the top strand at a distance of nine nucleotides from the 3′ end of the recognition site and another cut in the bottom strand at a distance of 13 nucleotides from the 5′ end of the recognition site. This mutation (376A>G) creates an additional cut site for FokI in exon 5. From Figure 19, it can be identified that in lanes 5 and 6, an additional cut site was generated. Whereas, in lanes 2, 3, 4, 7, 8, 9, 10, 11 and 12 no additional cut was observed. Lanes M and 1 represents 100bp DNA ladder and untreated PCR product of exon 5, respectively. A total of 17 samples were found to have this mutation.

4.6.4.4. Detection of A^{-202} variant (202G>A) in exon 4 by NlaIII restriction digestion analysis

Another mutation 202G>A (A^{-202} variant) co-exists with 376A>G (A^+ variant), which abolishes a cut site for restriction edonuclease NlaIII (5′CATG/3′) in exon 4. The 17 samples which were found to have the A^+ variant (376A>G) were examined for the presence of this additional mutation by treating with NlaIII at 37°C for 1-2 hours. NlaIII treated products are presented in Figure 20, where lane M represents 100bp DNA ladder and lane 1 represents untreated PCR product of exon 4. No cuts were observed in the samples (lanes 2-12). Hence, the additional mutation 202G>A was not present in the samples having 376A>G mutation.

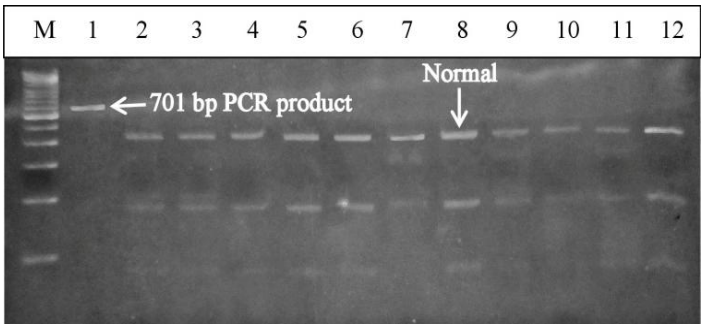


Figure 20. Agarose gel electrophoresis of PCR products digested with NlaIII (A^{-202} G>A). M → 100bp DNA ladder, lane 1 → undigested PCR product of exons 4-5, lanes 2-12 → mutation 202G>A was absent in all samples.

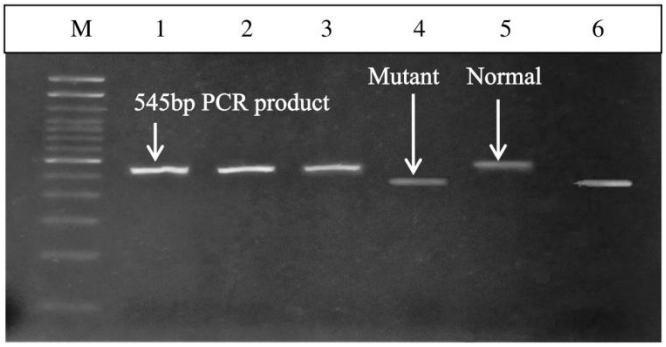


Figure 21. Agarose gel electrophoresis of PCR products digested with HindIII (Mahidol 487G>A). M → 100bp DNA ladder, lane 1 → undigested PCR product of exons 6-7, lanes 4, 6 → 487G>A mutation is present, lanes 2, 3, 5 → 487G>A mutation is absent.

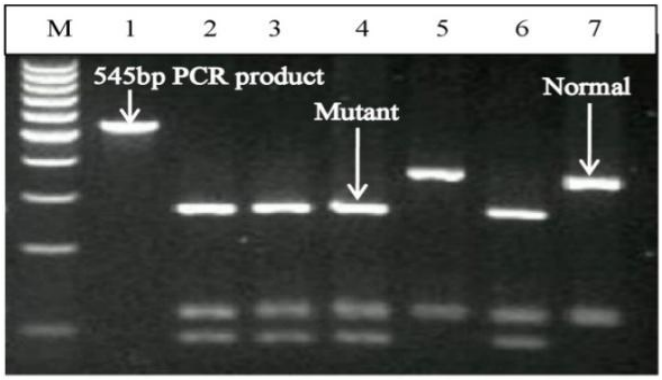


Figure 22. Agarose gel electrophoresis of PCR products digested with MboII (Mediterranean 563C>T). M → 100bp DNA ladder, lane 1 → undigested PCR product of exons 6-7, lanes 2-4 & 6 → 563C>T mutation is present, lanes 5 & 7 → 563C>T mutation is absent.

4.6.4.5. Detection of Mahidol variant (487G>A) in exon 6 by HindIII restriction digestion analysis

Mutation 487G>A in exon 6, was examined by digestion of PCR products of exon 6 with restriction endonuclease HindIII. HindIII cleaves the DNA at 5′A/AGCTT3′. The samples were kept for digestion with HindIII in 37°C for 1-2 hours. Exon 6 does not have a cut site for HindIII in its normal form, but the mutation 487G>A produces a cut site for HindIII in the exon. The agarose gel electrophoresis images of PCR products after digestion with HindIII are presented in Figure 21. In the figure, lane M represents the 100bp DNA ladder, lane 1 represents uncut PCR product of exons 6-7. Cuts were generated in lanes 2 and 3, indicating the presence of the mutation. We found two samples with this mutation.

4.6.4.6. Detection of Mediterranean variant (563C>T) in exon 6 by MboII restriction digestion analysis

All 144 G6PD deficient samples were checked for another mutation 563C>T in exon 6 via digestion of PCR product by restriction endonuclease MboII. This restriction endonuclease recognizes 5′GAAGA3′ and cuts the DNA eight nucleotides downstream from the recognition site. The samples were digested with MboII at 37°C for 1-2 hours. The mutation 563C>T creates an additional cut site for MboII in exon 6. Among 144 samples, 12 samples showed the presence of this mutation, presented in lanes 2, 3, 4 and 6 of Figure 22. The additional cut was not observed in lanes 5 and 7, and hence the mutation was absent in these two lanes. Lanes M and 1 represents 100bp DNA ladder and undigested PCR product of out of exons 6-7.

4.6.4.7. Detection of Acores variant (595A>G) in exon 6 by BstUI restriction digestion analysis

Another mutation 595A>G was also investigated in exon 6 by digestion of the

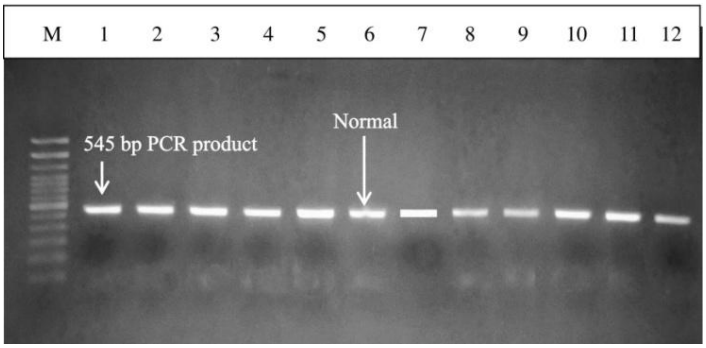


Figure 23. Agarose gel electrophoresis of PCR products digested with BstUI (Acores 595 A>G). M → 100bp DNA ladder, lane 1 → undigested PCR product of exons 6-7, lanes 2-12 → 595 A>G mutation is absent.

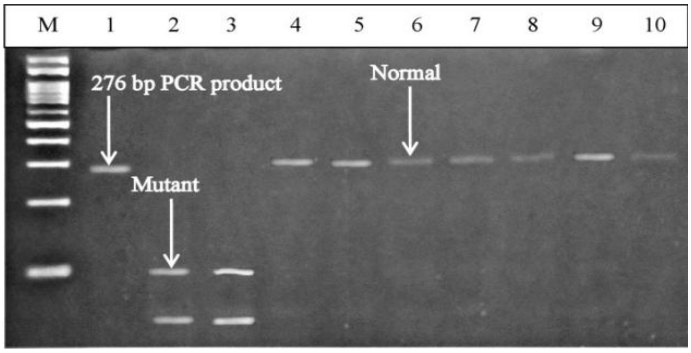


Figure 24. Agarose gel electrophoresis of PCR products digested with MnlI (Kalyan-Kerala/ Jamnagar / Rohini 949G>A). M → 100bp DNA ladder, lane 1 → undigested PCR product of exon 9, lanes 2, 3 → 949 G>A mutation is present, lanes 4-10 → 949G>A mutation is absent.

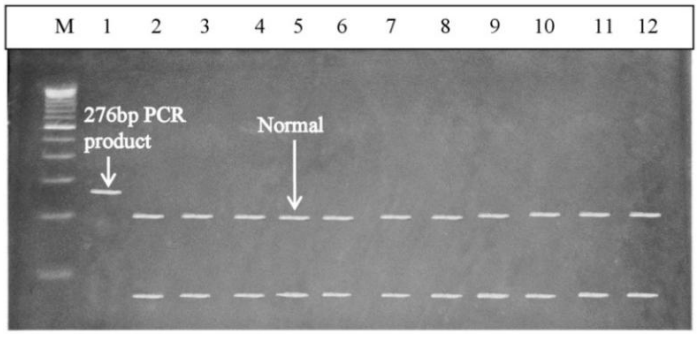


Figure 25. Agarose gel electrophoresis of PCR products digested with BstXI (Chatham (1003G>A). M → 100bp DNA ladder, lane 1 → undigested PCR product of exon 9, lanes 2-12 → 1003G>A mutation is absent.

PCR products by restriction endonuclease BstUI, which cuts the DNA at 5′CG/CG3′. The PCR products of exon 6-7 were digested with BstUI at 37°C for 1-2 hours. Exon 6 does not have a cut site for BstUI, but the mutation 595A>G creates a cut site in the exon. The agarose gel electrophoresis images of the digested samples are displayed in Figure 23. In the figure, lanes M and 1 represent the 100bp DNA ladder and uncut PCR product of exons 6-7, respectively; lanes 2-12 represents the digested products. But, the cut was not observed in any of the samples, thereby showing the absence of the mutation.

4.6.4.8. Detection of Kalyan-Kerala/Jamnagar/Rohini variant (949G>A) in exon 9 by MnlI restriction digestion analysis

The samples were examined for the presence of 949G>A mutation in exon 6 using restriction endonuclease MnlI. Digestion was carried out at 37°C for 1-2 hours. MnlI recognizes 5′CCTC3′ and creates cut seven nucleotides downstream the recognition site. In exon 9 there is no any cut site for MnlI, but this mutation (949G>A) creates a cut site for MnlI in exon 9. The agarose gel electrophoresis images demonstrating the digestion of PCR products (exon 9) by MnlI are presented in Figure 24. Out of 144 samples, 13 were detected to have the mutation, as shown in lanes 2 and 3 of the figure. In the other samples, the cut was not observed (lanes 4-10). Lane M represents 100bp DNA ladder and lane 1 represents 276bp uncut PCR product of exon 9.

4.6.4.9. Detection of Chatham variant (1003G>A) in exon 9 by BstXI restriction digestion analysis

Presence of the mutation 1003G>A was also examined in exon 9 by digestion with restriction endonuclease BstXI. The PCR products of exon 9 were digested with BstXI at 55°C for 1-2 hours. This mutation creates an additional cut site for BstXI in exon 9. The cut site of BstXI is 5′CCANNNNN/NTGG3′. The agarose gel electrophoresis images of the PCR products digested with BstXI are presented in Figure 25. Lane M represents 100bp DNA ladder and lane 1 represents uncut PCR product of exon 9. Lanes 2-12 represents the

digested PCR products, but the additional cut site was not observed in any of the samples. Hence, the mutation was absent in the samples.

4.6.4.10. Detection of Guadalajara variant (1159C>T) in exon 10 by HhaI restriction digestion analysis

The G6PD deficient samples were investigated for presence of 1159C>T mutation in exon 10 by digestion with restriction endonuclease HhaI. Hhai recognizes and cuts DNA at 5′GCG/C3′. Digestion of the PCR products was carried out at 37°C for 1-2 hours. Exon 10 has a cut site for HhaI in its normal state, but the mutation 1159C>T eliminates this cut site from the exon. The digested PCR products are displayed in Figure 26, where lanes M and 1 represent 100bp DNA ladder and uncut PCR product of exon 10, respectively. In lanes 2-12, cuts were observed, indicating the absence of 1159C>T mutation.

4.6.4.11. Detection of Union variant (1360C>T) in exon 11 by HhaI restriction digestion analysis

G6PD Union (1360C>T) in exon 11 was checked by digestion with restriction endonuclease HhaI at 37°C for 1-2 hours. Exon 11 has a cut site for HhaI (5′GCG/C3′) in its normal state, but the mutation 1360C>T eliminates the cut site for HhaI from the exon. The agarose gel electrophoresis images of PCR products of exon 10 digested with HhaI is presented in Figure 27. Lanes M and 1 represent the 100bp DNA ladder and uncut PCR product of exon 11, respectively; and lanes 2-9 represent the digested PCR products. The digested products showed the normal cuts, hence indicating the absence of the mutation in the samples.

4.6.4.12. Detection of Canton variant (1376G>T) in exon 12 by AflII restriction digestion analysis

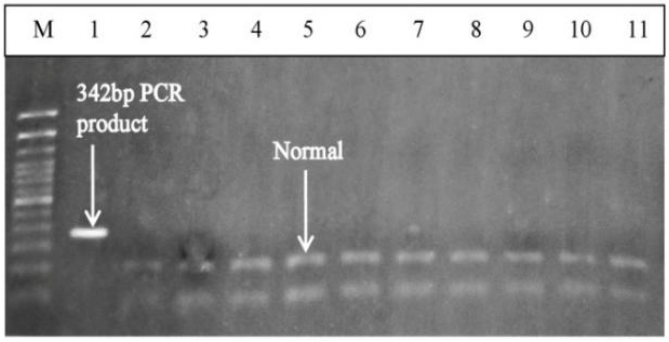


Figure 26. Agarose gel electrophoresis of PCR products digested with HhaI (Guadalajara 1159C>T). M → 100bp DNA ladder, lane 1 → undigested PCR product of exon 10, lanes 2-12 → 1159C>T mutation was absent in all samples.

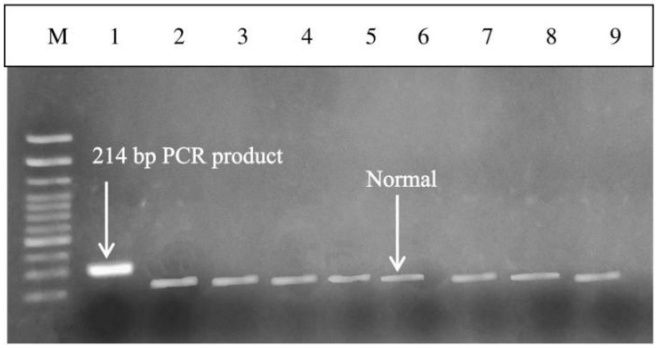


Figure 27. Agarose gel electrophoresis of PCR products digested with HhaI (Union 1360C>T). M → 100bp DNA ladder, lane 1 → undigested PCR product of exon 11, lanes 2-12 → 1360C>T mutation was absent.

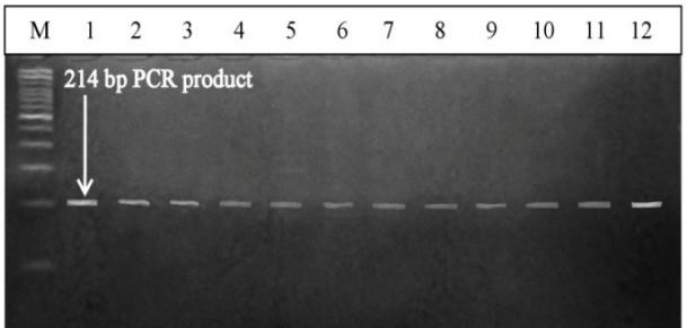


Figure 28. Agarose gel electrophoresis of PCR products digested with AflII (Canton 1376G>T). M → 100bp DNA ladder, lane 1 → undigested PCR product of exon 12, lanes 2-12 → 1376G>T mutation was absent.

PCR products of exon 12 of all the G6PD deficient samples were digested with restriction endonuclease AflII to check for the presence of 1376G>T mutation. The cut site

of AflII is 5′C/TTAAG3′. Digestion of the PCR products was done at 37°C for 1-2 hours. This mutation results in the creation of a cut site for AflII in exon 9. The images of the agarose gel electrophoresis of the samples treated with AflII are shown in Figure 28, where lanes M and 1 represent 100bp DNA ladder and uncut PCR product of exon 12, respectively; and lanes 2-12 represent the digested products. From the figure, it is evident that the mutation was not present in any of the samples as the cut was not seen.

4.6.4.13. Detection of Kaiping variant (1388G>A) in exon 12 by NdeI restriction digestion analysis

The samples were examined for another mutation 1388G>A in exon 12 by digestion with restriction endonuclease NdeI, which cuts the DNA at 5′CA/TATG3′. The digestion of the PCR products was carried out at 37°C for 1-2 hours. The mutation 1388G>A generates a cut site for NdeI in exon 12. Figure 29 represents the agarose gel electrophoresis images of samples treated with NdeI. Lanes M and 1 represent the 100bp DNA ladder and uncut PCR product of exon 12 respectively, and lanes 2-12 represent the digested products of exon 12. The cut was not observed in any of the samples. Hence, the mutation was absent.

4.6.5. Statistical analysis of Glucose-6-Phosphate Dehydrogenase variants

4.6.5.1. Variant-wise distribution among the Glucose-6-Phosphate Dehydrogenase deficient subjects

The frequency of the variants detected is presented in Figure 30. Majority of the deficient cases were found to be of Orissa variant (63.8%). The variant was detected exclusively among the Proto-Australoid population. Whereas, the variant Mahidol was

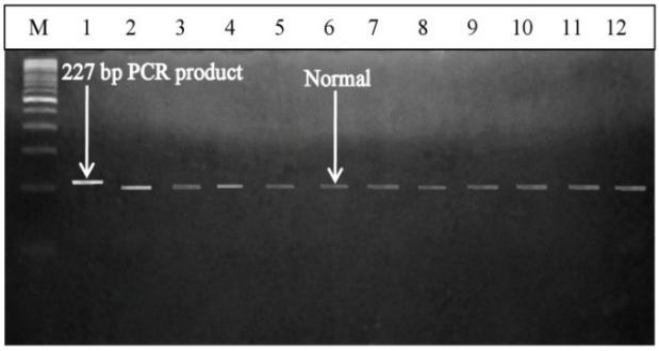


Figure 29. Agarose gel electrophoresis of PCR products digested with NdeI (Kaiping 1388G>A). M → 100bp DNA ladder, lane 1 → undigested PCR product of exon 12, lanes 2-12 → 1388G>A mutation was absent.

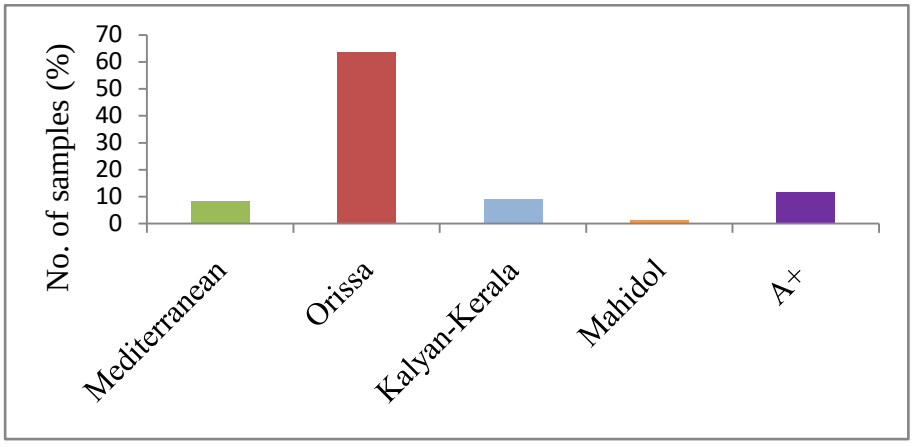


Figure 30. Frequency of G6PD variants detected in the study.

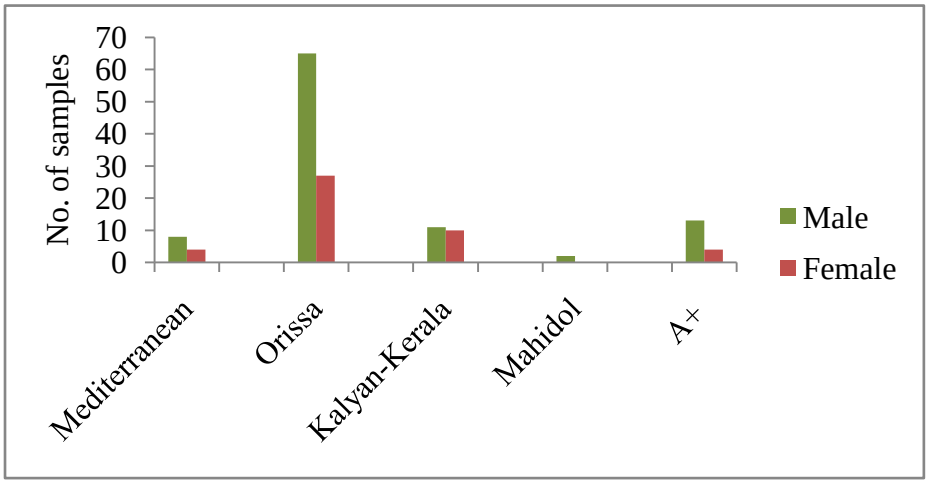


Figure 31. Gender-wise distribution of G6PD variants.

found in only 1.3% of the samples who belong to the Mongoloid race. The other three variants, namely Mediterranean, Kalyan-Kerala and A^+ were observed in samples from both the races.

4.6.5.2. Gender-wise distribution among the Glucose-6-Phosphate Dehydrogenase deficient subjects

The gender-wise distribution of the variants is depicted in Figure 31. In the variants Mediterranean, Orissa, Mahidol and A^+, the number of males were higher. Whereas, in case of Kalyan-Kerala almost equal number of males and females were found.

4.7. *In-silico* study of Glucose-6-Phosphate Dehydrogenase variants

4.7.1. Three dimensional structure modeling of detected Glucose-6-Phosphate Dehydrogenase variants

In the present study, altogether five variants were detected. These are Mediterranean, Orissa, Kalyan-Kerala, Mahidol and A^+. G6PD Mediterranean being a class II variant (less than 10% enzyme activity) and molecular dynamics simulation studies are performed by earlier workers; hence we excluded this variant from further *in silico* analysis. The three dimensional (3D) structures of the remaining four G6PD variants (Orissa, Kalyan-Kerala, Mahidol and A^+) were modeled using I-TASSER and validation of the generated structures were performed using ProTSAV. In the modeled 3D structures, the positions of the mutations have been represented in red colour. ProTSAV is a metaserver which incorporates Procheck, ProSA-Web, ERRAT, Verify3D, dDFire, Naccess, MolProbity, D2N, ProQ and PSN-QA to analyze the quality and accuracy of the modeled structure. In the result obtained from ProTSAV, the models in the green coloured region indicates that the structure is within an RMSD of 0-2Å, yellow region indicates RMSD between 2-5Å, orange region indicates an RMSD within the range 5-8Å and red color indicates an RMSD beyond 8Å. The results on the models and their validation are described below.

4.7.1.1. Three dimensional structure modeling and validation of G6PD Orissa (131C>G)

G6PD Orissa (131C>G) mutation lies in the catalytic NADP+ binding domain of G6PD. The 3D ribbon structure of the variant modeled using I-TASSER is shown in Figure 32a. The mutation was seen to affect the α-helix region of the protein structure. The validation by ProTSAV, as shown in Figure 32b showed the structure to have an RMSD of 2-5Å.

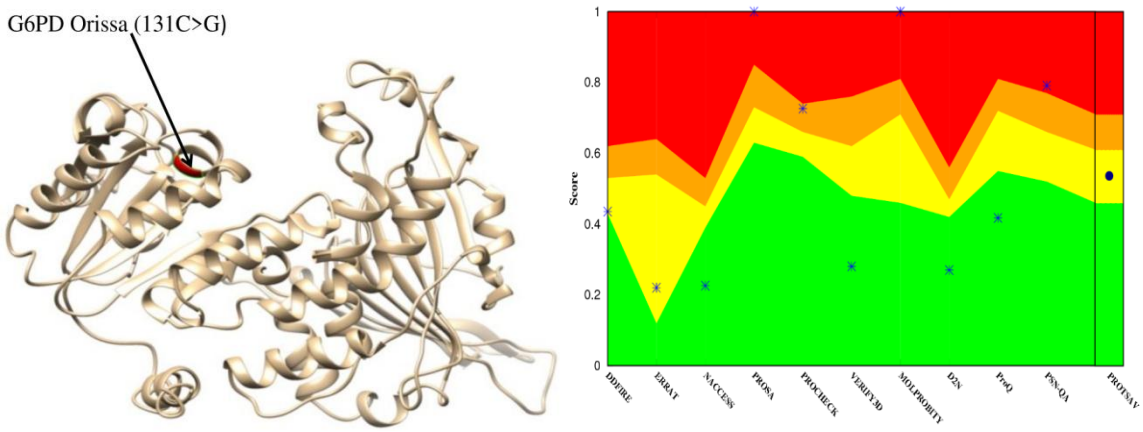
4.7.1.2. Three dimensional structure modeling and validation of G6PD Kalyan-Kerala (949G>A)

G6PD Kalyan-Kerala (949G>A) is located in the α-helix region of the protein structure, as depicted in Figure 33a. This mutation lies in the β + α domain of the protein. The modeled structure of Kalyan-Kerala also showed an RMSD within the range 2-5Å (Figure 33b).

4.7.1.3. Three dimensional structure modeling and validation of G6PD Mahidol (487G>A)

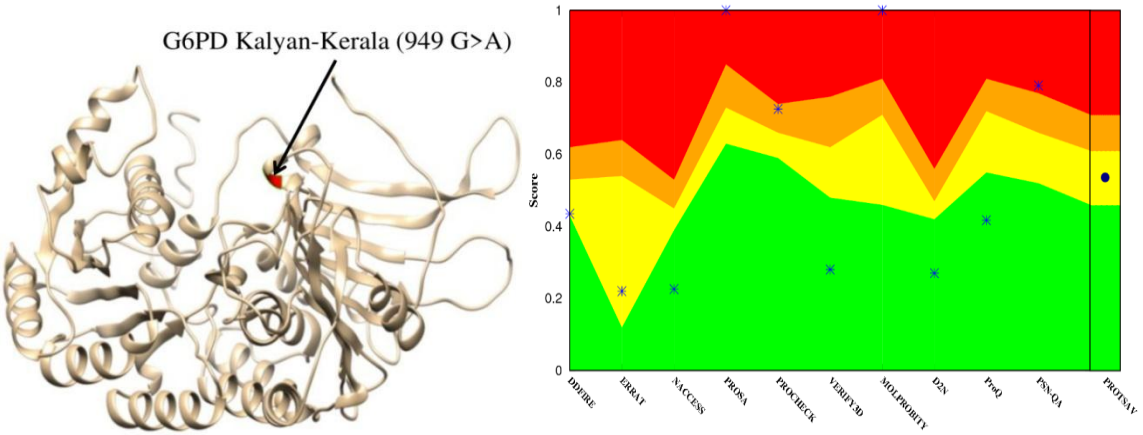
The mutation 487G>A also lies in the β + α domain of the protein structure. The mutation is seen to affect the loop region of the protein (Figure 34a). The modeled structure has an RMSD 2-5Å as shown in Figure 34b.

4.7.1.4. Three dimensional structure modeling and validation of G6PD A^+ (376A>G)



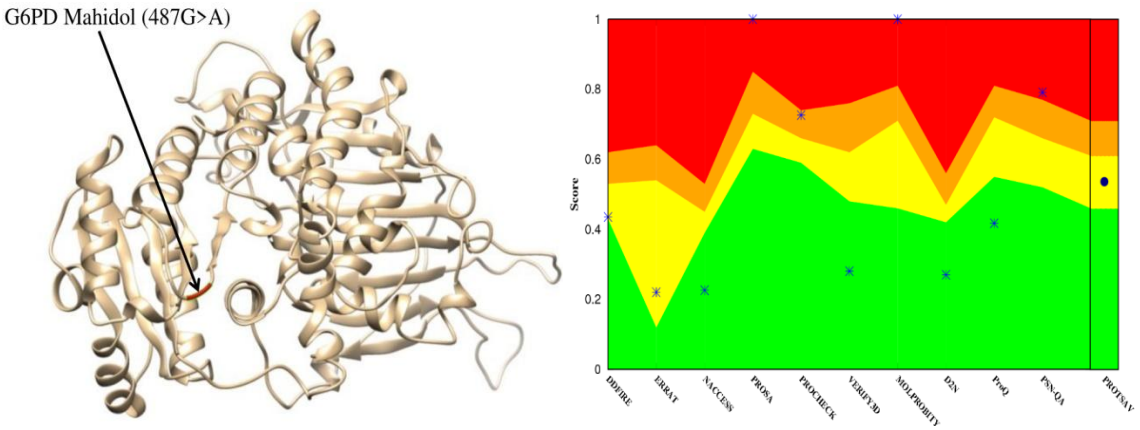
(a) (b)

Figure 32. Modeled structure and validation of G6PD Orissa variant. (a) G6PD Orissa modeled using I-TASSER. (b) Validation of G6PD Orissa by ProTSAV.



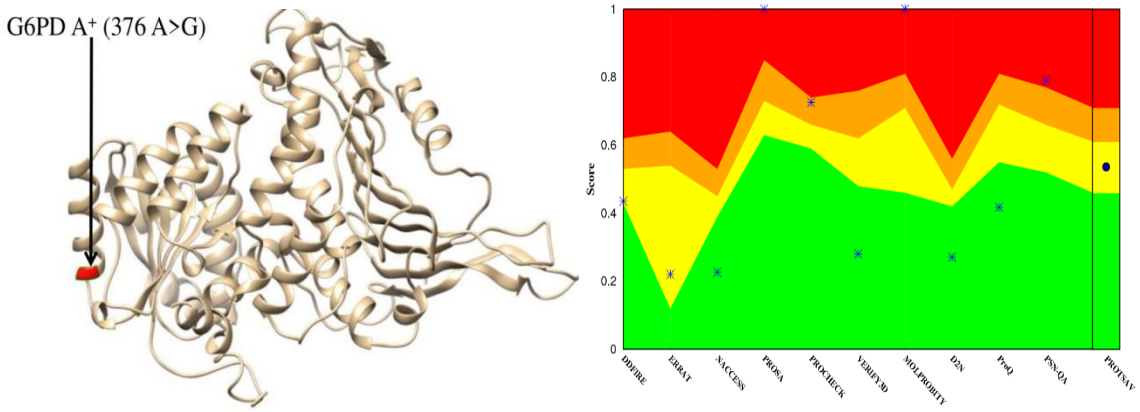
(a) (b)

Figure 33. Modeled structure and validation of G6PD Kalyan-Kerala/Jamnagar/Rohini variant. (a) G6PD Kalyan-Kerala/Jamnagar/Rohini modeled using I-TASSER. (b) Validation of G6PD Kalyan-Kerala/Jamnagar/Rohini by ProTSAV.



(a) (b)

Figure 34. Modeled structure and validation of G6PD Mahidol variant. (a) G6PD Mahidol modeled using I-TASSER. (b) Validation of G6PD Mahidol by ProTSAV.



(a) (b)

Figure 35. Modeled structure and validation of G6PD A^+ variant. (a) G6PD A^+ modeled using I-TASSER. (b) Validation of G6PD A^+ by ProTSAV.

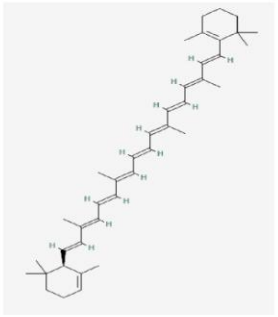
The mutation 376 A>G, lying in the β + α domain of the protein structure is shown in Figure 35a. This mutation was seen to affect the α-helix region of the protein. Figure 35b shows that the modeled structure of G6PD A^+ to have an RMSD between 2-5Å.

4.7.2. Retrieval of naturally available antioxidant compounds

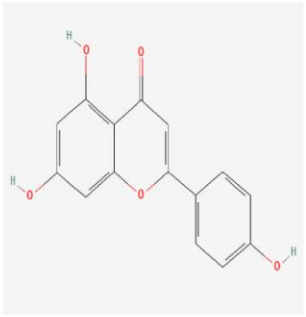
Naturally available antioxidant compounds were listed out from existing literature from different databases using appropriate keywords. The 3D conformers of the antioxidants were retrieved from Pubchem and stored in .sdf format. The compounds and their 2D structures are presented in Figure 36.

4.7.3. Molecular Docking of Glucose-6-Phosphate Dehydrogenase variants with natural antioxidants

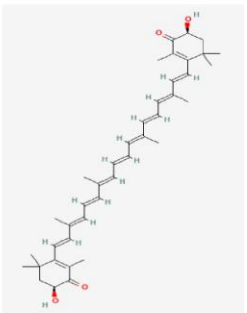
The natural antioxidants were evaluated for their binding affinity with the 3D structures of the variants by performing molecular docking. The binding pocket with the highest Molecular Surface (MS) volume (3998.4), pocket MS area (1724.7), openings (1), mouth MS area (676.2) and MS circumference sum (201.2) was selected for the study. The amino acid residues forming the binding pocket were Gly38, Ser40, Gly41, Asp42, Leu43, Ala44, Lys46, Lys47, Pro50, Thr51, Trp54, Ala71, Arg72, Phe87, Lys89, Tyr112, Leu140, Ala141, Leu142, Pro143, Pro144, Tyr147, Glu170, Lys171, Pro172, Phe173, Gly174, Arg175, Ile199, Asp200, His201, Tyr202, Lys205, Phe237, Glu239, Phe241, Gly242, Thr243, Glu244, Arg246, Tyr249, Phe250, Glu252, Phe253, Ile255, Arg257, Asp258, Val259, Gln261, Asn262, His263, Leu265, Gln266, Cys269, Leu292, Lys360, Arg365, Gln395, Asp435, Ala436, Tyr437, Val453, Glu457, Leu458, Ala461 and Trp462. The binding affinities obtained from docking study are presented in Table 21. The compounds that showed binding affinity ≤ -7.0kcal/mol and interacting with the active site amino acid residues were considered for further analysis.



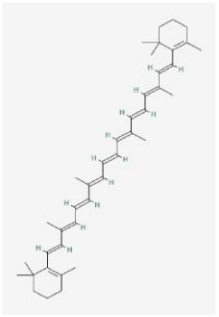
Alpha carotene



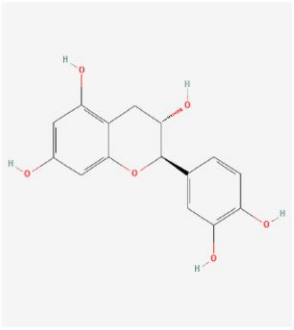
Apigenin



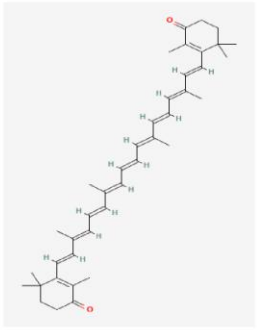
Astaxanthin



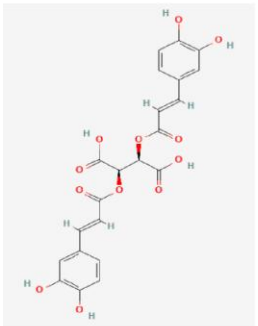
Beta carotene



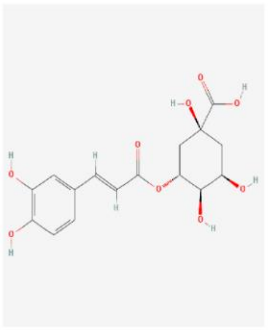
Catechin



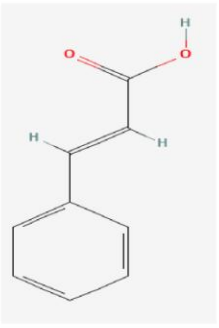
Canthaxanthin



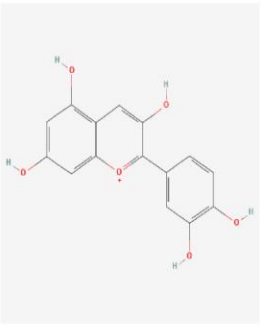
Chicoric acid



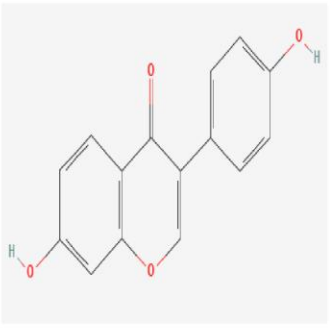
Chlorogenic acid



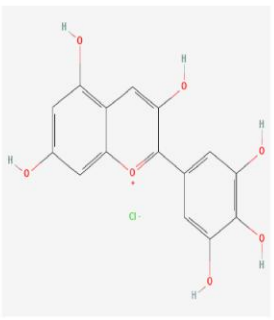
Cinnamic acid



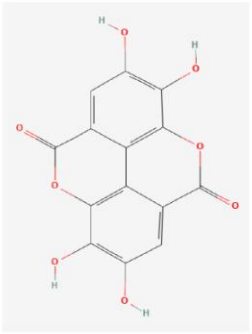
Cyanidin



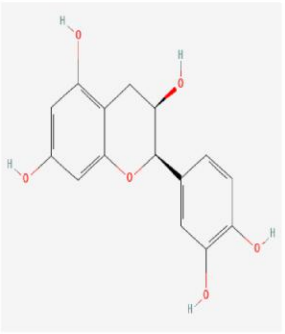
Daidzein



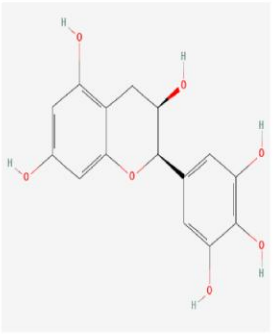
Delphinidin



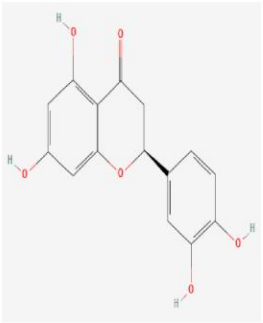
Ellagic acid



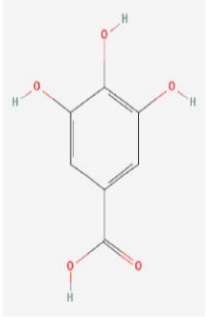
Epicatechin



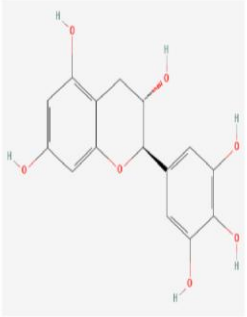
Epigallocatechin



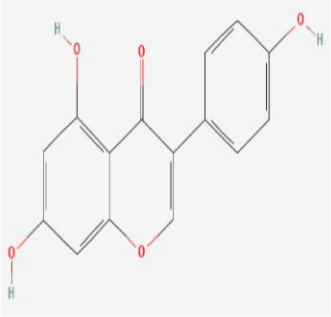
Eriodictyol



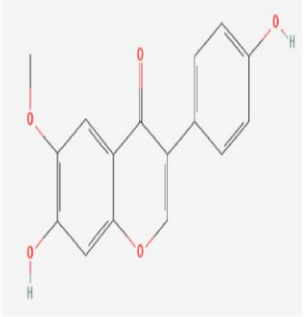
Gallic acid



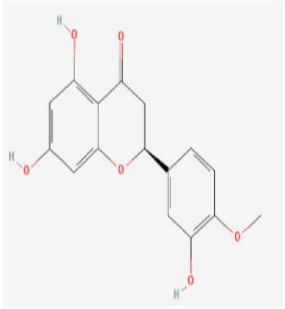
Gallocatechin



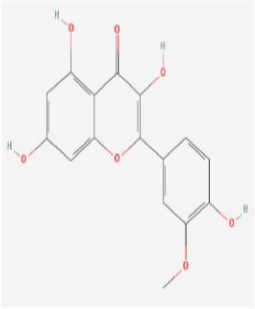
Genistein



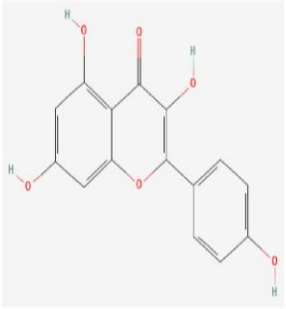
Glycitein



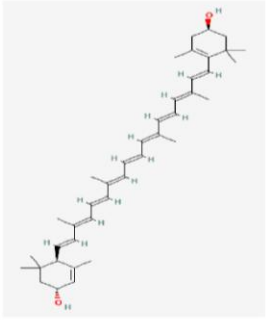
Hesperetin



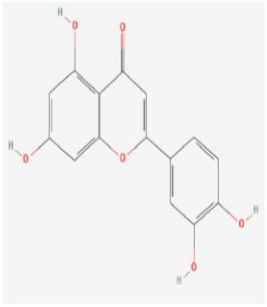
Isorhamnetin



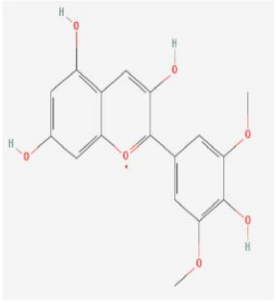
Kaempferol



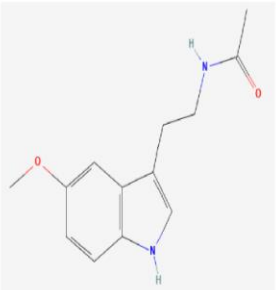
Lutein



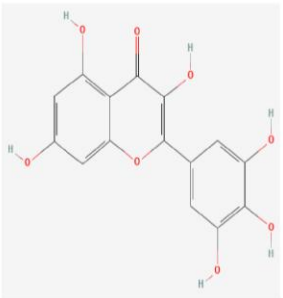
Luteolin



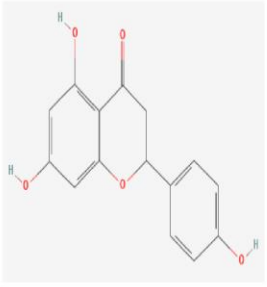
Malvidin



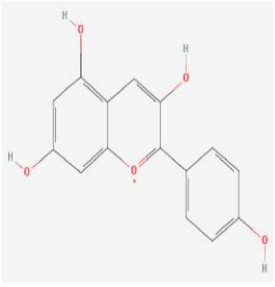
Melatonin



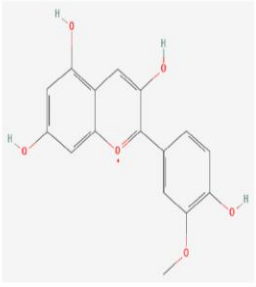
Myricetin



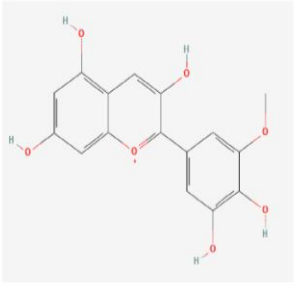
Naringenin



Pelargonidin



Peonidin



Petunidin

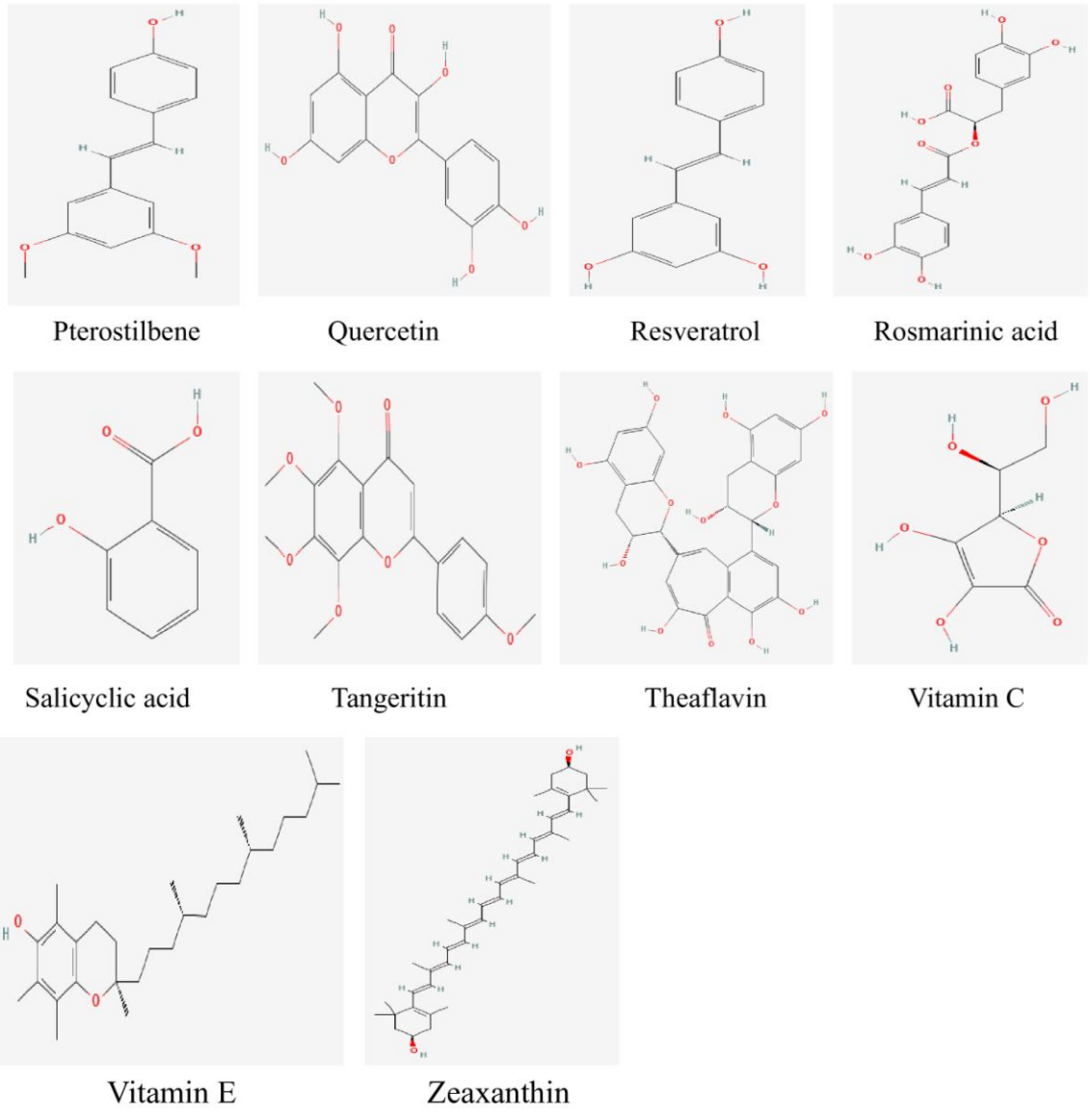


Figure 36. 2D structures of antioxidant compounds downloaded from Pubchem.

4.7.3.1. Molecular Docking of G6PD Orissa

In case of G6PD Orissa, the binding affinities of the ligands were observed to range between -5.2kcal/mol (with Salicylic acid) and -9.2kcal/mol (with Asthaxanthin). A total of 33 compounds out of 42, viz., Alpha-carotene, Apigenin, Astaxanthin, Beta-carotene, Canthaxanthin, Catechin, Chicoric acid, Cyanidin, Daidzein, Delphinidin,Ellagic acid, Epicatechin, Epigallocatechin, Eriodictyol, Gallocatechin, Genistein, Glycitein, Hesperetin, Isorhamnetin, Kaempferol, Lutein, Luteolin, Malvidin, Myricetin, Naringenin, Pelargonidin, Peonidin, Petunidin, Quercetin, Rosmarinic acid, Theaflavin, Vitamin E and Zeaxanthin were found to have binding affinity ≤ -7.0kcal/mol.

4.7.3.2. Molecular Docking of G6PD Kalyan-Kerala

The binding affinities of the compounds with Kalyan-Kerala were seen in the range between -5.1kcal/mol (with Salicylic acid) and -9.8kcal/mol (with Astaxanthin). Apigenin, Astaxanthin, Beta-carotene, Canthaxanthin, Catechin, Chicoric acid, Cyanidin, Daidzein, Delphinidin, Ellagic acid, Epicatechin, Epigallocatechin, Eriodictyol, Gallocatechin, Genistein, Hesperetin, Isorhamnetin, Kaempferol, Lutein, Luteolin, Myricetin, Naringenin, Pelargonidin, Peonidin, Petunidin, Quercetin, Rosmarinic acid, Theaflavin and Zeaxanthin showed binding affinities ≤ -7.0 kcal/mol.

4.7.3.3. Molecular Docking of G6PD Mahidol

G6PD Mahidol also showed the highest binding affinity with Astaxanthin (-9.4kcal/mol) and lowest with Salicylic acid (-4.8kcal/mol). The compounds found to have binding affinity ≤ -7.0 kcal/mol were Alpha-carotene, Apigenin, Astaxanthin, Beta-carotene, Canthaxanthin, Catechin, Cyanidin, Daidzein, Delphinidin, Ellagic acid, Epicatechin, Epigallocatechin, Eriodictyol, Gallocatechin, Genistein, Hesperetin, Isorhamnetin, Kaempferol, Lutein, Luteolin, Malvidin, Myricetin, Naringenin, Pelargonidin, Peonidin, Petunidin, Quercetin, Theaflavin and Zeaxanthin.

4.7.3.4. Molecular Docking of G6PD A^+

Astaxanthin and Vitamin C were seen to have highest (-10.5kcal/mol) and lowest binding affinities with G6PD A^+. Alpha-carotene, Apigenin, Astaxanthin, Beta-carotene, Canthaxanthin, Catechin, Chicoric acid, Daidzein, Delphinidin, Ellagic acid, Epicatechin, Epigallocatechin, Eriodictyol, Genistein, Glycitein, Hesperetin, Isorhamnetin, Kaempferol, Lutein, Luteolin, Myricetin, Naringenin, Petunidin, Quercetin, Theaflavin and Zeaxanthin were the compounds having binding affinity ≤ -7.0 kcal/mol.

Table 21. Binding affinities of the antioxidants with the 3D structures of the G6PD variants.

Sl. No.	Ligand	Binding affinity (kcal/mol)			
		Orissa	Kalyan-Kerala	Mahidol	A^+
1	Alpha-carotene	-8.3	-6.8	-7.6	-8.0
2	Apigenin	-7.8	-7.5	-7.3	-7.6
3	Astaxanthin	-9.2	-9.8	-9.4	-10.5
4	Beta-carotene	-8.1	-7.5	-7.9	-8.4
5	Canthaxanthin	-8.4	-7.7	-7.8	-8.1
6	Catechin	-7.8	-7.3	-7.1	-7.0
7	Chicoric acid	-8.3	-7.8	-6.7	-7.2
8	Chlorogenic acid	-6.9	-6.6	-6.6	-6.6
9	Cinnamic acid	-5.7	-5.6	-5.2	-6.0
10	Cyanidin	-7.8	-7.5	-7.2	-6.9
11	Daidzein	-7.4	-7.1	-7.6	-7.4
12	Delphinidin	-7.7	-7.7	-7.2	-7.3
13	Ellagic acid	-7.8	-7.7	-7.4	-7.5
14	Epicatechin	-7.9	-7.2	-7.3	-7.1
15	Epigallocatechin	-7.8	-7.0	-7.4	-7.3
16	Eriodictyol	-8.1	-7.5	-7.4	-7.6
17	Gallic acid	-5.7	-5.9	-5.7	-5.4

18	Gallocatechin	-7.8	-7.0	-7.2	-6.9
19	Genistein	-7.5	-7.0	-7.3	-7.0
20	Glycitein	-7.6	-6.9	-6.9	-6.9
21	Hesperetin	-8.1	-7.4	-7.7	-7.4
22	Isorhamnetin	-7.9	-7.2	-7.3	-7.0
23	Kaempferol	-7.8	-7.4	-7.2	-7.3
24	Lutein	-8.2	-8.8	-8.3	-7.6
25	Luteolin	-8.0	-7.7	-7.6	-7.1
26	Malvidin	-7.5	-6.5	-7.5	-6.8
27	Melatonin	-6.0	-6.2	-5.9	-6.3
28	Myricetin	-8.0	-7.4	-7.5	-7.3
29	Naringenin	-7.7	-7.2	-7.2	-7.3
30	Pelargonidin	-7.6	-7.2	-7.2	-6.8
31	Peonidin	-7.8	-7.0	-7.3	-6.8
32	Petunidin	-7.7	-7.7	-7.4	-7.4
33	Pterostilbene	-6.6	-6.4	-6.6	-6.2
34	Quercetin	-8.1	-7.7	-7.5	-7.2
35	Resveratrol	-5.9	-5.9	-6.1	-6.0
36	Rosmarinic acid	-7.4	-7.3	-5.9	-6.7
37	Salicyclic acid	-5.2	-5.1	-4.8	-5.1
38	Tangeritin	-6.4	-6.9	-6.4	-6.4
39	Theaflavin	-8.9	-9.1	-8.5	-8.4
40	Vitamin C	-5.4	-5.1	-5.2	-5.0
41	Vitamin E	-7.2	-6.6	-6.2	-6.5
42	Zeaxanthin	-8.0	-7.5	-7.7	-7.2

4.7.4. Analysis of drug-likeness and ADME properties of the natural antioxidants

The drug-likeness and ADME properties of the 42 compounds obtained using SwissADME are presented in Table 22. According to Lipinski's rule, the compounds that have molecular weight ≤500Da, number of H-bond acceptor ≤10, number of H-bond donor ≤5 and logP value ≤5 are considered to have drug-likeness property. Among these, 33 were found to possess drug-likeness properties. The compounds were Apigenin, Catechin, Chlorogenic acid, Cinnamic acid, Cyanidin, Daidzein, Delphinidin, Ellagic acid, Epicatechin, Epigallocatechin, Eriodictyol, Gallic acid, Gallocatechin, Genistein, Glycitein, Hesperetin, Isorhamnetin, Kaempferol, Luteolin, Malvidin, Melatonin, Myricetin, Naringenin, Pelargonidin, Peonidin, Petunidin, Pterostilbene, Quercetin, Resveratrol, Rosmarinic acid, Salicyclic acid, Tangeritin, Vitamin Cand Vitamin E. The Topological Surface Area (TPSA) values of the compounds were also less than 140 $Å^2$, indicating that the compounds have higher permeability through the cell membrane.

Table 22. Drug-likeness and ADME properties of the compounds.

Sl. No.	**Ligand**	**MW (Da)**	**TPSA**	**HbA**	**HbD**	**logP**	**Lipinski violation**	**Drug likeness**
1	Alpha carotene	536.87	0	0	0	8.96	2	No
2	Apigenin	270.24	90.9	5	3	0.52	0	Yes
3	Astaxanthin	596.84	74.6	4	2	5.09	2	No
4	Beta carotene	536.87	0	0	0	8.96	2	No
5	Canthaxanthin	564.84	34.14	2	0	6.82	2	No
6	Catechin	290.27	110.38	6	5	0.24	0	Yes
7	Chicoric acid	474.37	208.12	12	6	0.14	2	No
8	Chlorogenic acid	354.31	164.75	9	6	-1.05	1	Yes
9	Cinnamic acid	148.16	37.3	2	1	1.90	0	Yes
10	Cyanidin	287.24	114.29	6	5	0.32	0	Yes
11	Daidzein	254.24	70.67	4	2	1.08	0	Yes
12	Delphinidin	338.7	134.52	7	6	0.03	1	Yes

13	Ellagic acid	302.19	141.34	8	4	0.14	0	Yes
14	Epicatechin	290.27	110.38	6	5	0.24	0	Yes
15	Epigallocate-chin	306.27	130.61	7	6	-0.29	1	Yes
16	Eriodictyol	288.25	107.22	6	4	0.16	0	Yes
17	Gallic acid	170.12	97.99	5	4	-0.16	0	Yes
18	Gallocatechin	306.27	130.61	7	6	-0.29	1	Yes
19	Genistein	270.24	90.9	5	3	0.52	0	Yes
20	Glycitein	284.26	79.9	5	2	0.77	0	Yes
21	Hesperetin	302.28	96.22	6	3	0.41	0	Yes
22	Isorhamnetin	316.26	120.36	7	4	-0.31	0	Yes
23	Kaempferol	286.24	111.13	6	4	-0.03	0	Yes
24	Lutein	568.87	40.46	2	2	6.96	2	No
25	Luteolin	286.24	111.13	6	4	-0.03	0	Yes
26	Malvidin	331.3	112.52	7	4	0.28	0	Yes
27	Melatonin	232.28	54.12	2	2	0.97	0	Yes
28	Myricetin	318.24	151.59	8	6	-1.08	1	Yes
29	Naringenin	272.25	86.99	5	3	0.71	0	Yes
30	Pelargonidin	271.24	94.06	5	4	0.86	0	Yes
31	Peonidin	301.27	103.29	6	4	0.57	0	Yes
32	Petunidin	317.27	123.52	7	5	0.03	0	Yes
33	Pterostilbene	256.3	38.69	3	1	2.76	0	Yes
34	Quercetin	302.24	131.36	7	5	-0.56	0	Yes
35	Resveratrol	228.24	60.69	3	3	2.26	0	Yes
36	Rosmarinic acid	360.31	144.52	8	5	0.90	0	Yes
37	Salicyclic acid	138.12	57.53	3	2	0.99	0	Yes
38	Tangeritin	372.37	76.36	7	0	0.63	0	Yes
39	Theaflavin	564.49	217.6	12	9	-0.79	3	No

40	Vitamin C	176.12	107.22	6	4	-2.60	0	Yes
41	Vitamin E	430.71	29.46	2	1	6.14	1	Yes
42	Zeaxanthin	568.87	40.46	2	2	6.96	2	No

4.7.5. Analysis of toxicity of the natural antioxidants

The toxicity of the natural antioxidants analyzed using ProTox II are presented in Table 23. The toxicological parameters viz., carcinogenicity, mutagenicity, skin irritation, skin sensitization and biodegradability were analyzed. The LD50 values (mg/kg body weight), that is the median lethal dose at which 50% death of test subjects occur when exposed to a compound, were also studied. According to the globally recognized classification, compounds are classified into six classes on the basis of their LD50 values-

a) Compounds that have LD50 value ≤5 are classified as class 1 and they are fatal if swallowed.
b) LD50 values between 5 and 50 are classified as class 2 that are also fatal if swallowed.
c) Class 3 compounds show LD50 value of 50-300 and they are toxic if swallowed.
d) Class 4 compounds have LD50 value between 300-2000 and are harmful if swallowed.
e) LD50 value between 2000-5000 is classified as class 5 and may be harmful if swallowed.
f) Compounds with LD50 value >5000 are classified as class 6 that are non-toxic.

In the present study, 27 out of 42 compounds were found to be class 5 and 6 compounds. The compounds identified as class 5 were Apigenin, Astaxanthin, Chicoric acid, Chlorogenic acid, Cinnamic acid, Cyanidin, Daidzein, Delphinidin, Genistein, Glycitein, Isorhamnetin, Kaempferol, Luteolin, Malvidin, Pelargonidin, Peonidin, Petunidin, Rosmarinic acid, Tangeritin, Theaflavin, Vitamin C and Vitamin E. Whereas, Catechin, Canthaxanthin, Epicatechin, Epigallocatechin and Gallocatechin were identified as class 6.

Table 23. Toxicity analysis of the ligands.

Compound	Carcinog-enicity	Mutage-nicity	Skin Irritat-ion	Skin Sensiti-zation	Biodegra-dability	Toxicity	
						LD50 (mg/kg)	Toxicity class
Alpha carotene	Carcinogen	Non mutagen	Severe	Weak	Degradable	1510	4
Apigenin	Non carcinogen	Non mutagen	None	Strong	Non degradable	2500	5
Astaxanthin	Non carcinogen	Non mutagen	Mild	Weak	Degradable	4600	5
Beta carotene	Carcinogen	Non mutagen	Severe	Weak	Degradable	1510	4
Catechin	Non carcinogen	Non mutagen	None	Strong	Non degradable	10000	6
Canthaxanthin	Carcinogen	Non mutagen	Severe	Weak	Degradable	10000	6
Chicoric acid	Non carcinogen	Non mutagen	None	Strong	Degradable	5000	5
Chlorogenic acid	Non carcinogen	Non mutagen	Mild	-	Degradable	5000	5
Cinnamic acid	Carcinogen	Non mutagen	Mild	Strong	Degradable	2500	5
Cyanidin	Non carcinogen	Non mutagen	None	Strong	Non degradable	5000	5
Daidzein	Non carcinogen	Non mutagen	None	Strong	Non degradable	2430	5
Delphinidin	Non carcinogen	Non mutagen	None	Strong	Non degradable	5000	5
Ellagic acid	Non	Non	None	-	Non	2991	4

	carcinogen	mutagen			degradable		
Epicatechin	Non carcinogen	Non mutagen	None	Strong	Non degradable	10000	6
Epigallocatechin	Non carcinogen	Non mutagen	None	Strong	Non degradable	10000	6
Eriodictyol	Non carcinogen	Non mutagen	None	Strong	Non degradable	2000	4
Gallic acid	Non carcinogen	Non mutagen	None	Strong	Degradable	2000	4
Gallocatechin	Non carcinogen	Non mutagen	None	Strong	Non degradable	10000	6
Genistein	Non carcinogen	Non mutagen	None	Strong	Non degradable	2500	5
Glycitein	Non carcinogen	Non mutagen	None	Weak	Degradable	2500	5
Hesperetin	Non carcinogen	Non mutagen	None	Strong	Degradable	2000	4
Isorhamnetin	Non carcinogen	Mutagen	Severe	Strong	Non degradable	5000	5
Kaempferol	Non carcinogen	Mutagen	None	Strong	Non degradable	3919	5
Lutein	Non carcinogen	Non mutagen	Severe	Weak	Degradable	10	2
Luteolin	Non carcinogen	Non mutagen	None	Strong	Non degradable	3919	5
Malvidin	Non carcinogen	Non mutagen	None	Strong	Non degradable	5000	5
Melatonin	Non carcinogen	Non mutagen	None	None	Non degradable	963	4

Myricetin	Non carcinogen	Non mutagen	None	Strong	Non degradable	159	3
Naringenin	Non carcinogen	Non mutagen	None	Strong	Non degradable	2000	4
Pelargonidin	Non carcinogen	Non mutagen	None	Strong	Non degradable	3919	5
Peonidin	Non carcinogen	Non mutagen	None	Strong	Non degradable	5000	5
Petunidin	Non carcinogen	Non mutagen	None	Strong	Non degradable	5000	5
Pterostilbene	Non carcinogen	Non mutagen	None	None	Non degradable	1560	4
Quercetin	Non carcinogen	Non mutagen	None	Strong	Non degradable	159	3
Resveratrol	Non carcinogen	Non mutagen	None	Strong	Non degradable	1560	4
Rosmarinic acid	Non carcinogen	Non mutagen	None	None	Non degradable	5000	5
Salicyclic acid	Non carcinogen	Non mutagen	severe	None	Degradable	1034	4
Tangeritin	Non carcinogen	Non mutagen	severe	Weak	Degradable	5000	5
Theaflavin	Non carcinogen	Non mutagen	None	Strong	Non degradable	2500	5
Vitamin C	Non carcinogen	Non mutagen	None	None	Degradable	3367	5
Vitamin E	Non carcinogen	Non mutagen	severe	None	Non degradable	5000	5
Zeaxanthin	Non	Non	severe	Weak	Degradable	10	2

	carcinogen	mutagen					

4.7.6. Interaction of the ligands with the Glucose-6-Phosphate Dehydrogenase variants

The residues of the ligands interacting with each of the variants are presented in Table 24. Ligands showing interactions with the amino acid residues of binding pocket were taken into consideration for further analysis.

4.7.7. Visualization and analysis of best complexes

Based on the above observations, three parameters, binding affinity, drug-likeness and toxicity class were taken into consideration. Four best complexes, one for each of the variants were selected and visualized by LigPlot^{+}. The selected complexes were Orissa with Myricetin, Kalyan-Kerala with Apigenin, Mahidol with Catechin and A^{+} with Diadzen. Figure 37-40 represents the detailed interaction of the selected complexes. In the figures, the green dash represents hydrogen bonding, whereas the red arc with spokes represents hydrophobic interactions.

Table 24. Amino acid residues of the ligands interacting with the variants.

Ligand	Orissa	Kalyan-Kerala	Mahidol	A^+
Alpha-carotene	Arg57, Lys89, Lys94, Glu94, Lys205, Glu398, Gln395, Leu433	Val304, Pro477, Lys478, Ile480, Glu494, Leu495, Lys497, Arg498	Lys205, Glu206, Pro240, Phe241, Glu315, Arg365, Gln395, Glu398, Tyr484	Lys171, Phe241, Tyr249, Phe253, Glu315
Apigenin	Lys205, Glu206, Gln395, Glu398, Pro434, Ala436	Phe241, Glu244, Tyr249, Phe250, Phe253, Asp258	Val12, Ser8, Gln195, Arg166, Trp164	Thr236, Arg357, Arg370, Arg487, Met496, Phe501, Tyr503, Tyr507
Catechin	Lys403, Arg487, Met496, Phe501, Gln502, Tyr503, Glu504	Lys205, Glu206, Ala436	Glu206, Gln395, Lys205, His201, Ala436, Tyr437	Arg9, Arg192, Glu193, Asp194, Asp282, Arg454, Asp456
Chlorogenic acid	Arg370, Tyr401, Lys403, Glu504, Tyr507	Lys508, Trp509, Thr423, Arg393	Glu206, Leu433, Glu398, Pro396, Gln395, Lys205	Ser418, Glu419, Lys508, Trp509, Val510, Asn511, His513,
Cinnamic acid	Arg246, Tyr249, Phe250, Phe253	Arg246, Tyr249, Phe250, Phe253	Arg370, Tyr507	Glu419, Lys508, Trp509, Val510,

				Asn511, Pro512
Cyanidin	Lys403, Arg487, Phe501, Tyr503, Glu504,Tyr507,	Lys205, Glu208, Glu398, Ala436, Tyr437, Arg439	Gln11, Ile15, Asp194, Arg166, Trp164	Ser8, Val12, His27, Gln28, Asp30, Trp164, Arg166
Daidzein	Arg370, Lys403, Phe501, Tyr503, Glu504, Tyr507, Trp509	Phe153, Lys171, Glu244, Arg246, Tyr249, Phe250, Lys360	Glu416, Ser418, Glu419, Asp421, Trp509, Val510, Pro512, Lys514	Leu142, Pro144, Glu170, Lys171, Tyr249, Phe253
Delphinidin	Leu16, Asp24, Thr279, Ser281, Asp443, His451	Lys205, Glu206, Glu398, Pro434, Ala436, Tyr437,	Lys47, His201, Lys205, Glu206, Glu398, Pro434, Ala436, Tyr487	Arg357, Glu368, Arg370, Met406, Arg487, Phe501, Glu504, Tyr507
Ellagic acid	Arg393, Asn397, Tyr401, Thr423, Asn426, Tyr507,	Lys205, Glu206, Arg365, Gln395, Pro434, Ala436, Arg439	Lys205, Glu206, Arg385, Gln395, Pro434, Ala436	Ser418, Glu419, Asp421, Lys508, Trp509, Val510
Epigallocatechin	Glu317, Arg370, Gly372, Lys403, Arg487, Phe501, Tyr503, Glu504	Lys205, Glu206, Glu398, Asp435, Tyr437	Ser8, Gln11, Val12, Leu16, Arg166, Trp164, Gln195	Val6, Arg9, Glu193, Asp282, Asp286, Arg454

Eriodictyol	Arg370, Gly372, Tyr401, Lys403, Phe501, Gln502, Glu504, Tyr507	Thr236, Arg357, Glu364, Lys366, Arg370, Lys403, Phe501, Tyr503, Glu504	Glu417, Ser418, Asp421, Lys508, Pro512, Lys514	Phe26, His27, Gln28, Gln59, Pro62
Gallic acid	Lys238, Arg357, Lys366, Glu368, Met496	Lys205, Glu206, Glu398, Pro434, Ala436	Phe26, His27, Asp30, His32, Trp164, Arg166, Cys446	Thr279, Asn280, Ser281, Ser448
Gallocatechin	Arg370, Gln372, Lys403, Arg487, Phe501, Tyr503, Glu504	Lys205, Glu206, Glu398, Tyr437, Ala438	His20, Lys47, Lys205, Glu206, Gln395, Pro434, Ala436, Tyr437	Arg357, Arg370, Arg487, Tyr503, Glu504
Genistein	Lys205, Glu206, Val394, Gln395, Pro434, Ala436, Arg439	Lys171, Glu244, Arg246, Tyr249, Phe250, Phe253, Lys360	Glu417, Ser418, Glu419, Asp421, Lys508, Trp509, Val510, Lys514	Arg370, Phe501, Tyr503, Glu504
Glycitein	Arg393, Asn397, Pro512	Lys205, Glu206, Ala436, Tyr437	Asp421, Trp509, Val510, Lys514	Arg370, Glu504, Glu389, Tyr401, Trp509
Hesperetin	Cys13, Leu16, Arg17,	Lys205, Glu206,	Ser8, Val12, Arg166,	Phe501, Arg370,

	Leu20, Ser281, Leu442, Asp443, His451	Gln395, Glu398, Asp435, Tyr437	Gln195	Tyr503, Arg487, Lys238, Thr236
Isorhamnetin	Lys205, Glu206, Arg365, Gln395, Pro434, Ala436	His201, Lys205, Glu206, Glu398, Tyr437	Ser8, Val12, Ile15, Gln28, Asp30, Trp164, Arg166, Asp194, Gln195	Arg370, Glu389, Glu504, Tyr507, Trp509
Kaempferol	Lys47, Lys205, Glu208, Gln395, Pro434, Ala438	Lys205, Glu206, Glu395, Glu398, Pro434, Tyr437, Arg439	Gln11, Ile15, Trp164, Arg166, Asp194, Gln195	Asn280, His451, Leu16, Ala25, Leu20
Luteolin	Lys205, Glu208, Arg365, Gln395, Pro434, Ala436	Lys205, Glu206, Glu398, Pro434, Ala436, Tyr437	Ser8, Trp164, Arg166, Asp194, Gln195	Ser281, His451, Gln209, Arg439, Asp443, Leu18
Malvidin	Thr279, Ser281, His451	Gly171, Arg175, Ser179, Leu183, Glu252, Phe253	Gln11, Phe26, Trp164, Arg186, Asp194, Gln195	Lys238, Glu364, Arg370, Glu389, Arg487, Tyr503, Glu504
Melatonin	Lys47, Lys205, Glu206, Val394, Gln395, Leu433,	Lys205, Gln395, Pro396, Glu398, Tyr424, Leu433	Cys13, Leu20, Asn280, Ser281, Cys446, Met450, His451	Arg246, Thr243, Tyr249, Phe250, Phe253, Asp258,

	Pro434			Lys380
Myricetin	Lys205, Glu206, Arg365, Gln395, Pro434, Asp435, Ala436	Lys205, Glu206, Pro434, Ala436	Ser8, His32, Trp164, Arg166, Gln195	Val12, Ile15, His27, Gln28, Asp30, Trp164, Arg166
Naringenin	Lys171, Arg246, Tyr249, Phe250, Phe253, Asp258, Lys380	Lys205, Glu206, Gln395, Glu398, Tyr437	Ser8, Val12, Gln28, Asp30, Trp164, Arg166, Asp194	Thr236, Arg357, Arg370, Arg487, Met496, Phe501, Tyr503, Glu504
Peonidin	Lys403, Arg487, Met496, Tyr503, Glu504, Tyr507	Arg215, Phe216, Arg219, Lys275, Arg348	Gln307, Tyr308, Val309, Leu323, Tyr322, Lys478, Pro481	Glu419, Trp509, Val510, Pro512, His513
Petunidin	Val12, Leu16, Ala25, Ser281, Asp443, His451	His201, Lys205, Glu206, Ala436, Tyr437, Arg439	Gln307, Tyr308, Val309, Gly310, Gly321, Tyr322, Leu323, Lys478, Pro481	Arg357, Arg370, Lys403, Arg487, Phe501, Glu604
Pterostilbene	His201, Lys205, Glu206, Arg385, Gln395, Ala436	Arg393, Tyr401, Glu419, Tyr507, Trp509, Pro512	Glu417, Glu419, Lys508, Trp509, Pro512, Lys514	Val5, Arg192, Glu193, Asp194
Resveratrol	Tyr401, Trp509	Pro144, Thr145, Glu148, Gly174,	Gln11, Val12, Phe26, His27, Trp164, Arg166	Val5, Arg9, Glu193, Asp194, Ser281,

		Leu183, Phe253		Asp282, Gln447, Gln449
Rosmarinic acid	Val5, Ala6, Arg9, Val12, Asp194, Ser281, Asp282, Glu449, Arg454	Arg370, Gln372, Glu389, Tyr401, Lys403, Glu504, Gln502, Tyr503, Tyr507, Trp509	Arg9, Ser188, Glu193, Arg454, Ser455	Gln307, Tyr308, Val309, Tyr322, Leu323, Lys478, Pro479
Salicyclic acid	Lys238, Arg357, Lys366	Arg215, Phe216, Gln222, Trp225	---	Asp194, Arg285, Asp282
Tangeritin	Cys13, Leu16, Arg17, Asn280, Ser281, Leu442, Asp443, His451	Lys366, Arg393, Tyr401, Asp421, Thr423, Pro512	Phe241, Glu244, Tyr249, Phe250	Lys171, Tyr249, Phe250, Phe253, His263, Lys360
Zeaxanthin	Tyr401, Lys403, Glu419, Asp421, Tyr503, Tyr507, Trp509	Ile234, Thr236, Lys238, Arg357, Arg370, Phe501, Tyr507	Ile220, Phe221, Ile224, Asn229, Phe373, His374 Asn388	Ile234, Lys238, Arg357, Lys403, Met496, Tyr503, Tyr507, Trp509

4.7.7.1. Molecular docking of G6PD Orissa with Myricetin

Myricetin showed a binding affinity of -8.0kcal/mol with G6PD Orissa. Figure 37a and 37b shows the 3D and 2D interaction, respectively between the variant Orissa and Myricetin. Hydrogen bonding interactions were observed with Glu239, Arg365, Pro434 and Ala436 of Myricetin with bond lengths 3.16, 3.12, 2.74 and 3.05 respectively. Hydrophobic interactions with His201, Lys205, Glu206, Val394, Gln395, Glu398 and Asp435 were observed.

4.7.7.2. Molecular docking of G6PD Kalyan-Kerala with Apigenin

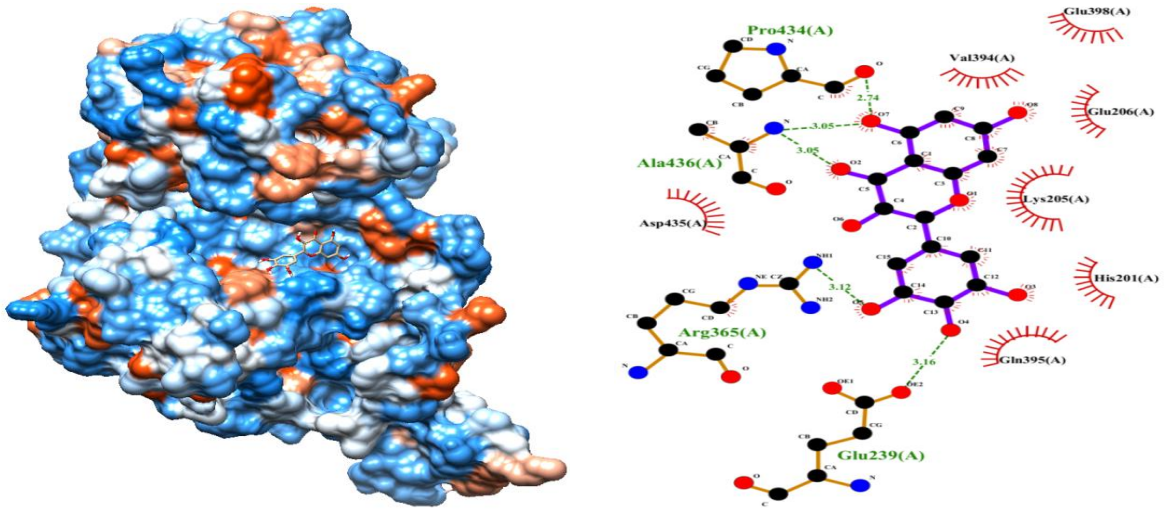
Molecular docking showed apigenin to have a binding affinity of -7.5kcal/mol with G6PD Kalyan-Kerala. The 3D and 2D interaction between Kalyan-Kerala and Apigenin is displayed in Figure 38a and 38b. Apigenin showed two hydrogen bonding interactions with Thr243 with bond lengths of 2.85 and 3.21. Another hydrogen bonding interaction was observed with Glu244 with a bond length of 2.99. Hydrophobic interactions with Phe241, Arg246, Tyr249, Phe250, Phe253 and Asp258 were also seen.

4.7.7.3. Molecular docking of G6PD Mahidol with Catechin

The interaction of Catechin with Mahidol with binding affinity of -7.1kcal/mol is shown in Figure 39a and 39b. Hydrogen bonding interactions of Catechin were observed with His201, Glu206, Pro434 and Tyr437. The bond lengths of the hydrogen bonds were 3.25, 3.10, 2.82 and 3.15 respectively. Catechin formed hydrophobic interactions with Lys47, Lys205, Gln395, Glu398 and Ala436.

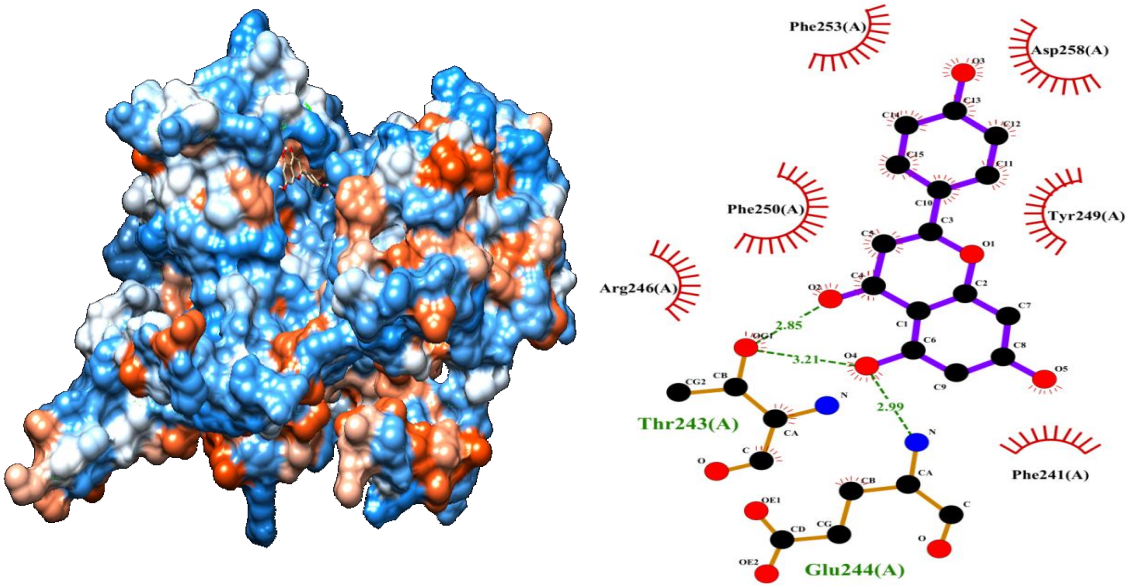
4.7.7.4. Molecular docking of G6PD A^+ with Daidzen

Figure 40a and 40b represents the 3D and 2D interaction of A^+ with Daidzen. Molecular docking revealed a binding affinity of -7.4kcal/mol with G6PD A^+. Only one hydrogen bond was formed with Leu142 of A^+ with a bond length of 2.98. Daidzen formed hydrophobic interactions with Leu43, Ala141, Pro144, Glu170, Lys171, Tyr249, Phe250, Phe253 and Asp258.



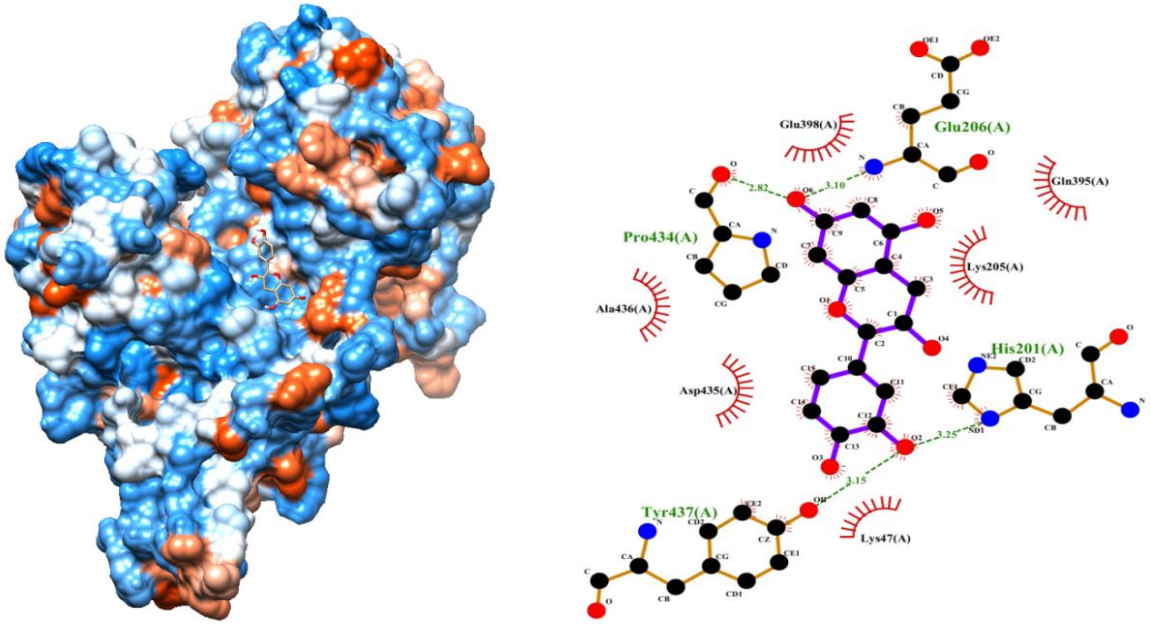
(a) (b)

Figure 37. 3D and 2D visualization of docking of Orissa and Myricetin. (a) 3D interaction between Orissa and Myricetin. (b) 2D interaction between Orissa and Myricetin.



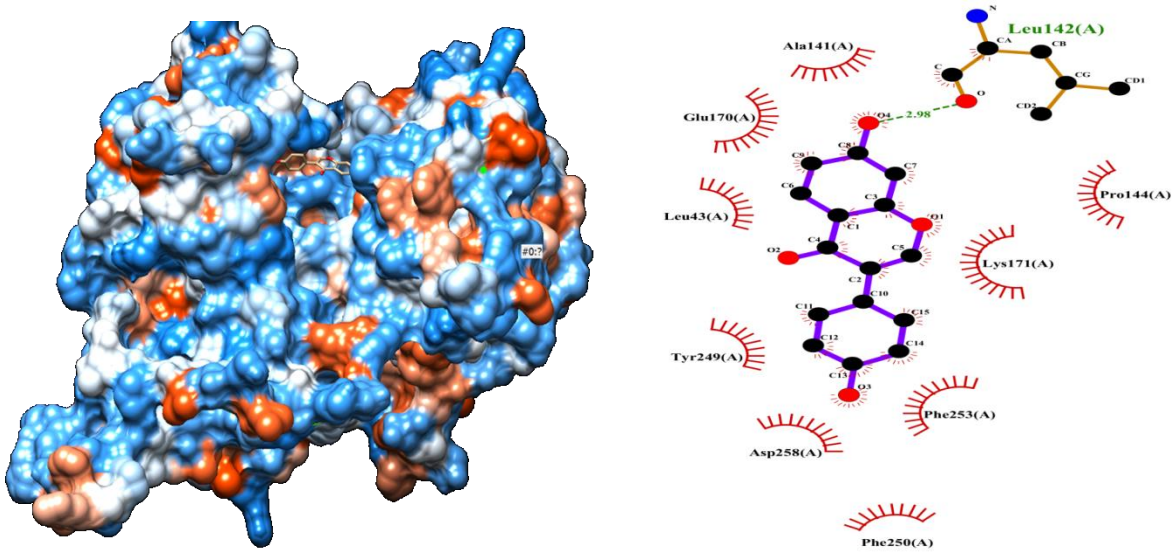
(a) (b)

Figure 38. 3D and 2D visualization of docking of Kalyan-Kerala and Apigenin. (a) 3D interaction between Kalyan-Kerala and Apigenin. (b) 2D interaction between Kalyan-Kerala and Apigenin.



(a) (b)

Figure 39. 3D and 2D visualization of docking of Mahidol and Catechin. (a) 3D interaction between Mahidol and Catechin. (b) 2D interaction between Mahidol and Catechin.



(a) (b)

Figure 40. 3D and 2D visualization of docking of A^+ and Daidzen. (a) 3D interaction between A^+ and Daidzen. (b) 2D interaction between A^+ and Daidzen.

4.7.8. Molecular Dynamics simulation

The Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Radius of Gyration (RG) and Solvent Accessibility Surface Area (SASA) for the four variants, viz., Orissa, Kalyan-Kerala, Mahidol and A^+ with reference to the WT G6PD protein were studied in molecular dynamics (MD) simulation. RMSD helps in understanding the equilibration of the simulated trajectory throughout the simulation. Higher RMSD implies greater deviation throughout the simulation. RMSF is a means for evaluation of differences in flexibility among the residues. Higher RMSF of the variants indicates relatively greater movements of residues with their average position. The RG provides information on the compactness of a protein structure during simulation. Increasing RG values indicates that the mutation has affected the protein structure. SASA is used to analyze whether the amino acid residues of a protein remain buried or exposed. Higher SASA scores indicate that more of the protein molecule is sticking out into water, whereas lower value indicates that more of the molecule is buried in the protein.

4.7.8.1. Molecular Dynamics simulation of Glucose-6-Phosphate Dehydrogenase variants

The overall structural stability of the WT protein and the variants are shown in Figure 41. An increasing RMSD upto 0.5 nm in the trajectory was observed in the WT protein which was stabilized at around 0.43nm after 65ns. Out of the four variants, Kalyan-Kerala and A^+ showed larger deviation compared to the WT protein. Fluctuating RMSDs of 0.5-0.6nm were observed in the Kalyan-Kerala variant which was stabilized at around 0.6nm after 77ns. Orissa showed lower RMSD than the WT protein between 0.4-0.5nm at 30-65ns but increased after 70ns. Increasing RMSD upto 0.55nm was observed in Mahidol which was after a while stabilized at around 0.4nm after 40ns.

RMSF for the backbone atoms of WT protein and the variants were studied to evaluate the changes in the enzyme's residual dynamics (Figure 42). The region of the protein between the residues 310-450 showed higher RMSF values. This region includes

the mutation 317 Glu→Lys (Kalyan-Kerala) indicating high fluctuations in the variant. The residues 80-140 also showed high RMSF values, in which lies the mutation 126 Asn→ Asp (A^+). The RMSF for the variants showed highest fluctuation upto 1.5nm in the residues 510-515 of G6PD A^+, compared to the other variants. Such increase in the flexibility indicates instability of the residues.

In Orissa and A^+, the RG values were similar to the WT protein upto 40000ps between 2.4-2.5nm, but increased after 60000ps and stabilized after 80000ps (Figure 43). Kalyan-Kerala and Mahidol showed larger RG values throughout the trajectory, which was between 2.5-2.6nm till 35000ps. Kalyan-Kerala was seen to have further higher RG after 50000ps which was stabilized after 80000ps at 2.5nm. The RG value of Mahidol declined after 50000ps and stabilized after 75000ps at 2.5nm.

All the four variants showed higher SASA throughout the trajectory (Figure 44). The WT protein showed a stable SASA value of 240nm^2 between 10-70ns, which was reduced and stabilized to 230nm^2 after 75ns. In all the four variants, it was increased to 265nm^2 till 75ns and reduced thereafter. Higher SASA scores indicate that greater portion of the variants is sticking out into water, decreasing the area that is accessible to the ligand or other proteins.

The overall Mean ± SD values of RMSD, RMSF, RG and SASA of the WT protein and the variants are presented in Table 25. G6PD Orissa showed a Mean ± SD RMSD value of 0.41 ± 0.06 which was equal to that of the WT protein. Mahidol and A^+ were found to have a Mean ± SD RMSD value of 0.46 ± 0.05 and 0.51 ± 0.05, respectively. Kalyan-Kerala was shown to have a higher Mean ± SD RMSD value of 0.57 ± 0.05. . G6PD Mahidol (0.18 ± 0.10) was seen to have similar Mean ± SD RMSF to that of the WT protein (0.18 ± 0.12). A lower Mean ± SD RMSF of 0.17 ± 0.11 was seen in Kalyan-Kerala, whereas, higher Mean ± SD RMSF of 0.19 ± 0.12 and0.19 ± 0.17 was observed in Orissa and A^+, respectively. A Mean ± SD value of RG 2.44 ± 0.01 was observed in the WT protein, which was increased in all the four variants, 2.45 ± 0.02 in Orissa, 2.52 ± 0.02 in Kalyan-Kerala, 2.50 ± 0.02 in Mahidol and 2.47 ± 0.02 in A^+. All of the variants showed

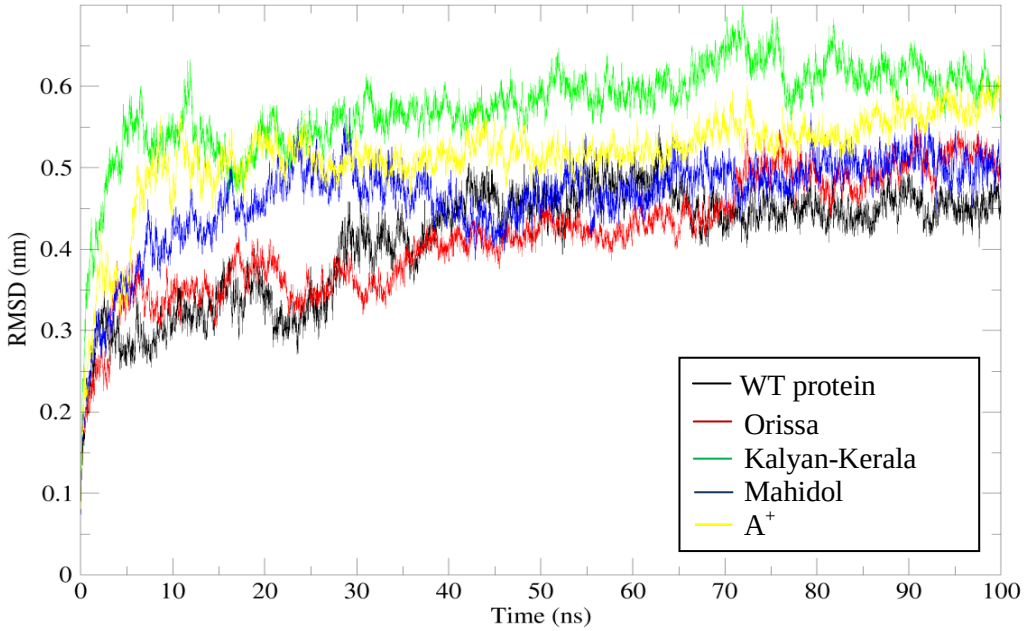


Figure 41. RMSD of the WT protein and the G6PD variants.

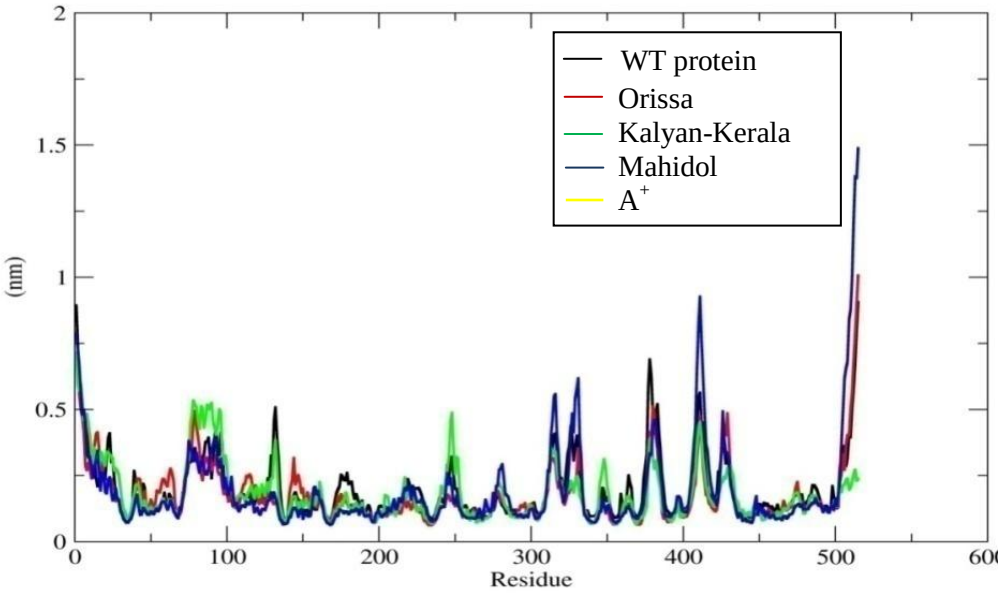


Figure 42. RMSF of the WT protein and the variants.

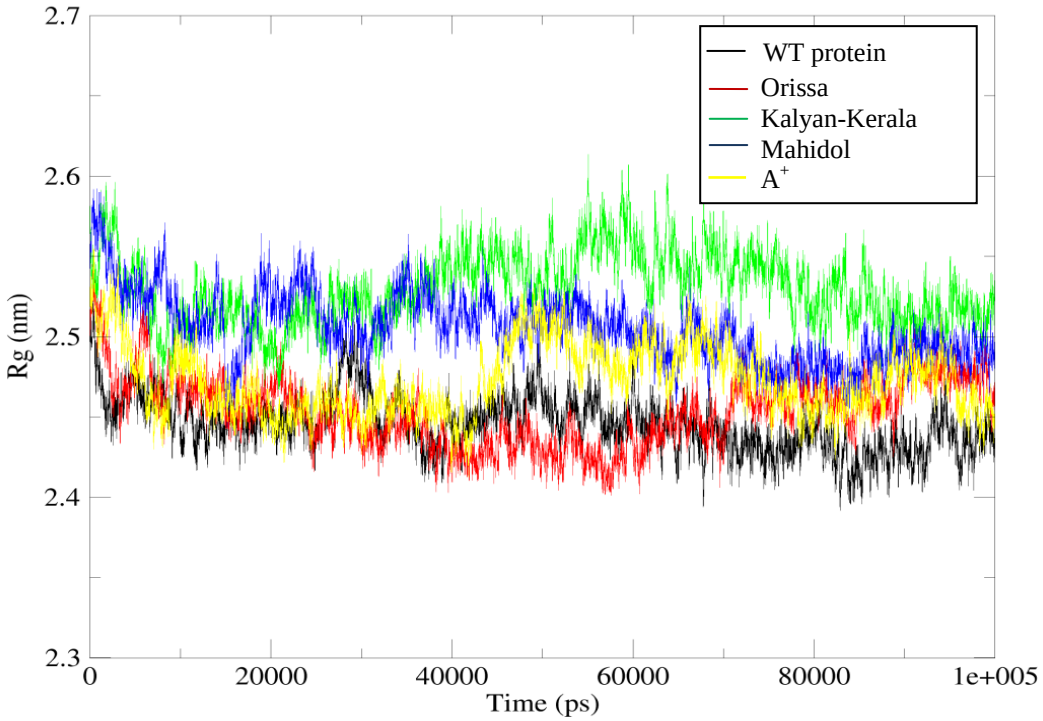


Figure 43. RG of the WT protein and the variants.

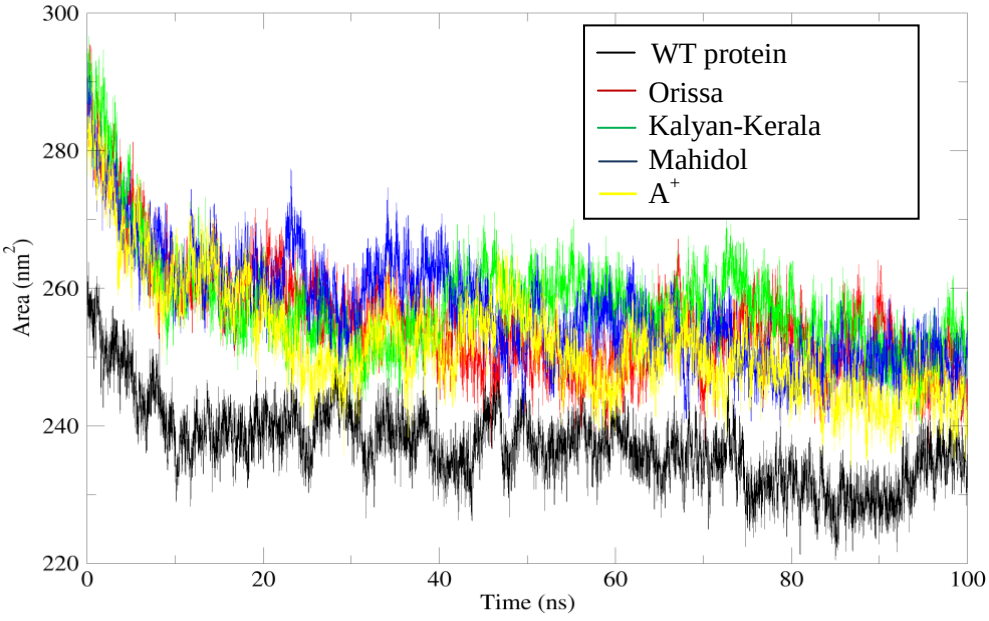


Figure 44. SASA of the WT protein and the variants.

almost similar Mean ± SD SASA values among them (Orissa 255.25 ± 7.89, Kalyan-Kerala 257.81 ± 7.45, Mahidol 256.78 ± 7.95 and A^+ 252.46 ± 8.34) but was higher compared to the WT protein (237.31 ± 5.98).

Table 25. Mean ± SD values of RMSD, RMSF, RG and SASA of G6PD variants.

Parameter	**WT G6PD**	**Orissa**	**Kalyan-Kerala**	**Mahidol**	**A^+**
RMSD (nm)	0.41 ± 0.06	0.41 ± 0.06	0.57 ± 0.05	0.46 ± 0.05	0.51 ± 0.05
RMSF (nm)	0.18 ± 0.12	0.19 ± 0.12	0.17 ± 0.11	0.18 ± 0.10	0.19 ± 0.17
RG (nm)	2.44 ± 0.01	2.45 ± 0.02	2.52 ± 0.02	2.50 ± 0.02	2.47 ± 0.02
SASA (nm^2)	237.31 ± 5.98	255.25 ± 7.89	257.81 ± 7.45	256.78 ± 7.95	252.46 ± 8.34

4.7.8.2. Molecular Dynamics simulation of Glucose-6-Phosphate Dehydrogenase variants

The ligand-variant complexes were compared to the WT and the respective variants. The Mean ± SD values of RMSD, RMSF, RG and SASA of the four complexes are depicted in Table 26.

4.7.8.2.1. Molecular Dynamics simulation of G6PD Orissa-Myricetin

The RMSD of the complex Orissa-Myricetin was seen to be higher than the WT protein and the Orissa variant (Figure 45), with a Mean ± SD value of 0.51 ± 0.08, which in case of both WT protein and Orissa was 0.41 ± 0.06. An increasing RMSD was observed till 60ns in the trajectory, after which it was stabilized at 0.45ns, 0.5ns and 0.6ns in WT protein, Orissa and Orissa-Myricetin, respectively. The RMSD trajectory showed larger deviation of Orissa in complex with Myricetin, indicating a decrease in its stability compared to the WT protein and the Orissa variant.

RMSF was seen to be higher in the Orissa variant in residues 80 - 100, which was later found to be similar between residues 150 and 200 in the WT protein, Orissa and

Orissa-Myricetin (Figure 46). Highest fluctuation in Orissa-Myricetin was observed in residues 310-315. The residue in which the mutation causing the Orissa variant lies (40-60), was seen to show similar fluctuations.

The RG scores of the three systems are presented in Figure 47. The WT protein and Orissa showed similar RG values of 2.5nm till 25000ps and fluctuated towards the later part of the trajectory, which was higher (2.5-2.55nm) in the Orissa-Myricetin complex.

In Figure 48, the SASA scores show that both Orissa and Orissa-Myricetin showed similar SASA scores, which was higher than the WT protein. In both the systems, the SASA scores dropped from 290nm^2 to 250nm^2 as evident from the figure. In the WT protein it was observed to decrease from 260 nm^2 to 240 nm^2. Thus, higher SASA of the Orissa-Myricetin provides evidence that lesser portion of the variant's structure is buried in the protein.

4.7.8.2.2. Molecular Dynamics simulation of G6PD Kalyan-Kerala-Apigenin

RMSD calculations of the WT protein, Kalyan-Kerala and Kalyan-Kerala-Apigenin shows that both Kalyan-Kerala and Kalyan-Kerala-Apigenin had overall higher RMSD compared to the WT protein (Figure 49). The RMSD of Kalyan-Kerala was higher throughout the trajectory, but in case of Kalyan-Kerala-Apigenin and WT protein, similar RMSD of 0.5nm were observed between 40-60ns. Thereafter, RMSD of Kalyan-Kerala-Apigenin increased and stabilized at 0.55nm, although it was higher than the WT protein which stabilized at 0.45nm. An overall higher stability was observed in Kalyan-Kerala-Apigenin in comparison to Kalyan-Kerala.

The RMS fluctuations, as shown in Figure 50 shows fluctuations in the residues 300-440, in which the mutation causing Kalyan-Kerala (residue 317) is present. Highest fluctuation of 1nm in Kalyan-Kerala-Apigenin was observed at amino acid 410-415, which was higher compared to both the WT protein and Kalyan-Kerala. The RG scores of the

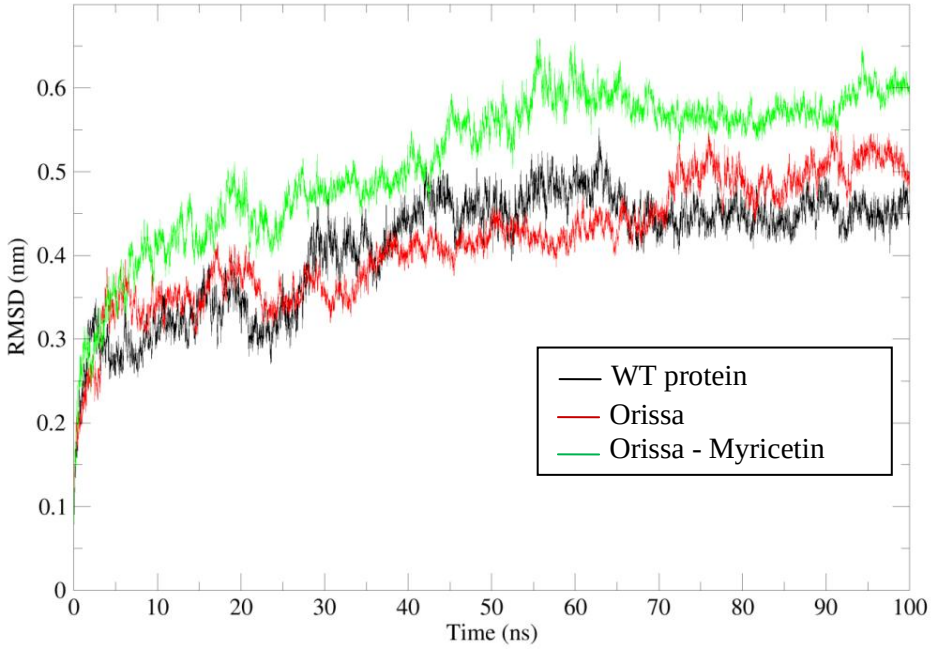


Figure 45. RMSD of the WT protein, Orissa and Orissa complexed with Myricetin.

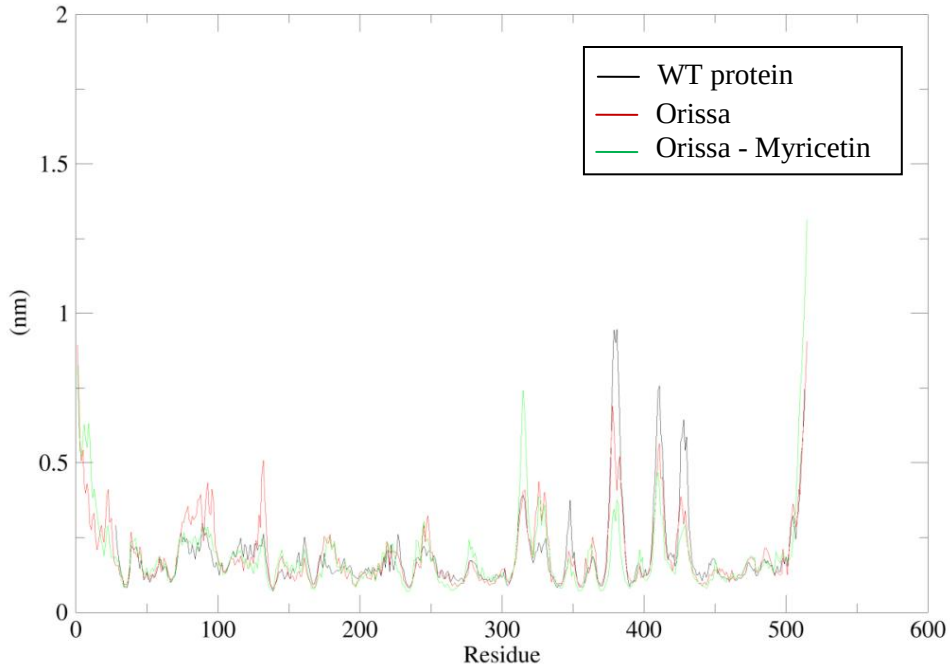


Figure 46. RMSF of the WT protein, Orissa and Orissa complexed with Myricetin.

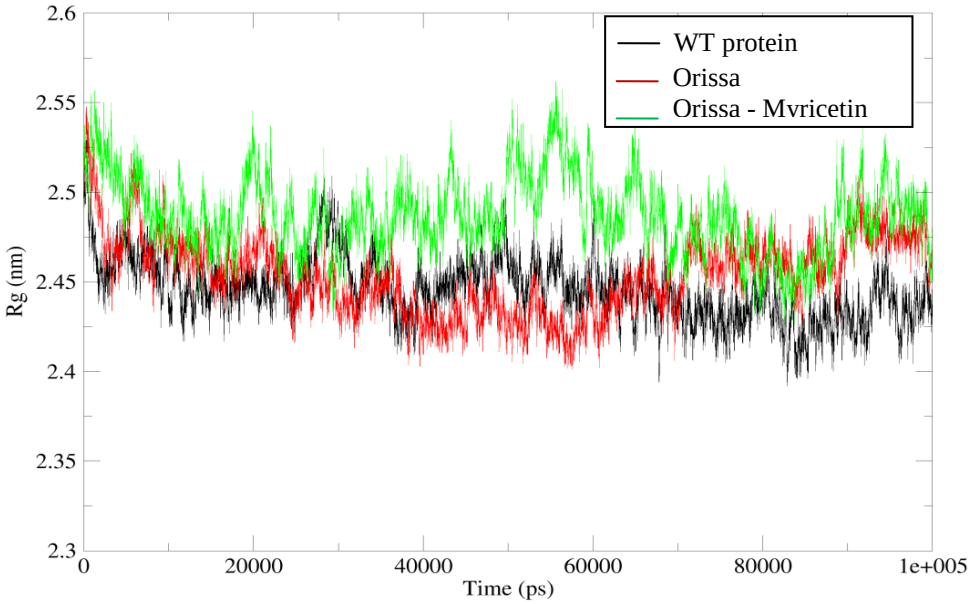


Figure 47. RG of the WT protein, Orissa and Orissa complexed with Myricetin.

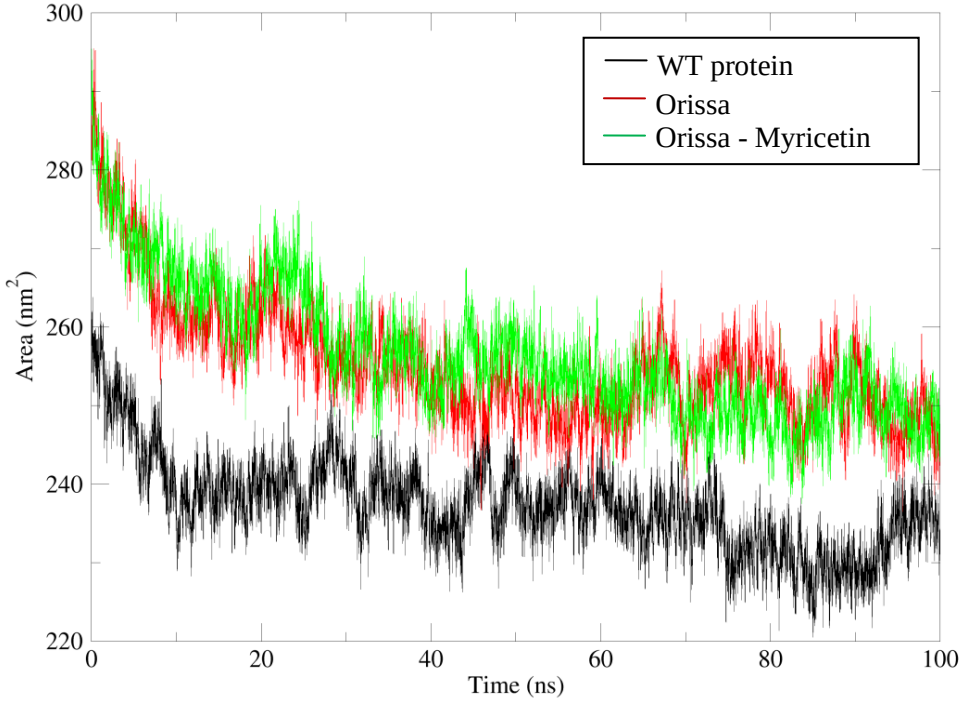


Figure 48. SASA of the WT protein, Orissa and Orissa complexed with Myricetin.

three systems, WT protein, Kalyan-Kerala and Kalyan-Kerala-Apigenin showed similar scores in the WT protein and the Kalyan-Kerala-Apigenin complex (Figure 51). Kalyan-Kerala showed an RG score upto2.6nm, but when complexed with Apigenin, it was reduced to 2.45-2.5nm, which was similar to the WT protein. This result is an indication that the compactness of Kalyan-Kerala was increased when it was complexed with Apigenin.

Calculation of the SASA values, as shown in Figure 52 shows an increase in both Kalyan-Kerala and Kalyan-Kerala-Apigenin. Similar SASA was observed in the two systems till 70ns, after which it was seen to be lower in Kalyan-Kerala-Apigenin in comparison to the Kalyan-Kerala. Although it was lower in the Kalyan-Kerala-Apigenin complex, it was still higher in comparison to the WT protein which showed SASA score less than 240nm^2 between 75-100ns.

4.7.8.2.3. Molecular Dynamics simulation of G6PD Mahidol-Catechin

RMSD analysis of the three systems, WT protein, Mahidol and Mahidol-Catechin showed slight higher deviations in the Mahidol and Mahidol-Catechin complex (Figure 53). The RMSD of the WT protein increased upto 0.5nm till 65ns, after which it was lowered to 0.45nm and stabilized thereafter. However, in case of the other two systems, Mahidol and Mahidol-Catechin, similar higher RMSD of 0.5nm was observed till 40ns, which then further increased in the Mahidol-Catechin complex.

Analysis of RMSF is shown in Figure 54. In the figure, higher fluctuation was observed in Mahidol in the residues 65-100, which was lowered in Mahidol-Catechin complex similar to that of the WT protein. Similar fluctuations were observed between Mahidol and Mahidol-Catechin at 150-360 residues, which then increased upto 1nm in Mahidol. Although fluctuations were observed in all the three systems, there was not much variation in the overall Mean ± SD values.

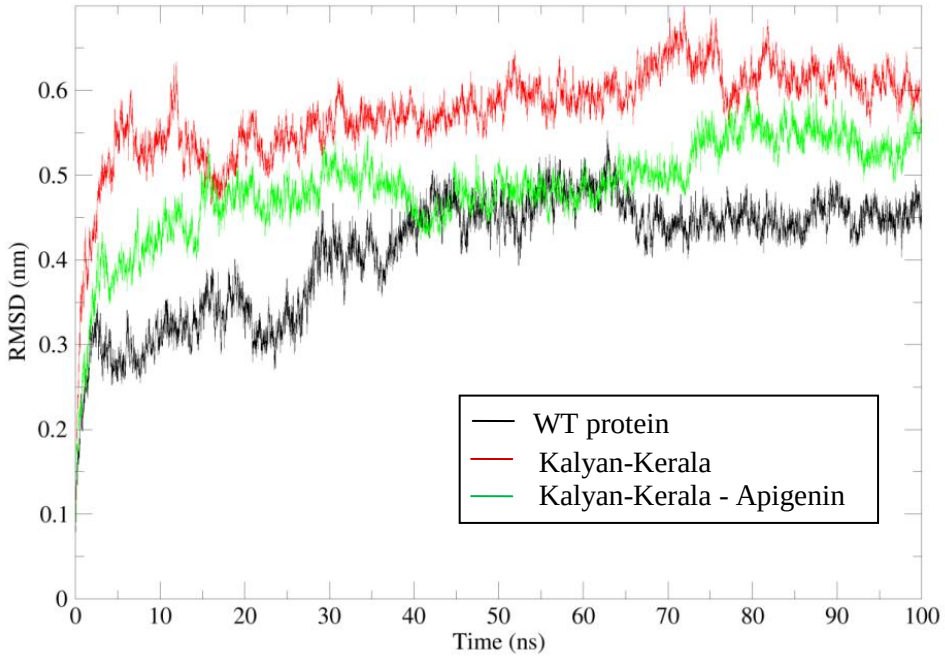


Figure 49. RMSD of the WT protein, Kalyan-Kerala and Kalyan-Kerala complexed with Apigenin.

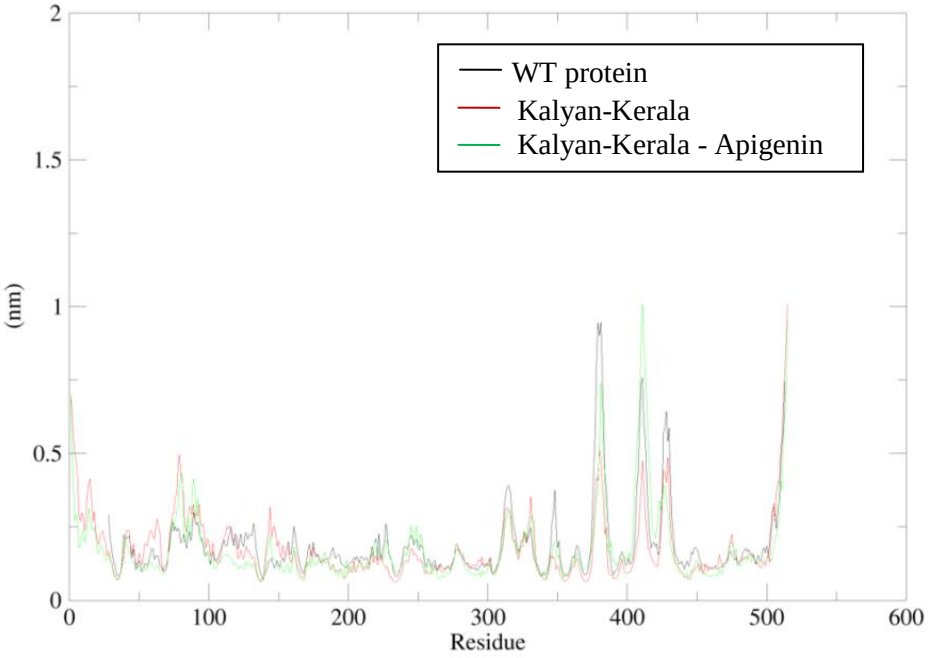


Figure 50. RMSF of the WT protein, Kalyan-Kerala and Kalyan-Kerala complexed with Apigenin.

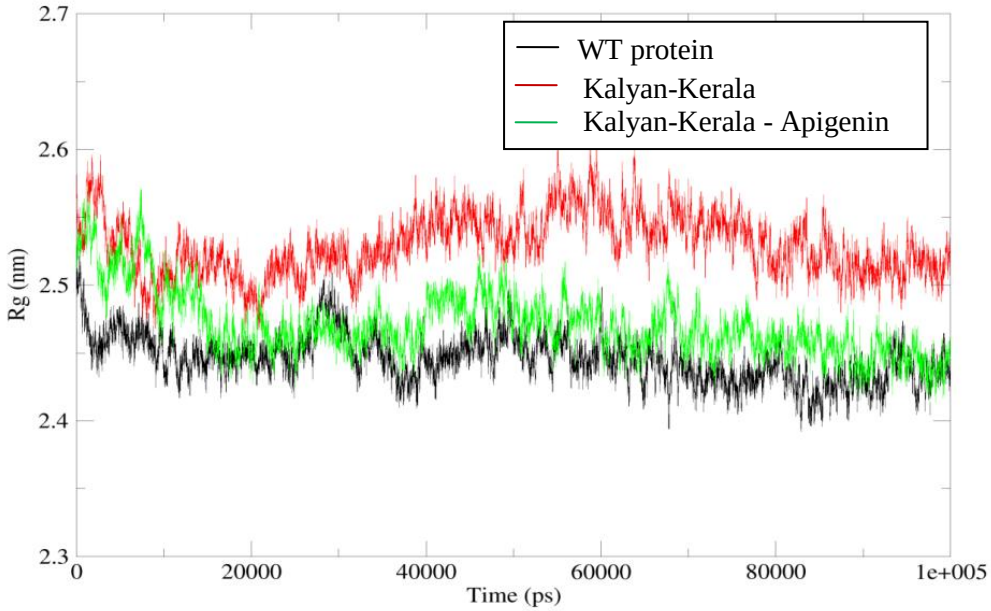


Figure 51. RG of the WT protein, Kalyan-Kerala and Kalyan-Kerala complexed with Apigenin.

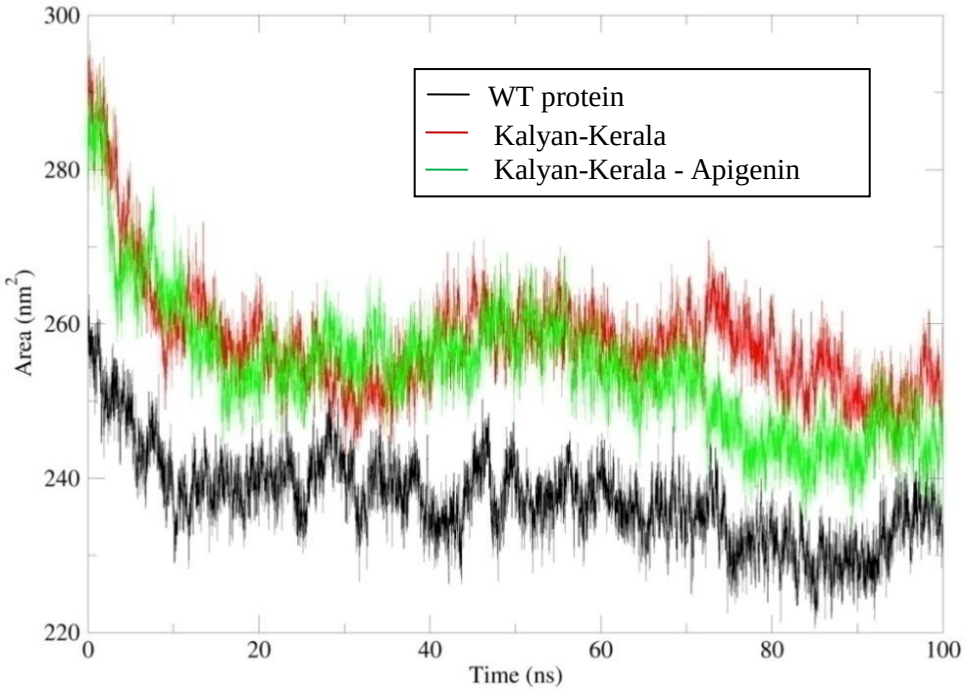


Figure 52. SASA of the WT protein, Kalyan-Kerala and Kalyan-Kerala complexed with Apigenin.

The RG scores, as shown in Figure 55 shows an increased RG in both Mahidol-Catechin and Mahidol-Catechin complex, higher than the WT protein. The RG of the Mahidol variant was similar to the Mahidol-Catechin complex upto 55000ps, after which it was further increased in the Mahidol-Catechin complex. This higher RG in the Mahidol-Catechin indicates a decrease in the compactness of the complex's structure.

Both the Mahidol and Mahidol-Catechin complex showed higher SASA values compared to the WT protein, as shown in Figure 56. The two systems showed a similar SASA score of 250-280nm^2 till 45ns, after which an increase was observed in the Mahidol-Catechin complex.

4.7.8.2.4. Molecular Dynamics simulation of G6PD A$^+$ - Diadzen

RMSD scores of the three systems, WT protein, A$^+$ and A$^+$-Daidzen revealed an increased instability in the A$^+$-Daidzen complex compared to the A$^+$ variant (Figure 57). RMSD of the A$^+$-Daidzen complex was higher throughout the trajectory with a Mean ± SD of 0.65 ± 0.10.

Higher RMS fluctuations were observed in the residues 1-252 of A$^+$-Daidzen compared to the other two systems (Figure 58). Although higher fluctuations were observed in the residues 320-440, it was lower in A$^+$-Daidzen compared to the other two systems. However, the overall Mean ± SD values of SASA were higher than the WT protein and the A$^+$-Daidzen complex.

Increased RG scores were also observed in the A$^+$-Daidzen complex, compared to the WT protein and the A$^+$ variant (Figure 59). In the trajectory, similar RG scores of 2.45nm were observed between the A$^+$ and A$^+$-Daidzen, which was higher in A$^+$-Daidzen complex. Although it was higher, it was stabilized at 2.55nm after 75000ps.

The SASA scores, as shown in Figure 60 showed similar values between A$^+$ and A$^+$-Daidzen, although it was higher than the WT protein. The two systems showed a

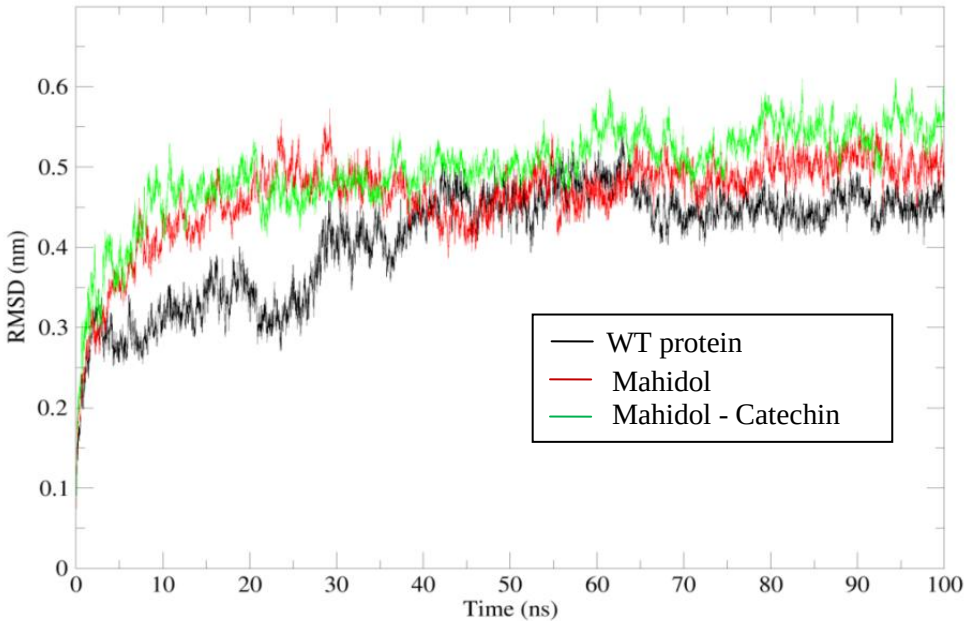


Figure 53. RMSD of the WT protein, Mahidol and Mahidol complexed with Catechin.

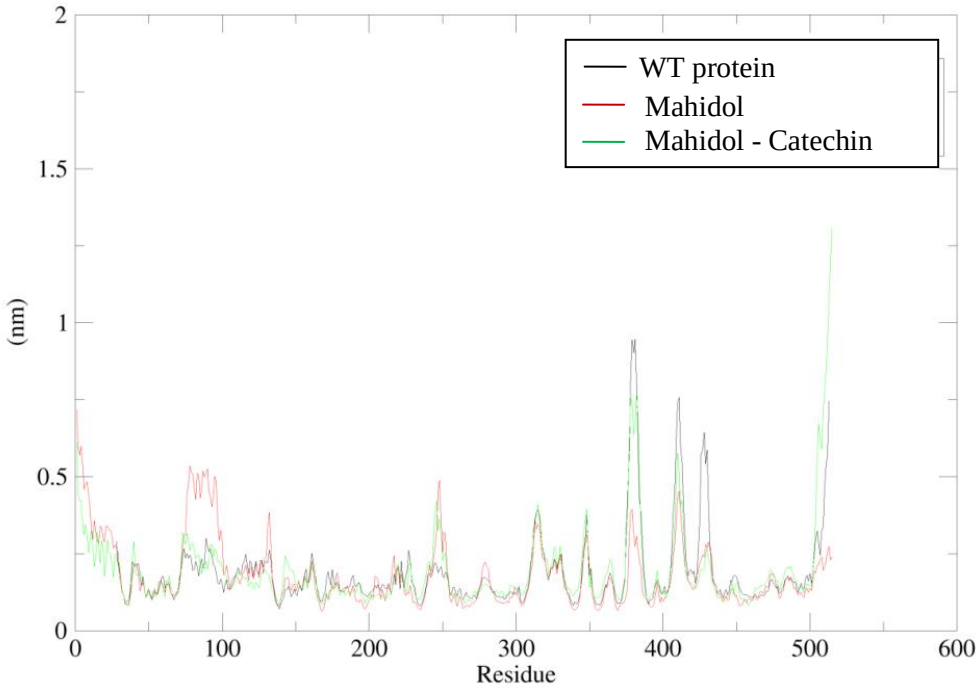


Figure 54. RMSF of the WT protein, Mahidol and Mahidol complexed with Catechin.

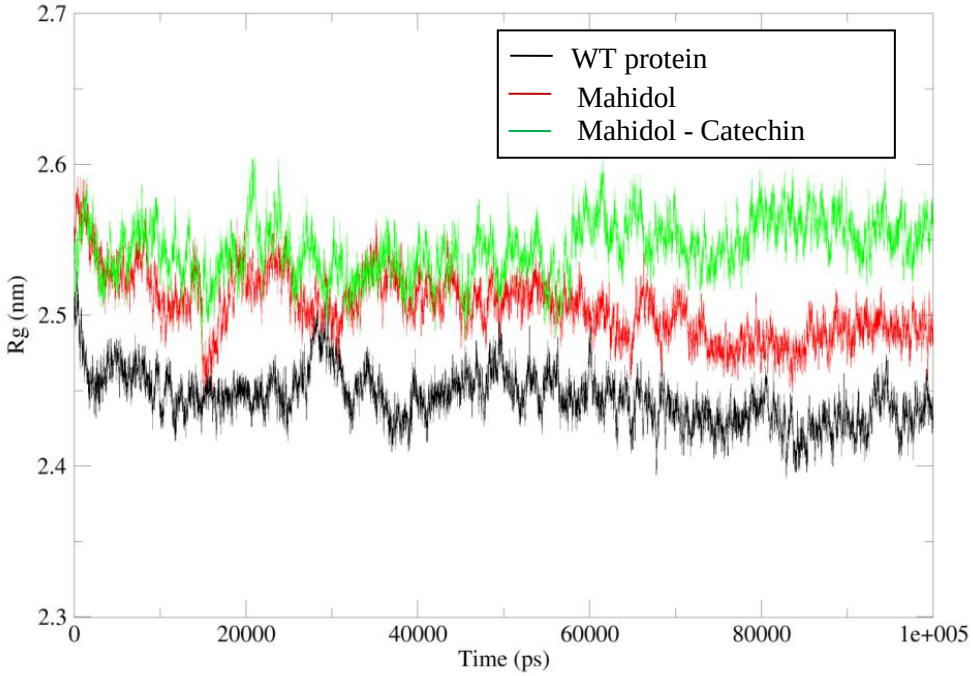


Figure 55. RG of the WT protein, Mahidol and Mahidol complexed with Catechin.

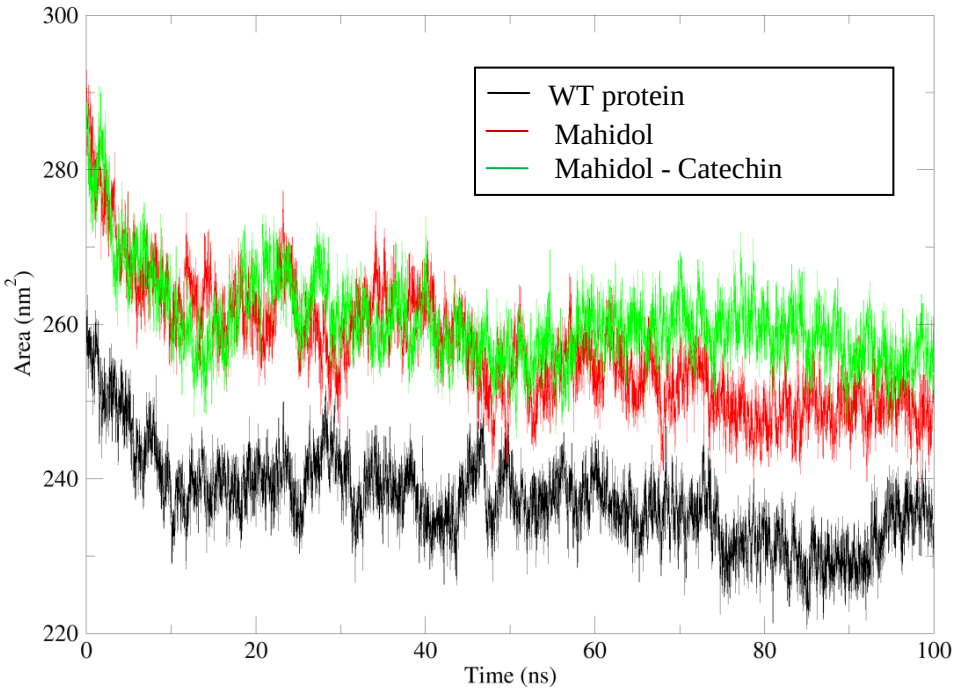


Figure 56. SASA of the WT protein, Mahidol and Mahidol complexed with Catechin.

decreasing SASA value from 280-250nm^2, however it was higher than the WT protein which showed SASA score of 240-260nm^2.

Table 26 presents the overall Mean ± SD values of RMSD, RMSF, RG and SASA of G6PD variant complexed with the ligands. Higher RMSD values were observed in all the four complexes compared to the WT protein, with the highest in A^+-Catechin (0.65 ± 0.10). Slightly higher Mean ± SD values of RMSF were observed in the four complexes, which were 0.19 ± 0.14 in Orissa-Myricetin, 0.17 ± 0.12 in Kalyan-Kerala-Apigenin, 0.19 ± 0.14 in Mahidol-Daidzen and 0.21 ± 0.18 in A^+-Catechin. The Mean ± SD values of RG were also higher in all the four complexes compared to the WT protein. Highest Mean ± SD RG was observed in Mahidol-Daidzen (2.54 ± 0.01), followed by A^+-Catechin (2.53 ± 0.02), Orissa-Myricetin (2.48 ± 0.02) and Kalyan-Kerala-Apigenin (2.47 ± 0.02). The Mean ± SD SASA values were also higher in all the four complexes (256 ± 8.56 in Orissa-Myricetin, 254.29 ± 8.65 in Kalyan-Kerala-Apigenin, 260.37 ± 5.74 in Mahidol-Daidzen and 253.09 ± 8.84 in A^+-Catechin).

Table 26. Mean ± SD values of RMSD, RMSF, RG and SASA of G6PD variant complexed with the ligand.

Parameter	**Orissa - Myricetin**	**Kalyan-Kerala - Apigenin**	**Mahidol - Catechin**	**A^+ - Diadzen**
RMSD (nm)	0.51 ± 0.08	0.48 ± 0.05	0.49 ± 0.05	0.65 ± 0.10
RMSF (nm)	0.19 ± 0.13	0.17 ± 0.12	0.19 ± 0.14	0.21 ± 0.18
RG (nm)	2.48 ± 0.02	2.47 ± 0.02	2.54 ± 0.01	2.53 ± 0.02
SASA (nm$^{2)}$	256.50 ± 8.56	254.29 ± 8.65	260.37 ± 5.74	253.09 ± 8.84

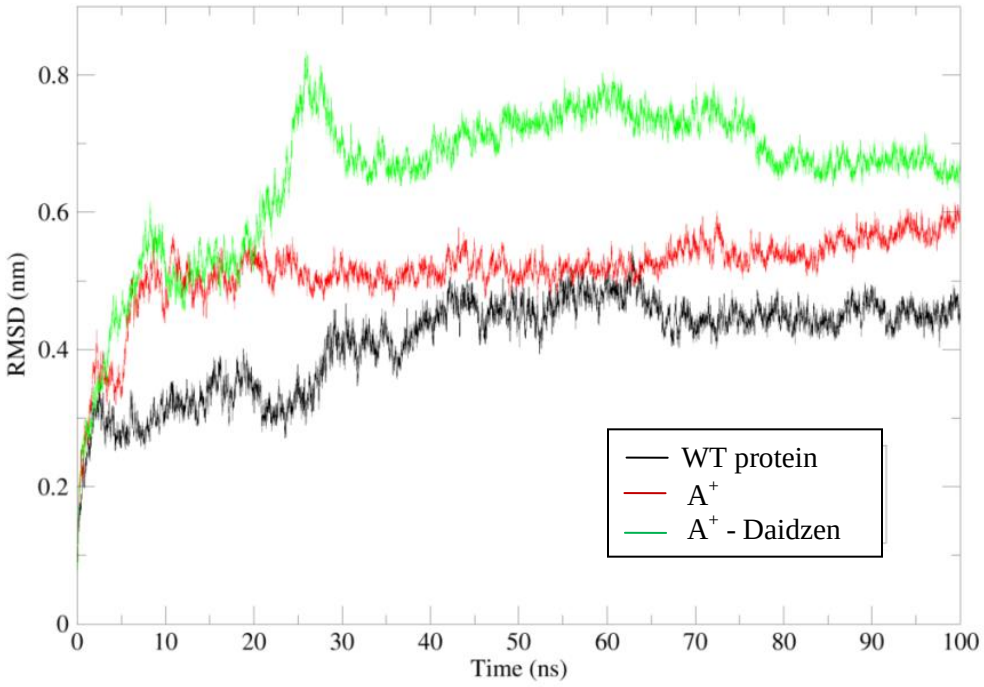


Figure 57. RMSD of the WT protein, A^+ and A^+ complexed with Daidzen.

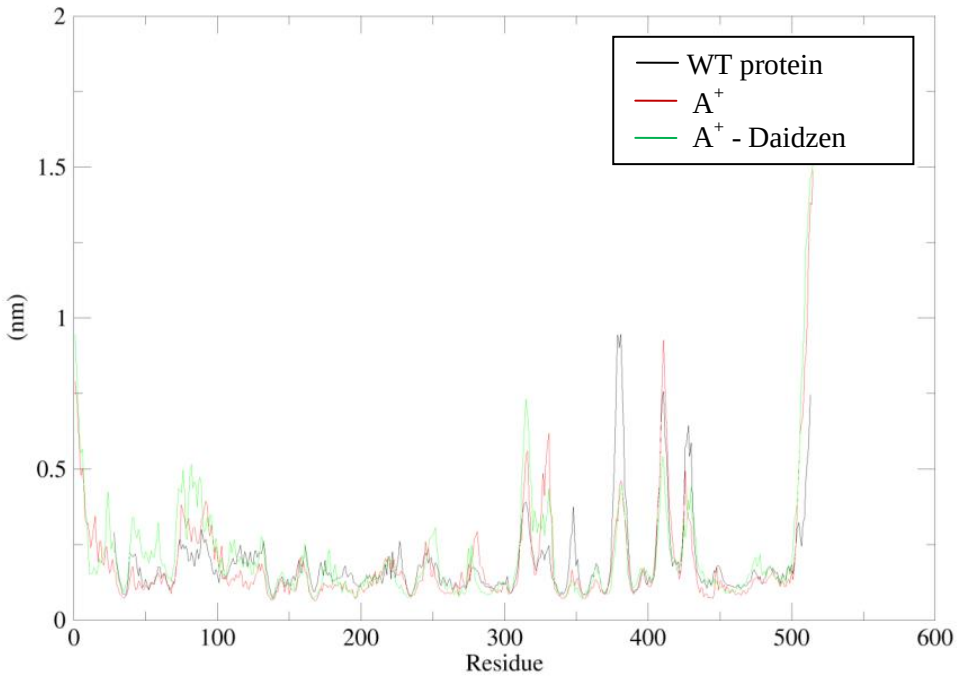


Figure 58. RMSF of the WT protein, A^+ and A^+ complexed with Daidzen.

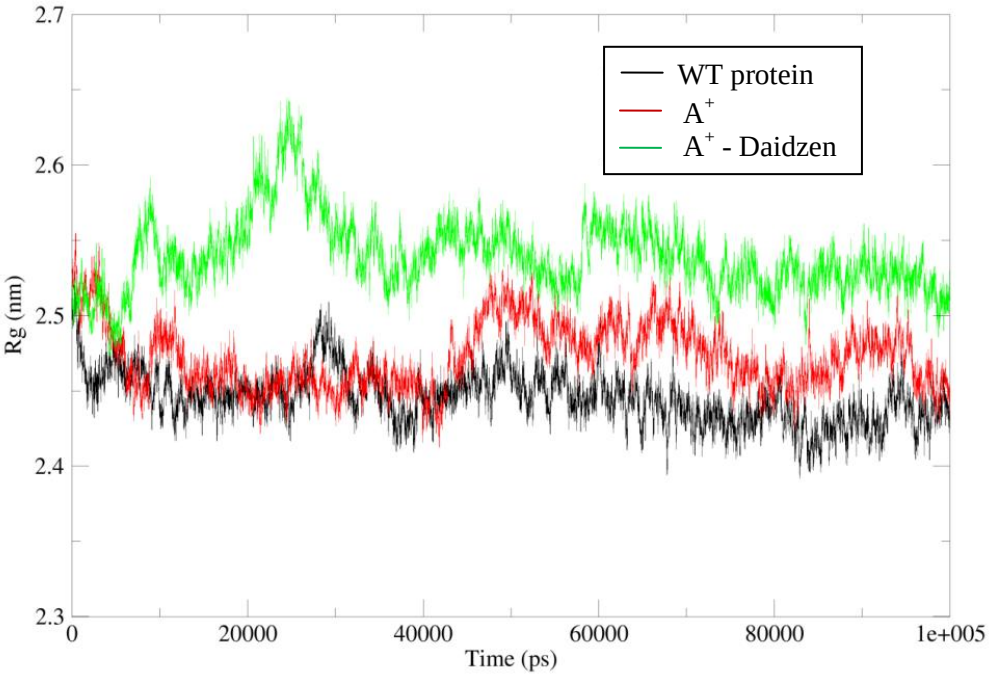


Figure 59. Rg of the WT protein, A^+ and A^+ complexed with Daidzen.

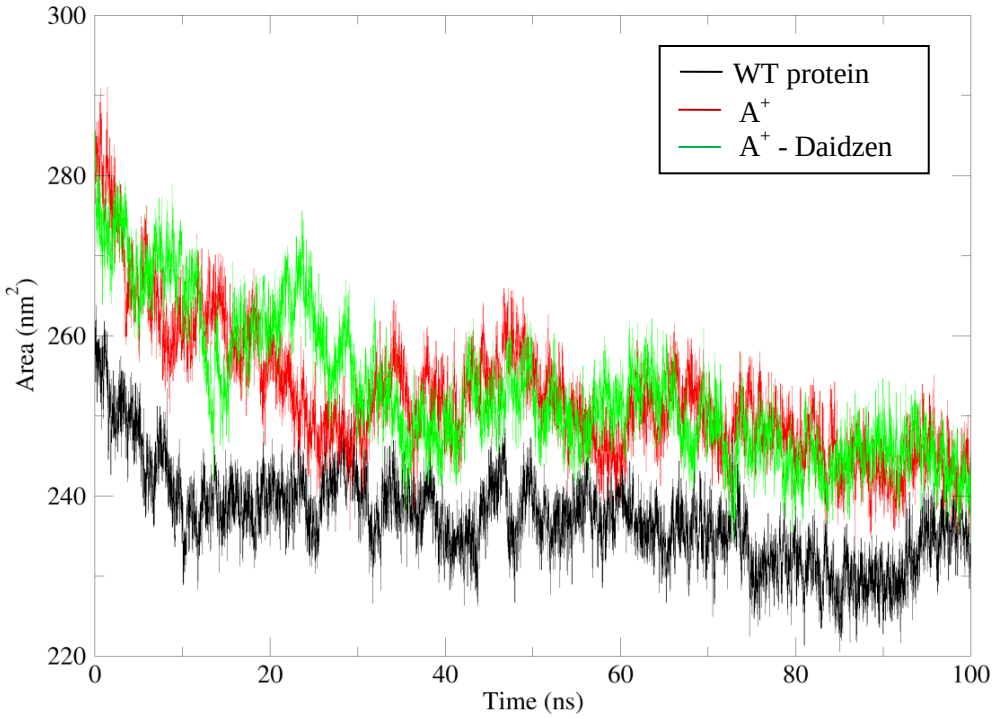


Figure 60. SASA of the WT protein, A^+ and A^+ complexed with Daidzen.