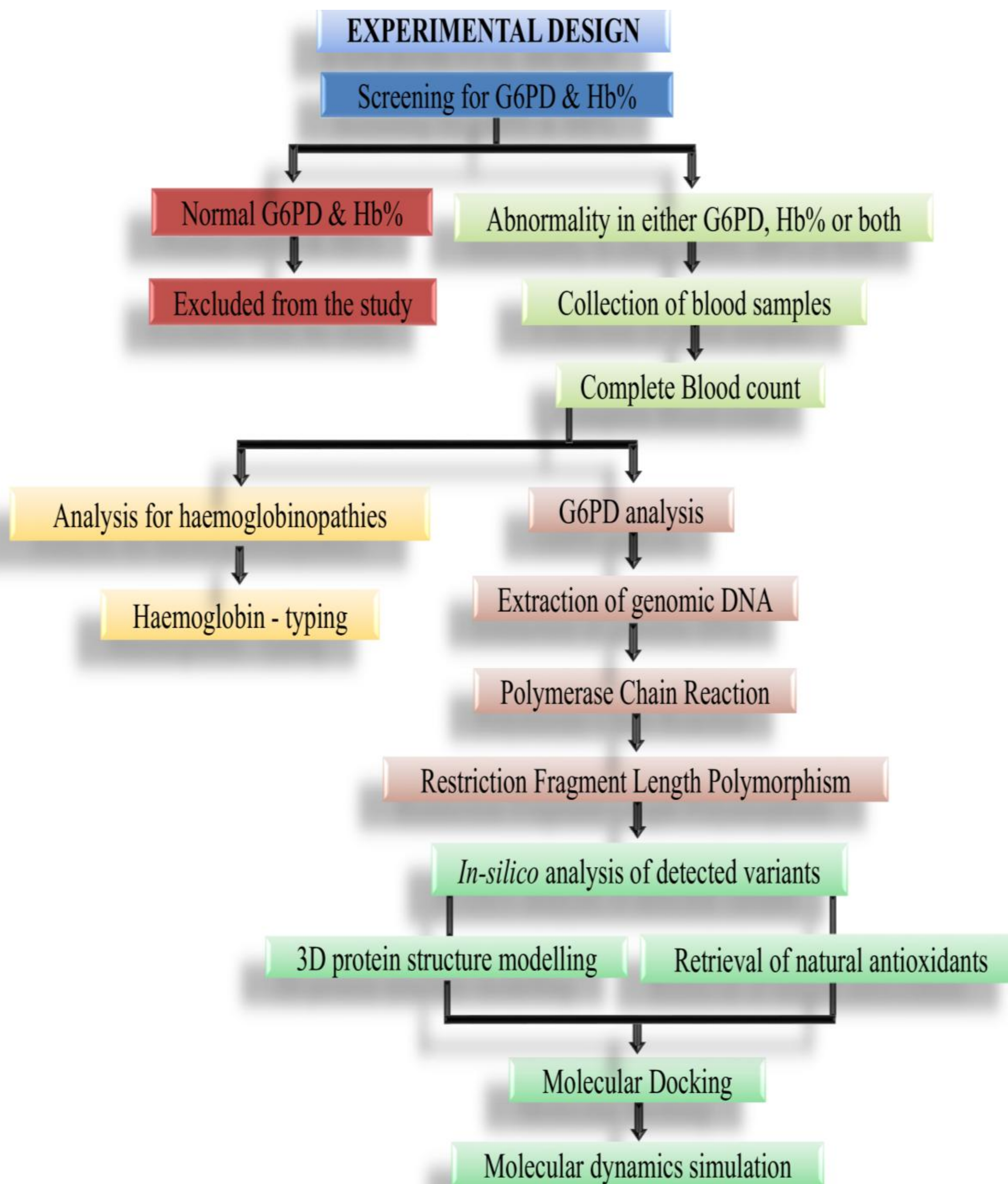
Chapter – 3

MATERIALS AND METHOD

Chapter – 3

MATERIALS AND METHOD



3.1. Site of the study

The study was conducted in four districts of Bodoland Territorial Region (BTR) of Assam, India (Figure 6). The districts are viz., Kokrajhar, Chirang, Baksa and Udalguri, located on the northern bank of the Brahmaputra River in the foothills of Bhutan and Arunachal Pradesh. The geographical area lies between 26° 7'12" N to 26° 47' 50" N latitude and 89° 47' 40" E to 92° 18' 30" E longitude, covering an area of 9,612 km^2 with a total population of 3,151,047 (Census of India, 2011). In the districts, the percentage of tribal population is 38%. The major tribes are the Bodos, Rabhas and Garos of Mongoloid race. A sizable population of tea tribe of Proto-Australoid race is also present.

Table 10. District wise villages and tea estates covered in the study.

Sl. No.	District	Villages	Tea estates
1	Kokrajhar	Amguri, Aflagaon, Bhumka, Choutara, Balagaon FV, Bollamjhora FV, Chandrapara, Debargaon, Dologaon, Garopara, Hokmabil, Joymaguri, Joleswari, Kachugaon, Kusumbil, Nayekgaon, Nasraibil FV, Rabhapara, Shiljani, Saraibil FV, Serfanguri.	Kokrajhar Tea Estate, Dalowabari and Banglabari Tea Estate and Mornoi Tea Estate
2	Chirang	Aminpara, Amguri, Ballamguri, Batabari, Beshorbari, Bengtol, Bamungaon, Bandoguri, Birgaon, Chourabari, Dhakingaon, Duttapur, Jaoliabari, Jitpur, Katribari, Kumguri, Hatigar, Khanthalmari, Dahalapara, Dorogaon, Khagrabari, Ouguri, Taikhajhar, Tukrajhar, Ravapara, Sialmari.	Nil.
3	Baksa	Bangalipara, Bangnabari, Barangabari, Barbari, Chepti, Chouraguri, Dangarigaon, Dongpar, Kadamtola, Karamtuli, Koklabari, Majarkhat, Mushalpur, Panimudi, Patkijuli,	Doomni Tea Estate, Nagrijuli Tea Estate and Menoka Tea Estate.

		Polashguri, Punia, Simla, Thamna and Tokankata.	
4	Udalguri	Bahjani, Balisitha, Batabari, Chaibari, Dakhin Kuiabil, Darrangipara, Dewargaon, Dumoruguri, Dwifang, Hapabari, Khachibari, Kalbari, Kundrubil, Merbangsuba, Naharbari, Nalkhamara, Rangamokha, Pakaribari, Tephakhat and Uttar Kuiabil.	Hatigarh Tea Estate, Majuli Tea Estate, Dimakuchi Tea Estate, Orang Tea Estate and Barangajuli Tea Estate.

3.2. Approval from ethical committee

The proposal for the study was submitted and presented before the Institutional Ethics Committee of Bodoland University, Kokrajhar, Assam for ethical approval prior to the commencement of the work. The committee has approved the study vide letter no. IEC/BU/ICMR/2019-2 (Appendix-I).

3.3. Informed consent

Prior to the screening and blood sample collection processes, the work was briefed to the participants and consent was obtained from them or their parents/guardians (Appendix-II).

3.4. Selection of study subjects

For the study, we randomly selected 20 villages and 3-4 tea estates from each of the districts viz., Kokrajhar, Baksa and Udalguri. In Chirang district, no tea estates exist, hence we selected Proto-Australoid population dominated villages from the district with same numbers of subjects. Around 20 adolescents from each village and 20-25 adolescents from each tea estates were selected. From each district, approximately 500 adolescents were screened for their hemoglobin and G6PD status.

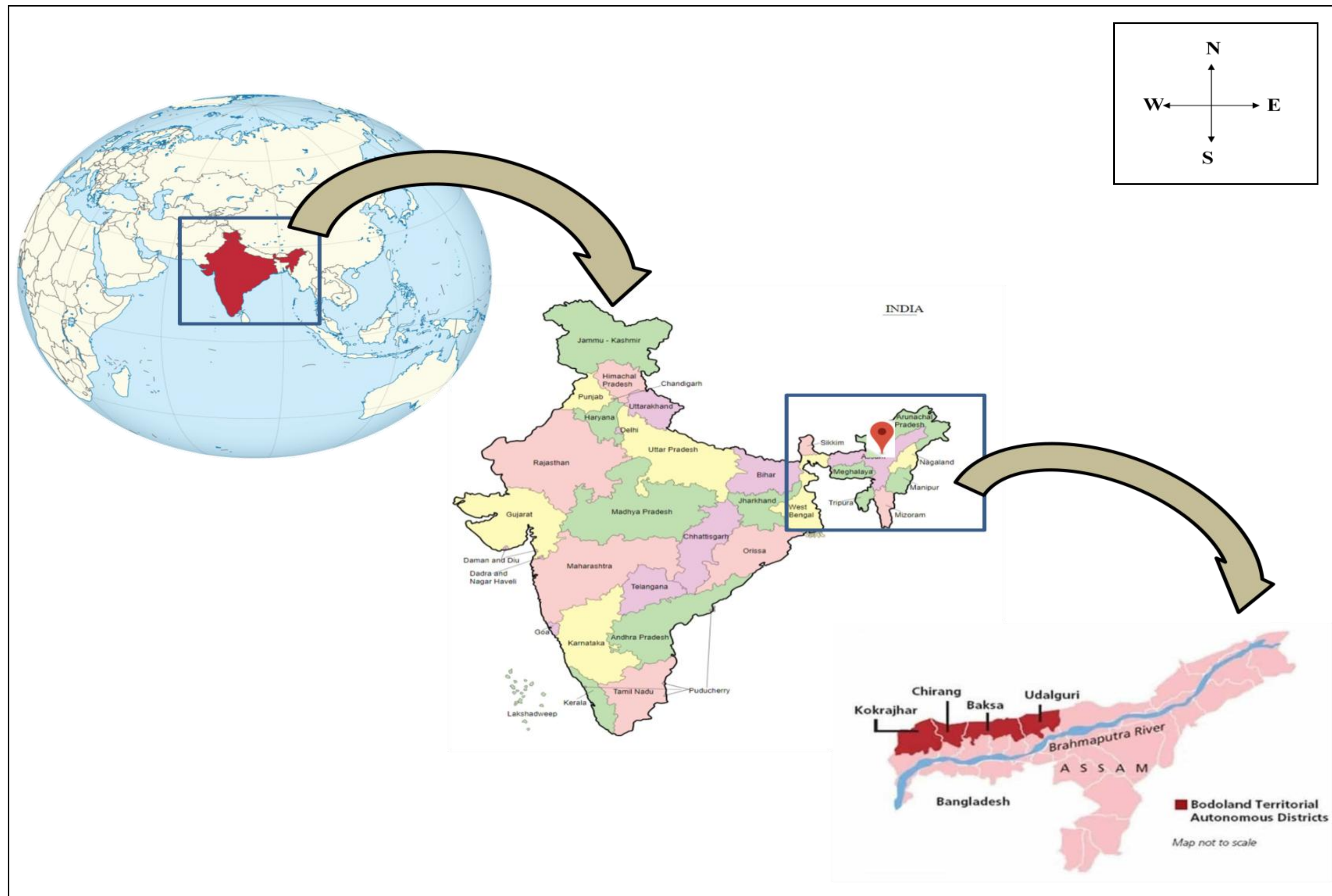


Figure 6. Map of the study site.

3.5. Screening for Glucose-6-Phosphate Dehydrogenase deficiency

G6PD deficiency was detected using STANDARD G6PD Analyzer (SD Biosensors) following the manufacturer's protocol. Briefly, 10µl of blood sample was added to extraction buffer (provided in the kit) and mixed thoroughly. From the mixture, 10µl is applied into the test strip. The device displays the amount of Hemoglobin (Hb) concentration in g/dL and quantitative value of G6PD level in U/g of Hb.

By following WHO recommendations, the status of Hb and G6PD is interpreted as follows:

Table 11. Interpretation of the device's reading as per WHO recommended range.

Sl. No.	Parameter	Interpretation
1	Hb	In male, ≤ 12g/dL → anaemic
		In female, ≤ 11g/dL → anaemic
2	G6PD	> 6U/g Hb → normal G6PD activity
		2-4U/g Hb → intermediate deficiency
		0-2U/g Hb → severe deficiency

3.6. Collection of blood samples

Approximately 2ml of blood samples were collected in heparinized vials from each individual who presented abnormality in either of the two values. The collected blood samples were used for detection of any form of haemoglobinopathy and further molecular analysis of G6PD deficiency.

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

Complete Blood Count (CBC) was done in all collected blood samples using Hematology Analyzer (Sysmex XP-100). Both Red Blood Cell (RBC) and White Blood Cell (WBC) indices along with Platelet counts were examined. The RBC indices were Hb

(g/dL), RBC (Million/Cumm), Mean Corpuscular Volume (MCV in fL), Mean Corpusculur Haemoglobin (MCH in pg), Mean Corpuscular Haemoglobin Concentration (MCHC in g/dL), Packed Cell Volume percentage (PCV) and Red Cell Distribution Width percentage (RDW). The parameters in WBC indices include Leucocyte (Cells/Cumm), Lymphocytes percentage, Monocytes percentage, Eosinophils percentage and Basophils percentage.

3.8. Screening for abnormal haemoglobin variants by Complete Blood Count & Haemoglobin-typing

Based on the CBC results, samples were selected for Hb-typing to detect presence of any forms of haemoglobinopathies. Hb-typing was done using Capillary Electrophoresis (Sebia Minicap).

3.9. Statistical analysis

Statistical analysis of the data obtained from G6PD screening, CBC and Hb-typing were done using SPSS (IBM Corp. released 2019, IBM SPSS for Windows, Version 26.0. Armonk, NY: IBM Corp.).

3.10. Molecular analysis of Glucose-6-Phosphate Dehydrogenase deficiency

3.10.1. Extraction of genomic DNA

Genomic DNA was extracted from the blood samples using QIAamp DNA Blood Mini Kit following their recommended protocol. The kit contains a Protease, Lysis buffer (Buffer AL), Wash buffers (AW1 and AW2) and an Elution buffer (Buffer AE). The protocol for extraction of DNA was as follows:

i. 20µl QIAGEN Protease was pipetted into a 1.5ml microcentrifuge tube and 200µl of sample was added.

ii. 200µl of buffer AL was added to the tube and vortexed for 15 seconds.

iii. The mixture was incubated at 56°C for 10 minutes.

iv. A brief spinning for few seconds was done to remove drops stuck in the inside of the lid.

v. 200µl of ethanol was added and mixed by vortexing for 15 seconds followed by a brief spinning to remove droplets from the lid.

vi. The mixture was transferred into a spin column placed in a 2ml collection tube and centrifuged at 8000 rpm for 1 minute.

vii. The collection tube was discarded and the spin column was placed in a new 2ml collection tube, 500µl of buffer AW1 was added and centrifuged at 8000 rpm for 1 minute.

viii. The collection tube was discarded and spin column was placed in a new 2ml collection tube.

ix. 500µl of buffer AW2 was added and centrifuged at 14000 rpm for 3 minutes.

x. The spin column was placed in a 2ml microcentrifuge tube and the old collection tube was discarded.

xi. 200µl of buffer AE was added and incubated at room temperature for 1 minute, followed by centrifugation at 8000 rpm for 1 minute.

xii. The spin column was discarded and the DNA was eluted.

The extracted DNA was run on 0.8% agarose gel electrophoresis containing ethidium bromide for confirmation and quality check. DNA was then stored in -20°C for further use.

3.10.2. Quantification of DNA

Quantification of DNA was done using UV-VIS Spectrophotometer (Shimadzu UV-1900i). The DNA samples were diluted using nuclease free water and absorbance at wavelengths 260nm (A_{260}) and 280nm (A_{280}) were recorded. A_{260} was recorded to determine the amount of DNA and ratio of A_{260} and A_{280} gives the purity of DNA.

The calculation of DNA concentration and purity was performed as follows:

DNA concentration (µg/ml) = $A_{260} \times 50 \times$ DF (dilution factor)

DNA purity = A_{260} / A_{280}

If $A_{260} / A_{280} \geq 1.8$, it indicates pure DNA.

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

The exonic region of *g6pd* gene was amplified using different primer sets for identification of mutations in the exons.

i) Primers used

The primer sets used for amplification of the exonic regions of *g6pd* were as given in Table 12:

Table 12. Details of primers used for exon-wise amplification of *g6pd* gene.

Exon no.	Primer sequence	Product size	Reference
3	Forward: 5'CAGCCACTTCTAACCACACACCT3' Reverse: 5'AGGGCAGGGCACAGCTGTAA3'	352 bp	Kaeda *et al.*, 1995
4	Forward: 5'CGGACTCAAAGAGAGGGGCT 3' Reverse: 5'GCACTGCCTGGGCCAGCCT3'	314 bp	Sobngwi *et al.*, 2005
4-5	Forward: 5′GTCTTCTGGGTCAGGGAT3′ Reverse: 5′GGAGAAAGCTCTCTCTCC3′	701 bp	Rajkhowa *et al.*, 2020
6-7	Forward: 5'ACTCCCCGAAGAGGGGTTCAAGG3' Reverse: 5'CCAGCCTCCCAGGAGAGAGGAAG3'	545 bp	Kaeda *et al.*, 1995
9	Forward: 5'ACCCAAGGAGCCCATTC3' Reverse: 5'ACACAGGGCATGCCCAGTTCTG3'	276 bp	Matsuoka *et al.*, 2007
10	Forward: 5'CTGAGAGAGCTGGTGCT3' Reverse: 5'CACCATGTGGAGTCCCCCGG3'	342 bp	Matsuoka *et al.*, 2007

11	Forward: 5'GTGAAAATACGCCAGGCCTTA3' Reverse: 5'GTGAAAATACGCCAGGCCTTA3'	214 bp	Charoenkwan *et al*., 2014
12	Forward: 5'GTGAAAATACGCCAGGCCTTA3' Reverse: 5'GTGCAGCAGTGGGGTGAACATA3'	227 bp	Charoenkwan *et al*., 2014

ii) Polymerase Chain Reaction conditions

Amplifications were carried out in 25µl reaction volumes containing 0.2mM of dNTP mix, 250nm each of forward and reverse primers, 2.5 units of Taq Polymerase and 1X PCR buffer containing 1.5mM $MgCl_2$. The temperature profiles for amplification of each of the exons were as provided in the Table 13. Amplification was done for 35 cycles and the products were kept on hold at 4°C until electrophoresis.

Table 13. Standardization of PCR settings for each exon.

Exon no.	**Initial denaturation** (5 min)	**Denaturation** (30 sec)	**Annealing** (45 sec)	**Extension** (45 sec)	**Final extension** (10 min)
3	95°C	95°C	60°C	72°C	72°C
4	95°C	95°C	64°C	72°C	72°C
4-5	95°C	95°C	61°C	72°C	72°C
6-7	95°C	95°C	66.5°C	72°C	72°C
9	95°C	95°C	61.6°C	72°C	72°C
10	95°C	95°C	63°C	72°C	72°C
11	95°C	95°C	68°C	72°C	72°C
12	95°C	95°C	62°C	72°C	72°C

iii) Agarose gel electrophoresis

The PCR products were run on 1.8-2% agarose gel containing ethidium bromide at 100V for 40-45 minutes. The gels were then visualized and photographed using UV-Transilluminator (Bio-Imaging Systems).

3.10.4. Restriction Fragment Length Polymorphism

The PCR products were digested with appropriate restriction endonucleases to detect the presence of mutations in the respective exon. For digestion, the reaction mixture was prepared by adding 5µl of PCR product, 5µl of 10X buffer (provided with the restriction endonuclease), 1µl of restriction endonuclease and 39µl of nuclease free water. The mixture was then incubated in hot water bath at the required temperature. Following incubation, heat inactivation at 65°C for 20 minutes was done. The details of the restriction endonucleases used were as provided in Table 14.

Table 14. Details of restriction endonucleases used for detection of mutation.

Sl. No.	Variant / mutation	Exon no.	RE	IT	Mutated product	Reference
1	Orissa 131C>G	3	HaeIII	37°C	HaeIII site abolished	Kaeda *et al.*, 1995
2	A^- 202G>A	4	NlaIII	37°C	NlaIII site abolished	Rajkhowa *et al.*, 2020
3	Namoru 208T>C	4	NlaIII	37°C	NlaIII site created	Chalvam *et al.*, 2007
4	A^+ 376A>G	5	FokI	37°C	FokI site created	Rajkhowa *et al.*, 2020
5	Mahidol 487G>A	6	Hind III	37°C	HindIII site created	Charoenkwan *et al.*, 2014
6	Mediterranean 563C>T	6	MboII	37°C	MboII site created	El-Gezeiry *et al.*, 2005
7	Acores 595A>G	6	BstUI	37°C	BstUI site created	Hung *et al.*, 2008
8	Kalyan-Kerala/ Jamnagar/ Rohini 949G>A	9	MnlI	37°C	MnlI site created	Devendra *et al.*, 2019

9	Chatham 1003G>A	9	BstXI	55°C	BstXI site created	Gandomani *et al*., 2011
10	Guadalajara 1159C>T	10	HhaI	37°C	HhaI site abolished	Minucci *et al*., 2012
11	Union 1360C>T	11	Hha I	37°C	Hha I site abolished	Charoenkwan *et al*., 2014
12	Canton 1376G>T	12	AflII	37°C	AflII site created	Charoenkwan *et al*., 2014
13	Kaiping 1388G>A	12	NdeI	37°C	NdeI site created	Charoenkwan *et al*., 2014

RE: Restriction endonuclease; IT: Incubation temperature

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

Statistical analysis of the data obtained from G6PD genotyping were done using SPSS (IBM Corp. released 2019, IBM SPSS for Windows, Version 26.0. Armonk, NY: IBM Corp.).

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

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

The three dimensional (3D) structure of WT G6PD was downloaded from RCSB PDB (ID: 6e08). Since the 3D structures of the mutants are not available, we modeled the 3D structures by using I-TASSER (https://zhanggroup.org/I-TASSER/), an online server for Protein structure and function predictions (Zhang, 2008; Roy *et al*., 2010; Yang *et al*., 2015). The server generates models by applying their I-TASSER based algorithms.

The structures generated by I-TASSER were validated by ProTSAV (http://www.scfbio-iitd.res.in/software/proteomics/protsav.jsp), an online protein structure analysis and validation tool (Singh *et al*., 2016). ProTSAV is a meta-server developed for analyzing the quality and precision of the modeled structure by incorporating the following assessment tools: Procheck (https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/) (Laskowski *et al*., 1993), ProSA-Web (https://www.came.sbg.ac.at/prosa.php) (Weiderstein & Sippl, 2007), ERRAT (http://services.mbi.ucla.edu/ERRAT/) (Colovos & Yeates, 1993), Verify3D (https://www.doe-mbi.ucla.edu/verify3d/) (Eisenberg *et al*., 1997), dDFire (https://mybiosoftware.com/ddfire-1-1-dfire-2-monomer-protein-energy-alculation.html) (Yang and Zhou, 2008), Naccess (http://www.bioinf.manchester.ac.uk/naccess/) (Hubbard and Thornton, 1993), MolProbity (https://molprobity.biochem.duke.edu/) (Chen *et al*., 2010), D2N (http://www.scfbio-iitd.res.in/software/d2n.jsp.) (Mishra *et al*., 2014), ProQ (https://proq.bioinfo.se/ProQ/ProQ.html) (Wallner & Elofsson, 2003) and PSN-QA (http://vishgraph.mbu.iisc.ernet.in/GraProStr/native_non_native_ranking.html) (Ghosh & Vishveshwara, 2014). The best model was then selected for further analysis.

3.11.2. Retrieval of naturally available antioxidant compounds from databases

Literature databases such as Pubmed, Pubmed Central, ResearchGate, Google Scholar and Semantic Scholar using “natural antioxidants”, “antioxidants”, “phytochemicals”, “natural compounds” as keywords were searched for natural antioxidant compounds. The 3D conformers of these antioxidants were downloaded from Pubchem in .sdf format.

3.11.3. Preparation of ligands and protein structures for docking

To perform molecular docking using AutoDockTools-1.5.7 (ADT), the supported file format of the ligands is .pdb (Trott & Olson, 2010). So, the ligands were converted from .sdf to .pdb formats using OpenBabel (O’Boyle *et al*., 2011). After generating the .pdb files, the ligands were prepared for docking using AutoDockTools-1.5.7 (ADT).

Ligand preparation includes addition of hydrogens, calculation of Gasteiger charges, detection of root and setting number of torsions. The prepared ligands were saved in .pdbqt formats.

For preparation of target, the modeled 3D structures were opened in ADT. The preparation steps include deleting the water, addition of polar hydrogens, check and repair (if any) missing atoms and addition of Kollman charges. The prepared target structures were saved in .pdbqt formats.

3.11.4. Identification of binding pocket and preparation of grid box

The binding pockets of G6PD were identified using CASTp 3.0 (Computer Atlas of Surface Topography of Proteins), an online resource for identification of binding pockets that are positioned in the interior of proteins (Tian *et al*., 2018). CASTp 3.0 identifies and presents all the possible pockets with information on Molecular Surface (MS) volume, pocket MS area, openings, mouth MS area and MS circumference sum. The pocket with largest MS volume and pocket MS area was selected. The amino acid residues forming this pocket were recorded and the grid parameters were set accordingly.

The grid parameters were set as x, y, z size- 90×90×90 with a spacing of 0.375 Å and centre-coordinates were set such that the interacting residues are covered.

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

Docking was performed using AutoDock Vina, one of the docking engines of the AutoDock Suite. The files required for docking such as configuration file having the grid parameters, number of modes at 20, energy range at 5 and exhaustiveness at 8; ligand file having the list of ligands and Vina execution file (Vina_windows.pl) were prepared. Docking was then performed by executing the commands in command prompt. After docking, the best poses with lowest binding energies were selected for further analysis.

3.11.6. Analysis of Molecular docking results

The best docking poses for each the detected variants were observed using LigPlot+, a program used to generate two dimensional (2D) interaction diagrams for protein-ligand complexes (Laskowski & Swindells, 2011). The 2D protein-ligand interaction diagram displays the hydrogen bond interactions and hydrophobic contacts between the ligands and the amino acid residues of the protein.

3.11.7. Analysis of drug-likeness and ADMET properties

The drug-likeness of the selected compounds was studied using SwissADME, an online tool for studying pharmacokinetics, drug-likeness and medicinal prospect of small molecules (Daina *et al*., 2017). The drug-likeness property of the compounds was evaluated based on Lipinski's rule. Analysis of ADMET properties includes Absorption, Distribution, Metabolism, Excretion and Toxicity.

Based on the binding energy and drug-likeness study, the best complex was selected and visualized using ADT and PyMOL (Schrodinger & Delano, 2020). In order to validate docking study and to understand the changes in the protein conformation, the complexes were subjected to molecular dynamics simulation using GROMACS 2022.1 (Bekker *et al*., 1993).

3.11.8. Molecular Dynamics simulation

To evaluate the structural deviation of the variants and to validate the docking results, we performed molecular dynamics simulation using GROMACS 2022 package. For simulation purposes, a simulation system was set up using GPU-enabled GROMACS 2022 package in a HP ProLiant Gen9 server. The GROMOS 54a7 force field was added to the system and the SPCE water system, along with a Cubic box, was used to model the solvent. The water-filled box was added with 1.5mM of NaCl ions to neutralize the systems. The system was then used for simulation time set up for 100 ns under normal NTP i.e., constant number of particles (N), pressure (P), and temperature (T) conditions where the model system was relaxed before the simulation. The simulation interactive diagram tool further

analyzed the simulation data in a graphical way, which provided information of protein-ligand complex properties during the simulation time. Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Radius of Gyration (RG) and Solvent Accessibility Surface Area (SASA) data were analyzed.