Proceedings of Science and Mathematics
Faculty of Science, Universiti Teknologi Malaysia

Vol. 6, 2022, Page 28-33

# Computational Analyses of Glucose–6–Phosphate Dehydrogenase Variants

**Naveen Eugene Louis[a], Muaawia Ahmed Hamza[b,c], Puteri Nur Sarah Diana Engku Baharuddin[a], Shamini Chandran[a], Nurriza Ab Latif[a], Mona Awad Alonazi[d] Joazaizulfazli Jamalis[e], Arjumand Warsy[f], and Syazwani Itri Amran[a]***

[a]Department of Biosciences, Faculty of Science, Universiti Teknologi Malaysia, Johor Bahru, Johor, Malaysia.
[b]Faculty of Medicine, King Fahad Medical City, Riyadh, Kingdom of Saudi Arabia.
[c]Research Center, King Fahad Medical City, Riyadh, Kingdom of Saudi Arabia.
[d]Department of Biochemistry, Science College, King Saud University, Riyadh, Kingdom of Saudi Arabia.
[e]Department of Chemistry, Faculty of Science, Universiti Teknologi Malaysia, Johor Bahru, Johor, Malaysia.
[f]Central Laboratory, Center for Science and Medical Studies for Girls, King Saud University, Riyadh, Saudi Arabia.
*Corresponding author: syazwaniitri@utm.my

**Abstract**
G6PD deficiency is the most common enzymopathy worldwide, with over 186 mutations that decrease enzyme activity. The structural mechanics as to how or why G6PD mutations affect enzyme activity remains ambiguous. Employing the use of molecular docking, a G6PD dimer in complex with ligands was constructed to simulate G6PD variants common to the Asian population using GROMACS. Trajectory analyses revealed that variants with high enzyme activity, were compact and exhibited low distance between each monomeric subunit of the dimer. A compact dimer led to the exposure of the tetramer salt bridge residues indicating strong protein–protein affinity in potential tetrameric form. Analyzing the protein– ligand complex showed that the inability to retain G6P and c.NADP might have hindered G6PD catalysis and enzyme activity.
**Keywords:** G6PD deficiency, Molecular dynamics simulation, Molecular docking.

## Introduction

Glucose–6–phosphate dehydrogenase (G6PD) deficiency is the most common enzymopathy worldwide. G6PD confers protection to red blood cells by producing anti–oxidative components nicotinamide adenine dinucleotide phosphate (NADP). Loss of G6PD enzyme activity leads to free radical induced hemolysis [1]. There are more than 400 G6PD mutations, of which approximately 47% are deleterious, which affect enzyme activity. The World Health Organization (WHO) has categorized different mutations based on their enzyme activity and clinical phenotype into five classes. Classes I, II and III retain <1%, <10% and 10–60% of enzyme activity respectively and present with hemolysis, whereas classes IV and V are asymptomatic and not harmful [1]. G6PD monomers have three ligands, one glucose–6–phosphate (G6P), one catalytic NADP (c.NADP), and one structural NADP (s.NADP). In an active state, G6PD exists as dimers and tetramers in dynamic equilibrium influenced by ligands. Tetramers are formed with increased NADP concentration, and dimers are disassociated into inactive monomers with increased G6P concentration. Dimerization is crucial for basic G6PD enzyme activity, whereas tetramerization increases the structural integrity of the protein [2, 3].

Molecular dynamics simulation (MDS) has been an invaluable tool for studying macromolecular structural and functional relationships [4]. Previous G6PD–MDS studies involved simulating G6PD monomers and dimers without ligands. There have not been any insights of the enzyme in its active state and in complex with ligands. Moreover, previous studies involved simulating G6PD variants common to the USA, German and Middle Eastern population [5, 6]. Even though 5–20% of global G6PD deficiency cases are found in Asia [7], there is a gap in the literature regarding structural insights of G6PD variants common to the Asian population. There are no structures of the G6PD dimer in complex with all of its ligands (G6P, c.NADP and s.NADP) available in the protein data bank (PDB) till date. This makes it challenging to study the enzyme in its active state which is highly influenced by its ligands.

To overcome the forementioned challenges, a complete G6PD dimer in complex with ligands was constructed by employing the use of molecular docking. Ten deleterious variants common to the Asian population were subjected for simulation and compared against the wild–type (WT) using trajectory analyses.

**Materials and methods**

Six ligands (2 X G6P, 2 X c.NADP and 2 X s.NADP) were docked onto 2BHL, after removing existing water molecules and ligands from the crystal structure, with controlled docking using AutoDock 4.2 and AutoDock Vina [8, 9]. Ten deleterious variants common to the Asian population were chosen for analysis. PyMOL was employed for in silico site directed mutagenesis for creating the selected variants. The WT and ten variants were subjected to 100 ns simulations using the GROMACS 2018.1 package with the GROMOS96 54a7 force field [10]. Trajectory analyses were performed to analyze changes at the mutation site, dimer and tetramer interfaces and protein–ligand complex of the variants with respect to the WT.

**Results and discussion**

TM–scores were calculated to evaluate structural deviations exhibited by the simulated variants against the WT, as depicted in Table 1. Greater the TM–score, higher the structural similarity [11]. It was noted that variants V291M, G163S and G131V with high enzyme activity exhibited the highest TM–scores > 0.49. Whereas G410D with low enzyme activity exhibited the lowest TM–score. Based on the WHO classification of enzyme activity, there exists a conundrum for R387C and H32R, as the reported enzyme activities listed in Table 1, do not fall in their respective threshold.

**Table 1:** Correlation between enzyme activity and structural similarity.

| Class | Variant | Enzyme activity (%) | Reference | TM–score (Variant against WT at 100 ns) |
|---|---|---|---|---|
| I | Shinagawa (G410D) | ≈ 0.1 | [12] | 0.4675 |
| | Bangkok (K275N) | Undetected | | 0.4776 |
| | Guadalajara (R387C) | ≈ 11.2 | [13] | 0.4801 |
| II | Viangchan (V291M) | ≈ 63 | [14] | 0.4935 |
| | Vanua Lava (L128P) | ≈ 61.7 | [14] | 0.4783 |
| | Canton (R459L) | ≈ 18 | [15] | 0.4814 |
| | Surabaya (V431M) | Undetected | | 0.478 |
| III | Gaohe (H32R) | ≈ 4 | [16] | 0.4708 |
| | Mahidol (G163S) | ≈ 82 | [17] | 0.4902 |
| | Quing Yan (G131V) | ≈ 70 | [18] | 0.4913 |

Figure 1.(A) depicts the presence of salt bridges between Lys 407 – Glu 206 (highly conserved) and Glu 419 –Arg 427 [2]. The dimer bridge, located at the βN – αI region is stabilized by hydrogen bonds between Asp 421 – Asp 421 [2]. As seen in Figure 1.(B), Lys 171 from the βE–αe loop is responsible for G6P and c.NADP positioning required for G6PD catalysis [15].

Docking two s.NADP ligands increased protein stability for the WT, R387C, V291M, and G163S, enabling it to establish additional hydrogen bonds between Glu 419 – Thr 423, which were not found in 2BHL. This validates the physiological significance of ligands on G6PD protein structure. Despite being unable to retain salt bridges and hydrogen bonds between βN at the dimer interface, G410D and K275N were able to retain the Lys 171–G6P–c.NADP hydrogen bond. G410D's ability to do so might have been due to s.NADP dependent enzyme reactivation, as exon 10 mutants tend to be reactivated by the addition of NADP [12]. H32R's low enzyme activity might be due the inability to retain the Lys 171–G6P–c.NADP hydrogen bond, indicating hindered G6PD catalysis.

Figure 1.(D–E) depict the radius of gyration (Rg) of the protein structure and solvent accessible surface area (SASA) of the tetramer salt bridge residues for the WT throughout the simulation. Lower the Rg, more compact the structure and higher the SASA, more exposed the residues are to the surrounding solvent respectively [19]. The presence of two s.NADPs at the dimer bridge was responsible for the additional hydrogen bonds between Glu 419–Thr 423 which increased the compactness of the protein structure, thereby reducing the βN – βN distance as shown in Table 2.

G6PD tetramers are stabilized by salt bridges between Lys275 – Glu347 and Glu287 – Lys290 of each dimer [2]. Greater the surface are of a protein interface, stronger the protein – protein affinity[20]. The presence of four NADP ligands on the dimer structures, shifted the equilibrium towards a tetrameric state characterized by increased SASA at the tetramer salt bridge residues for the WT by a factor of 1.77 nm2 as shown in Figure 1.(E).

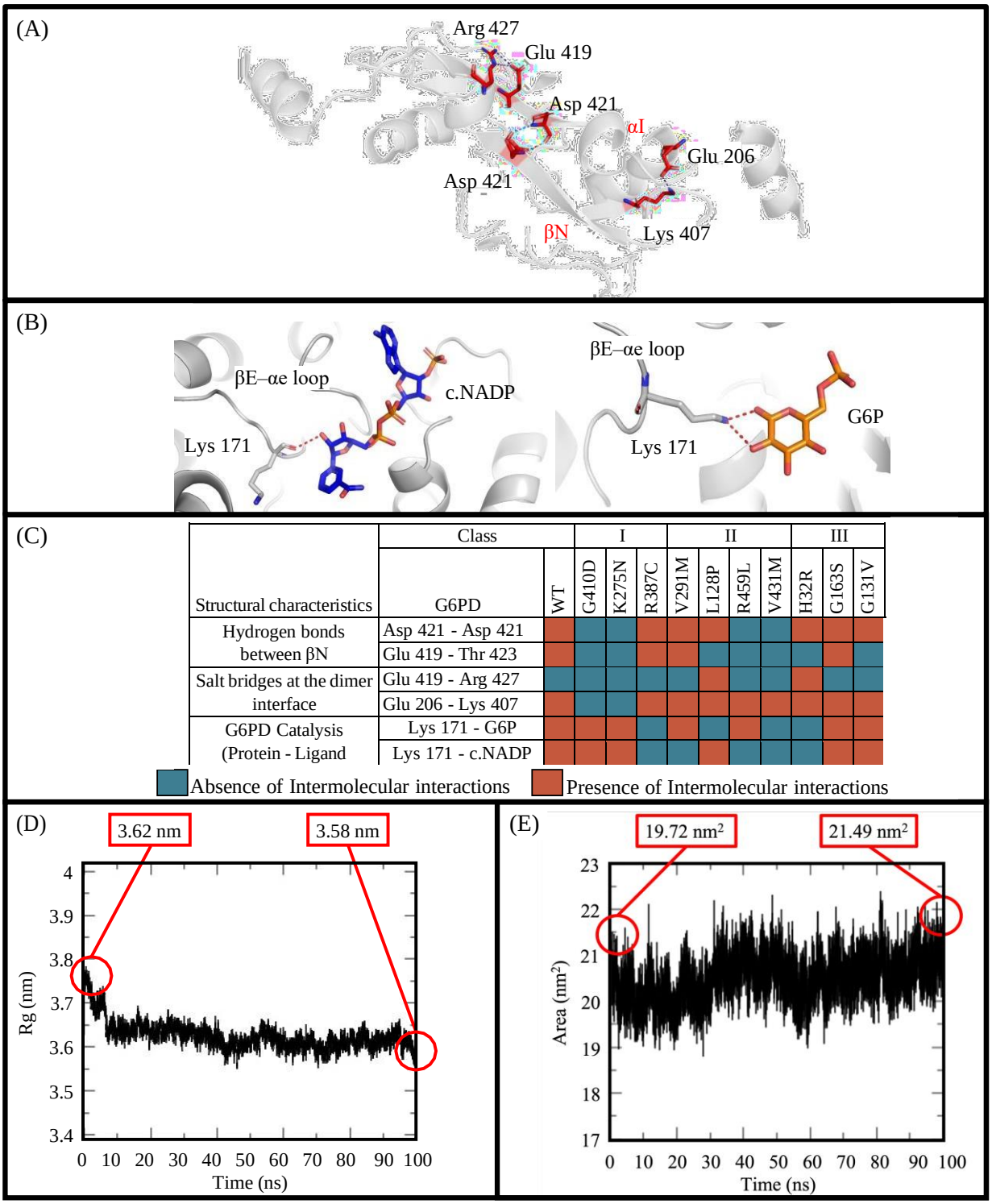


**Figure 1** (A) Stability of the G6PD dimer bridge is dependent on the formation of salt bridges and hydrogen bonds. (B) Lys 171 from the βE–αe loop is responsible for positioning G6P and c.NADP at their respective binding pockets. (C) Heatmap depicting the structural integrity of the dimer interface and ligand positioning capabilities for the simulated WT and variants. (D) Rg of the WT structure. (E) SASA of the tetramer salt bridge residues for the WT.

Based on the mutation sites, dimer interface originating variants, G410D, R459L and V431M were characterized with high βN – βN distance (> 0.26 Å), making the dimer structures loosely packed characterized by high Rg (>3.66 nm). However, R387C exhibited low βN – βN distance and Rg, despite being a dimer interface originating variant. This might have been due to s.NADP dependent reactivation similar to G410D, as R387C is also an exon 10 mutant.

Given that G6PD dimerization and tetramerization are interdependent, high stability at the dimer interface characterized by low βN – βN distance and low Rg values increases tetramerization capabilities characterized by high SASA values for the tetramer salt bridge residues.

The effects and properties of mutations are crucial for variant assessment. K275N's mutation renders it incapable of establishing the Lys275 – Glu347 salt bridge and was found to have the lowest tetramer salt bridge SASA and highest distance between βN – βN (Table 2). V291M's mutation exhibited high SASA at the salt bridge residues despite having high βN – βN distance and high Rg (Table 2), which might have been because methionine is often found at exposed regions of the protein core. R459L's mutation led to a loss of the αn – αe interhelical interactions concordant with the report by Hwang et al (2018) [15], and might have been the reason for its loosely packed structure characterized by a high Rg (Table 2).

High enzyme activity for G163S and G131V (>70%) might have been due to their ability to retain similar structural features as the WT as depicted in Figure1.(C) and Table 2, thus enabling them to exhibit high TM–scores (Table 1). The reason for L128P's hindered enzyme activity despite having similar characteristics to the WT with regards to βN – βN distance and Rg values, might have been due to low SASA (Table 2), similar to H32R. Interestingly both L128P and H32R were unable to retain the Lys 171-c.NADP hydrogen bond, indicative of hindered G6PD catalysis.

**Table 2:** Structural characteristics of the simulated WT and variants.

| Class | Protein | βN – βN distance (Å) | Rg (nm) | SASA ($nm^2$) |
|---|---|---|---|---|
| | WT | 0.19 | 3.62 | 20.44 |
| I | G410D | 0.26 | 3.71 | 20.38 |
| | K275N | 0.31 | 3.69 | 19.6 |
| | R387C | 0.20 | 3.59 | 20.75 |
| II | V291M | 0.23 | 3.67 | 20.68 |
| | L128P | 0.21 | 3.65 | 20.09 |
| | R459L | 0.28 | 3.72 | 20.33 |
| | V431M | 0.29 | 3.66 | 20.39 |
| III | H32R | 0.21 | 3.72 | 20.07 |
| | G163S | 0.20 | 3.63 | 20.38 |
| | G131V | 0.22 | 3.67 | 20.58 |

## Conclusion

This was the first G6PD–MDS study to relate in silico findings to previous biochemical and structural reports. This study was successful in unravelling the required structural dynamics for the G6PD protein to produce optimum enzyme activity. Variant assessment from this study allowed identifying prognostic markers which would be beneficial for future drug development.

## Acknowledgement

We are grateful for the financial support from Fundamental Research Grant Scheme MOHE (FRGS/1/2019/SKK08/UTM/02/1) provided by Malaysia Ministry of High Education awarded to SIA and Intramural Research Fund (Grant: 4B363) awarded to MAH from Ministry of Health, Kingdom of Saudi Arabia. We also acknowledge support from Universiti Teknologi Malaysia, RU grant 15J90.

**References**


[1] Lee J, Kim TI, Kang J–M, Jun H, Lê HG, Thái TL, et al. Prevalence of glucose–6–phosphate dehydrogenase (G6PD) deficiency among malaria patients in Upper Myanmar. BMC Infect Dis. 2018;18(1):131.

[2] Au SWN, Gover S, Lam VMS, Adams MJ. Human glucose–6–phosphate dehydrogenase: the crystal structure reveals a structural NADP+ molecule and provides insights into enzyme deficiency. Structure. 2000;8(3):293–303.

[3] Kotaka M, Gover S, Vandeputte–Rutten L, Au SW, Lam VM, Adams MJ. Structural studies of glucose–6–phosphate and NADP+ binding to human glucose–6–phosphate dehydrogenase. Acta Crystallogr D Biol Crystallogr. 2005;61(Pt 5):495–504.

[4] Hospital A, Goñi JR, Orozco M, Gelpí JL. Molecular dynamics simulations: advances and applications. Adv Appl Bioinform Chem. 2015; 8:37–47.

[5] Doss CGP, Alasmar DR, Bux RI, Sneha P, Bakhsh FD, Al–Azwani I, et al. Genetic Epidemiology of Glucose–6–Phosphate Dehydrogenase Deficiency in the Arab World. Sci Rep. 2016;6(1):37284.

[6] Nguyen H, Nguyen T, Le L. Computational Study of Glucose–6–phosphate–dehydrogenase deficiencies using Molecular Dynamics Simulation. Journal of Life Sciences. 2016; 4:32–9.

[7] Li Q, Yang F, Liu R, Luo L, Yang Y, Zhang L, et al. Prevalence and Molecular Characterization of Glucose–6–Phosphate Dehydrogenase Deficiency at the China–Myanmar Border. PLoS One. 2015;10(7):e 0134593–e.

[8] Mrris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem. 2009;30(16):2785–91.

[9] Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31(2):455–61.

[10] Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: High performance molecular simulations through multi–level parallelism from laptops to supercomputers. SoftwareX. 2015;1–2:19–25.

[11] Zhang Y, Skolnick J. TM–align: a protein structure alignment algorithm based on the TM–score. Nucleic Acids Res. 2005;33(7):2302–9.

[12] Hirono A, Miwa S, Fujii H, Ishida F, Yamada K, Kubota K. Molecular study of eight Japanese cases of glucose–6–phosphate dehydrogenase deficiency by non–radioisotopic single–strand conformation polymorphism (SSCP) analysis. Blood. 1994; 83:3363–8.

[13] Martínez–Rosas V, Juárez–Cruz MV, Ramírez–Nava EJ, Hernández–Ochoa B, Morales–Luna L, González–Valdez A, et al. Effects of Single and Double Mutants in Human Glucose–6–Phosphate Dehydrogenase Variants Present in the Mexican Population: Biochemical and Structural Analysis. International journal of molecular sciences. 2020;21(8):2732.

[14] Gómez–Manzo S, Quino J, Ortega–Cuellar D, Serrano–Posada H, Gonzalez–Valdez A, Vanoye Carlo A, et al. Functional and Biochemical Analysis of Glucose–6–Phosphate Dehydrogenase (G6PD) Variants: Elucidating the Molecular Basis of G6PD Deficiency. Catalysts. 2017; 7:135.

[15] Hwang S, Mruk K, Rahighi S, Raub AG, Chen C–H, Dorn LE, et al. Correcting glucose–6–phosphate dehydrogenase deficiency with a small–molecule activator. Nat Commun. 2018;9(1):4045.

[16] Chao LT, Du CS, Louie E, Zuo L, Chen E, Lubin B, et al. A to G substitution identified in exon 2 of the G6PD gene among G6PD deficient Chinese. Nucleic Acids Res. 1991;19(21):6056–.

[17] Huang Y, Choi MY, Au SW, Au DM, Lam VM, Engel PC. Purification and detailed study of two clinically different human glucose 6–phosphate dehydrogenase variants, G6PD(Plymouth) and G6PD(Mahidol): Evidence for defective protein folding as the basis of disease. Mol Genet Metab. 2008;93(1):44–53.

[18] Fu C, Luo S, Li Q, Xie B, Yang Q, Geng G, et al. Newborn screening of glucose–6–phosphate dehydrogenase deficiency in Guangxi, China: determination of optimal cutoff value to identify heterozygous female neonates. Sci Rep. 2018;8(1):833

[19] Kumar CV, Swetha RG, Anbarasu A, Ramaiah S. Computational Analysis Reveals the Association of Threonine 118 Methionine Mutation in PMP22 Resulting in CMT–1A. Adv Bioinformatics. 2014; 502618.

[20] Chen J, Sawyer N, Regan L. Protein–protein interactions: general trends in the relationship between binding affinity and interfacial buried surface area. Protein Sci. 2013;22(4):510–5.