



Molecular dynamics of G6PD variants from sub-Saharan Africa

Jorge Batista da Rocha^{a,b}, Houcemeddine Othman^a, Scott Hazelhurst^{a,c,*}

^a Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

^b Division of Human Genetics, National Health Laboratory Service, and School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

^c School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa

ARTICLE INFO

Keywords:

G6PD
Structural bioinformatics
Variants
Precision medicine
Africa
Molecular Dynamics

ABSTRACT

Precision medicine uses genomic guidance to improve drug treatment safety and efficacy. Prior knowledge of genetic variant impact can enable such strategies, but current knowledge of African variants remains scarce. *G6PD* variants are linked to haemolytic adverse effects for a number of drugs commonly used in African populations. We have investigated a set of *G6PD* variants with structural bioinformatics techniques to further characterise variants with known effect, and gain insights into variants with unknown impact. We observed wide variations in patterns of root-mean-square deviation between wild-type and variant structures. Variants with known, highly deleterious impact show structural effects which may likely result in the destabilisation of the *G6PD* homodimer. The V68M and N126D variants (which are both common across African populations, and together form the A- haplotype) induce large conformational shifts in the catalytic NADP⁺ binding domain. We observed a greater impact for the haplotype than for each of the individual variants in these cases. A novel African variant (M207T) shows the potential to disrupt interactions within the protein core, urging further investigation. We explore how characterising the molecular impact of African *G6PD* variants can enable advanced strategies for precision medicine, as well as impact the use of novel therapeutics aiming to treat *G6PD* deficiency. This knowledge can assist in bridging current knowledge gaps, and aid to facilitate precision medicine applications in African populations.

1. Introduction

Precision medicine, based on genomic guidance, could improve drug safety and efficacy in African populations. These strategies would be based on strong genomic knowledge of variants and how these impact drug reactions. There is currently a scarcity of such knowledge for African populations [1]. Genetic variation that leads to protein coding changes are known to have large impacts on function, which in turn may have implications for drug metabolism, interaction, or transport.

The glucose-6-phosphate dehydrogenase (*G6PD*) enzyme is critical to the reduction of oxidative stress in erythrocytes via the production of nicotinamide adenine dinucleotide phosphate (NADPH) [2]. As the citric acid cycle is unavailable in erythrocytes, only NADPH produced by *G6PD* is available for glutathione mediated reduction of reactive oxygenated species. *G6PD* binds glucose-6-phosphate and NADP⁺ and catalyses the formation of 6-phospho-D-glucono-1,5-lactone and NADPH. The structure of *G6PD* is formed of a homodimer, which can then form a tetramer. Each monomer consists of two domains, a catalytic domain which binds *G6P* and NADP⁺, and a c-terminal domain

which forms the dimer interface, and binds a structural NADP⁺ which is key to protein stability [3]. The enzyme is inactive as a monomer [4], but is active in its homodimeric and tetrameric forms [5,6].

Deleterious variants in the *G6PD* gene can lead to *G6PD* deficiency, which is linked to increased risk of haemolytic events. As the *G6PD* gene lies on the X chromosome, the deficiency is more common in males. Most variants with severe functional impact (activity < 90%) lie outside of the enzyme active site, and occur in dimeric interface regions, or proximal to residues which interact with the structural NADP⁺ [7]. As there are over 200 *G6PD* amino acid substitutions reported to exist [8], novel ways of interpreting functional impact are important, particularly for populations lacking genetic characterisation.

The *G6PD* gene is an important pharmacogene for numerous drugs used in African populations. These include commonly used antimalarial drugs such as primaquine [9]. Aminoquinolines such as primaquine and chloroquine increase oxidative stress via production of reactive oxygen species, which may contribute to their antimalarial effects [10]. Insufficient or nonfunctional *G6PD* may lead to a lack of NADPH in

* Corresponding author at: Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.

E-mail address: scott.Hazelhurst@wits.ac.za (S. Hazelhurst).

<https://doi.org/10.1016/j.bbrep.2022.101236>

Received 30 November 2021; Received in revised form 15 February 2022; Accepted 16 February 2022

2405-5808/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

erythrocytes, which no longer neutralise the reactive oxygen species produced by aminoquinolines, leading to risk of haemolysis. G6PD genotyping would be beneficial for primaquine radical curing in malaria endemic areas [11], and dosing models would benefit from more knowledge of the distribution and impact of G6PD variants [12]. Knowing more about variants which could lead to G6PD deficiency could be used to improve safety strategies of such drugs in African populations.

In previous work, we observed large differences in allele frequency for G6PD variants in African populations [13]. There were nine G6PD coding variants which induce missense type substitutions of amino acids identified across a large sample of African population data. Molecular dynamics is a computer simulation method to assess the physical movements of atoms and molecules. It can be used to gain insight into the behaviour of protein structures. Recent advances in protein force field and scoring functions have allowed for computationally efficient assessment of the impact of a variant on protein structure [14]. These could be employed to simulate the structural impact of G6PD variants.

In this study, we investigate G6PD variants with structural bioinformatics techniques to reveal insights into the molecular mechanisms of effect for variants of known outcome, and infer hypotheses for those variants with currently unknown functional effect.

2. Methodology

2.1. Genetic characterisation of G6PD variants in Africa

A set of nine African G6PD missense variants have been identified in a collection of African sequencing datasets. These data were formed of combined datasets from the Human Heredity and Health in Africa (H3Africa) project ($n=458$) and African data from the 1000 Genomes project (KGP) ($n=506$). We have applied the shorthand names to a selection of the variants as follows: “MEDI” (S188F), “ADEL” (V68M), ABEN (N126D); and “AMIN” to the V68M+N126D haplotype.

2.2. Structure preparation

The initial structure of G6PD was adapted from the crystal structure of the G6PD homodimer coupled with structural NADP+ (PDB ID: 6E07, chains C and L [15]). The 6E07 structure contains the Canton variant (Arg459Leu). The Canton structure is of very similar topology to wild-type G6PD. The wild-type structure was produced by reversing this variant in both monomer chains using the *tleap* program from AMBER 18 molecular dynamics package [14]. PDB2PQR [16] was used to establish protonation forms of charged amino acid residues with PROPKA [17,18] based predictions. *tleap* preparation included force field parameter files for NADP+ [16], and the ff14SB parameters for protein force fields.

An *in vacuo* minimisation of the structure was performed prior to solvation for 250 cycles of steepest descent minimisation (SDM), followed by 400 cycles of conjugate gradient minimisation (CGM), with a 2 kcal/mol/Å² restraint applied to all atoms excluding those in mutated residues. The resultant structure was then solvated in a 15Å octagonal box spaced from the edge of the outermost atoms in the structure. The solvated structure was fully minimised following three steps, with each step releasing restraints applied to replaced residue side chains, all residue side chains, and finally backbone atoms. During the first stage, 500 cycles of SDM followed by 500 cycles of CGM were used with 500 kcal/mol/Å² restraining force applied to all atoms except the replaced atoms of the variant side chain. In the second stage, 500 cycles of SDM and 10000 cycles of CGM were applied respectively, while using a 500 kcal/mol/Å² restraining force applied to the backbone atoms of the protein. Finally, in the third stage, 500 cycles of SDM, followed by 10000 cycles of CGM were used during which we lifted out all the restraining forces on any atoms. This was followed by heating of the system to 300 K. The SHAKE protocol was applied to restrain bonds involving hydrogen. This was performed step wise over a restraint ranging from 10 kcal/mol/Å² to 0 kcal/mol/Å², in an increment of 1 kcal/mol/Å², applied to all system residues.

2.3. Molecular dynamics simulation

Production of the system for wild-type and variant forms of G6PD (Table 1) was carried out for 510 ns, with a 2 fs timestep. Temperature was controlled via the weak-coupling ensemble method. The pressure periodic boundary was kept constant at 1 atmosphere. Pressure was maintained via isotropic position scaling. CPPTRAJ [19] was used to strip trajectories of water and ion molecules prior to analysis, and to merge the trajectory intervals. Root-mean-square-deviation (RMSD), root-mean-square-fluctuation (RMSF) and Principal Component Analysis (PCA) analyses were performed with MDtraj [20]. Dynamic cross correlation matrix (DCC/M) analysis was conducted with pytraj [19].

3. Results

3.1. G6PD Structure and variants

The selected G6PD structure (PDB – 6E07) is a homodimer, with each monomer consisting of 511 residues. Each monomer comprises two separate domains: the catalytic NADP+ (catNADP+) binding domain (residue 1–206), and the c-terminal domain (residue 207–511), which binds a structural NADP+ (strNADP+), and forms the dimer interface (Fig. 1). Variants of interest detected in African population datasets [13], are highlighted in Fig. 1. Of the nine missense variants identified, we have selected eight (Table 1) (excluding one that could not be mapped to the structure) and the G6PD Mediterranean variant for structural bioinformatics analysis. Classifications for G6PD variants were developed and applied by the World Health Organisation to reflect the resultant level of enzyme deficiency [21,22]. The MEDI variant is also included for comparison. Of these variants four are in the catNADP+ binding domain, while four are in the c-terminal domain. Two c-terminal domain variants (M207T, M212V) are close to the G6PD active site, which is located between residues 200–206. The D350H, M207T, M212V, and R104H variants are unclassified in terms of functional effect.

In the catalytic NADP+ binding region (residue 1–206), the range of RMSD is shorter as compared with those of the whole structure, however, different patterns are observed for variants. Notably, E156K has a sustained increase in deviation across the trajectory, and L323P an increase after 300 ns, while patterns for the overall structure for both of these variants did not diverge greatly from WT. This may indicate a region-specific impact induced by these variants. The novel M207T variant shows fluctuations in RMSD in this region, before returning to converge along WT-like patterns. RMSD for active site residues (200–206) closely resemble that of WT.

3.2. Local fluctuations in G6PD

RMSF for variant forms (Figure S2) show that the variant residue itself does not deviate compared to its WT counterpart. There is a trend of fluctuation in residues 80–120 (in the catNADP binding domain), with the largest (>1 nm) observed in the AMIN variant trajectory. A change in RMSF for residues 322–329 (c-terminal domain) was observed in all variants compared to WT structure, though this may be an artifact of the WT trajectory rather than a variant induced effect. This is likely as the RMSF for this section does not differentiate across each variant trajectory.

3.3. Essential dynamics analysis

Projections of C_{α} coordinates show that most variant forms occupy a wider space than that of WT G6PD (Fig. 2), based on first and second principal components. The MEDI variant shows a large deviation, which may indicate increased accessible conformations as it samples a wider degree of sub-spaces. Interestingly, the benign ABEN variant shows greater variance than its ADEL counterpart. Other variants with wide

Table 1

African G6PD variants used in molecular dynamics analyses. Class refers to WHO classification, which defines relative G6PD enzyme activity: I is < 10% and chronic anaemia; II is < 10% with risk of acute haemolytic anaemia; III is 10 – 60% and risk of acute haemolytic anaemia; IV = 60 < 100%; and Unk is Unknown. Distance AS: measured to the alpha carbon of residue 203 (central to the active site — AS). Distance strNADP+: to the alpha phosphate of the strNADP+ ligand.

Variant name	Identifier	Amino acid	Class	Domain	Dist. AS (Å)	Dist. strNADP+ (Å)
WT	N/A	N/A	IV	NA	N/A	N/A
MEDI	rs5030868	S188F	II	catNADP+ Binding	22.0	44.8
ABEN	rs1050829	N126D	IV	catNADP+ Binding	33.7	57.9
ADEL	rs1050828	V68M	III	catNADP+ Binding	26.0	48.1
D350H	rs34193178	D350H	Unk	cTerminal	25.7	25.5
M207T	chrX:g.153762577A>G	M207T	Unk	catNADP+ Binding	9.0	19.3
M212V	rs782754619	M212V	Unk	catNADP+ Binding	11.3	24.1
R104H	rs181277621	R104H	Unk	catNADP+ binding	29.0	49.9
E156K	rs137852313	E156K	III	catNADP+ binding	31.2	55.3
L323P	rs76723693	L323P	III	cTerminal	32.4	31.4

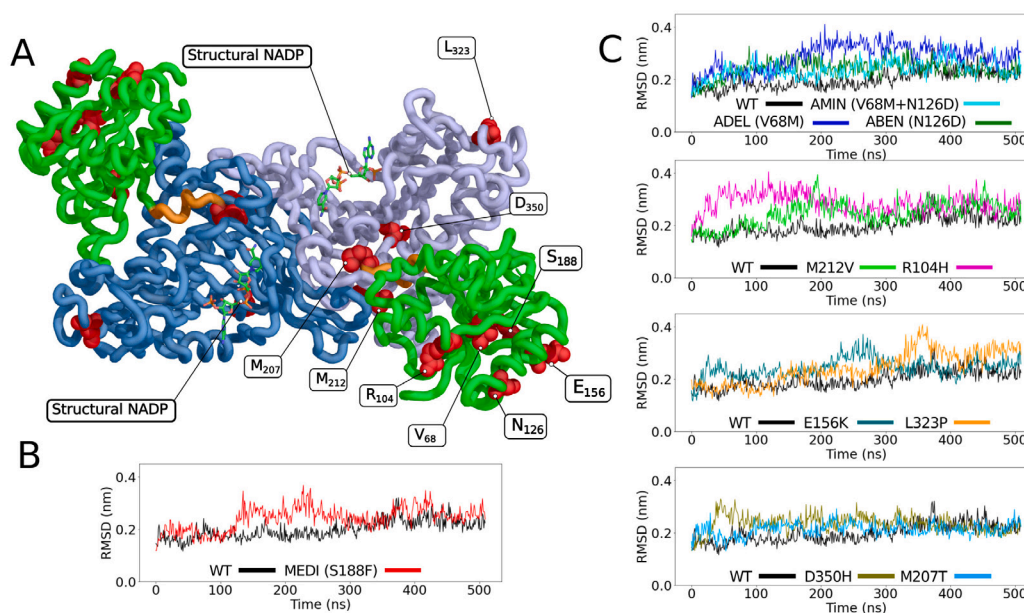


Fig. 1. (A) The G6PD homodimeric structure. WT form of G6PD as a reference (structural base - PDB - 6E07). The catalytic NADP+ binding domain is shown in green and the active site region is shown in orange. Blues indicate the c-terminal domains of each respective monomer. Variants are shown in red spheres. (B) RMSD patterns of G6PD WT (black) compared to the Mediterranean variant. (C) and variants detected in African population data.

divergences are M207T, M212V E156K and L323P. The D350H variant structure is sampling similar conformations to wild-type. No notable differences between variant and wild-type forms were observed in the third and fourth principal components (Figure S4).

We explored position-specific effects further by projection of principal components into porcupine plots (Figure S5) and by examining cross correlated motions (Figure S7) for G6PD wild-type and variant forms. The MEDI variant (WHO class II) and L323P variant (WHO class III) show similar patterns of dynamic cross-correlation (Fig. 3). The network of residue-residue pairwise correlations is more dense in the L323P trajectory than that of MEDI. Residues in the structural domain may be correlating in movement to those in the catalytic NADP+ domain in response to a shift of structural forces. Local views of the porcupine plots for these variants show a pattern where the catalytic NADP+ binding domain and the c-terminal are arcing apart from one another (Figure S6).

In contrast to the MEDI and L323P variants, the African A-variants show less correlation overall. Residue motions within the protein core for AMIN, ADEL, and ABEN resemble that of the wild-type structure (Fig. 4). There are changes to residues in the catNADP+ binding region (see porcupine plot — Figure S6) in terms of direction and impact. The differences appear to be in an additive manner, with ADEL showing larger changes than ABEN, and the combination (AMIN) showing

greater differences compared with ADEL alone. Interestingly, there are some correlated movements between residues of the ABEN variant (see cross correlation matrix plot — Figure S7), but these did not occur in the AMIN trajectory when this variant was together with ADEL.

The M207T variant is a novel variant identified only Nigerian population data from the H3Africa project. It is located in the G6PD c-terminal domain (Fig. 1), in an α -helix which interfaces the β -sheet which interacts with the strNADP ligand. Trajectory motions indicate this variant induces shifts in the core of the protein (Figure S5). Overall patterns of correlated amino acid motions are not different from those observed in the wild-type structure. The D350 variant was observed in datasets from Ghana and Burkina Faso. It is on the protein surface in the c-terminal domain. Motions do not indicate changes from wild type in the protein core, although motions in the catNADP+ binding domain appear to be reduced compared to those of wild-type G6PD (Figure S5). Patterns of amino acid correlation appear very similar to those displayed by the wild-type (Figure S7).

4. Discussion

The G6PD enzyme is a relevant pharmacogene for multiple drug types commonly used in Africa. Variation in the *G6PD* gene has been

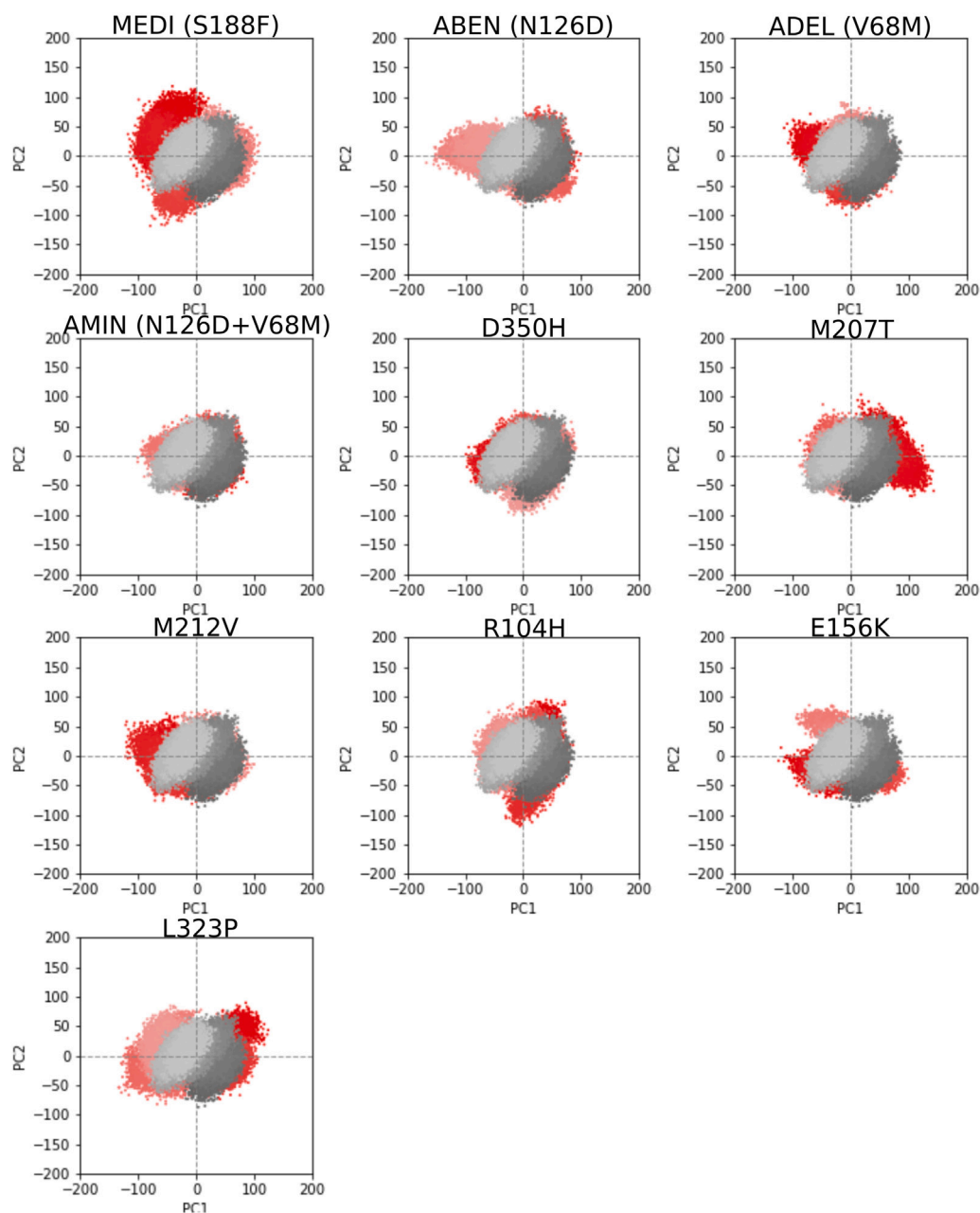


Fig. 2. Projection of G6PD C_{α} atom co-ordinates across the first and second principal components for the entire 510 ns trajectory. Projections for WT G6PD is shown in grey-black, and variant forms are shown in pinks-red.

linked to adverse drug reactions, but most studies to date have investigated variants common in Asian [23] and Mediterranean populations [24]. Previous work showed there are numerous variants which are diverse in sub-Saharan African populations [13]. This study has investigated the impact of African population variation on the structure of G6PD enzyme through molecular dynamics simulation.

Most of the variants we assessed induce large structural drifts as compared to the wild-type form. The structural changes induced by the variants are distinct for each variant. Comparing these changes may allow us to better understand the mechanism of variants with known damaging outcomes (i.e. linked to changes in drug response), and compare variants for which such evidence is lacking.

We have assessed a well-characterised variant, the Mediterranean (MEDI, S188F) class II variant. This variant induces a severe decrease in enzyme activity. Previous work has shown that this variant reduces solvent contact [24], which may lead to its deleterious impact. However, this was observed in a relatively short simulation time (30 ns)

for the size of the system. Longer simulation times allowed us to observe a large increase in correlated motion of amino acids between the two protein sub-units. These motions are distinct from wild type and other variant forms, and may result in an increase in molecular stresses in the enzyme, or induce a shift in contact mapping between G6PD domain segments. These stresses may lead to overall decreased stability of the dimer, resulting in the loss of catalytic activity due the dimer separating. Previous work has shown that the MEDI variant induces structural perturbations that may disrupt hydrogen bonding between $catNADP^{+}$ and a surface α helix near the active site [25]. A combination of overall molecular stresses and disruptions in ligand binding may contribute to the deleterious impact of the MEDI variant.

The L323P variant showed similar characteristics in terms of structural impact to those of MEDI, although it has been labelled a Class III variant. This variant has been shown to reduce risk of development of cerebral malaria in West African populations [26] with strong effect

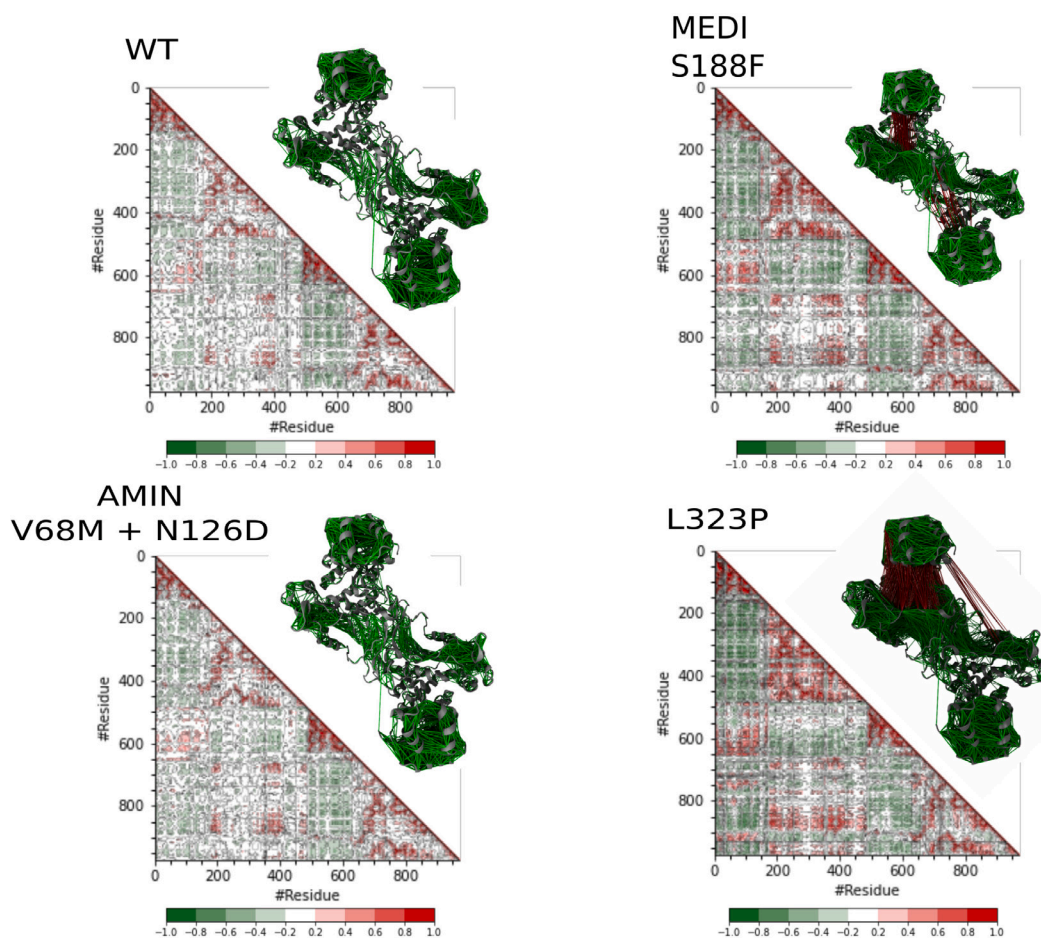


Fig. 3. Dynamic Cross-Correlation matrix analysis of G6PD wild-type and variant forms. $DCC(i, j)$ is mapped onto the structure if greater than 0.7 and lower than -0.7 . Red colours indicate correlation and greens indicate anti-correlation between residues.

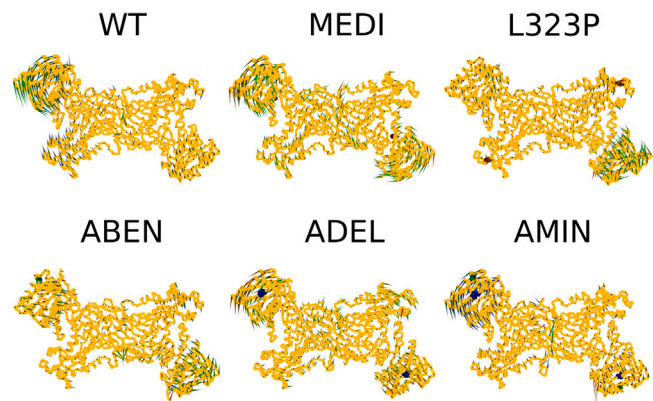


Fig. 4. Projections of first and second principle components (Ca atoms) for G6PD wild-type, Mediterranean, variants in the African A-(AMIN) haplotype. Arrows indicate the proportional movement of alpha-carbon atoms.

sizes, adding to evidence of a stronger functional impact on enzyme function. The novel M207T variant also shows disruptions in the protein core, which may have potential impact to dimer stability. Work evaluating the L323P and the M207T variants in terms of classification would be beneficial for pharmacogenomics strategies to be applied in West Africa.

In contrast, variants common in African populations (ABEN, ADEL, D350H) do not show differences in correlated motions as compared to

the wild type form. The ABEN, ADEL, and the combined AMIN form show their greatest impact on motions within the catNADP+ domain where they are located. The ABEN variant imparts a modest reduction on enzyme activity. It is likely that the ADEL variant emerged in African populations after ABEN, and positive selection effects lead to an increased frequency of the allele combination [27]. While the structural impact of ABEN is modest compared to that of wild type, when combined with ADEL we observed greater motions in the catNADP+ binding domain than in the trajectory of ADEL alone. This increase in motion may lead to increased molecular stresses which overall reduce stability. This is supported by previous work showing ABEN having a cooperative impact on the reduction of enzyme stability when in combination with ADEL, and other G6PD variants [28,29]. These molecular interpretations correlate with clinical evidence that showed the ABEN variant to be a genetic modifier of the impact induced by ADEL on enzyme activity and risk of G6PD deficiency [30].

Although the ADEL and AMIN variant forms show diverse molecular impact from variants like MEDI, it is difficult to characterise how this will lead to the decrease of G6PD enzyme activity. Sirdah et al. (2021) showed that AMIN induces less favourable binding conditions for catNADP+ and G6P, but these were not observed for the single ADEL and ABEN variants [25]. We observed increases in catNADP+ domain motions for both single and variant combinations. There is possibility, that in addition to the potential impact on catalytic ligand binding, these variants have a role in disruption of core protein stability. Previous work has shown that allosteric mechanisms of effect have been observed in deleterious G6PD variants [31] — where the variant induces modulation of sites in the protein far from where the variant itself is located. As dimerisation is critical to the function of

G6PD, a possibility is that changes in the motions of the catNADP+ domain could lead to distal stresses in the dimer interaction interface. This possibility is supported by previous work showing that the AMIN combination of variants, while having a modest impact on specific activity, had drastic impacts on enzyme yield and thermal stability [32]. Hwang et al. (2018) showed that the AMIN variants do induce a loss of dimerisation [15]. This work also showed that a small molecule activator “AG1” can be used to restore the dimeric states of AMIN, MEDI, the Kaiping (R463H) variants of G6PD. This supports the conjecture of allosteric impact of the AMIN variant and suggests future work to include longer simulation times, and detailed assessments of allosteric modulation. Recent work has shown that AG1 only has marginal effects on a diverse set of G6PD variants (Mahidol, Kaiping, Gaohe, Canton, Shoklo and combinations of these variants) [33]. This reinforces the need to investigate such molecules in diverse populations. In particular, for African populations who face a combination of high prevalence of G6PD deficiency, and the need for G6PD based antimalarial pharmacogenomic knowledge.

Interestingly, the AG1 activator showed reduction of oxidative stress in erythrocytes exposed to oxidative stressors, such as chloroquine. Chloroquine has recently been investigated as a repurposed drug to treat COVID-19 [34], which may pose significant risks to African populations where G6PD variants are common [13]. As multiple antimalarial drugs have haemolytic effect, and with the resurged interest in drugs such as chloroquine — it is of increasing relevance to investigate African G6PD variants in terms of functional impact, and the potential use of AG1 to rescue G6PD function for variants common in African populations. In addition to this, we foresee that multidisciplinary efforts will be used enable precision medicine, particularly in African populations. The molecular dynamics patterns we have generated could be used as a record to compare novel or yet uncharacterised variants.

5. Conclusion

Pharmacogenomic knowledge could have significant benefit for drug safety and efficacy in African populations. We have investigated African G6PD variants and shown these display diverse molecular impact on the G6PD structure. The molecular impact of the African AMIN variants are distinct from those of the MEDI variant, although future work would be beneficial to define the mechanism of these. We conjecture that the AMIN variant induces structural stresses that reduce G6PD dimer stability. We have provided evidence for M207T, a novel African variant which lacks in depth functional characterisation indicating potential deleterious outcomes for G6PD activity. We urge further work to characterise African G6PD variants to enable detailed pharmacogenomic strategies for antimalarial and other drugs types commonly used in African populations. Such work should consider the large genetic diversity seen in Africans, and that African variants may display different molecular mechanisms of effect.

6. Future work

Future work could include simulations of other well described variants which could help identify patterns common to deleterious or benign variants. These simulations would benefit from inclusion of catalytic NADP+ and G6P ligands, and investigation of the structural impacts on the catalytic electron transport chain

7. Limitations

The structure we used for simulation does not have G6P and catalytic NADP+ present, and thus we cannot investigate effects of variants on these ligands, and the potential implications for the electron transport chain in their catalysis.

Declaration of competing interest

Jorge E. Batista da Rocha reports financial support was provided by GlaxoSmithKline Inc. Houcemeddine Othman reports financial support was provided by GlaxoSmithKline Inc. Scott Hazelhurst reports financial support was provided by GlaxoSmithKline Inc. Our research group is financially supported by GlaxoSmithKline (GSK). GSK had no part in the design of this study or analysis presented here and the views and opinions expressed are not necessarily those of GSK.

Data availability

Data will be made available on request.

Acknowledgements

We would like to acknowledge Prof. Michèle Ramsay for her guidance and oversight of this work. We acknowledge the Centre for High Performance Computing (CHPC), South Africa, for providing computational resources used in this research project. This work was funded through a grant by GlaxoSmithKline Research & Development Ltd to the Wits Health Consortium. JEBDR was partially funded by the South African National Research Foundation (SFH170626244782).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.bbrep.2022.101236>.

References

- [1] L.G. Landry, N. Ali, D.R. Williams, H.L. Rehm, V.L. Bonham, Lack of diversity in genomic databases is a barrier to translating precision medicine research into practice, *Health Aff (Millwood)* 37 (5) (2018) 780–785.
- [2] P.J. Mason, J.M. Bautista, F. Gilsanz, G6PD deficiency: the genotype-phenotype association, *Blood Rev.* 21 (5) (2007) 267–283.
- [3] S.W. Au, S. Gover, V.M. Lam, M.J. Adams, Human glucose-6-phosphate dehydrogenase: the crystal structure reveals a structural NADP(+) molecule and provides insights into enzyme deficiency, *Structure* 8 (3) (2000) 293–303.
- [4] A. Bonsignore, R. Cancedda, A. Nicolini, G. Damiani, A. De Flora, Metabolism of human erythrocyte glucose-6-phosphate dehydrogenase. VI. Interconversion of multiple molecular forms, *Arch. Biochem. Biophys.* 147 (2) (1971) 493–501.
- [5] A. Bonsignore, I. Lorenzoni, R. Cancedda, A. De Flora, Distinctive patterns of NADP binding to dimeric and tetrameric glucose 6-phosphate dehydrogenase from human red cells, *Biochem. Biophys. Res. Commun.* 39 (1) (1970) 142–148.
- [6] A.D. Cunningham, D. Mochly-Rosen, Structural analysis of clinically relevant pathogenic G6PD variants reveals the importance of tetramerization for G6PD activity, *Matters (Zur)* 2017 (2017).
- [7] N. Horikoshi, S. Hwang, C. Gati, T. Matsui, C. Castillo-Orellana, A.G. Raub, A.A. Garcia, F. Jabbarpour, A. Batyuk, J. Broweleit, X. Xiang, A. Chiang, R. Broweleit, E. Vöhringer-Martinez, D. Mochly-Rosen, S. Wakatsuki, Long-range structural defects by pathogenic mutations in most severe glucose-6-phosphate dehydrogenase deficiency, *Proc. Natl. Acad. Sci.* 118 (4) (2021) <http://dx.doi.org/10.1073/pnas.2022790118>.
- [8] S. Gómez-Manzo, J. Marcial-Quino, A. Vanoye-Carlo, H. Serrano-Posada, D. Ortega-Cuellar, A. González-Valdez, R.A. Castillo-Rodríguez, B. Hernández-Ochoa, E. Sierra-Palacios, E. Rodríguez-Bustamante, R. Arreguin-Espinosa, Glucose-6-phosphate dehydrogenase: update and analysis of new mutations around the world, *Int. J. Mol. Sci.* 17 (12) (2016).
- [9] S.A. Shekalaghe, R. ter Braak, M. Daou, R. Kavishe, W. van den Bijllaardt, S. van den Bosch, J.B. Koenderink, A.J. Luty, C.J. Whitty, C. Drakeley, R.W. Sauerwein, T. Bousema, In Tanzania, hemolysis after a single dose of primaquine coadministered with an artemisinin is not restricted to glucose-6-phosphate dehydrogenase-deficient (G6PD A-) individuals, *Antimicrob. Agents Chemother.* 54 (5) (2010) 1762–1768.
- [10] P. Loria, S. Miller, M. Foley, L. Tilley, Inhibition of the peroxidative degradation of haem as the basis of action of chloroquine and other quinoline antimalarials, *Biochem. J.* 339 (Pt 2) (1999) 363–370.
- [11] J. Recht, E.A. Ashley, N.J. White, Use of primaquine and glucose-6-phosphate dehydrogenase deficiency testing: divergent policies and practices in malaria endemic countries, *PLoS Negl. Trop. Dis.* 12 (4) (2018) e0006230.
- [12] J. Watson, W.R. Taylor, D. Menard, S. Kheng, N.J. White, Modelling primaquine-induced haemolysis in G6PD deficiency, *Elife* 6 (2017).

- [13] J.E.B. da Rocha, H. Othman, C.T. Tiemessen, G. Botha, M. Ramsay, C. Masimirembwa, C. Adebamowo, A. Choudhury, J.-T. Brandenburg, M. Matshaba, G. Simo, F.-J. Gambo, S. Hazelhurst, J.E.B. da Rocha, Consortium, as members of the H3Africa, G6PD distribution in sub-saharan africa and potential risks of using chloroquine/hydroxychloroquine based treatments for covid-19, *Pharm. J.* (2021) <http://dx.doi.org/10.1038/s41397-021-00242-8>.
- [14] D.A. Case, T.E. Cheatham, T. Darden, H. Gohlke, R. Luo, K.M. Merz, A. Onufriev, C. Simmerling, B. Wang, R.J. Woods, The amber biomolecular simulation programs, *J. Comput. Chem.* 26 (16) (2005) 1668–1688.
- [15] S. Hwang, K. Mruk, S. Rahighi, A.G. Raub, C.H. Chen, L.E. Dorn, N. Horikoshi, S. Wakatsuki, J.K. Chen, D. Mochly-Rosen, Correcting glucose-6-phosphate dehydrogenase deficiency with a small-molecule activator, *Nat. Commun.* 9 (1) (2018) 4045.
- [16] N. Holmberg, U. Ryde, L. Bülow, Redesign of the coenzyme specificity in l-lactate dehydrogenase from bacillus stearotherophilus using site-directed mutagenesis and media engineering, *Protein Eng.* 12 (10) (1999) 851–856.
- [17] M.H. Olsson, C.R. Sndergaard, M. Rostkowski, J.H. Jensen, PROPKA3: consistent treatment of internal and surface residues in empirical pka predictions, *J. Chem. Theory Comput.* 7 (2) (2011) 525–537.
- [18] C.R. Sndergaard, M.H. Olsson, M. Rostkowski, J.H. Jensen, Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pka values, *J. Chem. Theory Comput.* 7 (7) (2011) 2284–2295.
- [19] D.R. Roe, T.E. Cheatham, PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data, *J. Chem. Theory Comput.* 9 (7) (2013) 3084–3095, <http://dx.doi.org/10.1021/ct400341p>, PMID: 26583988.
- [20] R.T. McGibbon, K.A. Beauchamp, M.P. Harrigan, C. Klein, J.M. Swails, C.X. Hernández, C.R. Schwantes, L.P. Wang, T.J. Lane, V.S. Pande, MDTraj: A modern open library for the analysis of molecular dynamics trajectories, *Biophys. J.* 109 (8) (2015) 1528–1532.
- [21] N. authors listed, Glucose-6-phosphate dehydrogenase deficiency. WHO working group, *Bull. World Health Organ.* 67 (6) (1989) 601–611.
- [22] A. Yoshida, E. Beutler, A.G. Motulsky, Human glucose-6-phosphate dehydrogenase variants, *Bull. World Health Organ.* 45 (2) (1971) 243–253.
- [23] H.R. Lu, Q.L. Tang, X. Wang, H.J. Li, D.Y. Li, Y.F. Yang, S.F. Tong, C.H. Zhang, Y.C. Zhu, Enzyme kinetics and molecular modeling studies of G6PD(Mahidol) associated with acute hemolytic anemia, *Indian J. Biochem. Biophys.* 48 (5) (2011) 316–324.
- [24] C.G. Doss, D.R. Alasmar, R.I. Bux, P. Sneha, F.D. Bakhsh, I. Al-Azwani, R.E. Bekay, H. Zayed, Genetic epidemiology of glucose-6-phosphate dehydrogenase deficiency in the Arab world, *Sci. Rep.* 6 (2016) 37284.
- [25] M. Sirdah, N.S. Reading, H. Vankayalapati, J.T. Prchal, A computational study of structural differences of binding of NADP⁺ and G6P substrates to G6PD mediterranean^{c.563T}, G6PD A^{—c.202A/e.376G}, G6PD Cairo^{c.404C} and G6PD Gaza^{c.536A} mutations, *Blood Cells Mol Dis* 89 (2021) 102572.
- [26] S.S. Shah, K.A. Rockett, M. Jallow, F. Sisay-Joof, K.A. Bojang, M. Pinder, A. Jeffreys, R. Craik, C. Hubbart, T.E. Wellems, D.P. Kwiatkowski, Heterogeneous alleles comprising G6PD deficiency trait in West Africa exert contrasting effects on two major clinical presentations of severe malaria, *Malar. J.* 15 (2016) 13.
- [27] M.A. Saunders, M. Slatkin, C. Garner, M.F. Hammer, M.W. Nachman, The extent of linkage disequilibrium caused by selection on G6PD in humans, *Genetics* 171 (3) (2005) 1219–1229.
- [28] A. Praoparotai, T. Junkree, M. Imwong, U. Boonyuen, Functional and structural analysis of double and triple mutants reveals the contribution of protein instability to clinical manifestations of G6PD variants, *Int. J. Biol. Macromol.* 158 (2020) 884–893.
- [29] V. Martínez-Rosas, M.V. Juárez-Cruz, E.J. Ramírez-Nava, B. Hernández-Ochoa, L. Morales-Luna, A. González-Valdez, H. Serrano-Posada, N. Cárdenas-Rodríguez, P. Ortiz-Ramírez, S. Centeno-Leija, R. Arreguin-Espinosa, M. Cuevas-Cruz, D. Ortega-Cuellar, V. Pérez de la Cruz, L.M. Rocha-Ramírez, E. Sierra-Palacios, R.A. Castillo-Rodríguez, I. Baeza-Ramírez, J. Marcial-Quino, S. Gómez-Manzo, Effects of single and double mutants in human glucose-6-phosphate dehydrogenase variants present in the Mexican population: biochemical and structural analysis, *Int. J. Mol. Sci.* 21 (8) (2020).
- [30] S.S. Shah, A. Macharia, J. Makale, S. Uyoga, K. Kivinen, R. Craik, C. Hubbart, T.E. Wellems, K.A. Rockett, D.P. Kwiatkowski, T.N. Williams, Genetic determinants of glucose-6-phosphate dehydrogenase activity in Kenya, *BMC Med. Genet.* 15 (2014) 93.
- [31] W.-V. Tee, E. Guarnera, I.N. Berezovsky, On the allosteric effect of nsSNPs and the emerging importance of allosteric polymorphism, *J. Mol. Biol.* 431 (19) (2019) 3933–3942, <http://dx.doi.org/10.1016/j.jmb.2019.07.012>.
- [32] M. Town, J.M. Bautista, P.J. Mason, L. Luzzatto, Both mutations in G6PD A- are necessary to produce the G6PD deficient phenotype, *Hum. Mol. Genet.* 1 (3) (1992) 171–174.
- [33] P. Pakparnich, S. Sudsumrit, M. Imwong, T. Suteewong, K. Chamchoy, D. Pakotiprapha, U. Leartsakulpanich, U. Boonyuen, Combined effects of double mutations on catalytic activity and structural stability contribute to clinical manifestations of glucose-6-phosphate dehydrogenase deficiency, *Sci. Rep.* 11 (1) (2021) 24307.
- [34] R.M. Ghazy, A. Almaghraby, R. Shaaban, A. Kamal, H. Beshir, A. Moursi, A. Ramadan, S.H.N. Taha, A systematic review and meta-analysis on chloroquine and hydroxychloroquine as monotherapy or combined with azithromycin in covid-19 treatment, *Sci. Rep.* 10 (1) (2020) 22139, <http://dx.doi.org/10.1038/s41598-020-77748-x>.