



ORIGINAL PAPER

Red Cells

Pathogenic *G6PD* variants: Different clinical pictures arise from different missense mutations in the same codon

Simonetta Costa^{1,2}  | Angelo Minucci³ | Amit Kumawat⁴ | Maria De Bonis³ |
 Giorgia Prontera¹ | Mariannita Gelsomino¹ | Milena Tana¹ | Eloisa Tiberi¹ |
 Alberto Romano¹ | Antonio Ruggiero^{1,2} | Stefano Mastrangelo^{1,2} | Giuseppe Palumbo⁵ |
 Valentina Giorgio^{1,2} | Maria Elisabetta Onori³ | Martino Bolognesi^{4,6} | Carlo Camilloni⁴ |
 Lucio Luzzatto^{7,8}  | Giovanni Vento^{1,2}

¹Dipartimento di Scienze della Salute della Donna, del Bambino e di Sanità Pubblica, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy

²Dipartimento di Scienze della Vita e Sanità Pubblica, Università Cattolica del Sacro Cuore, Rome, Italy

³Molecular and Genomic Diagnostics Unit, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy

⁴Department of Biosciences, University of Milano, Milan, Italy

⁵Department of Hematology/Oncology, Cell and Gene Therapy, Bambino Gesù Children's Hospital, IRCCS, Rome, Italy

⁶Centro di Ricerca Pediatrica Romeo ed Enrica Invernizzi, Università degli Studi di Milano, Milan, Italy

⁷Department of Haematology and Blood Transfusion, Muhimbili University of Health and Allied Sciences, Dar es Salaam, United Republic of Tanzania

⁸Department of Hematology, University of Florence, Firenze, Italy

Correspondence

Simonetta Costa, Dipartimento di Scienze della Salute della Donna, del Bambino e di Sanità Pubblica, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Largo A. Gemelli 8, Rome 00168, Italy.
 Email: simonetta.costa@unicatt.it

Summary

G6PD deficiency results from mutations in the X-linked *G6PD* gene. More than 200 variants are associated with enzyme deficiency: each one of them may either cause predisposition to haemolytic anaemia triggered by exogenous agents (class B variants), or may cause a chronic haemolytic disorder (class A variants). Genotype–phenotype correlations are subtle. We report a rare *G6PD* variant, discovered in a baby presenting with severe jaundice and haemolytic anaemia since birth: the mutation of this class A variant was found to be p.(Arg454Pro). Two variants affecting the same codon were already known: *G6PD Union*, p.(Arg454Cys), and *G6PD Andalus*, p.(Arg454His). Both these class B variants and our class A variant exhibit severe *G6PD* deficiency. By molecular dynamics simulations, we performed a comparative analysis of the three mutants and of the wild-type *G6PD*. We found that the tetrameric structure of the enzyme is not perturbed in any of the variants; instead, loss of the positively charged Arg residue causes marked variant-specific rearrangement of hydrogen bonds, and it influences interactions with the substrates G6P and NADP. These findings explain severe deficiency of enzyme activity and may account for p.(Arg454Pro) expressing a more severe clinical phenotype.

KEY WORDS

chronic haemolytic disorder, class A variant, *G6PD*, *G6PD* deficiency, molecular dynamics simulations

Simonetta Costa and Angelo Minucci contributed equally to this work and acted as first authors.

Lucio Luzzatto and Giovanni Vento contributed equally to this work and acted as last authors.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *British Journal of Haematology* published by British Society for Haematology and John Wiley & Sons Ltd.

INTRODUCTION

G6PD is a rate-limiting enzyme in the pentose phosphate pathway. G6PD transfers reducing equivalents from glucose-6-phosphate (G6P) to NADP⁺ to produce NADPH, which in turn is the key molecule for defence against reactive oxygen species. More than 200 pathogenic variants of the *G6PD* gene are known,^{1,2} and recently by searching publicly available databases 1041 variants have been listed,³ most of which, however, are non-pathogenic or not yet interpreted.⁴ The World Health Organisation (WHO) has recently revised their classification⁵: variants in class A (formerly class I) are the more severe group, as they are associated with life-long chronic non-spherocytic haemolytic anaemia; persons with variants in class B (formerly class II and class III) are asymptomatic in the steady state, but they are at risk of acute haemolytic anaemia, that may be triggered by eating fava beans, by infections or by certain drugs.^{2,6}

Genotype–phenotype correlation analysis has previously revealed that many but not all class A mutations underlie amino acid replacements located at the dimer interface, mostly from mutations in exon 10.⁷ Here, we report an infant diagnosed with G6PD deficiency soon after birth, who has a *c.1361G>C* mutation in exon 11 of the *G6PD* gene, producing an Arginine to Proline replacement at amino acid position 454 (p.Arg454Pro): we have found that this is a class A *G6PD* variant. This finding was intriguing, because two *G6PD* mutations affecting the same codon were already known: *G6PD Union* (*c.1360C>T/Arg454Cys*),⁸ a class B variant widely spread in the world; and *G6PD Andalus* (*c.1361G>A/Arg454His*), of which there have been only four cases reported, and the classification of which is not quite certain.⁹

The existence of three different amino acid replacements within the same codon has prompted us to investigate in detail their impact on the structure and function of the G6PD molecule. Our data provide a reasonable explanation as to why these mutations produce different phenotypes.

METHODS

Clinical report

A 2325-g-male infant, the second child of a healthy unrelated couple, was born at 36⁺⁵ weeks gestation by caesarean delivery, because of labour and previous caesarean section, to a 29-year-old mother. Apgar scores were 6 at 1 min and 8 at 5 min. Maternal history included gestational diabetes treated with insulin, and lymphedema in lower limbs. During pregnancy, the mother had been on monthly benzathine penicillin G therapy, aspirin and enoxaparin. Her TORCH serological screening was negative, and there was no evidence or history of haemolytic anaemia.

At 2 h of life, the baby was jaundiced. There was no blood group incompatibility and the direct antiglobulin test was negative. Phototherapy was started, but it was discontinued after 35 h, because unconjugated bilirubin was never above

12 mg/dL, although direct bilirubin (DB) had increased further: it peaked at 13 mg/dL at 88 h of life (Table 1). There were no signs of acute bilirubin encephalopathy. Splenomegaly without hepatomegaly was detected. Intravenous immunoglobulin was administered, but there was no improvement. A blood film on day of life 5 showed marked aniso-poikilocytosis, some spherocytes and nucleated red cells (Figure 1); by that time the haematocrit had fallen from 52% at birth to 32%, with marked reticulocytosis (Figure 2; Table 1), consistent with severe haemolytic anaemia: a transfusion of 40 mL of packed red cells was administered. Ursodeoxycholic acid and phenobarbital were also given.

Blood levels of α 1-antitrypsin, ammonia, ceruloplasmin, erythrocyte pyruvate kinase and thyroid hormones were normal. Screening for major metabolic diseases was negative. Molecular tests for hepatitis viruses were negative; blood cultures yielded no growth. On abdominal ultrasound, there was no evidence of biliary atresia. Liver biopsy showed preserved architecture, cholestasis and iron-loaded hepatocytes. Portal spaces were infiltrated by lymphocytes and neutrophils; and there were foci of haematopoiesis.

The baby required two more red cell transfusions and was discharged at the age of 8 weeks, when his haemoglobin was 9 g/L, serum total bilirubin 3.8 mg/dL and DB 2.5 mg/dL (Figure 2A). At the age of 1 year, the peripheral blood smear was normal (Figure 1B). Over the first 3 years of life the infant, in concomitance with urinary infections, experienced several further drops in haemoglobin, requiring red cell transfusions (total of 12: see Figure 2B).

The parents of the infant have provided written informed consent for investigations and genetic analyses, as required by our institutional review board.

Enzyme analysis

G6PD assays were carried out within 24 h of collection on white cell-depleted blood samples of the proband and his mother using a commercial kit (G6P-DH; Sentinel Diagnostics, Milano, Italy). The reference values for normal G6PD activity, in our laboratory, are 9.2–13.8 U/g Hb.

Sequencing and data analysis

Genomic DNA was collected from peripheral blood samples of the patient, his mother, and maternal grandparents using the QIAamp DNA Blood Kit (Qiagen, Germany), according to the manufacturer's instructions. Our diagnostic work-flow involves first screening for the six *G6PD* variants most frequent in Italy.¹⁰ Since the patient's DNA was negative for these, we proceeded to bidirectional Sanger sequencing of the full *G6PD* gene on ABI 3500 Genetic Analyzer (Applied Biosystems, Foster City, CA). The primers used were previously reported.¹¹ RefSeq NM_001042351.2 was used as the reference sequence. Sequence analysis was carried out with the SeqScape Software v2.5. Designation of the *G6PD* alterations was based on the

TABLE 1 Haematological and biochemical parameters of infants with G6PD deficiency during the first week of life.

Hours of life	2	8	12	18	22	29	34	37	49	53	60	68	80	88	102	109	116	138	168
sTB, mg/dL	10.2	12.1	13.8	14.0	15.1	17.3	19.1	18.2	20.0	20.9	24.1	22.6	21.3	23.5	21.6	20.4	18.7	17.6	16.9
DB, mg/dL	2.9	3.7	-	-	-	-	-	6.5	8.2	9.5	11.9	11.8	-	13	-	9.9	-	10.0	-
Haematocrit, %	52	35	40	42	47	42	44	33	38	32	33	30	30	28	24	50	36	34	38
Reticulocytes, $\times 10^9/L$	-	-	-	-	340	-	-	477	-	-	-	466	-	587	-	-	-	-	495
AST, iu/L	-	249	-	-	-	-	-	154	141	116	119	96	-	77	73	-	-	73	192
ALT, iu/L	-	42	-	-	-	-	-	28	24	22	-	24	-	24	23	-	-	27	63
GGT, iu/L	-	119	-	-	-	-	-	82	81	-	73	68	-	50	51	-	-	52	105

Abbreviations: ALT, alanine aminotransferase; AST, aspartate aminotransferase; DB, direct bilirubin; GGT, gamma glutamyl transferase; sTB, serum total bilirubin.

guidelines given by the Human Genome Variation Society (<http://www.hgvs.org/mutnomen/>). We also performed clinical exome sequencing (CES) using capture-based next-generation sequencing (NGS) Clinical Exome Solution® v3 kit (SOPHiA Genetics) that covers the coding regions and splicing junctions of 5500 genes related to rare and inherited conditions. The NGS CES protocol was performed in paired ends reads mode on the Illumina NextSeq550DX® NGS platform (Illumina). The sequencing FastQ data were analysed by the Sophia DDM® platform (Sophia Genetics) to detect single nucleotide variants, *indels*, and copy number variations. The variant calling strategy adopted consists of a filtering of the analysed genes using a phenotype-driver approach on the Sophia DDM® platform. A virtual panel of 220 genes was applied as a filter across the clinical-exome sequencing dataset using the features of ‘Human Phenotype Ontology’ (<https://hpo.jax.org>), Online Mendelian Inheritance in Man (<https://www.omim.org/>), Orphanet (<https://www.orpha.net/>); as well as those reported in the literature to have a causal role in bilirubin metabolism and cholestatic jaundice. In addition, we performed direct sequencing of the uridine-50-diphosphate glucuronosyltransferase isoform 1A1 (*UGT1A1*) gene promoter, as previously reported.¹²

Molecular dynamics simulations

Molecular dynamics (MD) simulations were performed for the wild-type G6PD and three mutant proteins with residue substitutions: Arg454Pro (G6PD *Salt Lake*), Arg454Cys (G6PD *Union*), Arg454His (G6PD *Andalus*). The starting point structure was the cryogenic electron microscopy (cryo-EM) model of the G6PD tetramer in the NADP+ and G6P bound state (PDB ID: 7SNI: the wild-type structure was obtained by back-replacing residue Asp200 to Asn).¹³ All simulations were performed with GROMACS 2021 software¹⁴ using the charmm36 force field and the TIP3P water model.¹⁵ The forcefield parameters for G6P and NADP+ were generated using CgenFF.¹⁶ The systems were neutralised, and energy was minimised using the steepest descent method. The temperature was maintained at 300K using the velocity rescale method and pressure was kept constant at 1 atm using the C-rescale barostat.¹⁷ All simulations were performed using periodic boundary conditions, and the particle mesh Ewald method was used for the calculation of electrostatic interactions. The cut-off distance for electrostatic and van der Waals interactions was set to 1.2 nm. Three independent simulations were run for each system for 100 ns each, with frames saved at every 10 ps.

RESULTS

Biochemical and molecular data

G6PD activity in red cells was 1.8 U/g at birth, 0.3 U/g at 4 months and 0.3 U/g again at 8 months of life. G6PD activity in mother and grandparents, who are all from Rome, was

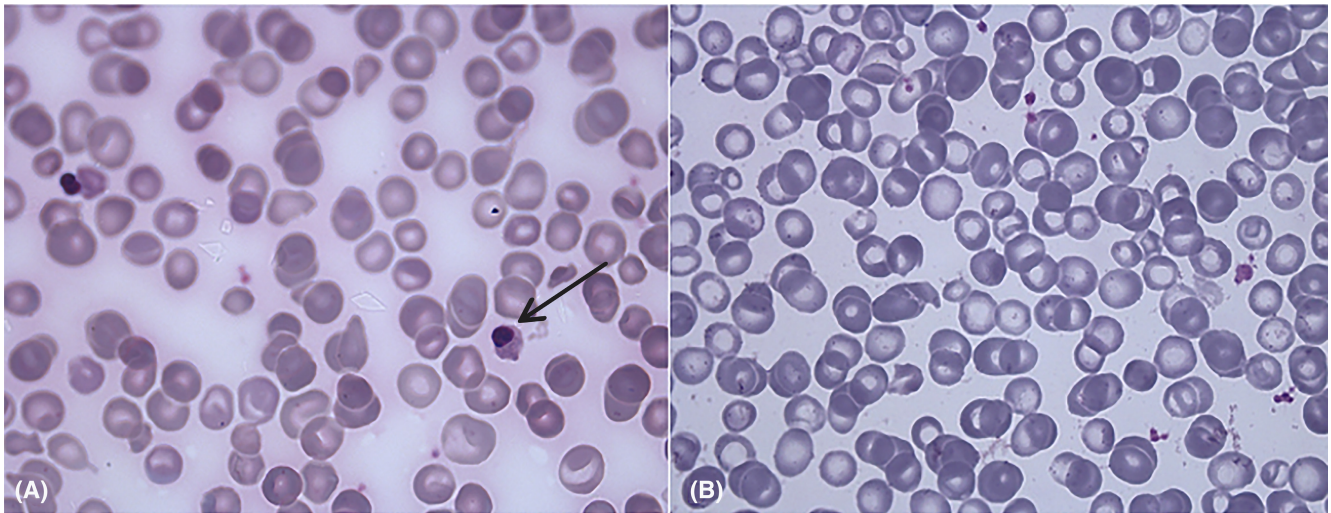


FIGURE 1 Blood smears of G6PD-deficient infants. (A) On day of life 5 there is marked aniso-poikilocytosis, spherocytes, and nucleated red cells (black arrow). (B) At age 1 year, there is still anisocytosis, but morphology is otherwise essentially normal (non-spherocytic).

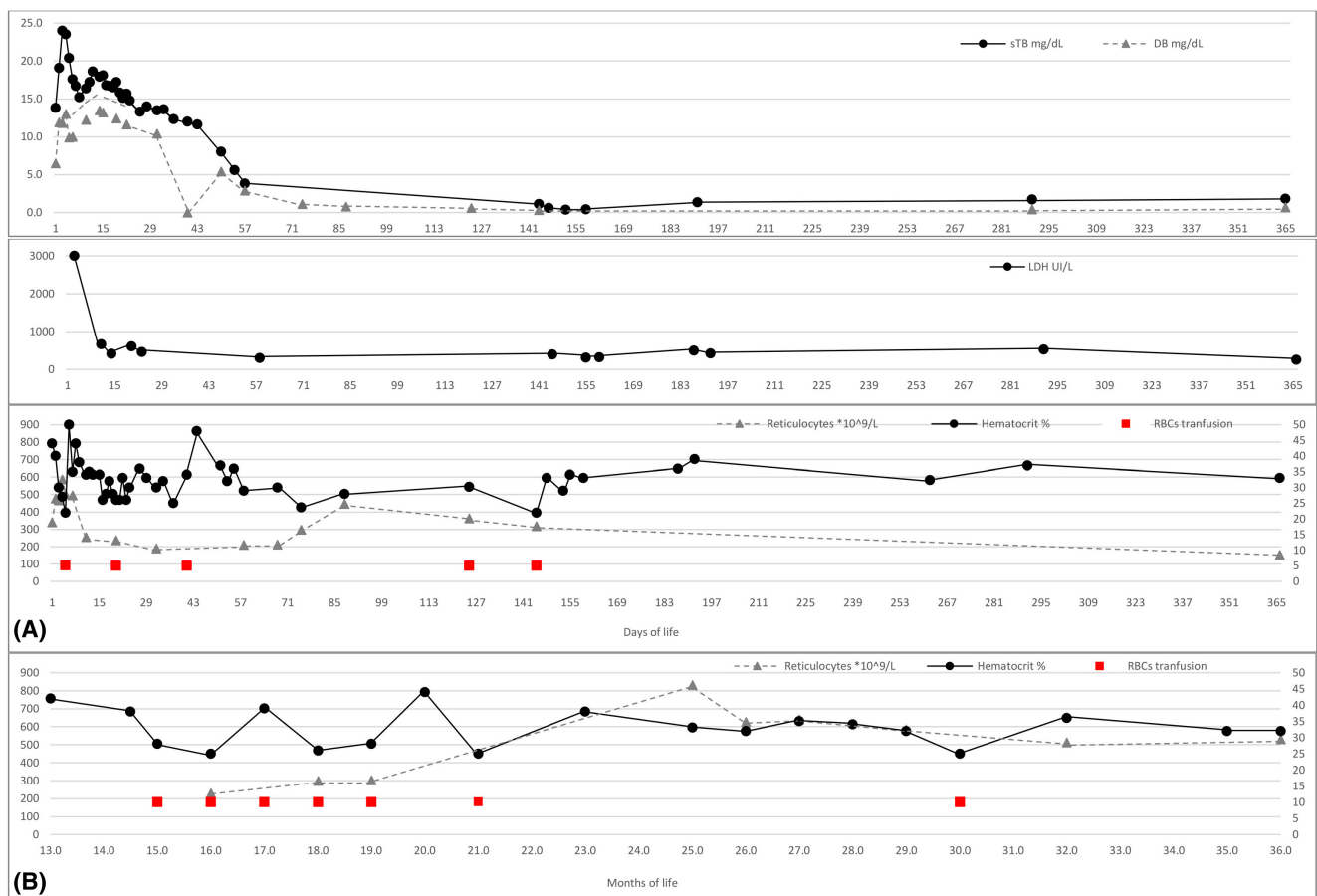


FIGURE 2 Clinical course of infants with G6PD deficiency: Haematological parameters. (A) First year of life. (B) Years 2–3.

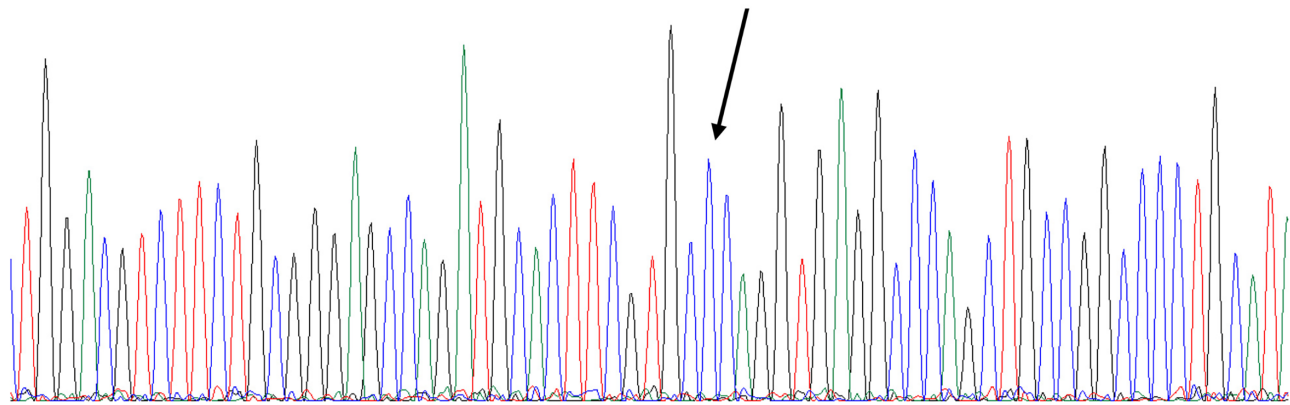
within the normal range. Direct sequencing analysis of the patient's *G6PD* gene exons and intron flanking regions revealed a single base change: *c.1361G>C* (NM_001042351.3) (Figure 3A), causing a Proline for Arginine amino acid replacement (p.Arg454Pro). The mutation was found in the patient's mother (Figure 3B), who was an only child; but not

in the maternal grandparents, suggesting a de novo event. The mutation is not found in the databases of the Exome Aggregation Consortium, 1000 Genomes (<http://www.internationalgenome.org/>), Single Nucleotide Polymorphism Database, ClinVar (<http://www.ncbi.nlm.nih.gov/clinvar/>) and LOVD (<https://www.lovd.nl/>); however, it was reported

G G A C G T C T T C T G C G G G A G C C A G A T G C A C T T C G T G C G C A G G T G A G G C C C A G C T G C C G G C C C C T G C A T A C (

(A)

Single base change
at cDNA position 1361



(B)

Dual peak
at cDNA position 1361

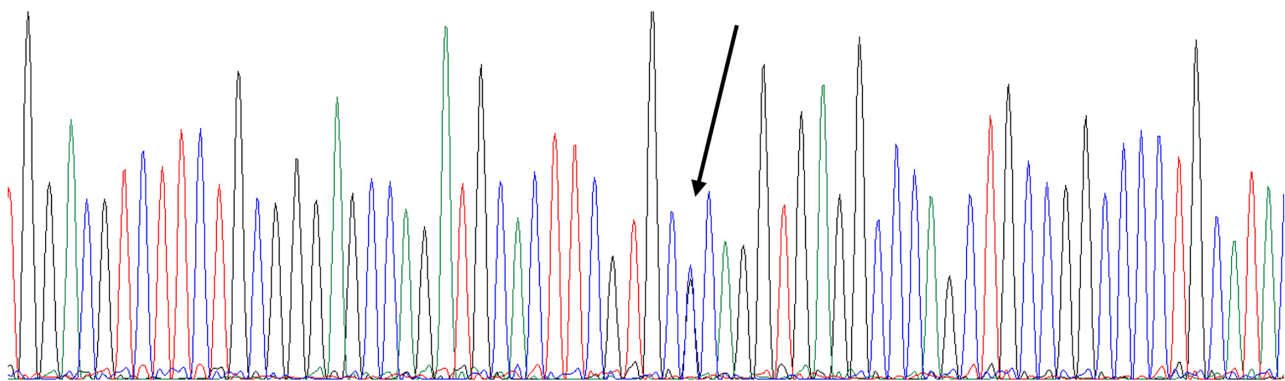


FIGURE 3 Chromatogram of cDNA sequence. (A) Proband. (B) Proband's mother: The double peak shows that the mother is heterozygous.

and referred to as G6PD *Salt Lake* by Powers et al.¹⁸ in one G6PD deficient person, on whom unfortunately no clinical information is available.

Analysis of the patient's *UGT1A1* gene promoter revealed homozygosity for the *UGT1A1**28 (rs8175347) allele, consistent with Gilbert's syndrome.

Structural analysis of G6PD variants with amino acid replacements at R454

The cryo-EM structure of the G6PD tetramer (PDB ID: 7SN1), recently published¹³ displays in each subunit (a) the bound catalytic NADP⁺ cofactor (NADP-c), (b) the G6P substrate, (c) the structure stabilising NADP⁺ molecule (NADP-s), far removed from the active site. Focusing on the mutation site, inspection of the enzyme three-dimensional structure shows that the positively charged Arg454 residue

is linked by a salt bridge to the negatively charged Asp286. Arg454 is part of a cluster of polar amino acids that, in each subunit, stabilise this G6PD region through intramolecular electrostatic interactions (e.g. there is a close-by salt bridge involving Glu193 and Arg285) and several hydrogen bonds (H-bonds) (Figure 4A).¹³ Since Arg454 is the first residue in a 20 amino-acid long α -helix (residues 454–476), replacement by proline may not be excessively disruptive of the helical structure. On the other hand, since residue 454 is at 24–28 Å from both G6P and NADP-c, and since the intervening protein matrix consists mostly of α -helices that are rigid, and the dielectric constant of the protein core is low, loss of a positive charge at residue 454 may be perceived at both the G6P and the NADP-c sites, such that the ligands binding affinities may be altered. Based on such structural considerations, the Arg454 replacement with the non-polar residue Pro454 is expected to affect the dynamics and electrostatic interactions of the

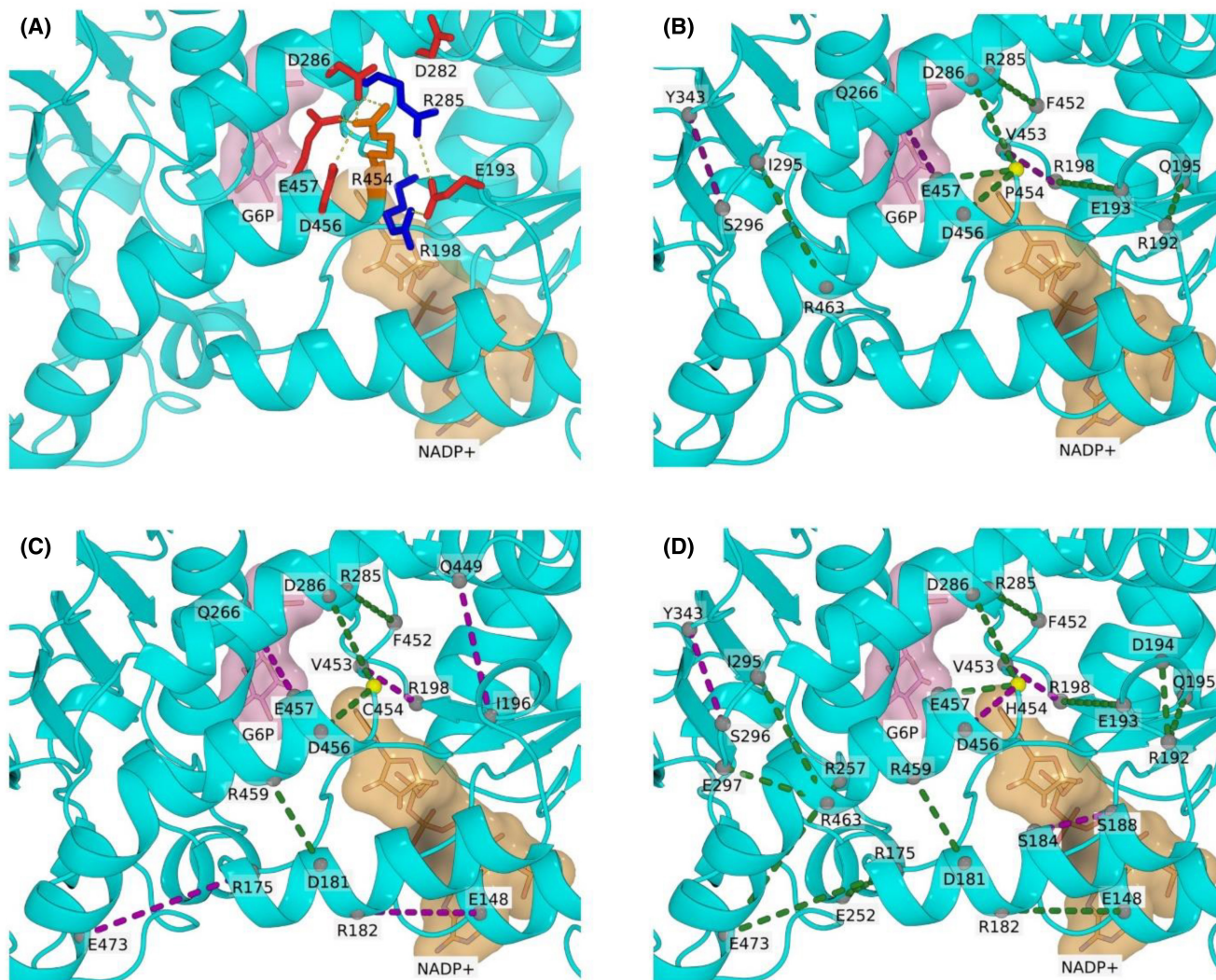


FIGURE 4 Structure of a region of normal G6PD and changes resulting from amino acid replacements in position 454. (A) Detailed view of the G6PD cryogenic electron microscopy (cryo-EM) structure (PDB ID: 7SNI) showing the cluster of charged residues surrounding Arg454 (sphere and stick model in orange). Neighbouring positively charged residues (Arg198, Arg285; Blue stick models) and neighbouring negatively charged residues (Glu193, Asp282, Asp286, Asp456, Glu457; Red stick models) are highlighted because electrostatic interactions extend radially through space. For residues that are very close (distance <0.35 nm) H-bonds (dotted lines) provide additional local stabilisation. The catalytic cofactor (coenzyme) NADP⁺ (light brown) and the substrate G6P (light purple) are shaped in the background. (B) Arg454 is replaced by Pro454 (patient reported in this paper). (C) Arg454 is replaced by Cys454 (G6PD *Union*). (D) Arg 454 is replaced by His454 (G6PD *Andalus*). In panels (B–D), we illustrate differences in H-bond populations (throughout the MD simulation time) that distinguish the three mutants from the wild type, and each one of the three mutants from each other: H-bonds that are less populated in a mutant protein compared to the wild type are shown by green-dashed lines, whereas H-bonds that are more populated in a mutant protein compared to the wild type are shown by purple-dashed lines. In all three mutants, throughout the cluster of polar residues that characterise this region of the G6PD molecule, there is an evident redistribution of hydrogen bonds that is reflected by changes in its dynamics. Individual amino acid sites are shown as small grey spheres, and for reasons of space all residues are identified by their one-letter code (e.g. D and E are the negatively charged aspartic acid and glutamic acid; R is the positively charged arginine). MD, molecular dynamics.

G6PD subunit in this and neighbouring regions, including the active site.

To test this hypothesis, MD simulations were performed in explicit solvent, for 100 ns each, in triplicate runs, on the full tetrameric wild-type G6PD, and on the Arg454Pro *Salt Lake* variant. In addition, we performed similar simulations for the G6PD *Union* (Arg454Cys) and the G6PD *Andalus* (Arg454His) variants, in order to assess whether the latter mutations might differently affect the protein dynamics.

By comparing the (local structure) root mean square fluctuation (calculated over the entire MD simulation period) of the wild type and of the mutant G6PD molecules (Figure S1), we found that each of the amino acid replacements at site 454 does not significantly perturb the overall 3D structure of the G6PD subunits; moreover, the tetrameric quaternary assembly is not perturbed throughout the 100 ns simulation time. On the other hand, close inspection of the mutated site highlights the rearrangement of several H-bonds around Pro454 (Figure 4B); such local structural readjustment is reflected

by the onset of marked collective motions (as defined by Principal Component Analysis throughout the MD trajectories) in the tetrameric assembly, affecting the catalytic region of the enzyme (Figure S2A). A different rearrangement of H-bonds can be observed in G6PD *Union* surrounding residue Cys454 (an amino acid replacement that also removes the positive charge from site 454: Figure 4C), and in G6PD *Andalus* surrounding residue His454 (an amino acid replacement that decreases the positive charge at site 454: Figure 4D), albeit with less increase in collective motions compared to the Pro454 mutation (Figure S2).

In terms of the compactness and rigidity of the catalytic domain (based on the dynamic behaviour of residues falling approximately within 0.8 nm of Arg454), the wild-type G6PD exhibits a narrow distribution with a peak at 0.97 nm (Figure 5). The peak is slightly shifted in G6PD *Union* and G6PD *Andalus*; and even more in G6PD *Salt Lake*, resulting in a non-uniform, broader distribution with two peaks at 1 and 1.05 nm. These findings suggest access of this enzyme region to a higher number of conformations in the mutants, in keeping with the collective motions reported above (Figure S2).

We have also analysed the binding stability of the NADP-s, NADP-c and G6P molecules, resulting from the MD simulations, by measuring the dynamic displacement from their locations (defined by the average root mean square deviation (RMSD) of their atomic coordinates) relative to those of the bound states in the reference cryo-EM structure. As expected, NADP-s, being rather distant from the active site,¹³ is not significantly displaced in either the wild type or in any of the variants, with 100% remaining bound throughout the simulations (Table 2). As for G6P

and NADP-c, our simulations suggest instead that they are more likely to be displaced and to unbind, in agreement with their affinities being lower than that for NADP-s. In the wild type, the fraction of simulation frames in which G6P is bound to the enzyme is 45%, and the fraction of simulation frames in which NADP-c is bound to the enzyme is 72% (Table 2). Thus, the simulated G6P (substrate) and NADP-c (cofactor) dissociation trends are consistent with their functional roles in the G6PD catalytic mechanism. The Arg454Pro mutant hosts a lower population of NADP-c (with RMSD value less than 0.5 nm, i.e. less mobile species) than the wild type. Moreover, compared to the wild-type, NADP-c has a deviation (Average bound RMSD) from the initial structure that is higher in G6PD *Salt Lake* and G6PD *Union*, and comparable to the wild-type protein in G6PD *Andalus*, that has instead a lower deviation with respect to G6P (Table 2).

DISCUSSION

Our patient with G6PD deficiency, whose clinical history was summarised above, presented with neonatal jaundice (NNJ) and severe neonatal anaemia; he has a persistent marked reticulocytosis and has required recurrent blood transfusions (12 in 36 months). Thus, he has a chronic non-spherocytic haemolytic disorder, and therefore his G6PD variant, G6PD *Salt Lake*, is clearly in class A (previously class I) of the WHO classification.² The normal G6PD activity in the proband's mother is not against this classification: in fact, it is in keeping with a previous observation, namely that in several class A variants somatic cell selection after X chromosome inactivation favours cells having the normal G6PD allele on their active X chromosome.¹⁹

Neonatal jaundice is a common manifestation of G6PD deficiency, and it is serious, because it may lead to kernicterus.²⁰ Babies with 7TA repeats in the TATAbox of the *UGT1A1* gene, for which our patient was homozygous, tend to have higher levels of unconjugated bilirubin²¹; however, this baby's jaundice was accounted for mainly by conjugated bilirubin, which may be a presenting feature of metabolic diseases,²² or biliary atresia, or liver disease, for none of which there was any evidence; and within 2 weeks the jaundice cleared. We have found three previous reports of cholestatic jaundice in the neonatal period associated with class A G6PD variants.^{20,23,24} All three, like our patient, were given ursodeoxycholic acid, and the jaundice subsequently cleared; interestingly, all four patients, including ours, had been delivered by caesarean section.

Remarkably, two G6PD variants affecting residue 454 are already known. G6PD *Union* (c.1360C>T; p.Arg454Cys) is a widespread polymorphic variant,⁸ belonging to class B (formerly class II). G6PD *Andalus* (c.1361G>A; p.Arg454His), instead, has been reported only in isolated instances: first, in a patient with favism in Spain⁹; later, as a class A (formerly class I) variant in Japan²⁵; later, in a baby with NNJ in Malaysia,²⁶ and in one patient in India.²⁷

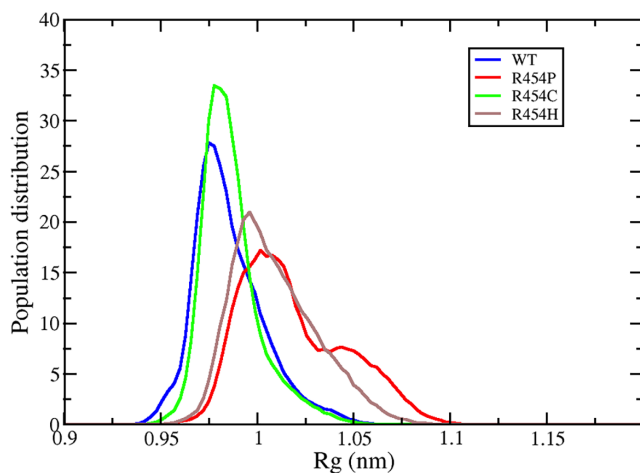


FIGURE 5 Size distribution of Radius of gyration (Rg) for the enzyme catalytic domain in wild-type G6PD and in the three Arg454 replacements. Rg was calculated for all the residues within 0.8 nm of residue 454, and it was averaged over all four G6PD subunits for all frames in the MD simulations. The distribution sharpness (a parameter linked to structural compactness of the catalytic, NADP-c bound, domain) is seen here to correlate with the extent of collective motions (see Figure S2): These are larger in the Pro454 variant than in the *Union* and *Andalus* variants. MD, molecular dynamics.

TABLE 2 Estimates of structural properties regarding ligand binding in Arg454 mutants, compared to wild type.

System	Ligand	Population where RMSD <0.5 nm, %	Average bound RMSD (nm)
Wild type	NADP-s	100	0.18 ± 0.02
	NADP-c	72 ± 10	0.35 ± 0.04
	G6P	45 ± 9	0.46 ± 0.04
Arg454Pro	NADP-s	100	0.19 ± 0.01
	NADP-c	50 ± 9	0.47 ± 0.06
	G6P	55 ± 11	0.44 ± 0.05
Arg454Cys	NADP-s	94 ± 4	0.24 ± 0.03
	NADP-c	57 ± 11	0.40 ± 0.06
	G6P	62 ± 8	0.38 ± 0.04
Arg454His	NADP-s	100	0.18 ± 0.02
	NADP-c	79 ± 7	0.32 ± 0.03
	G6P	74 ± 10	0.31 ± 0.04

Note: The variability figures after the ± sign are standard error of the mean. The values most conspicuously different in mutants compared to wild type are in bold.

The classification of *G6PD Andalus* is not clear, but it is likely to be class B.

In order to explore genotype–phenotype correlations we have performed MD simulations of the wild-type *G6PD* and of the three Arg454 variants. The first important result was that the overall structure of the tetrameric enzyme is not perturbed in any of them; consistent with this finding, the interaction with NADP-s (measured by RMSD) is unaffected (Table 2). This is of interest, because in the majority of class A variants the main cause of *G6PD* deficiency is thought to be in vivo instability⁷ specifically, in many class A variants that have amino acid replacements in the dimerisation interface (largely in exon 10), or in the structural NADP binding region.²⁸ Wang et al. have previously shown that *G6PD Union* and *G6PD Andalus*, obtained as recombinant proteins in *Escherichia coli*, have about one-tenth the activity of the wild-type enzyme, and they have a correspondingly low k_{cat} .²⁹

Molecular dynamics simulations have been used in order to analyse ligand binding and protein stability of a number of both class A variants^{30,31} and class B variants.³² These studies indicate that a variety of different mechanisms can ultimately cause *G6PD* deficiency, and therefore the molecular basis for enzyme deficiency must be assessed for each individual variant. In our case, we were able to focus on three mutants that share the Arg454 replacement: this causes the loss of a positive charge that will directly affect electrostatic interactions and dynamics in the surrounding region of the molecule. Our simulations show that, as a consequence, H-bonds are rearranged (Figure 4), and there is an increase in collective motions that will affect the catalytic region. Ultimately, these changes have a dramatic impact on enzyme activity, since the residual *G6PD* activity in red cells is only 2.7%

of normal in *G6PD Union*,³³ and less than 1% in *G6PD Andalus*⁸; in *G6PD Salt Lake* it was about 3%, in our patient, in the face of marked reticulocytosis; and about 7% in the case of Powers et al.¹⁸

The differences among the three mutants are subtle. The rearrangements of H-bonds are conspicuous in all three cases, although they are qualitatively different. The largest deviation from normal in the interaction with G6P (measured by RMSD) is in *G6PD Andalus*, whereas the largest deviation from normal in the interaction with NADP-c is in the Arg454Pro mutant. In this mutant, we have also found the largest deviation from normal in the population distribution of the radius of gyration (Figure 5). Thus, the adverse effects on catalytic activity may be greater, when compared to the *G6PD Union* and *Andalus* variants, explaining why not only our patient has severe *G6PD* deficiency, but why this causes a chronic haemolytic disorder: that is, *G6PD Salt Lake* is class A rather than class B.¹⁵ A major outstanding challenge for class A variants is that of therapeutic intervention: a detailed understanding of the mechanism causing this phenotype may help in the search for small molecules that may mitigate *G6PD* deficiency.

AUTHOR CONTRIBUTIONS

Simonetta Costa, Giorgia Prontera, Mariannita Gelsomino, Milena Tana, Eloisa Tiberi, Valentina Giorgio, Stefano Mastrangelo and Alberto Romano acquired clinical and haematological data for the research. Angelo Minucci, Amit Kumawat, Maria De Bonis, Maria Elisabetta Onori, Martino Bolognesi and Carlo Camilloni performed enzyme analysis, gene sequencing and molecular dynamics simulations. Simonetta Costa, Angelo Minucci and Lucio Luzzatto wrote the paper. Giovanni Vento, Antonio Ruggiero and Giuseppe Palumbo critically revised the paper. All authors approved the submitted and final version of the paper.

ACKNOWLEDGEMENTS

A.K. and C.C. were supported by Fondazione Telethon—Italy (Grant # GGP19134). A.K. and C.C. acknowledge the EuroHPC Joint Undertaking for awarding this project access to the EuroHPC supercomputer LUMI, hosted by CSC (Finland) and the LUMI consortium through a EuroHPC Regular Access call. Open access funding provided by BIBLIOSAN.

FUNDING INFORMATION

This research received no specific grant from any funding.

CONFLICT OF INTEREST STATEMENT

The authors have nothing to disclose.

PATIENT CONSENT STATEMENT

The parents of the infant have provided written informed consent for investigations and genetic analyses, as required by our institutional review board.

ORCID

Simonetta Costa  <https://orcid.org/0000-0003-0980-5297>

Lucio Luzzatto  <https://orcid.org/0000-0003-4663-8790>

REFERENCES

- Minucci A, Moradkhani K, Hwang MJ, Zuppi C, Giardina B, Capoluongo E. Glucose-6-phosphate dehydrogenase (G6PD) mutations database: review of the 'old' and update of the new mutations. *Blood Cells Mol Dis*. 2012;48(3):154–65. <https://doi.org/10.1016/j.bcmd.2012.01.001>
- Luzzatto L, Ally M, Notaro R. Glucose-6-phosphate dehydrogenase deficiency. *Blood*. 2020;136(11):1225–40. <https://doi.org/10.1182/blood.2019000944>
- Geck RC, Powell NR, Dunham MJ. Functional interpretation, cataloging, and analysis of 1,341 glucose-6-phosphate dehydrogenase variants. *Am J Hum Genet*. 2023;110(2):228–39. <https://doi.org/10.1016/j.ajhg.2023.01.003>
- Luzzatto L, Nannelli C, Notaro R. Potentially pathogenic and pathogenic G6PD variants. *Am J Hum Genet*. 2023;110(11):1983–5. <https://doi.org/10.1016/j.ajhg.2023.10.003>
- Luzzatto L, Bancone G, Dugué PA, Jiang W, Minucci A, Nannelli C, et al. New WHO classification of genetic variants causing G6PD deficiency. *Bull World Health Organ*. 2024;102(8):615–7. <https://doi.org/10.2471/BLT.23.291224>
- Minucci A, Giardina B, Zuppi C, Capoluongo E. Glucose-6-phosphate dehydrogenase laboratory assay: how, when, and why? *IUBMB Life*. 2009;61(1):27–34. <https://doi.org/10.1002/iub.137>
- Mason PJ, Bautista JM, Gilsanz F. G6PD deficiency: the genotype-phenotype association. *Blood Rev*. 2007;21(5):267–83. <https://doi.org/10.1016/j.blre.2007.05.002>
- Rovira A, Vulliamy TJ, Pujades A, Luzzatto L, Corrons JL. The glucose-6-phosphate dehydrogenase (G6PD) deficient variant G6PD Union (454 Arg → Cys) has a worldwide distribution possibly due to recurrent mutation. *Hum Mol Genet*. 1994;3(5):833–5. <https://doi.org/10.1093/hmg/3.5.833>
- Vives-Corrons JL, Kuhl W, Pujades MA, Beutler E. Molecular genetics of the glucose-6-phosphate dehydrogenase (G6PD) Mediterranean variant and description of a new G6PD mutant, G6PD Andalus1361A. *Am J Hum Genet*. 1990;47(3):575–9.
- Minucci A, Delibato E, Castagnola M, Concolino P, Meglio F, Zuppi C, et al. Identification of RFLP G6PD mutations by using microcapillary electrophoretic chips (Experion). *J Sep Sci*. 2008;31(14):2694–700. <https://doi.org/10.1002/jssc.200800216>
- Minucci A, Concolino P, Vendittelli F, Giardina B, Zuppi C, Capoluongo E. Glucose-6-phosphate dehydrogenase Buenos Aires: a novel de novo missense mutation associated with severe enzyme deficiency. *Clin Biochem*. 2008;41(9):742–5. <https://doi.org/10.1016/j.clinbiochem.2007.11.009>
- Minucci A, Ruggiero A, Canu G, Maurizi P, De Bonis M, Concolino P, et al. Co-inheritance of G6PD and PK deficiencies in a neonate carrying a novel UGT1A1 genotype associated to Crigler-Najjar type II syndrome. *Pediatr Blood Cancer*. 2015;62(9):1680–1. <https://doi.org/10.1002/pbc.25500>
- Wei X, Kixmoeller K, Baltrusaitis E, Yang X, Marmorstein R. Allosteric role of a structural NAD⁺ molecule in glucose-6-phosphate dehydrogenase activity. *Proc Natl Acad Sci U S A*. 2022;119(29):e2119695119. <https://doi.org/10.1073/pnas.2119695119>
- Abraham MJ, Murtola T d, Schulz R, Páll S, Smith JC, Hess B, et al. GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *Software*. 2015;1–2:19–25.
- Huang J, MacKerell AD Jr. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem*. 2013;34(25):2135–45. <https://doi.org/10.1002/jcc.23354>
- Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, et al. CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J Comput Chem*. 2010;31(4):671–90. <https://doi.org/10.1002/jcc.21367>
- Bernetti M, Bussi G. Pressure control using stochastic cell rescaling. *J Chem Phys*. 2020;153(11):114107. <https://doi.org/10.1063/5.0020514>
- Powers JL, Best DH, Grenache DG. Genotype-phenotype correlations of glucose-6-phosphate-deficient variants throughout an activity distribution. *J Appl Lab Med*. 2018;2(6):841–50. <https://doi.org/10.1373/jalm.2017.024935>
- Filosa S, Giacometti N, Wangwei C, De Mattia D, Pagnini D, Alfinito F, et al. Somatic-cell selection is a major determinant of the blood-cell phenotype in heterozygotes for glucose-6-phosphate dehydrogenase mutations causing severe enzyme deficiency. *Am J Hum Genet*. 1996;59(4):887–95.
- Ben Fredj D, Barro C, Joly P, Thomassin N, Collardeau-Frachon S, Plantaz D, et al. Transient liver injury and severe neonatal cholestasis in infant with glucose-6-phosphate dehydrogenase deficiency due to a new mutation. *Arch Pediatr*. 2019;26(6):370–3. <https://doi.org/10.1016/j.arcped.2019.05.005>
- Keffler S, Kelly DA, Powell JE, Green A. Population screening for neonatal liver disease: a feasibility study. *J Pediatr Gastroenterol Nutr*. 1998;27(3):306–11. <https://doi.org/10.1097/00005176-199809000-00007>
- Dellert SF, Balistreri WF. Neonatal cholestasis. In: Walker A, Durie P, Hamilton J, Walker Smith J, editors. *Pediatric gastroenterology*. 3rd ed. Lewiston, NY: B.C. Decker, Inc; 2000. p. 880–94.
- Kordes U, Richter A, Santer R, Schäfer H, Singer D, Sonntag J, et al. Neonatal cholestasis and glucose-6-P-dehydrogenase deficiency. *Pediatr Blood Cancer*. 2010;54(5):758–60. <https://doi.org/10.1002/pbc.22390>
- Mizukawa B, George A, Pushkaran S, Weckbach L, Kalinyak K, Heubi JE, et al. Cooperating G6PD mutations associated with severe neonatal hyperbilirubinemia and cholestasis. *Pediatr Blood Cancer*. 2011;56(5):840–2. <https://doi.org/10.1002/pbc.22744>
- 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 nonradioisotopic single-strand conformation polymorphism analysis. *Blood*. 1994;83(11):3363–8.
- Ainon O, Yu YH, Amir Muhriz AL, Boo NY, Cheong SK, Hamidah NH. Glucose-6-phosphate dehydrogenase (G6PD) variants in Malaysian Malays. *Hum Mutat*. 2003;21(1):101. <https://doi.org/10.1002/humu.9103>
- Arunachalam AK, Sumithra S, Maddali M, Fouzia NA, Abraham A, George B, et al. Molecular characterization of G6PD deficiency: report of three novel G6PD variants. *Indian J Hematol Blood Transfus*. 2020;36(2):349–55. <https://doi.org/10.1007/s12288-019-01205-7>
- Horikoshi N, Hwang S, Gati C, Matsui T, Castillo-Orellana C, Raub AG, et al. Long-range structural defects by pathogenic mutations in most severe glucose-6-phosphate dehydrogenase deficiency. *Proc Natl Acad Sci U S A*. 2021;118(4):e2022790118. <https://doi.org/10.1073/pnas.2022790118>
- Wang XT, Chan TF, Lam VM, Engel PC. What is the role of the second 'structural' NAD⁺-binding site in human glucose 6-phosphate dehydrogenase? *Protein Sci*. 2008;17(8):1403–11. <https://doi.org/10.1110/ps.035352.108>
- Rani S, Malik FP, Anwar J, Zafar PR. Investigating effect of mutation on structure and function of G6PD enzyme: a comparative molecular dynamics simulation study. *PeerJ*. 2022;10:e12984. Published 2022 Mar 29. <https://doi.org/10.7717/peerj.12984>
- Alakbaree M, Abdulsalam AH, Ahmed HH, Ali FH, Al-Hili A, Omar MSS, et al. A computational study of structural analysis of class I human glucose-6-phosphate dehydrogenase (G6PD) variants: elaborating the correlation to chronic non-spherocytic hemolytic anemia (CNSHA). *Comput Biol Chem*. 2023;104:107873. <https://doi.org/10.1016/j.compbiolchem.2023.107873>
- Sirdah M, Reading NS, Vankayalapati H, Prchal JT. A computational study of structural differences of binding of NAD⁺ and G6P substrates to G6PD Mediterraneanc.563T, G6PD A-c.202A/c.376G, G6PD Cairoc.404C and G6PD Gazac.536A mutations. *Blood Cells*

- Mol Dis. 2021;89:102572. <https://doi.org/10.1016/j.bcmd.2021.102572>
33. Nannelli C, Bosman A, Cunningham J, Dugué PA, Luzzatto L. Genetic variants causing G6PD deficiency: clinical and biochemical data support new WHO classification. Br J Haematol. 2023;202(5):1024–32. <https://doi.org/10.1111/bjh.18943>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Costa S, Minucci A, Kumawat A, De Bonis M, Prontera G, Gelsomino M, et al. Pathogenic *G6PD* variants: Different clinical pictures arise from different missense mutations in the same codon. Br J Haematol. 2024;00:1–10. <https://doi.org/10.1111/bjh.19775>