PNAS RESEARCH ARTICLE | BIOPHYSICS AND COMPUTATIONAL BIOLOGY

Allosteric role of a structural NADP$^+$ molecule in glucose-6-phosphate dehydrogenase activity

Xuepeng Wei[a,b,1], Kathryn Kixmoeller[c], Elana Baltrusaitis[c], Xiaolu Yang[d], and Ronen Marmorstein[a,b,c,1]

Edited by Anton Berns, Antoni van Leeuwenhoek Nederlands Kanker Instituut, Amsterdam, Netherlands; received October 28, 2021; accepted March 20, 2022

Human glucose-6-phosphate dehydrogenase (G6PD) is the main cellular source of NADPH, and thus plays a key role in maintaining reduced glutathione to protect cells from oxidative stress disorders such as hemolytic anemia. G6PD is a multimeric enzyme that uses the cofactors β-D-glucose 6-phosphate (G6P) and "catalytic" NADP$^+$ (NADP$^+$c), as well as a "structural" NADP$^+$ (NADP$^+$s) located ∼25 Å from the active site, to generate NADPH. While X-ray crystallographic and biochemical studies have revealed a role for NADP$^+$s in maintaining the catalytic activity by stabilizing the multimeric G6PD conformation, other potential roles for NADP$^+$s have not been evaluated. Here, we determined the high resolution cryo-electron microscopy structures of human wild-type G6PD in the absence of bound ligands and a catalytic G6PD-D200N mutant bound to NADP$^+$c and NADP$^+$s in the absence or presence of G6P. A comparison of these structures, together with previously reported structures, reveals that the unliganded human G6PD forms a mixture of dimers and tetramers with similar overall folds, and binding of NADP$^+$s induces a structural ordering of a C-terminal extension region and allosterically regulates G6P binding and catalysis. These studies have implications for understanding G6PD deficiencies and for therapy of G6PD-mediated disorders.

G6PD | NADPH | allostery | cryo-EM

Significance

Glucose-6-phosphate dehydrogenase (G6PD) is the main cellular source of NADPH and plays a key role in maintaining reduced glutathione to protect cells from oxidative stress disorders such as hemolytic anemia. High resolution cryo-electron microscopy structures of human G6PD were determined in the absence of bound ligands and bound to catalytic (NADP$^+$c) and structural (NADP$^+$s) NADP$^+$ in the absence or presence of G6P. A comparison of these structures with previously reported structures reveals that unliganded human G6PD forms a mixture of dimers and tetramers with similar overall folds, and binding of NADP$^+$s induces a structural ordering of a C-terminal extension region and allosterically regulates G6P binding and catalysis. These studies have implications for understanding G6PD deficiencies and for therapy of G6PD-mediated disorders.

Human glucose-6-phosphate dehydrogenase (G6PD) catalyzes the rate limiting and first step of the pentose phosphate pathway, converting β-D-glucose 6-phosphate (G6P) to 6-phosphoglucono-lactone and simultaneously reducing NADP$^+$ to NAPDH (1, 2). Since G6PD is the main cellular source of NADPH, it plays a key role in protecting cells from oxidative stress, including the use of NAPDH to reduce glutathione, the cell's primary defense against oxidative stress injury (3, 4). G6PD deficiency is the most common human enzymatic deficiency disorder, found primarily in males due to G6PD being found on the X-chromosome. G6PD deficiency primarily manifests with hemolytic anemia triggered by exposure of red blood cells (RBCs) to oxidizing agents (including anti-malarial drugs, other medications, or consumption of fava beans) (5–8). RBCs are particularly susceptible to the effects of G6PD deficiency since they lack mitochondria and therefore rely on G6PD as their sole source of NADPH. The persistence of G6PD deficiency in humans can be explained by the fact that the increased oxidative stress resulting from these mutations confer some resistance to the malaria parasite, and G6PD deficiency is primarily found in humans originating from areas in which malaria is endemic (9–11). Overexpression of G6PD has also been implicated in the pathogenesis of human cancer (12, 13). Thus, the fine balance of G6PD levels and activity are critical for cell viability.

Structural and biochemical studies of G6PD have provided important insights into its mode of action. Structure determination of human G6PD by X-ray crystallography demonstrate that it exists as a tetrameric dimer of dimers with each subunit containing a classical β+α coenzyme binding domain at the N terminus (residues 31–200) that mediates binding to a catalytic NADP$^+$ (NADP$^+$c) molecule and a larger β+α domain at the C terminus (residues 201–515) that binds a structural NADP+ (NADP$^+$s) molecule, proximal to the dimer interface (14). An X-ray crystal structure bound to G6P reveals that it binds at the interface of the N- and C-terminal subdomains proximal to NADP$^+$c (15). Many G6PD inactivating mutations map to the binding site for NADP$^+$s, and a recent study demonstrates that these mutations have long-range effects on G6P binding (16), although the detailed molecular mechanism for this has not been resolved. In solution, G6PD has been shown to be in equilibrium between monomeric, dimeric, and tetrameric forms. This equilibrium could be affected by ionic strength, pH, and the tetramer was reported to have maximal catalytic activity (17–19), leading to questions about the roles of G6PD multimerization and the structural NADP$^+$ molecule on enzymatic activity. Significantly, all high-resolution structures of G6PD have been determined with X-ray crystallography, with the potential of crystal packing forces playing a role in the observed multimerization properties. In order to address the issues noted above,

Author contributions: X.W., K.K., E.B., X.Y., and R.M. designed research; X.W., K.K., and E.B. performed research; X.W., K.K., E.B., X.Y., and R.M. analyzed data; and X.W. and R.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: marmor@upenn.edu or wei_xuepeng@gzlab.ac.cn.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2119695119/-/DCSupplemental.

Published July 12, 2022.

we set out to determine the cryo-electron microscopy (cryo-EM) structures of G6PD alone and in various liganded states. Here, we report the high-resolution cryo-EM structures of wild-type G6PD alone and a catalytically defective G6PD mutant (G6PD-D200N) in complex with structural and catalytic $NADP^+$, with or without G6P. A comparison of these structures to each other and to previously reported crystal structures of G6PD provides insights into G6PD multimerization and the allosteric role of the structural $NADP^+$ molecule in G6PD activity.

Results

Cryo-EM Structure of Unliganded G6PD Shows a Mixture of G6PD Dimers and Tetramers. To determine the cryo-EM structures of human G6PD, we first prepared recombinant unliganded wild-type protein (G6PD-WT) in bacteria for our initial analysis. Optimization of the cryo-EM grids resulted in well-dispersed particles resulting in the generation of high quality two-dimensional (2D) class averages from ~570,000 particles, confirming that G6PD formed both dimers and tetramers, as previously reported (17) (*SI Appendix*, Fig. S1*A* and Fig. 1*A* and *B*). Notably, G6PD monomers were not observed, despite a report that they may form in solution (17). Three-dimensional (3D) reconstructions of the 2D classes corresponding to G6PD dimers and tetramers were carried out using CRYOSPARC2 (*SI Appendix*, Fig. S1*B* and Table 1) resulting in G6PD C2-symmetric dimer and D2-symmetric tetramer structures to overall resolutions of 3.5 Å and 2.8 Å , respectively (*SI Appendix*, Fig. S2 and Table 1). As previously reported, the tetramer forms a side-by-side dimer of dimers (14), and a superposition of the dimer and tetramer reveals that there are no significant structural changes between the G6PD dimer and tetramer with an rmsd of 0.83 Å (Fig. 1*A*–*C*). Given

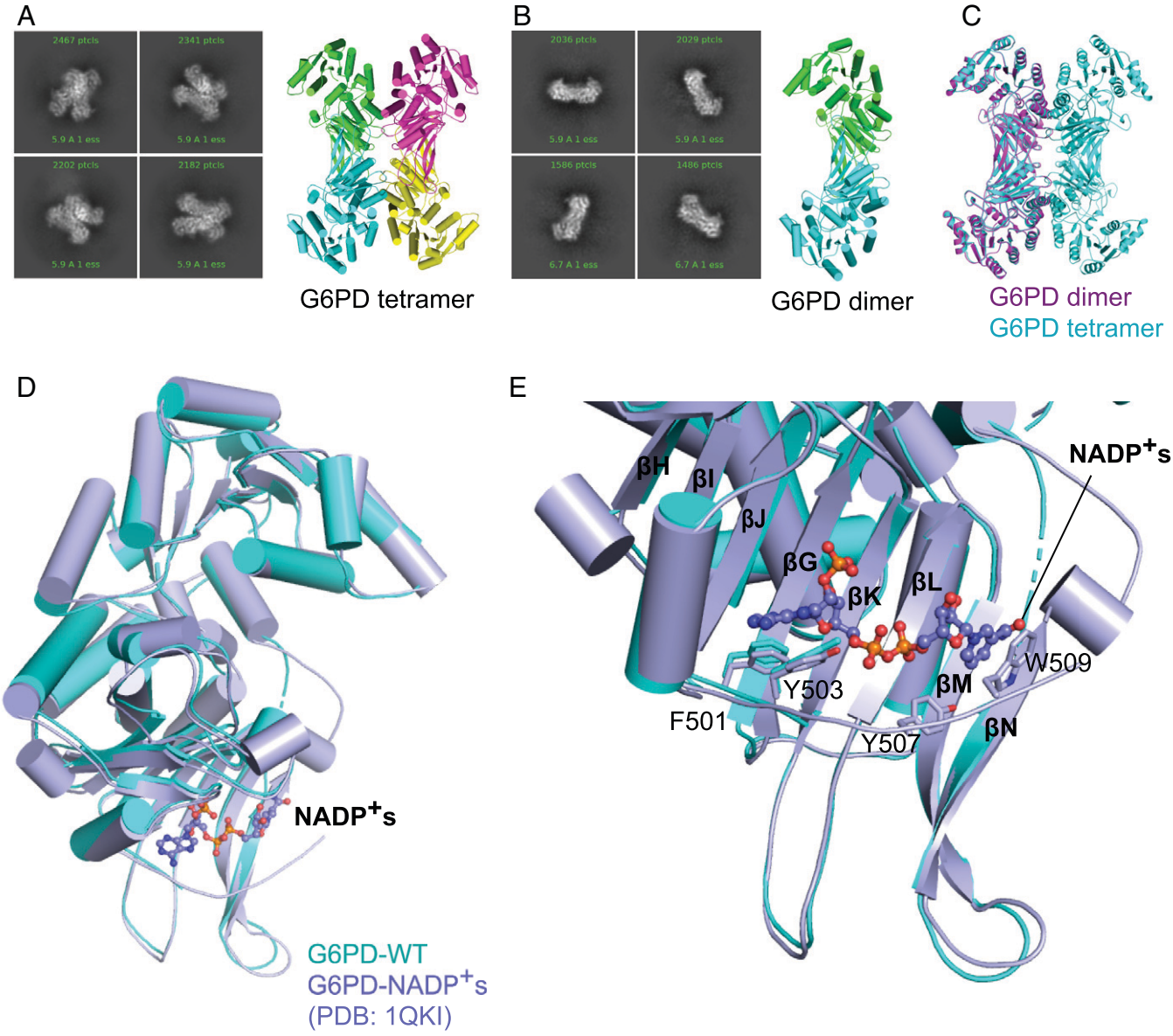

Fig. 1. Overall cryo-EM structure of G6PD without bound cofactors. (*A*) Selected 2D class averages of G6PD-WT tetramer (*Left*) and cartoon representation of G6PD tetramer structure (*Right*). The four monomers are colored in green, cyan, yellow and magenta. (*B*) Selected 2D class averages of G6PD-WT dimer (*Left*) and cartoon representation of G6PD dimer structure (*Right*). The two monomers are colored in green and cyan. (*C*) Superposition of G6PD dimer (purple) and tetramer (blue). (*D*) Overall superposition of unliganded G6PD-WT (cyan) and $NADP^+$s-bound G6PD subunit (light blue) crystal structure (PDB:1QKI). (*E*) Close-up of unliganded G6PD-WT and a structural $NADP^+$s-bound G6PD subunit crystal structure (PDB:1QKI). C-terminal extension residues that play roles in $NADP^+$s binding are highlighted. Stands of the C-domain β-sheet region (G–N) are labeled.

Table 1. Cryo-EM data collection, refinement, and validation statistics

	G6PD-WT-tetramer (EMDB: 25225; (PDB: 7SNG)	G6PD-WT-dimer (EMDB: 25224; PDB: 7SNF)	G6PD-D200N-NADP+ (EMDB: 25226; PDB: 7SNH)	G6PD-D200N-NADP+-G6P (EMDB: 25227; PDB: 7SNI)
Data collection and processing				
Magnification	10,500	10,500	10,500	10,500
Voltage (kV)	300	300	300	300
Electron exposure (e–/Å^2)	40	40	40	40
Defocus range (μm)	1.0–2.0	1.0–2.0	1.0–2.0	1.0–2.0
Pixel size (Å)	0.83	0.83	0.83	0.83
Symmetry imposed	D2	C2	D2	D2
Initial particle images (no.)	570,405	570,405	866,530	356,519
Final particle images (no.)	139,200	76,546	243,562	67,823
Map resolution (Å)	2.8	3.5	2.2	2.5
FSC threshold	0.143	0.143	0.143	0.143
Refinement				
Initial model used (PDB code)	2BH9	2BH9	2BH9	2BH9
Model resolution (Å)	2.9	4.0	2.2	2.8
FSC threshold	0.5	0.5	0.5	0.5
Map sharpening *B* factor (Å^2)	−123.8	−156.0	−76.2	−92.2
Model composition				
Nonhydrogen atoms	15,204	7,588	16,078	16,533
Protein residues	1,868	938	1,904	1,952
Ligands	0	0	8	12
B factors (Å^2)				
Protein	44.17	52.82	32.66	37.54
Ligand	—	—	37.89	37.36
rmsd				
Bond lengths (Å)	0.005	0.004	0.009	0.012
Bond angles (°)	0.713	0.663	0.701	0.842
Validation				
MolProbity score	1.70	2.17	1.44	1.47
Clashscore	10.14	17.53	6.37	8.17
Poor rotamers (%)	0.43	0	1.20	0.06
Ramachandran plot				
Favored (%)	96.98	93.44	98.25	97.89
Allowed (%)	3.02	6.45	1.75	2.11
Disallowed (%)	0	0.11	0	0

the concordance between the dimer and tetramer structures and the higher resolution afforded by the tetramer structure, we will focus the rest of our analysis on the tetramer structure.

Structural NADP$^+$ Binding Does Not Significantly Alter Overall G6PD Structure But Instead Orders the C-Terminal Extension Region. A surprise of the G6PD-WT tetramer and dimer structures is that they do not contain any bound ligands, thus representing unliganded human G6PD structures. This was confirmed by analyzing the G6PD tetramer and dimer structures with D2 and C2 symmetry, respectively (*SI Appendix*, Fig. S2*C*), or without imposing symmetry (*SI Appendix*, Fig. S3 and Table S1). This is in contrast to previous reports demonstrating that recombinant G6PD is typically isolated with a tightly bound structural NADP$^+$ (NADP$^+$s) molecule (14, 15). A superposition of G6PD-WT with a previously reported crystal structure of a tetramer G6PD Canton R459L variant bound to NADP$^+$s (1QKI) (14) reveals a nearly identical overall subunit structure (Fig. 1*D*) with an rmsd of 0.98 Å for residues 28–504, demonstrating that the structural NADP$^+$ molecule does not play a significant role in the overall G6PD subunit structure. Also of note, C-terminal extension residues 505–511 of G6PD-WT are disordered, while these residues are well ordered and play an important role in NADP$^+$s binding in 1QKI, from residues Phe501, Tyr503, Tyr507, and Trp509 (Fig. 1*E*), while residues from the G6PD C-domain β-sheet region (strands G-N) also participate in structural NADP$^+$ binding. This demonstrates that the C-terminal extension of G6PD adopts structure as a function of NADP$^+$s binding and is consistent with the proposed role of NADP$^+$s in G6PD stability (14, 16, 20).

Recently reported crystal structures of G6PD harboring class I G6PD deficiency mutant associated with more than a 90% loss of G6PD activity in hemolytic anemias (P396L, R393H, V394L, W509A) also reveal a disordered C-terminal extension and the absence of a NADP$^+$s molecule, along with an altered C-terminal β-sheet region for the P396L, R393H, and V394L mutants (16). This observation is consistent with the importance of both the C-terminal β-sheet region and ordering of the C-terminal extension for NADP$^+$s binding.

Structure of a NADP$^+$ Catalytic Mutant Reveals That Binding of Catalytic G6P and NADP$^+$ Cofactors Do Not Significantly Change the G6PD Structure. While structures of G6PD bound to NADP$^+$s with either the catalytic NADP$^+$ (NADP$^+$c) (14) or G6P (15) cofactors have been determined, the absence of a structure with structural NADP$^+$ (NADP$^+$s) and both cofactors bound has not been determined, thus hampering a molecular

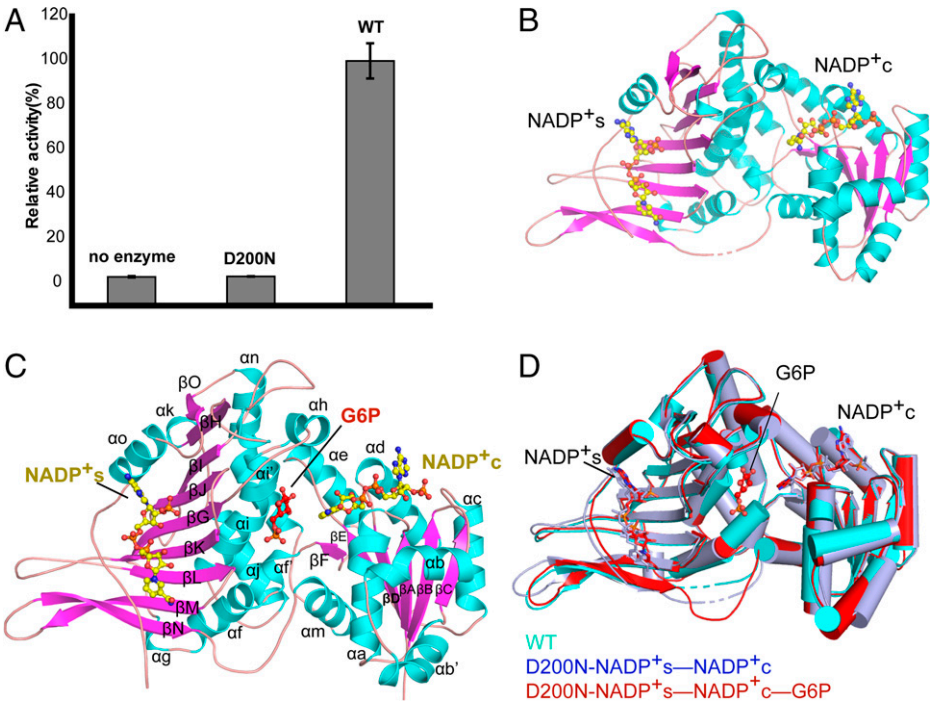

Fig. 2. Overall cryo-EM structures of G6PD-D200N mutant with bound cofactors. (*A*) The enzymatic activity of G6PD-D200N mutant relative to G6PD-WT. The no-enzyme control is indicated. (*B*) Cartoon representation of G6PD-D200N bound to NADP$^+$s and NADP$^+$c. (*C*). Cartoon representation of G6PD-D200N bound to NADP$^+$s, NADP$^+$c and G6P. (*D*) Overall superposition of unliganded G6PD-WT (cyan), G6PD-D200N-NADP$^+$s-NADP$^+$c (blue) and D200N-NADP$^+$s-NADP$^+$c-G6P (red).

understanding of the interdependence of NADP$^+$s and the catalytic cofactors. To address this complication, we prepared a catalytic G6PD-D200N mutant to facilitate the assembly of a fully liganded G6PD complex that is "trapped" in a state bound to substrates since it is unable to process a reaction to products. D200 was previously shown to be part of the catalytic dyad (together with His263) required for catalysis (1, 15). The recombinant G6PD-D200N mutant showed the same elution profile from gel filtration as G6PD-WT, consistent with its native-like folding properties, and an enzymatic assay was used to confirm that G6PD-D200N was significantly compromised for catalytic activity (Fig. 2*A*).

We prepared G6PD-D200N (13 μM) with saturating concentrations of NADP$^+$ (3 mM) and two different concentrations of G6P (3 mM and 100 mM) for cryo-EM structure determination. Incubation with the lower G6P concentration resulted in a structure containing both NADP$^+$s and NADP$^+$c molecules but no bound G6P (G6PD-D200N-NADP$^+$s-NADP$^+$c), while incubation with the higher G6P concentration resulted in a complex with all ligands bound (G6P-D200N-NADP$^+$s-NADP$^+$c-G6P), to resolutions of 2.2 Å and 2.5 Å for the tetrameric complexes refined with D2 symmetry, respectively (*SI Appendix*, Fig. S4 and Fig. 2*B* and *C* and Table 1). Notably, analysis of the data from the G6P-D200N-NADP$^+$s-NADP$^+$c-G6P complex to 2.9 Å resolution without imposing symmetry, confirmed the presence of all three ligands bound to each of the subunits of the tetramer (*SI Appendix*, Fig. S5 and Table S1). Unlike unliganded G6PD-WT, both liganded complex particles showed predominately the tetrameric form, likely due to the lower pH created by the addition of 3 mM NADP$^+$ to the liganded complexes, consistent with previous findings (17–19). A superposition of these structures with G6PD-WT reveals no significant overall structural changes (Fig. 2*D* and *SI Appendix*, Table S2), demonstrating that the binding of both structural and catalytic NADP$^+$ molecules and G6P do not alter the overall G6PD structure. Importantly, an overlay of the evolutionarily conserved R_{198}IDHYLGK$_{205}$ sequence of G6PD-D200N-NADP$^+$s-NADP$^+$c-G6P and a G6PD Canton R459L variant bound to NADP$^+$s (1QKI) superimposes almost perfectly demonstrating no significant local alterations caused by the D200N mutation (*SI Appendix*, Table S2). Interestingly, in the G6PD-D200N-NADP$^+$s-NADP$^+$c-G6P structure, the nicotinamide of NADP$^+$c is too far for hydride transfer from G6P (~5.5 Å) (Fig. 2*C*), as it is for other reported G6P structures with bound NADP$^+$c with modeled G6P, or bound G6P with modeled NADP$^+$c, suggesting that additional structural changes may be required for catalysis.

G6P Is Bound in a Position That Could Be Influenced by the Binding of Structural NADP$^+$. The structure of the G6P-D200N-NADP$^+$s-NADP$^+$c allows for evaluation of the molecular protein interactions with both the structural and catalytic NADP$^+$ molecules. What we observe is largely consistent with previously reported crystal structures of G6PD bound to NADP$^+$ molecules (14, 15). Briefly, NADP$^+$s, sits on top and across the C-terminal β-sheet region (strands G-N), which mediates several hydrogen bonding (Arg357, Lys366, Arg370, Arg393, Tyr401, Asp42, and Thr423) and a few van der Waals (Lys238 and Ala399) interactions (Fig. 3*A*). In addition, NADP$^+$s is surrounded by additional van der Waals interactions from the C-terminal extension (Phe501, Tyr503, Tyr507, Tyr507, and Trp509), hydrogen bonding (Arg487) and van der Waals (Met496) interactions from the O helix, and a hydrogen bond from a preceding loop (Asp493) (Fig. 3*A*).

The NADP$^+$c molecule, located within the coenzyme binding domain, is ~30 Å away from NADP^+_S and also makes extensive interactions with G6PD. Interactions are made by residues at the N-terminal end of helix a, harboring the GxxGDLX dinucleotide binding fingerprint. This region mediates hydrogen bonding (Ser40, Asp42, and Leu43) and van der Waals (Gly41 and Val46) interactions. Several interactions are made through strand-helix

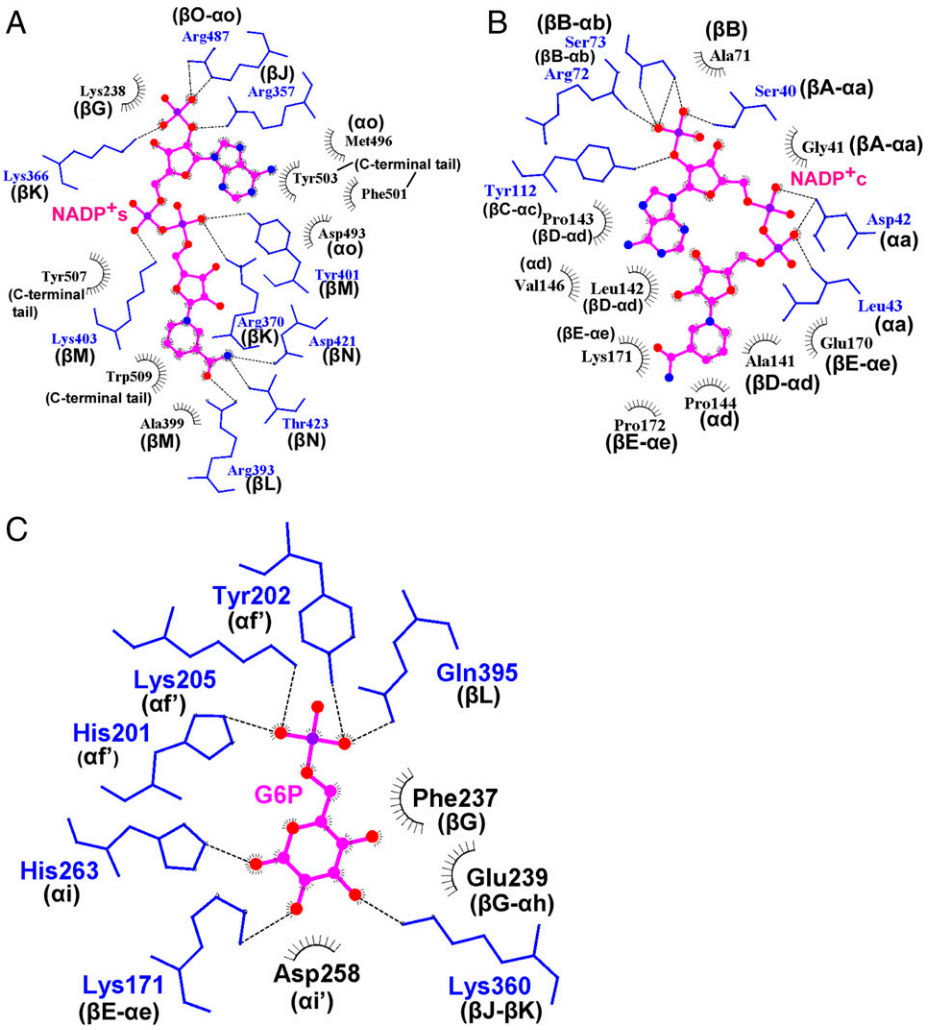

Fig. 3. Detailed interactions of G6PD with bound cofactors. Ligplot representation highlighting protein residue interactions with NADP$^+$s (*A*), NADP$^+$c (*B*), and G6P (*C*). Relevant secondary structural elements are also indicated.

connecting loops via hydrogen bond (Arg72, Ser73, and Try112) and van der Waals (Als71, Ala141, Leu142, Pro143, Pro144, Glu170, and Pro172) interactions (Fig. 3*B*).

The structure of the G6PD-D200N-NADP$^+$s-NADP$^+$c-G6P complex shows that G6P sits between the two NADP$^+$ molecules, ~7 Å away from NADP$^+$c and ~19 Å away from NADP$^+$s. Despite being further away from NADP$^+$s than NADP$^+$c, G6P is in a position to be more influenced by NADP$^+$s binding since all but one (hydrogen bond from Lys171) of the G6P interacting residues sit on loops of the C-terminal β-sheet that directly interacts with NADP$^+$s or secondary structural elements that sit against the C-terminal β-sheet (helices i and f′) (Fig. 3*C*). Within helix f′, hydrogen-bonds (His201, Tyr202, and Lys205) are mediated by the evolutionarily conserved R_{198}IDHYLGK$_{205}$ sequence. Within helix i, Asp258 makes a van der Waals interaction and His263 (an important catalytic residue) makes a hydrogen bond. Phe237 and Glu239 from a strand of the β-sheet and the following loop, respectively, make van der Waals interactions with G6P; and Lys360 and Glu395 from loops of the β-sheet also make hydrogen bond interactions with G6P.

Structural NADP$^+$ Binding Allosterically Regulates G6P Binding and Catalysis. The G6PD-WT and G6PD-D200N-NADP$^+$s-NADP$^+$c-G6P structures reported here, together with the previously reported crystal structure of G6PD bound to NADP$^+$s (1QKI) (14), allow us to directly evaluate the role of NADP$^+$s binding on the binding of NADP$^+$c and G6P cofactors. A close-up view of the NADP$^+$c binding site reveals that the position of NADP$^+$c interacting residues are largely maintained (Fig. 4*A*), demonstrating that NADP$^+$s binding does not significantly influence the binding of NADP$^+$c.

The same superposition as above reveals significant differences around the G6P binding site. Most dramatically, loop residues 201–204 in G6PD-WT transition into a short f′ helix in the NADP$^+$s-bound structures, shifting the position of residues His201, Tyr202, and Lys205 from being incompatible with G6P binding to positions that allow these residues to interact directly with G6P in the NADP$^+$s-bound structures (Fig. 4*B* and *SI Appendix*, Fig. S6). Specifically, in the transition from GSPD-WT to the NADP$^+$s-bound structures His201 and Lys-205 flip by ~ 90° and 120°, respectively, from a position away from G6P to G6P interacting positions, and Tyr202 flips by ~ 90° from a position that would sterically clash with G6PD to a position that interacts with G6P. Apparently correlated with the position of Tyr202 in G6PD-WT, the catalytic His263 flips by ~ 90° from a catalytically incompetent to competent position. Notably, each of these residues is strictly conserved across evolution, consistent with their importance for G6PD activity regulation.

To better understand the molecular basis of how NADP$^+$s binding may propagate these structural changes around the G6P binding site, we looked more closely at the back side of

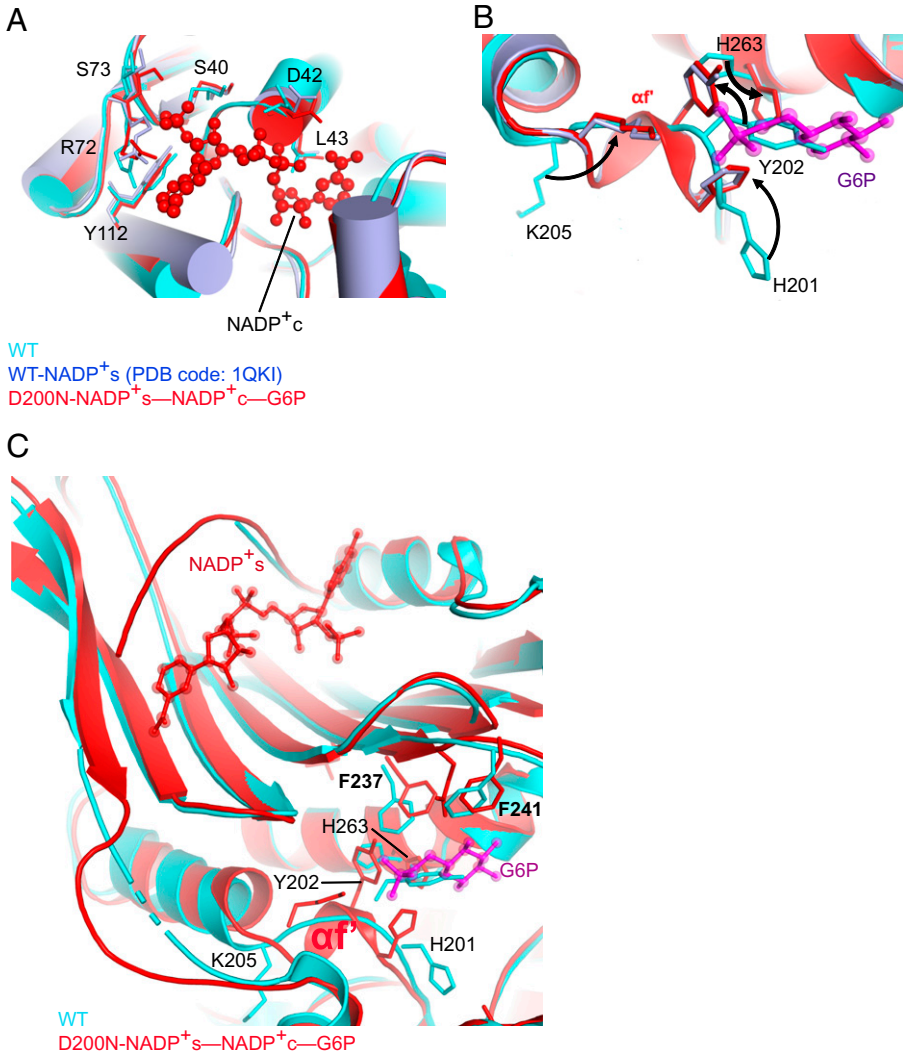

Fig. 4. Molecular basis for allosteric regulation of G6PD by $NADP^+$s. (*A*) Close-up view of the $NADP^+$c binding site superimposing key interaction residues from G6PD-WT (cyan), $NADP^+$s bound G6PD (PDB:1QKI, blue), and G6PD-D200N-$NADP^+$s-$NADP^+$c-G6P (red). (*B*) Superposition of G6PD-D200N-$NADP^+$s-$NADP^+$c-G6P, G6PD-WT, and 1QKI, highlighting residue interactions with G6P. The arrows indicate movement of key residues from G6PD-WT to the $NADP^+$s bound forms. (*C*). Superposition of G6PD-WT and G6PD-D200N-$NADP^+$s-$NADP^+$c-G6P, highlighting residues that reorient upon cofactor binding; His201, Tyr202, Lys205, Phe237, Phe241, and His263. We propose that $NADP^+$s-mediated reorientation of residues Phe237 and Phe241 nucleate the repositioning of G6P binding residues His201, Tyr202, and Lys205.

the C-terminal domain β-sheet region that interacts with $NADP^+$s. Residues in this region that adopted notably different positions in G6PD-WT relative to the $NADP^+$s-bound structures were Phe237 and Phe241, which are both positioned closer to the G6P binding site in the G6PD-WT structure (Fig. 4*C*). This movement appears to nucleate the rearrangement of loop residues 201–205 in G6PD-WT, which precludes G6P binding. Phe237 appears to play a particularly important role in this as it would directly clash with the Tyr202 position that is observed when G6P is bound (Fig. 4C). Given that G6PD mutations associated with hemolytic anemias typically retain some residual G6PD activity, the fact that mutations of Phe237, Phe241, and loop residues 201–205 are not observed is consistent with their essential roles in G6PD activity (21, 22). Taken together, these structural comparisons suggest that $NADP^+$s-binding reorients the β-sheet region of the C-terminal domain in a way that allosterically reorients key G6P binding residues within the 201–205 loop to form a short f′ helix, placing residues His201, Tyr202, and Lys205 in positions that are compatible with G6P binding and subsequent catalysis. Given that the binding site for $NADP^+$c is unchanged in the presence or absence of $NADP^+$s, we propose that $NADP^+$c binds independently of $NADP^+$s, although later due to its lower relative affinity. A schematic model for this regulation is depicted in Fig. 5.

Discussion

Here, we determined high resolution cryo-EM structures of G6PD, including the structure of human G6PD in the absence of bound ligands. We also exploited a catalytic G6PD-D200N mutant to capture the cryo-EM structure of human G6PD bound to all ligands required for activity, catalytic $NADP^+$c and G6P and structural $NADP^+$s. A comparison of these human G6PD structures, together with the previously reported structures (14, 15), reveals that unliganded G6PD forms a mixture of dimers and tetramers, and binding of $NADP^+$s induces a structural ordering of a C-terminal extension region and allosterically regulates G6P binding and catalysis.

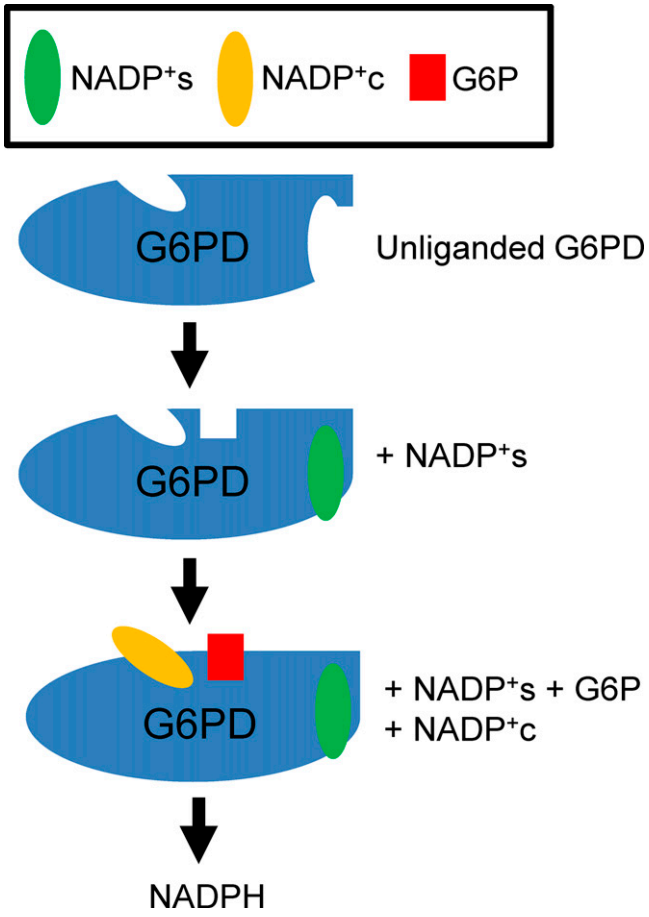

Fig. 5. Schematic for allosteric regulation of G6PD by structural NADP$^+$ (NADP$^+$s) G6PD first binds the higher affinity NADP$^+$s (green), which engages the G6PD C-terminal extension and allosterically promotes G6P (red) binding. The binding site for NADP$^+$c (yellow) remains accessible regardless of the binding of the other cofactors but likely binds after NADP$^+$s due to its lower relative affinity. The order of binding between NADP$^+$c and G6P is unknown.

The molecular role of G6PD multimerization has been previously investigated with the monomer form shown to be inactive (17, 18) and dimer and tetramer forms shown to be active, although tetramers that are prevented from dissociation due to disulfide bond crosslinking have been shown to be at least fourfold more active (19). Our cryo-EM studies reveal that G6PD dimers and tetramers have similar overall structures (Fig. 1*C*) and that the residues that interact with NADP$^+$s, NADP$^+$c, and G6P are fully contained within each subunit of the multimer (Fig. 1). These observations are most consistent with a role of G6PD multimerization on protein stability instead of multimerization playing a direct role on catalysis. An additional intriguing possibility is that multimerization may have a role in interprotein interaction, for example, with a recently proposed interaction with NAD$^+$ kinase (23).

Horikoshi et al. (16) recently reported on the crystal structures of several class I G6PD deficiency mutants (P396L, R393H, V394L, and F381L) bound to NADP$^+$c, with related biochemical and biophysical studies. Class I mutants are the most severe of the five classes of G6P deficiency mutants showing less than 10% activity relative to wild-type (22, 24). The authors focused on a subset of mutants located around the NADP$^+$s binding site and dimerization interface and found that these mutants resulted in a disordering of the C-terminal extension and dimerization interface involving strands M, N, and part of L of the NADP$^+$s binding β-sheet, the associated absence of a bound NADP$^+$s, and shifted helices f and f′. A molecular dynamics simulation of the unliganded protein and X-ray crystal structure of a W509A NADP$^+$s binding site also showed a disordered C-term extension and β-sheet dimerization interface. Together, these studies led the authors to a propose that class I mutants cause a destabilization of the NADP$^+$s binding site, which in turn leads to a disordering of the C-terminal extension and dimerization region of the β-sheet NADP$^+$s binding region and shifting of the f and f′ helices leading to a disruption of f′ helix G6P binding residues.

A superposition of our G6PD-WT, G6PD-D200N-NADP$^+$s-NADP$^+$c-G6P and G6PD-D200N-NADP$^+$s-NADP$^+$c structures, for comparison with the class I mutants reported by Horikoshi et al. (16), reveals that while the C-terminal end of strand M of the β-sheet and the following loop shows significant variability, the rest of the sheet and helix f superimpose well in our structures (*SI Appendix*, Fig. S7). This suggests that the lack of NADP$^+$s binding does not unfold the M and N strands of the β-sheet region or shift the f helix as suggested by Horikoshi et al. and implies that the more significant alterations observed by Horikoshi et al. are more directly nucleated by the class I G6P deficiency mutants. As the studies reported here and by Horikoshi et al. (16) both point to the importance of NADP$^+$s binding and the proper orientation of the f′ helix for combating G6PD deficiency mutants resulting in oxidative stress disorders including hemolytic anemias, the identification of small molecule compounds that might mimic the effects of NADP$^+$s binding and promote the active conformation of the f′ helix in particular might be an attractive therapeutic avenue to pursue.

Materials and Methods

Expression and Purification of Proteins. The recombinant G6PD protein expressed in the BL21(DE3) strain was fused with a hexahistidine tag at its C terminus. The *E. coli* cells were cultured in LB-kanamycin medium. When the cell density reached A_{600nm} = 0.6–0.8, protein expression was induced with the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside for 16–20 h at 16 °C. After expression, the cells were harvested by centrifugation at 4700g for 10 min, and the pellets were frozen and stored at −80 °C. Subsequently, the cell pellets were resuspended in a lysis buffer containing 25 mM Hepes, pH 7.5, 150 mM NaCl, and subjected to sonication. The cell lysate was centrifuged at 48,900g for 30 min and the supernatant was collected. The protein was purified through Cobalt Chelating Resin. The buffers used for column purification were as follows. Equilibration buffer contained 25 mM Hepes, pH 7.5, 150 mM NaCl. Washing buffer contained 25 mM Hepes, pH 7.5, 500 mM NaCl and 25 mM imidazole. Elution buffer contained 25 mM Hepes, pH 7.5, 150 mM NaCl and 200 mM imidazole. After elution from the cobalt resin, the protein sample was further purified through gel filtration chromatography using a Superdex 200 10/300 GL column (GE Healthcare) in a buffer containing 25 mM Hepes, pH 7.5, 150 mM NaCl. The protein was concentrated to 10–20 mg/mL and flash frozen at −20 °C until further use.

Enzymatic characterization. The Glucose-6-Phosphate Dehydrogenase activity colorimetric assay kit (Abcam) was used to analyze the activity of G6PD-WT and G6PD-D200N. Each reaction was performed in in a 96-well plate with proteins incubated with substrates at 30 °C for 5 min, the absorbance of each sample was measured at 600 nm (A_{600}) using a plate reader.

Cryo-EM data sample preparation and data collection. For G6PD-WT, G6PD protein was diluted to 0.8 mg/mL for cryo-EM grids preparation. For G6PD-D200N mutant, initially, the same concentration of the mutant protein was incubated with 3 mM NADP$^+$ and 3 mM G6P at room temperature for at least 15 min. Only NADP$^+$ cofactors were observed in the resultant structure after EM data analysis. The subsequent incubation of 13 μM G6PD-D200N with 3 mM NADP$^+$ and 100 mM G6P resulted in the observation of NADP$^+$ and G6P cofactors in EM map following EM data analysis. For cryo-EM data collection,

1 μL 0.1% (wt/vol) DDM was added to 20 μL protein sample and 3 μL of the protein/detergent mixture was applied to glow discharged Ultraufoil R1.2/1.3 300 mesh gold grids and blotted for 9 s (blot force = 5) under 100% humidity at 16 °C and plunged into liquid ethane using an FEI Vitribot Mark IV. These grids were screened using a FEI TF20 microscope equipped with a Falcon 3 detector. Data collection was performed using a Titan Krios equipped with a K3 direct detector (Gatan). A defocus range of 1.0–2.0 μm was applied in image acquisition. Details of the data collections are shown in Table 1.

Image Processing. All cryo-EM data were processed using the following workflow. First, beam-induced motions were corrected using MotionCor2. After motion correction, all micrographs were imported into Cryosparc2 and subsequent data processing was performed in Cryosparc2. Ctffind4 was used to determine the defocus value of each micrograph. Micrographs, which contained CTF information poorer than 5 Å, were discarded. A diameter of 160–200 Å was used for blob particle picking to pick ∼20,000 particles from ∼150 micrographs. 2D class averages were generated from ∼20,000 particles, and good classes were selected as a reference for template picking for all micrographs.

For 3D reconstruction of G6PD-WT, G6PD-D200N with 3 mM NADP and 3 mM G6P (G6PD-D200N-NADP$^+$s-NADP$^+$c), and G6PD-D200N with 3 mM NADP and 100 mM G6P (G6PD-D200N-NADP$^+$s-NADP$^+$c-G6P), we used 3,365,325 particles (2,945 micrographs), 1,904,781 particles (3,116 micrographs), and 1,450,216 particles (3,631 micrographs), respectively. All particles were classified into 200 classes. After 2D classification of G6PD-WT, G6PD-D200N-NADP$^+$s-NADP$^+$c, and G6PD-D200N-NADP$^+$s-NADP$^+$c-G6P, 570,405, 866,530, and 385,501 particles were kept, respectively. Eight initial models were generated from all particles followed by multiple rounds of heterogeneous refinement. Refinement of G6PD-WT resulted in one tetramer class with D2 symmetry, one dimer class with C2 symmetry, and some bad classes. A map of the tetramer class with an overall resolution of 3.1 Å was generated from 139,200 G6PD-WT particles with no symmetry applied, and the resolution was improved to 2.8 Å with D2 symmetry applied. A map of the dimer class with an overall resolution of 3.7 Å was generated from 96,567 G6PD-WT particles with no symmetry imposed, and the resolution was improved to 3.5 Å with C2 symmetry applied.

Refinement of G6PD-D200N-NADP$^+$s-NADP$^+$c resulted in one tetramer class and some bad classes. A map of the G6PD-D200N-NADP$^+$s-NADP$^+$c tetramer class with an overall resolution of 2.5 Å was generated from 243,006 G6PD-WT particles with no symmetry applied, and the resolution was improved to 2.2 Å with D2 symmetry applied. Refinement of G6PD-D200N-NADP$^+$s-NADP$^+$c-G6P resulted in one tetramer class, one bad class with severe preferred orientation and some bad classes. A map of the G6PD-D200N-NADP$^+$s-NADP$^+$c-G6P tetramer class with an overall resolution of 2.9 Å was generated from 67,823 G6PD-WT tetramer particles with no symmetry applied, and the resolution was improved to 2.5 Å with D2 symmetry applied.

All representations of cryo-EM density and structural models were prepared with Chimera (25) and PyMol (26), respectively.

Data Availability. Structural coordinates and electron microscopy map data have been deposited in the Protein Data Bank and Electron Microscopy Data Bank: G6PD-WT-tetramer (EMDB: 25225 (27); PDB: 7SNG (28)); G6PD-WT-dimer (EMDB: 25224 (29); PDB: 7SNF (30)); G6PD-D200N-NADP$^+$ (EMDB: 25226 (31); PDB: 7SNH (32)); and G6PD-D200N-NADP$^+$-G6P (EMDB: 25227 (33); PDB: 7SNI (34)).

ACKNOWLEDGMENTS. This work was supported by NIH (R35GM118090 and P01AG031862 to R.M. and R01CA235760 and R01CA243520 to X.Y.). We thank Stefan Steimle from the Beckman Center for Cryo-EM at the University of Pennsylvania for technical assistance on data collection.

Author affiliations: [a]Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA, 19104; [b]Abramson Family Cancer Research Institute, University of Pennsylvania, Philadelphia, PA, 19104; [c]Graduate Group in Biochemistry and Molecular Biophysics, University of Pennsylvania, Philadelphia, PA, 19104; and [d]Department of Cancer Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, 19104

1. M. S. Cosgrove, C. Naylor, S. Paludan, M. J. Adams, H. R. Levy, On the mechanism of the reaction catalyzed by glucose 6-phosphate dehydrogenase. *Biochemistry* **37**, 2759–2767 (1998).
2. R. C. Stanton, Glucose-6-phosphate dehydrogenase, NADPH, and cell survival. *IUBMB Life* **64**, 362–369 (2012).
3. H. Y. Ho, M. L. Cheng, D. T. Chiu, Glucose-6-phosphate dehydrogenase–Beyond the realm of red cell biology. *Free Radic. Res.* **48**, 1028–1048 (2014).
4. J. Lemire, A. Alhasawi, V. P. Appanna, S. Tharmalingam, V. D. Appanna, Metabolic defence against oxidative stress: The road less travelled so far. *J. Appl. Microbiol.* **123**, 798–809 (2017).
5. E. Beutler, Study of glucose-6-phosphate dehydrogenase: History and molecular biology. *Am. J. Hematol.* **42**, 53–58 (1993).
6. L. Luzzatto, C. Nannelli, R. Notaro, Glucose-6-phosphate dehydrogenase deficiency. *Hematol. Oncol. Clin. North Am.* **30**, 373–393 (2016).
7. M. D. Cappellini, G. Florlli, Glucose-6-phosphate dehydrogenase deficiency. Lancet 371, 64-74 (2008).
8. K. Ryan, B. L. Tekwani, Current investigations on clinical pharmacology and therapeutics of glucose-6-phosphate dehydrogenase deficiency. *Pharmacol. Ther.* **222**, 107788 (2021).
9. C. Ruwende, A. Hill, Glucose-6-phosphate dehydrogenase deficiency and malaria. *J. Mol. Med. (Berl.)* **76**, 581–588 (1998).
10. G. Manganelli, U. Masullo, S. Passarelli, S. Filosa, Glucose-6-phosphate dehydrogenase deficiency: Disadvantages and possible benefits. *Cardiovasc. Hematol. Disord. Drug Targets* **13**, 73–82 (2013).
11. G. Bancone, C. S. Chu, G6PD variants and haemolytic sensitivity to primaquine and other drugs. *Front. Pharmacol.* **12**, 638885 (2021).
12. K. C. Patra, N. Hay, The pentose phosphate pathway and cancer. *Trends Biochem. Sci.* **39**, 347–354 (2014).
13. A. Stincone *et al.*, The return of metabolism: Biochemistry and physiology of the pentose phosphate pathway. *Biol. Rev. Camb. Philos. Soc.* **90**, 927–963 (2015).
14. 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**, 293–303 (2000).
15. M. Kotaka *et al.*, Structural studies of glucose-6-phosphate and NADP+ binding to human glucose-6-phosphate dehydrogenase. *Acta Crystallogr. D Biol. Crystallogr.* **61**, 495–504 (2005).
16. N. Horikoshi *et al.*, Long-range structural defects by pathogenic mutations in most severe glucose-6-phosphate dehydrogenase deficiency. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2022790118 (2021).
17. 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**, 493–501 (1971).
18. 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**, 142–148 (1970).
19. A. T. Ranzani, A. T. Cordeiro, Mutations in the tetramer interface of human glucose-6-phosphate dehydrogenase reveals kinetic differences between oligomeric states. *FEBS Lett.* **591**, 1278–1284 (2017).
20. X. T. Wang, T. F. Chan, V. M. Lam, P. C. Engel, What is the role of the second "structural" NADP+-binding site in human glucose 6-phosphate dehydrogenase? *Protein Sci.* **17**, 1403–1411 (2008).
21. S. Gómez-Manzo *et al.*, Glucose-6-phosphate dehydrogenase: Update and analysis of new mutations around the world. *Int. J. Mol. Sci.* **17**, 2069 (2016).
22. A. Minucci *et al.*, Glucose-6-phosphate dehydrogenase (G6PD) mutations database: Review of the "old" and update of the new mutations. *Blood Cells Mol. Dis.* **48**, 154–165 (2012).
23. Y. Zhang *et al.*, Upregulation of antioxidant capacity and nucleotide precursor availability suffices for oncogenic transformation. *Cell Metab.* **33**, 94–109.e8 (2021).
24. Anonymous; WHO Working Group, Glucose-6-phosphate dehydrogenase deficiency. *Bull. World Health Organ.* **67**, 601–611 (1989).
25. E. F. Pettersen *et al.*, UCSF Chimera–A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
26. W. L. DeLano, *The PyMOL Molecular Graphics System* (DeLano Scientific, San Carlos, CA, 2002).
27. X. Wei, K. Kixmoeller, E. Baltrusaitis, X. Yang, R. Marmorstein, Allosteric role of a structural NADP$^+$ molecule in glucose-6-phosphate dehydrogenase activity. EMDB: 25225, https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-25225. Deposited 28 October 2021.
28. X. Wei, K. Kixmoeller, E. Baltrusaitis, X. Yang, R. Marmorstein, Allosteric role of a structural NADP$^+$ molecule in glucose-6-phosphate dehydrogenase activity. PDB: 7SNG, https://www.rcsb.org/structure/7SNG. Deposited 28 October 2021.
29. X. Wei, K. Kixmoeller, E. Baltrusaitis, X. Yang, R. Marmorstein, Allosteric role of a structural NADP$^+$ molecule in glucose-6-phosphate dehydrogenase activity. EMDB: 25224, https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-25224. Deposited 28 October 2021.
30. X. Wei, K. Kixmoeller, E. Baltrusaitis, X. Yang, R. Marmorstein, Allosteric role of a structural NADP$^+$ molecule in glucose-6-phosphate dehydrogenase activity. PDB: 7SNF, https://www.rcsb.org/structure/7SNF. Deposited 28 October 2021.
31. X. Wei, K. Kixmoeller, E. Baltrusaitis, X. Yang, R. Marmorstein, Allosteric role of a structural NADP$^+$ molecule in glucose-6-phosphate dehydrogenase activity. EMDB: 25226, https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-25226. Deposited 28 October 2021.
32. X. Wei, K. Kixmoeller, E. Baltrusaitis, X. Yang, R. Marmorstein, Allosteric role of a structural NADP$^+$ molecule in glucose-6-phosphate dehydrogenase activity. PDB: 7SNH, https://www.rcsb.org/structure/7SNH. Deposited 28 October 2021.
33. X. Wei, K. Kixmoeller, E. Baltrusaitis, X. Yang, R. Marmorstein, Allosteric role of a structural NADP$^+$ molecule in glucose-6-phosphate dehydrogenase activity. EMDB: 25227, https://www.ebi.ac.uk/pdbe/entry/emdb/EMD-25227. Deposited 28 October 2021.
34. X. Wei, K. Kixmoeller, E. Baltrusaitis, X. Yang, R. Marmorstein, Allosteric role of a structural NADP$^+$ molecule in glucose-6-phosphate dehydrogenase activity. PDB: 7SNI, https://www.rcsb.org/structure/7SNI. Deposited 28 October 2021.