PNAS
www.pnas.org

Supplementary Information for

Long-range structural defects by pathogenic mutations in most severe glucose-6-phosphate dehydrogenase deficiency

Naoki Horikoshi[a,b,c], Sunhee Hwang[d], Cornelius Gati[b,c], Tsutomu Matsui[e], Carlos Castillo-Orellana[f], Andrew G. Raub[d], Adriana A. Garcia[d], Fatemeh Jabbarpour[b,c], Alexander Batyuk[b], Joshua Broweleit[c], Xinyu Xiang[c], Andrew Chiang[c], Rachel Broweleit[c], Esteban Vöhringer-Martinez[f], Daria Mochly-Rosen[d,1], Soichi Wakatsuki[b,c,1]

[a]Life Science Center for Survival Dynamics, University of Tsukuba, Tsukuba, Ibaraki, Japan. [b]Biological Sciences Division, SLAC National Accelerator Laboratory, Menlo Park, CA, USA. [c]Department of Structural Biology, Stanford University School of Medicine, Stanford, CA, USA. [d]Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA, USA. [e]Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Menlo Park, CA, USA. [f]Departamento de Físico-Química, Facultad de Ciencias Químicas, Universidad de Concepción, Concepción, Chile [1]Corresponding authors

Corresponding authors: Soichi Wakatsuki and Daria Mochly-Rosen
Email: soichi.wakatsuki@stanford.edu; mochly@stanford.edu

This PDF file includes:

Other supplementary materials for this manuscript include the following:

Supplementary Methods

Crystallization of G6PD^{WT} in the condition in which the Class I mutants were crystallized.

To investigate if the structural changes observed in the Class I mutants are possibly due to the crystallization condition or crystal packing, G6PD^{WT} was crystallized in the crystallization condition in which G6PD^{P396L}, G6PD^{R393H}, and G6PD^{W509A} were crystallized (which will be referred to as G6PD^{WT}* crystals). The wild-type crystals obtained in this way were significantly different from those of the G6PD mutants although only small crystals whose sizes were approximately 5x10x5 μm^3 were obtained, possibly because the crystallization condition was not suited for G6PD^{WT}. The crystal shape of the G6PD mutants was bipyramid; on the other hand, that of the G6PD^{WT}* crystals obtained using the mutant-crystallization condition were oval. Besides, the space group of the crystals of the G6PD mutants was $P4_12_12$, whereas that for the new G6PD^{WT}* was $P6_422$. Although we were only able to collect a dataset up to 6.5Å from the G6PD^{WT}*, as a result of the molecular replacement using the G6PD^{WT} or G6PD^{P396L} structure as a search model, the LLG and TFZ scores by Phaser-MR in the PHENIX suite (1) were 155.9 and 16.3, in case of the G6PD^{WT} as a search model, and 17.4 and 5.9, in case of G6PD^{P396L}. After the initial refinement, R_{free} score was 0.379 for G6PD^{WT} and 0.598 for G6PD^{P396L}, indicating that the G6PD^{WT}* structure is similar to that of the G6PD^{WT} and does not resemble the G6PD^{P396L} structure. Therefore, we conclude that the crystallization condition does not affect the structures of the Class I mutants of G6PD.

Kinetics of catalytic activity of G6PD mutants in the presence of G6P and NADP$^+$.

The k_{cat} and V_{max} values of G6PD^{P396L}, G6PD^{W509A}, G6PD^{R393H}, G6PD^{V394L}, and G6PD^{F381L} were significantly lower than those of G6PD^{WT} (SI Appendix, Table S4). Values of G6PD^{P396L} and G6PD^{V394L} were too low to estimate. For G6PD^{W509A}, G6PD^{R393H}, and G6PD^{F381L}, the values were much lower than those of G6PD^{WT} by factors of 3, 12, and 30, respectively. The K_M (G6P) value of G6PD^{W509A} was slightly higher than that of G6PD^{WT}. Although the K_M (G6P) values of G6PD^{F381L} and G6PD^{R393H} were lower than that of G6PD^{WT}, it does not seem to be simply comparable because the k_{cat} and V_{max} values are significantly different. Since the V_{max} values of G6PD^{F381L} and G6PD^{R393H} were quite low, lower G6P concentrations might be enough to saturate the G6P-binding sites of G6PD^{F381L} and G6PD^{R393H}, compared to G6PD^{WT}.

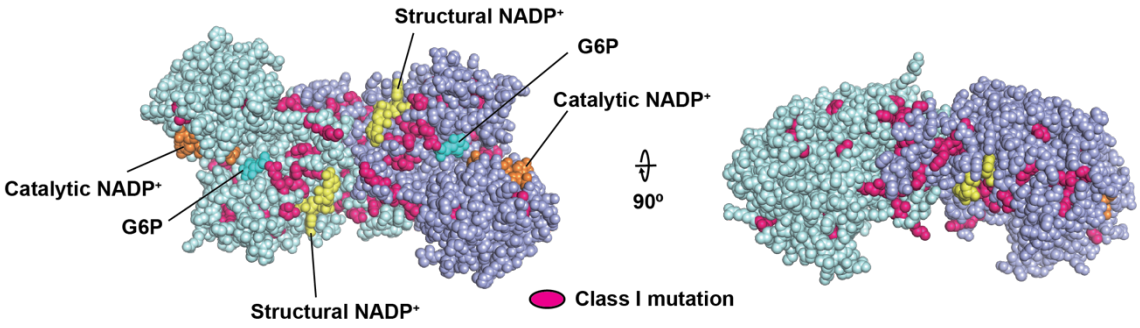
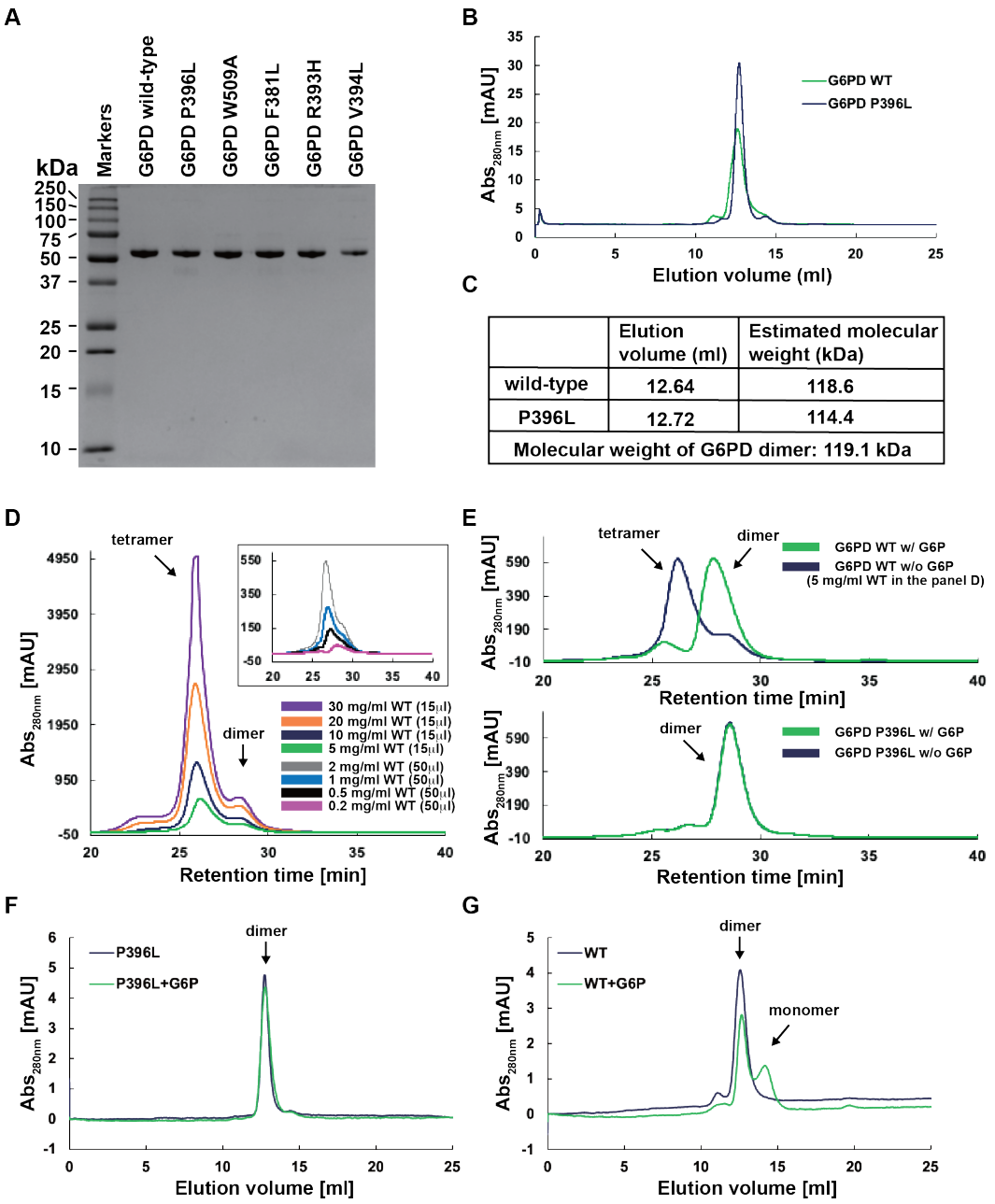


Fig. S1. Class I clinical mutations in the G6PD dimer structure. The model structure of the G6PD dimer with $NADP^+$ and G6P was made using two G6PD structures containing $NADP^+$ molecules or G6P molecules (PDB IDs: 2BH9 and 2BHL). Class I pathogenic mutations are shown in pink.



	Elution volume (ml)	Estimated molecular weight (kDa)
wild-type	12.64	118.6
P396L	12.72	114.4
Molecular weight of G6PD dimer: 119.1 kDa		

Fig. S2. Oligomerization state of G6PD. (*A*) Purified G6PD^{WT}, G6PD^{P396L}, G6PD^{W509A}, G6PD^{F381L}, G6PD^{R393H}, G6PD^{V394L} are analyzed by 15% SDS-PAGE with coomassie brilliant blue staining. (*B*) Purified dimer G6PD^{WT} and G6PD^{P396L} were analyzed by Superdex 200 Increase 10/300 GL. (*C*) Molecular weights of G6PD^{WT} and G6PD^{P396L} were estimated based on the gel filtration standard (Bio-Rad). (*D*) The SEC profile of G6PD^{WT} at different concentrations (0.2~30 mg/ml). The inset shows the SEC profiles of 0.2~2 mg/ml G6PD^{WT}. (*E*) The SEC profile of 5 mg/ml G6PD^{WT} and G6PD^{P396L} with or without G6P. The SEC profile of G6PD^{WT} without G6P is identical to that of 5 mg/ml G6PD^{WT} in the panel *D*. In panels *D* and *E*, the samples were analyzed with Superdex 200 Increase PC 3.2/300. (*F* and *G*) The SEC analyses of G6PD^{P396L} and G6PD^{WT} with or without G6P using Superdex 200 10/300 GL.

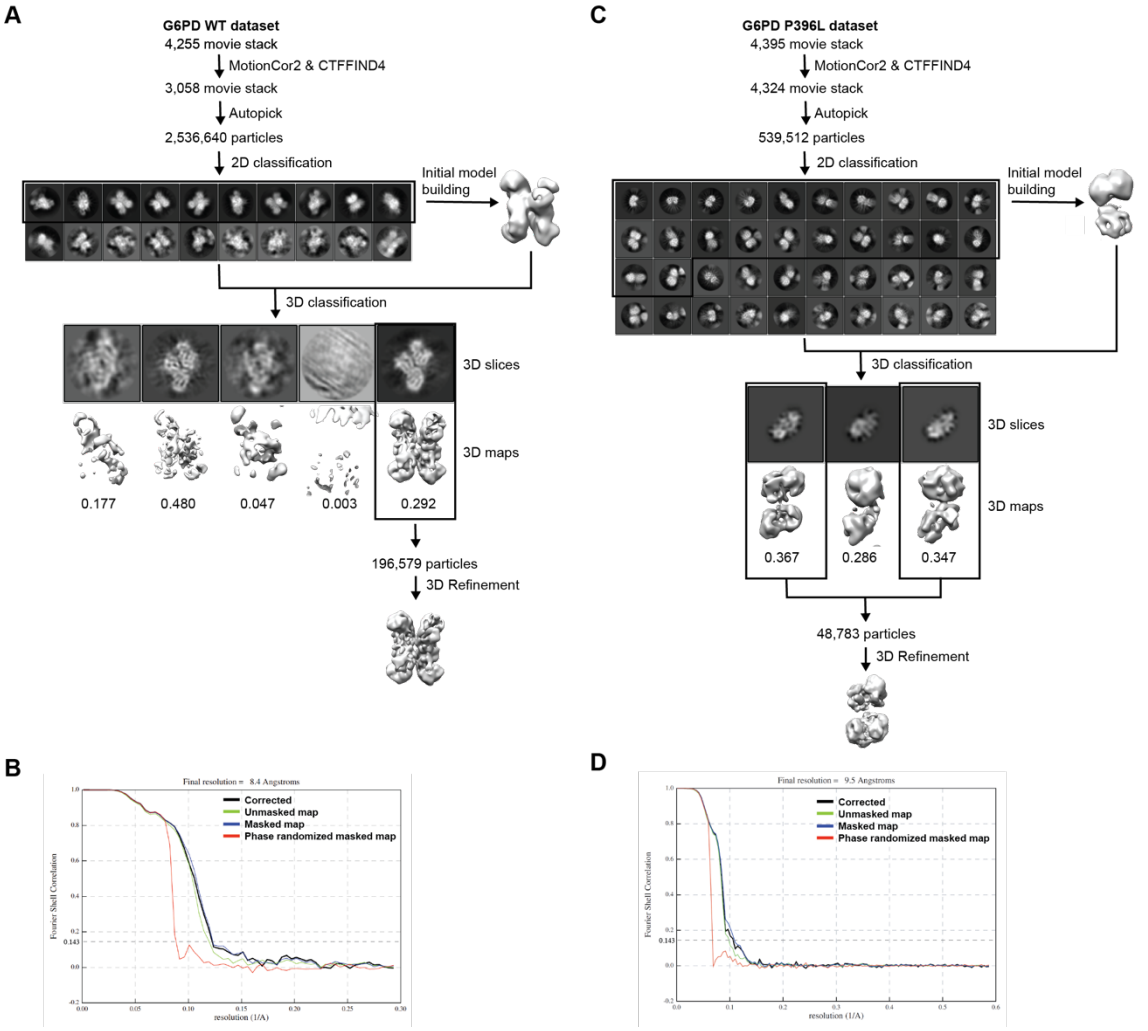


Fig. S3. Data processing of the Cryo-EM datasets of the G6PD^{WT} and G6PD^{P396L}. (*A*) The scheme of the data processing of G6PD^{WT}. (*B*) The Fourier shell correlation (FSC) plots of G6PD^{WT}. (*C*) The scheme of the data processing of G6PD^{P396L}. (*D*) The FSC plots of G6PD^{P396L}.

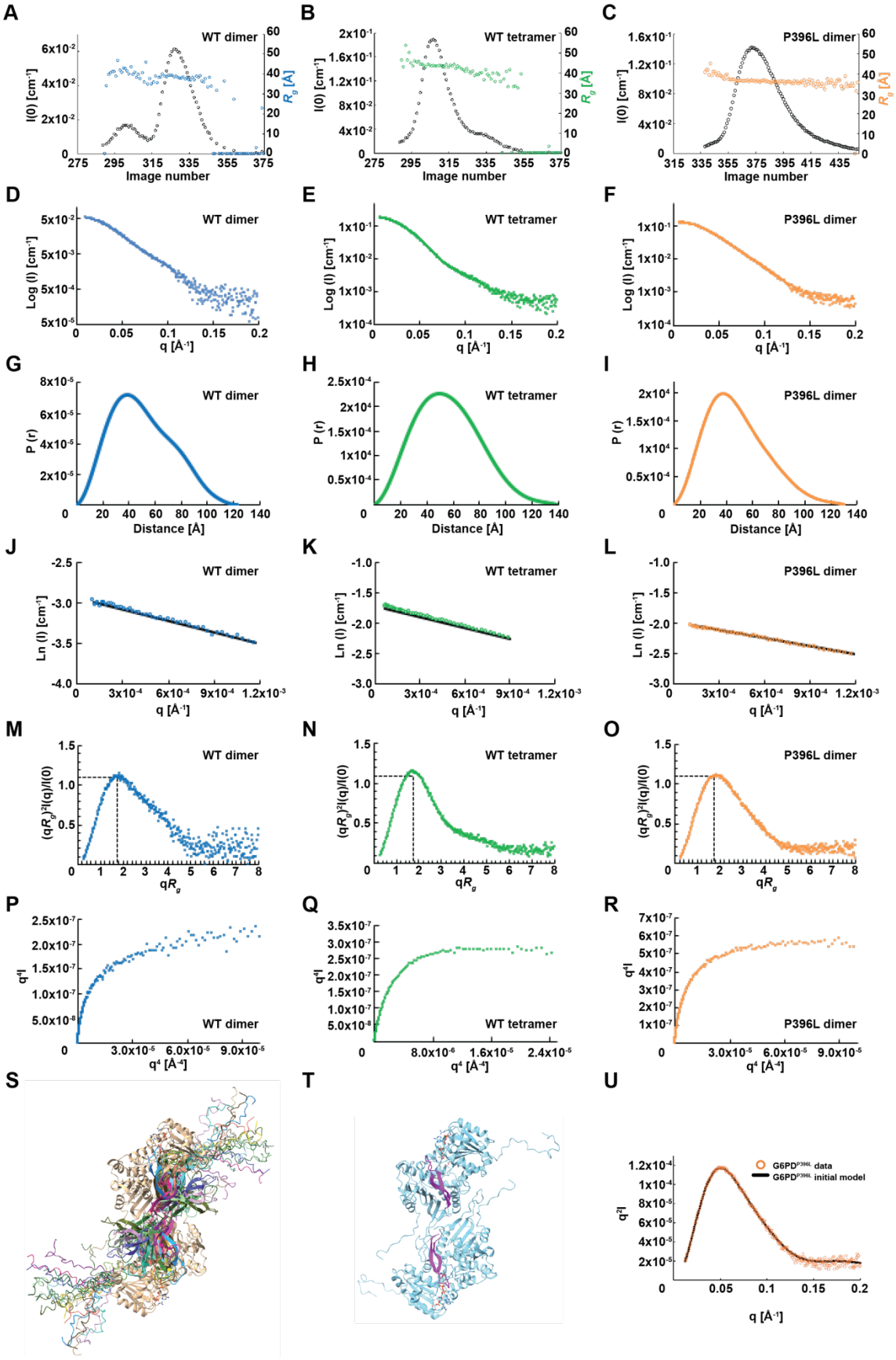

A
WT dimer
I(0) [cm⁻¹]
R_g [Å]
Image number
B
WT tetramer
I(0) [cm⁻¹]
R_g [Å]
Image number
C
P396L dimer
I(0) [cm⁻¹]
R_g [Å]
Image number
D
WT dimer
Log (I) [cm⁻¹]
q [Å⁻¹]
E
WT tetramer
Log (I) [cm⁻¹]
q [Å⁻¹]
F
P396L dimer
Log (I) [cm⁻¹]
q [Å⁻¹]
G
WT dimer
P (r)
Distance [Å]
H
WT tetramer
P (r)
Distance [Å]
I
P396L dimer
P (r)
Distance [Å]
J
WT dimer
Ln (I) [cm⁻¹]
q [Å⁻¹]
K
WT tetramer
Ln (I) [cm⁻¹]
q [Å⁻¹]
L
P396L dimer
Ln (I) [cm⁻¹]
q [Å⁻¹]
M
WT dimer
$(qR_g)^3I(q)/I(0)$
qR_g
N
WT tetramer
$(qR_g)^3I(q)/I(0)$
qR_g
O
P396L dimer
$(qR_g)^3I(q)/I(0)$
qR_g
P
WT dimer
q^4I
q^4 [Å^{-4}]
Q
WT tetramer
q^4I
q^4 [Å^{-4}]
R
P396L dimer
q^4I
q^4 [Å^{-4}]
S
T
U
G6PD^{P396L} data
G6PD^{P396L} initial model
q^2I
q [Å⁻¹]

Fig. S4. SEC-SAXS analysis and *CORAL* modeling. (*A*-*C*) The SEC-SAXS profiles. $I(0)$ (black open circle) and R_g (blue, green, orange circle) are plotted along with the SAXS image number. (*D*-*F*) $I(q)$ versus q as log-linear plots. (*G*-*I*) $P(r)$ functions. (*J*-*L*) *Guinier* plots. Open circles and black lines indicate experimental data and *Guinier* fit, respectively. (*M*-*O*) Dimensionless Kratky plots. The dashed lines indicate $(qR_g)^2I(q)/I(0) = 1.1$ and $qR_g = 1.73$, respectively. (*P*-*R*) *Porod-Debye* plots. (*S*) 20 runs of *CORAL* modeling based on the experimental SAXS curve of $G6PD^{P396L}$ were performed. All 20 structures are superimposed. (*T*) A representative structure of the *CORAL* modeling. The models of the two β-strands disordered in the crystal structure are shown in magenta. (*U*) The theoretical Kratky plot of the initial model of $G6PD^{P396L}$ is fitted with the experimental data.

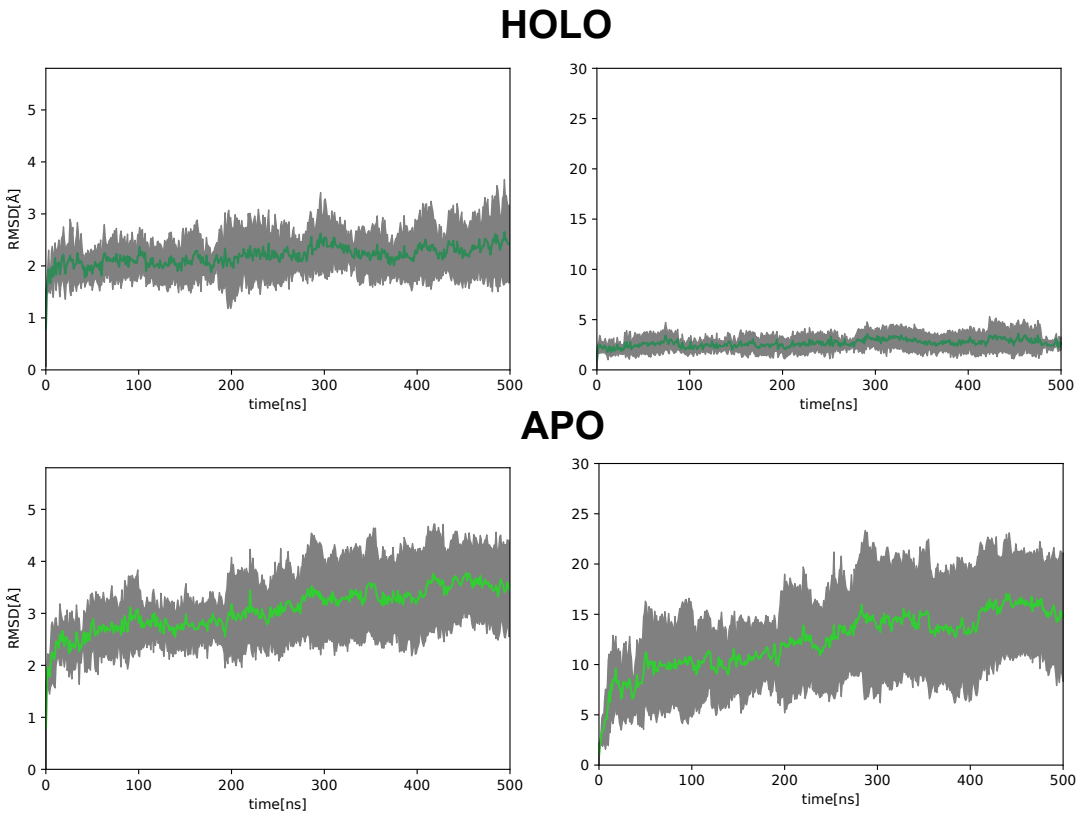


Fig. S5. Left: Average root mean square deviation (RMSD) of the Cα atom positions in the wild type G6PD dimer with (top) and without (bottom) the structural $NADP^+$ calculated from six independent MD simulations. The grey shadow shows the standard deviation of the mean of six independent trajectories. **Right:** RMSD of the Cα atom positions only in the C-terminal tail (residues 501-513). The comparison between the holo (top) and apo (bottom) form clearly shows the increased flexibility and the disordered state of the C-terminal tail of the apo form due to the absence of the structural $NADP^+$. This disordered state of the C-terminal tail is responsible for the increasing total RMSD of the whole apo dimer at the left side. For the RMSD calculation of the C-terminal tail, we first removed the translational and rotational motion of all the other atoms not contained in the C-terminal tail.

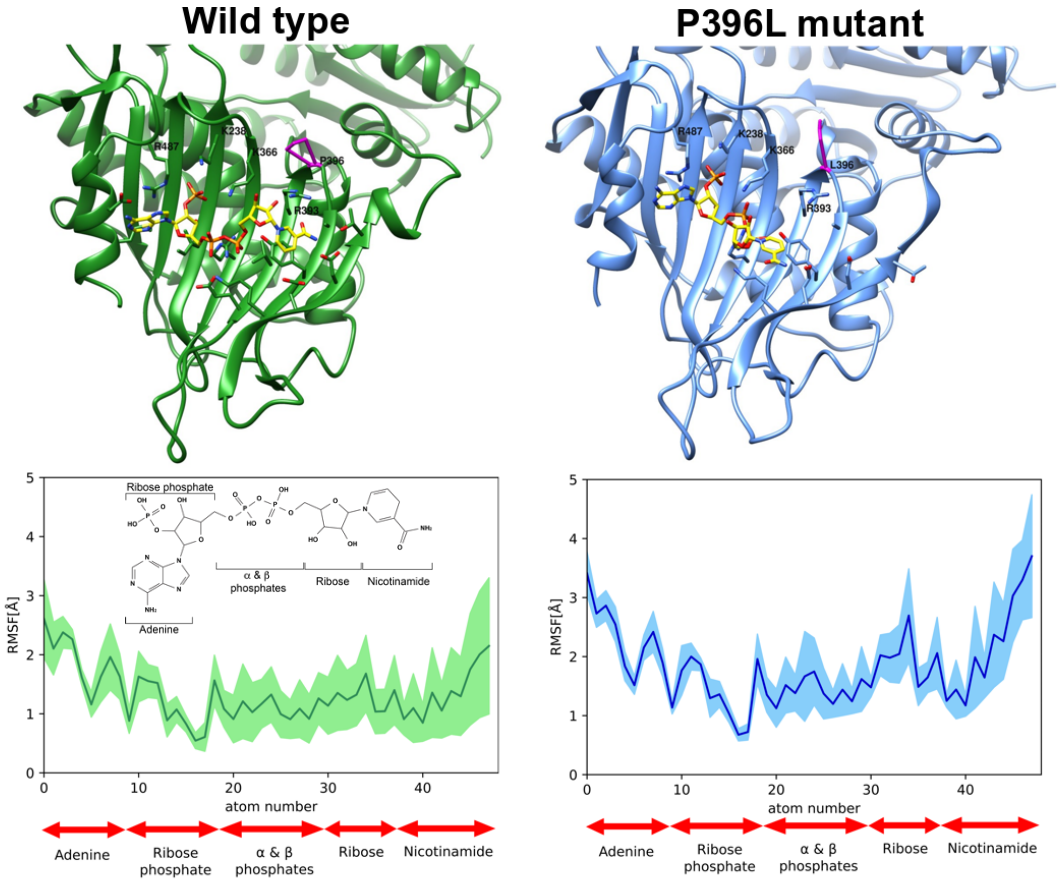


Fig. S6. Top: Representative snapshots from the MD simulations of the *in-silico* created C-terminal tail truncated dimer of the wild-type (green) and P396L mutant (blue) (residues 501-513 were deleted from the wild type X-ray structure before simulation). Bottom: Average root mean square fluctuations (RMSF) of $NADP^{+1}$'s non-hydrogen atoms taken from six independent MD simulations. Atom numbers were assigned to characteristic functional groups in $NADP^{+}$. Atoms involving phosphate in the wild type and the P396L mutant present the lowest fluctuations because of the protein's strong electrostatic interactions (see interactions with K366 and K238 in the representative snapshots at the top). The nicotinamide group and the ribose ring of $NADP^{+}$ display almost twice larger RMSF for the P396L mutant compared to the wild-type. This considerable fluctuation difference between the two is due to the mutant induced partial unzipping of the two β-strands and loss of the interaction to R393 (see SI Appendix, movie S4 and representative snapshots at the top). Shadows represent one standard deviation of the mean RMSF calculated from six independent simulations.

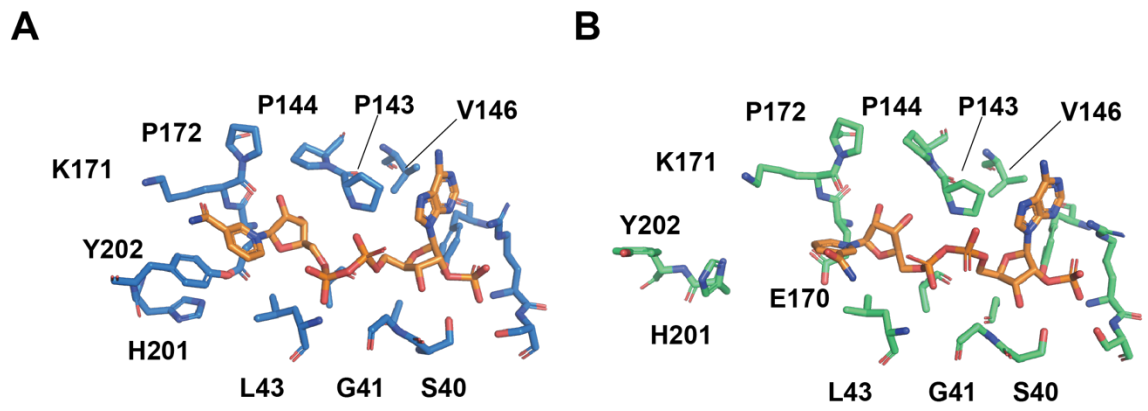


Fig. S7. Catalytic NADP$^+$ in the G6PD^{P396L} and G6PD^{WT} structures. (*A*) The catalytic NADP$^+$ binding site in the G6PD^{P396L} structure. Residues of G6PD^{P396L} around the catalytic NADP$^+$ are shown in blue. (*B*) The catalytic NADP$^+$ binding site in G6PD^{WT} structure. Residues of G6PD^{WT} around the catalytic NADP$^+$ are shown in green. (*A* and *B*) The catalytic NADP$^+$ are shown in orange.

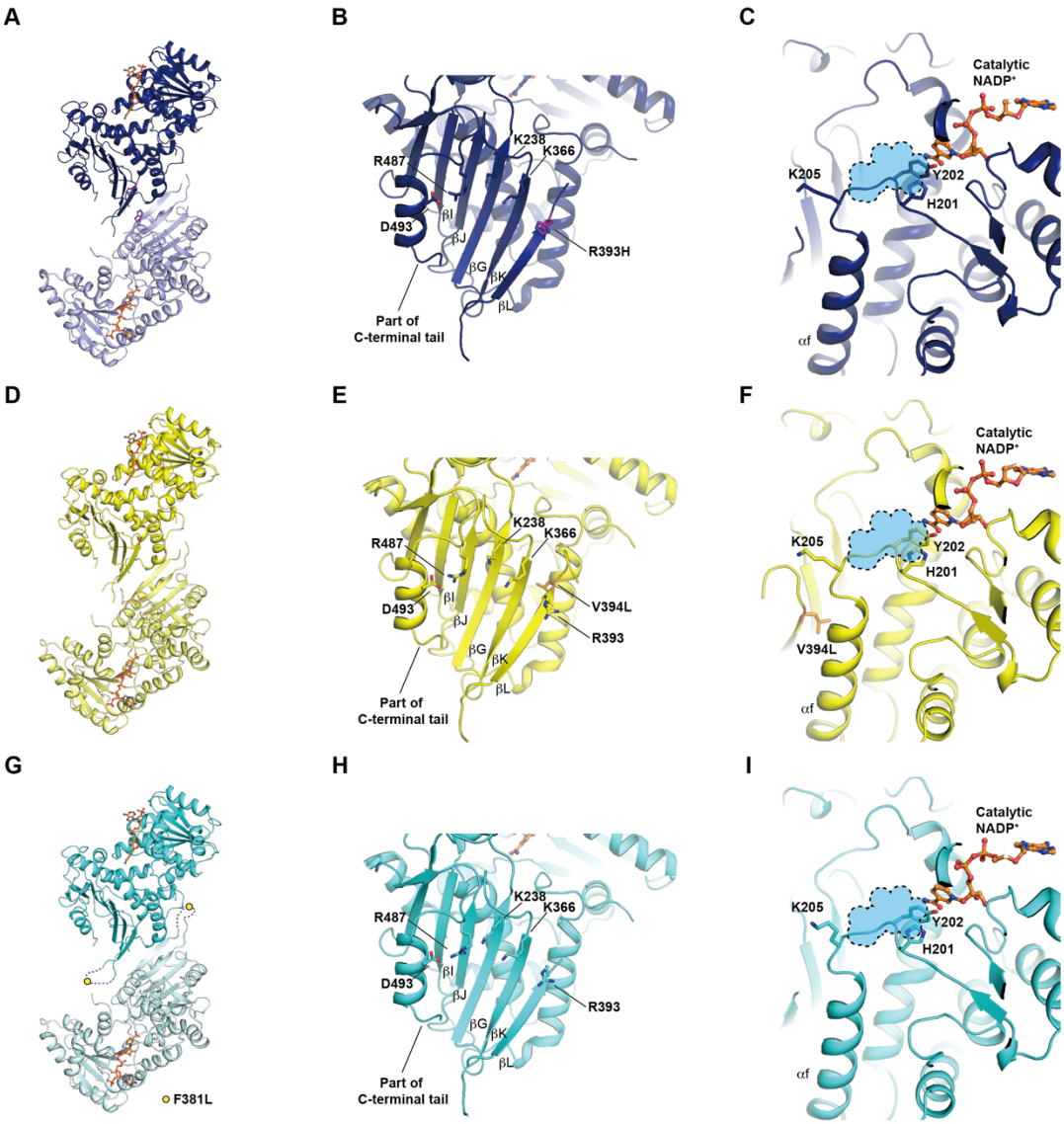


Fig. S8. Structures of G6PD^{R393H}, G6PD^{V394L}, and G6PD^{F381L}. (*A*) Overall structure of G6PD^{R393H}. Two G6PD molecules are shown, one in dark and the other in light blue. (*B*) Close-up view of the structural NADP$^+$ binding site in G6PD^{R393H}. The R393H residue is shown in purple. (*C*) Close-up view of the G6P binding site in G6PD^{R393H}. (*D*) Overall structure of G6PD^{V394L}. Two G6PD molecules are shown in yellow and light yellow. (*E*) Close-up view of the structural NADP$^+$ binding site in G6PD^{V394L}. The V394L residue is shown in light orange. (*F*) Close-up view of the G6P binding site in G6PD^{V394L}. (*G*) Overall structure of G6PD^{F381L}. G6PD molecules are shown in cyan and pale cyan, respectively. The purple dotted line and yellow circle indicate a disordered loop and the location of the F381L residue, respectively. (*H*) Close-up view of the structural NADP$^+$ binding site in G6PD^{F381L}. (*I*) Close-up view of the G6P binding site in G6PD^{F381L}. (*C*, *F*, and *I*) The G6P binding site is shown in sky blue. The catalytic NADP$^+$ is shown in orange.

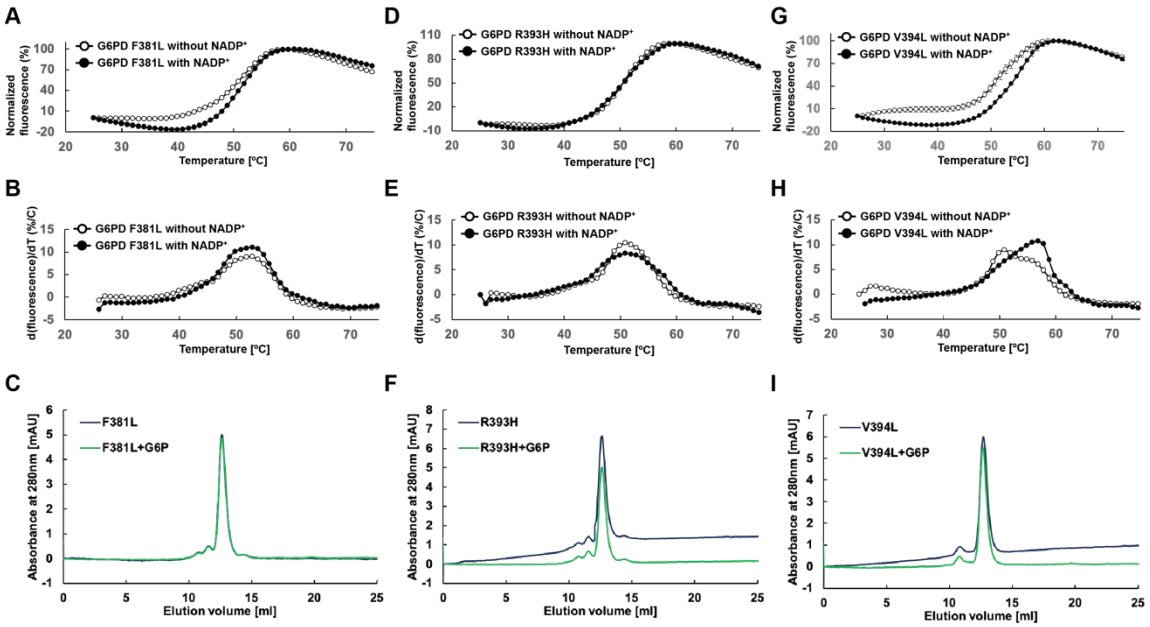


Fig. S9. Biophysical properties of G6PD^{F381L}, G6PD^{R393H}, and G6PD^{V394L}. (*A*, *D*, and *G*) The thermal denaturation curves of the G6PD^{F381L}, G6PD^{R393H}, and G6PD^{V394L} with or without NADP^{+}. (*B*, *E*, and *H*) The derivative values of the thermal denaturation curves in panels A, D, G, respectively. The error bars indicate the standard deviations (n=4). (*C*, *F*, and *I*) The SEC analysis of G6PD^{F381L}, G6PD^{R393H}, and G6PD^{V394L} with or without G6P. 0.5 mg/ml G6PD mutants with or without 1 mM G6P are analyzed using Superdex 200 Increase 10/300 GL.

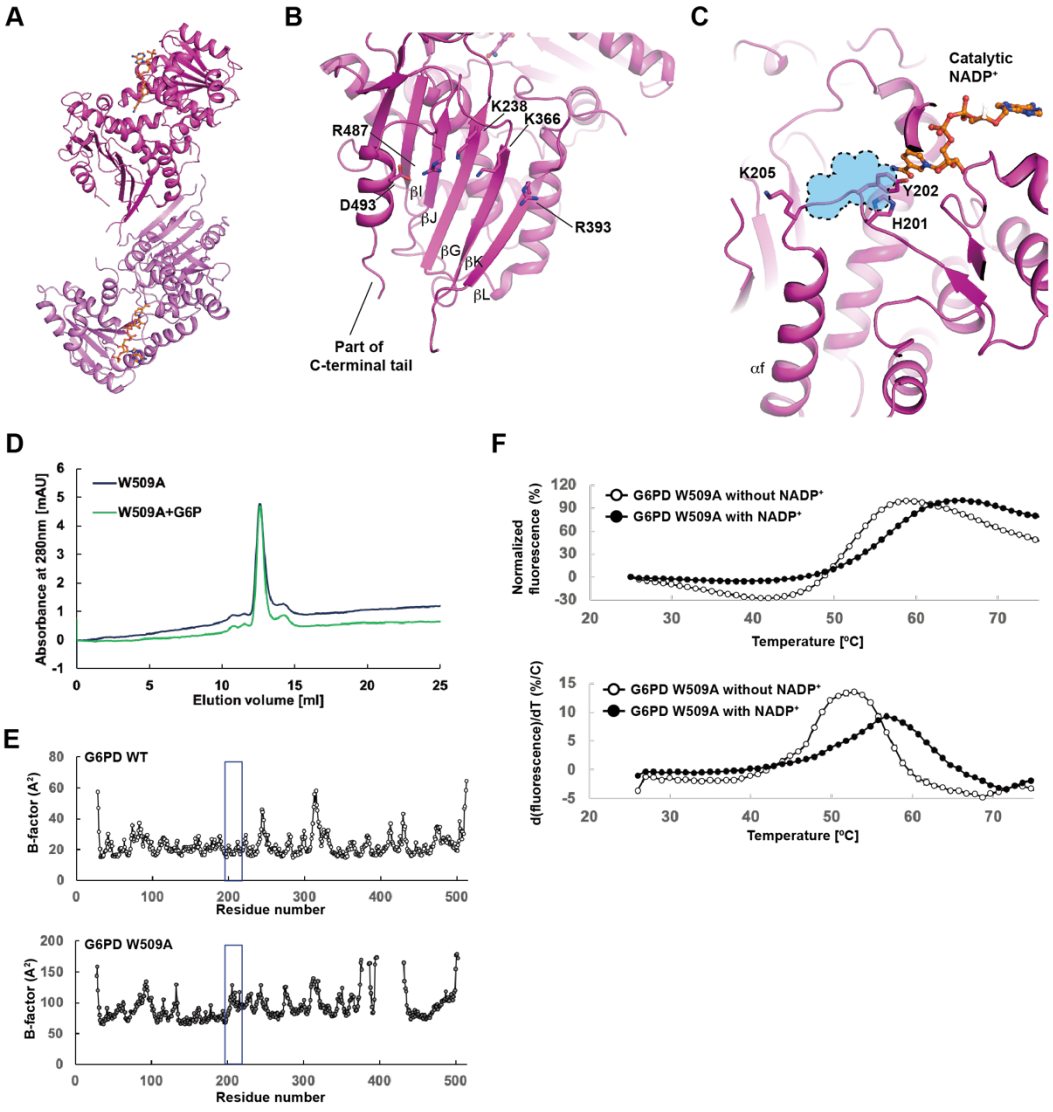


Fig. S10. Structure and biophysical property of the G6PD W509A mutant. (*A*) Overall structure of G6PD^{W509A}. G6PD molecules are shown in bright magenta and light purple, respectively. (*B*) Close-up view of the structural NADP$^+$ binding site in G6PD^{W509A}. (*C*) Close-up view of the G6P binding site in G6PD^{W509A}. The G6P binding site is shown in sky blue. The catalytic NADP$^+$ is shown in orange. (*D*) The SEC analysis of G6PD^{W509A} with or without G6P. (*E*) B-factors of G6PD^{WT} (PDB ID: 6E08) and G6PD^{W509A} are plotted. Boxed regions indicate αf and αf′ helices. (*F*) The thermal denaturation curves of the G6PD^{W509A} with or without NADP$^+$. The bottom panel shows the derivative values of the thermal denaturation curves. The error bars indicate the standard deviations (n=4).

Table S1. Data collection and refinement statistics

Supplementary Table S1. Data collection and refinement statistics

	G6PD^{P396L}	G6PD^{W509A}	G6PD^{R393H}	G6PD^{V394L}	G6PD^{F381L}
Data collection					
Space group	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P4_12_12$
Cell dimensions					
a, *b*, *c* (Å)	157.35, 157.35, 113.62	157.36, 157.36, 113.57	157.44, 157.44, 113.02	158.97, 158.97, 113.19	158.62, 158.62, 113.91
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Resolution (Å)	50-3.07 (3.26-3.07) *	50-3.10 (3.29-3.10)	50-3.95 (4.19-3.95)	50-2.95 (3.13-2.95)	50-3.95 (4.19-3.95)
R_{sym} or R_{merge}	0.114 (1.596)	0.143 (1.772)	0.365 (1.944)	0.108 (1.102)	0.289 (1.591)
$I / \sigma I$	21.94 (1.61)	34.75 (2.59)	7.39 (1.59)	19.29 (2.11)	8.31 (1.55)
Completeness (%)	99.8 (99.3)	97.2 (99.8)	99.8 (99.5)	99.7 (98.4)	99.6 (99.2)
Redundancy	13.28 (13.64)	42.19 (44.62)	12.86 (13.42)	13.16 (13.59)	10.16 (10.81)
$CC_{1/2}$	0.999 (0.647)	1.000 (0.850)	0.995 (0.702)	0.999 (0.814)	0.999 (0.814)
Refinement					
Resolution (Å)	49.8-3.07 (3.18-3.07)	47.6-3.10 (3.23-3.10)	47.6-3.95 (4.25-3.95)	39.7-2.95 (3.05-2.95)	48.0-3.95 (4.26-3.95)
No. reflections	27084	25682	12917	30842	13156
R_{work} / R_{free}	0.203/0.225	0.232/0.240	0.191/0.220	0.209/0.222	0.206/0.224
No. atoms					
Protein	3493	3505	3478	3485	3447
Ligand/ion	48	48	48	48	48
B-factors					
Protein	94.1	94.7	132.3	80.7	122.8
Ligand/ion	81.2	80.0	122.9	68.5	118.9
R.m.s. deviations					
Bond lengths (Å)	0.004	0.004	0.003	0.004	0.003
Bond angles (°)	0.767	0.802	0.617	0.749	0.623

*Values in parentheses are for highest-resolution shell.

Table S2. SAXS data collection and analysis of G6PD^{WT} and G6PD^{P396L}

	WT (Dimer)	WT (Tetramer)	P396L
Data collection parameters			
Instrument		SSRL BL4-2	
Type of Experiment		SEC-SAXS	
Defining slits size (H mm × V mm)		0.3 × 0.30	
Detector distance (m)		1.7	
Detector		Pilatus3 X 1M	
Beam energy (keV)		11.0	
q range (Å^{-1})		0.0072–0.506	
Sample cell		Quartz capillary (D=~1.2mm)	
Temperature (K)		298	
Exposure time/frame (s)		1	
Frames per SEC-SAXS data set		500	
Number of blank images used for averaging		100	
Number of sample images used for averaging		5	
Image numbers used for averaging	330-334	305-309	375-379
SEC column		Superdex 200 Increase PC 3.2/300	
HPLC flow rate (mL/min)		0.05	
Sample concentration (mg/ml)	5	10	5
SEC injection volume (µL)	15	15	100
Buffer		20 mM Tri-HCl (pH=8.0), 150 mM NaCl	
Software employed			
Primary data reduction		*SasTool/SECPipe*	
Data processing		*PRIMUS*	
P(r) analysis		*GNOM*	
Atomistic modeling		*CORAL*	
Structural parameters			
Guinier analysis			
$I(0)$	0.053 ± 0.00025	0.18 ± 0.00047	0.14 ± 0.00069
R_g (Å)	37.80 ± 0.27	43.31 ± 0.16	36.37 ± 0.28
q_{min} (Å^{-1})	0.0103	0.00821	0.0109
qR_g range	0.39 - 1.29	0.36 - 1.30	0.40 - 1.28
P(r) analysis			
$I(0)$	0.05	0.18	0.14
R_g (Å)	38.89	43.24	37.39
D_{max} (Å)	122.87	139.74	130.00
q range (Å^{-1})	0.0103 – 0.211	0.00821 – 0.184	0.0109 – 0.220

Porod volume estimate ($Å^3$)	187621	325948	178097
Atomistic Modeling (CORAL)			
q range ($Å^{-1}$)	0.0109 – 0.249	0.00927 – 0.250	0.0125 – 0.249
Number of repetition	20	20	20
χ^2 range	1.74 – 1.90	2.66 – 3.15	0.23 – 0.34
Predicted R_g (Å)	37.24	42.48	35.92
Predicted D_{max} (Å)	126.1	145.5	154.6

Table S3. Structural features of wild-type and pathogenic mutants of G6PD

	G6P	catalytic $NADP^+$	structural $NADP^+$	βM, βN strands	C-terminal tail	Note	Reference	PDB ID
WT with the structural $NADP^+$	empty	empty	$NADP^+$	ordered	ordered		2	6E08
WT with the catalytic and structural $NADP^+$s	empty	$NADP^+$	$NADP^+$	ordered	ordered	N-terminal 1-25 is truncated.	3	2BH9
WT with G6P	G6P	empty	empty	ordered	disordered	N-terminal 1-25 is truncated.	3	2BHL
R459L	occluded	$NADP^+$	empty	disordered	disordered		2, 4	1QKI, 6E07
P396L	occluded	$NADP^+$	empty	disordered	disordered			6VA7
R393H	occluded	$NADP^+$	empty	disordered	disordered			6VA9
W509A	occluded	$NADP^+$	empty	disordered	disordered			6VA0
V394L	occluded	$NADP^+$	empty	disordered	disordered			6VAQ
F381L	occluded	$NADP^+$	empty	disordered	disordered			6VA8

Table S4. Kinetic analysis of the G6PD mutants

	K_M NADP$^+$ (μM)	K_M G6P (μM)	k_{cat} (s^{-1})	V_{max} (μMs^{-1})	k_{cat}/K_M (μM$^{-1}s^{-1}$)
WT	429 ± 36	225 ± 14	1105 ± 13	9.29 ± 0.11	4.95 ± 0.35
F381L	440 ± 40	109 ± 17	35.5 ± 0.5	0.298 ± 0.004	0.343 ± 0.056
R393H	403 ± 56	161 ± 17	90.0 ± 7.9	0.756 ± 0.067	0.567 ± 0.053
V394L	ND	ND	ND	ND	ND
P396L	ND	ND	ND	ND	ND
W509A	445 ± 59	267 ± 19	368 ± 1.5	3.09 ± 0.01	1.39 ± 0.10

ND: not determined

Movie S1 (separate file). Molecular dynamics simulation of the tetramer $G6PD^{WT}$ without the structural $NADP^+$.

Movie S2 (separate file). Molecular dynamics simulation of the dimer $G6PD^{WT}$ with the structural $NADP^+$.

Movie S3 (separate file). Molecular dynamics simulation of the dimer $G6PD^{WT}$ without the structural $NADP^+$.

Movie S4 (separate file). Molecular dynamics simulation of the C-terminal end truncated dimer $G6PD^{P396L}$ showing partial unbinding of structural $NADP^+$ and opening of the βN-strand.

SI References

1. P. D. Adams, *et al.*, PHENIX : a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).
2. S. Hwang, *et al.*, Correcting glucose-6-phosphate dehydrogenase deficiency with a small-molecule activator. *Nat. Commun.* **9**, 4045 (2018).
3. M. Kotaka, *et al.*, Structural studies of glucose-6-phosphate and NADP + binding to human glucose-6-phosphate dehydrogenase. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **61**, 495–504 (2005).
4. S. W. N. Au, S. Gover, V. M. S. 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 (2000).