

Structural studies of glucose-6-phosphate and NADP⁺ binding to human glucose-6-phosphate dehydrogenase

Masayo Kotaka,^{a,b*‡} Sheila Gover,^a Lucy Vandeputte-Rutten,^a§ Shannon W. N. Au,^{a,b¶} Veronica M. S. Lam^{b‡‡} and Margaret J. Adams^{a*}

^aLaboratory of Molecular Biophysics, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, England, and ^bDepartment of Biochemistry, University of Hong Kong, Faculty of Medicine Building, Sassoon Road, Pokfulam, Hong Kong SAR, People's Republic of China

‡ Current address: Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, England.

§ Current address: Department of Crystal and Structural Chemistry, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands.

¶ Current address: Department of Biochemistry, Chinese University of Hong Kong, New Territories, Hong Kong SAR, People's Republic of China.

‡‡ Dr Veronica Lam is missing following the Indian Ocean tsunami of 26 December 2004. We wish to record here her important contribution to research on human G6PD.

Correspondence e-mail:
masayo.kotaka@strubi.ox.ac.uk,
margaret.adams@biop.ox.ac.uk

Human glucose-6-phosphate dehydrogenase (G6PD) is NADP⁺-dependent and catalyses the first and rate-limiting step of the pentose phosphate shunt. Binary complexes of the human deletion mutant, ΔG6PD, with glucose-6-phosphate and NADP⁺ have been crystallized and their structures solved to 2.9 and 2.5 Å, respectively. The structures are compared with the previously determined structure of the Canton variant of human G6PD (G6PD_{Canton}) in which NADP⁺ is bound at the structural site. Substrate binding in ΔG6PD is shown to be very similar to that described previously in *Leuconostoc mesenteroides* G6PD. NADP⁺ binding at the coenzyme site is seen to be comparable to NADP⁺ binding in *L. mesenteroides* G6PD, although some differences arise as a result of sequence changes. The tetramer interface varies slightly among the human G6PD complexes, suggesting flexibility in the predominantly hydrophilic dimer–dimer interactions. In both complexes, Pro172 of the conserved peptide EKPxG is in the *cis* conformation; it is seen to be crucial for close approach of the substrate and coenzyme during the enzymatic reaction. Structural NADP⁺ binds in a very similar way in the ΔG6PD–NADP⁺ complex and in G6PD_{Canton}, while in the substrate complex the structural NADP⁺ has low occupancy and the C-terminal tail at the structural NADP⁺ site is disordered. The implications of possible interaction between the structural NADP⁺ and G6P are considered.

1. Introduction

Human glucose-6-phosphate dehydrogenase (G6PD; EC 1.1.1.49) is an X-linked housekeeping enzyme that catalyses the first and rate-limiting step of the pentose phosphate shunt. It converts β-D-glucose 6-phosphate (G6P) to 6-phosphoglucono-δ-lactone with the reduction of NADP⁺ to NADPH, providing cells with pentoses for nucleic acid synthesis and NADPH as the principal modulator of intracellular redox potential. G6PD has been identified as the only NADPH-producing enzyme that is activated during oxidative stress (Filosa *et al.*, 2003). In erythrocytes, the role of G6PD is particularly important as the sole source of NADPH.

Active human G6PD exists in a rapid dimer–tetramer equilibrium, which is affected by ionic strength and pH (Cohen & Rosemeyer, 1969). Little is known about the physiological significance of the interconversion. Kinetically homogenous recombinant human G6PD has been shown recently to obey a rapid-equilibrium random-order mechanism (Wang *et al.*, 2002).

Multiple sequence alignment of over 100 currently known G6PDs from different organisms shows sequence identity varying from 30 to 94% (data not given). The human enzyme and *Leuconostoc mesenteroides* G6PD (LM G6PD), the

Received 27 August 2004

Accepted 21 January 2005

PDB References: ΔG6PD–G6P, 2bhl, r2bhlsf; ΔG6PD–NADP⁺, 2bh9, r2bh9sf.

structures of which are known, have 35% sequence identity. The sequence alignment reveals three conserved regions: a nine-residue peptide (RIDHYLGKE, residues 198–206 of the human enzyme), a nucleotide-binding fingerprint (GxxG-GDLA, residues 38–44 of the human enzyme) and the sequence EKPxG (residues 170–174 of the human enzyme). In the nine-residue peptide, the aspartate, histidine and lysine have been shown to be important in G6P binding and catalysis in LM G6PD (Cosgrove *et al.*, 2000) and Lys205 has been implicated in binding and catalysis in the human enzyme (Bautista *et al.*, 1995). The nucleotide fingerprint has been associated with coenzyme binding (Levy *et al.*, 1996) and the function of these residues has been elucidated by the structures of LM G6PD (Rowland *et al.*, 1994; Naylor *et al.*, 2001).

The 3 Å resolution structure of the human G6PD Canton (R459L) variant (G6PD_{Canton}) crystallized in the presence of NADP⁺ and glycollate ions has been described previously (Au *et al.*, 2000). Human G6PD is similar to LM G6PD, with a ‘Rossmann-fold’ coenzyme-binding domain and a β+α domain which forms the dimer interface (Fig. 1). The human G6PD tetramer has 222 symmetry. The tetramer interface is small and interactions are primarily electrostatic. No NADP⁺ was found in the expected coenzyme site; ‘structural’ NADP⁺ was found to bind in the β+α domain of G6PD_{Canton} between the β-sheet and the C-terminus. The site is close to the dimer interface, but all contacts are made to side-chain atoms of one subunit. There is no structural NADP⁺ site in LM G6PD.

G6PD deficiency is one of the most common human genetic defects. Common clinical manifestations include acute haemolytic anaemia, neonatal jaundice and favism, symptoms which are normally triggered by oxidative agents (Beutler, 1990). To date, over 150 different human G6PD mutants with a spectrum of clinical severity have been identified. The rarer severe variants (class I mutants) give rise to chronic nonspherocytic haemolytic anaemia. In all cases, some residual enzyme activity is found; there is no known human null mutant. It has recently been shown that the total loss of G6PD activity is an embryonic lethal in mouse (Longo *et al.*, 2002). Levels of NADPH are reduced in erythrocytes containing low-

activity variants (Mareni & Gaetani, 1976). Among the 71 reported class I mutants, 28 are clustered within 4.5 Å of the dimer interface and the structural NADP⁺ site; the integrity of these regions is important for enzyme stability and *in vivo* activity.

This paper describes G6P and NADP⁺ complexes of the human G6PD deletion mutant ΔG6PD. This mutant was created in the hope of growing higher quality crystals (Au *et al.*, 1999). It is not a natural variant; the 25 N-terminal residues of human G6PD that precede the dinucleotide-binding fold have been removed. Not only were these residues poorly ordered in the G6PD_{Canton} structure, but there is also very poor conservation among species of the region N-terminal of the Rossmann fold and great variation in length. This mutant enzyme is here shown to have similar initial rate kinetic properties to those of the wild type. The structures of the ΔG6PD complexes are described and the proteins and structural NADP⁺ sites are compared with G6PD_{Canton} (Au *et al.*, 2000). The coenzyme and substrate sites are compared with those previously described in LM G6PD (Cosgrove *et al.*, 2000; Naylor *et al.*, 2001). The contributions of this study to the understanding of the kinetic mechanism and the stability of human G6PD are discussed.

2. Materials and methods

2.1. Cloning, expression and purification of ΔG6PD

ΔG6PD was constructed as described previously (Au *et al.*, 1999). Expression of ΔG6PD was carried out using the G6PD-deficient *Escherichia coli* strain DF213 [Δ(eda-zwf)15, hisG1, rpsL115, metA28, mu⁺] (*E. coli* Stock Centre, Yale University). Terrific Broth (500 ml) containing ampicillin (200 μg ml⁻¹) was inoculated with 5 ml of an overnight culture of *E. coli* containing the recombinant plasmid pTrc/ΔG6PD. When the OD₆₀₀ of the culture reached 0.7–0.8, 0.5 mM isopropyl thio-β-D-galactoside was added to induce protein expression. Expression was carried out for 16–18 h at 310 K. Harvested cells were resuspended in 0.1 M Tris–HCl pH 7.6 and 5 mM EDTA with Complete EDTA-free protease-inhibitor cocktail (Roche) and lysed by sonication. After centrifugation, the supernatant was loaded onto a 2',5'-ADP Sepharose column (Amersham). The enzyme was eluted with 75 μM NADP⁺ in 0.1 M Tris–HCl pH 7.6, 5 mM EDTA and Complete EDTA-free protease-inhibitor cocktail and concentrated to 5 mg ml⁻¹ using Amicon stirred cells (Millipore).

2.2. Measurement of initial rate kinetic parameters

The fluorimetric measurement and calculation of initial rate kinetics parameters of ΔG6PD were carried out as described previously (Wang *et al.*, 2002).

2.3. G6P-bound ΔG6PD in C222₁

2.3.1. Crystallization. Initial screening for crystallization conditions was carried out by hanging-drop vapour diffusion using Molecular Dimensions Structure Screens 1 and 2.

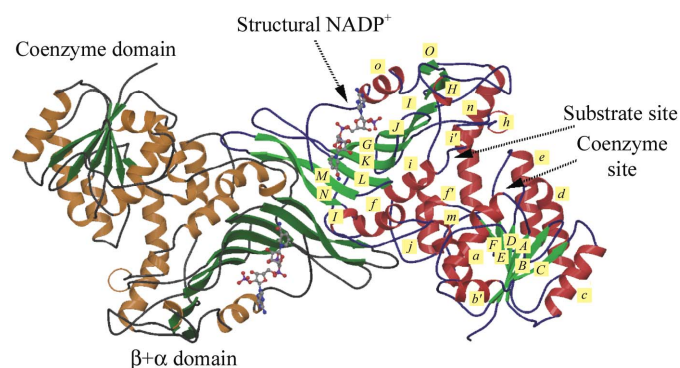


Figure 1
The human G6PD_{Canton} dimer. Helices and sheet strands of the A subunit are shown in red and green, respectively, and each of the secondary-structure elements identified in the G6PD_{Canton} structure is labelled. The figure was prepared using BOBSCRIPT (Kraulis, 1991; Esnouf, 1997) and RASTER3D (Merrit & Murphy, 1994; Bacon & Anderson, 1988), as were all other diagrams in this paper.

Table 1Data-processing statistics of G6P-bound and NADP⁺-bound Δ G6PD.

Values in parentheses are for the last shell.

	G6P-bound Δ G6PD	NADP ⁺ -bound Δ G6PD
Wavelength (Å)	1.488	0.87
Space group	<i>C</i> 222 ₁	<i>F</i> 222
Unit-cell parameters (Å)		
<i>a</i>	117.55	60.76
<i>b</i>	179.53	172.48
<i>c</i>	137.94	217.22
G6PD molecules per AU	2	1
Resolution (Å)	30–2.9 (3.06–2.9)	30–2.5 (2.6–2.5)
Oscillation angle (°)	1	1
Completeness (%)	98.9 (92.9)	73.3 (59.3)
No. of reflections	503036	115667
No. of unique reflections	32261	18730
Multiplicity	6.8	2.9
<i>I</i> > 3 σ (<i>I</i>) (%)	76.4 (31.3)	66.3 (35.6)
<i>R</i> _{merge} (%)	11.8 (41.2)	12.5 (31.4)

Crystals were found to grow in solution 33 (0.1 M Tris–HCl pH 8.5, 0.2 M MgCl₂, 30% PEG 4000). These conditions were optimized using sitting-drop vapour diffusion. Drops made by mixing equal volumes of 5 mg ml⁻¹ Δ G6PD in elution buffer (0.1 M Tris–HCl pH 7.5, 5 mM EDTA, 75 μ M NADP⁺), 60 mM G6P and well buffer (0.1 M Tris–HCl pH 8.5, 0.2 M MgCl₂, 12% PEG 4000, 5% glycerol) were equilibrated for two weeks at 291 K. The crystals grew to approximately 0.6 \times 0.2 \times 0.2 mm.

2.3.2. Data collection. The data set was collected at 100 K using an Oxford Cryosystems Cryostream on station 14.1 at the CCLRC Synchrotron Radiation Source, Daresbury Laboratory, England. The crystal was flash-frozen in cryoprotectant containing 0.1 M Tris–HCl pH 8.5, 0.2 M MgCl₂, 12% PEG 4000 and 30% glycerol. Diffraction did not extend beyond 2.9 Å. Images were processed using *MOSFLM* (Leslie, 1999) and the data were scaled using *SCALA* from the *CCP4* package (Collaborative Computational Project, Number 4, 1994). The space group is *C*222₁, with unit-cell parameters *a* = 117.6, *b* = 179.5, *c* = 137.9 Å ($\alpha = \beta = \gamma = 90^\circ$). There are two subunits in the asymmetric unit; the solvent content is approximately 60%. Data-collection statistics are given in Table 1. A randomly selected set of 5% of the terms was set aside for *R*_{free}-factor validation and was not involved in any refinement or electron-density map calculation (Brünger, 1992a).

2.3.3. Molecular replacement, refinement and model building. Molecular replacement was performed using *MOLREP* (Vagin & Teplyakov, 1997) from the *CCP4* package with a preliminary incomplete data set. The dimer of subunits *A* and *B* of G6PD_{Canton} was used as a search model. The solutions for the two possible orientations of the dimer had correlation coefficients of 0.53 and 0.52. The next highest correlation coefficient was 0.39. A partly refined model gave correlation coefficients of 0.63 and 0.60 for the two dimer orientations using the complete data set. After rigid-body refinement with *REFMAC5* (Murshudov *et al.*, 1997), the *R* and *R*_{free} factors of the initial solution were 41.3 and 41.9%, respectively. σ_A -weighted 2|*F*_o| – |*F*_c| and |*F*_o| – |*F*_c| electron-

Table 2Refinement statistics and quality indicators of G6P-bound and NADP⁺-bound Δ G6PD.

Data set	G6P-bound Δ G6PD	NADP ⁺ -bound Δ G6PD
Resolution (Å)	30–2.9	30–2.5
Observations in refinement working set	30621	13822
Observations in validation set	1640	726
<i>R</i> factor (%)	21.2	19.6
<i>R</i> _{free} (%)	26.1	29.6
Mean <i>B</i> factors (Å ²)		
Main-chain atoms	47.5	29.4
Side-chain atoms	54.2	36.2
Water	36.9	30.2
G6P	49.0	—
Coenzyme NADP	—	34.0
Structural NADP	—	29.5
Glycerol	54.7	33.0
R.m.s.d. distances (Å)	0.006	0.009
R.m.s.d. bond angles (°)	0.95	1.03
Ramachandran plot†, residues in		
Most favoured regions	85.3	86.8
Additionally allowed regions	14.7	13.2
Generously allowed regions	0	0
Disallowed regions	0	0

† As defined by *PROCHECK* (Laskowski *et al.*, 1993).

density maps (Read, 1986) were generated using the *CCP4* program suite. Several cycles of manual rebuilding using *O* (Jones & Kjeldgaard, 1997) followed by refinement using the program *BUSTER-TNT* (Blanc *et al.*, 2004) were carried out. Water molecules were included at stereochemically sensible positions in the structure where the electron density was at least 1 σ in the 2|*F*_o| – |*F*_c| map and at least 3 σ in the |*F*_o| – |*F*_c| map. The final model, referred to as Δ G6PD–G6P, contains 956 G6PD residues, two G6P molecules, six molecules of glycerol and 82 water molecules. At convergence, the *R* and *R*_{free} factors are 21.2 and 26.1%, respectively, and 85.3% of the residues lie within the most favoured regions of the Ramachandran plot, as defined by *PROCHECK* (Laskowski *et al.*, 1993). Refinement details and quality indicators are given in Table 2.

2.4. NADP⁺-bound Δ G6PD in *F*222

2.4.1. Crystallization, data collection and molecular replacement. Expression, crystallization, data collection to 2.5 Å and structure determination by molecular replacement of Δ G6PD–NADP⁺ have previously been reported (Au *et al.*, 1999). Data-collection statistics are repeated in Table 1.

2.4.2. Refinement and model building. Several cycles of manual rebuilding using *O* and *X-PLOR* (v.3.851; Brünger, 1992b) energy minimization were carried out. In the later stages of refinement, the program *BUSTER-TNT* was used. Water molecules were included at stereochemically sensible positions in the structure where the electron density was at least 1 σ in the 2|*F*_o| – |*F*_c| map and at least 3 σ in the |*F*_o| – |*F*_c| map. The final model contains 489 G6PD residues, one structural NADP⁺ molecule, one coenzyme NADP⁺ molecule, 162 water molecules and two molecules of glycerol. At convergence, the *R* and *R*_{free} factors are 19.8 and 29.6%, respectively, and 87% of the residues lie within the most

Table 3

Dalziel parameters and their ratios for the reaction catalysed by wild-type G6PD (G6PD WT; Wang *et al.*, 2005) and Δ G6PD.

Enzyme	φ_0 (s)	φ_{NADP} ($\mu\text{M s}$)	φ_{G6P} ($\mu\text{M s}$)	$\varphi_{\text{NADP-G6P}}$ ($\mu\text{M}^2 \text{s}$)	k_{cat} (s^{-1})	K_{mNADP} (μM)	K_{mG6P} (μM)	$k_{\text{cat}}/K_{\text{mNADP}}$ ($\mu\text{M}^{-1} \text{s}^{-1}$)	$k_{\text{cat}}/K_{\text{mG6P}}$ ($\mu\text{M}^{-1} \text{s}^{-1}$)
G6PD WT	0.0037	0.0259	0.191	1.61	275	7.07	52	39.7	5.31
Δ G6PD	0.0042	0.0271	0.175	1.55	208	6.45	42	32.3	5.02

favoured regions of the Ramachandran plot, as defined by PROCHECK. Refinement details and quality indicators are given in Table 2.

3. Results and discussion

3.1. Initial rate kinetic properties of Δ G6PD

Δ G6PD has been cloned and expressed in *E. coli* and purified by one-step affinity chromatography using 2',5'-ADP Sepharose. The specific activity of the purified mutant ($\sim 190 \text{ IU mg}^{-1}$) is similar to that of the recombinant human wild-type G6PD ($\sim 180 \text{ IU mg}^{-1}$). k_{cat} and K_{m} for NADP⁺ and K_{m} for G6P obtained fluorimetrically for Δ G6PD were also comparable to those of the recombinant human wild-type enzyme (Wang *et al.*, 2005; Table 3).

3.2. G6P-bound Δ G6PD structure in C222₁

Crystals of Δ G6PD grown in 0.1 M Tris-HCl pH 8.5, 0.2 M MgCl₂ and 10–14% PEG 4000 in the presence of G6P only diffracted to 6 Å. Addition of 5% glycerol to the crystallization buffer greatly increased the resolution limit of the X-ray diffraction images to 2.9 Å. The two subunits in the asymmetric unit are related by a non-crystallographic twofold axis in the *bc* plane and at 31° to *c*. The dimers are related by a

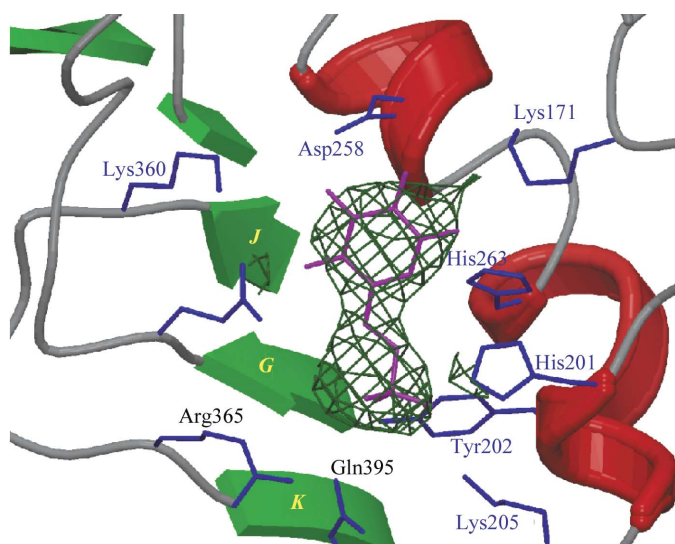


Figure 2
Electron density for bound G6P seen in the initial Δ G6PD-G6P map. The $2|F_o| - |F_c|$ density is contoured at 1σ . Neighbouring residues that are conserved are labelled in blue. Sheet strands are identified as in Fig. 1.

crystallographic twofold axis parallel to *a* to form a 222 tetramer.

The crystal was found to contain bound G6P; the electron density was clearly seen in each subunit in the initial map (Fig. 2 shows subunit *B*). The C-terminus of Δ G6PD-G6P is disordered and the final ten residues of both subunits have been omitted from the refinement. The conserved prolines 172 and

396 are both found here in the *cis* conformation. The structural NADP⁺ site in the $\beta+\alpha$ domain is no more than partly occupied in either subunit. Since there was no improvement in R_{free} when the structural NADP⁺ was included, no structural NADP⁺ has been included in the final refinement cycle; however, the three largest peaks on the final $|F_o| - |F_c|$ map are close to the anticipated bisphosphate and nicotinamide of structural NADP⁺ in subunit *B* and to the expected bisphosphate in subunit *A*.

There is some evidence that the coenzyme NADP⁺ site is also partly occupied in subunit *B*. The final $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ electron-density maps show significant density in the anticipated adenine site, but this is lower than the peaks in the structural NADP⁺ site. Electron-density maps for subunit *A* gave no indication of coenzyme NADP⁺ binding, although the coenzyme-binding site, while more exposed to solvent than that of subunit *B*, remains relatively ordered. No coenzyme NADP⁺ has been included in refinement.

The bound G6P is well ordered and its displacement parameters are similar to those of its protein ligands. The hydrogen-bonding network in the substrate-binding site of Δ G6PD-G6P is shown in Fig. 3. In common with the G6P-bound LM G6PD (LM G6PD-G6P) complex (Cosgrove *et al.*, 2000), G6P is bound in the pocket between domains. Most of the residues of Δ G6PD that interact with G6P are conserved and correspond to those in LM G6PD-G6P.

A distinction can be made between the edge of G6P that approaches the coenzyme domain, which includes the hydroxyl atom (O1) to be oxidized, and the edge that approaches the large mixed β -sheet. In addition to atom O1, phosphate atoms O6 and O8, ring atom O5 and hydroxyl atom O2 face the coenzyme domain. Contacts are preserved between human Δ G6PD and LM G6PD for these atoms. Tyr202 (179)¹, His201 (178) and Lys205 (182) from the highly conserved peptide sequence at the domain boundary interact with phosphate O atoms, while Asp258 (235) contacts O2 and O3, and His263 (240) N^{ε2} contacts O1.

The mutant LM G6PD H178N showed dramatic increases in K_{m} and K_{d} for G6P and a loss of the ability to discriminate between G6P and D-glucose (Cosgrove *et al.*, 1998). Mutations K205R and K205T were shown to affect k_{cat} primarily and thus to be essential for catalysis in human G6PD (Bautista *et al.*, 1995). A different result was found for LM G6PD, where K_{m} was primarily affected on mutagenesis of Lys182 (Vought *et al.*, 2000). The two-point interaction of Asp258 is also seen in

¹ Residue numbers in roman type are for the human enzyme; those in italics are for the *L. mesenteroides* enzyme.

LM G6PD–G6P (Cosgrove *et al.*, 2000). In Δ G6PD–G6P, atom O1 of G6P is 2.9 Å from the N^{ε2} atom of His263; N^{δ1} of this residue hydrogen bonds with Asp200 O^{δ1}. This interaction is the same as that seen in LM G6PD, where His240 is the general base for catalysis and Asp177 and His240 form the catalytic dyad (Cosgrove *et al.*, 2000). The mutation of H240N in LM G6PD caused a decrease in k_{cat} of the enzyme by a factor of 10^5 (Cosgrove *et al.*, 1998). There is, however, one

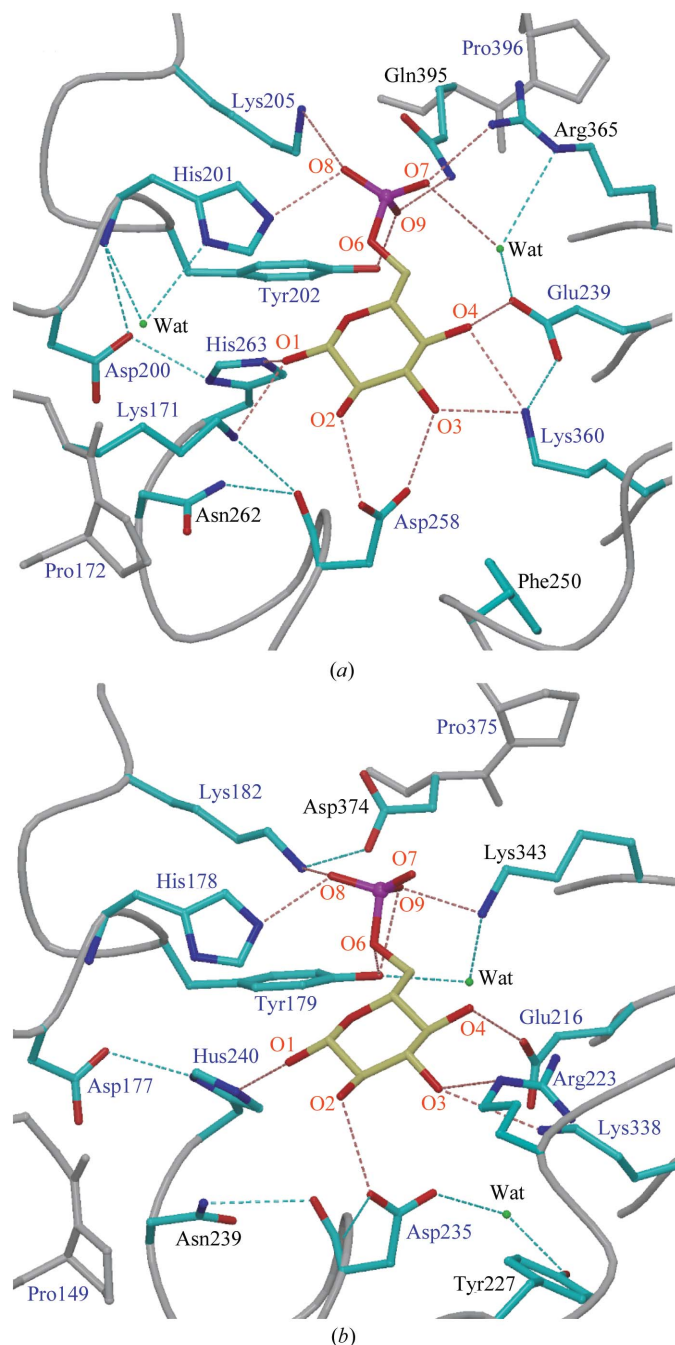


Figure 3
The hydrogen-bonding network in the substrate site of (a) Δ G6PD–G6P and (b) LM G6PD. C atoms of G6P are drawn in cream; those of residues interacting with G6P are drawn in cyan. Potential hydrogen bonds are indicated by dotted lines. Those that directly involve atoms of the substrate are drawn in pink; others are in cyan. Neighbouring residues that are conserved are labelled in blue.

significant difference in the contacts with O1 in the two species. In Δ G6PD, Lys171, the second residue of the conserved EKPxG peptide, is seen to interact with O1 and O2 of G6P in subunit B, although it is not well ordered. In contrast, no density can be seen for the side chain of Lys149 of LM G6PD–G6P.

Differences are seen between the two species in binding the edge of G6P, which interacts with the mixed β -sheet, even though some of the residues involved are conserved. In Δ G6PD–G6P, Arg365 and Gln395 interact with phosphate O atoms. Arg365 hydrogen bonds directly with phosphate atom O7 and interacts indirectly with phosphate atom O9 via a water molecule, while in LM G6PD–G6P the equivalent Lys343 interacts with phosphate atom O9 (Cosgrove *et al.*, 2000). The interaction of Gln395 (Asp374) with the phosphate of G6P is not conserved. Lys360 (338) bridges atoms O3 and O4 in Δ G6PD, while only atom O3 is involved in LM G6PD. Glu239 (216) interacts with O4 in both species, but Arg246 (223) is disordered in Δ G6PD while it is a ligand in LM G6PD. The main chain of Arg246 is far enough away from G6P to preclude an equivalent interaction. Glu239, Lys360 and Arg365 are close to the ends of the sheet strands β G, β I and β K, respectively.

3.3. NADP⁺-bound Δ G6PD structure in F222

The Δ G6PD–NADP⁺ binary complex crystallizes in space group F222 with one G6PD subunit in the asymmetric unit. The molecule forms a tetramer with crystallographic 222 symmetry. Prolines 172 and 396 are again in the *cis* conformation.

NADP⁺ is bound to both the coenzyme site and the structural NADP⁺ site, even though no additional NADP⁺ was present in the crystallization solution. Both the structural and the coenzyme NADP⁺s have full occupancy and are well ordered with displacement parameters close to those of the surrounding protein; those for the coenzyme NADP⁺ are higher than for the structural NADP⁺. Coenzyme NADP⁺ is shown in the final $2|F_o| - |F_c|$ electron-density map of the region in Fig. 4(a). The molecule is positioned similarly to NADP⁺ in the LM G6PD–NADP⁺ complex (Naylor *et al.*, 2001). NADP⁺ binds with the adenine ring in the *anti* conformation and the two ribose rings and the bisphosphate span the C-terminus of the β -sheet of the coenzyme domain. In the built conformation, the nicotinamide approaches O1 of G6P. In a possible alternative conformation, the nicotinamide ribose is rotated and the nicotinamide is rotated again so that it remains in the entrance to the active-site cleft and is not in the correct position for catalysis. In its built position, the coenzyme conformation resembles that in LM G6PD; in the alternate position, it is less open.

Protein–coenzyme NADP⁺ interactions are shown in Fig. 4(b). There are hydrogen bonds from the adenine ribose 3'-hydroxyl and the bisphosphate O atoms to the main-chain amino groups of Gly41 and Asp42 of the nucleotide-binding fingerprint (GASGDLA) in the β A– α a turn. The nicotinamide ribose 2'-hydroxyl group forms a hydrogen bond to the

carbonyl group of Lys171, one of the residues of the βE - α turn (residues 170–176). Arg72 interacts with the 2'-phosphate, with hydrogen bonds made by both the N⁶ and N⁷ atoms of the arginine to a phosphate O atom. The hydroxyl group of Tyr112 also hydrogen bonds to this phosphate O atom; in LM G6PD, the equivalent residue is Val86 and this interaction cannot occur. A second phosphate O atom makes a hydrogen bond to Ser73 O^γ; the interaction is not possible for Gln47 of LM G6PD. A hydrogen bond to Ser40 O^γ mirrors that with the equivalent Thr14 O^γ1 of LM G6PD–NADP⁺ (Naylor *et al.*, 2001).

The environment of the adenine is surprisingly different between the Δ G6PD–NADP⁺ and LM G6PD–NADP⁺ complexes (Fig. 4c). Two regions shield the adenine ring from solvent in LM G6PD (Naylor *et al.*, 2001). The first, proximal to the 2'-phosphate, comprises Ala45 (71) and Arg46 (72) and is retained to a large extent. However, Gln47 (Ser73) is poorly conserved (60% Ser or Thr and only 5% Gln) and the differing side-chain lengths affect packing around the adenine and adenine ribose. The face of the adenine distal to the 2'-phosphate is affected by the region 110–112 (84–86). In human G6PD, the sequence is Gly-Gln-Tyr, while in LM G6PD it is His-Asp-Val. The hydrogen bond between Tyr112 O^η and the 2'-phosphate forces the adenine ring to approach the solvent, while in the LM G6PD–NADP⁺ complex Val86 makes a hydrophobic surface for the adenine and further contacts are made with the side chain of Asp85 and the main chain of His84. In both species there is some contribution from Leu142 and Val146 (Val118 and Phe122) in the βD - αD turn; Pro143 (Ala119) is also involved in Δ G6PD but not in LM G6PD. The more hydrophobic environment for

the adenine ring in the LM G6PD complex may be a significant contribution to dual coenzyme specificity in that species.

3.4. The G6PD tetramer

Mean coordinate differences between the crystallographically independent subunits of these structures and the subunits of the ABCD tetramer in G6PD_{Canton} are given in Table 4. The two subunits of Δ G6PD–G6P are very similar. The main differences at the subunit level between either Δ G6PD or G6PD_{Canton}, other than at the N- and C-termini, correspond to the changes consequent on the change of Pro172 from *trans* in G6PD_{Canton} to *cis* in both Δ G6PD structures. Additionally, helix αb is longer and better ordered in Δ G6PD.

The larger differences seen for dimer-to-dimer alignments relate to slightly different hinge angles between coenzyme and $\beta + \alpha$ domains, particularly when coenzyme is present; the coenzyme domains are at the greatest distance from the 222 centre. The hinge-angle differences are much smaller than those that were seen in LM G6PD complexes (Naylor *et al.*, 2001).

The relatively small tetramer contact involves two surfaces close to the 222 centre. Although the interface surfaces are not identical in the three space groups, the contact regions are very similar and many of the same residues, at the junction of αi and αj , on αj and in the βI - βJ and βK - βL loops, are important. The contacts are predominantly polar; each subunit contains eight basic and ten acidic residues with functional groups less than 10 Å apart across the interface.

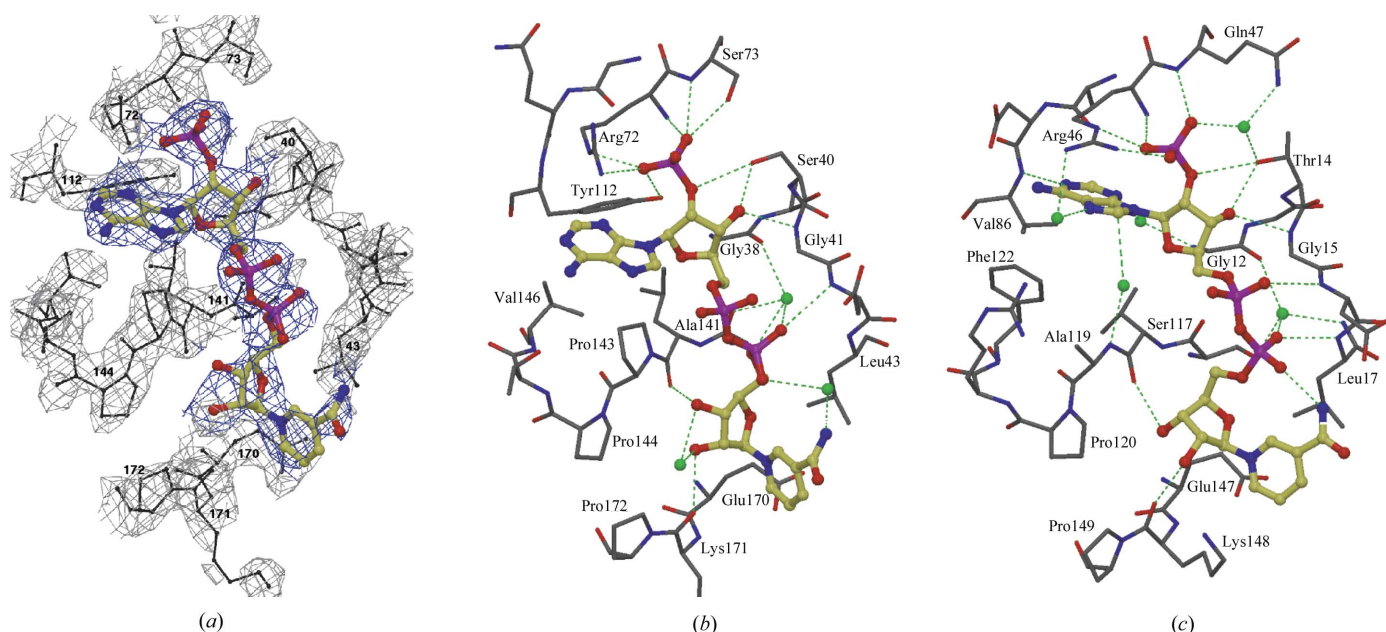


Figure 4
 (a) Final $2|F_o| - |F_c|$ map for Δ G6PD–NADP⁺ in the region of the coenzyme NADP⁺. Electron density corresponding to the protein is in grey, contoured at 1.2σ . Density for NADP⁺ is in blue, contoured at 0.9σ . (b) Potential hydrogen-bond interactions for the coenzyme NADP⁺ in Δ G6PD–NADP⁺. (c) Potential hydrogen bonds for the coenzyme NADP⁺ in LM G6PD–NADP⁺. Interactions and direct contacts are to both backbone and side chains of surrounding residues. Interactions mediated by water molecules (in green) are also shown. The coenzyme lies in a classic ‘Rossmann-fold’, with residues 38–44 as the binding fingerprint. The G6PD sequence does not contain the extended motif recently identified in other nucleotide-binding proteins (Kleiger & Eisenberg, 2002).

The salt-bridge interactions between Lys275 and Glu347 and between Lys290 and Glu287 observed in G6PD_{Canton} are not preserved, although these residues remain close in both Δ G6PD complexes. The dimers of Δ G6PD–G6P have moved apart slightly to accommodate a glycerol which contacts Glu347 and this change probably accounts for both the increased order in the presence of glycerol and the increase in mean coordinate differences when the tetramer is aligned with either G6PD_{Canton} or Δ G6PD–NADP⁺. In Δ G6PD–NADP⁺, there are ordered water molecules in the interface and an alternative salt bridge is made between subunits by Asp283 and Lys293, one turn earlier and one turn later in the helices α _j from the two subunits.

Table 4

Mean coordinate differences (Å), for all equivalent atoms, between different molecules.

	Alignment	Δ G6PD– G6P	Δ G6PD– NADP ⁺	G6PD _{Canton}
Δ G6P–G6P	Monomer–monomer	0.58	0.89	0.95
Δ G6P–G6P	Dimer–dimer		1.50	1.20
Δ G6P–G6P	Tetramer–tetramer		2.43	1.91
Δ G6PD–NADP ⁺	Dimer–dimer			1.40
Δ G6PD–NADP ⁺	Tetramer–tetramer			1.58

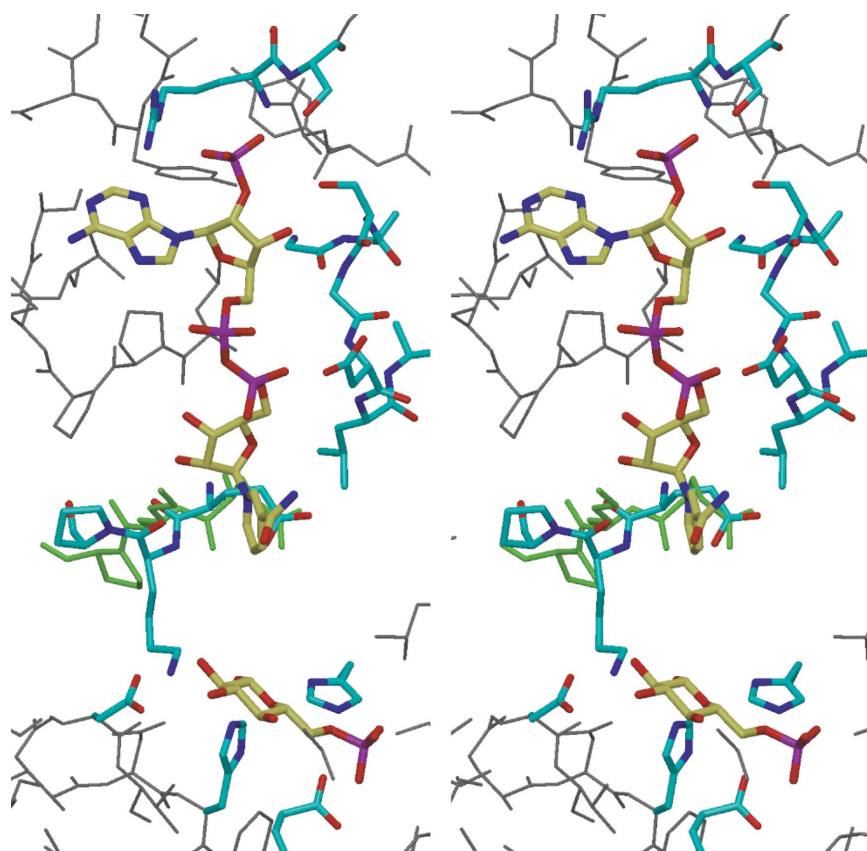


Figure 5

Stereoview of the proximity of coenzyme and substrate. The coenzymes seen in Δ G6PD–NADP⁺ and G6P, superimposed from the Δ G6PD–G6P structure, are drawn with C atoms coloured cream. Residues important in binding or catalysis are drawn in cyan; other portions of the chain are drawn in black. Also superimposed (in green) are residues 170–172 (EKP) from subunit A of G6PD_{Canton}, in which Pro172 is *trans*; the effect on Lys171 of the change to *cis* Pro172 in Δ G6PD–NADP⁺ and Δ G6PD–G6P is evident.

Our different structures suggest the flexible polar interactions are responsible for the dimer–tetramer equilibrium and its sensitivity to ionic strength (Bonsignore *et al.*, 1971; Wrigley *et al.*, 1972). The biological importance of this equilibrium remains obscure; human G6PD is active as either dimer or tetramer.

3.5. Proline 172

The conserved Pro172 (149), the central residue of the conserved EKP_xG peptide, has been found in the *cis* conformation in both Δ G6PD complexes, while in the uncomplexed G6PD_{Canton} structure the residue is *trans* in seven out of eight subunits. In LM G6PD, Pro149 has been shown to be *cis* in all binary complexes (Cosgrove *et al.*, 1998, 2000; Naylor *et al.*, 2001), but *trans* for one subunit of the dimer in the absence of substrate or coenzyme (Rowland *et al.*, 1994). The approach of the substrate and the coenzyme and the proximity of the conserved EKP_xG peptide in the Δ G6PD structures are seen in Fig. 5. The *trans*–*cis* isomerization of Pro172 results in a movement of the helix α _e, allowing Lys171 to interact with G6P O1 through the terminal amino group and with the nicotinamide ribose 3'-hydroxyl of NADP⁺ through its carbonyl group. In LM G6PD the mutations P149G and P149V greatly reduced the *k*_{cat} values of the mutant and affected all the kinetic constants and the dissociation constant for G6P (Vought *et al.*, 2000). The human mutant P172S (Volendam) exhibits class 1 deficiency, while the isolated enzyme has reduced specific activity and the highest *K*_m value for G6P for any G6PD mutant (Roos *et al.*, 1999). These results suggest that Pro172 is directly involved in the correct positioning of the substrate and coenzyme pockets. Mutation to residues that are unable to adopt a *cis* conformation would lead to an extremely high activation-energy barrier for the enzyme to adopt the correct transition state during the enzymatic reaction.

3.6. The structural NADP⁺ site

The structural NADP⁺ site is only found and conserved in higher organisms. The increased number of replacements and the shortened C-terminal tail in prokaryotic G6PD implies the absence of the site (data not shown); it is not present in LM G6PD. The evolutionary significance of this site is still unknown. However, the large number of class 1 mutants clustered in the structural NADP⁺ site and also the presence of several mutants that can be reactivated by increased NADP⁺ concentration (Beutler *et al.*, 1991) suggested that the binding of NADP⁺ is important for the enzymatic

activity in higher organisms and that this region is sensitive to changes in NADP⁺ concentration. Many studies (*e.g.* Scopes *et al.*, 1998) have shown the lower thermal stabilities and differing sensitivity to low NADP⁺ concentration of mutant enzymes with changes close to the structural NADP⁺ site.

In Δ G6PD–NADP⁺, well ordered NADP⁺ is seen in the structural NADP⁺ site between the β -sheet and C-terminus of the monomer. It is shown in the final electron-density map of the region (Fig. 6*a*). As in G6PD_{Canton} (Au *et al.*, 2000), the NADP⁺ is effectively buried in the protein, occupying a crevice which is positively charged. The contacts of the protein to the NADP⁺ (Fig. 6*b*) are essentially the same as those seen in G6PD_{Canton} and are all made by side-chain atoms of the subunit. There are six water molecules that make contact with the NADP⁺; the G6PD_{Canton} structure was of too low a resolution for them to have been included in the refinement. Two contacts are changed from G6PD_{Canton}: the interaction of both Arg357 and Glu364 with the 2'-phosphate is mediated by water molecules in Δ G6PD. Lys238 N^ε, while close to the 2'-phosphate, interacts with Asn363, which itself interacts with a third water molecule bound to the 2'-phosphate.

In Δ G6PD–G6P, the structural NADP⁺ is bound in less than half the molecules. There is little difference between the subunits either in the nature of the site or in the evidence for bound NADP⁺. Both structural NADP⁺s were omitted from the final structure, although the final electron-density maps gave evidence of some binding with a small movement from its

position in G6PD_{Canton} or in Δ G6PD–NADP⁺. A modelled NADP⁺ molecule, optimally positioned in the Δ G6PD–G6P structure, is shown superposed on the final $2|F_o| - F_c$ and the difference density of subunit *B* in Fig. 7.

The absence of an ordered C-terminal region and the poor binding of the structural NADP⁺ are correlated. The nicotinamide ring could not be stabilized in Δ G6PD–G6P since Trp509 is not ordered. In the absence of the nicotinamide ring, the two Asp421 side chains of the dimer have rotated and interact across the interface through a single water molecule. Much of the binding pocket for this part of the structural NADP⁺ has been lost in this complex and any bound NADP⁺ is exposed to solvent. The low occupancy of the structural NADP⁺ site suggests a dynamic equilibrium and that the molecule may migrate to the coenzyme site if NADP⁺ in solution has been depleted. This may result in the observed flexibility of the C-terminal ten residues or the potential for flexibility may be necessary for NADP⁺ migration. Movement of the C-terminal tail is possible with Gly505 acting as a hinge; a group of negatively charged residues (Glu416, Glu417 and Glu419) is exposed to the surface when the tail is mobile (Figs. 6*a* and 7). In a closed system with a large excess of G6P, where the NADP⁺ is essentially all reduced to NADPH, some NADPH may bind at the structural NADP⁺ site; however, the binding will be less tight than that of NADP⁺ as the reduced nicotinamide ring, with a pyramidal ring N atom and limited π -electron delocalization, is unable to form π - π interactions with Tyr401 and Trp509. It follows that, if the C-terminal ten residues are disordered in the absence of the structural NADP⁺, NADPH is not likely to reorder them. Furthermore, the exposed negatively charged patch may attract the positive charge on the oxidized nicotinamide ring to the general area of the binding site while the more negatively charged NADPH is not attracted to it.

The tightness of binding of NADP⁺ to the structural NADP⁺ site may also be affected by G6P binding at the distant G6P site. Although the G6P site is situated some 15 Å from the structural NADP⁺, the two sites are on opposite faces of the large mixed β -sheet (Fig. 8). The most effective way in which the presence of G6P can be signalled to the NADP⁺ site is through groups of residues which are close in the sequence but span from one binding site to the other. There are three sets of residues of this kind. The most extensive cross-connection is in the loop between sheet strands βJ and βK and at the beginning of that strand. In this loop, Asn363 and Glu364 contact bound water molecules, which are themselves

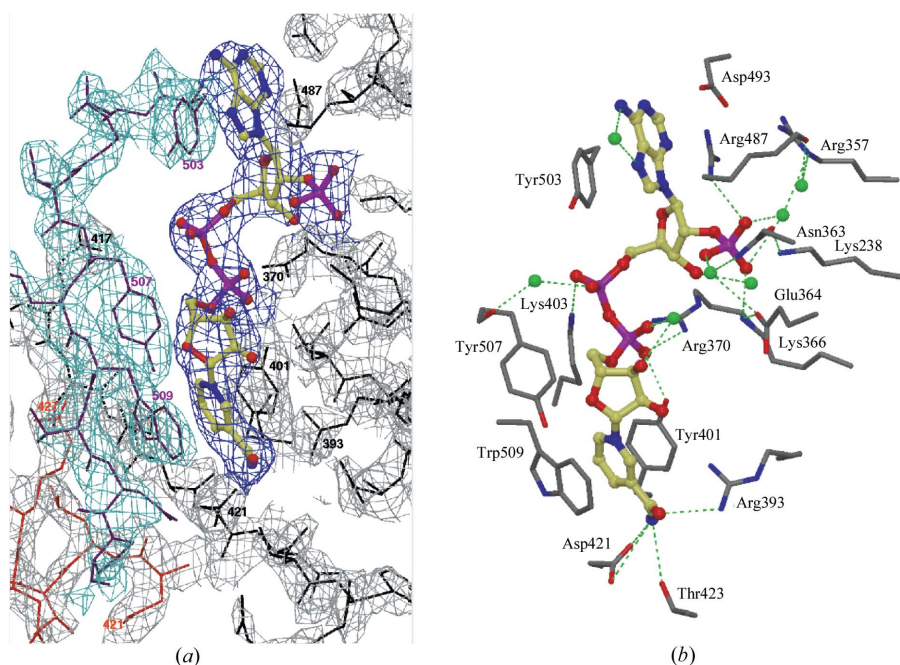


Figure 6
(a) Final $2|F_o| - |F_c|$ map for Δ G6PD–NADP⁺ in the region of the structural NADP⁺. Electron density for most of the molecule is in grey and is contoured at 1.25σ . Density for the C-terminal region from 503 is in cyan and is contoured at 1σ . Density for NADP⁺ is in blue and is contoured at 1.0σ . Residues shown in black are in one subunit of the dimer; those in red are in the other. *(b)* Potential hydrogen bonds for the structural NADP⁺ in Δ G6PD–NADP⁺. Interactions and direct contacts are either with the side chains of surrounding residues or to water molecules (shown in green). The changes in the conformation of NADP⁺ in Δ G6PD when bound in the structural site or in the coenzyme site (Fig. 4*b*) are clear.

ligands to the 2'-phosphate of the structural NADP⁺; Arg365 binds the phosphate of G6P while Lys366 at the beginning of the sheet again interacts with the 2'-phosphate of the structural NADP⁺. At the end of the adjacent sheet strand β G, Lys238 also interacts indirectly with the structural NADP⁺, while Glu239 forms a hydrogen bond to atom O4 of G6P. In the third strand, Gln395, at the end of β L, is a G6P phosphate ligand, while Arg393, within the sheet strand, interacts with the NADP⁺ nicotinamide amide O atom. The three groups of residues on opposite faces of the adjacent sheet strands β G, β K and β L are thus well placed to relay the presence of G6P across the mixed β -sheet.

It has already been noted that the two residues 365 and 395 are not conserved for all species. Of the residues discussed above, only Glu239 and Pro396 are fully conserved, while

Gln395 is 96% conserved; the importance of these last three residues in binding G6P and in defining the shape of the binding site is thus underlined. The specificity for structural NADP⁺ is apparent in that four of the residues involved in the cross-connection are 2'-phosphate ligands: Lys366 directly and Lys238, Asn363 and Glu364 through water molecules.

The significance of the region connecting the two binding sites in fine-tuning the activity of human G6PD is also indicated by the number of severe variants found among the above residues, particularly those relevant to NADP⁺ binding. Durham K238R, Aachen and Loma Linda N363K (nucleotide 1089 C to G and C to A, respectively), Wisconsin R393G, Nashville (Anaheim, Calgary, Portici) R393H, Alhambra V394L and Bari P396L are all class I variants. While our

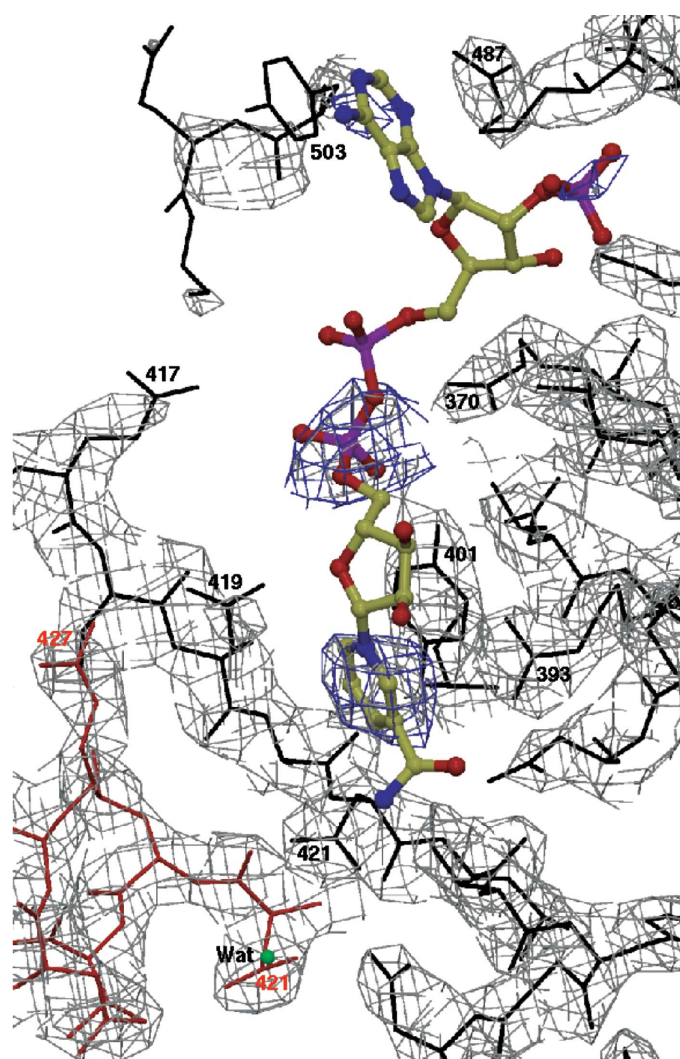


Figure 7
Modelled structural NADP⁺ molecule in subunit *B* of Δ G6PD-G6P. The final $2|F_o| - |F_c|$ map, drawn in grey, is contoured at 1.25σ . The difference electron density, in blue, is contoured at 2.5σ . NADP⁺ was not included in F_c . Residues shown in black are in subunit *B*; those in red are in subunit *A*. Comparison with Fig. 6(a) shows the exposure of charged residues in sheet strand β N when the residues of the tail are removed and the change in position of the side chain of both A421 and B421 in the absence of well bound NADP⁺.

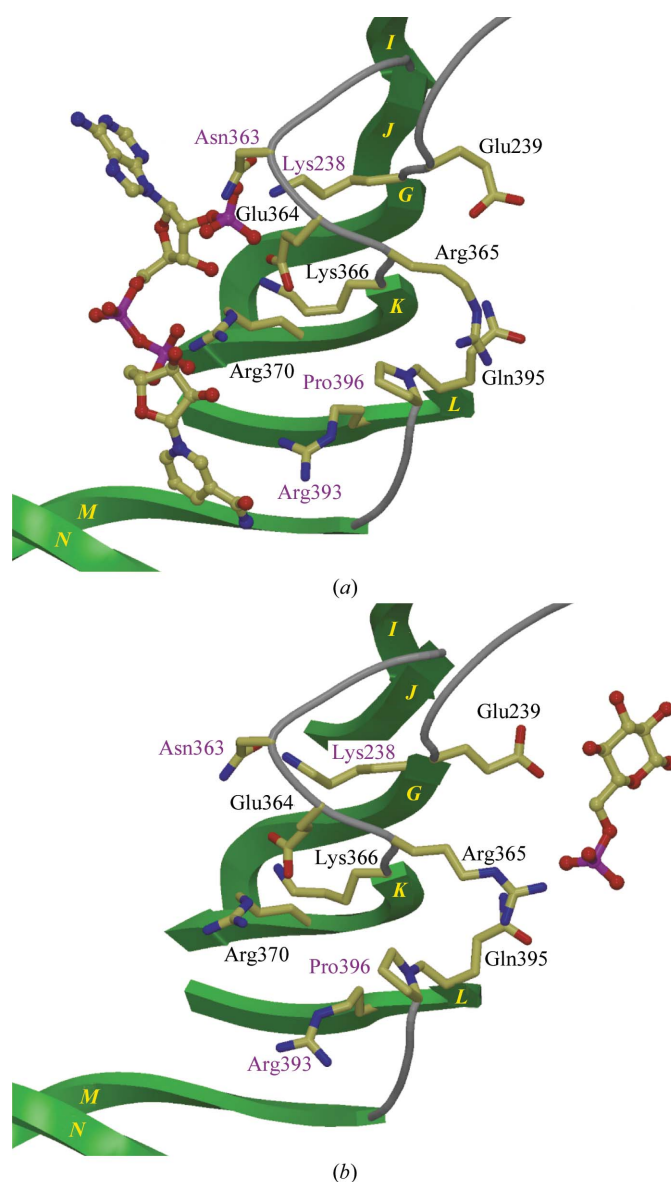


Figure 8
Residues bridging from the G6P site to the structural NADP⁺ site. The same view is used to allow comparison of this region in (a) Δ G6PD-G6P and (b) Δ G6PD-NADP⁺. Sheet strands are drawn in green and identified as in Fig. 1. Class I variants arise on mutation of those residues labelled in magenta.

results can be explained by reduction of all NADP⁺ to NADPH, the possibility of an interaction between bound G6P and the NADP⁺ site, lowering the affinity for NADP⁺ at the structural site, cannot be discounted. However, there is little movement of the side chains of the residues involved in the communication between G6P and structural NADP upon binding to the substrates. Measurements of the dissociation constant of the structural NADP⁺ in the wild-type enzyme and in certain mutants are under way (personal communication from P. C. Engel) and are to be followed by an investigation of NADPH. Any change of affinity of NADP⁺ or NADPH upon G6P binding will shed further light on the details of human G6PD stability and activity.

This study is funded by Hong Kong RGC grant Nos. HKU 7272/98M and HKU7243/00M to VL and MJA. We are grateful to Professor Dame Louise Johnson FRS for support and facilities in the Laboratory of Molecular Biophysics. We thank the support staff of stations 9.6 and 14.1 at CCLRC Daresbury Laboratory, Warrington, England for facilities and support. We thank Drs Elspeth Garman and Rick Lewis for their help in cryocrystallography techniques, Dr Pietro Roversi for his help with the *BUSTER-TNT* refinement program, and Dr Petra Lukacik and Mr James Murray for helpful discussions. MK was supported by a postdoctoral fellowship partly funded by the University of Hong Kong. MJA is the Dorothy Hodgkin–E. P. Abraham Fellow of Somerville College, Oxford and an associate member of the Oxford Centre for Molecular Sciences.

References

- Au, S. W. N., Gover, S., Lam, V. M. S. & Adams, M. J. (2000). *Structure*, **8**, 293–303.
- Au, S. W. N., Naylor, C. E., Gover, S., Vandeputte-Rutten, L., Scopes, D. A., Mason, P. J., Luzzatto, L., Lam, V. M. S. & Adams, M. J. (1999). *Acta Cryst. D* **55**, 826–834.
- Bacon, D. J. & Anderson, W. F. (1988). *J. Mol. Graph.* **6**, 219–220.
- Bautista, J. M., Mason, P. J. & Luzzatto, L. (1995). *FEBS Lett.* **366**, 61–64.
- Beutler, E. (1990). *Hematology*, edited by W. J. Williams, E. Beutler, A. J. Erslev & M. A. Lichtman, pp. 591–606. New York: McGraw-Hill.
- Beutler, E., Kuhl, W., Gelbart, T. & Forman, L. (1991). *J. Biol. Chem.* **266**, 4145–4150.
- Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S. M. & Bricogne, G. (2004). *Acta Cryst. D* **60**, 2210–2221.
- Bonsignore, A., Cancedda, R., Nicolini, A., Damiani, G. & DeFlora, A. (1971). *Arch. Biochem. Biophys.* **147**, 493–501.
- Brünger, A. T. (1992a). *Nature (London)*, **355**, 472–475.
- Brünger, A. T. (1992b). *X-PLOR. Version 3.1. A System for X-ray Crystallography and NMR*. New Haven, CT, USA: Yale University Press.
- Cohen, P. & Rosemeyer, M. A. (1969). *Eur. J. Biochem.* **8**, 8–15.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst. D* **50**, 760–763.
- Cosgrove, M. S., Gover, S., Naylor, C. E., Vandeputte-Rutten, L., Adams, M. J. & Levy, H. R. (2000). *Biochemistry*, **39**, 15002–15011.
- Cosgrove, M. S., Naylor, C. E., Paludan, S., Adams, M. J. & Levy, H. R. (1998). *Biochemistry*, **37**, 2759–2767.
- Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–134.
- Filosa, S., Fico, A., Paglialunga, F., Balestrieri, M., Crooke, A., Verde, P., Abrescia, P., Bautista, J. M. & Martini, G. (2003). *Biochem. J.* **370**, 935–943.
- Jones, T. A. & Kjeldgaard, M. (1997). *Methods Enzymol.* **277**, 173–208.
- Kleiger, G. & Eisenberg, D. (2002). *J. Mol. Biol.* **323**, 69–76.
- Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
- Leslie, A. G. W. (1999). *Acta Cryst. D* **55**, 1696–1702.
- Levy, H. R., Vought, V. E., Yin, X. & Adams, M. J. (1996). *Arch. Biochem. Biophys.* **326**, 145–151.
- Longo, L., Vanegas, O. C., Patel, M., Rosti, V., Li, H., Waka, J., Merghoub, T., Pandolfi, P. P., Notaro, R., Manova, K. & Luzzatto, L. (2002). *EMBO J.* **21**, 4229–4239.
- Mareni, C. & Gaetani, G. F. (1976). *Biochim. Biophys. Acta*, **430**, 395–398.
- Merritt, M. A. & Murphy, M. E. P. (1994). *Acta Cryst. D* **50**, 869–873.
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst. D* **53**, 240–255.
- Naylor, C. E., Gover, S., Basak, A. K., Cosgrove, M. S., Levy, H. R. & Adams, M. J. (2001). *Acta Cryst. D* **57**, 635–648.
- Read, R. J. (1986). *Acta Cryst. A* **42**, 140–149.
- Roos, D., van Zwieten, R., Wijnen, J. T., Gomez-Gallego, F., de Boer, M., Stevens, D., Pronk-Admiraal, C. J., de Rijk, T., van Noorden, C. J., Weening, R. S., Vulliamy, T. J., Ploem, J. E., Mason, P. J., Bautista, J. M., Khan, P. M. & Beutler, E. (1999). *Blood*, **94**, 2955–2962.
- Rowland, P., Basak, A. K., Gover, S., Levy, H. R. & Adams, M. J. (1994). *Structure*, **2**, 1073–1087.
- Scopes, D. A., Bautista, J. M., Naylor, C. E., Adams, M. J. & Mason, P. J. (1998). *Eur. J. Biochem.* **251**, 382–388.
- Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
- Vought, V., Ciccone, T., Davino, M. H., Fairbairn, L., Lin, Y., Cosgrove, M. S., Adams, M. J. & Levy, H. R. (2000). *Biochemistry*, **39**, 15012–15021.
- Wang, X. T., Au, S. W. N., Lam, V. M. S. & Engel, P. C. (2002). *Eur. J. Biochem.* **269**, 3417–3424.
- Wang, X. T., Lam, V. M. S. & Engel, P. C. (2005). In the press.
- Wrigley, N. G., Heather, J. V., Bonsignore, A. & DeFlora, A. (1972). *J. Mol. Biol.* **68**, 483–499.