Expression, crystallization and preliminary X-ray crystallographic analysis of glucose-6-phosphate dehydrogenase from the human pathogen *Trypanosoma cruzi* in complex with substrate

Cecilia Ortíz,[a] Nicole Larrieux,[b] Andrea Medeiros,[a,c] Horacio Botti,[b] Marcelo Comini[a]* and Alejandro Buschiazzo[b,d]*

[a]Redox Biology of Trypanosomes Laboratory, Institut Pasteur de Montevideo, Mataojo 2020, 11400 Montevideo, Uruguay, [b]Unit of Protein Crystallography, Institut Pasteur de Montevideo, Mataojo 2020, 11400 Montevideo, Uruguay, [c]Biochemistry Department, Universidad de la República, Avenida General Flores 2125, 11800 Montevideo, Uruguay, and [d]Department of Structural Biology and Chemistry, Institut Pasteur, 25 Rue du Dr Roux, 75015 Paris, France

Correspondence e-mail: mcomini@pasteur.edu.uy, alebus@pasteur.edu.uy

An N-terminally truncated version of the enzyme glucose-6-phosphate dehydrogenase from *Trypanosoma cruzi* lacking the first 37 residues was crystallized both in its apo form and in a binary complex with glucose 6-phosphate. The crystals both belonged to space group $P2_1$ and diffracted to 2.85 and 3.35 Å resolution, respectively. Self-rotation function maps were consistent with point group 222. The structure was solved by molecular replacement, confirming a tetrameric quaternary structure.

1. Introduction

The protozoan organisms of the family Trypanosomatidae encompass parasitic species that cause highly disabling and fatal diseases in humans and animals. *Trypanosoma cruzi* is a human pathogen that is responsible for Chagas disease, which has a prevalence of approximately 12–16 million cases in Latin America (Dias, 2009). New therapeutic approaches are urgently needed, given that the available treatments rely on drugs with severe side effects and typically low efficacy in the chronic stage of the disease. In this scenario, the identification and detailed characterization of essential biomolecules for the survival and/or virulence of the parasites, particularly those that are substantially different or even absent in the mammalian host, may contribute to improving the current situation. Several enzymes of the trypanosomatid pentose-phosphate pathway (PPP) have received increasing attention owing to their potential as drug targets (Igoillo-Esteve *et al.*, 2007). The first branch of the PPP involves the stepwise oxidation of glucose 6-phosphate (G6P) to ribulose 5-phosphate, with the concomitant generation of two NADPH molecules. The first reaction

$$NADP^+ + G6P \rightarrow \text{6-phosphogluconolactone} + NADPH$$

is catalyzed by glucose-6-phosphate dehydrogenase (G6PDH). The reducing power resulting from this reaction fuels anabolic pathways (lipid and nucleotide biosyntheses) and contributes to cellular redox homeostasis. In *T. cruzi*, G6PDH is encoded by three different sequences that share >97% amino-acid identity (Igoillo-Esteve & Cazzulo, 2006). Levels of G6PDH have been reported to be elevated in virulent strains and infective stages of the parasite, as well as in response to oxidative insult (Mancilla & Naquira, 1964; Igoillo-Esteve & Cazzulo, 2006; Mielniczki-Pereira *et al.*, 2007). This led to consideration of the enzyme as a virulence factor that plays an important role within the antioxidant defence systems of the parasite (Igoillo-Esteve & Cazzulo, 2006). More recently, the orthologue from *T. brucei brucei* (the causative agent of the cattle disease nagana) has been shown to be essential for the growth of infective parasites (Cordeiro *et al.*, 2009). Despite the high conservation of catalytic residues and substrate-binding motifs between the *T. cruzi* and human enzymes, they display only 46% overall amino-acid identity (Fig. 1). A detailed atomic study of the parasite protein is deemed to be essential to reveal regions that might be targets for selective inhibition. Although the three-dimensional structures of several PPP enzymes from trypanosomatids have been elucidated (Phillips *et al.*, 1998; Veitch *et al.*, 2004; Delarue *et al.*, 2007), to date no structural information is available for G6PDH from a protist. With the final aim of performing a thorough structural characterization of *T. cruzi*

crystallization communications
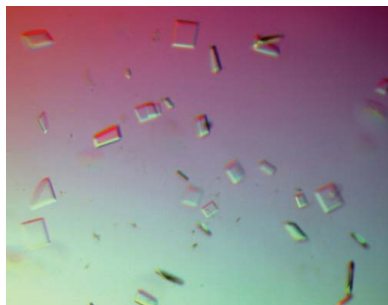
G6PDH, we report the heterologous expression, crystallization and X-ray diffraction data collection of an N-terminally truncated form of G6PDH from *T. cruzi*, which retains full catalytic activity, in both apo and G6P-bound forms.

2. Materials and methods

2.1. Expression and purification of His-tagged Δ37N G6PDH from *T. cruzi*

All of the chemicals and reagents used were of analytical grade and were purchased from Sigma unless stated otherwise. Attempts to obtain the recombinant full-length form of *T. cruzi* G6PDH (accession No. Q4E0B2) from vector pET28a (Igoillo-Esteve & Cazzulo, 2006) resulted in low yields and unstable enzyme, precluding its use in crystallization screening. Instead, a truncated form of TcG6PDH lacking the first 37 amino acids (TcG6PDHΔ37N; referred to in the following as Δ37N; Fig. 1) but displaying enzyme activity and kinetic parameters similar to those of the full-length form (Igoillo-Esteve & Cazzulo, 2006) was selected. In addition, Δ37N retains 94% of the overall sequence and overlaps the full length of the human enzyme, making it a suitable model for future structural studies (Fig. 1). Δ37N was expressed as an N-terminally 6×His-tagged protein using vector pET28a (the expression construct was a kind gift from Professor Cazzulo, Universidad de San Martín, Argentina; Igoillo-Esteve & Cazzulo, 2006) and *Escherichia coli* Tuner DE3 as a heterologous expression host. The transformed bacteria were grown in Terrific Broth medium with 10 g l^{-1} glucose and 50 mg l^{-1} kanamycin at 310 K and 200 rev min^{-1} agitation. When the optical density at 600 nm reached 0.8, the cultures were incubated at 277 K without shaking. After 30 min, $0.5 \text{ m}M$ isopropyl β-D-1-thiogalactopyranoside (Euromedex) was added to induce recombinant protein expression and growth of the culture was immediately resumed at 298 K and 180 rev min^{-1} for $\sim$16 h. The cells were harvested by centrifugation at 5000g for 10 min at 277 K. The pellets were suspended at a ratio of 1 g wet weight per 5 ml cold buffer A (50 mM Tris pH 8.0, 500 mM NaCl, 5 mM MgCl$_2$) with 0.3 g l^{-1} lysozyme, 150 nM pepstatin, 4 nM cystatin, 1 mM phenylmethanesulfonylfluoride and then incubated on ice for 60 min. Complete cell lysis was achieved by three cycles of sonication (30 pulses per minute at 55% power) using a microtip and a Digital Sonifier 450 (Branson). The cell lysate was centrifuged at 12 000g (15 min, 277 K) and the supernatant was further cleared of debris by centrifugation at 18 000g (45 min, 277 K) and filtration through a 0.7 µm cellulose acetate/cellulose nitrate mixed-ester filter (Sigma). This sample was applied onto a 1 ml HisTrap column (GE Healthcare) pre-equilibrated in buffer A. The column was washed with 20 ml buffer A followed by 20 ml buffer A containing 80 mM imidazole and the protein was eluted with buffer A containing 500 mM imidazole. The purification was performed at 277 K using a flow rate of 1.0 ml min^{-1}. The elution fractions containing active enzyme (see below) were pooled, concentrated by centrifugation on an Amicon Ultra-15 30K NMWL filter (Millipore; 4000g, 277 K) and purified by gel-filtration chromatography on Superdex G200 (Tricorn 10/300 column, GE Healthcare) equilibrated in buffer A. The chromatography was performed at room temperature using an ÄKTA-purifier system (GE Healthcare) and a flow rate of 0.5–0.7 ml min^{-1}. Elution samples showing the highest purity of Δ37N, as assessed by Coomassie Blue-stained SDS–PAGE gels, were pooled, tested for enzyme specific activity and stored at 277 K until further use. The enzyme activity was measured spectrophotometrically at 340 nm and 298 K in 50 mM Tris pH 7.5, 5 mM MgCl$_2$ buffer using 500 µM NADP and 1 mM glucose 6-phosphate (G6P). The protein concentration was determined using the bicinchoninic acid assay (Sigma). The final yield of pure recombinant protein was estimated to be $\sim$1.7 mg per litre of culture medium.

2.2. Crystallization

Prior to crystallization, buffer-exchange of purified Δ37N was performed using an Amicon Ultra-15 30K NMWL filter (Millipore). Briefly, 0.5 ml protein solution was diluted to 5 ml with 20 mM Tris–HCl pH 8.0 and reconcentrated three times by centrifugation at 4000g and 277 K. To obtain sample containing substrate, the final concentration step was performed in the presence of 20 mM Tris–HCl pH 8.0 supplemented with 5 mM G6P. The proteins were concentrated by centrifugation (15 000g and 277 K) with a 10 kDa cutoff IVSS Vivaspin 500 filter device (Sigma). Initial crystallization screenings were performed by the sitting-drop vapour-diffusion method with Crystal Screen and Crystal Screen 2 (Hampton Research) using a robotic pipetting station (Honeybee 963, Digilab). Drops consisting
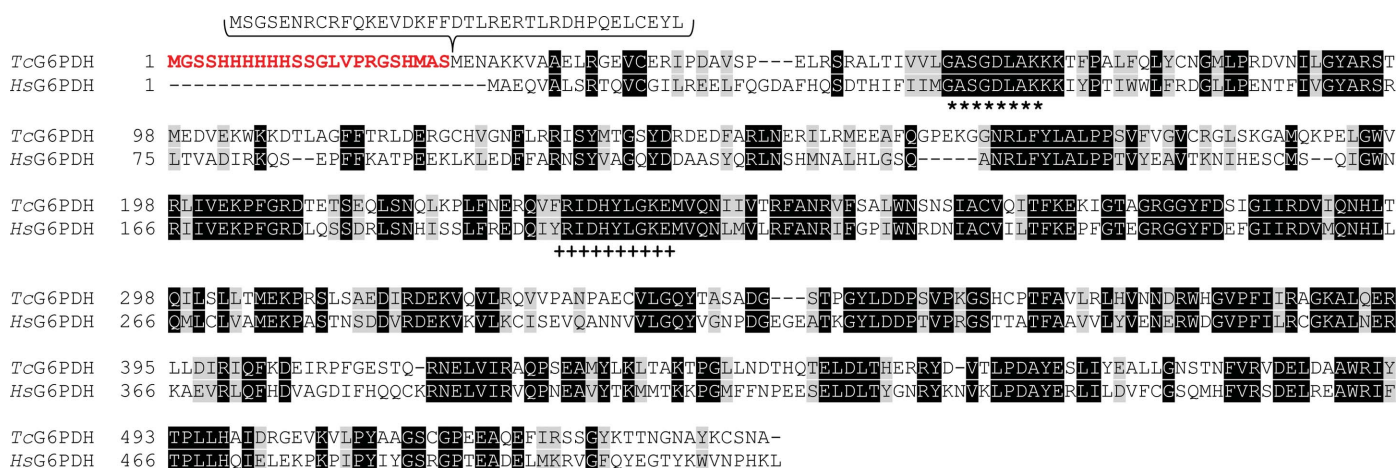

Figure 1
Amino-acid sequence of the G6PDH from *T. cruzi* (TcG6PDH, accession No. Q4E0B2) used for crystallization. The N-terminal extension, which is absent in the truncated mutant Δ37N, is shown as an insertion fragment at the top of the figure. The N-terminal stretch incorporated as a fusion from the expression vector is indicated in bold red. Identical residues with respect to the human enzyme (accession No. P11413) are shown on a black background, while conservative substitutions are highlighted with a grey background. The cofactor-binding and substrate-binding sites are indicated by asterisks and crosses, respectively.

Table 1
Data-collection statistics for Δ37N G6PDH.

Values in parentheses are for the highest resolution shell.

	Apoprotein	G6P–protein complex
No. of crystals	1	1
Crystal dimensions (μm)	$50 \times 50 \times 200$	$50 \times 50 \times 100$
Beamline	ALS 5.0.2	ALS 5.0.2
Wavelength (Å)	1.00	1.00
Detector	ADSC Quantum 315	ADSC Quantum 315
Crystal-to-detector distance (mm)	390	385
Rotation range per image (°)	0.5	0.5
Total rotation range (°)	170.5	167.5
Exposure time per image (s)	0.5	0.5
Resolution range (Å)	30.00–2.85 (3.00–2.85)	67.36–3.35 (3.53–3.35)
Space group	$P2_1$	$P2_1$
Unit-cell parameters (Å, °)	$a = 96.6$, $b = 133.0$, $c = 107.8$, $\beta = 100.1$	$a = 96.8$, $b = 133.0$, $c = 107.7$, $\beta = 100.3$
Mosaicity (°)	0.94	0.88
Measured reflections	214946	134881
Unique reflections	62028	38508
Multiplicity	3.5 (3.4)	3.5 (3.6)
Mean $I/\sigma(I)$	7.1 (2.3)	5.5 (2.9)
Completeness (%)	99.1 (99.2)	99.5 (99.9)
R_{merge}† (%)	0.138 (0.501)	0.209 (0.450)
R_{meas}† (%)	0.163 (0.594)	0.246 (0.530)
Matthews coefficient‡ ($Å^3 Da^{-1}$)	2.79	2.80
Solvent content‡ (%)	56.0	56.1

† $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ and $R_{meas} = \sum_{hkl} [N/(N-1)]^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $\langle I(hkl) \rangle$ is the average of symmetry-related (or Friedel-related) observations of a unique reflection. ‡ These figures correspond to the inclusion of four molecules in the asymmetric unit. In principle, a Matthews coefficient of 2.24 $Å^3 Da^{-1}$ (45% solvent) would be compatible with five protomers. Structure determination confirmed that a tetramer is the correct choice (see text).

of 0.4 μl protein solution mixed with 0.4 μl reservoir solution were dispensed into the wells of a CrystalQuick 96-well plate (plate model No. 609820, Greiner) containing 90 μl reservoir solution. After 4 d, crystals developed under Crystal Screen condition No. 39. Crystallization conditions were optimized manually using a hanging-drop vapour-diffusion setup, testing the following variables: protein and precipitant concentration, pH, additives and seeding. Optimal crystallization of the apo form (20 mg ml^{-1}) was achieved in the presence of 6% PEG 400, 1.6 M ammonium sulfate, 0.1 M Na HEPES pH 7.5. In the presence of 5 mM G6P, the crystallization conditions were adjusted to 33 mg ml^{-1} protein, 4% PEG 400, 1.8 M ammonium sulfate, 0.1 M Na HEPES pH 7.5. Optimizations were carried out using 4 μl drops (2 μl protein with or without substrate plus 2 μl reservoir solution) in VDX 24-well crystallization plates (Hampton Research) filled with 1 ml reservoir solution per well. All crystallization assays were performed at 293 K.

2.3. X-ray diffraction data collection and processing and structure determination

The crystals were mounted in nylon loops of 0.1–0.3 μm and flash-cooled in liquid nitrogen after addition of 22%(v/v) glycerol to the mother liquor. In the case of the G6P complex the PEG 400 concentration was raised to 6% during cryopreservation. Data collection was performed remotely using the robotic facilities at beamline 5.0.2 of the Advanced Light Source synchrotron (Lawrence Berkeley National Laboratory). Diffraction data were indexed and integrated with *MOSFLM* v.7.0.7/*iMOSFLM* v.1.0.5 (Battye *et al.*, 2011) and then scaled and reduced with *SCALA*/*TRUNCATE* within the *CCP*4 suite (Winn *et al.*, 2011; see Table 1 for full statistics). A monomer of human G6PDH (PDB entry 1qki) was used as a search probe for molecular replacement with the program *AMoRe* (Navaza, 1994).

3. Results and discussion

A truncated mutant of *T. cruzi* G6PDH lacking the nonconserved N-terminal end (the first 37 residues), which had previously been subjected to biochemical characterization (Igoillo-Esteve & Cazzulo, 2006), was expressed in a soluble form in *E. coli* and purified to homogeneity using immobilized metal-affinity and size-exclusion chromatography. The active enzyme displayed a molecular weight of ∼260 kDa in solution, which is consistent with the protein being a tetramer (protomer molecular weight of 61 kDa; Fig. 2). Screening of 96 crystallogenesis conditions using recombinant protein alone or attempting cocrystallization in complex with the substrate G6P yielded positive hits that were further optimized. Crystallization was manually optimized using the hanging-drop vapour-diffusion method. X-ray diffraction analysis was performed on Δ37N single crystals, which grew in a few days under the following conditions: (i) 6% PEG 400, 1.6 M ammonium sulfate, 0.1 M Na HEPES pH 7.50 and (ii) 4% PEG 400, 1.8 M ammonium sulfate, 0.1 M Na HEPES pH 7.5, 5 mM G6P (Fig. 3). Data sets for these crystals were collected to 2.85 and 3.35 Å resolution, respectively, using synchrotron radiation (Fig. 4a). The crystals both belonged to space group $P2_1$; they proved to be isomorphous and appeared to contain four or five molecules in the asymmetric unit (Table 1). A self-rotation function (SRF) was calculated using the scaled amplitudes (Fig. 4b), allowing the identification of three orthogonal twofold noncrystallographic (NCS) axes, compatible with a tetramer possessing 222 (D2) point-group symmetry. The NCS dyad located at $(\theta, \varphi, \chi) = (54, 180, 180°)$ in polar coordinates (peak 1 in Fig. 4b) is stronger than the other two peaks, probably because it is perpendicular to the unique crystallographic axis, which doubles the number of Patterson vectors contributing to the peak. The two weaker peaks (peaks 2 and 3) are both orthogonal to the first peak. None of the three NCS axes are parallel to the monoclinic dyad, which is consistent with the absence of peaks on the
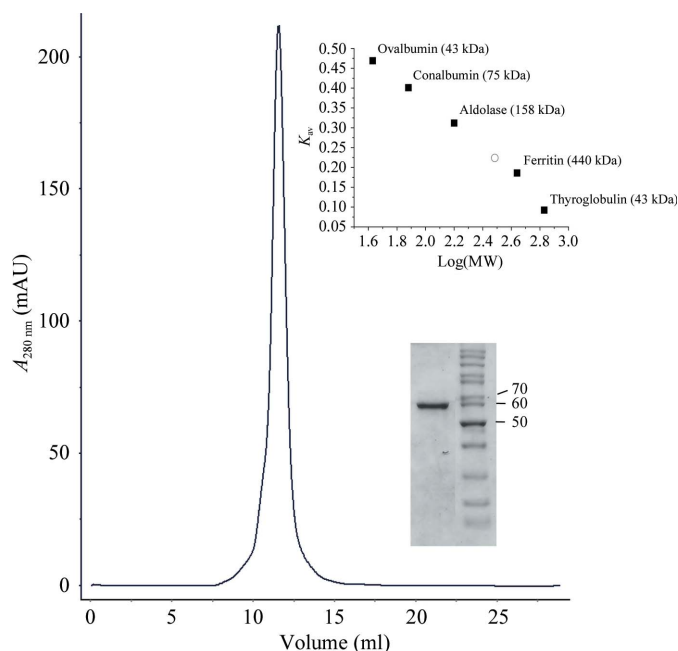
Figure 2
Purification and oligomeric state analysis of *T. cruzi* Δ37N G6PDH. Isocratic elution chromatogram of the recombinant protein separated by size-exclusion chromatography (SEC) on a Superdex G200 column. Bottom inset, 12% SDS–PAGE of the SEC peak fraction stained with Coomassie Blue; the right lane contains molecular-weight markers labelled in kDa. Top inset, plot employed to determine the apparent molecular mass of Δ37N (filled squares, molecular-weight standards; circle, recombinant TcG6PDHΔ37N).

crystallization communications

$v = 0.5$ section of the native Patterson map. There is nevertheless a non-origin peak in these Patterson maps ($u = 0.13$, $v = 0$, $w = 0.13$) displaying ∼20% of the origin peak signal. Typically, such a peak may reveal pseudo-translational symmetry, *e.g.* derived from near-perfect alignment of different protomers in the asymmetric unit. However, in this case it is very likely to be the consequence of a crystal-packing defect (see below), with no simple interpretation, and to be related in some way to the almost continuous streak of twofold peaks that connect NCS peak 1 with the crystallographic *y*-axis poles in the SRF section (Fig. 4*b*). Different crystals showed variable heights of these streaking peaks, which sometimes even became undetectable over the noise. The crystals that we report here represent the best crystals in

terms of the resolution limit and were hence selected for further work. In order to confirm the preliminary analyses, we solved the structures by molecular replacement (Navaza, 1994), initially using the apo data set with one monomer of the human orthologue (PDB entry 1qki, chain *A*) as the search probe (the solution of the iso-morphous G6P complex data set was straightforward starting from the apo crystal phases and difference Fourier electron-density maps). Although refinements are still in progress, with an R_{free} of ∼25% we can already safely confirm the space-group assignment, the presence of four monomers in the asymmetric unit and the correctness of the essential features of the SRF analysis presented above. In particular, NCS peak 1 corresponds to the axis that relates one dimer to the
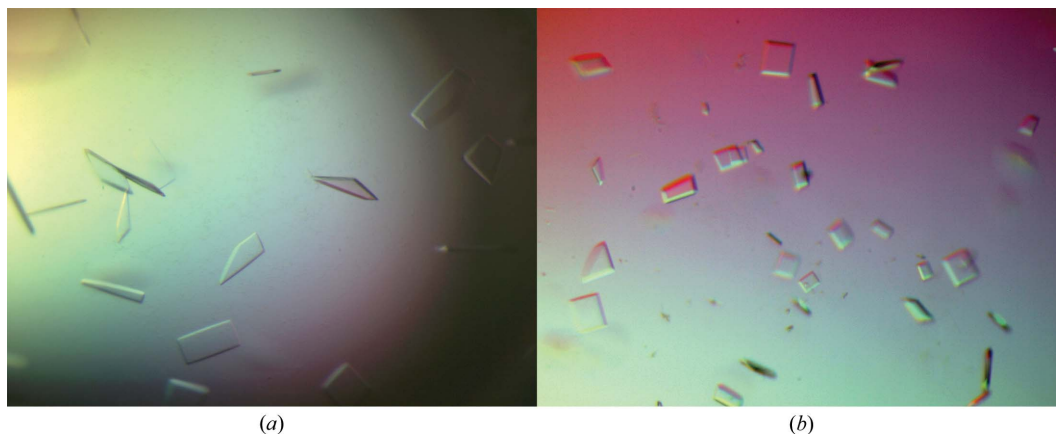

Figure 3
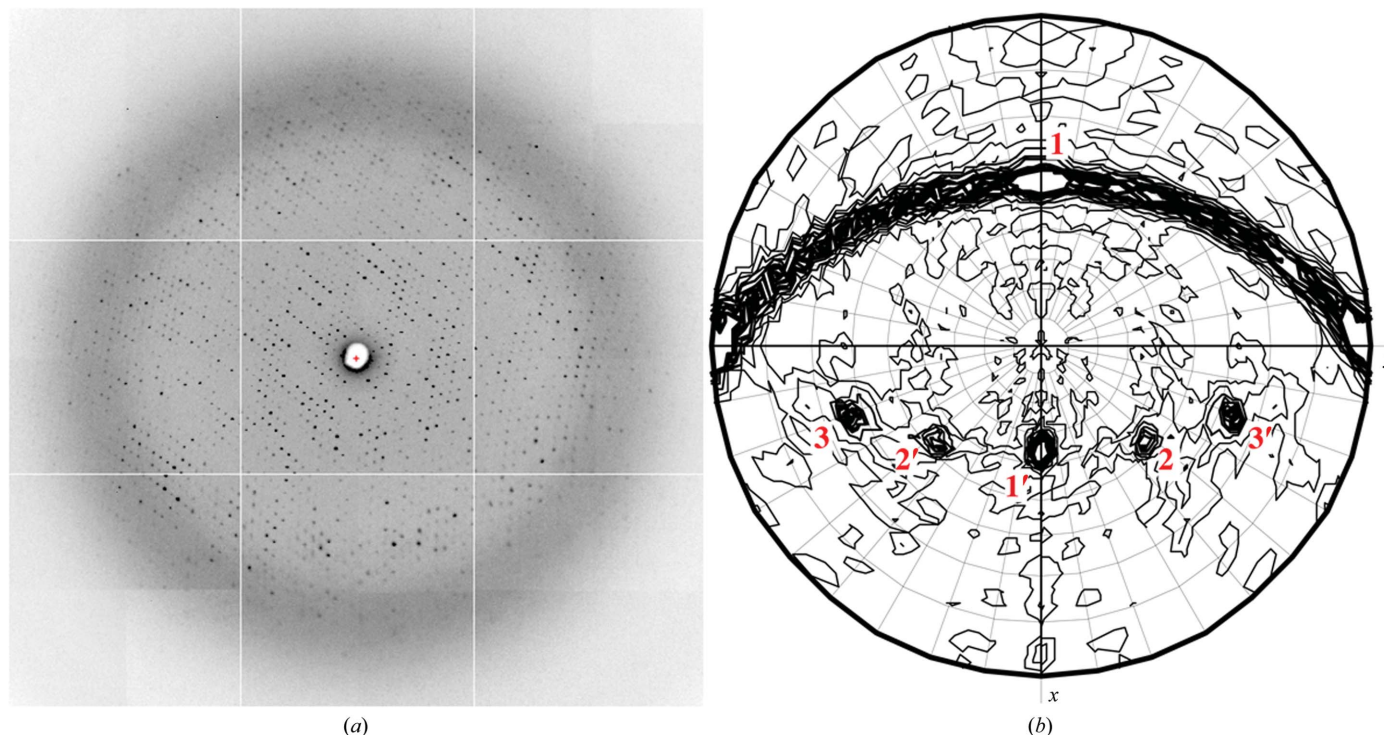Crystals of *T. cruzi* Δ37N G6PDH: (*a*) apo form, (*b*) cocrystallized in the presence of G6P.

Figure 4
X-ray diffraction of *T. cruzi* Δ37N G6PDH apo-form crystals. (*a*) A typical frame picked out from the data set, showing diffraction to 2.85 Å resolution. (*b*) $\chi = 180°$ section of the self-rotation function calculated with observed structure-factor amplitudes using *MOLREP* (Vagin & Teplyakov, 2010). The unique crystallographic axis is parallel to *y* [$I/\sigma(I) = 31.8$]. The major [$I/\sigma(I) = 26.9$] noncrystallographic symmetry (NCS) twofold axis [marked as peak 1 in the stereographic projection, polar coordinates (53.8, 180, 180°)], corresponds to the proper 180° rotation that relates one dimer to the other in the asymmetric unit. This axis is orthogonal to the crystallographic axis, as well as to the weaker NCS dyads marked as peaks 2 [(47.2, 47.3, 180°), $I/\sigma(I) = 7.7$] and 3 [(64.1, 290.7, 180°), $I/\sigma(I) = 7.5$] which relate one monomer to the other within each dimer and to the non-equivalent monomer of the other dimer, respectively. 1′, 2′ and 3′ correspond to the symmetrically equivalent NCS axes. Note the streak of twofold peaks observed connecting 1 to the *y*-axis poles.

other, with an architecture that is similar overall to that seen in the human 'dimer of dimers' tetramer (Au *et al.*, 2000), in which each dimer has much larger inter-monomeric protein–protein interactions compared with the fewer contacts between dimers. NCS peaks 2 and 3 correspond to the axes relating one monomer to the other within and between dimers, respectively. Recomputation of the SRF with the calculated amplitudes results in clearer sections (no signals from the streak of NCS dyads) and elimination of the non-origin peak in the native Patterson maps.

Clear electron density for bound glucose 6-phosphate in the second crystal constitutes further strong evidence in support of the determined structure solution. Full refinement of the structures reported here will ultimately allow full analysis of the quaternary-structure features and the potential effects of bound ligands in terms of internal symmetry.

References

Au, S. W., Gover, S., Lam, V. M. & Adams, M. J. (2000). *Structure*, **8**, 293–303.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Cordeiro, A. T., Thiemann, O. H. & Michels, P. A. (2009). *Bioorg. Med. Chem.* **17**, 2483–2489.
Delarue, M., Duclert-Savatier, N., Miclet, E., Haouz, A., Giganti, D., Ouazzani, J., Lopez, P., Nilges, M. & Stoven, V. (2007). *J. Mol. Biol.* **366**, 868–881.
Dias, J. C. P. (2009). *Mem. Inst. Oswaldo Cruz*, **104**, Suppl. 1, 41–45.
Igoillo-Esteve, M. & Cazzulo, J. J. (2006). *Mol. Biochem. Parasitol.* **149**, 170–181.
Igoillo-Esteve, M., Maugeri, D., Stern, A. L., Beluardi, P. & Cazzulo, J. J. (2007). *Ann. Acad. Bras. Cienc.* **79**, 649–663.
Mancilla, R. & Naquira, C. (1964). *J. Protozool.* **11**, 509–513.
Mielniczki-Pereira, A. A., Chiavegatto, C. M., López, J. A., Colli, W., Alves, M. J. & Gadelha, F. R. (2007). *Acta Trop.* **101**, 54–60.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Phillips, C., Dohnalek, J., Gover, S., Barrett, M. P. & Adams, M. J. (1998). *J. Mol. Biol.* **282**, 667–681.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Veitch, N. J., Maugeri, D. A., Cazzulo, J. J., Lindqvist, Y. & Barrett, M. P. (2004). *Biochem. J.* **382**, 759–767.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.