STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Structures of prostaglandin F synthase from the protozoa *Leishmania major* and *Trypanosoma cruzi* with NADP

Spencer O. Moen,[a,b] James W. Fairman,[a,b] Steve R. Barnes,[a,b] Amy Sullivan,[a,b] Stephen Nakazawa-Hewitt,[a,c] Wesley C. Van Voorhis,[a,d] Bart L. Staker,[a,e] Donald D. Lorimer,[a,b]* Peter J. Myler[a,e,f] and Thomas E. Edwards[a,b]

[a]Seattle Structural Genomics Center for Infectious Disease, USA, [b]Beryllium, Bainbridge Island, WA 98110, USA, [c]CERID, University of Washington, Seattle, WA 98109, USA, [d]Department of Medicine, Division of Allergy and Infectious Diseases, University of Washington, Seattle, WA 98195, USA, [e]Seattle Biomedical Research Institute, Seattle, WA 98109, USA, and [f]Department of Global Health and Medical Education and Biomedical Bioinformatics, University of Washington, Seattle, WA 98109, USA. *Correspondence e-mail: dlorimer@be4.com

Received 15 December 2014
Accepted 6 April 2015

Edited by W. N. Hunter, University of Dundee, Scotland

Keywords: prostaglandin; infectious diseases; leishmaniasis; Chagas disease; SSGCID.

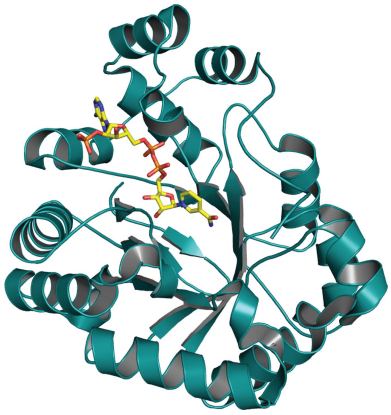
PDB references: *T. cruzi* prostaglandin F synthase, apo, 4fzi; complex with NADP, 4gie; *L. major* prostaglandin F synthase, apo, 4f40; complex with NADPH, 4g5d

Supporting information: this article has supporting information at journals.iucr.org/f

The crystal structures of prostaglandin F synthase (PGF) from both *Leishmania major* and *Trypanosoma cruzi* with and without their cofactor NADP have been determined to resolutions of 2.6 Å for *T. cruzi* PGF, 1.25 Å for *T. cruzi* PGF with NADP, 1.6 Å for *L. major* PGF and 1.8 Å for *L. major* PGF with NADP. These structures were determined by molecular replacement to a final *R* factor of less than 18.6% (R_{free} of less than 22.9%). PGF in the infectious protozoa *L. major* and *T. cruzi* is a potential therapeutic target.

1. Introduction

The Seattle Structural Genomics Center for Infectious Disease (SSGCID) is a consortium funded by NIAID to elucidate solutions of protein structures from biodefense organisms, as well as those causing emerging and re-emerging diseases. Prostaglandin F synthase (PGF) from the infectious protozoa *Leishmania major* (UniProt P22045) and *Trypanosoma cruzi* (UniProt Q4DJ07) have been identified by the TDR target database as being essential proteins for both organisms and potential therapeutic targets (Magariños *et al.*, 2012). PGF is an oxidoreductase enzyme that catalyzes the reaction (5*Z*,13*E*)-(15*S*)-9α,11α,15-trihydroxyprosta-5,13-dienoate + $NADP^+ \rightleftharpoons$ (5*Z*,13*E*)-(15*S*)-9α,15-dihydroxy-11-oxoprosta-5,13-dienoate + NADPH + H^+. PGF specifically acts on the CH–OH of the proton donor with NADP as the acceptor (Watanabe *et al.*, 1981). In humans, PGF (UniProt P42330) can interconvert active androgens, estrogens and progestins with their cognate inactive metabolites (Qin *et al.*, 1993). In protozoa, PGF is involved in essential lipid-metabolism pathways. The protozoa *L. major*, the causative agent of leishmaniasis, and *T. cruzi*, the causative agent of Chagas disease, both affect millions of people and represent major public health issues. Both diseases have very limited treatment options and drug resistance is prevalent (Minodier & Parola, 2007; Buckner *et al.*, 1998). Here, we describe the structures of PGF with and without NADP in both organisms and the structural differences between the bound and the unbound state of the enzyme with regard to its cofactor. Previous structural work on PGFs has been performed for *Homo sapiens* PGF (PDB entry 1ry8; Komoto *et al.*, 2004) and



© 2015 International Union of Crystallography

Table 1
Crystallization.

Sample	PDB entry	Protein concentation (mg ml^{-1})	Cofactor concentration (m*M*)	Crystallization screen condition	Reservoir composition
TrcrA.00019.a.B1, apo	4fzi	37	—	Morpheus H1	0.1 *M* MES–imidazole pH 6.5, 10% PEG 20 000, 20% PEG 550 MME, 0.02 *M* glutamic acid, glycine, serine, alanine and lysine
TrcrA.00019.a.B1 + NADP	4gie	37	4	Index G6	0.2 *M* ammonium acetate, 0.1 *M* bis-tris pH 5.5, 25% PEG 3350; cryoprotectant 15% ethylene glycol
LemaA.00019.a.B1, apo	4f40	25	—	JCSG+ A2	0.1 *M* sodium citrate pH 5.50, 20% PEG 3000
LemaA.00019.a.B1 + NADPH	4g5d	25	2	JCSG+ B8	200 m*M* magnesium chloride, 100 m*M* Tris pH 7.0, 10% PEG 8000

PGF from *T. brucei*, another protozoan (PDB entry 1vbj; T. Inoue, unpublished work). PDB entry 1ry8 is a useful structure for determining the selectivity of potential compounds against human PGF, which will need to be taken into consideration (Komoto *et al.*, 2004). PDB entry 1vbj was used as a model for molecular replacement and has sequence identities of 60% to *T. cruzi* PGF and 61% to *L. major* PGF (Fig. 1).

2. Materials and methods

2.1. Cloning, expression and purification

The gene encoding PGF from *L. major* strain Friedlin (SSGCID ID LemaA.00019.a.B1) was PCR-amplified from genomic DNA that was kindly provided by Frederick S. Buckner. The gene was amplified using the following primer sequences: FWD primer 5′-CTCACCACCACCACCACCATATGGCTGGCGTTGATAAGGCAAT-3′ and REV primer 5′-ATCCTATCTTACTCACTTAGAACTGCGCCTCATCAGGGTC-3′. Thermal cycling conditions were 371 K for 180 s followed by 34 cycles of 371 K for 30 s, 346 K for 330 s, 360 K for 360 s and 360 K for 360 s followed by a final extension at 351 K for 300 s.

Likewise, the gene encoding PGF from *T. cruzi* strain CL Brener (SSGCID ID TrcrA.00019.a.B1) was PCR-amplified using the following primer sequences: FWD primer 5′-CTCACCACCACCACCACCATATGAATTGCAATTACAACTGTGTGAC-3′ and REV primer 5′-ATCCTATCTTACTCACTTACTCCTCTCCACCAGGGAAAAAAT-3′. Thermal cycling conditions were 367 K for 180 s followed by 30 cycles of 367 K for 30 s, 333 K for 60 s and 345 K for 120 s followed by a final extension at 345 K for 600 s.

Purified PCR products from both PGF constructs were cloned into a BG1861 expression vector (pET-14b derivative)

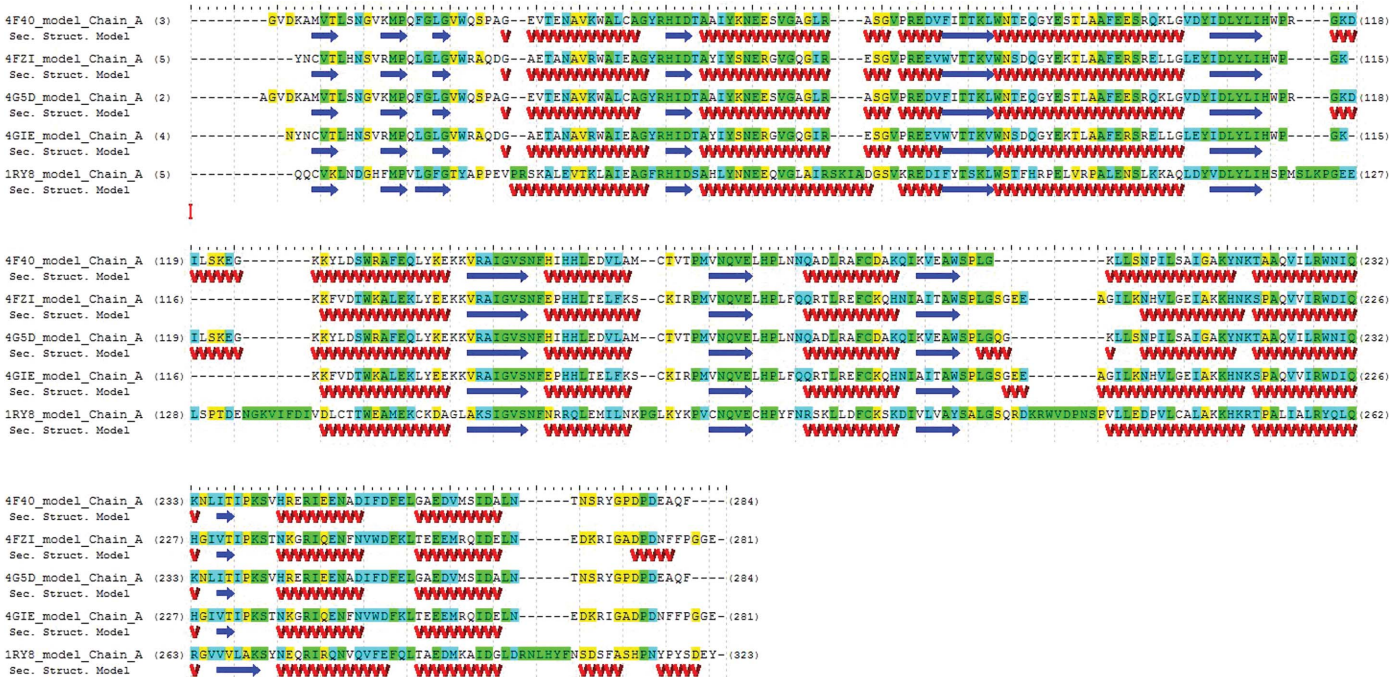


Figure 1
An alignment of PGFs from *H. sapiens* (PDB entry 1ry8), *T. cruzi* (PDB entries 4fzi and 4gie) and *L. major* (PDB entries 4f40 and 4g5d) with residues that match highlighted in green, residues that are conserved highlighted in blue and residues that are similar highlighted in yellow; residues that are different are not highlighted. The two protozoan proteins have 74% sequence identity to each other. This figure was generated using *GeneComposer* (Lorimer *et al.*, 2011).

Table 2
Data collection and processing.

Values in parentheses are for the outer shell.

	TrcrA.00019.a.B1		LemaA.00019.a.B1	
Target ID	Apo	NADPH	Apo	NADPH
PDB code	4fzi	4gie	4f40	4g5d
Space group	$P2_1$	$C2$	$P2_1$	$P2_1$
a, b, c (Å)	54.46, 66.62, 87.31	155.54, 50.34, 37.60	94.55, 34.59, 107.20	94.66, 34.59, 106.95
α, β, γ (°)	90, 94.23, 90	90, 94.23, 90	90, 103.22, 90	90, 102.98, 90
Beamline	5.0.3, ALS	5.0.1, ALS	5.0.1, ALS	Rigaku FR-E + SuperBright
Wavelength (Å)	0.9765	0.9774	0.9774	1.54
Data-collection temperature (K)	100	100	100	100
Resolution range (Å)	50.0–2.60 (2.67–2.60)	50.0–1.25 (1.28–1.25)	50.0–1.60 (1.64–1.60)	50.0–1.80 (1.85–1.80)
Unique reflections	19336 (1416)	75887 (4767)	89147 (5975)	61422 (4020)
Completeness (%)	99.7 (99.5)	94.7 (81.0)	98.7 (90.2)	96.5 (86.7)
Multiplicity	4.82 (4.95)	7.61 (7.11)	4.06 (3.52)	5.08 (3.30)
$\langle I/\sigma(I)\rangle$	17.75 (3.35)	23.41 (5.29)	13.33 (2.37)	13.36 (2.62)
$R_{\rm merge}$	0.080 (0.529)	0.062 (0.369)	0.083 (0.578)	0.096 (0.515)

Table 3
Structure solution and refinement.

Values in parentheses are for the outer shell.

PDB code	4fzi	4gie	4f40	4g5d
Resolution range (Å)	50.0–2.60 (2.70–2.60)	29.9–1.25 (1.28–1.25)	46.2–1.60 (1.64–1.60)	46.2–1.80 (1.85–1.80)
Completeness (%)	99.7 (99.5)	94.7 (81.0)	98.7 (90.2)	96.5 (86.7)
σ Cutoff	$F > 0.0\sigma(F)$	$F > 0.0\sigma(F)$	$F > 0.0\sigma(F)$	$F > 0.0\sigma(F)$
No. of reflections, working set	19335 (1298)	75882 (4543)	89114 (2444)	61405 (2307)
No. of reflections, test set	988 (67)	3826 (218)	4466 (135)	3097 (131)
Final $R_{\rm cryst}$	0.186 (0.267)	0.119 (0.170)	0.154 (0.2151)	0.172 (0.2219)
Final $R_{\rm free}$	0.229 (0.354)	0.136 (0.184)	0.180 (0.2228)	0.218 (0.2890)
No. of non-H atoms				
Protein	4252	2309	4366	4402
Ion	5	4	1	0
Ligand	0	48	98	96
Water	84	413	674	860
Total	4361	2774	5139	5358
R.m.s. deviations				
Bonds (Å)	0.013	0.007	0.014	0.009
Angles (°)	1.497	1.448	1.540	1.332
Average B factors (Å^2)				
Wilson B	39.8	12.3	19.7	17.6
Protein	36.1	7.4	13.1	12.1
Ion	65.6	5.8	49.9	—
Ligand	—	4.3	27.4	9.9
Water	25.6	20.6	26.8	27.8
Ramachandran plot				
Most favoured (%)	95.74	97.43	98.33	97.60
Allowed (%)	4.07	2.57	1.30	1.99
Asymmetric unit content	2 chains	1 chain	2 chains	2 chains
B-factor refinement	Isotropic	Anisotropic	Isotropic	Isotropic
TLS refinement	Yes, two groups	No	Yes, ten groups	Yes, ten groups

using ligation-independent cloning (LIC; Aslanidis & de Jong, 1990). The BG1861 vector contains a T7 promoter, an ampicillin-resistance gene and an N-terminal His_6 tag. Plasmid DNA was transformed into *Escherichia coli* strain BL21 (DE3) for recombinant protein expression.

2 l cultures were grown for both LemaA.00019.a.B1 and TrcrA.00019.a.B1 in a LEX bioreactor (Harbinger) at 293 K in Novagen Overnight Express autoinduction medium. Following 72 h of growth, the cultures were harvested *via* centrifugation at 5000*g* for 30 min. The supernatant was discarded and the cell pellets were collected and flash-frozen in liquid nitrogen.

Cell paste for TrcrA.00019.a.B1 and LemaA.00019.a.B1 was resuspended in buffer *A* [25 m*M* HEPES pH 7.0, 500 m*M* NaCl, 5%(*v*/*v*) glycerol, 30 m*M* imidazole, 0.5% CHAPS, 10 m*M* $MgCl_2$, 1 m*M* TCEP, 0.01 mg ml^{-1} lysozyme, 125 U benzonase, 1× EDTA-free protease-inhibitor cocktail from Roche] by stirring at 277 K for 1 h. The cell suspension underwent cell lysis by sonication on ice for 30 min (100 W; cycles of 15 s pulse on and 15 s pulse off). The lysed samples were then clarified *via* centrifugation at 18 000*g* for 1 h at 277 K. The supernatant was collected and applied onto a 5 ml HisTrap HP column (GE Healthcare) that had been pre-equilibrated with buffer *B* [25 m*M* HEPES pH 7.0, 500 m*M*

NaCl, 5%(*v*/*v*) glycerol, 30 mM imidazole, 1 m*M* TCEP] using an ÄKTAexplorer (GE Healthcare). The column was then washed with buffer *B* until the A_{280} stabilized at background levels. Following the wash step, buffer *C* [25 m*M* HEPES pH 7.0, 500 m*M* NaCl, 5%(*v*/*v*) glycerol, 500 m*M* imidazole, 1 m*M* TCEP] was applied to elute the target protein from the column. Fractions from nickel-affinity chromatography were analyzed by SDS–PAGE and pooled. Each target was then subjected to further purification by size-exclusion chromatography (SEC) on a Superdex 75 (GE Healthcare) SEC column using buffer *D* [25 m*M* Tris pH 8.0, 200 m*M* NaCl, 1%(*v*/*v*) glycerol, 1 m*M* TCEP]. Fractions were analyzed *via* SDS–PAGE to determine fractions of interest. Fractions of interest were collected and concentrated to 20 mg ml^{-1} using a polyethersulfone concentrator with an appropriate molecular-weight cutoff. Protein concentrations were determined with a UV spectrophotometer using the theoretical extinction coefficients determined using the online tool at http://web.expasy.org/protparam/. The final samples were divided into 100 µl aliquots and flash-frozen in liquid nitrogen.

2.2. Protein crystallization

All protein crystallization experiments were performed using the sitting-drop vapor-diffusion method in Compact 300 (Rigaku Reagents) crystallization trays at 289 K. Crystallization experiments consisted of 400 nl protein solution and 400 nl precipitant solution in the sample well equilibrated against 80 µl precipitant solution in the reservoir well. Detailed crystallization conditions including protein concentrations and cofactor concentrations are available in Table 1. All crystals formed in 2–4 weeks.

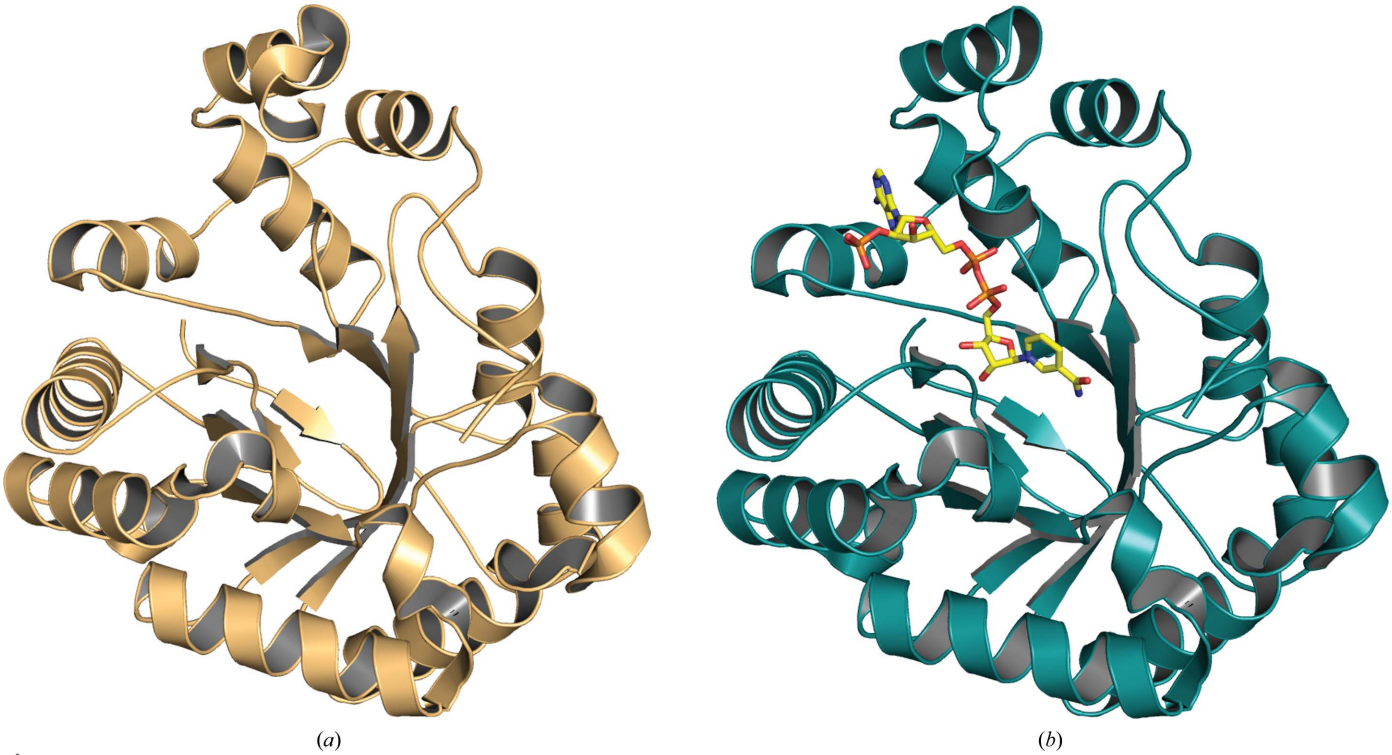
The crystals were harvested using mounted cryoloops (Hampton Research) and flash-cooled in pucks immersed in liquid nitrogen for storage until X-ray diffraction data collection. Data were either collected on our in-house FR-E+ SuperBright X-ray source (Rigaku) or on Advanced Light Source (ALS) beamlines 5.0.1 and 5.0.3. All X-ray reflection data were indexed, integrated and scaled using *XDS* and *XSCALE* (Kabsch, 2010*a*,*b*). Data statistics for all data sets are available in Table 2.

2.3. Data collection and processing

The diffraction data sets for PDB entries 4gie and 4f40 were collected on ALS beamline 5.0.1 with an ADSC Quantum 315 CCD detector. The diffraction data set for PDB entry 4fzi was collected on ALS beamline 5.0.3 with an ADSC Quantum 315 CCD detector. The diffraction data set for PDB entry 4g5d was collected in-house with a Rigaku FR-E+ SuperBright generator using a Rigaku 944+ CCD detector. All data sets were collected at 100 K. The X-ray data were reduced with *XDS* and *XSCALE*. Details of the data collections are summarized in Table 2.

2.4. Structure solution and refinement

Phases for structure determination were obtained *via* molecular replacement with *Phaser* (McCoy *et al.*, 2007) using PDB entry 1vbj as a search model for PDB entry 4f40. After preparing the starting models with *CHAINSAW* (Stein, 2008),



(*a*) (*b*)

Figure 2
The overall fold of (*a*) *T. cruzi* PGF (PDB entry 4fz1) and (*b*) *T. cruzi* PGF with NADP bound (PDB entry 4gie). The appearance of the disordered loop in the NADP-bound structure can be seen.

monomers were used as the search model in molecular replacement. All other structures were solved by molecular replacement using PDB entry 4f40 as the search model. Initial molecular model building was performed using *ARP/wARP* (Langer *et al.*, 2008). Models were refined against the X-ray reflection data using either *PHENIX* (Adams *et al.*, 2010) and *REFMAC* (Murshudov *et al.*, 2011) interspersed with rounds of model building using *Coot* (Emsley *et al.*, 2010). TLS groups were chosen using *phenix.find_tls_groups* within the *PHENIX* suite (Afonine *et al.*, 2012). Figures containing molecular graphics were prepared using *PyMOL* (Schrödinger). Solution and refinement statistics can be found in Table 3.

3. Results and discussion

As part of the efforts of the SSGCID as a high-throughput structure-determination consortium our goal is to help enable, through structure determination, therapeutic development or function determination of infectious disease proteins (Myler *et al.*, 2009). For each of the two proteins two crystallization trials were set up before finding usable crystals. For LemaA.00019.a.B1 JCSG+ and PACT were used. For TrcrA.00019.a.B1 the Morpheus and Index screens were used. The prostaglandin F synthases from both *L. major* and *T. cruzi* represent classic NADP-binding Rossmann-fold structural motifs, as are common in oxidoreductases (Rao & Rossmann, 1973). These high-resolution structures show that the same loop [residues 188–196 in TrcrA.00019.a.B1 and residues 201–205 (disordered) in LemaA.00019.a.B1] has significant movement between the apoenzyme and holoenzyme (Fig. 2). The *L. major* loop becomes ordered and a key interaction with one of the phosphates of NADP at Gln202 is likely to stabilize this region (Figs. 3*a* and 3*c*). Similarly, *T. cruzi* Ser193 makes an interaction with one of the phosphates of NADP that is likely to stabilize the loop in a significantly different orientation (Figs. 3*b* and 3*d*). In the human PGF structure (PDB entry

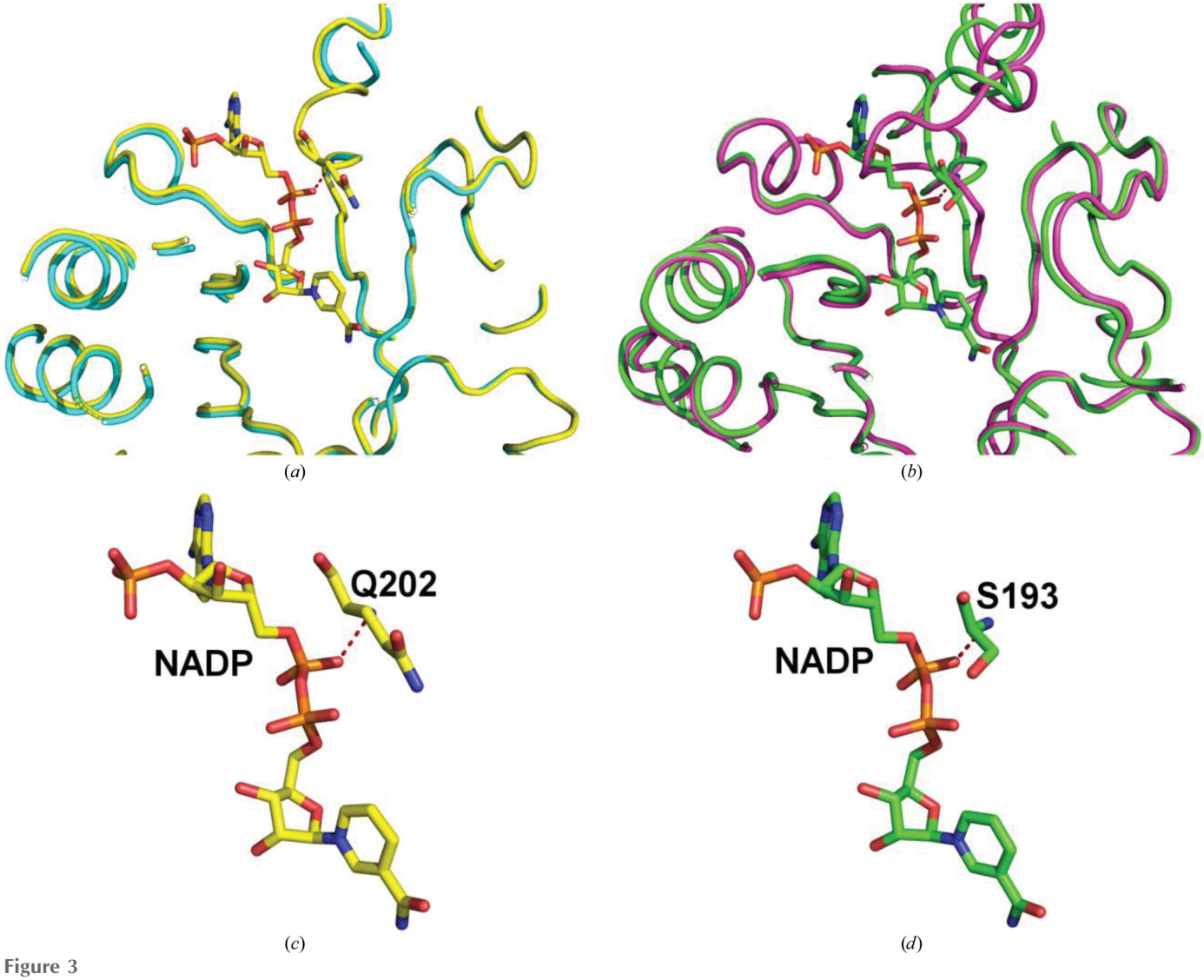


Figure 3
(*a*) Overlay of PDB entry 4f40 (blue) and PDB entry 4g5d (yellow) from *L. major* showing the NADP binding-site area. (*b*) Overlay of PDB entry 4fzi (purple) and PDB entry 4gie (green) from *T. cruzi* showing the NADP binding-site area. (*c*) NADP from PDB entry 4g5d and its interaction with Gln202. (*d*) NADP from PDB entry 4gie and its interaction with Ser193.

1ry8), Ser221 makes a hydrogen-bond interaction with one of the phosphates of NADP *via* the backbone amide and is the equivalent residue to Ser193 and Gln202 in the *T. cruzi* and *L. major* structures, respectively. The pairwise C^{α} r.m.s.d. between the two apo structures of protozoan PGFs is 0.400 Å. The r.m.s.d.s between the two cofactor-bound structures of protozoan PGFs is 0.409 Å. The structures of these proteins with and without their cofactor provide a more robust and clearer understanding of potential binding sites for not only the cofactor but also for the substrate and any subsequent molecules designed for these proteins. Their structural similarity could be useful as a template for structure-based drug-design efforts for therapeutics that target both *L. major* and *T. cruzi*. However, the high structural similarity of human PGF (PDB entry 1ry8), with a C^{α} r.m.s.d. of 0.656 Å for PDB entry 4gie and 0.729 Å for PDB entry 4g5d, represents a potential hurdle for the design of selective molecules even though the sequence identity between human PDF and the *T. cruzi* and *L. major* PGFs is 36 and 37%, respectively. As the interaction of residue Ser221 with the cofactor takes place through the backbone amide, the difference in the amino acid at this location in the binding pocket is not likely to be very important.

Acknowledgements

The authors wish to thank the entire SSGCID team and David M. Dranow. This research was funded by the National Institute of Allergy and Infectious Diseases, National Institute of Health, Department of Health and Human Services under Federa Contract Nos. HHSN272201200025C and HHSN272200700057C. The expression plasmids and surplus protein samples can be obtained from either the contact authors or through http://www.ssgcid.org and the raw X-ray diffraction data can be obtained through http://www.csgid.org.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* D**68**, 352–367.
Aslanidis, C. & de Jong, P. J. (1990). *Nucleic Acids Res.* **18**, 6069–6074.
Buckner, F. S., Wilson, A. J., White, T. C. & Van Voorhis, W. C. (1998). *Antimicrob. Agents Chemother.* **42**, 3245–3250.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Kabsch, W. (2010*a*). *Acta Cryst.* D**66**, 125–132.
Kabsch, W. (2010*b*). *Acta Cryst.* D**66**, 133–144.
Komoto, J., Yamada, T., Watanabe, K. & Takusagawa, F. (2004). *Biochemistry*, **43**, 2188–2198.
Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature Protoc.* **3**, 1171–1179.
Lorimer, D., Raymond, A., Mixon, M., Burgin, A., Staker, B. & Stewart, L. (2011). *Acta Cryst.* F**67**, 985–991.
Magariños, M. P., Carmona, S. J., Crowther, G. J., Ralph, S. A., Roos, D. S., Shanmugam, D., Van Voorhis, W. C. & Agüero, F. (2012). *Nucleic Acids Res.* **40**, D1118–D1127.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Minodier, P. & Parola, P. (2007). *Travel Med. Infect. Dis.* **5**, 150–158.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Myler, P. J., Stacy, R., Stewart, L., Staker, B. L., Van Voorhis, W. C., Varani, G. & Buchko, G. W. (2009). *Infect. Disord. Drug Targets*, **9**, 493–506.
Qin, K.-N., New, M. I. & Cheng, K.-C. (1993). *J. Steroid Biochem. Mol. Biol.* **46**, 673–679.
Rao, S. T. & Rossmann, M. G. (1973). *J. Mol. Biol.* **76**, 241–256.
Stein, N. (2008). *J. Appl. Cryst.* **41**, 641–643.
Watanabe, K., Shimizu, T. & Hayaishi, O. (1981). *Biochem. Int.* **2**, 603–610.