Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Structure of cyclophilin from *Leishmania donovani* at 1.97 Å resolution

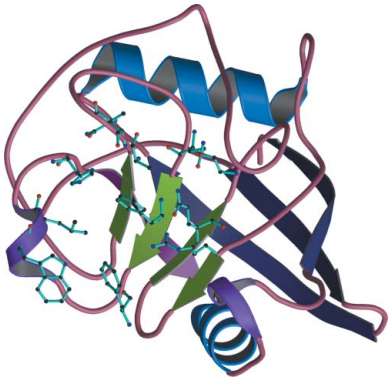
V. Venugopal,[a] Banibrata Sen,[b] Alok K. Datta[b] and Rahul Banerjee[a]*

[a]Crystallography and Molecular Biology Division, Saha Institute of Nuclear Physics, Sector 1, Block AF, Bidhan Nagar, Kolkata 700 064, India, and [b]Indian Institue of Chemical Biology, 4 Raja S. C. Mullick Road, Kolkata 700 032, India

Correspondence e-mail: rahul.banerjee@saha.ac.in

Received 13 November 2006
Accepted 26 December 2006

PDB Reference: cyclophilin, 2haq, r2haqsf.



The crystal structure of cyclophilin from *Leishmania donovani* (LdCyp) has been determined and refined at 1.97 Å resolution to a crystallographic *R* factor of 0.178 (R_{free} = 0.197). The structure was solved by molecular replacement using cyclophilin from *Trypanosoma cruzi* as the search model. LdCyp exhibits complete structural conservation of the cyclosporin-binding site with respect to the homologous human protein, as anticipated from LdCyp–cyclosporin binding studies. Comparisons with other cyclophilins show deviations primarily in the loop regions. The solvent structure encompassing the molecule has also been analyzed in some detail.

1. Introduction

Cyclophilins (Cyps) are an ubiquitous class of proteins functioning primarily as peptidylprolyl *cis–trans* isomerases (PPIases) and are also known to be intracellular receptors for the drug cyclosporin (CsA). The cyclophilin–cyclosporin complex (Cyp–CsA) inhibits calcineurin (Cn), a phosphatase that is involved in signalling. Other biological functions reported for Cyp include involvement in cell division (Schreiber, 1991), signalling pathways (Freeman *et al.*, 1996; Duina *et al.*, 1996), cell-surface recognition (Anderson *et al.*, 1993), protein folding (Schmid, 1993) and heat-shock response (Weisman *et al.*, 1996; Sykes *et al.*, 1993). Human cyclophilin A (CypA) has been reported to bind HIV-1 Gag protein and is essential for the infectious activity of HIV-1 virions (Khan *et al.*, 2004; Thali *et al.*, 2002; Vajdos *et al.*, 1997; Franke *et al.*, 1994; Luban *et al.*, 1993).

Several X-ray crystal structures of CypA in ligated and unligated forms have been reported (Kallen *et al.*, 1998; Ke, 1992) and are now well characterized. The molecule adopts an eight-stranded anti-parallel β-barrel structure capped with two α-helices on either end. The hydrophobic core is tightly packed and inaccessible to ligands. The CsA-binding site is a hydrophobic crevice comprising 11 amino acids on one face of the β-barrel. While neither CypA nor CsA (or its derivatives) can bind Cn on their own, the CypA–CsA complex binds and inhibits Cn (Mikol *et al.*, 1998). Cn is a unique Ca^{2+}/calmodulin-dependent serine/threonine protein phosphatase that is highly conserved from yeast to human (Sugiura *et al.*, 2002; Rusnak & Mertz, 2000) and is a heterodimeric molecule with a catalytic subunit A (CnA) and a regulatory subunit B (CnB). Two crystal structures of human CypA–CsA–Cn ternary complexes have also been solved (Jin & Harrison, 2002; Huai *et al.*, 2002).

While most protozoan and helminthic parasites are susceptible to CsA, *Leishmania major* and *L. donovani* are relatively insensitive to this drug (Page *et al.*, 1995). *Leishmania* spp. are intracellular protozoa that infect macrophages in the vertebrate host and cause leishmaniasis, a broad spectrum of diseases, in humans. Leishmaniasis is currently a major health problem in the tropical countries of the world, with around 350 million people estimated to be at risk (Dedet, 2002). Visceral leishmaniasis, popularly known as kala-azar, caused by *L. donovani* is fatal if not treated in time. In recent years, the re-emergence of drug-resistant strains of *L. donovani* in India (Croft *et al.*, 2006) has called for renewed efforts to identify novel drugs against this pathogen.

Sequence and structural alignment (Fig. 1) of cyclophilin from *L. donovani* (LdCyp) with that from human confirms the complete conservation of the residues binding to CsA, whereas the residues involved in interaction with Cn are only partially conserved. Trp121 in CypA, which was found to be essential for Cn binding (Zydowsky *et al.*, 1992), is conserved in LdCyp. Arg148 of CypA was also thought to play an important role in forming the ternary complex. However, the CsA-sensitivity of the cyclophilins from *Toxoplasma gondii* and *Plasmodium falciparum*, both of which lack Arg at the position corresponding to 148 in the human enzyme (High *et al.*, 1994; Berriman & Fairlamb, 1998), have cast doubt on this hypothesis. Given the ability of LdCyp to bind to CsA, structural studies have been undertaken to elucidate the mechanism of drug (CsA) resistance in *L. donovani*. With this overall aim, we report as a first step the high-resolution crystal structure of LdCyp at 1.97 Å.

2. Materials and methods

2.1. Expression and purification

As reported previously (Dutta *et al.*, 2001; Banerjee *et al.*, 2002), recombinant cyclophilin from *L. donovani* (LdCyp) was overexpressed in *Escherichia coli* M15 cells. The expression vector pQE32 (Qiagen) contained the LdCyP gene (residues 22–187) fused to a N-terminal His_6 tag. Amino acids 1–21 are a signal sequence that is post-translationally cleaved in the native protein. Residues 22–187 thus constitute the native LdCyp polypeptide chain. The protein was purified to homogeneity using a nickel–nitrilotriacetic acid agarose column and was extensively dialysed against buffer containing 0.02 *M* Tris pH 8.5, 0.02% NaN_3 prior to setting up crystallization drops.

2.2. Crystallization and data collection

The crystallization protocol reported previously (Banerjee *et al.*, 2002) was slightly modified and optimized to obtain crystals which diffracted to higher resolution (1.97 Å). Crystals were obtained by the hanging-drop method with 10 µl drops containing 10 mg ml^{-1} protein, 7.5% PEG 3350 inverted over a 1 ml reservoir of 40% PEG 3350 at 293 K. Crystals suitable for data collection appeared in about two weeks. Data were collected at 293 K from crystals mounted in locally produced glass capillary tubes using a MAR image-plate scanner and an in-house Rigaku Cu $K\alpha$ rotating-anode X-ray source. The crystal-to-detector distance was set to 145.0 mm with an oscillation range of 1°. The unit-cell parameters of the crystal, processed by *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997) in Laue symmetry 4/*mmm*, were almost identical to previously reported values. The overall R_{merge} (15.0–1.97 Å) was 6.8%, with a mosacity of 0.314 as output by *SCALEPACK*. Data-collection statistics are summarized in Table 1.

2.3. Crystal structure solution and refinement

The crystal structure based on the new data set (Table 1) was solved by molecular replacement using the program *AMoRe* (Navaza, 1994) from the *CCP*4 program suite with cyclophilin from *Trypanosoma cruzi* (TcCyp; PDB code 1xo7) as the search model. The sequence identity between the two polypeptide chains is 73%. Despite the fact that a low-resolution crystal structure of LdCyp had already been solved (Banerjee *et al.*, 2004), it was decided to re-solve the structure independently as there were several gaps in the previous structure of LdCyp arising from the inferior quality of the low-resolution data. Three models derived from the homologous molecule from *T. cruzi* were used for the molecular-replacement calculations: (i) the complete set of coordinates from 1xo7, (ii) an all-alanine model, the sequence being truncated to Ala with the exception of glycine residues, and (iii) a hybrid model in which residue positions which have sequence identity with LdCyp were included and the remainder were truncated to Ala. Model (iii) gave the best solution in terms of correlation coefficient (72.8; 48.4 for the next best solution) and *R* factor (35.7%; 46.6% for the next best solution) in the previously determined space group $P4_32_12$ (resolution range 15.00–3.5 Å) for one molecule in the asymmetric unit. A total of 735 reflections (6%) were randomly selected from the entire resolution range 15.00–1.97 Å for the calculation of R_{free} (Brünger, 1992). After a few cycles of rigid-body refinement using *CNS* (Brünger *et al.*, 1998) subsequent to the placement of model (iii) in the cell, a $2F_o - F_c$ map was calculated and manual model building was performed using the program *O* (Jones *et al.*, 1991) on a Silicon Graphics workstation. A few iterations of model building and refinement enabled the resolu-

```
Human           --MVNPTVFFDIAVDGEPLGRVSFELFADKVPKTAENFRALSTGE------KGFGYKGSC 52
L. donovani     EPEVTAKVYFDVMIDSEPLGRITIGLFGKDAPLTTENFRQLCTGE------HGFGYKDSI 75
T. cruzi        MPVVTDKVYFDITIGDEPVGRVVIGLFGNDVPKTVENFKQLASGE------NGFGYKGSI 54
Yeast           --MS--QVYFDVEADGQPIGRVVFKLYNDIVPKTAENFRALCTGE------KGFGYAGSP 50
Plasmodium      --SKRSKVFFDISIDNSNAGRIIFELFSDITPRTCENFRALCTGEKIGSRGKNLHYKNSI 58
E. coli         -----------MVTFHTNHGDIVIKTFDDKAPETVKNFLDYCREG---------FYNNTI 40

                  *    *  *        *                              ***       *
Human           FHRIIPGFMCQGGDFTRHNGTGGKSIYGEKFEDENFILKHTGPGILSMANAGPNTNGSQF 112
L. donovani     FHRVIQNFMIQGGDFTNFDGTGGKSIYGEKFADENLNVKH-FVGALSMANAGPNTNGSQF 134
T. cruzi        FHRVIRNFMIQGGDFTNFDGTGGKSIYGTRFDDENLKIKH-FVGAVSMANAGPNSNGSQF 113
Yeast           FHRVIPDFMLQGGDFTAGNGTGGKSIYGGKFPDENFKKHHDRPGLLSMANAGPNTNGSQF 110
Plasmodium      FHRIIPQFMCQGGDITNGNGSGGESIYGRSFTDENFNMKHDQPGLLSMANAGPNTNSSQF 118
E. coli         FHRVINGFMIQGGGFEP--GMKQKATKEPIKNEANNGLKNTRGTLAMARTQAPHSATAQF 98

                *       **           *                          *
Human           FICTAKTEWLDG--------KHVVFGKVKEGMNIVEAMERFGS-----RNGKTSKKITIA 159
L. donovani     FITTAPTPWLDG--------RHVVFGKVLDGMDVVLRIEKTKTN----SHDRPVKPVKIV 182
T. cruzi        FVTTAPTPWLDG--------RHVVFGKVVEGMDVVKKVENTKTG----LNDKPKKAVKIN 161
Yeast           FITTVPCPWLDG--------KHVVFGEVVDGYDIVKKVESLGS-----PSGATKARIVVA 157
Plasmodium      FITLVPCPWLDG--------KHVVFGKVIEGMNVVREMEKEGA-----KSGYVKRSVVIT 165
E. coli         FINVVDNDFLNFSGESLQGWGYCVFAEVVDGMDVVDKIKGVATGRSGMHQDVPKEDVIIE 158

Human           DCGQLE 165
L. donovani     ASGEL- 187
T. cruzi        DCGVL- 166
Yeast           KSGEL- 162
Plasmodium      DCGEL- 170
E. coli         SVTVSE 164
```

Figure 1
Sequence alignment of cyclophilins from human (PDB code 2cpl), *L. donovani* (2haq), *T. cruzi* (1xo7), yeast (1vdn), *P. falciparum* (1qng) and *E. coli* (2nul). The cyclosporin-binding site is marked with a star. Cyp from *E. coli* does not bind to the drug.

Table 1
Summary of diffraction data and crystallographic refinement statistics.

Values in parentheses are for the highest resolution shell (2.02–1.97 Å).

PDB code	2haq
Space group	$P4_32_12$
Unit-cell parameters (Å, °)	$a = b = 48.59$, $c = 141.19$, $\alpha = \beta = \gamma = 90$
Wavelength (Å)	1.5418
Resolution (Å)	15.00–1.97
Average redundancy	8.1 (5.2)
R_{merge}† (%)	6.8 (29.3)
Completeness (%)	99.0 (86.6)
Unique reflections	12516
σ cutoff (F)	2.00
$\langle I\rangle/\langle\sigma(I)\rangle$	30.1 (4.5)
No. of reflections (working set)	11402
No. of reflections (test set)	735
No. of protein atoms	1274
No. of water molecules	102
R_{cryst}/R_{free}‡ (%)	17.8/19.7
R.m.s. deviation from ideal values	
Bond lengths (Å)	0.007
Bond angles (°)	1.40
Dihedrals (°)	25.6
Impropers (°)	0.78
Mean B value (Å^2)	28.5
Ramachandran plot§	
Most favoured (%)	85.3
Additional allowed (%)	14.0
Generously allowed (%)	0.7
Residues truncated to Ala	Glu22, Gln81, Asn170, Ser171

† $R_{merge} = \sum\sum|I(k) - \langle I\rangle|/\sum I(k)$, where $I(k)$ and $\langle I\rangle$ represent the diffraction-intensity values of the individual measurements and the corresponding mean values. The summation is over all unique measurements. ‡ $R_{cryst} = \sum||F_{obs}| - |F_{calc}||/\sum|F_{obs}|$; R_{free} is R_{cryst} for a 6% cross-validated test data set. § The Ramachandran plot was obtained using *PROCHECK* (Laskowski *et al.*, 1993).

tion of the data to be raised from 3.5 to 2.25 Å, at which stage simulated annealing (maximum-likelihood amplitude, MLF target) was performed following the slow-cooling molecular-dynamics scheme with a starting temperature of 5000 K. It was possible to model almost all the side chains of LdCyp, as well defined electron density (Fig. 2) was observed even for those residues that had been

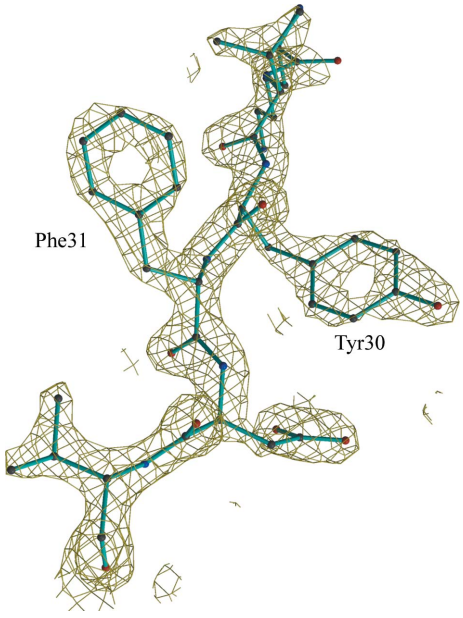


Figure 2
Electron density contoured at the 2.5σ level for residues Val29–Val33 of LdCyp from a $2F_o - F_c$ map. This figure and Figs. 3–5 were generated using *MOLSCRIPT*/*BOBSCRIPT* (Kraulis, 1991; Esnouf, 1999) and *RASTER3D* (Merritt & Bacon, 1997).

Table 2
Sequential and structural conservation amongst cyclophilins overall and for the residues constituting the hydrophobic core.

R.m.s.d. was calculated with respect to LdCyp using C$^\alpha$ atoms alone.

		Overall		Core residues	
Source	PDB code	Sequence identity (%)	R.m.s.d. (Å)	Sequence identity (%)	R.m.s.d. (Å)
T. cruzi	1xo7	73	0.5	79	0.28
Human	2cpl	60	1.2	66	0.34
Yeast	1vdn	59	1.1	70	0.37
P. falciparum	1qng	58	1.2	75	0.37
E. coli	2nul	37	1.6	87	0.57

truncated to alanine. The R and R_{free} at this stage of refinement were 25% and 28%, respectively. Subsequently, the full data set (15.00–1.97 Å) was used, with R and R_{free} falling to 26% and 27%, respectively. Waters were identified automatically from an $F_o - F_c$ map based on the hydrogen-bond criteria using the *CNS* task file `water_pick.inp`. Waters with unrealistic B factors (>60 Å^2) or irregular electron density were excluded. Individual B factors were refined in the final stage with an initial value of 20 Å^2. The final fitting was performed on an OMIT map. A total of 1274 protein atoms and 102 solvent water molecules were included in the final structure, which had R and R_{free} values of 17.8% and 19.7%, respectively. Residues with missing electron density (Glu22, Gln81, Asn170 and Ser171) were truncated to Ala. The coordinates were deposited in the Protein Data Bank with PDB code 2haq after validation using *PROCHECK* (Laskowski *et al.*, 1993). The final model contained no residues in the disallowed region of the Ramachandran plot and had r.m.s. deviations in bond lengths and bond angles of 0.007 Å and 1.40°, respectively. The mean coordinate errors for the working set estimated from the Luzzati plot (Luzzati, 1952) and σ_A (Read, 1986) were 0.20 and 0.15 Å, respectively. The overall B factor from the Wilson plot and that averaged from the finally refined coordinates were 24.40 and 28.50 Å^2, respectively, and were thus in good agreement. The refinement and structure-validation statistics are summarized in Table 1.

3. Results and discussion

3.1. Overall structure

LdCyp adopts an eight-stranded β-barrel structure composed of antiparallel β-sheets similar to those of other cyclophilins (Kallen *et al.*, 1991). The barrel is composed of two antiparallel sheet systems, Sheet-1 and Sheet-2, which are approximately orthogonal to each other, with two α-helices, H1 (Ala52–Gly65) and H3 (Gly157–Lys166), closing the barrel at either end. Four strands, namely B3 (Arg78–Leu80), B4 (Met84–Gly87), B5 (Ala119–Met122) and B6 (Phe134–Thr137), comprise Sheet-1, which is approximately planar. The other strands, B1 (Val29–Ile35), B2 (Glu38–Leu47) and B8 (Val179–Glu186), form Sheet-2. The β-strands making up Sheet-2 are considerably longer than those that make up Sheet-1. Strand B7 (Val150–Asp156) is twisted and interacts with both the sheet systems (B5 and B2) *via* hydrogen bonds. The β-strands and α-helices are connected by β-turns and loops. The only hydrophobic core of the protein lies within the barrel constituted by an extended cluster of amino acids (Val29, Phe31, Val33, Ile45, Leu47, Phe48, Ala52, Thr55, Asn58, Phe59, Leu62, Cys63, Tyr71, Phe76, Val79, Ile85, Leu120, Met122, Phe134, Phe151, Val161, Ile164, Val179 and Ile181) which are highly conserved in cyclophilins (Ellis *et al.*, 2000; Mikol *et al.*, 1994; Peterson *et al.*, 2000; Fejzo *et al.*, 1994). In addition to the two

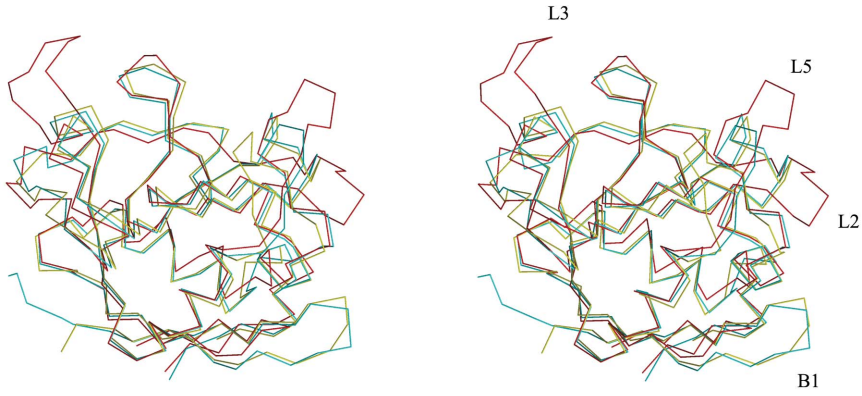


Figure 3
Superposition of C^{α}-backbone structures of the Cyps from *L. donovani* (cyan), human (yellow) and *E. coli* (red). Insertions in *E. coli* Cyp (corresponding to regions L2, L3 and L5 in LdCyP) and deletions (B1) are indicated.

extended α-helices (H1 and H3), there are also two short 3_{10}-helices H2 (Thr141–Asp145) and H4 (Asn170–Arg174) within surface-exposed loops of the protein. Three type IV β-turns (Wilmot & Thornton, 1998) connect B1 to B2, B2 to H1 and B3 to B4. A somewhat unusual feature in a type II turn (Tyr71–Ser74) is the position of Asp73 (with φ, ψ = 70.5, 18.8°) at position $i + 2$ generally occupied by Gly in a typical type II turn. In contrast to human cyclophilin (Ke, 1992), which contains six β-bulges, LdCyp contains only three: between strands B2 and B1 (Leu40–Gly41 hydrogen bonded to Val33), B7 and B5 (Phe151–Gly152 to Leu120) and B8 and B1 (Val182–Ala183 to Asp32). There is only one cysteine residue in LdCyp. The loop regions L1 (His67–His77), L2 (Gly88–Gly118), L3 (Ala123–Gln133), L4 (Thr138–Val149) and L5 (Thr167–Pro178) are considerably flexible and contain most of the residues involved in the functional activity of the protein.

Structural alignment and superposition of LdCyp with Cyps from other sources using the *DALI* server (Holm & Park, 2000) shows a high degree of conservation in the β-strands and helices and variability in the loop regions. The highest structural variability is observed in L5 (Fig. 3), with the r.m.s.d.s between individual C^{α} atoms ranging from 1.5 to 5 Å (with respect to human, yeast and *P. falciparum* Cyps). The bacterial Cyp has the largest structural deviation from LdCyp (r.m.s.d. 1.6 Å), which is partly caused by several insertions and deletions along the polypeptide chain of the bacterial enzyme and a shorter β-strand corresponding to B1 in LdCyp. The helices corresponding to H1, H2 and H4 in LdCyp exhibit r.m.s. deviations of between 1 and 2 Å in the bacterial protein, in addition to large deviations in L2. The hydrophobic core is highly conserved both sequentially and structurally compared with the rest of the molecule (Table 2).

3.2. Hydration shell

102 water molecules were initially identified in the asymmetric unit. A distance criterion of 3.6 Å was employed between solvent waters and polar atoms of the protein molecule in order to generate hydration shells by appropriate application of crystallographic symmetry (Baker & Hubbard, 1984; Kodandapani *et al.*, 1990; Radha Kishan *et al.*, 1995). The first and second hydration shells contained 113 and 17 waters, respectively. In the first hydration shell, 56, 26, 16, nine and four water molecules make one, two, three, four and five contacts with the protein, respectively. Two waters were found to make six protein contacts and one water molecule (Wat11) exhibited a maximum number of ten surrounding polar protein atoms (but only made four actual hydrogen bonds). A total of 230 contacts were found for the 113 water molecules in the first hydration shell, involving 86 main-chain CO groups, 52 main-chain NH groups, 62 side-chain O atoms and 30 side-chain N atoms. Waters making more than two or three contacts have been postulated to play a significant role in stabilizing local structural elements (Madhusudan & Vijayan, 1991). Almost all the waters making three or more protein contacts were found to bridge distinct secondary-structural elements in LdCyp (Fig. 4). Waters bridging β-strands and loops (Wat24, L1–B8; Wat32, L1–B4; Wat42, L2–B7), neighbouring loops (Wat2 and Wat13, L1–L2; Wat15, L2–L4) and helices and loops (Wat73, H1–L2; Wat66, H3–L4) were fairly common. Waters also occupied voids within helical and β-strand packing interfaces (Wat31, B2–H3; Wat19, H1–B8).

3.3. Active site

The CsA-binding site is constituted of Arg78, Phe83, Gln86, Gly95, Ala123, Asn124, Gln133, Phe135, Trp143, Leu144 and His148 in LdCyp. The residues corresponding to His148 and Arg78 in human Cyp have been postulated to play a primary role in PPIase activity (Kallen *et al.*, 1991; Ke *et al.*, 1993). The active-site residues cluster on one face of the protein contributed by the secondary-structural elements B3, B4, B6, T5, L2, L3 and L4. It will be recalled that strands B3, B4 and B6 are from Sheet-2 (Fig. 5) and form the base of the binding site, with the rest of the residues being from surrounding loops and turns that form the crevice. This is unlike most instances, in which the active site is within the barrel (Ke, 1992). As expected,

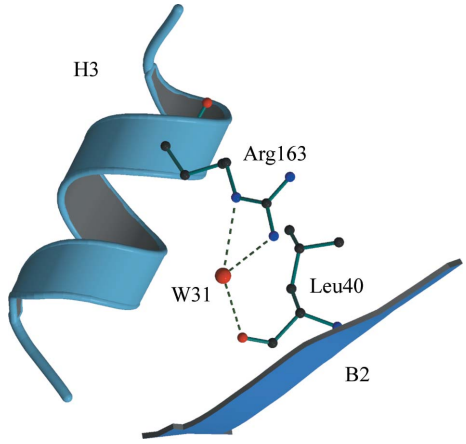


Figure 4
Wat31 (shown in red) bridging helix H3 and β-strand B2. Hydrogen bonds are shown as dashed lines. Residues Leu40 from B2 and Arg163 from H3 are indicated.

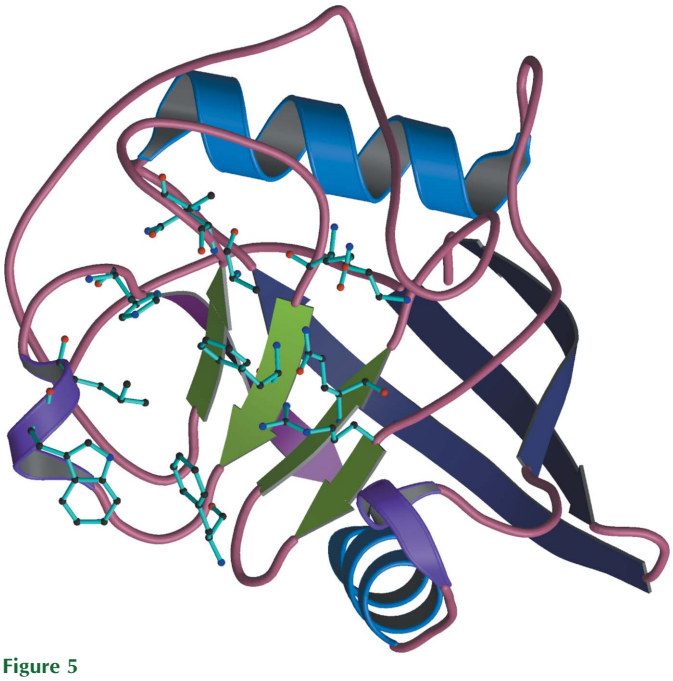


Figure 5
Ribbon diagram of the overall three-dimensional structure of cyclophilin from *L. donovani*. The active-site residues are shown in ball-and-stick representation.

these residues are structurally and sequentially conserved in the human and *T. cruzi* enzymes, with the r.m.s.d.s for active-site residues being 0.41 and 0.35 Å, respectively. Thus, from structural considerations LdCyP is expected to bind CsA with high affinity, which has also been confirmed experimentally (Dutta *et al.*, 2001).

4. Conclusions

The high-resolution crystal structure of LdCyp exhibits complete structural conservation of the CsA-binding site with respect to human Cyp. Thus, the high affinity of CsA for LdCyp, which has been demonstrated experimentally, is structurally validated. The insensitivity of *L. donovani* to CsA is thus not a consequence of the inability of LdCyp to bind to CsA. In addition, there is evidence that low expression of cytoplasmic cyclophilin coupled with efflux of the drug in *L. donovani* (Dutta *et al.*, 2001) could contribute to drug resistance in the parasite. The status of calcineurin in the parasite is presently unknown. The cloning of calcineurin from *L. donovani* has been initiated in order to structurally and functionally characterize the interaction between the three molecules.

This work was funded by grants from the Council of Scientific and Industrial Research, Government of India. Professor Udayaditya Sen is thanked for his technical assistance.

References

Anderson, S. K., Gallinger, S., Roder, J., Frey, J., Young, H. A. & Ortaldo, J. R. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 542–546.
Baker, E. N. & Hubbard, R. E. (1984). *Prog. Biophys. Mol. Biol.* **44**, 97–179.
Banerjee, R., Dutta, M., Sen, M. & Datta, A. K. (2002). *Acta Cryst.* D**58**, 1846–1847.
Banerjee, R., Dutta, M., Sen, M. & Datta, A. K. (2004). *Curr. Sci.* **86**, 319–322.
Berriman, M. & Fairlamb, A. H. (1998). *Biochem. J.* **334**, 437–445.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Croft, S. L., Sundar, S. & Fairlamb, A. H. (2006). *Clin. Microbiol. Rev.* **19**, 111–126.
Dedet, J.-P. (2002). *World Class Parasites*, Vol. 4, *Leishmania*, edited by J. P. Farrell, pp. 1–10. Boston: Kluwer Academic Publishers.
Duina, A. A., Chang, H. C., Marsh, J. A., Lindquist, S. & Caber, R. F. (1996). *Science*, **274**, 1713–1715.
Dutta, M., Delhi, P., Sinha, K. M., Banerjee, R. & Dutta, A. K. (2001). *J. Biol. Chem.* **276**, 19294–19300.
Ellis, P. J., Carlow, C. K. S., Ma, D. & Kuhn, P. (2000). *Biochemistry*, **39**, 592–598.
Esnouf, R. M. (1999). *Acta Cryst.* D**55**, 938–940.
Fejzo, J., Etzkorn, F. A., Clubb, R. T., Shi, Y., Walsh, C. T. & Wagner, G. (1994). *Biochemistry*, **33**, 5711–5720.
Franke, E. K., Yuan, H. E. H. & Luban, J. (1994). *Nature (London)*, **372**, 359–362.
Freeman, B. C., Toft, D. O. & Morimoto, R. I. (1996). *Science*, **274**, 1718–1720.
High, K. P., Joiner, K. A. & Handschumacher, R. E. (1994). *J. Biol. Chem.* **269**, 9105–9112.
Holm, L. & Park, J. (2000). *Bioinformatics*, **16**, 566–567.
Huai, Q., Kim, H. Y., Liu, Y., Zhao, Y., Mondragon, A., Liu, J. O. & Ke, H. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 12037–12042.
Jin, L. & Harrison, S. C. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 13522–13526.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kallen, J., Mikol, V., Taylor, P. & Walkinshaw, M. D. (1998). *J. Mol. Biol.* **283**, 435–449.
Kallen, J., Spitzfaden, C., Zurini, M., Wider, G., Widmer, H., Wuthrich, K. & Walkinshaw, M. D. (1991). *Nature (London)*, **353**, 276–279.
Ke, H. (1992). *J. Mol. Biol.* **228**, 539–550.
Ke, H., Mayrose, D. & Cao, W. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 3324–3328.
Khan, M., Jin, L., Huang, M. B., Miles, L., Bond, V. C. & Powell, M. D. (2004). *J. Virol.* **78**, 1843–1850.
Kodandapani, R., Suresh, C. G. & Vijayan, M. (1990). *J. Biol. Chem.* **265**, 16126–16131.
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.
Luban, J., Bossolt, K. L., Franke, E. K., Kalpana, G. V. & Goff, S. P. (1993). *Cell*, **73**, 1067–1078.
Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.
Madhusudan & Vijayan, M. (1991). *Curr. Sci.* **60**, 165–170.
Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
Mikol, V., Kallen, J. & Walkinshaw, M. D. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 5183–5186.
Mikol, V., Paul, T., Kallen, J. & Walkinshaw, M. D. (1998). *J. Mol. Biol.* **283**, 451–461.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–374.
Page, A. P., Kumar, S. & Carlow, C. K. S. (1995). *Parasitol. Today*, **11**, 385–388.
Peterson, M. R., Hall, D. R., Berriman, M., Nunes, J. A., Leonard, G. A., Fairlamb, A. H. & Hunter, W. N. (2000). *J. Mol. Biol.* **298**, 123–133.
Radha Kishan, K. V., Chandra, N. R., Sudarsanakumar, C., Suguna, K. & Vijayan, M. (1995). *Acta Cryst.* D**51**, 703–710.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Rusnak, F. & Mertz, P. (2000). *Phys. Rev.* **80**, 1483–1521.
Schmid, F. X. (1993). *Annu. Rev. Biophys. Biomol. Struct.* **22**, 123–143.
Schreiber, S. L. (1991). *Science*, **251**, 283–287.
Sugiura, R., Sio, S. O., Shuntoh, H. & Kuno, T. (2002). *Genes Cells*, **7**, 619–627.
Sykes, K., Gething, M. J. & Sambrook, J. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 5853–5857.
Thali, M., Bukovsky, A., Kondo, E., Rosenwirth, B., Walsh, C. T., Sodroski, J. & Göttlinger, H. G. (2002). *Nature (London)*, **372**, 363–365.
Vajdos, F. F., Yoo, S., Houseweart, M., Sundquist, W. I. & Hill, C. P. (1997). *Protein Sci.* **6**, 2297–2307.
Weisman, R., Creanor, J. & Fantes, P. (1996). *EMBO J.* **15**, 447–456.
Wilmot, C. M. & Thornton, J. M. (1998). *J. Mol. Biol.* **203**, 221–232.
Zydowsky, L. D., Etzkorn, F. A., Chang, H. Y., Ferguson, S. B., Stolz, L. A., Ho, S. I. & Walsh, C. T. (1992). *Protein Sci.* **1**, 1092–1099.