The Crystal Structure of the Leishmania major Deoxyuridine Triphosphate Nucleotidohydrolase in Complex with Nucleotide Analogues, dUMP, and Deoxyuridine

Glyn R. Hemsworth, Olga V. Moroz, Mark J. Fogg, Benjamin Scott, Cristina Bosch-Navarrete, Dolores González-Pacanowska, and Keith S. Wilson

From the Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York YO10 5DD, United Kingdom and the Instituto de Parasitología y Biomedicina “López-Neyra,” Consejo Superior de Investigaciones Científicas, Granada, 18100 Armilla, Spain

Members of the Leishmania genus are the causative agents of the life-threatening disease leishmaniasis. New drugs are being sought due to increasing resistance and adverse side effects with current treatments. The knowledge that dUTPase is an essential enzyme and that the all α-helical dimeric kinetoplastid dUTPases have different structures compared with the trimeric β-sheet type dUTPase possessed by most organisms, including humans, make the dimeric enzymes attractive drug targets. Here, we present crystal structures of the Leishmania major dUTPase in complex with substrate analogues, the product dUMP and a substrate fragment, and of the homologous Campylobacter jejuni dUTPase in complex with a triphosphate substrate analogue. The metal-binding properties of both enzymes are shown to be dependent upon the ligand identity, a previously unseen characteristic of this family. Furthermore, structures of the Leishmania enzyme in the presence of dUMP and deoxyuridine coupled with tripyran fluorescence quenching indicate that occupation of the phosphate binding region is essential for induction of the closed conformation and hence for substrate binding. These findings will aid in the development of dUTPase inhibitors as potential new lead antitrypanosomal compounds.

Various members of the genus Leishmania cause leishmaniasis, which threatens ∼350 million people worldwide and gives rise to about two million clinical cases each year, of which ∼25% are of the fatal visceral form (1). The disease is largely endemic to developing countries, and current treatments are expensive and can result in undesirable side effects for the patient (1). This, combined with the increasing drug resistance that is developing in the Leishmania species, means that new and novel anti-parasitic drug targets are urgently required.

Deoxyuridine triphosphate nucleotidohydrolase (dUTPase) represents such a target. dUTPases catalyze the hydrolysis of dUTP to dUMP and pyrophosphate (2). This provides the starting material for the synthesis of dTMP by thymidylate synthase and in addition maintains the ratio of dTTP:dUTP in the cell at a high enough level to prevent excessive misincorporation of dUMP into the genome during DNA replication (3). dUTPase activity is essential as was shown by gene knock-outs in Escherichia coli and Saccharomyces cerevisiae. Inhibition of dUTPase activity results in futile cycles of DNA repair that ultimately cause the fragmentation of DNA and cell death (4, 5). dUTPase activity also appears to be essential in L. major (6) and the related parasite Trypanosoma brucei, where an RNAi approach was used to knock down dUTPase expression, resulting in decreased cell proliferation and growth in both procyclic and bloodstream forms of the organism (7).

dUTPases have been characterized extensively both biochemically and structurally in many species. Most organisms, including humans, have trimeric dUTPases with structures largely consisting of β-sheeted sheet (8). Three active sites are formed at the trimer interfaces by five characteristic sequence motifs. Each subunit contributes to every active site present in the trimer, with a single divalent metal ion bound in each of the active sites in the presence of substrate (9, 10). This metal ion has been implicated in substrate binding and catalysis as the presence of Mg2+ reduces the Km of the E. coli enzyme by a 0.5 to 1 order of magnitude and increases the value of kcat 2- to 3-fold in the Drosophila melanogaster enzyme (11, 12). Monomeric dUTPases have also been identified that appear to have arisen by gene duplication merging two protomers of the trimeric enzyme into a single polypeptide (13). These monomeric enzymes are found only in mammalian herpes viruses and have the same five characteristic sequence motifs as the trimeric enzymes but in a different order.

The enzymes from L. major (LmdUTPase), Trypanosoma cruzi (TcdUTPase), and Trypanosoma brucei (TbdUTPase) as

*This work was funded by European Union Grant agreement 222328, Spanish Plan Nacional Grant SAF2007-62596, FIS Network Grant RD06/0021, and the Junta de Andalucía Grant CVI-199.

The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. 1–3.

The atomic coordinates and structure factors (codes 2CJE, 2CIC, 2YAY, 2YAZ, and 2Y80) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

1 To whom correspondence should be addressed: Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York YO10 5DD, UK. E-mail: keith@ysbl.york.ac.uk.

2 The abbreviations used are: dUTPase, deoxyuridine triphosphate nucleotidohydrolase; LmdUTPase, dUTPase from L. major; TcdUTPase, dUTPase from T. cruzi; TbdUTPase, dUTPase from T. brucei; Cj dUTPase, dUTPase from C. jejuni; dUTP, 2′-deoxyuridine 5′-triphosphate; dUDP, 2′-deoxyuridine 5′-diphosphate; dUMP, 2′-deoxyuridine 5′-monophosphate; dUDP, a,β-imino-deoxyuridine triphosphate; dUpNp, a,β-imino-deoxyuridine diphosphate; r.m.s.d., root mean square deviation(s); TLS, Translation Libration Screw-motion.
well as some bacterial species show markedly different structural and biochemical properties to the trimeric enzymes (14). These enzymes act as dimers and are able to hydrolyze not only dUTP but also dUDP (15), whereas dUDP has been shown to be an inhibitor of the trimeric enzyme in E. coli (16). They are, in addition, subject to product inhibition by dUMP (17). The structure of TcdUTPase revealed a dimeric all-α-helical structure that is very distinct from the trimeric enzymes (18), with which they share no sequence similarity. The dimeric dUTPases also have five sequence motifs responsible for forming the active site, but these are quite different from those in the trimeric dUTPases. The structure of the T. cruzi enzyme has been determined in the apo open form and in the closed form in complex with dUDP (18). These structures revealed a large conformational change in the protein upon substrate binding with the mobile domain moving as much as 20 Å with the concomitant rearrangement of secondary structure elements in this domain. Subsequently, a structure of the closed form of the dimeric dUTPase from Campylobacter jejuni (CjdUTPase) was determined in the presence of the non-hydrolyzable substrate analogue dUpNp (19). This revealed the presence of three Mg²⁺ ions in the active site adjacent to the phosphate groups of the substrate. The metals were implicated in substrate binding and led to the proposal that these dimeric dUTPases utilize a mechanism of reaction similar to that used in the phosphorolysis transfer reactions catalyzed by DNA polymerases (20, 21). Indeed, the dimeric enzyme, like the trimeric one, requires divalent metal ions for activity (17).

The essential nature of dUTPases and the marked differences between the two forms of the enzyme in different species make the dimeric dUTPases an attractive target for anti-parasitic drug development. With this in mind, we have undertaken a structural characterization of the L. major dUTPase in complex with various substrate fragments to better understand the substrate binding determinants and the requirements to induce closure of this family of enzymes. Here, we present crystal structures of the dimeric dUTPase from Campylobacter jejuni (CjdUTPase) and the trimeric dUTPase from T. cruzi (TcDUT3C) in complex with dUTP but also dUDP (15), whereas dUDP has been shown to be an inhibitor of the trimeric enzyme in E. coli (16). They are, in addition, subject to product inhibition by dUMP (17). The structure of TcdUTPase revealed a dimeric all-α-helical structure that is very distinct from the trimeric enzymes (18), with which they share no sequence similarity. The dimeric dUTPases also have five sequence motifs responsible for forming the active site, but these are quite different from those in the trimeric dUTPases. The structure of the T. cruzi enzyme has been determined in the apo open form and in the closed form in complex with dUDP (18). These structures revealed a large conformational change in the protein upon substrate binding with the mobile domain moving as much as 20 Å with the concomitant rearrangement of secondary structure elements in this domain. Subsequently, a structure of the closed form of the dimeric dUTPase from Campylobacter jejuni (CjdUTPase) was determined in the presence of the non-hydrolyzable substrate analogue dUpNp (19). This revealed the presence of three Mg²⁺ ions in the active site adjacent to the phosphate groups of the substrate. The metals were implicated in substrate binding and led to the proposal that these dimeric dUTPases utilize a mechanism of reaction similar to that used in the phosphorolysis transfer reactions catalyzed by DNA polymerases (20, 21). Indeed, the dimeric enzyme, like the trimeric one, requires divalent metal ions for activity (17).

The essential nature of dUTPases and the marked differences between the two forms of the enzyme in different species make the dimeric dUTPases an attractive target for anti-parasitic drug development. With this in mind, we have undertaken a structural characterization of the L. major dUTPase in complex with various substrate fragments to better understand the substrate binding determinants and the requirements to induce closure of this family of enzymes. Here, we present crystal structures of the closed L. major dUTPase in the presence of the non-hydrolyzable substrate analogues α,β'-imino-deoxyuridine triphosphate (dUpNpp) and α,β'-imino-deoxyuridine diphosphate (dUpNp) and divalent metal ions supporting the proposed mechanism for these dimeric enzymes. Subsequent structures with deoxyuridine monophosphate (dUMP) and deoxyuridine (dU) bound reveal a completely closed conformation with sulfate ions bound to the enzyme, showing the importance of the negatively charged 5’ region of the substrate to induce enzyme closure. We also present the structure of the CjdUTPase in complex with dUpNpp with important implications for the metal-binding properties of this family of enzymes.

**EXPERIMENTAL PROCEDURES**

*Cloning and Expression of L. major dUTPase—*The coding sequence for the dimeric dUTPase from L. major 252 was amplified by PCR, using genomic DNA as template, and cloned. For expression, two constructs were created using the forward primers LmDUT3C-F (5’-CCAGGGACCAGCAATGAAGCGCGCTCGCAGC-3’) and LmDUTLIC-F (5’-CACCCACCACCATGAAGCGCGCTCGCAGC-3’) with the reverse primer LmDUT-R (5’-GAGGAGAAGGGCGCGTTATGCT-

TGATGCAGCCAGCCG-3’) common to both constructs. The underlined overhangs were included to allow cloning into the pET-YSB3C and pET-YSBLLIC vectors, respectively (22). 0.2 pmol of PCR product were treated with T4 DNA polymerase (New England Biolabs) in the presence of dATP to produce large overhangs for cloning into the Ligation Independent Cloning vectors. The Ligation Independent Cloning vector was linearized with the restriction enzyme BseRI and treated with T4 DNA polymerase in the presence of 2’-deoxothymidine 5’-triphosphate to produce complementary overhangs to the PCR product. The PCR product and the vector were then mixed in a 2:1 ratio of insert to vector and allowed to anneal at room temperature for 30 min. The mixture was then used to transform NovaBlue singles competent cells (Novagen). Colonies were picked and grown in 5-ml cultures overnight. Plasmids were isolated from cells using the Qiagen mini-prep kit, and the presence of insert in the vector confirmed by sequencing from the T7 promoter. The vectors were designated pET-LmDUT3C and pET-LmDUTLIC. BL-21 (DE3) competent cells were then transformed with these vectors to allow protein expression. Six 500-ml cultures derived from a single colony of BL-21 (pET-LmDUT3C) or BL-21 (pET-LmDUTLIC) were grown in LB medium containing 30 µg/ml kanamycin at 37 °C to an A₆₀₀ of 0.6–0.8. Expression of dUTPase was then induced by the addition of isopropyl 1-thio-β-D-galactopyranoside to a final concentration of 1 mM, and the temperature was lowered to 25 °C for the LmDUT3C construct but maintained at 37 °C for the LmDUTLIC construct, as these were the conditions determined to be optimal for soluble expression of protein in small scale expression trials. After 4 h, cells were harvested by centrifugation at 8000 rpm in a Sorvall F-10 rotor for 20 min. The cell paste was stored at −20 °C until required for purification.

*Purification of L. major dUTPase—*Typically ~2.5 g of cell paste were resuspended in 10× volumes of lysis buffer (50 mM HEPES, pH 7.0, 50 mM NaCl, 5 mM MgCl₂, 1 mM PMSE), and cells were lysed using three 10-s pulses of sonication at maximum amplitude using an MSE Soniprep 150 sonicator. Cell debris was removed by centrifugation at 18,000 rpm in a Sorvall SS-34 rotor for 20 min at 4 °C. Supernatants were removed and applied to a 5-ml GE Healthcare HiTrap nickel-nitrilotriacetic acid column that had been equilibrated in Buffer A (50 mM HEPES, pH 7.0, 50 mM NaCl, 5 mM MgCl₂, 1 mM imidazole). The column was washed with 3 column volumes of Buffer A before a 50-ml gradient of Buffer B (Buffer A + 300 mM imidazole) was applied to elute the protein. 5-ml fractions were collected along the gradient. For the LmDUT3C construct, peak fractions containing the enzyme were pooled for treatment with human rhinovirus 3C protease (Novagen). 1 unit of protease was added per mg of dUTPase, and the solutions were incubated at 4 °C overnight. The protein was concentrated using a 10-kDa molecular mass cut off Millipore concentrator by centrifugation at 4 × g to a volume of ~1 ml. The protein was diluted 10-fold with Buffer A to dilute the imidazole, and the protein was applied back onto the nickel column used in the previous step. The flow-through was collected, and the column was washed with 2 column volumes of Buffer A. The untagged protein was in the flow-through, which was concentrated to ~1 ml for gel filtration. The LmDUTLIC construct lacks the 3C
protease site and therefore was concentrated following the initial nickel column for gel filtration. Gel filtration was performed in the same way for both constructs using a GE Healthcare 16/60 HiLoad Superdex S75 column, which had been equilibrated with gel filtration buffer (10 mM HEPES, pH 7.0, 200 mM NaCl, 5 mM MgCl₂). Following a void volume of 45 ml, 2.5-ml fractions were collected. Fractions containing LmDUTPase were combined, and the buffer was exchanged into 10 mM HEPES, pH 7.0, 5 mM MgCl₂ on a Millipore concentrator. The LmDUT3C and LmDUTLIC protein samples were concentrated to 80 mg/ml and 15 mg/ml, respectively, as judged by the absorbance at 280 nm using an extinction coefficient of 53,900 M⁻¹ cm⁻¹ for both proteins. 50-μl aliquots of protein were flash-frozen under liquid nitrogen and stored at −80 °C for future use.

**Crystallization of L. major dUTPase in Complex with dUpNp and dUpNpp**—The LmDUTLIC sample was used for crystallization with the dUpNp substrate analogue (Jena Biosciences). Prior to crystallization, dUpNp was added to the protein to a final concentration of 2 mM. Initial crystallization trials were performed using a Mosquito robot (TTP Labtech) to lead to crystals in the Index screen (Hampton Research), condition G10: 0.2 M MgCl₂, 0.1 M Bis-Tris, pH 5.5, 25% PEG-3350. However, these were small clusters of needles that were unsuitable for X-ray analysis and step-by-step seeding was used to increase the size. The clusters were crushed using a needle and were used to seed fresh hanging drops set up in similar conditions with a precipitant range of 20–25%. Bigger crystals grew following this procedure, which was repeated until diffraction quality crystals were obtained. Finally, single crystals were obtained in 20% PEG MME 3350, 0.2 M MgCl₂, 0.1 M Bis-Tris, pH 5.5.

For the dUpNpp complex, LmDUT3C protein at 70 mg/ml was mixed with an equal volume of 10 mM dUpNpp to give final concentrations of 35 mg/ml protein and 5 mM ligand. Initial crystallization trials were set up as for the dUpNp complex leading to crystals in the PACT screen (Molecular Dimensions), condition B11: 0.2 M CaCl₂, 0.1 M MES, pH 6.0, 20% PEG-6000. Repeating the condition in larger hanging drops varying the concentration of PEG-6000 from 15–25% yielded quality crystals suitable for data collection.

**Crystallization of L. major dUTPase in Complex with dUMP and dU**—The LmDUT3C protein was used for crystallization of both the dUMP and dU complexes. To form the 2'-dUMP complex, 2.5 μl of a 200 mM nucleotide stock was diluted 20-fold with 80 mg/ml protein to yield a final concentration of 10 mM dUMP and 76 mg/ml protein. To obtain the dU complex, protein was mixed with a 200 mM stock of 2'-dU and a 0.5 mM stock of sodium pyrophosphate (PP₃) to give final concentrations of 10 mM dU, 10 mM PP₃, and 75 mg/ml dUTPase. Initial crystallization trials were set up using a Mosquito robot as for the other complexes. Crystals of both complexes grew in the same condition in the index screen (Hampton Research), condition C6: 0.1 M NaCl, 0.1 M Bis-Tris, pH 6.5, 1.5 M (NH₄)₂SO₄. Larger crystals were then obtained in hanging drop vapor diffusion experiments using the same condition but varying the (NH₄)₂SO₄ concentration over the range of 1.0–2.0 M.
Crystallization of C. jejuni dUTPase in Complex with dUpNpp—Protein was purified as described in Moroz et al. (19). Briefly, recombinant protein was purified in a two-step procedure consisting of a phosphocellulose column followed by size exclusion chromatography. Purified protein was concentrated to 15 mg/ml, and buffer was exchanged into 10 mM HEPES, pH 7.0, with either 5 mM magnesium acetate to produce crystallization screens in the same way as for the L. major dUTPase. Initial microcrystals were obtained in the index screen, condition C1: 3.5 M sodium formate, pH 7.0. These crystals were optimized manually with the addition of 20 mM magnesium acetate to produce diffraction quality crystals.

X-ray Data Collection, Processing, and Structure Determination for C. jejuni dUpNpp Complex—Crystals were cryo-cooled to 120 K in a nitrogen stream from an Oxford Cryosystems device using 3.5 M sodium formate, pH 7.0, 12.5% glycerol as a cryo-protectant. Crystals belonged to the space group I4122, with cell dimensions a = b = 93.4 Å and c = 143.6 Å, and data were collected to 1.7 Å resolution at the European Synchrotron Radiation Facility beamline BM14. There was a single subunit in the asymmetric unit with a V\text{M} of 2.9 Å\textsuperscript{3}/Da and a solvent content of 57%. Data were processed using DENZO and SCALEPACK (31) before structure determination by molecular replacement in MOLREP using the C. jejuni dUpNp complex as the search model. The model was completed manually using the XFIT option of the program QUANTA (Accelrys, San Diego, CA) and refined using REFMAC with the TLS option. The rigid and mobile domains of CjUTPase were used as two separate TLS groups. Processing and refinement statistics are shown in Table 1.

Tryptophan Fluorescence Quenching—Titrations monitoring the tryptophan fluorescence of L. major dUTPase were performed using a Horiba Jobin Yvon Fluoromax-4 spectrofluorometer with a 5-mm path length quartz cuvette. The dUTPase was present in solution at a concentration of 1 μM to which 1 μl of concentrated ligand was added per injection. The solutions were mixed thoroughly and allowed to equilibrate for 1 min before measurements were taken. An excitation wavelength of 295 nm was used, and fluorescence was measured at 330 nm using excitation and emission slit widths of 2 and 3 nm, respectively. The buffer used throughout the experiments was 10 mM HEPES, pH 7.0, with either 5 mM MgCl\textsubscript{2} or 1 mM EDTA. When pyrophosphate was used in the measurements, it was present at a concentration of 1 mM. Data for the dU and dUMP titrations were plotted and analyzed using SigmaPlot in a similar manner to that used by Ortiz-Salmeron et al. (32). No significant inner filter effects were experienced due to the ligands used during the course of this work.

For the dUpNpp titrations where the ligand to protein concentration ratio was lower, data were fitted to the quadratic equation below,

\[ F_{\text{norm}} = \frac{B_m - B_q}{(L_0 + M_0 + K_d) - \sqrt{(L_0 + M_0 + K_d)^2 - 4L_0M_0}} \]

where \( F_{\text{norm}} \) is the normalized protein fluorescence, \( B_m \) is the maximum protein fluorescence, \( B_q \) is the fluorescence change upon ligand binding, \( L_0 \) is the ligand concentration corrected for dilution, \( M_0 \) is the protein concentration corrected for dilution, and \( K_d \) is the dissociation constant for the ligand.

RESULTS

L. major dUTPase Overall Fold—The structure was determined in two different crystal forms. The dUdNp and dUpNp complexes were in space group P6\textsubscript{5}22 with a single protomer in the asymmetric unit (Table 1). Structures of the dUMP and dUTPase complexes were in space group P3\textsubscript{2}1, with two dimers in the asymmetric unit (Table 1). The structures in both space groups are very similar with the protomer in P6\textsubscript{5}22 superimposing with each protomer in P3\textsubscript{2}1 with r.m.s.d. of 0.67, 0.60, 0.78, and 0.65 Å over 257 residues for chains A, B, D, and E, respectively.

The LmdUTPase protomer is made up of 10 α-helices (residues 11–29, 32–35, 38–53, 70–92, 135–150, 154–168, 172–188, 218–222, 224–242, and 252–262) linked by flexible loops (Fig. 1A). Two domains were identified previously in the TcdUTPase structure, a “rigid” domain that mediates the dimerization and a “mobile” domain that undergoes significant domain movement upon substrate binding (18). Similar to CjUTPase, LmdUTPase structures could only be determined in the closed form, but the domains can be defined similarly (Fig. 1A). The rigid domain is formed by residues 32–169 and the mobile domain by residues 8–31 and 170–265. In the TcdUTPase structure, there were two regions that could not be modeled in the mobile domain, residues 94–103 and 119–139 (18). There was good density for the equivalent loops (residues 94–103 and 117–125, respectively) in L. major, so most of the sequence could be modeled into the electron density with the exception of the first eight N-terminal residues in most of the structures and two residues at the very C terminus that are likely to be flexible.

LmdUTPase, similar to the T. cruzi (18) and C. jejuni (19) enzymes, is a dimer (Fig. 1B), with dimerization largely mediated via the rigid domains of the protein. In the P6\textsubscript{5}22 crystals, the dimer is formed around a crystallographic 2-fold axis burying 5530 Å\textsuperscript{2} and 5700 Å\textsuperscript{2} of surface area in the dUdNpp and dUpNp complexes, respectively. In P3\textsubscript{2}1, there are two dimers in the asymmetric unit with 5380 Å\textsuperscript{2} (chains A and B) and 5790 Å\textsuperscript{2} (chains C and D) surface area buried in the dUDP complex and 5790 Å\textsuperscript{2} (chains A and B) and 5500 Å\textsuperscript{2} (chains C and D) surface area buried in the dU complex. In all of the structures, a loop from one subunit reaches over to contact the ligand in the other member of the dimer. The interactions at the dimer interface in the CjUTPase were slightly different to those of the T. cruzi enzyme (18). The interactions at the interface in L. major resemble those of the T. cruzi enzyme with many of the residues here conserved between the two species (Fig. 1C).

CjUTPase dUpNpp Complex Structure—We sought to obtain further information on binding of the triphosphate moiety by determining the CjUTPase dUpNp complex, following our earlier study of the closed structure of the dUpNp complex (19). The
structure in complex with the triphosphate substrate analogue determined here belonged to a different crystal form (Table 1) with a single subunit in the asymmetric unit, the dimer being formed about a crystallographic 2-fold symmetry axis. The protomer in the Cj dUTPase dUpNpp complex can be superimposed onto that of the diphosphate analogue complex with an r.m.s.d. of 0.67Å over 223-carbons showing that the overall fold of the protein is essentially identical in the two complexes. The biggest difference between the two complexes is the presence of only two Mg2+ ions in the dUpNpp active site compared with three in the dUpNp complex. This will be discussed in more detail below.

Protein Ligand Interactions—Structures of Lm dUTPase were determined with four ligands bound (dUpNpp, dUpNp, dUMP, and dU). The interactions formed between the five sequence motifs of this family of enzymes and their substrates have been described previously (18, 19), and so will not be discussed in detail here. The five motifs form the active site in the same way as in the closed TcdUTPase structure (Fig. 2A). The Lm dUTPase and Cj dUTPase dUpNpp complexes are the first structures for this family of enzymes to contain the triphosphate moiety. Both reveal that there are surprisingly few direct hydrogen bonds between the γ-phosphate of the substrate and the protein. Indeed, the only residue within hydrogen bonding distance of the γ-phosphate in Lm dUTPase is Glu51, which is also involved in coordinating the catalytically important metal ions. The γ-phosphate makes two contacts in Cj dUTPase hydrogen bonding to Asn202 and Lys194. These differences do not appear to be due to movements in side chain positions as these are well conserved between the two species but rather due to a different positioning of the γ-phosphate of the substrate (Fig. 2B). Therefore, it would appear that there is degree of flexibility in the positioning of the γ-phosphate in the active site. This is additionally supported by the fact that the B-factors for the γ-phosphate atoms are higher than for the rest of the ligand, indicating a substantial degree of mobility.

The only other difference in ligand binding between Cj dUTPase and Lm dUTPase is at the bottom of the active site pocket adjacent to the O4 group of the uracil ring. In the Cj dUTPase complex, the O4 group hydrogen-bonds to the Nε2 group of Asn21. The equivalent residue in L. major is Asp29, which has its side chain pointing away from the active site. This side chain is involved in a hydrogen-bonding network through water molecules to the O4 group of the uracil ring, resulting in the presence of a small water filled cavity at the bottom of the pocket, which is also seen in the TcdUTPase. This feature might be exploitable in the development of new inhibitors of the kinetoplastid enzymes.

Sulfate Ions in Active Site—The aim of obtaining the dU- and dUMP-bound Lm dUTPase crystal structures was to determine whether the phosphate groups were required for closure of the enzyme upon binding. These complexes reveal that the interactions between the protein and the uracil and deoxyribose moieties are identical when these ligands are bound compared with dUpNpp and dUpNp. However, large tetrahedral electron density peaks were observed, which were interpreted as SO42− ions from the crystallization medium (Fig. 2, C and D).
In both the dUMP and dU complexes, there was electron density for a tetrahedral group adjacent to the 5’ end of the ligand. These were modeled as sulfate ions (Fig. 2, C and D), based on the presence of ammonium sulfate as the crystallization precipitant. In the dUMP complex, a single sulfate was modeled into each of the four protomers. These sulfate ions

FIGURE 1. A, structure of the LmdUTPase protomer. Secondary structures are numbered in order from the N to C terminus. The rigid domain is colored green, and the mobile domain is colored blue. Orange spheres represent the Ca2+ ions present in the active site adjacent to the dUpNpp ligand. (Sticks are colored by atom type.) B, the LmdUTPase dUpNpp complex dimer formed about a crystallographic 2-fold. C, structure-based sequence alignment of Lm-, Tc- and CjUTPases. Residues that interact across the dimer interface are highlighted with green stars, orange squares, and red circles in Lm-, Tc-, and CjUTPases, respectively, illustrating that the dimer interface in LmdUTPase resembles more closely the interface in T. cruzi than in C. jejuni. All molecular graphics figures were generated using CCP4mg (35), and sequence alignments were made using ALINE (36).
hydrogen-bond to the side chains of Lys59, Lys62, and Asn63, except in chain A, where the sulfate is positioned slightly differently and hydrogen-bonds to the side chain of Lys62, and the main chain amide groups of Lys198 and Val199; all of which are conserved among the kinetoplastid dimeric dUTPases (supplemental Fig. 1). Superposition of the various chains reveals that the loop formed by residues 198–205 occupies a different position in chain A, leading to this difference in positioning of the sulfate ion in this chain, a result of crystal packing as this loop mediates interactions with chain D of a symmetry related molecule (supplemental Fig. 2).

In the dU co-crystallizations, PPi was added, as this was expected to aid binding of the dU moiety. Additional electron density present in the active site could not be modeled as PPi and was again interpreted as sulfate ions (Fig. 2D). There were more sulfates than in the dUMP complex. Chains B, D, and E have two sulfates bound. One occupies a position intermediate to that taken up by the α- and β-phosphate groups of the longer substrate. This hydrogen bonds to a water molecule bridging it to the 5′-OH group of the dU moiety and the sidechains of Lys59, Trp61, Lys62, Tyr191, and Lys198. The second sulfate binds in a similar position to that taken up by those in the equivalent chains of the dUMP complex. As in the dUMP complex, chain A differs from the other three protomers as it has three sulfate ions bound. Two of these are equivalent to those in the other chains, with the third occupying a similar position to that taken
up by the sulfate in chain A of the dUMP complex, which
hydrogen-bonds to the main chain amide groups of Lys198 and
Val199 and the side chain of Lys42.

Identification of Metal Binding Sites—Use of non-hydrolyzable substrate analogues resulted in the binding of metal ions in the active sites. As the equivalent CjdTUTPase complex (19), the LmdUTPase dUpNp complex contains three metal ions adjacent to the phosphate tail of the substrate. These were assigned as Mg2+ based on the octahedral coordination geometry and bond lengths (33), and are coordinated by the side chains of four conserved acidic residues, Glu48 and Glu76 from motif II and Glu76 and Asp79 from motif IV. The phosphate groups of the dUpNp and water molecules provide the remaining metal coordination.

Both the Lm- and CjdTUTPases dUpNpp triphosphate analogue complexes surprisingly contained only two metal ions. The two metals in CjdTUTPase were again identified as Mg2+ on the basis of octahedral coordination geometry and bond lengths. There was an electron density peak ~1 Å from the third metal binding site in the dUpNp complex in both species, but it did not have coordination geometry consistent with a metal ion and was thus interpreted as a water molecule (Fig. 3, A and B). In the LmdUTPase structure, the metal ions showed octahedral coordination geometry, but the bond lengths were slightly longer and were assigned as Ca2+ from the crystallization solution. Calculation of anomalous difference Fourier maps revealed two strong electron density peaks at these positions, confirming them as Ca2+ with no peak present that could represent a third calcium (supplemental Fig. 3). It would therefore appear that the presence of the γ-phosphate displaces the third metal ion from the active site.

A single Mg2+ ion was modeled into the active sites of three of the four protomers in the asymmetric unit in the dUMP complex. The electron density for these was somewhat weak, and the bond lengths were longer than would be expected for an Mg2+ ion. However, superposition of the dUMP bound structure with the dUpNp-bound structure shows that the Mg2+ ion is flipped away from the active site in the presence of the triphosphate substrate analogue dUpNp complex. The Mg2+ ion was positioned in line with the scissile bond, leading to the suggestion that this was a likely nucleophile with the reaction proceeding by a similar two ion mechanism to that proposed for the phosphoryl transfer reactions performed by DNA polymerases and many other enzymes (20, 21).

We describe here the structure of L. major dUTPase in complex with two substrate analogues dUpNpp and dUpNp (Fig. 3A). The positions of essentially all of the ligand atoms are close to identical in both complexes. There is a small rotation in the position of the β-phosphate. More importantly, there are only two metal ions (Ca2+) in the dUpNp complex. In the dUpNpp complex, the γ-phosphate causes a 0.8 Å movement of Glu51, distorting the active site slightly, which is perhaps related to the loss of the third metal ion. The two structures are both in keeping with the previously proposed mechanism, with the catalytic water molecule positioned in the same place in both structures, but 0.2/0.4 Å from that in the C. jejuni dUpNp complex (Fig. 3C) (19).

It is curious that in the presence of the triphosphate substrate analogue, only two metal ions are bound in the active site. The C. jejuni dUpNp complex reported here also only has two metal ions, although these are both Mg2+ ions. There is a more dramatic difference between the C. jejuni dUpNpp and dUpNp complexes, where the side chain of Glu49 (equivalent to Glu51 in LmdUTPase) is flipped away from the active site in the presence of the triphosphate, such that it no longer coordinates any metal ions (Fig. 3B). This is quite different to its position in the C. jejuni dUpNp complex, and indeed, the two Leishmania complexes, where a much smaller movement of Glu51 was observed between the equivalent structures. The 0.8 Å shift in the position of Glu51 in the LmdUTPase complex could result in catalyzing that it must play a role in dU binding, even though it was not seen in the crystal structure. The presence of pyrophosphate had a significant influence on the affinity of the enzyme for dU, reducing the dissociation constant from an immeasurably high value to 140 μM. In contrast, pyrophosphate had very little influence upon dUMP binding, with closely similar Kd values of 7.3 μM without PPi, and 17 μM in its presence. This suggests that the α-phosphate plays a major role in inducing closure of the enzyme, and from the point of view of inhibitor design, the presence of a negatively charged moiety, which would interact with the phosphate binding region of the protein, will be required to lock the protein into the higher affinity closed conformation.

**DISCUSSION**

**Active Site and a Likely Mechanism of Action**—The active sites of all of the dimeric dUTPases determined to date are very similar. The T. cruzi dUDP complex lacked metals in the active site, and so a mechanism of reaction could not be proposed (18). The subsequent structure of the C. jejuni enzyme bound to the diphosphate substrate analogue dUpNp revealed the presence of three Mg2+ ions adjacent to the phosphate groups of the substrate (19). Metals are clearly important for substrate binding (17), supported by the tryptophan fluorescence quenching experiments on the L. major enzyme (Fig. 4) and data reported previously (34). In the C. jejuni dUpNp complex (19), one of the water molecules coordinating the central Mg2+ ion was positioned in line with the scissile bond, leading to the suggestion that this was a likely nucleophile with the reaction proceeding by a similar two ion mechanism to that proposed for the phosphoryl transfer reactions performed by DNA polymerases and many other enzymes (20, 21).

Representative binding curves are shown in Fig. 4. As would be expected, the affinity of the protein ligand interactions follows the pattern dUpNpp > dUMP > dU in the absence of pyrophosphate and in the presence of Mg2+. Binding of all of the ligands is weakened severely in the absence of Mg2+, indicating that it must play a role in dU binding, even though it was not seen in the crystal structure. The presence of pyrophosphate had a significant influence on the affinity of the enzyme for dU, reducing the dissociation constant from an immeasurably high value to 140 μM. In contrast, pyrophosphate had very little influence upon dUMP binding, with closely similar Kd values of 7.3 μM without PPi, and 17 μM in its presence. This suggests that the α-phosphate plays a major role in inducing closure of the enzyme, and from the point of view of inhibitor design, the presence of a negatively charged moiety, which would interact with the phosphate binding region of the protein, will be required to lock the protein into the higher affinity closed conformation.
from the presence of Ca$^{2+}$ rather than Mg$^{2+}$ in this structure, with the longer coordinating bond lengths causing the loss of the third metal site. However, the loss of the third metal site in the presence of the triphosphate substrate analogue is common to both enzymes, and so this appears to be a general feature of this family of enzymes. In addition, the presumed nucleophilic water is displaced by 1.2/1.5 Å from that observed in the two *Leishmania* structures.

The differing metal-binding properties depending upon the identity of the substrate may be necessary for the correct positioning of the nucleophile for attack on the scissile bond. In the presence of the diphosphate substrate, three metals are bound with the proposed catalytic water molecule coordinated by metals 2 and 3, aligning it for an in-line attack on the α-β-phosphate bond. In the presence of the triphosphate, the third metal is lost, but the proposed nucleophile is...
still present, albeit in a slightly different position to that in the diphosphate complex, ~0.3 Å away in the Leishmania and ~1.3 Å in the C. jejuni enzyme. The third metal ion may no longer be required, as the $\gamma$-phosphate itself aids in aligning the nucleophile for an in-line attack on the scissile bond. This certainly seems to hold for the C. jejuni enzyme; however, in the LmdUTPase dUpNpp complex, the third phosphate group occupies a slightly different position to that taken up in the C. jejuni enzyme and is too far away from the proposed catalytic water to interact with it.

From an inhibitor design point of view, there is little space in or around the active site in the closed form of the enzyme to allow binding of substrate derivatives (Fig. 3D). One avenue that may be exploited is at the bottom of the ligand binding pocket. In the C. jejuni enzyme, there is an Asn residue that hydrogen-bonds to the O4 group of the uracil ring. In the kinetoplastids, this residue is replaced by an Asp, which is rotated away from the active site with the hydrogen bond to the O4 group of the substrate provided by a water molecule in both the L. major structures presented here and the T. cruzi dUDP complex. Thus, this deeper cavity could be exploited by a modification to the O4 group of the uracil ring. Another site for modification could be the water-occupied pocket adjacent to the deoxyribose moiety of the substrate (Fig. 3D). However, both of these positions pose certain difficulties in terms of synthetic chemistry. In principle, there is the possibility of modifying the phosphate moiety, but this was rather tightly packed against the surrounding protein.

**Importance of Phosphate Groups for Enzyme Closure**—Dimeric dUTPases represent a potential drug target and a better understanding of the requirements for enzyme closure will clearly aid in the design of new inhibitors. The structures of the LmdUTPase in complex with the substrate fragment dU revealed a completely closed enzyme with sulfate ions from the crystallization liquor present in and near the active site. A single sulfate was observed in the dUMP complex, binding at a distinct site from that taken up by the other phosphate.
groups of the larger substrate analogues. These sulfate ions are positioned 5–6.5 Å away from the phosphate group of the dUMP molecule and are bound by residues that are conserved among the dimeric dUTPases of the kinetoplastids (supplemental Fig. 1). This positioning may represent the site occupied by the pyrophosphate moiety following the reaction.

Sulfates occupy the same sites in the dU as in the dUMP complex with the exception of an additional sulfate, which binds in a position intermediate to that taken up by the α- and β-phosphate groups of the larger dUpNp and dUpNpp ligands (Fig. 2D). The presence of the sulfate ion here suggested that occupation of the phosphate binding region by a negatively charged group is likely to be important to induce closure of the enzyme. Tryptophan fluorescence quenching with the LmddTUTPase revealed that the presence of pyrophosphate had a marked influence on the dissociation constant for dU reducing the Kd from an immeasurable number to 140 μM (Fig. 4C). In contrast, the effect of pyrophosphate on the affinity for dUMP was negligible suggesting that a single phosphate group is sufficient to induce closure of the enzyme, as would be expected given that dUMP is an inhibitor of the dimeric dUTPases. Inspection of the interactions between the protein and dUpNpp reveals that of the 13 hydrogen bonds between protein and ligand, four are directed toward the α-phosphate group, three toward the β-phosphate, and none toward the γ-phosphate, with the remainder involved in binding to the uracil ring and deoxyribose moieties. Further contribution comes from the active site metals, one of which coordinates to the α-phosphate group, and this region does indeed appear to be very important for substrate binding. It would therefore appear that any compounds that are likely to induce closure and inhibit dimeric dUTPases must bear a suitable negative charge that would enable them to interact efficiently with the phosphate-binding region.

Conclusions—The fact that structures of the dUTPases from trypanosomes are vastly different from the human trimeric enzymes makes them an attractive drug target for the development of new drugs against the diseases caused by these organisms. We have determined structures of the enzyme from L. major in complex with substrate analogues (dUpNpp and dUpNp), the product dUMP, and a substrate fragment (dU). The resulting information reveals a very tight ligand-binding pocket in this family of dUTPases with little room for modifications of the uracil and ribose rings without perturbation of binding features. The structures in complex with dUMP and dU revealed the presence of sulfate ions in the active site. In the dU complex, a sulfate was observed binding in a position between that taken up by the α- and β-phosphates of the larger substrate analogues, and, in line with this observation, binding data showed that affinity for dU was dependent on PPi. Taken together, the structural and biophysical evidence indicates that the presence of a single phosphate or similarly charged group is important to induce closure and formation of a high affinity conformation of the enzyme. Work is ongoing to try and develop new inhibitors of these enzymes that exploit these observations and are amenable to use as anti-trypanosomal.

Acknowledgments—We thank Dr. Andrew Leech for assistance with the tryptophan fluorescence quenching data analysis. We acknowledge the Diamond Light Source (beamline IO4) and the European Synchrotron Radiation Facility (beamlines ID29, ID14.2, and BM14) for provision of x-ray data collection facilities. We thank Johan Turkenburg and Sam Hart for assistance with data collection.

REFERENCES
1. Desjeux, P. (1996) Clin. Dermatol. 14, 417–423
2. Shlomai, J., and Kornberg, A. (1978) J. Biol. Chem. 253, 3305–3312
3. Pearl, L. H., and Savva, R. (1996) Nat. Struct. Biol. 3, 485–487
4. el-Hajj, H. H., Zhang, H., and Weiss, B. (1988) J. Bacteriol. 170, 1069–1075
5. Gaedelen, M. H., McIntosh, E. M., Game, J. C., Wilson, P. J., and Haynes, R. H. (1993) EMBO J. 12, 4425–4431
6. Hidalgo-Zarco, F., and González-Pazanowska, D. (2001) Curr. Protein Pept. Sci. 2, 389–397
7. Castillo-Acosta, V. M., Estévez, A. M., Vidal, A. E., Ruiz-Perez, L. M., and González-Pacanowska, D. (2002) Int. J. Biochem. Cell Biol. 40, 2901–2913
8. Persson, R., Cedergren-Zeppezauer, E. S., and Wilson, K. S. (2001) Curr. Protein Pept. Sci. 2, 287–300
9. Barabás, Ö., Pongrácz, V., Kovári, J., Wilmanns, M., and Vertessy, B. G. (2004) J. Biol. Chem. 279, 42907–42915
10. Dauter, Z., Persson, R., Rosengren, A. M., Nyman, P. O., Wilson, K. S., and Cedergren-Zeppezauer, E. S. (1999) J. Mol. Biol. 285, 655–673
11. Kovári, J., Barabás, Ö., Takács, E., Békési, A., Dubrovay, Z., Pongrácz, V., Zágyva, I., Imre, T., Szabó, P., and Vertessy, B. G. (2004) J. Biol. Chem. 279, 17932–17944
12. Mustafi, D., Bekesi, A., Vertessy, B. G., and Makinen, M. W. (2003) Proc. Natl. Acad. Sci. U.S.A. 100, 5670–5675
13. McGeoch, D. J. (1990) Nucleic Acids Res. 18, 4105–4110
14. Galperin, M. Y., Moroz, O. V., Wilson, K. S., and Murzin, A. G. (2006) Mol. Microbiol. 59, 5–19
15. Hidalgo-Zarco, F., Camacho, A. G., Bernier-Villamor, V., Nord, J., Ruiz-Pérez, L. M., and González-Pacanowska, D. (2001) Protein Sci. 10, 1426–1433
16. Larsson, G., Nyman, P. O., and Kvassman, J. O. (1996) J. Biol. Chem. 271, 24010–24016
17. Camacho, A., Hidalgo-Zarco, F., Bernier-Villamor, V., Ruiz-Pérez, L. M., and González-Pacanowska, D. (2000) Biochem. J. 346, 163–168
18. Harkiolaki, M., Dodson, E. J., Bernier-Villamor, V., Turkenburg, J. P., González-Pacanowska, D., and Wilson, K. S. (2004) Structure 12, 41–53
19. Moroz, O. V., Harkiolaki, M., Galperin, M. Y., Vagin, A. A., González-Pacanowska, D., and Wilson, K. S. (2004) J. Mol. Biol. 342, 1583–1597
20. Reese, L. S., and Steitz, T. A. (1991) EMBO J. 10, 25–33
21. Steitz, T. A. (1993) Curr. Opin. Struct. Biol. 3, 31–38
22. Fogg, M. J., and Wilkinson, A. J. (2008) Biochem. Soc. Trans. 36, 771–775
23. Vagin, A., and Teplyakov, A. (1997) J. Appl. Crystallogr. 30, 1022–1025
24. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. D Biol. Crystallogr. 53, 240–255
25. Lamzin, V. S., and Wilson, K. S. (1993) Acta Crystallogr. D Biol. Crystallogr. 49, 129–147
26. Emsley, P., and Cowtan, K. (2004) Acta Crystallogr D Biol. Crystallogr. 60, 2126–2132
27. Leslie, A. G. (1999) Acta Crystallogr D Biol. Crystallogr. 55, 1696–1702
28. CCP4 (1994) Acta Crystallogr. D 50, 760–763
29. Painter, J., and Merritt, E. A. (2006) Acta Crystallogr. D Biol. Crystallogr. 62, 439–450
30. Davis, I. W., Lever-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W.,arendall, W. B., 3rd, Snoeyink, J., Richardson, J. S., and Richardson, D. C. (2007) Nucleic Acids Res. 35, W375–383
31. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
Crystal Structures of *L. major* dUTPase Complexes

32. Ortiz-Salmerón, E., Yassin, Z., Clemente-Jimenez, M. J., Las Heras-Vazquez, F. J., Rodríguez-Vico, F., Barón, C., and García-Fuentes, L. (2001) *Eur. J. Biochem.* **268**, 4307–4314
33. Harding, M. M. (2001) *Acta Crystallogr. D Biol. Crystallogr.* **57**, 401–411
34. Téllez-Sanz, R., Yassin, Z., Bernier-Villamor, V., Ortiz-Salmerón, E., Musso-Buendia, J. A., Barón, C., Ruíz-Pérez, L. M., González-Pacanowska, D., and García-Fuentes, L. (2007) *Biochimie* **89**, 972–980
35. Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G. N., Cohen, S., Perrakis, A., and Noble, M. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2288–2294
36. Bond, C. S., and Schüttelkopf, A. W. (2009) *Acta Crystallogr. D Biol. Crystallogr.* **65**, 510–512