Structural basis for UTP specificity of RNA editing TUTases from *Trypanosoma brucei*

Junpeng Deng^{1,2,3}, Nancy Lewis Ernst⁴,
Stewart Turley^{1,2,3}, Kenneth D Stuart^{4,5}
and Wim GJ Hol^{1,2,3,*}

¹Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA, ²Department of Biochemistry, Biomolecular Structure Center, University of Washington, Seattle, WA, USA, ³Department of Biological Structure, Biomolecular Structure Center, University of Washington, Seattle, WA, USA, ⁴Seattle Biomedical Research Institute, Seattle, WA, USA and ⁵Department of Pathobiology, University of Washington, Seattle, WA, USA

Trypanosomatids are pathogenic protozoa that undergo a unique form of post-transcriptional RNA editing that inserts or deletes uridine nucleotides in many mitochondrial pre-mRNAs. Editing is catalyzed by a large multiprotein complex, the editosome. A key editosome enzyme, RNA editing terminal uridylyl transferase 2 (TUTase 2; RET2) catalyzes the uridylyl addition reaction. Here, we report the 1.8 Å crystal structure of the *Trypanosoma brucei* RET2 apoenzyme and its complexes with uridine nucleotides. This structure reveals that the specificity of the TUTase for UTP is determined by a crucial water molecule that is exquisitely positioned by the conserved carboxylates D421 and E424 to sense a hydrogen atom on the N3 position of the uridine base. The three-domain structure also unveils a unique domain arrangement not seen before in the nucleotidyltransferase superfamily, with a large domain insertion between the catalytic aspartates. This insertion is present in all trypanosomatid TUTases. We also show that *TbRET2* is essential for survival of the bloodstream form of the parasite and therefore is a potential target for drug therapy.

The EMBO Journal (2005) **24**, 4007–4017. doi:10.1038/sj.emboj.7600861; Published online 10 November 2005

Subject Categories: RNA; structural biology

Keywords: kinetoplastid RNA editing; nucleotidyl transferases; poly (A) polymerase; TUTase

Introduction

Trypanosomatids are the causative agents of a number of diseases such as African sleeping sickness, Chagas disease and leishmaniasis that together cause about one million deaths per year (World Health Organization, 2003). Currently, drug therapy for these diseases is inadequate and often is toxic. These organisms transform many of their mitochondrial pre-mRNAs into translatable mature mRNAs

by extensive post-transcriptional RNA editing (Stuart *et al*, 1997, 2002; Madison-Antenucci *et al*, 2002; Simpson *et al*, 2003). This unique process is accomplished by insertion and deletion of uridine nucleotides (U) in the pre-mRNA that are specified by the precise sequence of small mitochondrially encoded guide RNAs (gRNAs). Each gRNA consists of three regions: (i) a 5' anchor region that can duplex with its cognate pre-mRNA via base-pairing; (ii) an adjacent central region that specifies where Us are to be inserted or deleted; and (iii) a 3' oligo (U) tail that may provide for interactions between the gRNA, pre-mRNA and the editing enzyme complex. Multiple gRNAs are used to edit an mRNA with each gRNA specifying the editing of several sites that in many cases creates new sequences that are complementary to the next gRNA that is used. Strikingly, in some cases, editing can account for more than half of the sequence of the mature mRNA (Stuart *et al*, 1997).

The editing of each site occurs by a coordinated series of enzymatic steps. The process begins with the formation of an 'anchor' duplex between the gRNA and the pre-mRNA. The pre-mRNA is then cleaved by an endonuclease at the editing site, apparently at a mismatch between the pre-mRNA and the gRNA. Us are then either added to or removed from the pre-mRNA by a 3'-terminal uridylyl transferase (TUTase) or a U-specific exoribonuclease (exoUase), respectively. The final step of editing is the religation of the processed RNA fragments by RNA ligases. The RNA editing enzymes in *Trypanosoma brucei* are organized into a ~1.6 MDa multiprotein complex, the editosome, containing ~20 different proteins (Rusche *et al*, 1997; Madison-Antenucci and Hajduk, 2001; Panigrahi *et al*, 2001a,b, 2003; Aphasizhev *et al*, 2003a). To date, two TUTases (Aphasizhev *et al*, 2002, 2003b; Ernst *et al*, 2003), two ExoUases (Schnauffer *et al*, 2003; Kang *et al*, 2005; NL Ernst *et al*, unpublished) and two editing RNA ligases have been reported (Huang *et al*, 2001; McManus *et al*, 2001; Rusche *et al*, 2001; Schnauffer *et al*, 2001; Gao and Simpson, 2003). In addition, deletion (Trotter *et al*, 2005) and insertion (Carnes *et al*, 2005) endonucleases have been identified. The crystal structure of the catalytic domain of RNA editing ligase 1 (REL1) from *T. brucei* has recently been described (Deng *et al*, 2004b), representing the first crystal structure of a kinetoplastid editosome protein.

The two TUTases that are involved in RNA editing in trypanosomes (Aphasizhev *et al*, 2002, 2004; Ernst *et al*, 2003) differ significantly in many aspects such as polypeptide size, RNA substrate specificity, processivity and quaternary structure. RNA editing TUTase 1 (RET1) from *Leishmania tarentolae* was shown to be a tetramer with 121 kDa subunits and has been implicated in the addition of the nonencoded 3'-oligo (U) tail to gRNAs (Aphasizhev *et al*, 2003b) and in RNA turnover in trypanosome mitochondria (Ryan and Read, 2005). RET1 is not a component of the editosome, but appears to be in a separate complex that can interact with the editosome via RNA (Aphasizhev *et al*, 2002). RNA editing TUTase 2 (RET2) from *T. brucei* is a 57 kDa monomeric

*Corresponding author. Department of Biochemistry, Biomolecular Structure Center, University of Washington, Seattle, WA 98195, USA. Tel.: +1 206 685 7044; Fax: +1 206 685 7002; E-mail: wghol@u.washington.edu

Received: 2 August 2005; accepted: 10 October 2005; published online: 10 November 2005

protein and is a key integral component of the editosome (Ernst *et al*, 2003). RET2 interacts directly with another editosome protein, KREPA1 (previously known as *TbMP81*), and its activity is enhanced by this interaction (Ernst *et al*, 2003). *TbRET2* primarily adds a single U to single-stranded (ss)RNA, and adds the number of Us specified by the gRNA to precleaved double-stranded (ds)RNA editing substrates. In contrast, *TbRET1* adds hundreds of Us to an ssRNA primer *in vitro* and does not precisely add the correct number of Us to dsRNA (Aphasizhev *et al*, 2002). Despite these differences, the two TUTases share substantial homology between the overlapping regions, with 24% sequence identity for a 554-amino-acid region of overlap between the proteins according to ClustalW (Ernst *et al*, 2003).

Both RET1 and RET2 are essential for editing and survival of the insect stage of the parasite (Aphasizhev *et al*, 2003b). Although TUTase activity is present in a number of organisms (Aphasizhev *et al*, 2004), no homologs of the trypanosomatid TUTases have been found in other organisms. PFAM database searches place the TUTases within the nucleotidyltransferase (NT) superfamily that has the signature motif G[G/S]X₉-13DX[D/E] (Holm and Sander, 1995), with usually three, or in the case of some class II CCA-adding enzymes (Yue *et al*, 1998), two metal-coordinating invariant carboxylates that are crucial for activity. The NT family has been divided into two classes based on their sequence similarities (Yue *et al*, 1996; Aravind and Koonin, 1999). The class I subfamily includes the archaeal CCA-adding enzyme, eukaryotic poly(A) polymerase, DNA polymerase β , terminal deoxynucleotidyl transferase and kanamycin NT. Class II enzymes include eubacterial and eukaryotic CCA-adding enzymes and eubacterial PAPs. Crystal structures from the two classes of enzymes demonstrate that they all share a homologous polymerase domain and contain a variety of additional domains with different structures and functions (Sakon *et al*, 1993; Pelletier *et al*, 1994; Bard *et al*, 2000; Martin *et al*, 2000; Delarue *et al*, 2002; Li *et al*, 2002; Augustin *et al*, 2003; Okabe *et al*, 2003; Xiong *et al*, 2003; Tomita *et al*, 2004; Xiong and Steitz, 2004). They all share the same nucleotidyl transfer mechanism: a two-metal-coordinated in-line attack by the 3'-OH of the RNA or DNA primer on the α -phosphate of the incoming NTP/dNTP, incorporating the monophosphate moiety into the 3'-end of the primer and releasing pyrophosphate without forming a covalent enzyme-nucleotide intermediate. Nevertheless, these enzymes are all different with respect to the substrates recognized.

Sequence analysis of the trypanosomatid TUTases indicates that these two enzymes have the conserved signature of the NT superfamily, except that they have a large insertion of more than 100 amino acids between the conserved catalytic aspartate residues (Aphasizhev *et al*, 2002; Worthey *et al*, 2003) (Figure 1). Given the low sequence identities of other NT family members, with less than 14% even in the catalytic domain, the correct topology remained elusive. Here, we report the crystal structures of the essential RNA editing enzyme *TbRET2*, in complex with UTP, UMP and divalent metal ions. The structures reveal several conserved amino acids that form a potential 'TUTase signature', which allows for selectivity of UTP ribonucleotides. A crucial water molecule that interacts with the carboxylates of two conserved residues is the key factor in determining uridine (U)-specificity. We also show that *TbRET2* is essential for survival of the

bloodstream form of the parasite. As no close homologs are found among human host proteins, *TbRET2* appears to be an excellent potential drug target. The structures not only increase our insight into the architectural and evolutionary relationships with other members of the NT superfamily but also provide a platform to better understand the enzymatic mechanism, which may also serve as a basis for drug design.

Results and discussion

TbRET2 is essential for the growth of bloodstream form of trypanosomes

The importance of RET2 for the survival of bloodstream form trypanosomes was assessed with a cell line in which both endogenous RET2 alleles were replaced in the presence of a tetracycline-inducible ectopic RET2 allele using an established targeted gene replacement strategy (Wirtz *et al*, 1999; Schnauffer *et al*, 2001; and Supplementary data). In these cells, RET2 expression is dependent on the presence of tetracycline (tet) (Figure 2A). Western analysis of cell lysates in which RET2 was expressed in the presence of tet or repressed for 3 days in the absence of tet using a RET2 polyclonal antiserum showed that RET2 expression was essentially abolished in the absence of tet (Figure 2). The growth of the cells in the absence of tet was slightly reduced after 2–3 days and ceased after 4 days. A few cells were able to survive in the absence of tet and began to proliferate after about 8 days. These cells are likely revertants that contain mutations in the tet-regulated PARP promoter or have reduced expression of the tet repressor resulting in re-expression of RET2 (Chen *et al*, 2003). These results indicate that RET2 is essential for the growth of bloodstream form trypanosomes. This is consistent with the fact that we were unable to generate a RET2 null mutant by direct replacement of both RET2 endogenous alleles (data not shown).

The structure of the *TbRET2*

Wild-type recombinant *TbRET2* was soluble and active (Figure 2B), but was not stable in solution and did not produce crystals. A triple mutant of *TbRET2* (W94R/W122R/W207R) that had mutations of three potentially surface-exposed tryptophan residues to arginines (see also the Supplementary data) showed no altered activity compared to the wild-type protein (Figure 2B). Both the wild-type and triple mutant *TbRET2* proteins were assayed for the addition of nucleotides to 5'CL18, a short ssRNA and a precleaved RNA editing substrate that specifies the addition of two Us (Igo *et al*, 2000). Both proteins were specific for UTP and primarily added a single U to 5'CL18 but also yielded products with three and five Us added, as has been observed previously with recombinant *TbRET2* protein (Ernst *et al*, 2003). Both proteins added two Us to a precleaved editing substrate as specified by the gRNA.

The *TbRET2* triple mutant successfully produced high-quality crystals. The crystal structures of apo *TbRET2* and its complexes with UTP, UMP and Mn^{2+} or Mg^{2+} have been determined at resolutions between 1.8 and 2.1 Å. The apo structure was initially solved by the single wavelength anomalous dispersion method using data from a crystal of SeMet-containing *TbRET2* triple mutant protein. Although 2 mM UTP and 10 mM Mg^{2+} were present in all the protein solutions, there was no UTP ligand found in the apo crystal

Figure 1 Multiple sequence alignment of selected terminal uridylyl transferases. Depicted are RET2 and RET1 from *T. brucei*, *L. major*, RET2 from *T. cruzi* and poly(A) polymerase from *Saccharomyces cerevisiae* (PDB code: 1fa0). The red secondary structure symbols on top represent those observed in the TbRET2 structure. Those at the bottom, in blue, are from the poly(A) polymerase structure. The NTD, MD and CTD are shown as green, yellow and blue bars above the secondary structure elements, respectively. The unique MD insertion in trypanosomatid TUTases is indicated by a yellow-shaded box. The residues involved in UTP binding at site A are shown with red asterisks. The three conserved aspartates, D97, D99 and D267, are indicated by red inverted triangles and labeled. The figure was generated by the program ESPrnt (Gouet et al, 1999).

structure, presumably due to the high salt concentration of the crystallization conditions (for details see Materials and methods). After crosslinking with glutaraldehyde, the apo crystals were successfully transferred into a PEG-containing solution with low salt concentration, without disrupting the crystals. The structures of the nucleotide complexes were obtained by subsequent soaking. Unbiased $F_o - F_c$ difference Fourier methods before the ligands were included in the model showed that the procedure was successful (see below). All structures had good geometry and refinement statistics (Supplementary Table I). In all, 10 residues from the N-terminus as well as 14 residues from the C-terminus were disordered and not incorporated into the structures.

The TbRET2 structure consists of three domains: the N-terminal domain (NTD), the middle domain (MD) and the C-terminal domain (CTD) (Figure 3A). These domains are arranged in a manner that has not been seen before in nucleotidyl transferases. The molecule displays a wedge-shaped architecture and is ~80 Å in length and ~50 Å in width at the blunt end. The noncontiguous NTD (residues 53–152 and 263–272) contains a five-stranded antiparallel β -sheet flanked by three α -helices. The contiguous MD is

a unique structural insertion when compared to all other members of the NT superfamily and contains 110-amino-acid residues (residues 153–262) that are inserted between the two NTD β strands (β 4 and β 9) via two short flexible loops and corresponds to the insertion between the second and third conserved catalytic aspartates (Figure 1). This MD domain is mainly composed of six helices and a four-stranded antiparallel β -sheet. The noncontiguous CTD contains seven helices and four short β strands, spanning residues 273–473 in addition to the N-terminal α 1 (residues 30–52). There is a wide and deep cleft between the NTD and CTD (Figure 3A), formed by the antiparallel β -sheet of the NTD and the two helices α 13 and α 14 in the CTD. Three aspartic acids in the NTD are clustered in the cleft, forming a catalytic aspartate triad, the core of the active site of the enzyme. The two conserved NT motif aspartates are D97, situated just prior to strand β 2, and D99 in β 2. The third catalytically essential aspartate D267 (Aphasizheva et al, 2004) is located in β 9, immediately following the MD insertion. The three domains are organized in an extended linear architecture with extensive interactions between the domains. The NTD and CTD share a large interface burying 2923 Å² solvent-accessible

Figure 2 *TbRET2* expression is essential for the growth of bloodstream form trypanosomes. (A) *TbRET2* conditional knockout cells were generated in which both endogenous alleles were replaced in the presence of a tetracycline-regulated ecotopic RET2 allele (see Materials and methods for details). Duplicate cultures of cells were grown in the presence (RET2 expressed, filled line) or absence (RET2 repressed, dashed line) of tetracycline. The cumulative cell number was calculated by multiplying the number of cells by the dilution. Inset: Western blot of glycerol gradient purified lysates of cells in which *TbRET2* was expressed (E) or repressed for 3 days (R) using *TbRET2* antiserum. *TbRET2* is indicated by the arrow. The upper band is a protein from the whole-cell lysates that crossreacts with the polyclonal antiserum. (B) Recombinant *TbRET2* protein activity. Wild-type (WT) and triple mutant (W94R, W1222R and W207R) recombinant RET2 proteins were assayed for nucleotide addition to a 5'-labeled ssRNA (left panel) or a precleaved editing dsRNA substrate (right panel). Diagrams of the RNA substrates are shown and the asterisk denotes the radiolabel. Each assay contained 25 ng protein and 100 μ M of the indicated nucleotide. Labeled input RNA is indicated by (I) and input RNA with 1, 2, 3 or 5 added Us are indicated by +1U, +2U, +3U and +5U, respectively.

surface and provide all essential catalytic residues. Therefore, we refer to these domains together as the 'catalytic bidomain'. Extensive interactions between the MD and CTD are also observed, burying 2595 $\text{\AA}^2$ solvent-accessible surface (Figure 3A).

UTP- and UMP-binding sites

In addition to the unliganded *TbRET2* structure, we have determined the structures of *TbRET2* in complexes with UTP, UTP/UMP and divalent metal ions Mn^{2+} or Mg^{2+} in both cases. The ligands were soaked into the apo crystals. At lower concentrations of UTP (1 or 10 mM), only one UTP- Mg^{2+} or UTP- Mn^{2+} molecule is observed in the deep binding pocket

(site A) between the NTD and CTD (Figures 3A and 4). At a higher concentration of UTP (100 mM), two additional sites are occupied (sites B and C) (Figures 3B–D and 5), which will be described below. Soaking experiments with ATP, GTP or CTP individually at 10 mM concentrations revealed no densities at sites B and C, and only densities of the triphosphate moieties at site A, without clearly defined ribose or bases (Figure 6). We also probed potential nucleotide-binding sites with UMP in combination with UTP occupying the active site A. In a soaking experiment with 10 mM UTP together with 100 mM UMP, two UMP molecules, occupying sites B and C, were observed in addition to the UTP- Mg^{2+} at site A. In all of the structures, the site A UTP- Mg^{2+} molecules adopt identical conformations. Similarly, the bases, sugar moieties and the phosphates of the UTP or UMP molecules at sites B and C have the same geometries.

UTP-binding site A: the triphosphate and divalent metal ions

The UTP molecule, together with a divalent metal ion at site A, is located in the cleft of the catalytic bidomain, and the ligand has extensive interactions with both the NTD and CTD (Figures 3A and 4). The triphosphate groups of the UTP together with one Mg^{2+} are intimately linked to the conserved aspartates, D97 and D99 from the NTD. In soaking experiments with Mn^{2+} , the Mn^{2+} ion appears to be located at the identical position as the Mg^{2+} . The Mn^{2+} could be identified unambiguously by its distinct 7σ electron density peak in the anomalous Fourier difference map (Figure 4B). The divalent metal ion adopts a near-perfect octahedral geometry involving six direct ligands: (i) three nonbridging phosphate oxygen atoms, one each from the α -, β - and γ -phosphate groups, respectively, (ii) one oxygen atom from each of the carboxylate side chains of the conserved aspartates D97 and D99 and (iii) one water molecule (Figure 4). The $\text{O}^{\gamma 2}$ atom on the γ -phosphate is directly recognized by the side chains of K300 from the CTD and S96 from the NTD, respectively, through bifurcated hydrogen bonding. The third nonbridging oxygen atom on the γ -phosphate group forms direct bifurcated hydrogen bonds with the side chains of K304 and S318 from the CTD. The nonbridging oxygen atom on the β -phosphate group ($\text{O}^{\beta 1}$), which is coordinating the divalent metal ion, forms hydrogen bonds with the backbone nitrogen atom of S85 from the NT family signature motif in the NTD. The $\text{O}^{\alpha 1}$ atom of the α phosphate, which is also a direct ligand of the divalent metal ion, is engaged in water-mediated interactions with the carboxylate side chains of conserved D99 and D267 (Figure 4).

The sugar moiety of the UTP, which displays C3'-endo geometry, is recognized by residues from both the NTD and CTD through extensive interactions (Figures 4 and 5A). The side chain of C83, immediately preceding the NT signature motif, as well as the aromatic ring of Y319 from the CTD, stacks onto the sugar ring from both sides, stabilizing the sugar ring via hydrophobic interactions. The 2'-hydroxyl is involved in hydrogen bonding with the side-chain atoms, $\text{O}^{\delta 1}$ of N277 and O^{γ} of S278 (Figure 5A). This might suggest that N277 and S278 are the key residues that discriminate a ribo-UTP from a dUTP incoming nucleotide. The 3'-hydroxyl is hydrogen bonded to the O^{γ} of S278 as well as a water molecule, which in turn is hydrogen bonded to the $\text{O}^{\beta 2}$ nonbridging oxygen of the UTP β -phosphate group.

Figure 3 Overall structure of *TbRET2*. (A) Ribbon presentation of the molecule. The three domains are shown in green (NTD), yellow (MD) and blue (CTD) (as in Figure 1). The UTP- Mg^{2+} at site A is shown as ball-and-stick in a cleft between the NTD and the CTD. (B–D) Electrostatic potential surface of *TbRET2*. The positive and negative regions are colored in blue and red, respectively. Nucleotide-binding sites are labeled in yellow. UTP/UMP and Mg^{2+} are shown as ball-and-stick. Notice the extended blue patch across three domains that suggests an RNA binding region. (B) Front view, UTP-binding site A with a UTP and nucleotide-binding site B with a UMP molecule; (C) side view (90° upward rotation), nucleotide-binding site B with a UMP molecule; (D) back view (180° upward rotation), nucleotide-binding site C with a UMP molecule. Figure generated by PYMOL (DeLano, 2002).

U-specificity

The uridine base adopts an *anti*-conformation and is inserted deeply into the bottom of the cleft between the NTD and CTD domains. It is involved in extensive interactions exclusively with the CTD (Figures 4A and 5A). The uridine base has stacking interactions with the aromatic ring of Y319. Its O2 atom forms a hydrogen bond with the $\text{N}^{\delta 2}$ atom of N277, as well as with a water molecule. Through two additional water molecules, the O4 atom is hydrogen bonded to $\text{O}^{\gamma 1}$ of T317 and $\text{N}^{\eta 1}$ of R435. E424 and R144 form a salt bridge and close one end of the cleft between the NTD and CTD domains. Interestingly, the two carboxylate side chains from D421 and E424 are clustered around the N3 amine group on the U base. A conserved water molecule (Wat1) bridges the N3 amine group with two crucial carboxylate oxygens, $\text{O}^{\delta 1}$ of D421 and $\text{O}^{\epsilon 1}$ of E424 (Figure 5B). This water molecule was present in all electron densities obtained from more than 20 crystals, and had an average B-factor of $\sim 24 \text{ \AA}^2$, while that of the uridine ring was $\sim 20 \text{ \AA}^2$. The angle between the E424 $\text{O}^{\epsilon 1}$ to Wat1 vector and the Wat1 to D421 $\text{O}^{\delta 1}$ vector is 109° on average. Thus, the two hydrogens of the unique water molecule are donated exclusively to the two carboxylate oxygens through hydrogen bonding, rendering the water molecule oxygen only able to accept hydrogen bond donors, in this case, the hydrogen atom on the N3 amine group of the uridine base. This suggests that a uridine base is more favorable than a cytidine base as the latter has a hydrogen acceptor at the N3 position, which would be repelled by the

lone pairs of the precisely coordinated Wat1 molecule. Clearly, a bulky purine base will not fit in the position, simply because of geometric restrictions. In fact, soaking experiments using the same conditions with 10 mM CTP, ATP or GTP clearly displayed no trace of base densities at the active site of *TbRET2*. In contrast, at 1 mM UTP, the fully occupied UTP density was distinctly observed (Figure 6).

Additional nucleotide-binding sites B and C

The UMP/UTP-binding site B is located in the vicinity of site A, with a $4.5 \text{ \AA}$ closest distance between the two O4 atoms on the uridine bases (Figure 3B and C). The uridine base is hydrogen bonded mainly with the backbone atoms of A314, M315 and T317 (Figure 5C). The C3'-*endo* conformational sugar moiety and the phosphate group are located on the surface of the protein that might be a potential RNA-binding site, as it displays a uniformly positively charged surface potential across the catalytic bidomain and part of the MD (see below). The loop that binds the U base at site B (V308–S318) is poorly defined in the apo *TbRET2* structure.

Site C is located on the 'back side' of the protein (Figure 3D). The sugar displays C2'-*endo* geometry and the base atoms are mainly involved in hydrogen-bonding interactions with the backbone atoms from Y352 and D354 (Figure 5D). The hydroxyl groups on the sugar moiety are recognized by the side chain of R282. In addition, stacking interactions occur between the uridine base and P350, as well as the side chain of D354. As site C is far away from the active

Figure 4 Stereoview of the UTP-binding site. **(A)** The UTP-Mg²⁺ as well as the essential interacting residues are shown as ball-and-stick. The U-specifying Wat1 molecule is shown as a white sphere. The hydrogen bonds are shown as red dashed lines. The secondary structure elements are colored with the same scheme as in Figure 1. **(B)** 2mF_o-DF_c electron density map covering the UTP molecule and an Mn²⁺ ion. The anomalous Fourier map of Mn²⁺ is shown in red at a 7σ contour level.

site as well as from the putative RNA binding region, the binding of the nucleotide at this site may be nonspecific.

Comparisons with the NT family: unique structural motif shuffling

A Dali search (Holm and Sander, 1993) revealed that the fold of the NTD-CTD catalytic bidomain of *TbRET2* most closely resembles the catalytic domains from the yeast poly(A) polymerase (Bard *et al*, 2000; PDB code: 1fa0), the archaeal tRNA CCA-adding enzyme (Okabe *et al*, 2003; PDB code: 1uet) and the 2'-5'-oligoadenylate synthetase (Hartmann *et al*, 2003; PDB code: 1px5), even though the shared amino-acid sequence identity is less than 14% in each pairwise comparison. The topology of *TbRET2* is unique in that there is a large insertion within the NTD, which forms the MD. This insertion is conserved among all trypanosomatid TUTases (Figures 1 and 7). In contrast to all other structures of the NT superfamily known to date, whose topologies are always linearly sequential with a contiguous two-domain catalytic core, the catalytic core of *TbRET2* is composed of both the NTD and CTD. The r.m.s. deviation of 269 equivalent C^α atoms between the *TbRET2* noncontiguous catalytic bidomain and the top Z-scoring yeast poly(A) polymerase

(PDB code: 1fa0, Z-score = 20.6) is about 3.5 Å. The NTD and CTD domains of the current structure most likely serve as one functional unit, which is reminiscent of the two contiguous bidomains in poly(A) polymerase structures (Figure 7).

Within the known NT family structures, an insertion of an extra domain near the end of the NTD has not been seen so far. A Dali search reveals that the fold of this MD is topologically similar to the binding domains of several RNA-binding proteins such as the RNA-binding domain of the U1A spliceosomal protein (Oubridge *et al*, 1994; PDB code: 1urn, Z-score = 5.6, r.m.s. deviation = 2.7 Å over 74 equivalent C^α atoms with 8% sequence identity), the RRM domain of the human La protein (Alfano *et al*, 2004; PDB code: 1s79, Z-score = 5.6, r.m.s. deviation = 3.2 Å over 77 equivalent C^α atoms with 18% sequence identity) and the CTD of an archaeal CCA-adding enzyme (Xiong and Steitz, 2004; PDB code: 1tfy, Z-score = 5.6, r.m.s. deviation = 4.2 Å over 80 equivalent C^α atoms with 16% sequence identity). The CTD of the archaeal CCA-adding enzyme has been shown to bind double-stranded tRNA stem substrate through the α-helices regions (Xiong and Steitz, 2004). The corresponding parts of the *TbRET2* MD display extensive electron-positive surface potential, which constitute part of the potential RNA binding

Figure 5 UTP-specificity and UMP binding. (A) Interactions of the UTP sugar moiety at site A with the 2'-OH and the side chains of N277 and S278. (B) The uridine base is specifically recognized by Wat1 and D421 and E424. The two hydrogen atoms of Wat1 shown as white spheres makes exquisite hydrogen bonds with O^{δ1} of D421 and O^{ε1} of E424. The latter two residues are conserved among all trypanosomatid TUTases (Figure 1). (C) UMP binding at site B; (D) UMP binding at site C, which is located far from the active site.

path (Figure 3, see also below). All this evidence implicates a potential RNA binding activity of the MD of *TbRET2*. In fact, deletion of this highly conserved insertion region from *L. tarentolae* RET1 (*LtRET1*) abolished enzyme activity, highlighting the importance of this domain (Aphasizheva *et al*, 2004). However, the major RNA binding region of *LtRET1* was mapped by these authors to the C-terminal portion of the protein that is not found in RET2. Therefore, the MD domain might have a function in protein–protein interactions, yet it seems unlikely that the MD would be a heterologous protein–protein interaction domain as this region is conserved in both RET1s and RET2s, and these two TUTases interact with different proteins and associate with different complexes. In addition, a size-exclusion chromatography study on *TbRET2* protein in solution showed a monomeric state of the molecule (data not shown). Obviously, additional experiments will have to be carried out to reveal the RNA and/or protein binding functions of the MDs in RET2 and RET1.

A wealth of information has been obtained during the past several decades illuminating the enzymatic mechanism of the nucleotidyl transferases. It is believed that nucleotide addition is catalyzed by a two-metal ion mechanism (Steitz, 1998) that is the same for all polymerases. Multiple divalent metal ions have been located in several crystal structures such as those of PAP, CCA-adding enzymes (Bard *et al*, 2000; Martin *et al*, 2000; Okabe *et al*, 2003; Xiong *et al*, 2003; Tomita *et al*, 2004; Xiong and Steitz, 2004) and DNA polymerase β (Sawaya *et al*, 1997; Steitz, 1998). In these structures, two or three divalent metal ions are coordinated by three structurally invariant aspartates in the catalytic domains. When the catalytic bidomain of *TbRET2* is superimposed onto that of

CCA-adding enzyme complexed with dsRNA (Xiong and Steitz, 2004), the three aspartates overlay well with an r.m.s. deviation within 1 Å. However, in the current *TbRET2* structure, only one single divalent metal ion is unambiguously located and directly coordinates with the UTP triphosphates at site A and two of the invariant aspartates, D97 and D99. In *TbRET2*, the third essential aspartate, D267, is not directly involved in interactions with divalent metal ions or phosphate groups of UTP. A single metal ion has also been observed binding in the active site of DNA polymerase β and a transposase in the absence of a nucleotide substrate, displaying an asymmetric nature of the active site with respect to the divalent metal ions (Pelletier, 1994; Davies *et al*, 2000). Nevertheless, the third conserved aspartate D267 is located in the vicinity of the α -phosphate group of the UTP and its O^{δ1} is hydrogen bonded to the UTP O^{ε1} via a bridging water molecule (Figure 4). This water position might be occupied by a second metal ion in the presence of an RNA primer and D267 might be essential during the uridine monophosphate moiety transfer step. Indeed, mutation of the equivalent aspartate in *LtRET1* completely inactivated TUTase activity (Aphasizheva *et al*, 2004).

Implications for RNA binding and uridylyl transferring process

The *TbRET2* molecule displays a large T-shaped concave basic electrostatic surface, suggesting the potential for interaction with RNA substrates (Figure 3B and C). The short arm of the 'inverted' T-shaped potential RNA path runs 'vertically' between the NTD and CTD, directly above the primary UTP binding pocket (site A). The pronounced long 'horizontal'

Figure 6 *TbRET2* in complex with NTPs. (A) 1 mM UTP, 10 mM Mg^{2+} , 5 min soaking time, 2.2 Å resolution. (B) 10 mM CTP, 10 mM Mg^{2+} , 30 min soaking time, 3.2 Å resolution. (C) 10 mM ATP, 10 mM Mg^{2+} , 30 min soaking time, 3.2 Å resolution. (D) 10 mM GTP, 10 mM Mg^{2+} , 30 min soaking time, 3.2 Å resolution. Nonmodel-biased Fourier difference electron density maps are shown in red at a 3.0σ contour level. The positioned NTP models shown, except for UTP, are not refined.

arm of this putative RNA binding region is composed of the extremely basic front surfaces of all three domains. In the crystal structures of *TbRET2* soaked in 100 mM UTP or UMP, the secondary nucleotide-binding site B is located in this hypothetical RNA path, making contacts with R435 and R144 (Figures 3C and 5C).

TUTases may adopt the same mechanism that is used by the NT superfamily. The two divalent metal ions neutralize the negative charges of the triphosphate groups of the incoming UTP, while being coordinated by the three invariant aspartates. The 3'-hydroxyl group of the RNA primer then attacks the α -phosphorus of the UTP in an in-line manner, releasing pyrophosphate without a covalent intermediate state. Similar to DNA polymerase β , *TbRET2* acts in a step-wise manner rather than processively (Ernst *et al*, 2003). In contrast, RET2 is not strictly template-dependent since Us can be added opposite guiding Cs instead of As or Gs (Igo *et al*, 2000). The guiding As or Gs appear to be important for stabilizing the added Us through base-pairing, which subsequently enhances the efficiency of the ligation step. DNA polymerase β uses an 'induced-fit' mechanism to select specifically the correct incoming nucleotide that forms the correct Watson-Crick base-pairing with the template base at the active site dynamically (Sawaya *et al*, 1997). In contrast, in the present *TbRET2* structure, the U-specificity comes from

the U-signature residues (D421 and E424) that position an invariant water molecule with exquisite precision (Figures 4A and 5B). In addition, N277 and S278 discriminate an incoming RNA nucleotide from a DNA nucleotide by recognizing the 2'-hydroxyl group on the sugar moiety (Figure 5A). This U-signature is present in all RET1s and RET2s in all of the kinetoplastid orthologs (Figure 1).

In the structures of human DNA polymerase β complexed with gapped or nicked DNA, the DNA substrates bind with a 90° kink occurring precisely at the 5'-phosphodiester linkage of the templating residue (Sawaya *et al*, 1997). As *TbRET2* inserts Us into an RNA editing site with gaps, the substrate RNA might be kinked, and it is possible that the RNA binding mode of *TbRET2* may adopt a similar pattern as that observed with DNA polymerase β complexed with DNA, given the potential T-shaped RNA binding surface on *TbRET2* (Figure 3). However, the directions and precise binding mode of the RNA substrates to *TbRET2* remains to be determined.

Key features of *T. brucei* RET2

Numerous key residues of *TbRET2* were identified in the current structure. These include three invariant catalytic aspartates, D97, D99 and D267, which coordinate a divalent metal ion and interact indirectly with the triphosphate of the

Figure 7 Superimposition of the structures of *TbRET2* and poly(A) polymerase (PAP) from *Saccharomyces cerevisiae* (Bard *et al*, 2000). (A) Side view. The PAP structure is shown in magenta, the *TbRET2* structure is colored according to the same scheme as in Figure 1. Note that, despite a similar topology (see text), the MD of *TbRET2* (in orange dashed circle) displays an entirely different position compared to the C-terminal domain of PAP (in magenta dashed circle), with respect to the equivalent catalytic bidomains that are common to both enzymes. (B) Top view of the catalytic bidomain. The UTP at site A in the *TbRET2* structure is shown in cyan ball-and-stick, while the ATP in the PAP structure is shown in red. The UTP and ATP occupy similar positions, particularly for the triphosphate moieties.

incoming UTP. Four residues from the catalytic bidomain, S96, K300, K304 and S318, recognize the γ -phosphate, while S85 contacts the β -phosphate. C83 and Y319 stabilize the sugar ring of the UTP via stacking interactions. S278 and N277 might discriminate an incoming ribo-U from a 2'-deoxy nucleotide by forming hydrogen bonds with the O2' hydroxyl of the sugar moiety. A major conclusion from our structure determination is that the lone pair of a crucial water molecule, exquisitely positioned due to the interactions of the carboxylates of D421 and E424 with the hydrogens of this water molecule, is key to the U-specificity of the N3 on the base. R435 and T317 may contribute additional specificity at the O4 site through water-mediated hydrogen bonding. A salt bridge formed by R144 and E424 at the bottom site of the UTP pocket may provide essential structural support. An extensive basic electrostatic potential surface implies an RNA-binding site centered around the active site of *TbRET2*.

Comparisons of the crystal structure of *TbRET2* with that of PAP, the CCA-adding enzyme, as well as other NT superfamily members, show that *TbRET2* adopts a unique domain architecture. The *TbRET2* MD is formed by a large insertion between the second and third catalytic aspartates, a feature not seen before in any member of the superfamily. Finally, *TbRET2* is essential for survival of the bloodstream form of *T. brucei* and is likely to be essential for other trypanosomatid pathogens as well and, hence, is a potential target for drug therapy.

Materials and methods

Site-directed protein surface engineering

Wild-type *TbRET2* protein appeared to be not stable and failed to yield crystals. We reasoned that this might be due to the hydrophobic surface residues involved in protein-protein interactions in the editosome, and that without partners, such hydrophobic patches result in poor behavior of the wild-type *TbRET2* on its own. Therefore, a systematic surface mutagenesis from Trp or Phe to Arg was carried out based on the available structure of a distantly related homologous protein, the yeast poly(A) polymerase (Bard *et al*, 2000; PDB code: 1fa0) aiming for *TbRET2* variants with better solubility and stability in solution. About seven single mutant variants were constructed, followed by combining those single mutations that showed better solubility compared to the wild-type *TbRET2* into double and then triple mutants. Eventually, a specific triple mutant of *TbRET2*, W94R/W122R/W207R, readily produced crystals with good qualities that ultimately led to the solution of the structure. Enzymatic activities of double and triple mutants were analyzed to check the nonaltered functionalities of the protein.

Protein purification

The *TbRET2* triple mutant, 94R/122R/207R, was expressed in *Escherichia coli* as N-terminal his-tagged recombinant protein. The protein was purified by Ni-NTA column and TEV cleavage of the his-tag, followed by size-exclusion chromatography. SeMet-substituted protein was expressed as described previously (Duyne *et al*, 1993). The purified *TbRET2* protein was immediately flash-frozen and stored at -80°C (Deng *et al*, 2004a). More details have been provided in the Supplementary data.

Trypanosome cell lines and culture

The conditional *TbRET2* knockout experiments were performed as described for other proteins (Wirtz *et al*, 1999; Schnauffer *et al*, 2001). More details can be found in the Supplementary data.

TUTase activity assays

Purified recombinant *TbRET2* proteins were assayed as described previously (Ernst *et al*, 2003). More details can be found in the Supplementary data.

Crystallization and data collection

Flash-frozen protein aliquots were thawed in hand immediately before crystallization experiments. Sitting-drop, vapor-diffusion experiments at 4°C with a 3 M KCl reservoir solution yielded crystals with excellent diffraction quality. A solution containing 1.8 M KCl, 2 mM Na-UTP, 10 mM Mg^{2+} plus 25% glycerol was used as the cryoprotectant. The data collection was carried out at beamline 19-ID at APS. Single wavelength SAD data at the peak position of the Se absorption K-edge were collected to 2.0 Å resolution at 100 K. Data (1.8 Å) were collected at the Se remote position using another larger crystal. The crystal belongs to space group $P4_12_12$, with cell dimensions of $a = b = 91.9 \text{ Å}$ and $c = 163.4 \text{ Å}$ with one molecule in the asymmetric unit. After crosslinking by 0.1% glutaraldehyde at 4°C for 30 min, crystals complexed with UTP were obtained by soaking the apo crystals in a soaking solution containing 20 mM Tris, 20% PEG 8K, 25% glycerol, pH 8, 1–100 mM UTP, 10 mM Mg^{2+} or Mn^{2+} . Crystals complexed with UTP/UMP were obtained in the same way as in the above buffer containing 10 mM UTP and 100 mM UMP. ATP, GTP and CTP soaking experiments were carried out in the same manner as above with all nucleotides at 10 mM concentration. Data from the

TbRET2–nucleotide complex crystals were collected at beam-line 8.2.2 at ALS. All the data were processed and reduced by HKL2000 (Otwinowski and Minor, 1997).

Structure determination and refinement

The apo TbRET2 structure was solved by the single wavelength anomalous dispersion method using data collected from a crystal from SeMet-substituted TbRET2 protein. The program SOLVE (Terwilliger and Berendzen, 1999) was used to locate the positions of the Se atom sites. After density modification by RESOLVE (Terwilliger and Berendzen, 1999), most of the residues were automatically built by Arp/Warp (Perrakis *et al*, 1999). The final model was built with the program COOT (Emsley and Cowtan, 2004) and refined with the program REFMAC5 (Murshudov *et al*, 1999). The uridine nucleotide complexed structures were determined using unbiased $F_o - F_c$ difference Fourier methods before the ligands were included in the model. The program PROCHECK

(Laskowski *et al*, 1993) was used for the final model analysis. The current models are of excellent geometry and refinement statistics (Supplementary Table 1).

Supplementary data

Supplementary data are available at *The EMBO Journal* Online.

Acknowledgements

We gratefully acknowledge the staff of beam-line 19ID at the Advanced Photon Source, and the staff of beam-line 8.2.2 at the Advanced Light Source for their support. This work received support from NIH Grant GM42188 and built on work supported by NIH Grant AI14102 to KDS. Coordinates and structure factors have been deposited with the Protein Data Bank with Accession codes 2B4V, 2B51 and 2B56.

References

- Alfano C, Sanfelice D, Babon J, Kelly G, Jacks A, Curry S, Conte M (2004) Structural analysis of cooperative RNA binding by the La motif and central RRM domain of human La protein. *Nat Struct Mol Biol* **11**: 323–329
- Aphasizhev R, Aphasizheva I, Nelson RE, Gao G, Simpson AM, Kang X, Falick AM, Sbicego S, Simpson L (2003a) Isolation of a U-insertion/deletion editing complex from *Leishmania tarentolae* mitochondria. *EMBO J* **22**: 913–924
- Aphasizhev R, Aphasizheva I, Simpson L (2003b) A tale of two TUTases. *Proc Natl Acad Sci USA* **100**: 10617–10622
- Aphasizhev R, Aphasizheva I, Simpson L (2004) Multiple terminal uridylyltransferases of trypanosomes. *FEBS Lett* **572**: 15–18
- Aphasizhev R, Sbicego S, Peris M, Jang S, Aphasizheva I, Simpson A, Rivlin A, Simpson L (2002) Trypanosome mitochondrial 3' terminal uridylyl transferase (TUTase): the key enzyme in U-insertion/deletion RNA editing. *Cell* **108**: 637–648
- Aphasizheva I, Aphasizhev R, Simpson L (2004) RNA-editing terminal uridylyl transferase 1: identification of functional domains by mutational analysis. *J Biol Chem* **279**: 24123–24130
- Aravind L, Koonin EV (1999) DNA-binding proteins and evolution of transcription regulation in the archaea. *Nucleic Acids Res* **27**: 4658–4670
- Augustin M, Reichert A, Betat H, Huber R, Morl M, Steegborn C (2003) Crystal structure of the human CCA-adding enzyme: insights into template-independent polymerization. *J Mol Biol* **328**: 985–994
- Bard J, Zhelkovsky A, Helmling S, Earnest T, Moore C, Bohm A (2000) Structure of yeast poly(A) polymerase alone and in complex with 3'-dATP. *Science* **289**: 1346–1349
- Chen Y, Hung C-H, Burdener T, Gwo-Shu ML (2003) Development of RNA interference revertants in *Trypanosoma brucei* cell lines generated with a double stranded RNA expression construct driven by two opposing promoters. *Mol Biochem Parasitol* **126**: 275–279
- Carnes J, Trotter JR, Ernst NL, Steinberg A, Stuart K (2005) An essential RNase III insertion editing endonuclease in *Trypanosoma brucei*. *PNAS Early Edition* pnas 0506133102
- Davies D, Goryshin I, Reznikoff W, Rayment I (2000) Three-dimensional structure of the Tn5 synaptic complex transposition intermediate. *Science* **289**: 77–85
- DeLano WL (2002) The PyMOL molecular graphics system. <http://www.pymol.org>
- Delarue M, Boule J, Lescar J, Expert-Bezancon N, Jourdan N, Sukumar N, Rougeon F, Papanicolaou C (2002) Crystal structures of a template-independent DNA polymerase: murine terminal deoxynucleotidyltransferase. *EMBO J* **21**: 427–439
- Deng J, Davies DR, Wisedchaisri G, Wu M, Hol WG, Mehlin C (2004a) An improved protocol for rapid freezing of protein samples for long-term storage. *Acta Crystallogr Sect D* **60**: 203–204
- Deng J, Schnauffer A, Salavati R, Stuart KD, Hol WGJ (2004b) High resolution crystal structure of a key editosome enzyme from *Trypanosoma brucei*: RNA editing ligase 1. *J Mol Biol* **343**: 601–613
- Drozdz M, Palazzo SS, Salavati R, O'Rear J, Clayton C, Stuart K (2002) TbMP81 is required for RNA editing in *Trypanosoma brucei*. *EMBO J* **21**: 1791–1799
- Duyn GDV, Standaert RF, Karplus PA, Schreiber SL, Clardy J (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J Mol Biol* **229**: 105–124
- Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr Sect D* **60**: 2126–2132
- Ernst NL, Panicucci B, Igo RP, Panigrahi AK, Salavati R, Stuart K (2003) TbMP57 is a 3' terminal uridylyl transferase (TUTase) of the *Trypanosoma brucei* editosome. *Mol Cell* **11**: 1525–1536
- Gao G, Simpson AM (2003) Is the *Trypanosoma brucei* REL1 RNA ligase specific for U-deletion RNA editing, and is the REL2 RNA ligase specific for U-insertion editing? *J Biol Chem* **278**: 27570–27574
- Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**: 305–308
- Hartmann R, Justesen J, Sarkar S, Sen G, Yee V (2003) Crystal structure of the 2'-specific and double-stranded RNA-activated interferon-induced antiviral protein 2'-5'-oligoadenylate synthetase. *Mol Cell* **12**: 1173–1185
- Holm L, Sander C (1993) Protein structure comparison by alignment of distance matrices. *J Mol Biol* **233**: 123–138
- Holm L, Sander C (1995) DNA polymerase beta belongs to an ancient nucleotidyltransferase superfamily. *Trends Biochem Sci* **20**: 345–347
- Huang CE, Cruz-Reyes J, Zhelonkina AG, O'Hearn S, Wirtz E, Sollner-Webb B (2001) Roles for ligases in the RNA editing complex of *Trypanosoma brucei*: band IV is needed for U-deletion and RNA repair. *EMBO J* **20**: 4694–4703
- Igo RJ, Palazzo S, Burgess M, Panigrahi A, Stuart K (2000) Uridylate addition and RNA ligation contribute to the specificity of kinetoplast insertion RNA editing. *Mol Cell Biol* **20**: 8447–8457
- Kang X, Rogers K, Gao G, Falick AM, Zhou S, Simpson L (2005) Reconstitution of uridine-deletion precleaved RNA editing with two recombinant enzymes. *Proc Natl Acad Sci USA* **102**: 1017–1022
- Li F, Xiong Y, Wang J, Cho H, Tomita K, Weiner A, Steitz T (2002) Crystal structures of the *Bacillus stearothermophilus* CCA-adding enzyme and its complexes with ATP or CTP. *Cell* **111**: 815–824
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* **26**: 283–291
- Madison-Antenucci S, Grams J, Hajduk SL (2002) Editing machines: the complexities of trypanosome RNA editing. *Cell* **108**: 435–438
- Madison-Antenucci S, Hajduk SL (2001) RNA editing-associated protein 1 is an RNA binding protein with specificity for preedited mRNA. *Mol Cell* **7**: 879–886
- Martin G, Keller W, Doublet S (2000) Crystal structure of mammalian poly(A) polymerase in complex with an analog of ATP. *EMBO J* **19**: 4193–4203
- McManus MT, Shimamura M, Grams J, Hajduk SL (2001) Identification of candidate mitochondrial RNA editing ligases from *Trypanosoma brucei*. *RNA* **7**: 167–175

- Murshudov GN, Vagin AA, Lebedev A, Wilson KS, Dodson EJ (1999) Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr Sect D* **55**: 247–255
- Okabe M, Tomita K, Ishitani R, Ishii R, Takeuchi N, Arisaka F, Nureki O, Yokoyama S (2003) Divergent evolutions of trinucleotide polymerization revealed by an archaeal CCA-adding enzyme structure. *EMBO J* **22**: 5918–5927
- Otwinowski Z, Minor W (1997) DENZO. Processing of X-ray diffraction data collected in oscillation mode, In *Methods in Enzymology*, Sweet RM, Carter, CWJ (eds) Vol. 276, pp 307–326. New York: Academic Press
- Oubridge C, Ito N, Evans P, Teo C, Nagai K (1994) Crystal structure at 1.92 Å resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **372**: 432–438
- Panigrahi AK, Gygi SP, Ernst NL, Igo Jr RP, Palazzo SS, Schnauffer A, Weston DS, Carmean N, Salavati R, Aebersold R, Stuart KD (2001a) Association of two novel proteins, TbMP52 and TbMP48, with the *Trypanosoma brucei* RNA editing complex. *Mol Cell Biol* **21**: 380–389
- Panigrahi AK, Schnauffer A, Carmean Jr N, Igo Jr RP, Gygi SP, Ernst NL (2001b) Four related proteins of the *Trypanosoma brucei* RNA editing complex. *Mol Cell Biol* **21**: 6833–6840
- Panigrahi AK, Schnauffer A, Ernst NL, Wang B, Carmean N, Salavati R, Stuart K (2003) Identification of novel components of *Trypanosoma brucei* editosomes. *RNA* **9**: 484–492
- Pelletier H (1994) Polymerase structures and mechanism. *Science* **266**: 2025–2026
- Pelletier H, Sawaya M, Kumar A, Wilson S, Kraut J (1994) Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer, and ddCTP. *Science* **264**: 1891–1903
- Perrakis A, Morris R, Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* **6**: 458–463
- Pollard VW, Harris ME, Hajduk SL (1992) Native mRNA editing complexes from *Trypanosoma brucei* mitochondria. *EMBO J* **11**: 4429–4438
- Rusche LN, Cruz-Reyes J, Piller KJ, Sollner-Webb B (1997) Purification of a functional enzymatic editing complex from *Trypanosoma brucei* mitochondria. *EMBO J* **16**: 4069–4081
- Rusche LN, Huang CE, Piller KJ, Hemann M, Wirtz E, Sollner-Webb B (2001) The two RNA ligases of the *Trypanosoma brucei* RNA editing complex: cloning the essential band IV gene and identifying the band V gene. *Mol Cell Biol* **21**: 979–989
- Ryan CM, Read LK (2005) UTP-dependent turnover of *Trypanosoma brucei* mitochondrial mRNA requires UTP polymerization and involves the RET1 TUTase. *RNA* **11**: 763–773
- Sakon J, Liao HH, Kanikula A, Benning M, Rayment I, Holden H (1993) Molecular structure of kanamycin nucleotidyltransferase determined to 3.0-Å resolution. *Biochemistry* **32**: 11977–11984
- Sawaya M, Prasad R, Wilson S, Kraut J, Pelletier H (1997) Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism. *Biochemistry* **36**: 11205–11215
- Schnauffer A, Ernst NL, Palazzo SS, O'Rear J, Salavati R, Stuart K (2003) Separate insertion and deletion subcomplexes of the *Trypanosoma brucei* RNA editing complex. *Mol Cell* **12**: 307–319
- Schnauffer A, Panigrahi AK, Panicucci Jr B, Igo Jr RP, Wirtz E, Salavati R, Stuart K (2001) An RNA ligase essential for RNA editing and survival of the bloodstream form of *Trypanosoma brucei*. *Science* **291**: 2159–2162
- Simpson L, Sbicego S, Aphasizhev R (2003) Uridine insertion/deletion RNA editing in trypanosome mitochondria: a complex business. *RNA* **9**: 265–276
- Steitz T (1998) A mechanism for all polymerases. *Nature* **391**: 231–232
- Stuart K, Allen TE, Heidmann S, Seiwert SD (1997) RNA editing in kinetoplastid protozoa. *Microbiol Mol Biol Rev* **61**: 105–120
- Stuart K, Panigrahi AK, Schnauffer A, Drozd M, Clayton C, Salavati R (2002) Composition of the editing complex of *Trypanosoma brucei*. *Philos Trans R Soc Ser B* **357**: 71–79
- Terwilliger TC, Berendzen J (1999) Automated MAD and MIR structure solution. *Acta Crystallogr Sect D* **4**: 849–861
- Tomita K, Fukai S, Ishitani R, Ueda T, Takeuchi N, Vassilyev D, Nureki O (2004) Structural basis for template-independent RNA polymerization. *Nature* **430**: 700–704
- Trotter JR, Ernst NL, Carnes J, Panicucci B, Stuart K (2005) A deletion site editing endonuclease in *Trypanosoma brucei*. *Mol Cell* (in press)
- Wirtz E, Leal S, Ochatt C, Cross GA (1999) A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol Biochem Parasitol* **99**: 89–101
- World Health Organization (2003) *WHO Report in Global Surveillance of Epidemic-Prone Infectious Diseases-African Trypanosomiasis*. Rome: WHO
- Worthey EA, Schnauffer A, Mian IS, Stuart K, Salavati R (2003) Comparative analysis of editosome proteins in trypanosomatids. *Nucleic Acids Res* **31**: 6392–6408
- Xiong Y, Li F, Wang J, Weiner A, Steitz T (2003) Crystal structures of an archaeal class I CCA-adding enzyme and its nucleotide complexes. *Mol Cell* **12**: 1165–1172
- Xiong Y, Steitz T (2004) Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature* **430**: 640–645
- Yue D, Maizels N, Weiner A (1996) CCA-adding enzymes and poly(A) polymerases are all members of the same nucleotidyltransferase superfamily: characterization of the CCA-adding enzyme from the archaeal hyperthermophile *Sulfolobus shibatae*. *RNA* **2**: 895–908
- Yue D, Weiner A, Maizels N (1998) The CCA-adding enzyme has a single active site. *J Biol Chem* **273**: 29693–29700