170118

ATP synthase from *Trypanosoma brucei* has an elaborated canonical F_1-domain and conventional catalytic sites

Martin G. Montgomery[a, 1], Ondřej Gahura[a, b, 1], Andrew G. W. Leslie[c], Alena Zíková[b], and John E. Walker[a, 2]

[a]*The Medical Research Council Mitochondrial Biology Unit, University of Cambridge, Cambridge Biomedical Campus, Hills Road, Cambridge CB2 0XY, United Kingdom;* [b]*Biology Centre, Czech Academy of Sciences, Institute of Parasitology, České Budějovice, 37005, Czech Republic;* [c]*The Medical Research Council Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Ave, Cambridge, CB2 0QH, United Kingdom*

[1]M. G. M. and O. G. contributed equally

[2]To whom correspondence should be addressed. e-mail: walker@mrc-mbu.cam.ac.uk

Running title: ATP synthase from *Trypanosoma brucei*

The authors declare no conflict of interest

PDB code: 6F5D

Author contributions: J. E. W. and A. Z. designed research. J. E. W. supervised project; and M. G. M. and O. G. performed research; M. G. M, A. G. W. L. and J. E. W. analyzed data, and J. E. W. prepared the manuscript.

Classification: BIOLOGICAL SCIENCES, Biochemistry

Key words: ATP synthase; *Trypanosoma brucei*; p18 subunit; catalytic domain; structure; mechanism; conservation.

Significance

Mitochondria generate the cellular fuel, adenosine triphosphate, or ATP, to sustain complex life. Production of ATP depends on the oxidation of energy rich compounds to produce the proton motive force (pmf), a chemical potential difference for protons, across the inner membrane. The pmf drives the ATP synthase to synthesize ATP by a mechanical rotary mechanism. The structures and functions of the protein components of this molecular machine, especially those involved directly in the catalytic formation of ATP, are widely conserved in metazoans, fungi and eubacteria. Here, we show that a proposal that this conservation does not extend to the ATP synthase from *Trypanosoma brucei*, a member of the euglenozoa, and the causative agent of sleeping sickness in humans, is incorrect. (120 words)

Abstract

The structures and functions of the components of ATP synthases, especially those subunits involved directly in the catalytic formation of ATP, are widely conserved in metazoans, fungi, eubacteria and plant chloroplasts. On the basis of a map at 32.5 Å resolution determined *in situ* in the mitochondria of *Trypanosoma brucei* by electron cryo-tomography, it has been proposed that the ATP synthase in this species has a non-canonical structure and different catalytic sites, where the catalytically essential arginine-finger is provided, not by the α-subunit adjacent to the catalytic nucleotide binding site as in all species investigated to date, but by a protein called p18 found only in the euglenozoa. A crystal structure at 3.2 Å resolution described here of the catalytic domain of the same enzyme shows that this proposal is incorrect. In many respects, the structure is closely similar to those of F_1-ATPases determined previously. The $\alpha_3\beta_3$-spherical portion of the catalytic domain where the three catalytic sites are found, plus the

central stalk, are highly conserved, and the arginine finger is provided conventionally by the α-subunits adjacent to each of the three catalytic sites found in the β-subunits. Thus, the enzyme has a conventional catalytic mechanism. The structure differs from earlier ones by having a p18-subunit, identified only in the euglenozoa, associated with the external surface of each of the three α-subunits, thereby elaborating the F_1-domain. Subunit p18 is a pentatricopeptide repeat (PPR) protein with three PPRs and appears to have no function in the catalytic mechanism of the enzyme. [250 words]

\body

Introduction

The ATP synthases, also known as F-ATPases or F_1F_o-ATPases, are multisubunit enzyme complexes found in energy transducing membranes in eubacteria, chloroplasts and mitochondria (1, 2). They make ATP from ADP and phosphate under aerobic conditions using a proton-motive force, pmf, generated by respiration or photosynthesis, as a source of energy. Hitherto, studies of the subunit compositions, structures and mechanism of the ATP synthases have been confined mainly to the vertebrates, especially to man and cows, to various fungi, eubacteria and chloroplasts of green plants. These studies have established the conservation of the central features of these rotary machines. They are all membrane bound assemblies of multiple subunits organised into membrane intrinsic and membrane extrinsic sectors. The membrane extrinsic sector, known as F_1-ATPase, is the catalytic part where ATP is formed from ADP and inorganic phosphate. It can be detached experimentally from the membrane domain in an intact state, and retains the ability to hydrolyze, but not to synthesize ATP. The membrane intrinsic sector, sometimes called F_o, contains a rotary motor driven by pmf, and is connected to the extrinsic domain by a central stalk and a peripheral stalk. The enzyme's rotor is constituted from the central stalk, and an associated ring of c-subunits in the membrane

domain. The central stalk lies along an axis of six-fold pseudo-symmetry and penetrates into the $\alpha_3\beta_3$-domain, where the catalytic sites of the enzyme are found at three of the interfaces between α- and β-subunits. The penetrant region of the central stalk is an asymmetric α-helical coiled-coil, and its rotation inside the $\alpha_3\beta_3$-domain takes each catalytic site through a series of conformational changes that lead to the binding of substrates, and to the formation and release of ATP. During ATP hydrolysis in the experimentally detached F_1-domain, the direction of rotation, now driven by energy released from the hydrolysis of ATP, is opposite to the synthetic sense. Extensive structural analyses, mostly by X-ray crystallography at atomic resolution, have shown that the F_1-domains of the enzymes from bovine (3–23) and yeast (24–30) mitochondria, chloroplasts (31, 32) and eubacteria (33–39) are highly conserved. Not only is there conservation of the subunit compositions of the $\alpha_3\beta_3$-domain and the central stalk ($\gamma_1\varepsilon_1$ in eubacteria and chloroplasts, and in mitochondria orthologues $\gamma_1\delta_1$, plus an additional unique subunit, confusingly called ε, attached to the δ-subunit), but also the sequences of subunits are either highly conserved or absolutely conserved in many key residues. This extensive conservation includes residues in catalytic interfaces and in the catalytic sites themselves. In the β-subunits, they include a hydrophobic pocket, where the adenine ring of ADP (or ATP) is bound, a P-loop sequence that interacts with the α-, β- and γ-phosphates of ATP, and provides residues involved either directly, or indirectly via water molecules, in the binding of a hexa-co-ordinate magnesium ion, and, in the adjacent α-subunit, an "arginine finger" residue, which senses whether ADP or ATP is bound to the catalytic site. Indeed, these catalytic features are common to a wide range of NTPases (40, 41), and together with conserved structural features they are characteristics of the canonical ATP synthase.

Based on a structural model at 32.5 Å resolution derived by electron cryo-tomography (ECT), it has been suggested recently that the structure of the F_1-catalytic domain and its catalytic mechanism in the ATP synthase from *Trypanosoma brucei* have diverged extensively

from the canonical complex in an unprecedented manner (42). It was proposed that the structure of this F_1-domain is much more open than those that have been described in other species, and that the "arginine finger" is provided, not by the α-subunit, but by an additional p18-subunit found only in the euglenozoa (43–49). Here, we examine this proposal in the context of a structure of the F_1-domain of the *T. brucei* ATP synthase determined by X-ray crystallography at 3.2 Å resolution.

Results and Discussion

Structure Determination. The crystals of the *T. brucei* F_1-ATPase have the unit cell parameters a=124.2 Å, b=206.4 Å, c=130.2 Å, with α=γ=90.0°, β=104.9°, and they belong to space group $P2_1$, with one F_1-ATPase in the asymmetric unit. Data processing and refinement statistics are presented in Table S1. The final model of the complex contains the following residues: α_E, 20-125, 137-416 and 423-560; α_{TP}, 22-127, 137-414, and 421-560; α_{DP}, 22-125, 137-416, and 424-560; β_E, 6-492; β_{TP}, 7-494; β_{DP}, 8-488; γ, 2-58 and 66-285; δ, 5-16 and 32-165; ε, 1-66; and three copies of p18, residues, 6-169, 6-167, and 6-170, respectively attached to the α_{TP}-, α_{DP}- and α_E-subunits (see below). An ADP molecule and a magnesium ion are bound to each of the three α-subunits, and to the β_{TP}- and β_{DP}-subunits, whereas the β_E-subunit has a bound ADP molecule without a magnesium ion. A similar nucleotide occupancy of catalytic and non-catalytic sites has been observed previously in the bovine F_1-ATPase crystallised in the presence of phosphonate (20), and in the F_1-ATPase from *Caldalkalibacillus thermarum* (38). These structures are interpreted as representing a post-hydrolysis state where the ADP molecule has not been released from the enzyme. An unusual feature of the *T. brucei* F_1-ATPase is that the DP (diphosphate) catalytic interface is more open than the TP (triphosphate) catalytic interface, similar to the F_1-ATPase from *Saccharomyces cerevisiae* (24), whereas the converse is observed in all other structures (see Table S2). As usual, the E (empty) interface is the most open of the three catalytic interfaces (see Table S2). The rotational

position of the γ-subunit (determined by superposition of crown regions of structures) is +23.1° relative to the bovine phosphate release dwell, which is at or close to the catalytic dwell at +30° in the rotary catalytic cycle (6).

Structure of the F_1-ATPase from *T. brucei*. The structure consists of an $\alpha_3\beta_3$-complex with α- and β-subunits arranged in alternation around an anti-parallel α-helical coiled-coil in the γ-subunit (Fig. 1). The rest of the γ-subunit sits beneath the $\alpha_3\beta_3$-complex, and here it is associated with the δ- and ε-subunits. Together, these three subunits form the central stalk. Thus, the overall structure of this particular catalytic domain of the ATP synthase complex is extremely similar to structures of canonical F_1-ATPases determined in the mitochondria of other species, and in eubacteria and chloroplasts. For example, in a comparison of back-bone atoms with the bovine F_1-ATPase crystallised in the presence of phosphonate (20), the rmsd is 3.24 Å. As in these other canonical structures, each of the α- and β-subunits in the *T. brucei* F_1-ATPase has three domains. The N-terminal domain (residues 1-103 and 1-88 in α- and β-subunits, respectively) consists of a six-stranded β-barrel in both α- and β-subunits, and these six β-domains are associated in a stable annulus known as the "crown". The central domain (residues 104-389 and 89-365 in α- and β-subunits, respectively) provides the nucleotide binding sites (see Fig. S1). The C-terminal domain consists of a bundle of seven and four α-helices, in α- and β-subunits, respectively. The crown stabilizes the entire F_1-domain, and, during rotary catalysis, the rest of the α- and β-subunits swing from this crown in response to the rotation of the asymmetrical α-helical coiled-coil region of the γ-subunit.

The six bound ADP molecules occupy nucleotide binding sites that are very similar in structure to those in other ATP synthases. They retain the conventional features of a hydrophobic pocket to bind the adenine ring, and a characteristic P-loop sequence (GDRQTGKT in the α-subunit, residues 182-189; GGAGVGKT in the β-subunit, residues 162-169) interacting with the α- and β-phosphates of ADP or ATP (Fig. 2). The five

magnesium ions are hexa-coordinated by a threonine residue (residues 189 and 169 in α- and β-subunits, respectively) and four water molecules in each case. In the canonical enzymes, the nucleotides bound to the β-subunits participate in catalysis and exchange during a catalytic cycle, whereas those bound to the α-subunits are permanently bound to the enzyme and do not participate in catalysis. The close similarity of the structures of the *T. brucei* and bovine F_1-ATPases suggests strongly that the α- and β-subunits in the *T. brucei* enzyme have the same, or very similar, roles to those in the bovine enzyme. Thus, the nucleotide binding sites in the β-subunits are part of the catalytic sites of the enzyme, the other important catalytic feature being αArg-386, the arginine finger residue, which is positioned in the catalytic site in the β_{DP}-subunit from *T. brucei*, for example, in exactly the same position occupied by the equivalent residue, αArg-373, in the bovine enzyme (Fig. 2).

Despite the general conservation of the structure and mechanism of the *T. brucei* F_1-ATPase, the euglenozoan enzyme is elaborated relative to the bovine enzyme, for example. First, the α-subunit in *T. brucei* is cleaved *in vivo* by proteolysis at two adjacent sites, removing residues 128-135 (Fig. S2) (50). The cleavage of α-subunits has been noted also in other euglenozoan ATP synthases (48, 51–53), although the sites of cleavage have not been characterized precisely. In the bovine enzyme, the equivalent region (residues 117-123) forms an external loop (Fig. S2). These cleavages have no evident impact on the stability of either the α-subunit or the F_1-ATPase complex itself. Second, the α-, β-, δ- and ε-subunits of the *T. brucei* enzyme have additional surface features that are not found in the known structures of other F_1-ATPases (Fig. 1). The most extensive are residues 483-498 and 536-560 in the C-terminal region of the α-subunit, and their significance is discussed below. The additional surface features in the β-, δ- and ε-subunits are residues 485-499, 1-17 and 39-50, respectively. Those in the β- and ε-subunits have no obvious functions. The resolved residues of the additional sequence in the δ-subunit increases its area of interaction with the γ-subunit from

1000 Å^2 to 1700 Å^2. The C-terminal region of the γ-subunit from residues 286-304, although not resolved in the structure, is 19 residues longer than in the bovine enzyme, for example, and in the intact ATP synthase, it could extend beyond the crown region, possibly making contacts, permanently or transiently, during rotary catalysis with the OSCP (oligomycin sensitivity conferral protein), a component of the peripheral stalk. In other species, the OSCP is bound to the F_1-domain by the N-terminal regions of the three α-subunits (19, 29, 37, 54).

Third, and most significantly from a structural view-point, the *T. brucei* F_1-ATPase has an additional p18-subunit bound to each of its three α-subunits (50). The buried surface areas of interaction of the p18-subunits with their partner α_E-, α_{TP}-, and α_{DP}-subunits are 2500, 2600 and 2500 Å^2, respectively. All three p18-subunits are folded into seven α-helices, H1-H7, with an unstructured C-terminal region from residues 151-170. The subunit is bound via H2 and H4 to the surface of the nucleotide binding domain of an α-subunit, and via H5 and H6 to the surface of its C-terminal domain; H7 is not in contact with the α-subunit (Fig. S2), but it is bound to H6; the unstructured C-terminal tail interacts with the C-terminal domain of the α-subunit, travelling towards, but not entering, the non-catalytic interface with the adjacent β-subunit (Figs. 1, 3 and S3). In this region, the extended C-terminal element of the p18-subunit interacts with the two additional segments of sequence (residues 483-498 and 536-560) found in the *T. brucei* α-subunit (Fig. S3). The first additional segment is largely extended, starting with one α-helical turn (residues 483-485). The second additional segment starts with one α-helical turn (residues 536-539) followed by an extended region (residues 540-544) and terminates with an α-helix (residues 546-558) that doubles back into the non-catalytic interface, and interacts with the extreme C-terminal end of the p18-subunit.

Role of the p18-Subunit. As noted before, the sequence of the p18-subunit is related to the PPR (pentatricopeptide repeat) proteins (55), which are found in association with RNA molecules primarily in mitochondria and chloroplasts, and also in some bacterial species. These

proteins are characterized by a degenerate sequence motif 35 amino acids long, related to, but distinct from the motif in the TPR (tetratricopeptide repeat) proteins (56). The PPR repeat is folded into a helix-turn-helix motif, and PPR proteins usually contain several tandem repeats associated into a super-helix, with a concave groove on one face that serves as a binding surface for RNA ligands. The p18-subunit of the F_1-ATPase from *T. brucei* is predicted to be a PPR protein with three PPRs, whereas previously, it has been thought to have two PPRs (50, 55). Although the probability score (49%) is rather low, as reflected in the weak correspondence of the sequences of the three predicted PPRs to the PPR consensus (Fig. S4), the topography of p18 follows closely those of other well-predicted and well-established PPR proteins, such as the RNA binding PPR protein, PPR10, from *Zea mays* (57) (see Fig. S4). This structural comparison (rmsd 2.3 Å) illustrates that, as predicted, the p18-subunit has three PPRs consisting of H1 plus H2, H3 plus H4, and H5 plus H6. H7 could be the relic of the first element of a fourth PPR, where H8 has evolved into the extended C-terminal tail region of the p18-subunit (see Fig. 3*B*). However, p18 does not have the equivalent of the RNA binding site in PPR10, and other residues required for RNA binding in α-helix-7 of PPR10 have been substituted in H7 of p18. Therefore, there is no evidence to suggest that p18 has any role in binding an RNA molecule, and its role in the *T. brucei* F_1-ATPase remains obscure, although its presence is essential for the assembly of the enzyme (50). The sequences of p18-subunits, including the PPR repeats, are highly conserved across the euglenozoa, suggesting that the structure and the mode of interaction of the various p18-proteins with their cognate F_1-ATPases are conserved also (Fig. S5).

Structure of the *T. brucei* F_1-ATPase and the ECT Map. The structure of the F_1-ATPase from *T. brucei* described above at 3.2 Å resolution was docked into the map of the ATP synthase complex from the same organism at 32.5 Å resolution derived by ECT of mitochondrial membranes (Fig. 4). The structure of the catalytic domain described here fits the

region of the map with the mushroom shape, characteristic of the catalytic F_1-domain of other ATP synthases, very well. Thus, this correspondence is also consistent with the *T. brucei* ATP synthase having a canonical catalytic domain elaborated by the attachment of the three p18-subunits. It does not support the proposal in Fig. 4 (*C*) and (*D*), where the map has been interpreted as having a catalytic domain where the nucleotide binding and C-terminal domains of the α-subunits are displaced outwards away from the central stalk, and where the role of the p18-subunit, bound in an unspecified position, is to provide the catalytically essential arginine finger residue (42).

Two other features in the ECT map can also be interpreted in terms of the characterized structures of canonical ATP synthases. First, the uninterpreted region of density above the F_1-domain in Fig. 4 corresponds to the upper part of the peripheral stalk in other ATP synthases. In the eukarya, this region is occupied by the OSCP (oligomycin sensitivity conferral protein), and, in eubacterial and chloroplast enzymes, by the orthologous δ-subunit. As the ATP synthase in *T. brucei* and other euglenozoa that have been examined contain orthologs of the OSCP (49, 58, 59), it is highly probable that the *T. brucei* OSCP provides this feature in the ECT map, and that, as in the well characterized ATP synthases, it is attached to the F_1-domain via interactions with the N-terminal regions of the three α-subunits, which extend from the "top" of the crown domain. The role of the peripheral stalk in ATP synthases is to provide the stator of the enzyme with integrity by connecting the $\alpha_3\beta_3$-domain to the essential ATP6 (a-subunit in eubacteria; subunit IV in chloroplasts) in the membrane domain. ATP6 and orthologs, together with the c-ring in the rotor, provide the transmembrane pathway for protons (2). In order to maintain the integrity of this pathway, and to keep ATP synthesis coupled to proton motive force, the static ATP6 and the rotating c-ring have to be kept in contact by the action of the peripheral stalk. The peripheral stalk is the most divergent of the essential features of the ATP synthase (2). Apart from the OSCP and orthologous δ-subunits, their subunit

compositions, sequences and structures of the related and structurally simpler eubacterial and chloroplast peripheral stalks on the one hand, differ extensively from the more complex structurally characterized peripheral stalks in mitochondrial enzymes, although they are all dominated by approximately parallel, antiparallel, and apparently rigid, long α-helical structures connecting the OSCP to the ATP6-subunit (and orthologs), running alongside the catalytic domains. The peripheral stalks of ATP synthases in the mitochondria of euglenozoa (59) and also in the green alga, *Polytomella*, (60) appear to be even more diverged than those in characterized mitochondrial enzymes. Their subunit compositions are more complex, and as is evident in the map feature to the left of the F_1-domain in Fig. 4, and in other published images, they are thicker and apparently more robust than structurally characterized peripheral stalks. More details are likely to emerge in the near future, most likely from the application of electron cryomicroscopic imaging of individual particles of these enzymes. These endeavors are driven by the imperative to use knowledge of the structure of the ATP synthase from *T. brucei* (61, 62) to aid in the development of new drugs to treat patients with sleeping sickness by finding selective inhibitors of its activity.

Materials and Methods

Crystallization of F_1-ATPase from *T. brucei.* The F_1-ATPase purified from *T. brucei* (50) was crystallized at 4°C by the microbatch method under paraffin oil. The enzyme was dissolved at a protein concentration of 9.0 mg/ml in buffer consisting of 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 10 mM $MgSO_4$ and 1 mM ADP. This protein solution was mixed in wells in micro-batch plates with an equal volume of 7.7% (w/v) polyethyleneglycol 10000 dissolved in a buffer containing 100 mM 2-(N-morpholino)-ethanesulfonic acid, pH 6.0, under a layer of paraffin oil. The plates were kept at 4°C. Crystals appeared after 48 hours, and were harvested 8 days later. They were cryo-protected by addition to each well of 15 µl of a solution containing 10 mM Tris-HCl, pH 8.0, 10 mM NaCl, 5 mM $MgSO_4$, 0.5 mM ADP, 50 mM 2-(N-

morpholino)-ethanesulfonic acid, pH 6.0, 5% (w/v) polyethylene glycol 12000 and 30% (v/v) glycerol. After 5 min, the crystals were harvested with a microloop (Mitegen, Ithaca, U. S. A.), flash-frozen and stored in liquid nitrogen.

Data Collection and Structure Determination. X-Ray diffraction data were collected at 100 K from cryoprotected crystals with a PILATUS3 2M detector (Dectris, Baden-Daettwil Switzerland), at a wavelength of 0.966 Å at the European Synchrotron Radiation Facility, Grenoble, France using the MXPressE automated screening protocol (63, 64). Diffraction images were integrated with iMOSFLM (65), and the data were reduced with AIMLESS (66). Anisotropic correction was applied using STARANISO (http://staraniso.globalphasing.org). Molecular replacement using the $\alpha_3\beta_3$-domain from the structure of the ground state structure of bovine F_1-ATPase (Protein Data Bank ID code 2JDI) was carried out with PHASER (67). Nucleotides, magnesium ions and water molecules were removed from the model. Rigid body refinement and restrained refinement were performed with REFMAC5 (68). Manual rebuilding was performed with COOT (69), alternating with refinement performed with REFMAC5. For calculations of R_{free}, 5% of the diffraction data were excluded from the refinement. Additional electron density features, adjacent to the α-subunits, were attributed to p18. Initially, poly-Ala α-helices were fitted into this additional density with COOT (69), and the assignment of the direction of the α-helices, was guided by secondary structure predictions performed with PSIPRED (70). This prediction also detected structural homology of p18 with PPR10 from *Z. mays* (pdb code 4M59). Stereochemistry was assessed with MolProbity (71), and images of structures and electron density maps were prepared with PyMOL (72). Structural comparisons of *T. brucei* F_1-ATPase with bovine F_1-ATPase inhibited with dicyclohexylcarbodiimide (DCCD) (pdb code 1E79) (12), bovine F_1-ATPase crystallised in the presence of phosphonate (pdb code 4ASU) (20), bovine F_1-ATPase inhibited with ADP-AlF_4 (pdb code 1H8E) (16), the ground state structure of yeast F_1-ATPase (pdb code 2HLD) (24) and of the p18-subunit from

T. brucei with PPR10 from *Z. mays* (pdb code 4M59) (57) were made with COOT (69) and PyMOL (72). The p18-subunit was assessed for the presence of PPR and TPR sequences with TPRPred (73), and α-helices were assigned according to PyMol.

Acknowledgements We thank staff at beam-line ID30A-1 MASSIF-1 at the European Synchrotron Radiation Facility, Grenoble, France for their help. This work was supported by the Medical Research Council, UK by grants MC_U105663150 and MR/M009858/1 to J. E. W., and MC_U105184325 A. G. W. L.; the Czech Republic Ministry of Education, Youth and Sports (European Research Council CZ Grant LL1205 to A. Z.); the Postdok_BIOGLOBE project, cofinanced by the European Social Fund and the Czech Republic (Grant CZ.1.07/2.3.00/30.0032 to A.Z. and O.G.); and EMBO (Short-term fellowship ASTF 81-2016 to O.G.).

References

1. Walker JE (2013) The ATP synthase: the understood, the uncertain and the unknown. *Biochem Soc Trans* 41(1):1–16.
2. Walker JE (2017) Chapter 13 Structure, mechanism and regulation of ATP synthases. *Mechanisms of Primary Energy Transduction in Biology*, ed Wikström M (The Royal Society of Chemistry, London), pp 338–373.
3. Abrahams JP, Leslie AGW, Lutter R, Walker JE (1994) Structure at 2.8 Å resolution of F_1-ATPase from bovine heart mitochondria. *Nature* 370(6491):621–628.
4. Abrahams JP *et al.* (1996) The structure of bovine F_1-ATPase complexed with the peptide antibiotic efrapeptin. *Proc Natl Acad Sci U S A* 93(18):9420–9424.
5. Bason JV, Montgomery MG, Leslie AGW, Walker JE (2014) Pathway of binding of the intrinsically disordered mitochondrial inhibitor protein to F_1-ATPase. *Proc Natl Acad Sci U S A* 111(31):11305–11310.

6. Bason JV, Montgomery MG, Leslie AGW, Walker JE (2015) How release of phosphate from mammalian F_1-ATPase generates a rotary substep. *Proc Natl Acad Sci U S A* 112(19):6009–6014.

7. Bowler MW, Montgomery MG, Leslie AGW, Walker JE (2006) How azide inhibits ATP hydrolysis by the F-ATPases. *Proc Natl Acad Sci U S A* 103(23):8646–8649.

8. Bowler MW, Montgomery MG, Leslie AGW, Walker JE (2007) Ground state structure of F_1-ATPase from bovine heart mitochondria at 1.9 Å resolution. *J Biol Chem* 282(19):14238–14242.

9. Braig K, Menz RI, Montgomery MG, Leslie AGW, Walker JE (2000) Structure of bovine mitochondrial F_1-ATPase inhibited by Mg^{2+} ADP and aluminium fluoride. *Structure* 8(6):567–573.

10. Cabezón E, Montgomery MG, Leslie AGW, Walker JE (2003) The structure of bovine F_1-ATPase in complex with its regulatory protein IF_1. *Nat Struct Biol* 10(9):744–750.

11. Dickson VK, Silvester JA, Fearnley IM, Leslie AGW, Walker JE (2006) On the structure of the stator of the mitochondrial ATP synthase. *EMBO J* 25(12):2911–2918.

12. Gibbons C, Montgomery MG, Leslie AGW, Walker JE (2000) The structure of the central stalk in bovine F_1-ATPase at 2.4 Å resolution. *Nat Struct Biol* 7(11):1055–1061.

13. Gledhill JR, Montgomery MG, Leslie AGW, Walker JE (2007) How the regulatory protein, IF_1, inhibits F_1-ATPase from bovine mitochondria. *Proc Natl Acad Sci U S A* 104(40):15671–15676.

14. Gledhill JR, Montgomery MG, Leslie AGW, Walker JE (2007) Mechanism of inhibition of bovine F_1-ATPase by resveratrol and related polyphenols. *Proc Natl Acad Sci U S A* 104(34):13632–13637.

15. Kagawa R, Montgomery MG, Braig K, Leslie AGW, Walker JE (2004) The structure of bovine F_1-ATPase inhibited by ADP and beryllium fluoride. *EMBO J* 23(14):2734–2744.

16. Menz RI, Walker JE, Leslie AGW (2001) Structure of bovine mitochondrial F_1-ATPase with nucleotide bound to all three catalytic sites: implications for the mechanism of rotary catalysis. *Cell* 106(3):331–341.

17. Menz RI, Leslie AGW, Walker JE (2001) The structure and nucleotide occupancy of bovine mitochondrial F_1-ATPase are not influenced by crystallisation at high concentrations of nucleotide. *FEBS Lett* 494(1-2):11–14.

18. Orriss GL, Leslie AGW, Braig K, Walker JE (1998) Bovine F_1-ATPase covalently inhibited with 4-chloro-7-nitrobenzofurazan: the structure provides further support for a rotary catalytic mechanism. *Structure* 6(7):831–837.

19. Rees DM, Leslie AGW, Walker JE (2009) The structure of the membrane extrinsic region of bovine ATP synthase. *Proc Natl Acad Sci U S A* 106(51):21597–21601.

20. Rees DM, Montgomery MG, Leslie AGW, Walker JE (2012) Structural evidence of a new catalytic intermediate in the pathway of ATP hydrolysis by F_1-ATPase from bovine heart mitochondria. *Proc Natl Acad Sci U S A* 109(28):11139–11143.

21. van Raaij MJ, Abrahams JP, Leslie AGW, Walker JE (1996) The structure of bovine F_1-ATPase complexed with the antibiotic inhibitor aurovertin B. *Proc Natl Acad Sci U S A* 93(14):6913–6917.

22. Watt IN, Montgomery MG, Runswick MJ, Leslie AGW, Walker JE (2010) Bioenergetic cost of making an adenosine triphosphate molecule in animal mitochondria. *Proc Natl Acad Sci U S A* 107(39):16823–16827.

23. Zhou A *et al.* (2015) Structure and conformational states of the bovine mitochondrial ATP synthase by cryo-EM. *Elife* 4:e10180.

24. Kabaleeswaran V, Puri N, Walker JE, Leslie AGW, Mueller DM (2006) Novel features of the rotary catalytic mechanism revealed in the structure of yeast F_1 ATPase. *EMBO J* 25(22):5433–5442.

25. Kabaleeswaran V *et al.* (2009) Asymmetric structure of the yeast F_1 ATPase in the absence of bound nucleotides. *J Biol Chem* 284(16):10546–10551.

26. Stock D, Leslie AGW, Walker JE (1999) Molecular architecture of the rotary motor in ATP synthase. *Science* 286(5445):1700–1705.

27. Robinson GC *et al.* (2013) The structure of F_1-ATPase from *Saccharomyces cerevisiae* inhibited by its regulatory protein IF_1. *Open Biol* 3(2):120164.

28. Arsenieva D, Symersky J, Wang Y, Pagadala V, Mueller DM (2010) Crystal structures of mutant forms of the yeast F1 ATPase reveal two modes of uncoupling. *J Biol Chem* 285(47):36561–36569.

29. Vinothkumar KR, Montgomery MG, Liu S, Walker JE (2016) Structure of the mitochondrial ATP synthase from *Pichia angusta* determined by electron cryo-microscopy. *Proc Natl Acad Sci U S A* 113(45):12709–12714.

30. Hahn A *et al.* (2016) Structure of a complete ATP synthase dimer reveals the molecular basis of inner mitochondrial membrane morphology. *Mol Cell* 63(3):445–456.

31. Groth G, Pohl E (2001) The structure of the chloroplast F1-ATPase at 3.2 A resolution. *J Biol Chem* 276(2):1345–1352.

32. Groth G (2002) Structure of spinach chloroplast F_1-ATPase complexed with the phytopathogenic inhibitor tentoxin. *Proc Natl Acad Sci U S A* 99(6):3464–3468.

33. Shirakihara Y *et al.* (1997) The crystal structure of the nucleotide-free alpha 3 beta 3 subcomplex of F_1-ATPase from the thermophilic *Bacillus* PS3 is a symmetric trimer. *Structure* 5(6):825–836.

34. Cingolani G, Duncan TM (2011) Structure of the ATP synthase catalytic complex F_1 from *Escherichia coli* in an autoinhibited conformation. *Nat Struct Mol Biol* 18(6):701–707.

35. Roy A, Hutcheon ML, Duncan TM, Cingolani G (2012) Improved crystallization of

Escherichia coli ATP synthase catalytic complex (F_1) by introducing a phosphomimetic mutation in subunit Ɛ. *Acta Cryst F* 68(Pt 10):1229–1233.

36. Shirakihara Y *et al.* (2015) Structure of a thermophilic F_1-ATPase inhibited by an ε-subunit: deeper insight into the ε-inhibition mechanism. *FEBS J* 282(15):2895–2913.
37. Morales-Rios E, Montgomery MG, Leslie AGW, Walker JE (2015) The structure of ATP synthase from *Paracoccus denitrificans* determined by X-ray crystallography at 4.0 Å resolution. *Proc Natl Acad Sci U S A* 112(43):13231–13236.
38. Ferguson SA, Cook GM, Montgomery MG, Leslie AGW, Walker JE (2016) Regulation of the thermoalkaliphilic F_1-ATPase from *Caldalkalibacillus thermarum*. *Proc Natl Acad Sci U S A* 113(39):10860–10865.
39. Sobti M *et al.* (2016) Cryo-EM structures of the autoinhibited E. coli ATP synthase in three rotational states. *Elife* 5
40. Walker JE, Saraste M, Runswick MJ, Gay NJ (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* 1(8):945–951.
41. Ramakrishnan C, Dani VS, Ramasarma T (2002) A conformational analysis of Walker motif A [GXXXXGKT (S)] in nucleotide-binding and other proteins. *Protein Eng* 15(10):783–798.
42. Mühleip AW, Dewar CE, Schnaufer A, Kühlbrandt W, Davies KM (2017) In situ structure of trypanosomal ATP synthase dimer reveals a unique arrangement of catalytic subunits. *Proc Natl Acad Sci U S A* 114(5):992–997.
43. Balabaskaran Nina P *et al.* (2010) Highly divergent mitochondrial ATP synthase complexes in Tetrahymena thermophila. *PLoS Biol* 8(7):e1000418.
44. Balabaskaran Nina P *et al.* (2011) ATP synthase complex of Plasmodium falciparum: dimeric assembly in mitochondrial membranes and resistance to genetic disruption. *J Biol*

Chem 286(48):41312–41322.

45. Cardol P *et al.* (2005) The mitochondrial oxidative phosphorylation proteome of Chlamydomonas reinhardtii deduced from the Genome Sequencing Project. *Plant Physiol* 137(2):447–459.

46. van Lis R, Mendoza-Hernández G, Groth G, Atteia A (2007) New insights into the unique structure of the F0F1-ATP synthase from the chlamydomonad algae Polytomella sp. and Chlamydomonas reinhardtii. *Plant Physiol* 144(2):1190–1199.

47. Vaidya AB, Mather MW (2009) Mitochondrial evolution and functions in malaria parasites. *Annu Rev Microbiol* 63:249–267.

48. Perez E *et al.* (2014) The mitochondrial respiratory chain of the secondary green alga Euglena gracilis shares many additional subunits with parasitic Trypanosomatidae. *Mitochondrion* 19 Pt B:338–349.

49. Zíková A, Schnaufer A, Dalley RA, Panigrahi AK, Stuart KD (2009) The F_oF_1-ATP Synthase Complex Contains Novel Subunits and Is Essential for Procyclic *Trypanosoma brucei*. *PLoS Pathog* 5(5):e1000436.

50. Gahura O *et al.* (2018) The F_1-ATPase from *Trypanosoma brucei* is elaborated by three copies of an additional p18-subunit. *FEBS J* In press

51. Speijer D *et al.* (1997) Characterization of the respiratory chain from cultured Crithidia fasciculata. *Mol Biochem Parasitol* 85(2):171–186.

52. Nelson RE, Aphasizheva I, Falick AM, Nebohacova M, Simpson L (2004) The I-complex in *Leishmania tarentolae* is an uniquely-structured F_1-ATPase. *Mol Biochem Parasitol* 135(2):221–224.

53. Dean S, Gould MK, Dewar CE, Schnaufer AC (2013) Single point mutations in ATP synthase compensate for mitochondrial genome loss in trypanosomes. *Proc Natl Acad Sci U S A* 110(36):14741–14746.

54. Carbajo RJ *et al.* (2007) How the N-terminal domain of the OSCP subunit of bovine F1Fo-ATP synthase interacts with the N-terminal region of an alpha subunit. *J Mol Biol* 368(2):310–318.

55. Pusnik M, Small I, Read LK, Fabbro T, Schneider A (2007) Pentatricopeptide repeat proteins in Trypanosoma brucei function in mitochondrial ribosomes. *Mol Cell Biol* 27(19):6876–6888.

56. Small ID, Peeters N (2000) The PPR motif - a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem Sci* 25(2):46–47.

57. Yin P *et al.* (2013) Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature* 504(7478):168–171.

58. Cano-Estrada A *et al.* (2010) Subunit-subunit interactions and overall topology of the dimeric mitochondrial ATP synthase of Polytomella sp. *Biochim Biophys Acta* 1797(8):1439–1448.

59. Yadav KN *et al.* (2017) Atypical composition and structure of the mitochondrial dimeric ATP synthase from Euglena gracilis. *Biochim Biophys Acta* 1858(4):267–275.

60. Allegretti M *et al.* (2015) Horizontal membrane-intrinsic α-helices in the stator a-subunit of an F-type ATP synthase. *Nature* 521(7551):237–240.

61. Šubrtová K, Panicucci B, Zíková A (2015) ATPaseTb2, a unique membrane-bound FoF1-ATPase component, is essential in bloodstream and dyskinetoplastic trypanosomes. *PLoS Pathog* 11(2):e1004660.

62. Panicucci B, Gahura O, Zíková A (2017) Trypanosoma brucei $TbIF_1$ inhibits the essential F_1-ATPase in the infectious form of the parasite. *PLoS Negl Trop Dis* 11(4):e0005552.

63. Svensson O, Malbet-Monaco S, Popov A, Nurizzo D, Bowler MW (2015) Fully automatic characterization and data collection from crystals of biological macromolecules. *Acta Cryst D* 71(Pt 8):1757–1767.

64. Bowler MW *et al.* (2015) MASSIF-1: a beamline dedicated to the fully automatic characterization and data collection from crystals of biological macromolecules. *J Synchrotron Radiat* 22(6):1540–1547.

65. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AGW (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Cryst D* 67(Pt 4):271–281.

66. Evans PR, Murshudov GN (2013) How good are my data and what is the resolution. *Acta Cryst D* 69(Pt 7):1204–1214.

67. McCoy AJ *et al.* (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658–674.

68. Murshudov GN *et al.* (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst D* 67(Pt 4):355–367.

69. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Cryst D* 66(Pt 4):486–501.

70. Buchan DW, Minneci F, Nugent TC, Bryson K, Jones DT (2013) Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res* 41(Web Server issue):W349–57.

71. Chen VB *et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Cryst D* 66(Pt 1):12–21.

72. Schrodinger, LLC (2015) The PyMOL Molecular Graphics System, Version 1.8.

73. Alva V, Nam SZ, Söding J, Lupas AN (2016) The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. *Nucleic Acids Res* 44(W1):W410–5.

Figures

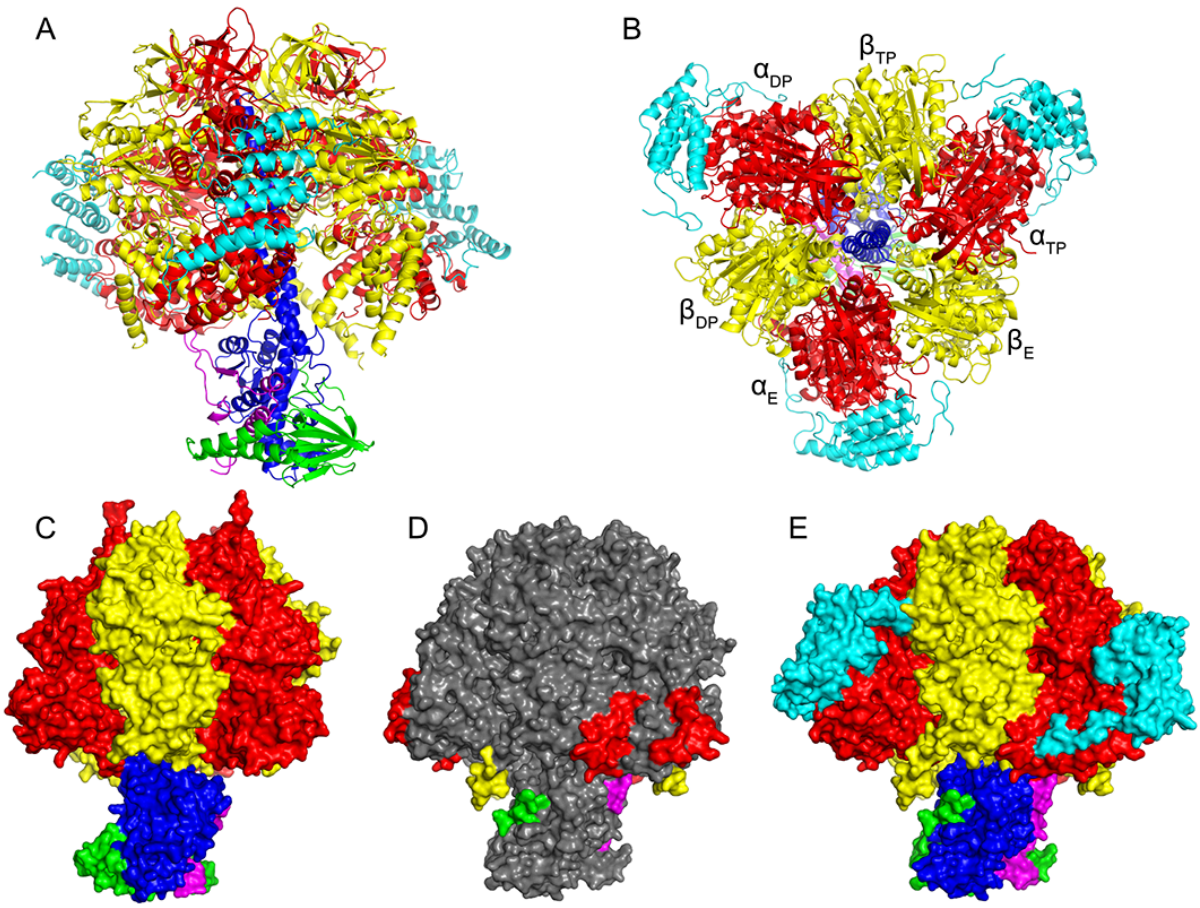

Fig. 1. Structure of the F_1-ATPase from *T. brucei*. The α-, β-, γ-, δ-, ε- and p18-subunits are red, yellow, blue, green, magenta and cyan, respectively. (*A*) Side and (*B*) top views in cartoon representation. (*C*), (*D*) and (*E*), side views in surface representation rotated 180° relative to (*A*). In (*C*), the bovine enzyme (12). (*D*) and (*E*), the *T. brucei* enzyme. In (*D*) p18 has been omitted, and only additional regions not found in the bovine enzyme are colored; the rest of the structure is gray. The two additional sections in the α-subunit (red) interact with the p18-subunit. (*E*), p18 is present and is shown interacting with the α-subunit.

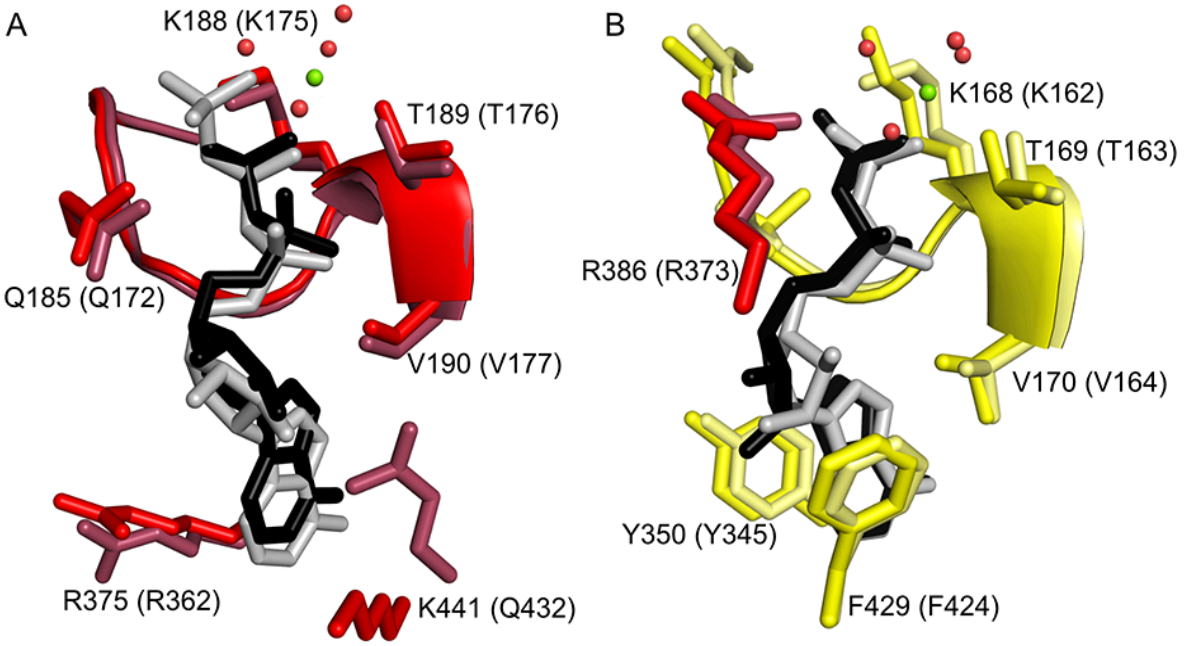

Fig. 2. Conservation of the non-catalytic and catalytic nucleotide binding sites in the F_1-ATPase from *T. brucei*. (*A*), The non-catalytic site in the α_{DP}-subunit superposed onto the equivalent site in the bovine enzyme (12); (*B*), the catalytic site in the β_{DP}-subunit superposed onto the equivalent site in the bovine enzyme. Residue αR386 is the catalytically essential arginine-finger (equivalent to αR373 in the bovine protein). Residues contributed by α- and β-subunits are red and yellow, respectively (with the bovine residues in muted colors), and the bound ADP molecules are black and grey in the *T. brucei* and bovine enzymes, respectively. The green and red spheres represent magnesium ions and water molecules, respectively (in *T. brucei* only). The residue numbers in parentheses denote the equivalent bovine residues.

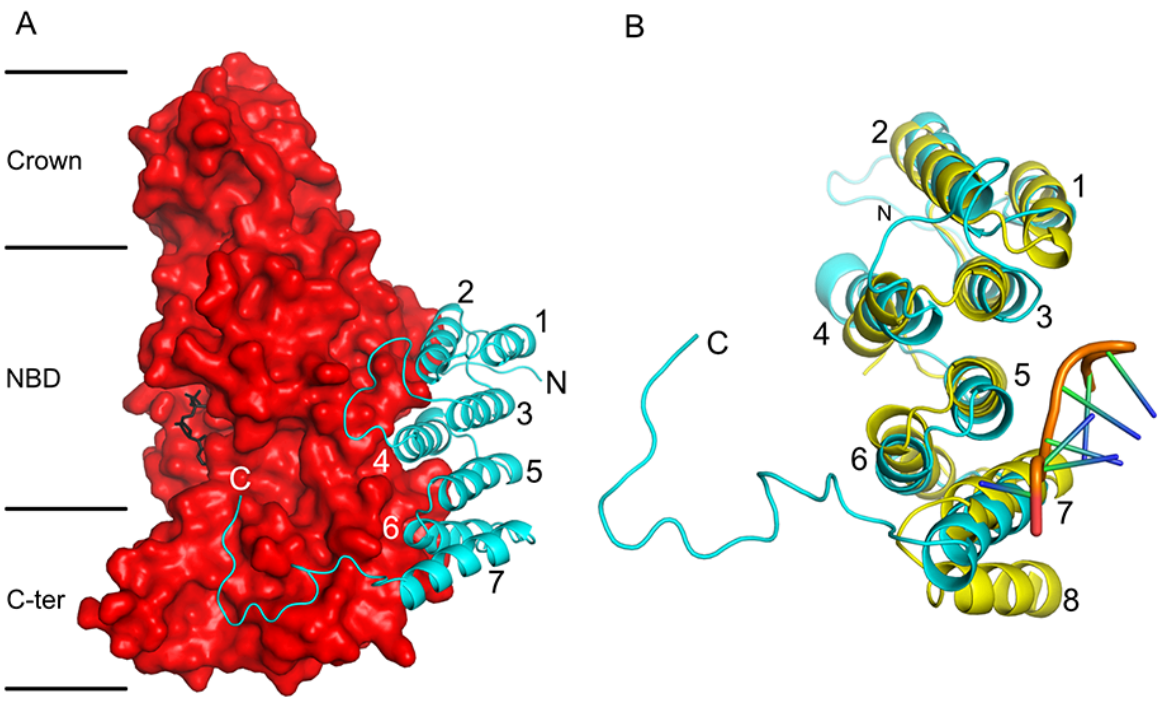

Fig. 3. Structure of the p18-subunit of the F_1-ATPase from *T. brucei*, and its relation to a PPR protein. (*A*), a p18 subunit (cyan) in cartoon representation, folded into α-helices H1-H7, with an extended C-terminal region from residues 151-170, bound to the α_{DP}-subunit in solid representation (red). The N-terminal, nucleotide binding and C-terminal domains of the α-subunit are indicated by Crown, NBD and C-ter, respectively; the bound ADP molecule is black; (*B*), comparison of the p18-subunit with an example of a PPR protein, the PPR10 protein from *Zea mays* (57) (yellow). PPR10 has eighteen PPRs, and the structures of PPRs 11-14 are shown (see Fig. S3). The orange region represents the back-bone of an 8 residue ribonucleotide bound to PPR10. The three PPRs in the p18-subunit correspond to H1 plus H2 (residues 20-28 and 33-45), H3 plus H4 (residues 52-64 and 78-93) and H5 plus H6 (residues 99-112 and 115-126). PPR10, has an additional α-helix, labelled 8, which, together with α-helix 7, constitutes a fourth PPR.

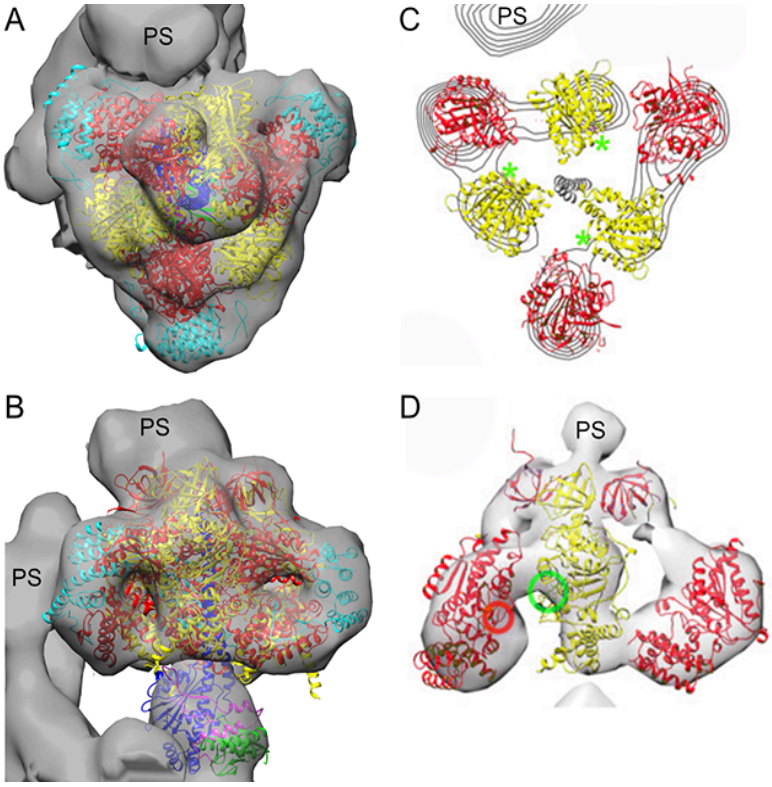

Fig. 4. Relationship of the crystallographic structure of the F_1-domain of the ATP synthase from *T. brucei* to an ECT map of the intact ATP synthase *in situ* in mitochondrial membranes from *T. brucei.* The subunits of the F_1-domain are coloured as in Fig. 1. (*A*) and (*B*), top and side views, respectively, of the ECT map (gray), determined independently at 32.5 Å resolution with the crystallographic structure of the F_1-domain determined at 3.2 Å resolution docked manually inside the ECT map with subunits α_{DP} and β_{TP} proximal to the peripheral stalk. (*C*) and (*D*), a published interpretation of the same ECT map proposing a structure of *T. brucei* F_1-ATPase where the α-subunits are opened away from the central stalk. It was proposed that the p18-subunit, which was not identified in the earlier published interpretation, contributes to the catalytic site by providing the arginine finger residue. The position of the conventional arginine-finger residue, Arg-386, in this interpretation, is indicated by a red circle in (*D*) (42). The catalytic sites are indicated by green asterisks in (*C*) and by a green circle in (*D*). PS, peripheral stalk of the enzyme. Panels *C* and *D* were adapted from Fig. 4 *B* and *E* (42).

Supplementary information for:

ATP synthase from *Trypanosoma brucei* has an elaborated canonical F_1-domain and conventional catalytic sites

Martin G. Montgomery[a, 1], Ondřej Gahura[a, b, 1], Andrew G. W. Leslie[c], Alena Zíková[b], and John E. Walker[a, 2]

[a]The Medical Research Council Mitochondrial Biology Unit, University of Cambridge, Cambridge Biomedical Campus, Hills Road, Cambridge CB2 0XY, United Kingdom; [b]Biology Centre, Czech Academy of Sciences, Institute of Parasitology, České Budějovice, 37005, Czech Republic; [c]The Medical Research Council Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Ave, Cambridge, CB2 0QH, United Kingdom

Table S1. Data collection and refinement statistics for the F_1-ATPase from *T. brucei.* Parentheses denote the statistics of the high-resolution bin.

Parameter	Value
Space group	$P2_1$
Unit cell dimensions a, b, c (Å); β	124.2, 206.3, 130.2; 104.9°
Resolution range (Å)	3.20-90.51
High-resolution bin (Å)	3.20-3.25
No. of unique reflections	102391 (5076)
Multiplicity	3.5 (3.7)
Completeness (%)	98.2 (97.6)
R_{merge}[1]	0.123 (0.565)
<I/σ (I)>	8.1 (2.8)
B factor, from Wilson plot (Å)^2	46.5
R factor[2] (%)	27.2
Free R factor[3] (%)	29.7
rmsd of bonds (Å)	0.007
rmsd of angles (°)	1.07

[1] $R_{merge} = \sum_h\sum_i|I(h)-I(h)_i|/\sum_h\sum_i I(h)_i$, where $I(h)$ is the mean weighted intensity after rejection of outliers

[2] R factor = $\sum_{hkl} ||F_{obs}| - k|F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively.

[3] $R_{free} = \sum_{hkl \subset T} ||F_{obs}| - k|F_{calc}|| / \sum_{hkl \subset T} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and the calculated structure factor amplitudes, respectively, and T is the test set of data omitted from refinement.

Table S2. Buried surface areas of catalytic interfaces in selected structures of F_1-ATPases.

Structure	Buried surface area of catalytic interface ($Å^2$)		
	DP	TP	E
T. brucei	2300	2400	1900
S. cerevisiae (2HLD; molecule 1)	2000	2300	1900
Bovine phosphate release dwell (2JDI)	3000	2200	1900
Bovine catalytic dwell (4ASU)	2500	2200	1800
Bovine inhibited with ADP-AlF_4 (1H8E)	2800	2100	2300
C. thermarum wild type (5HKK)	2900	2000	1900

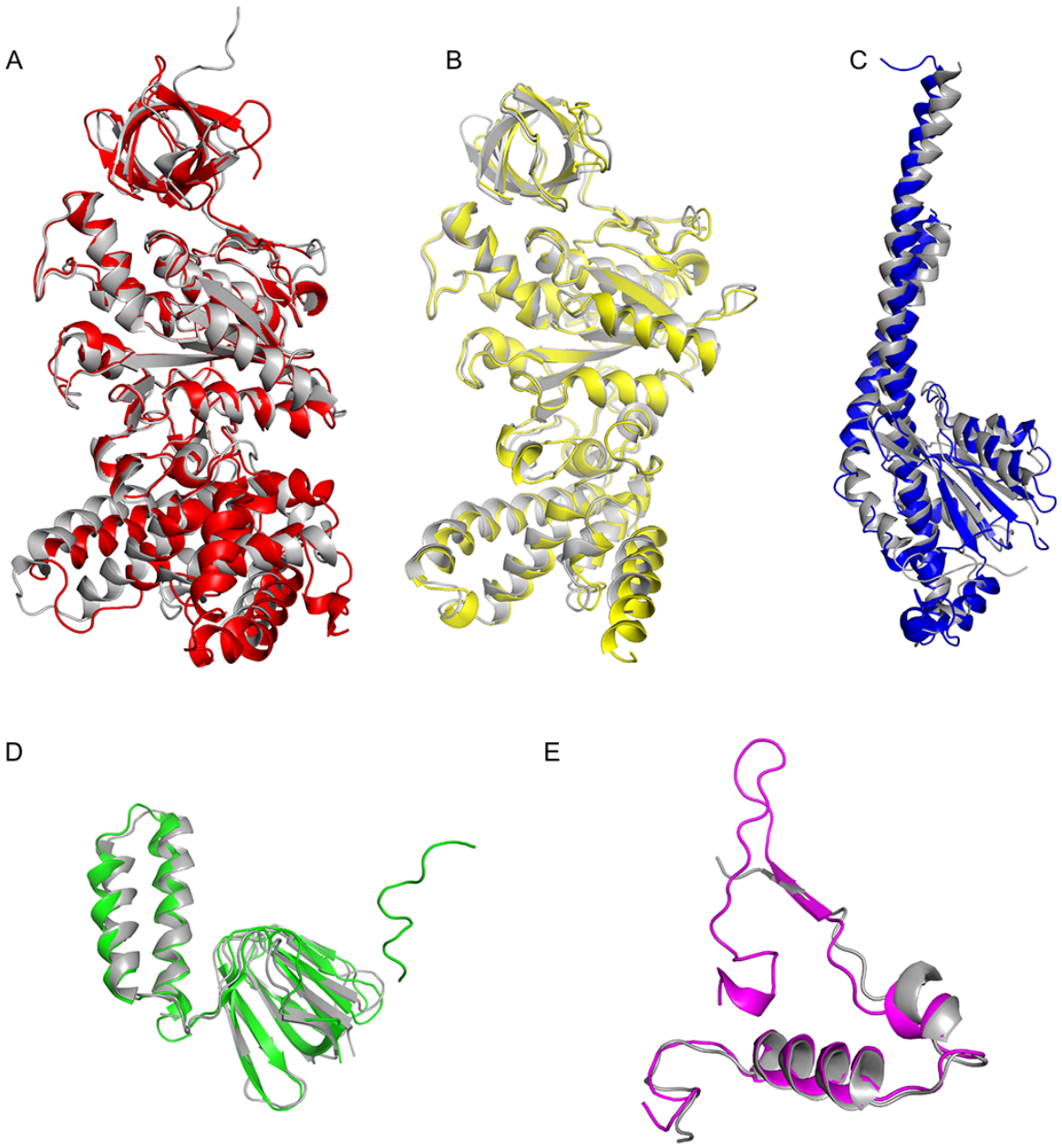

Fig. S1. Comparison of individual subunits of the F_1-ATPase from *T. brucei* with orthologs in the bovine F_1-ATPase inhibited by dicyclohexylcarbodiimide (12). The *T. brucei* subunits are colored and bovine subunits are grey. (*A*) α_{DP}-subunit (red; rmsd 2.3 Å); (*B*) β_{DP}-subunit (yellow; rmsd 2.0 Å); (*C*) the γ-subunit (blue; rmsd 3.8 Å); (*D*) the δ-subunit (green; rmsd 2.1 Å); (*E*) the ε-subunit (magenta; rmsd 1.8 Å).

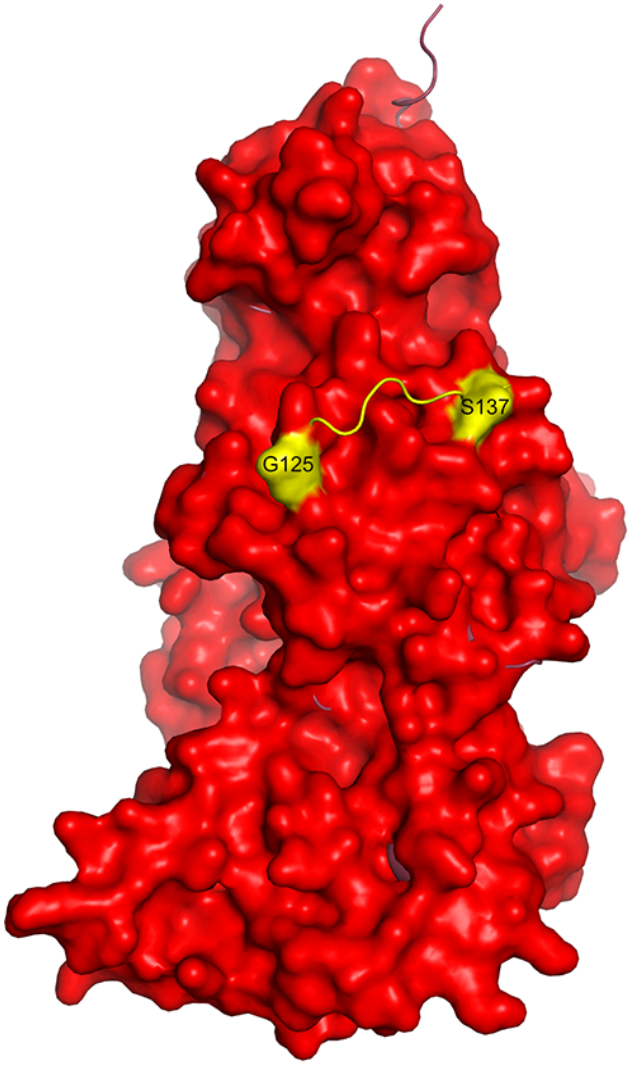

Fig. S2. The sites of proteolytic cleavage in the α-subunits of the F_1-ATPase from *T. brucei*. The α_{DP}-subunit from *T. brucei*, in surface representation, overlaid onto the equivalent subunit from the bovine enzyme inhibited by dicyclohexylcarbodiimide (12), in cartoon representation (mainly hidden by the *T. brucei* surface). The proteolytic cleavages in the *T. brucei* subunit follow residues 127 and 135 removing residues 128-135. Residues G125 and S137 are yellow. The equivalent region in the bovine α_{DP}-subunit, residues 117-123 is not cleaved by proteolysis and forms a solvent exposed loop in yellow.

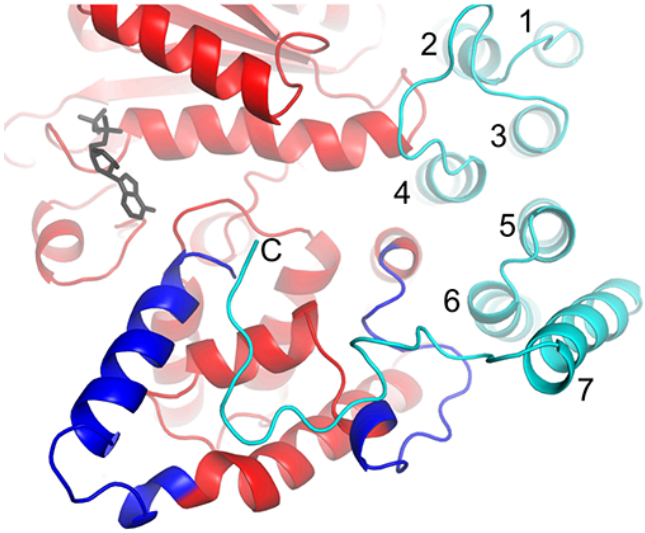

Fig. S3. The roles of the additional segments of sequence in the α-subunits of the F_1-ATPase from *T. brucei*. The parts of the nucleotide binding and C-terminal domains of the α_E-subunit that interact with the p18-subunit (cyan) are shown in red, except for the additional segments (residues 483-498 and 536-560), which are royal blue for emphasis. A bound ADP molecule is black. α-Helices in subunit p18 are numbered 1-7.

```
      18  -TNTAPWIEKIKKCKYYDEAGEVLVNMNVSNCPPDI--  52
p18   57  -ATLQCIYQSPSKQSTPVDNESKFCAMMDLLEEMQH--  91
      100 -ESWTWVMKECVKSGQFRLGYCIQQVMETECKGCPA-- 142
          --TYNALINAYAK-G--EEA--LY--M---G--PN---   PPRcon
      71  -SALEMVVRALGREGQHDAVCALLDETPLPPGSRLD-- 105
      107 -RAYTTVLHALSRAGRYERALELFAELRRQGVAPTL-- 141
      142 -VTYNVVLDVYGRMGRSWPRIVALLDEMRAAGVEPD-- 176
      178 -FTASTVIAACSRDGLVDEAVAFFEDLKARGHAPSV-- 212
      213 -VTYNALLQVFGKAGNYTEALRVLGEMEQNGCQPDA-- 247
      248 -VTYNELAGTYARAGFFEEAARCLDTMASKGLLPNA-- 282
      283 -FTYNTVMTAYGNVGKVDEALALFDQMKKTGFVPNV-- 317
      318 -NTYNLVLGMLGKKSRFTVMLEMLGEMSRSGCTPNR-- 352
PPR10 353 -VTWNTMLAVSGKRGMEDYVTRVLEGMRSSGVELSR-- 387
      388 -DTYNTLIAAYGRCGSRTNAFKMYNEMTSAGFTPCI-- 422
      423 -TTYNALLNVLSRQGDWSTAQSIVSKMRTKGFKPNE-- 457
      458 -QSYSLLLQCYAKGGNVAGIAAIENEVYGSGAVFPS-- 492
      494 -VILRTLVIANFKCRRLDGMETAFQEVKARGYNPDL-- 528
      529 -VIFNSMLSIYAKNGMYSKATEVFDSIKRSGLSPDL-- 563
      564 -ITYNSLMDMYAKCSESWEAEKILNQLKCSQTMKPD-- 598
      600 -VSYNTVINGFCKQGLVKEAQRVLSEMVADGMAPCA-- 634
      635 -VTYHTLVGGYSSLEMFSEAREVIGYMVQHGLKPME-- 669
      670 -LTYRRVVESYCRAKRFEEARGFLSEVSETDLDFDK-- 704
```

Fig. S4. Comparison of the PPR sequences in the p18-subunit of the F_1-ATPase from *T. brucei* and the 18 PPRs in the PPR10 protein from *Z. mays* with the PPR consensus sequence (red) (56). The sequences of the four PPR10 domains used in the structural comparison shown in Fig. 3 are blue. The PPRs in subunit p18 were predicted with TPRPred (73). The alignment was produced manually.

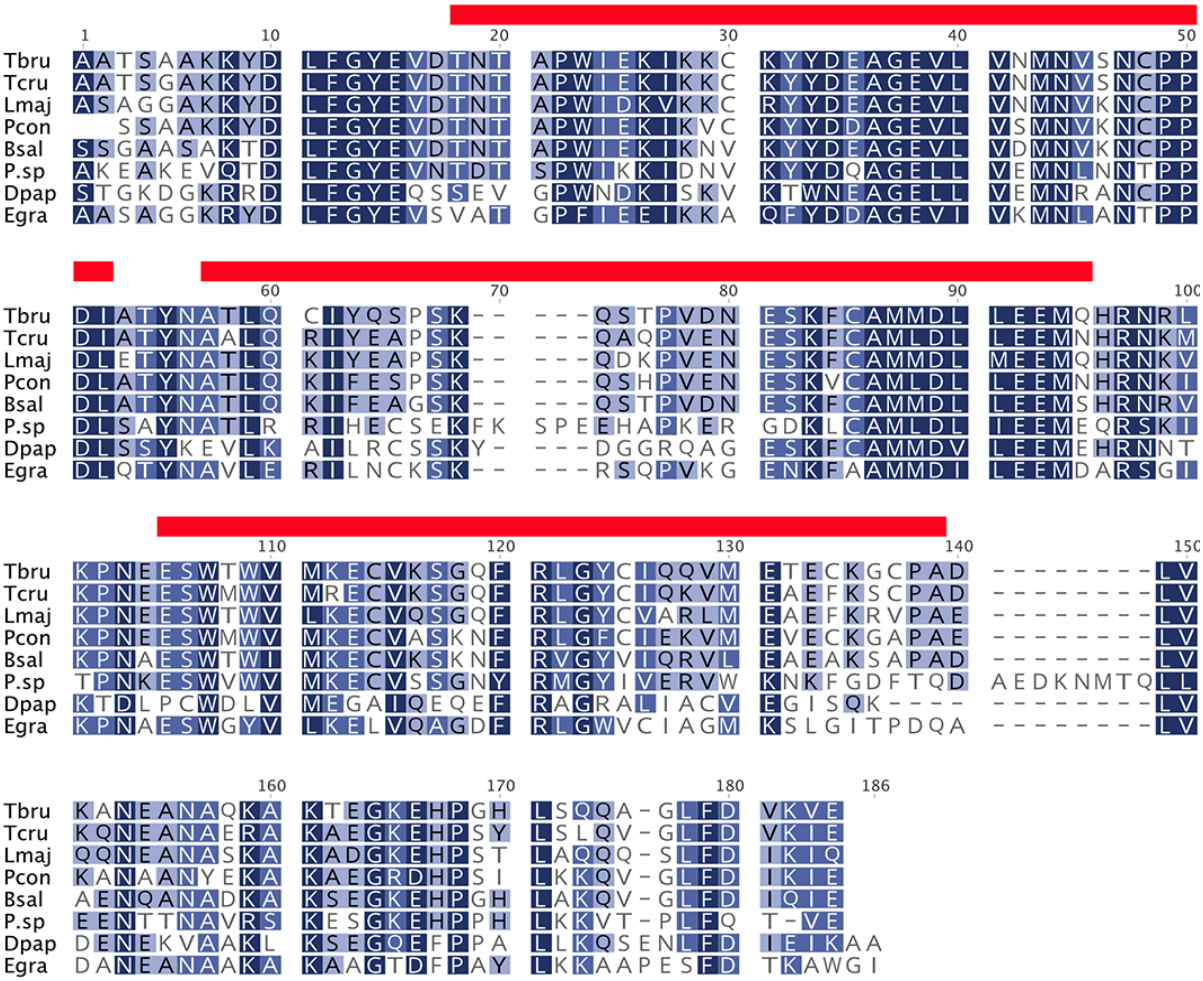

Fig S5. Conservation of sequences of p18-subunits of ATP synthases from euglenozoa. Tbru, *T. brucei*; Tcru, *T. cruzi*; Lmaj; *Leishmania major*; Pcon, *Paratrypanosoma confusum;* Bsal, *Bodo saltans*; P. sp, *Perkinsela sp*; Dpap, *Diplonema papillatum*; Egra, *Euglena gracilis*. Identities and conservative substitutions are dark and light blue, respectively. The red bars indicate the PPR domains predicted by TPRPred (73).