MS#2018-17615P (re-revised)

The structure of the catalytic domain of the ATP synthase from *Mycobacterium smegmatis* is a target for developing antitubercular drugs

Alice Tianbu Zhang[a1], Martin G. Montgomery[a1], Andrew G. W. Leslie[b], Gregory M. Cook[a,c], 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]*The Medical Research Council Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Ave, Cambridge, CB2 0QH, United Kingdom;* [c]*Department of Microbiology and Immunology, University of Otago, Dunedin 9016, New Zealand*

[1]Equal contributions

[2]To whom correspondence should be addressed. J. E. W., Tel.: +44-1223-252701; e-mail: walker@mrc-mbu.cam.ac.uk

Running title: Mycobacterial F_1-ATPase

The authors declare no conflict of interest

Author contributions: J. E. W. and G. M.C. designed research; J. E. W. supervised project; A.T. Z. and M. G. M. performed research; A.T. Z., M. G. M., G. M. C., A. G. W. L. and J. E. W. analyzed data; J. E. W. wrote the text and M. G. M. prepared the figures.

Key words: *Mycobacterium smegmatis*; F_1-ATPase; structure; inhibition; tuberculosis

PDB accession code: 6FOC

Significance statement

Last year, 1.6 million people died from tuberculosis (TB) and about 10 million became infected with the causative bacterium, *Mycobacterium tuberculosis*, 460,000 of them with multidrug resistant bacteria. Bedaquiline, a new anti-TB drug, developed to combat multi-drug resistant TB, kills *M. tuberculosis* by preventing the operation of its molecular machine for generating adenosine triphosphate (ATP), the fuel of life, without affecting the equivalent human machine. Here, we describe the molecular structure of the module in the mycobacterial machine where ATP is generated. Differences between this module and the equivalent human module can now be exploited to develop new anti-TB drugs, unrelated to bedaquiline, that also may help to prevent and cure TB by inhibiting the formation of ATP.

[118 words]

Summary

The crystal structure has been determined of the F_1-catalytic domain of the ATP synthase from *Mycobacterium smegmatis* which hydrolyzes adenosine triphosphate (ATP) very poorly. The structure of the $\alpha_3\beta_3$-component of the catalytic domain is similar to those in active F_1-ATPases in *Escherichia coli* and *Geobacillus stearothermophilus*. However, its ε-subunit differs from those in these two active bacterial F_1-ATPases as an ATP molecule is not bound to the two α-helices forming its C-terminal domain, probably because they are shorter than those in active enzymes and they lack an amino acid that contributes to the ATP binding site in active enzymes. In *E. coli* and *G. stearothermophilus*, the α-helices adopt an "up" state where the α-helices enter the $\alpha_3\beta_3$-domain and prevent the rotor from turning. The mycobacterial F_1-ATPase is most similar to the F_1-ATPase from

Caldalkalibacillus thermarum, which also hydrolyzes ATP poorly. The β_E-subunits in both enzymes are in the usual "open" conformation, but appear to be occupied uniquely by the combination of an ADP molecule with no magnesium ion, plus phosphate. This occupation is consistent with the finding that their rotors have been arrested at the same point in their rotary catalytic cycles. These bound hydrolytic products are probably the basis of inhibition of ATP hydrolysis. It can be envisaged that specific as yet unidentified small molecules might bind to the F_1-domain in *M. tuberculosis*, prevent ATP synthesis and inhibit the growth of the pathogen.

[247 words]

\body

Introduction

In 2017, around 1.6 million people died from tuberculosis (TB), and *Mycobacterium tuberculosis*, the causative bacterium, is now the second greatest killer of mankind by a single infectious agent, surpassed in its lethal impact only by HIV/AIDS (1). Of the 10 million people estimated to have developed TB in that single year, 4.6% were resistant to both rifampicin and isoniazid and are classed as multidrug-resistant (MDR), and 8.5% of the MDR-TB cases were extensively drug-resistant (XDR). Only 55% of the MDR-TB and 30% of the XDR-TB cases were treated successfully. These alarming statistics serve to emphasize the urgent need to develop new drugs that are effective and fast-acting against drug-resistant strains of *M. tuberculosis*. Preferably, they should be effective also against latent *M. tuberculosis* infections, where the bacteria lie dormant in infected humans in a non-replicating state before emerging as an active infection. It has been estimated that between a quarter and a third of the world's population are latently infected (2). However, the impact of this latency has been questioned recently as nearly everyone who falls

seriously ill with TB does so within two years of being infected, and latent infections rarely become active, even in old age (3).

In 2012, the US Food and Drug Administration approved the use of a novel oral drug bedaquiline (SIRTURO™) for the treatment of MDR-TB (4, 5), and bedaquiline received fast-track approval as a component of a combination therapy for the treating MDR-TB in adults. Its potential to shorten dramatically the treatment time for MDR-TB is highlighted by two recent studies. In mouse models of TB, a combination of bedaquiline with PA-824, an anti-mycobacterial drug with a complex mode of action (6) and linezolid, a repurposed protein synthesis inhibitor, significantly improved efficacy and relapse rates compared to the frontline regimen of rifampicin, isoniazid and pyrazinamide (7, 8). In the Nix-TB phase III clinical trial using this three-drug combination, 74% of the patients with MDR-TB were culture negative in 8 weeks*. The most recent recommendations for treatment of MDR-TB, based on the balance of effectiveness and harm and a preference for oral administration, now include bedaquiline (9). Bedaquiline is effective against both actively growing and non-replicating cells of *M. tuberculosis*, and acts by inhibiting the ATP synthase (10, 11) thereby shutting off the supply of cellular energy in the bacterium, without noticeably affecting the human enzyme found in the inner membranes of mitochondria. Thus, these observations provide proof of principle that the mycobacterial ATP synthase is a suitable target for developing new drugs to combat tuberculosis. A rational approach to the design of new drugs in addition to bedaquiline to inhibit the mycobacterial ATP synthase requires ideally that the structures and mechanistic and regulatory mechanisms of the human and mycobacterial ATP synthases be understood. The human enzyme is very similar in both respects to the well-studied bovine enzyme, which therefore provides an excellent surrogate. Mycobacterial ATP synthases have been less

studied, and only the structure of the c-ring in the membrane domain of the enzyme's rotor in the enzyme from the non-pathogenic organism, *M. phlei*, has been established (12). It is here that bedaquiline binds (12), presumably impeding the turning of the rotor in the intact enzyme. It has been proposed that it also binds at a secondary site in the ε-subunit (13, 14). Before, the work described here, the structure of its F_1-catalytic domain was not known in any mycobacterial ATP synthase and there was no molecular understanding of why the mycobacterial enzymes are barely capable of hydrolyzing ATP (15), whereas, for example the enzymes from facultative anaerobes such as *Escherichia coli* can both synthesize and hydrolyze ATP. Here, we describe the structure of the inhibited state of the catalytic domain of the ATP synthase from another non-pathogenic mycobacterium, *Mycobacterium smegmatis*. It is an excellent surrogate for the catalytic domain of the F_1-ATPase from *M. tuberculosis*, as a comparison of the sequences of the subunits from various mycobacterial species demonstrates (*SI Appendix*, Fig. S1 and Table S1).

Results and Discussion

Characterization of F_1-ATPase from *M. smegmatis*. The nine subunits of the ATP synthase in *M. smegmatis* are encoded by the *atp* operon, which includes the cluster of genes *atpAGDC* encoding the constituent α-, γ-, β- and ε-subunits, respectively, of the F_1-ATPase complex. This cluster was amplified by PCR, modified to encode the N-terminus of the β-subunit fused to a hexahistidine tag with an intervening protease cleavage site, and the vector containing the four genes was introduced into *M. smegmatis*. Attempts to over-express the *M. tuberculosis* orthologs in *M. smegmatis* in the same way failed as the genes from the two mycobacterial species recombined. The over-expressed purified F_1-ATPase (*SI Appendix*, Fig. S2) was a single complex with a mass of 380 kDa, composed of the

expected complement of α-, β-, γ- and ε-subunits with their characteristic molecular masses (*SI Appendix*, Fig. S3 and Table S2). It had a very low ATP hydrolase activity (0.07 U/mg), and, in contrast to some other latent F_1-ATPases, for example the enzyme from *Caldalkalibacillus thermarum* (16), this low activity could not be stimulated by lauryldimethylamine oxide (LDAO). However, when the mycobacterial enzyme was treated briefly with trypsin, the specific activity increased by one hundred-fold to 7 U/mg. Although characterization of the proteolytic fragments (*SI Appendix*, Table S3) did not provide a clear indication of the mechanism of activation, it is worth noting that the ε-subunit had been degraded almost completely after 2 min, with a corresponding significant increase in activity. In *E. coli* F_1-ATPase and also F_1F_o-ATPase from *M. smegmatis*, activation by trypsinolysis has been attributed to removal of the ε-subunit (17). The activity of F_1-ATPase from *M. smegmatis* uncovered by trypsinolysis was doubled by the addition of LDAO (*SI Appendix*, Fig. S4).

Structure Determination. Hexagonal crystals of the complex containing all four subunits (*SI Appendix*, Fig. S5) have the unit-cell parameters a=b=105.2 Å, c=628.6 Å, with α=β=90.0°, γ=120.0° and belong to space group $P3_121$ with one F_1-ATPase in the asymmetric unit (see *SI Appendix*, Table S4 for summary of data collection and refinement statistics). The quality of the electron density map is indicated in *SI Appendix*, Fig. S6, where representative segments and their interpretation are shown. The structure (Fig. 1A) contains the following residues: α_E, 31-190, 202-511, and 1512-1522 (corresponding to the C-terminal extension, where the register is unclear. Here, the residue numbers have been increased by 1000 to indicate uncertainty, as required by the PDB); α_{TP}, 30-190, 202-406 and 414-511; α_{DP}, 30-190, 202-409 and 412-511; β_E, 9-41, 47-108, 116-132 and 136-471;

β_{TP}, 8-41, 47-108 and 114-471; β_{DP}, 9-41, 47-108 and 116-471; γ, 4-57, 84-108, 119-129, 139-163, 188-198 and 238-304; ε, 3-115. Overall, the structure is similar to those of F_1-ATPases determined previously in other species (*SI Appendix*, Fig. S7 and Table S5), and especially to the F_1-ATPase from *C. thermarum* (18) (Fig. 1B). Although the structures of the α_E-, α_{TP}- and α_{DP}-subunits terminate at residue 511, the sequences of the subunits extend to residue 548. This C-terminal extension is characteristic of α-subunits in mycobacteria, and is not found in other eubacterial, chloroplast or mitochondrial sequences (*SI Appendix*, Fig. S8). By two independent programmes this extension is predicted to be intrinsically disordered (*SI Appendix*, Fig. S9), and it was possible to build a segment of 10 residues of extended structure immediately following the C-terminal α-helix of the α_E-subunit (Fig. 1C). However, the sequence register could not be determined unambiguously, and so this segment is modelled as UNK (unknown) and numbered 1512-1522. In a peptide representing residues 521-540 it has been shown by solution nuclear magnetic resonance (NMR) that residues 526-539 are α-helical, (but no coordinates are available), and on the basis of structure prediction that this α-helical structure prevailed in the intact protein (19). However, the current prediction of intrinsic disorder in the entire C-terminal region from residues 512-548 (*SI Appendix*, Fig. S9), and the secondary structures of the subunit in seven mycobacterial species predicted with PSIPRED (20) (*SI Appendix*, Fig. S10), are not in accord with this proposal. Another program, Predator (21), used previously (19), gives an ambiguous answer, predicting an α-helix in the region of residues 530-540 in four out of the seven species including *M. tuberculosis*, but not *M. smegmatis* (Fig S10), *M. phlei* or *M. ulcerans*. In conclusion, on the basis of the current structure predictions, (and acknowledging that NMR studies show that residues 526-539 in the isolated segment from residues 521-540 are α-helical), it appears on balance to be unlikely that an α-helix forms

in this region of the intact α-subunit. Nonetheless, it remains possible that this additional region of the mycobacterial subunits could play a role in the regulation of the enzyme (19).

The $\alpha_3\beta_3$-Domain. In the nucleotide binding sites of the three α-subunits and the β_{TP}-subunit, additional electron density is compatible with them being occupied by an ADP molecule, plus a magnesium ion (Fig. 2). There is also additional electron density associated with the nucleotide binding sites in both the β_{DP}- and β_E-subunits. Although the additional density is discontinuous in the former, it can be interpreted plausibly as ADP plus a magnesium ion also. In the latter site, the additional density is associated with the region of the P-loop where the α- and β- phosphates of a nucleotide would be bound, as for example in the β_E-subunit of the F_1-ATPase from *C. thermarum* (18), (where an ADP molecule with no magnesium ion is bound at 50-100% occupancy when crystals were grown in the presence of 500 μM ADP). Both ADP and a single phosphate ion were tested near the P-loop at various occupancies, but neither refined well into this site. However, there is, also additional density above the P-loop where a phosphate ion has been modelled, as in the β_E-subunit of the *C. thermarum* enzyme where a phosphate ion sits above the ADP. Although this density has been modelled as phosphate, it could possibly be a sulfate introduced from the crystallization buffer. Superimposition of the structures of the F_1-ATPases from *M. smegmatis* and *C. thermarum* via their $\alpha_3\beta_3$-domains showed that the two structures are very similar (rmsd value 0.91 Å). The occupancy of nucleotide binding sites in the *C. thermarum* enzyme is ADP and a magnesium ion in the three α-subunits and in the β_{TP}- and β_{DP}-subunits, and ADP and a phosphate without a magnesium ion in the β_E-subunit. Therefore, although the additional density in the P-loop region of the β_E-subunit in the *M. smegmatis* enzyme remains uninterpreted, the close similarity of the structure to

the structure of the inhibited complex in *C. thermarum* suggests that the site is probably occupied by an ADP molecule (with no accompanying magnesium ion) at low occupancy.

The γ-Subunit. The structure of the γ-subunit is the least well-resolved of the eight constituent subunits of the enzyme, probably because it is not constrained by any contacts with other F_1-ATPase complexes in the crystal lattice. In *C. thermarum*, for example, where the subunit is constrained in the crystal lattice, it is resolved entirely apart from residues 1-3. The *C. thermarum* subunit is folded into two α-helices in its N- and C-terminal regions with an intervening Rossmann fold, as in other species that have been studied (Fig. 3) (18, 22–25). The N- and C-terminal á-helices make an antiparallel coiled-coil occupying the central axis of the $\alpha_3\beta_3$-domain, and the Rossmann fold has five β-strands, with α-helices between strands 1 and 2, 2 and 3, and 3 and 4. In the mycobacterial enzyme, the N- and C-terminal α-helices are well resolved except for approximately three α-helical turns at the N-terminus of the C-terminal α-helix. The intervening α-helices 2, 3 and 4 were also resolved, but none of the five β-strands and connecting loops in the Rossmann fold could be built. However, superposition of the fragmentary structure of the mycobacterial γ-subunit upon the *C. thermarum* γ-subunit is consistent with the structures of the two orthologs being closely similar. This structural similarity extends to the γ-subunits from *E. coli* (22), *P. denitrificans* (23) and spinach chloroplasts (24) (Fig. 3). The overall fold of these γ-subunits is also similar to the γ-subunits from the enzymes from bovine (25) (Fig. 3) and yeast (26) mitochondria, except that the N-terminal α-helices of the bacterial γ-subunits extend further in a C-terminal direction and they are less curved. However, the γ-subunits of the *M. smegmatis* and *C. thermarum* enzymes have an additional highly significant similarity that distinguishes them from the γ-subunits in the other F_1-ATPases, and supports the view that the two determined structures represent the same inhibited state

of the enzyme. In the *M. smegmatis* and *C. thermarum* γ-subunits, residues 22-33, the "rigid-body" regions (see Materials and Methods) are rotated approximately to the same extent (10.5° in *M. smegmatis*, and 9° and 13° in the two copies in the asymmetric unit of the crystals of the *C. thermarum* enzyme, respectively), whereas the rotation angles in the *E. coli*, *P. denitrificans*, and spinach chloroplast enzymes, which have different nucleotide occupancies in their catalytic sites to the *C. thermarum* enzyme, are 50°, 27° and 22° respectively.

A feature, discussed before, that distinguishes mycobacterial γ-subunits from non-mycobacterial species that might be a target for drug design, is that 12-14 amino acids are inserted in the region from residues 165-169 in the aligned sequences (*SI Appendix*, Fig. S11). This insertion is at the C-terminus of α-helix 4 (Fig. 3A), and is predicted to form a random coil (*SI Appendix*, Fig. S12) that might extend to the bacterial membrane surface (27). As this region is unresolved, it is not known whether this suggestion is correct.

The ε-Subunit. As in ATP synthases from other eubacteria, chloroplasts, and mitochondria (where the orthologous protein is known as the δ-subunit) (28), the *M. smegmatis* ε-subunit has two domains (Fig. 4). The N-terminal domain is folded into an eight-stranded β-sandwich, and is very similar to those in other species. For example, the rmsd values for the comparisons of the N-terminal domain of the ε-subunit from *M. smegmatis* with those from *E. coli* and *C. thermarum* are 1.2 and 1.0 Å, respectively. In contrast, the C-terminal domain differs substantially from those in orthologues. In *E. coli*, *C. thermarum, G. stearothermophilus*, and in bovine and yeast mitochondria, this region is folded into two α-helices, approximately 23 and 30 Å long. In *E. coli* (22, 29–32) and *G. stearothermophilus* (33–35), the α-helices adopt two different states, referred to as "down" and "up". In the "down" state of the F-ATPase from *G. stearothermophilus* (35), the α-

helices of the ε-subunit bind an ATP molecule, and are associated with the β-sandwich. In the absence of bound ATP, the α-helices assume the "up" position, where they lie alongside the γ-subunit and interact with the $\alpha_3\beta_3$-domain, inhibiting ATP hydrolysis. "Up" positions have been captured in structures of the F_1-domain from *E. coli* (22) and in the intact *E. coli* ATP synthase complex (32), but the isolated *E. coli* ε-subunit adopts a "down" conformation although ATP is not bound to it suggesting that ATP does not influence the position of the ε-subunit (29–31). In the F_1-ATPase from *C. thermarum*, even in the absence of a bound ATP molecule, the α-helices remain in the "down" position and the "up" state has not been observed (18). In mitochondria, the two C-terminal α-helices of the orthologous δ-subunit are also permanently "down", and the site where the ATP molecule is bound in *E. coli*, *G. stearothermophilus* and *C. thermarum* and is occupied by the single α-helix of a small protein not found in bacteria and chloroplasts, known confusingly as the ε-subunit (25). In the mycobacterial ε-subunits, the sequences of their C-terminal regions are shorter than in the other species where the structure of the subunit is known, and in the *M. smegmatis* ε-subunit, a C-terminal α-helical hairpin also forms next to the N-terminal domain in the "down" state, but the α-helices are truncated relative to *E. coli*, *G. stearothermophilus* and *C. thermarum*. However, despite the shorter α-helical hairpin in the *M. smegmatis* ε-subunit, the general appearance of the interaction of the truncated C-terminal α-helix with the N-terminal domain of the protein is conserved (Fig. 4B), as are the number of interactions (eight in each case), and also their approximate positions in the N- and C-terminal domains, although none of the side-chains of these residues is conserved significantly in *M. smegmatis* (*SI Appendix*, Fig. S13). Furthermore, there is no evidence of an ATP molecule bound to the subunit. This was anticipated as Arg-94, one of the key residues involved in binding the nucleotide in *E. coli*, *G.*

stearothermophilus and *C. thermarum* is substituted in the equivalent site by an alanine residue in *M. smegmatis* (*SI Appendix*, Fig. S13), and superposition of the *M. smegmatis* and *E. coli* F_1-ATPases demonstrated that the former ε-subunit cannot assume the "up" position in the structure of the inhibited enzyme as its C-terminal α-helix would clash with the DELSEED region in the C-terminal domain of the β_{DP}-subunit (*SI Appendix*, Fig. S14). Moreover, there is no evidence for the presence of the "up"-state in the current electron density map, and in an NMR structure of the isolated ε-subunit from *M. smegmatis* the protein is in the "down" position (14). This structure resembles the ε-subunit described here, but as the deposited co-ordinates have not been released, a precise comparison with the current structure was not possible. Thus, there is currently no structural evidence that the ε-subunit plays a role in the regulation of the hydrolytic activity of the *M. smegmatis* ATP synthase complex.

Regulation of ATP hydrolysis. It is becoming clear that a variety of mechanisms operates to regulate the ATP hydrolytic activity of ATP synthases in eubacteria, mitochondria and chloroplasts. In α-proteobacteria, exemplified by *P. denitrificans*, ATP hydrolysis appears to be inhibited by a protein called the ζ-subunit (23, 36), where the N-terminal inhibitory region binds to a catalytic interface under hydrolytic conditions in a closely related fashion to the inhibitory action of the orthologous mitochondrial inhibitory protein IF_1 on the mitochondrial ATP synthase (37–39). In the case of mammalian IF_1, inhibition of ATP hydrolysis of the protein is activated by a fall in the pH (40) such as would occur in the mitochondrial matrix accompanying an increased reliance by cells on provision of ATP by glycolysis. In the chloroplasts of green plants and algae, during the hours of darkness when the proton motive force is low, ADP-Mg remains bound to one of the three catalytic sites of the enzyme forming an inactive ADP inhibited state of the enzyme (41, 42). This

inhibited state is stabilized by the formation of an intermolecular disulfide bond in the γ-subunit of the enzyme. The formation of this disulfide is proposed to stabilize a β-hairpin structure formed by a unique additional sequence in the γ-subunit (residues 198-233 in *SI Appendix*, Fig. S11; see *SI Appendix*, Fig. S15) that wedges between the β-subunit and the central stalk, thereby blocking the rotation of the γ-subunit and preventing futile ATP hydrolysis. (43). With daylight and a rising proton motive force, the synthetic activity of the enzyme is restored by reduction of the disulfide bond by thioredoxin. The γ-subunits in cyanobacterial ATP synthases, also contain a related insertion (44), but it lacks the nine residue sequence containing the two cysteine residues (*SI Appendix*, Fig. S11), and, although the residual additional loop appears to inhibit ATP hydrolysis, it is not regulated by the redox mechanism found in chloroplasts (45).

The ATPases in the aerobic bacterium *G. stearothermophilus* (35) and in the facultative anaerobe *E. coli* (22, 32, 46) appear to be regulated by their ε-subunit. For *G. stearothermophilus*, it has been proposed that when the proton-motive force and ATP concentration are low, this ATP molecule is released from the ε-subunit, allowing its two C-terminal α-helices to assume the "up" inhibitory position where they penetrate into a catalytic site alongside the rotary γ-subunit, and impede the turning of the rotor (33, 47, 48). However, there is no evidence for the operation of a similar inhibitory mechanism in the thermoalkaliphile, *C. thermarum*, where it appears that ATP hydrolysis is prevented either by the failure to release the products of ATP hydrolysis from one catalytic site, or less likely, for those products to be released and re-bound (18).

A characteristic feature of bacterial F-ATPases with latent hydrolytic activity is that ATP hydrolysis can be activated artificially *in vitro*. For example, LDAO activates the hydrolytic activity of F_1-ATPase from *C. thermarum* 30-fold and maximum activation was

achieved by removal of the C-terminal domain of the ε-subunit (49). The hydrolytic activity of this enzyme is not activated by proteolysis and its ε-subunit is resistant to such treatment (49). In contrast, the F_1-ATPase from *M. smegmatis* is activated by trypsinolysis, and its activity is stimulated further by the addition of LDAO. However, it is not activated by LDAO before trypsinolysis has taken place (*SI Appendix*, Fig. S4). In other bacterial species, the activation of hydrolytic activity by LDAO activation has been attributed to either release of an ADP molecule from a catalytic site (50, 51) or perturbation of the interaction between the □-subunit and $\alpha_3\beta_3$-domain (52). However, the molecular basis of the activation of the hydrolytic activity of F_1-ATPases, by trypsinolysis and/ or LDAO, including the enzyme from *M. smegmatis*, remains unclear.

The current structure of the F_1-ATPase from *M. smegmatis*, albeit at modest resolution, is very similar to that of the inhibited complex from *C. thermarum* in terms of the protein structure itself (apart from the C-terminal extension of α-subunits, which could also be involved in regulation of ATP hydrolysis). Especially, the rotational state of the γ-subunit suggests that the ATP hydrolytic activities of the two enzymes have been arrested at the same point in the rotary cycle. In *C. thermarum*, phosphate and ADP (at 50-70% occupancy) without a magnesium ion are bound to the site (18), whereas the occupancy of the β_E-subunit in the *M. smegmatis* enzyme is likely to be similar. The order of release of the products of ATP hydrolysis by F-ATPases has not been established firmly, although it appears that the magnesium ion leaves first as the catalytic site opens (18, 26, 53, 54). The data about whether the subsequent release of ADP precedes that of phosphate are conflicting (26, 53–58), although in other NTPases, phosphate leaves first (59–61). The current structure of the F_1-ATPase from *M. smegmatis* can be interpreted as being consistent with such an order.

Mycobacteria are obligate aerobes with an extraordinary ability to survive for prolonged periods of hypoxia. A key element of their survival is their ability to keep their respiratory chain energized and thereby to maintain their energy requirements by continuing to make ATP (62, 63). The membrane potential used by mycobacteria to drive ATP synthesis under hypoxia is low (-65 to -75 mV) (62, 63) and they are faced with the thermodynamic challenge of inhibiting ATPase activity, whilst at the same time remaining competent for ATP synthesis. If the ATP synthase were freely reversible, the cells would become depleted of ATP rapidly in order to re-establish the membrane potential, and they would die. Thus, the extreme latency of the enzyme in the direction of ATP hydrolysis is a characteristic feature of ATP synthases from fast and slow growing mycobacteria (15), and the mechanism of ATP inhibition is an intrinsic feature of the F_1-domain. The structure of the mycobacterial F_1-domain reported here is a big step towards uncovering the molecular basis of this inhibitory mechanism, and it provides a framework for the structure-based design of small-molecule that might activate ATP hydrolysis or inhibit ATP synthesis specifically in the pathogen.

Materials and Methods

The F_1-ATPase from *M. smegmatis* (subunits α, β, γ and ε) was over-expressed from a plasmid in *M. smegmatis* strain mc2 4517, purified by nickel affinity chromatography via a His_6-tag attached to the β-subunit, size exclusion chromatography, crystallized by vapor diffusion and its structure solved from X-ray diffraction data by molecular replacement with the F_1-ATPase from *C. thermarum* (PDB5hkk). Images of structures and electron density maps were prepared with PyMOL (64). For further details, see Supplementary Information.

Footnotes

* Conradie F *et al.* (2017) The NIX-TB trial of pretomanid, bedaquiline and linezoid to treat XDR-TB. *Conference on RetroViruses and Opportunistic Infections 2017*

Acknowledgements

This work was supported by the Medical Research Council, U. K. by grants MC_U105663150 and MR/M009858/1 to J. E. W., and MC_U105184325 to A. G. W. L.; and by the European Drug Initiative on Channels and Transporters via contract HEALTH-F4-2007-201924 (EDICT) to J. E. W. G. M. C. was supported by a James Cook Fellowship from The Royal Society of New Zealand. We thank the beamline staff at the Swiss Light Source, the European Synchrotron Radiation Facility and at beamline I04, the Diamond Light Source for their help, and Drs I. M. Fearnley and S. Ding for recording mass spectra.

REFERENCES

1. World Health Organization (2018) Global tuberculosis report 2018. http://www.who.int/tb/publications/global_report/en/
2. Houben RM, Dodd PJ (2016) The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. *PLoS Med* 13(10):e1002152.
3. Behr MA, Edelstein PH, Ramakrishnan L (2018) Revisiting the timetable of tuberculosis. *BMJ* 362:k2738.
4. Diacon AH *et al.* (2012) Randomized pilot trial of eight weeks of bedaquiline (TMC207) treatment for multidrug-resistant tuberculosis: long-term outcome, tolerability, and effect on emergence of drug resistance. *Antimicrob Agents Chemother* 56(6):3271–3276.
5. Diacon AH *et al.* (2014) Multidrug-resistant tuberculosis and culture conversion with bedaquiline. *N Engl J Med* 371(8):723–732.

6. Manjunatha U, Boshoff HI, Barry CE (2009) The mechanism of action of PA-824: Novel insights from transcriptional profiling. *Commun Integr Biol* 2(3):215–218.
7. Tasneen R *et al.* (2015) Contribution of the nitroimidazoles PA-824 and TBA-354 to the activity of novel regimens in murine models of tuberculosis. *Antimicrob Agents Chemother* 59(1):129–135.
8. Tasneen R *et al.* (2016) Contribution of Oxazolidinones to the Efficacy of Novel Regimens Containing Bedaquiline and Pretomanid in a Mouse Model of Tuberculosis. *Antimicrob Agents Chemother* 60(1):270–277.
9. World Health Organization (2018) Rapid communication: key changes to treatment of multidrug- and rifampicin-resistant tuberculosis (MDR/RR-TB). http://www.who.int/tb/publications/2018/rapid_communications_MDR/en/
10. Andries K *et al.* (2005) A diarylquinoline drug active on the ATP synthase of Mycobacterium tuberculosis. *Science* 307(5707):223–227.
11. Koul A *et al.* (2014) Delayed bactericidal response of Mycobacterium tuberculosis to bedaquiline involves remodelling of bacterial metabolism. *Nat Commun* 5:3369.
12. Preiss L *et al.* (2015) Structure of the mycobacterial ATP synthase Fo rotor ring in complex with the anti-TB drug bedaquiline. *Sci Adv* 1(4):e1500106.
13. Kundu S, Biukovic G, Grüber G, Dick T (2016) Bedaquiline Targets the ε Subunit of Mycobacterial F-ATP Synthase. *Antimicrob Agents Chemother* 60(11):6977–6979.
14. Joon S *et al.* (2018) The NMR solution structure of Mycobacterium tuberculosis F-ATP synthase subunit ε provides new insight into energy coupling inside the rotary engine. *FEBS J*

15. Haagsma AC, Driessen NN, Hahn MM, Lill H, Bald D (2010) ATP synthase in slow- and fast-growing mycobacteria is active in ATP synthesis and blocked in ATP hydrolysis direction. *FEMS Microbiol Lett* 313(1):68–74.

16. Cook GM *et al.* (2003) Purification and biochemical characterization of the F_1F_o-ATP synthase from thermoalkaliphilic *Bacillus sp*. strain TA2.A1. *J Bacteriol* 185(15):4442–4449.

17. Gavilanes-Ruiz M, Tommasino M, Capaldi RA (1988) Structure-function relationships of the Escherichia coli ATP synthase probed by trypsin digestion. *Biochemistry* 27(2):603–609.

18. 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.

19. Ragunathan P *et al.* (2017) The uniqueness of subunit α of mycobacterial F-ATP synthases: An evolutionary variant for niche adaptation. *J Biol Chem* 292(27):11262–11279.

20. 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.

21. Frishman D, Argos P (1996) Incorporation of non-local interactions in protein secondary structure prediction from the amino acid sequence. *Protein Eng* 9(2):133–142.

22. 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.

23. 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.

24. Hahn A, Vonck J, Mills DJ, Meier T, Kühlbrandt W (2018) Structure, mechanism, and regulation of the chloroplast ATP synthase. *Science* 360(6389)

25. 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.

26. 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.

27. Priya R *et al.* (2013) Solution structure of subunit γ (γ(1-204)) of the Mycobacterium tuberculosis F-ATP synthase and the unique loop of γ(165-178), representing a novel TB drug target. *J Bioenerg Biomembr* 45(1-2):121–129.

28. Walker JE, Runswick MJ, Saraste M (1982) Subunit equivalence in *Escherichia coli* and bovine heart mitochondrial F_1F_o-ATPases. *FEBS Lett* 146(2):393–396.

29. Uhlin U, Cox GB, Guss JM (1997) Crystal structure of the ε subunit of the proton-translocating ATP synthase from *Escherichia coli*. *Structure* 5(9):1219–1230.

30. Wilkens S, Dahlquist FW, McIntosh LP, Donaldson LW, Capaldi RA (1995) Structural features of the ε subunit of the *Escherichia coli* ATP synthase determined by NMR spectroscopy. *Nat Struct Biol* 2(11):961–967.

31. Wilkens S, Capaldi RA (1998) Solution structure of the ε subunit of the F_1-ATPase from *Escherichia coli* and interactions of this subunit with β subunits in the complex. *J Biol Chem* 273(41):26645–26651.

32. Sobti M *et al.* (2016) Cryo-EM structures of the autoinhibited E. coli ATP synthase in three rotational states. *Elife* 5

33. Yagi H *et al.* (2007) Structures of the thermophilic F_1-ATPase ε subunit suggesting ATP-regulated arm motion of its C-terminal domain in F_1. *Proc Natl Acad Sci U S A* 104(27):11233–11238.

34. Kato S, Yoshida M, Kato-Yamada Y (2007) Role of the ε subunit of thermophilic F_1-ATPase as a sensor for ATP. *J Biol Chem* 282(52):37618–37623.

35. 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.

36. Morales-Ríos E *et al.* (2010) A novel 11-kDa inhibitory subunit in the F_1F_o ATP synthase of *Paracoccus denitrificans* and related α-proteobacteria. *FASEB J* 24(2):599–608.

37. 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.

38. 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.

39. 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.

40. Cabezon E, Butler PJ, Runswick MJ, Walker JE (2000) Modulation of the oligomerization state of the bovine F_1-ATPase inhibitor protein, IF_1, by pH. *J Biol Chem* 275(33):25460–25464.

41. Nalin CM, McCarty RE (1984) Role of a disulfide bond in the γ subunit in activation of the ATPase of chloroplast coupling factor 1. *J Biol Chem* 259(11):7275–7280.

42. Ketcham SR, Davenport JW, Warncke K, McCarty RE (1984) Role of the gamma subunit of chloroplast coupling factor 1 in the light-dependent activation of photophosphorylation and ATPase activity by dithiothreitol. *J Biol Chem* 259(11):7286–7293.

43. Murakami S *et al.* (2018) Structure of the γ-ε complex of cyanobacterial F_1-ATPase reveals a suppression mechanism of the γ subunit on ATP-hydrolysis in phototrophs. *Biochem J*

44. Cozens AL, Walker JE (1987) The organization and sequence of the genes for ATP synthase subunits in the cyanobacterium *Synechococcus* 6301: Support for an endosymbiotic origin of chloroplasts. *J Mol Biol* 194(3):359–383.

45. Werner S, Schumann J, Strotmann H (1990) The primary structure of the gamma-subunit of the ATPase from Synechocystis 6803. *FEBS Lett* 261(1):204–208.

46. Shah NB, Hutcheon ML, Haarer BK, Duncan TM (2013) F_1-ATPase of *Escherichia coli*: the ε- inhibited state forms after ATP hydrolysis, is distinct from the ADP-inhibited state, and responds dynamically to catalytic site ligands. *J Biol Chem* 288(13):9383–9395.

47. Feniouk BA, Suzuki T, Yoshida M (2007) Regulatory interplay between proton motive force, ADP, phosphate, and subunit epsilon in bacterial ATP synthase. *J Biol Chem* 282(1):764–772.

48. Suzuki T *et al.* (2003) F_oF_1-ATPase/synthase is geared to the synthesis mode by conformational rearrangement of ε subunit in response to proton motive force and ADP/ATP balance. *J Biol Chem* 278(47):46840–46846.

49. Keis S, Stocker A, Dimroth P, Cook GM (2006) Inhibition of ATP hydrolysis by thermoalkaliphilic F_1F_o-ATP synthase is controlled by the C terminus of the epsilon subunit. *J Bacteriol* 188(11):3796–3804.

50. Paik SR, Jault JM, Allison WS (1994) Inhibition and inactivation of the F1 adenosinetriphosphatase from Bacillus PS3 by dequalinium and activation of the enzyme by lauryl dimethylamine oxide. *Biochemistry* 33(1):126–133.

51. Jault JM *et al.* (1995) The $\alpha_3\beta_3\gamma$ complex of the F_1-ATPase from thermophilic *Bacillus* PS3 containing the $\alpha D_{261}N$ substitution fails to dissociate inhibitory MgADP from a catalytic site when ATP binds to noncatalytic sites. *Biochemistry* 34(50):16412–16418.

52. Dunn SD, Tozer RG, Zadorozny VD (1990) Activation of *Escherichia coli* F_1-ATPase by lauryldimethylamine oxide and ethylene glycol: relationship of ATPase activity to the interaction of the ε and β subunits. *Biochemistry* 29(18):4335–4340.

53. 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.

54. 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.

55. Watanabe R, Iino R, Noji H (2010) Phosphate release in F_1-ATPase catalytic cycle follows ADP release. *Nat Chem Biol* 6(11):814–820.

56. Okazaki K, Hummer G (2013) Phosphate release coupled to rotary motion of F_1-ATPase. *Proc Natl Acad Sci U S A* 110(41):16468–16473.

57. Watanabe R, Noji H (2014) Timing of inorganic phosphate release modulates the catalytic activity of ATP-driven rotary motor protein. *Nat Commun* 5:3486.

58. Nam K, Pu J, Karplus M (2014) Trapping the ATP binding state leads to a detailed understanding of the F_1-ATPase mechanism. *Proc Natl Acad Sci U S A* 111(50):17851–17856.

59. Gulick AM, Rayment I (1997) Structural studies on myosin II: communication between distant protein domains. *BioEssays* 19(7):561–569.

60. Wittinghofer A, Vetter IR (2011) Structure-function relationships of the G domain, a canonical switch motif. *Annu Rev Biochem* 80:943–971.

61. Milic B, Andreasson JO, Hancock WO, Block SM (2014) Kinesin processivity is gated by phosphate release. *Proc Natl Acad Sci U S A* 111(39):14136–14140.

62. Rao SP, Alonso S, Rand L, Dick T, Pethe K (2008) The protonmotive force is required for maintaining ATP homeostasis and viability of hypoxic, nonreplicating Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* 105(33):11945–11950.

63. Berney M, Cook GM (2010) Unique flexibility in energy metabolism allows mycobacteria to combat starvation and hypoxia. *PLoS ONE* 5(1):e8614.

64. Schrodinger, LLC (2018) The PyMOL Molecular Graphics System, Version 2.2.2.

FIGURES

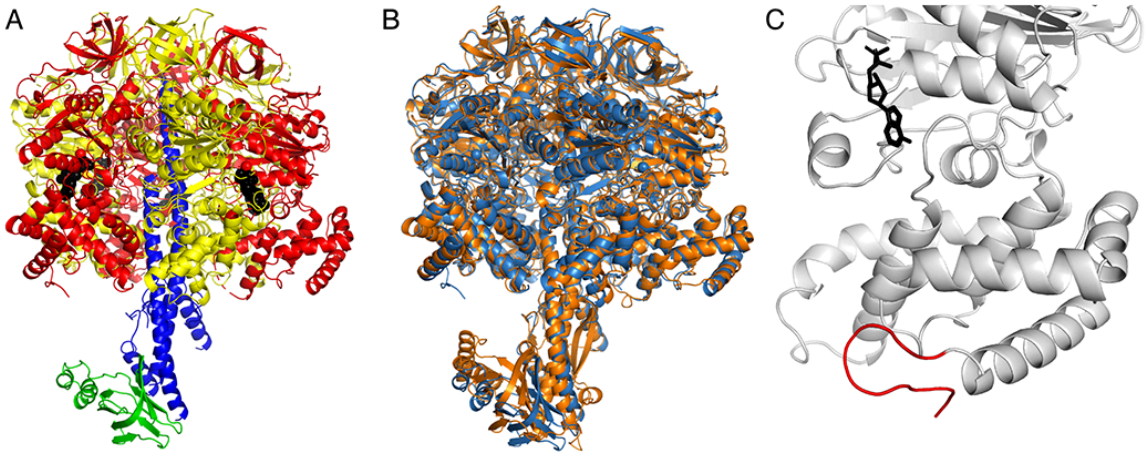


Fig. 1. The structure of the F_1-ATPase from *M. smegmatis*. Part (A), side view in ribbon representation with the α-, β-, γ- and ε-subunits in red, yellow, blue and green, respectively, and bound nucleotides in black. Part (B), comparison of the F_1-ATPase complexes from *M. smegmatis* (sky blue) and *C. thermarum* (18) (orange). The structures were superimposed via their $\alpha_3\beta_3$-domains. Part (C), the C-terminal extension in the α_E-subunit of F_1-ATPase from *M. smegmatis*. Density extending from the C-terminal helix (residues 494-511) was modelled from residues 1512-1522 (red), but the register is uncertain and the following 27 residues are unresolved. An ADP molecule (black) is shown bound to the nucleotide binding site.

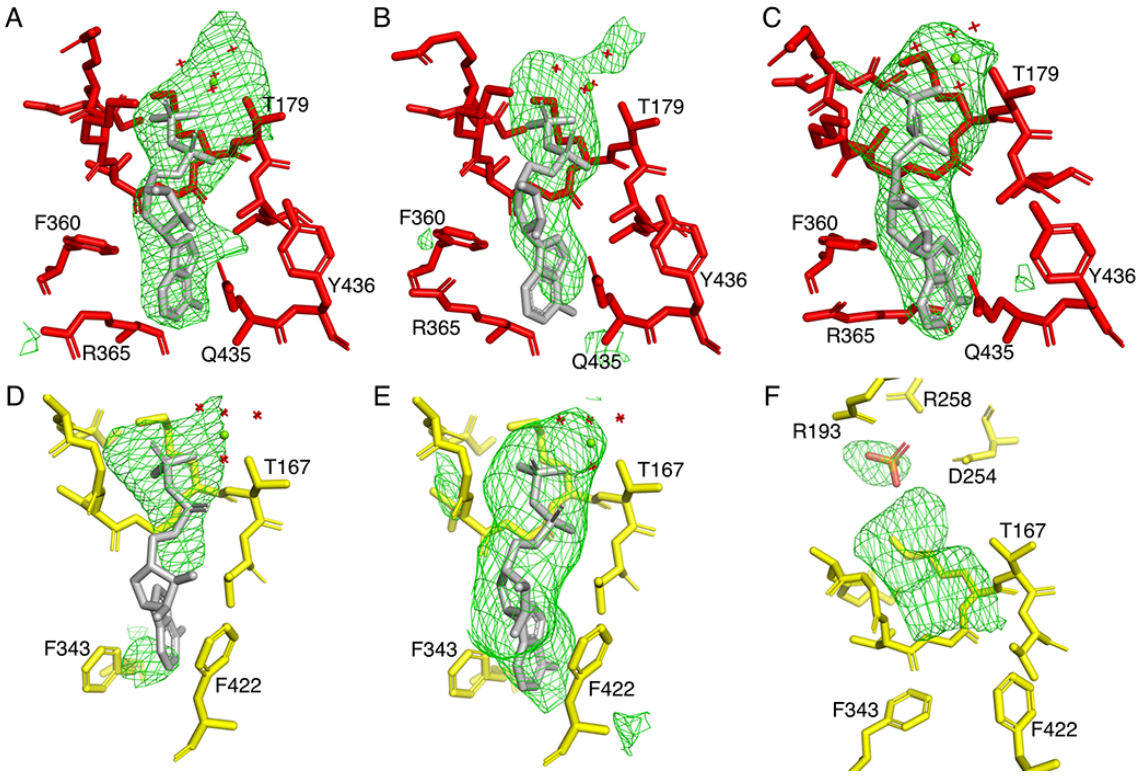


Fig. 2. Occupancy of nucleotide binding sites in the α- and β-subunits of the F_1-ATPase from *M. smegmatis*. An Fo-Fc difference density map was calculated for the complex with the nucleotides, phosphate, Mg^{2+} and water molecules at zero occupancy. The difference density is shown as green mesh contoured to 2.5 σ. Parts (A-C), the α_{DP}-, α_{TP}- and α_E-subunits; parts (D-F), the β_{DP}-, β_{TP}- and β_E-subunits. In parts (A-E), the sites are occupied by ADP and an accompanying magnesium ion (black sphere) with four water ligands (red crosses); the fifth and sixth ligands are provided by O2B of the ADP and the hydroxyl of either αThr-179 or βThr-167. In part (F), the upper region of the catalytic site is occupied by a phosphate (or sulfate) ion (orange and red). Although the electron density beneath it in the vicinity of the P-loop, cannot be interpreted with confidence, it probably can be accounted for by an ADP molecule (without a magnesium ion) at partial occupancy.

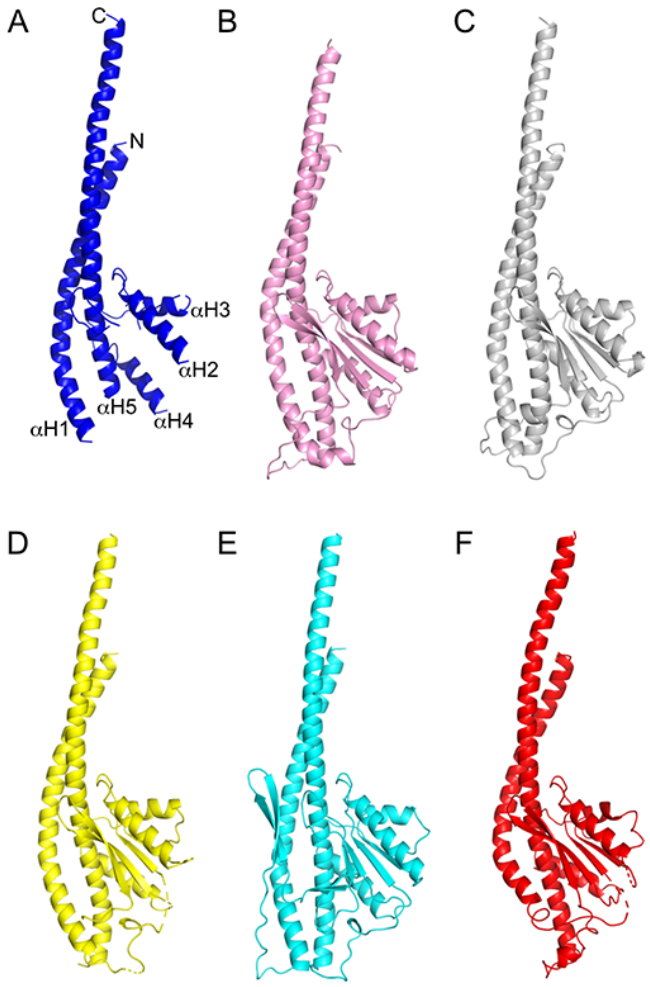


Fig. 3. Comparison of the structure of the γ-subunits of the F-ATPases from *M. smegmatis* with those of orthologues. Part (A), *M. smegmatis* with the five α-helices numbered 1-5 from N- to C-terminus; part (B), *E. coli* (22); part (C), *C. thermarum* (18); part (D), *P. denitrificans* (23); part (E), chloroplast from spinach (24); part (F), bovine mitochondria (25).

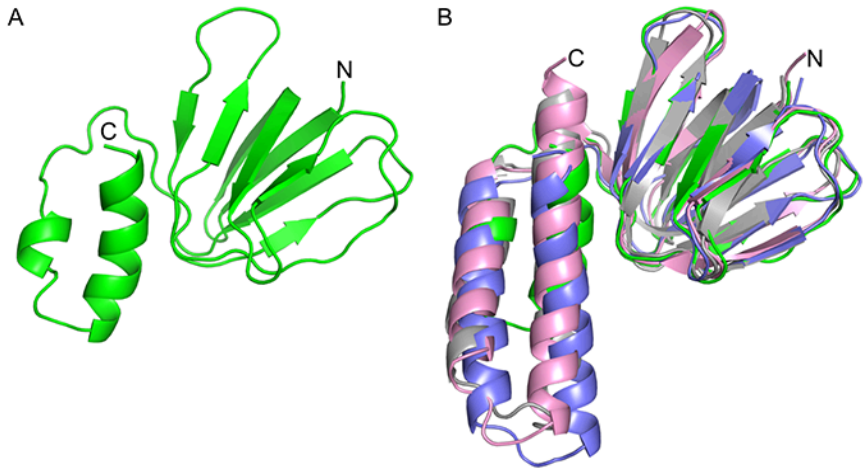


Fig. 4. The structure of the ε-subunit of the F_1-ATPase from *M. smegmatis* compared with orthologs. Part (A), the mycobacterial ε-subunit showing the N-terminal β-sandwich and the C-terminal α-helical domains; part (B), superimposition of the ε-subunits from *M. smegmatis* (green), *C. thermarum* (18) (grey), *E. coli* (29) (pink) and the bovine δ-subunit (25) (slate blue). The N-terminal domains are very similar, but both C-terminal helices of the *M. smegmatis* protein are shorter than in the other examples.

Supplementary Appendix for:

Structure of the catalytic domain of the ATP synthase from *Mycobacterium smegmatis*: a surrogate target for development of anti-tubercular drugs

Alice Tianbu Zhang[a1], Martin G. Montgomery[a1], Andrew G. W. Leslie[b], Gregory M. Cook[a,c], 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]The Medical Research Council Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Ave, Cambridge, CB2 0QH United Kingdom; [c]Department of Microbiology and Immunology, University of Otago, Dunedin 9016, New Zealand.

Materials and Methods

General Procedures. Protein concentrations were measured with bicinchoninic acid (Thermo Fisher Scientific). Proteins were analyzed by SDS-PAGE in 12-22% Tris-glycine gels and by blue native polyacrylamide gel electrophoresis (BN-PAGE) in 3-12% Bis-tris native gels (Life Technologies). They were detected by staining with Coomassie-Blue dye or silver.

Overexpression of the F_1-ATPase from *M. smegmatis*. The genes encoding the α-, β-, γ- and ε-subunits of the F_1-ATPase were amplified by PCR with DNA from *M. smegmatis* strain mc2 155 as template. A hexahistidine tag and an intervening site for cleavage with the protease from tobacco etch virus were introduced at the N-terminus of the β-subunit. The genes were cloned into the *E. coli*-mycobacterium shuttle vector pYUB1049. The resulting expression shuttle plasmid with a T7 promoter was transformed into cells of *M. smegmatis* expression strain mc2 4517 encoding T7 polymerase (1). The transformed cells were cultured at 37°C in 2xTY medium supplemented with hygromycin B (100 μg/ml), kanamycin (50 μg/ml) and Tween-80 (0.05%, v/v). When the OD_{600} of the culture had reached ca. 1.0, expression of the F_1-ATPase was induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside. The cells were grown for 24 h at 30°C, and then harvested.

Purification of the Enzyme. The *M. smegmatis* cells (30 g) were resuspended in buffer (100 ml) consisting of 50 mM Tris-HCl, pH 7.5, 150 mM sodium chloride, 20% (v/v) glycerol, 20 mM imidazole, 5 mM magnesium sulfate, 1 mM tris-(2-carboxyethyl)-phosphine, 1 mM ADP and an EDTA-free protease inhibitor cocktail (Roche). They were sonicated, passed at 4°C through a cell disruptor (Constant Systems Limited) at 30,000 psi. and centrifuged (151,000xg, 45 min, 4°C). All subsequent steps were performed at room temperature. The supernatant was applied to a nickel Sepharose HisTrap HP column (5 ml; GE Healthcare Life Sciences), which was washed successively with 40 mM and 250 mM imidazole (Fig. S2). Fractions containing the F_1-ATPase were pooled and TEV protease (2) was added at a protein:protease ratio of 50:1 (w:w). This solution was dialyzed for 1 h at room temperature in a Spectra/Por dialysis membrane (molecular weight cut off 3500 Da) against 1 l of buffer consisting of 50 mM Tris-HCl, pH 7.5, 20% (v/v) glycerol, 5 mM magnesium sulfate, 1 mM tris-(2-carboxyethyl)-phosphine, 1 mM ADP and an EDTA-free

protease inhibitor cocktail. The dialyzed sample was applied to two nickel and two Q-Sepharose HiTrap HP columns (1 ml each; GE Healthcare Life Sciences) connected in series. Once the bound proteins had flowed through, the nickel columns were removed, and the Q-Sepharose columns were eluted with a linear gradient of sodium chloride from 0 to 1 M (Fig. S2). Fractions containing F_1-ATPase were pooled, concentrated and passed through a Superdex 200 10/300 column (GE Healthcare Life Sciences) equilibrated with buffer consisting of 20 mM Tris-HCl, pH 7.5, 10% (v/v) glycerol, 2 mM magnesium sulfate, 1 mM ADP and an EDTA-free protease inhibitor cocktail (Fig. S2). The yield of pure enzyme was 15 mg.

Enzyme Assay. The ATP hydrolytic activity of the *M. smegmatis* F_1-ATPase was measured at 37ºC in the presence of an ATP regenerating system (3). The latent enzyme was activated by digestion for 10 min at 37°C with bovine trypsin (1 mg/ml) at a trypsin:F_1-ATPase ratio of 1:50 (w/w). Digestion was terminated with a 2-fold molar excess of bovine pancreatic trypsin inhibitor (5 mg/ml). Further activation was achieved by the addition of LDAO to a concentration of 0.4 %.

Mass Spectrometry. Proteins in stained bands from SDS-PAGE gels were identified by mass mapping of tryptic peptides by matrix assisted laser desorption ionization-time of flight mass spectrometry. The purified F_1-ATPase and the products of 10 min of trypsinolysis were fractionated by reverse phase HPLC, passed directly into the electrospray ionization inlet of a Quattro Ultima quadrupole mass spectrometer (Waters-Micromass), and the masses of the intact subunits and their fragments were measured (Table S2).

Crystallization of the Enzyme. Purified F_1-ATPase was concentrated with a VivaSpin concentrator (molecular weight cut off 100 kDa) and centrifuged (168,000xg, 30 min, 21°C). Crystals were grown by vapor diffusion in 400 nl or 4 µl sitting drops in 96- or 48-well plates (Swissci) at a protein concentration of 20-30 mg/ml in buffer containing 300 mM magnesium formate and 100 mM Tris-HCl, pH 7.5, and 23% (w/v) polyethylene glycol (PEG) 4000 (the volume ratio of the protein solution and precipitant was 1:1) After 2-3 weeks, the largest hexagonal crystals had dimensions of ca. 250 x 200 x 20 µm (Fig. S5). They were cryo-protected with 20% PEG 400, harvested with 60° angled MicroLoops (MiTeGen, U. S. A.) and vitrified in liquid nitrogen.

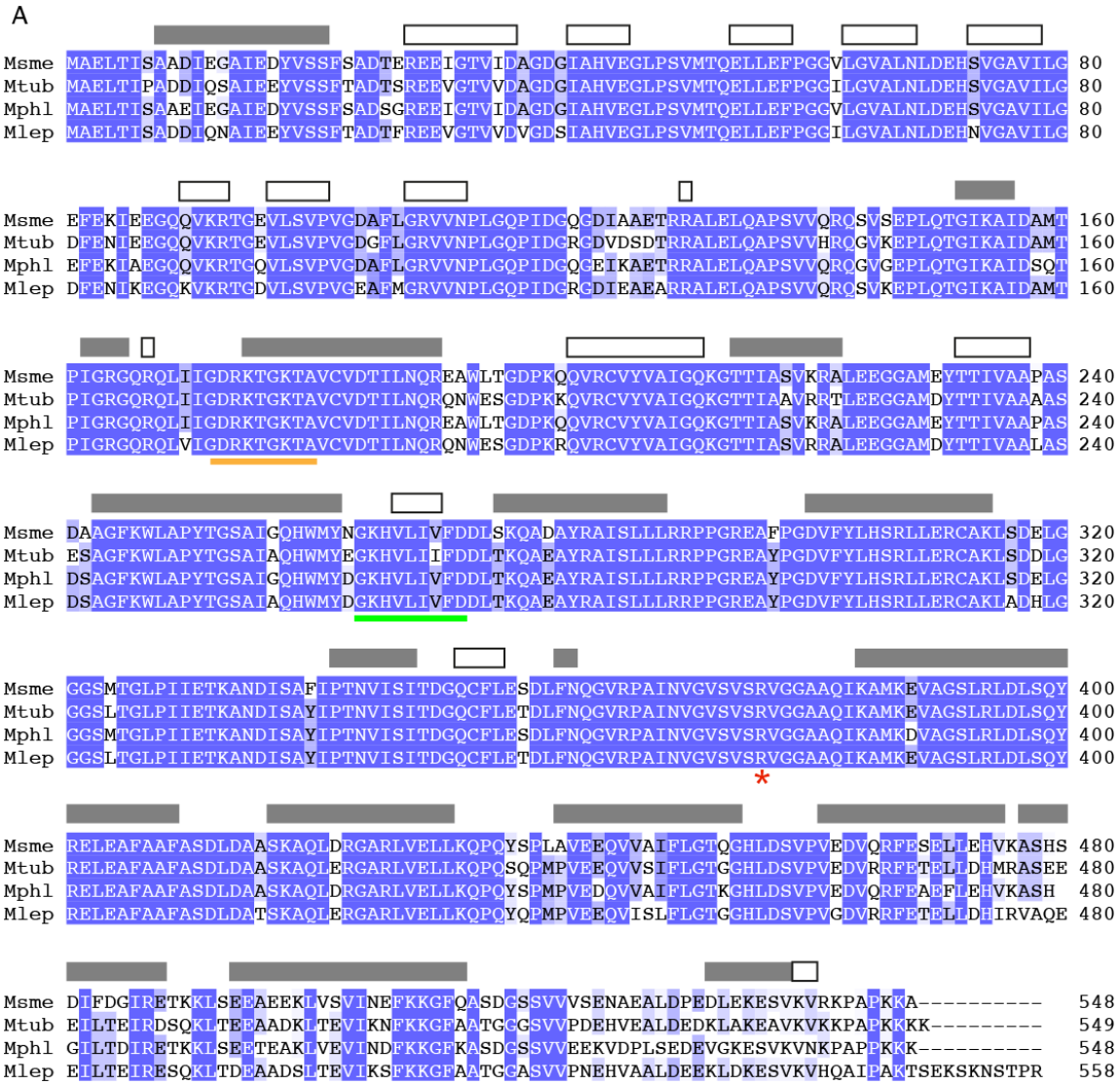
Data Collection. Initial X-ray diffraction data were collected at the Swiss Light Source (Villigen, Switzerland) and the European Synchrotron Radiation Facility (Grenoble, France). The definitive data were collected at 100K at 1.0 Å wavelength on beamline I04 with a Pilatus 6M-F detector (Dectris, Switzerland) at The Diamond Light Source (Didcot, U. K.). In order to obtain maximum separation of reflections, vitrified crystals were mounted on a mini-κ-goniometer, orientated with the longest axis of the unit cell nearly parallel to the data collection axis.

Data processing. Data were processed with programs from the CCP4 suite (4). Diffraction images were integrated with iMOSFLM (5) and the data reduced with AIMLESS (6) and CTRUNCATE (4). Molecular replacement was carried out with PHASER (7). The F_1-ATPase from *C. thermarum* (PDB5hkk) (8), but without nucleotides, phosphate or water

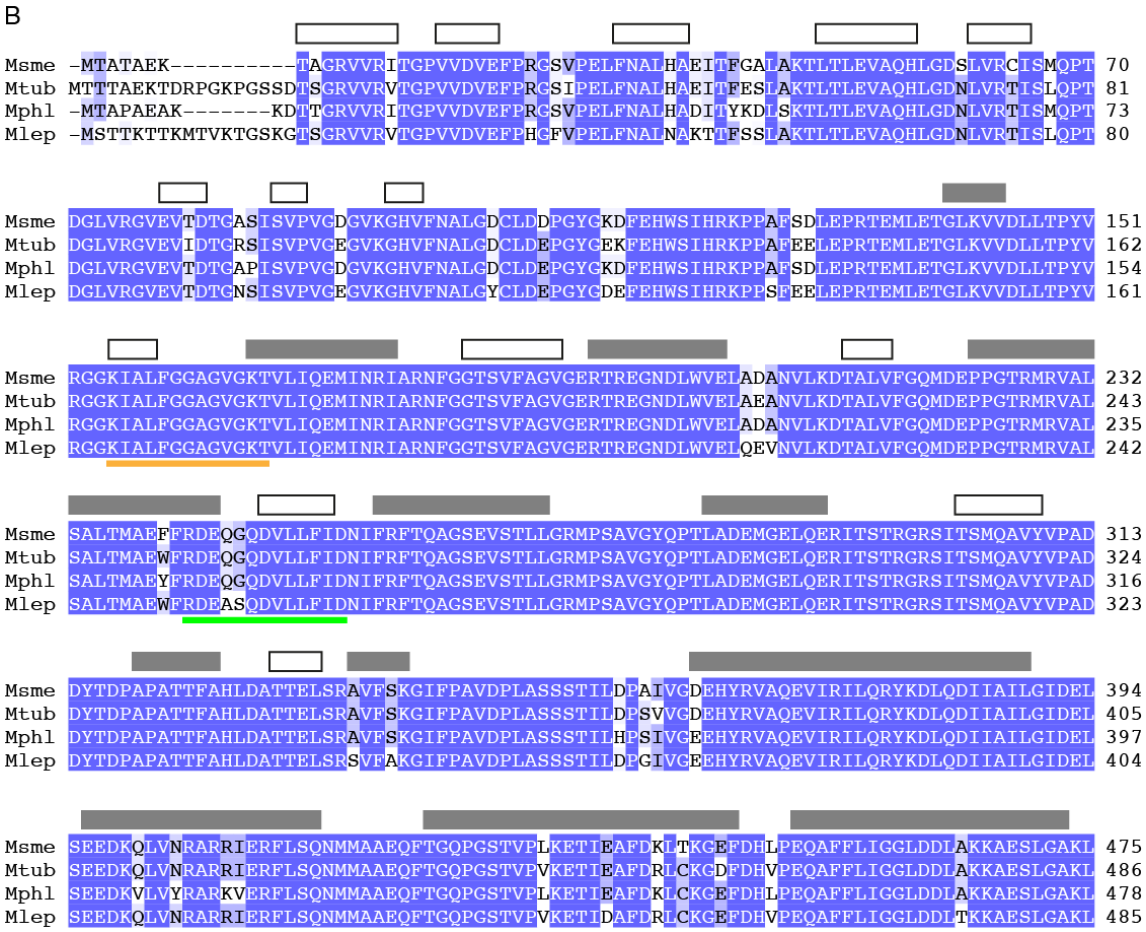
molecules, was used as an initial model for molecular replacement. The model was refined in REFMAC5 (9) using additional restraints provided by PROSMART (10) derived from the F_1-ATPase from *C. thermarum* (8). Alternate rounds of refinement with REFMAC5 (9) and manual rebuilding with COOT (11) were carried out. Following each round of refinement, the stereochemistry of the model was assessed with COOT and MolProbity (12). Electron density maps presented in figures were made by FFT (fast Fourier transform) in the CCP4 package (4). Grid sampling was doubled for the Fo-Fc map (Fig. 2) and tripled for the 2Fo-Fc map (Fig. S6) to produce a tighter mesh for improved image quality in PyMol (13). Images of structures and electron density maps were prepared with PyMOL (13).

Protein analyses. Sequences of proteins were aligned with CLUSTAL O via the UniProt website (14), and the secondary structures of proteins and intrinsic disorder were predicted with PSIPred (15) and PrDOS (16). RMSDs were calculated in PyMol (13) comparing Cα only, with 0 cycles of refinement. The total atom pairs for each subunit are listed in Table S5.

Rotation of the γ-Subunit. Residues 24-34 of the *M. smegmatis* γ-subunit (and equivalent regions in other F_1-ATPases) interact with the C-terminal domains of the α- and β-subunits, and this segment acts a rigid body uninfluenced by contacts in the crystal lattice between adjacent F_1-ATPase complexes. In contrast, residues 33-226 of the γ-subunit, and the associated δ- and ε-subunits lie outside the $\alpha_3\beta_3$-domain, where their positions may be subject to such influences. Therefore, the rotations of residues 24-34 of the γ-subunit in the various aligned structures were measured relative to the position of the same segment (residues 22-32) in the ground state structure of azide-free bovine F_1-ATPase (17). These measurements were made by aligning the structures via the crown domains at the N-termini of α- and β-subunits, and then calculating the centre of mass of residues 22-33 of the γ-subunit and determining the rotation angle required to match its position with that of the equivalent segment in the bovine azide-free ground state structure, which was taken as the reference point, set as 0° (18).


A
Msme MAELTISAADIEGAIEDYVSSFSADTEREEIGTVIDAGDGIAHVEGLPSVMTQELLEFPGGVLGVALNLDEHSVGAVILG 80
Mtub MAELTIPADDIQSAIEEYVSSFTADTSREEVGTVVDAGDGIAHVEGLPSVMTQELLEFPGGILGVALNLDEHSVGAVILG 80
Mphl MAELTISAAEIEGAIEDYVSSFSADSGREEIGTVIDAGDGIAHVEGLPSVMTQELLEFPGGVLGVALNLDEHSVGAVILG 80
Mlep MAELTISADDIQNAIEEYVSSFTADTFREEVGTVVDVGDSIAHVEGLPSVMTQELLEFPGGILGVALNLDEHNVGAVILG 80
Msme EFEKIEEGQQVKRTCEVLSVPVGDAFLGRVVNPLGQPIDGQCDIAAETRRALELQAPSVVQRQSVSEPLQTGIKAIDAMT 160
Mtub DFENIEEGQQVKRTCEVLSVPVGDGFLGRVVNPLGQPIDGRCDVDSDTRRALELQAPSVVHRQGVKEPLQTGIKAIDAMT 160
Mphl EFEKIAEGQQVKRTCQVLSVPVGDAFLGRVVNPLGQPIDGQCEIKAETRRALELQAPSVVQRQGVGEPLQTGIKAIDSQT 160
Mlep DFENIKEGQKVKRTCDVLSVPVGEAFMGRVVNPLGQPIDGRCDIEAEARRALELQAPSVVQRQSVKEPLQTGIKAIDAMT 160
Msme PIGRGQRQLIIGDRKTGKTAVCVDTILNQREAWLTGDPKQQVRCVYVAIGQKGTTIASVKRALEEGGAMEYTTIVAAPAS 240
Mtub PIGRGQRQLIIGDRKTGKTAVCVDTILNQRQNWESGDPKKQVRCVYVAIGQKGTTIAAVRRTLEEGGAMDYTTIVAAAAS 240
Mphl PIGRGQRQLIIGDRKTGKTAVCVDTILNQREAWLTGDPKQQVRCVYVAIGQKGTTIASVKRALEEGGAMEYTTIVAAPAS 240
Mlep PIGRGQRQLVIGDRKTGKTAVCVDTILNQRQNWESGDPKRQVRCVYVAIGQKGTTIASVRRALEEGGAMDYTTIVAALAS 240
Msme DAAGFKWLAPYTGSAIGQHWMYNGKHVLIVFDDLSKQADAYRAISLLLRRPPGREAFPGDVFYLHSRLLERCAKLSDELG 320
Mtub ESAGFKWLAPYTGSAIAQHWMYEGKHVLIIFDDLTKQAEAYRAISLLLRRPPGREAYPGDVFYLHSRLLERCAKLSDDLG 320
Mphl DSAGFKWLAPYTGSAIGQHWMYDGKHVLIVFDDLTKQAEAYRAISLLLRRPPGREAYPGDVFYLHSRLLERCAKLSDELG 320
Mlep DSAGFKWLAPYTGSAIAQHWMYDGKHVLIVFDDLTKQAEAYRAISLLLRRPPGREAYPGDVFYLHSRLLERCAKLADHLG 320
Msme GGSMTGLPIIETKANDISAFIPTNVISITDGQCFLESDLFNQGVRPAINVGVSVSRVGGAAQIKAMKEVAGSLRLDLSQY 400
Mtub GGSLTGLPIIETKANDISAYIPTNVISITDGQCFLETDLFNQGVRPAINVGVSVSRVGGAAQIKAMKEVAGSLRLDLSQY 400
Mphl GGSMTGLPIIETKANDISAYIPTNVISITDGQCFLESDLFNQGVRPAINVGVSVSRVGGAAQIKAMKDVAGSLRLDLSQY 400
Mlep GGSLTGLPIIETKANDISAYIPTNVISITDGQCFLETDLFNQGVRPAINVGVSVSRVGGAAQIKAMKEVAGSLRLDLSQY 400
Msme RELEAFAAFASDLDAASKAQLDRGARLVELLKQPQYSPLAVEEQVVAIFLGTQGHLDSVPVEDVQRFESELLEHVKASHS 480
Mtub RELEAFAAFASDLDAASKAQLERGARLVELLKQPQSQPMPVEEQVVSIFLGTGGHLDSVPVEDVRRFETELLDHMRASEE 480
Mphl RELEAFAAFASDLDAASKAQLDRGARLVELLKQPQYSPMPVEDQVVAIFLGTKGHLDSVPVEDVQRFEAEFLEHVKASH 480
Mlep RELEAFAAFASDLDATSKAQLERGARLVELLKQPQYQPMPVEEQVISLFLGTGGHLDSVPVGDVRRFETELLDHIRVAQE 480
Msme DIFDGIRETKKLSEEAEEKLVSVINEFKKGFQASDGSSVVVSENAEALDPEDLEKESVKVRKPAPKKA---------- 548
Mtub EILTEIRDSQKLTEEAADKLTEVIKNFKKGFAATGCGSVVPDEHVEALDEDKLAKEAVKVKKPAPKKKK--------- 549
Mphl GILTDIRETKKLSEETEAKLVEVINDFKKGFKASDGSSVVEEKVDPLSEDEVGKESVKVNKPAPPKKK---------- 548
Mlep EILTEIRESQKLTDEAADSLTEVIKSFKKGFAATGCASVVPNEHVAALDEEKLDKESVKVHQAIPAKTSEKSKNSTPR 558

B


Msme -MTATAEK----------TAGRVVRITGPVVDVEFPRGSVPELFNALHAEITFGALAKTLTLEVAQHLGDSLVRCISMQPT 70
Mtub MTTTAEKTDRPGKPGSSDTSGRVVRVTGPVVDVEFPRGSIPELFNALHAEITFESLAKTLTLEVAQHLGDNLVRTISLQPT 81
Mphl -MTAPAEAK-------KDTTGRVVRITGPVVDVEFPRGSVPELFNALHADITYKDLSKTLTLEVAQHLGDNLVRTISMQPT 73
Mlep -MSTTKTTKMTVKTGSKGTSGRVVRVTGPVVDVEFPHGFVPELFNALNAKTTFSSLAKTLTLEVAQHLGDNLVRTISLQPT 80
Msme DGLVRGVEVTDTGASISVPVGDGVKGHVFNALGDCLDDPGYGKDFEHWSIHRKPPAFSDLEPRTEMLETGLKVVDLLTPYV 151
Mtub DGLVRGVEVIDTGRSISVPVGEGVKGHVFNALGDCLDEPGYGEKFEHWSIHRKPPAFEELEPRTEMLETGLKVVDLLTPYV 162
Mphl DGLVRGVEVTDTGAPISVPVGDGVKGHVFNALGDCLDEPGYGKDFEHWSIHRKPPAFSDLEPRTEMLETGLKVVDLLTPYV 154
Mlep DGLVRGVEVTDTGNSISVPVGEGVKGHVFNALGYCLDEPGYGDEFEHWSIHRKPPSFEELEPRTEMLETGLKVVDLLTPYV 161
Msme RGGKIALFGGAGVGKTVLIQEMINRIARNFGGTSVFAGVGERTREGNDLWVELADANVLKDTALVFGQMDEPPGTRMRVAL 232
Mtub RGGKIALFGGAGVGKTVLIQEMINRIARNFGGTSVFAGVGERTREGNDLWVELAEANVLKDTALVFGQMDEPPGTRMRVAL 243
Mphl RGGKIALFGGAGVGKTVLIQEMINRIARNFGGTSVFAGVGERTREGNDLWVELADANVLKDTALVFGQMDEPPGTRMRVAL 235
Mlep RGGKIALFGGAGVGKTVLIQEMINRIARNFGGTSVFAGVGERTREGNDLWVELQEVNVLKDTALVFGQMDEPPGTRMRVAL 242
Msme SALTMAEFFRDEQGQDVLLFIDNIFRFTQAGSEVSTLLGRMPSAVGYQPTLADEMGELQERITSTRGRSITSMQAVYVPAD 313
Mtub SALTMAEWFRDEQGQDVLLFIDNIFRFTQAGSEVSTLLGRMPSAVGYQPTLADEMGELQERITSTRGRSITSMQAVYVPAD 324
Mphl SALTMAEYFRDEQGQDVLLFIDNIFRFTQAGSEVSTLLGRMPSAVGYQPTLADEMGELQERITSTRGRSITSMQAVYVPAD 316
Mlep SALTMAEWFRDEASQDVLLFIDNIFRFTQAGSEVSTLLGRMPSAVGYQPTLADEMGELQERITSTRGRSITSMQAVYVPAD 323
Msme DYTDPAPATTFAHLDATTELSRAVFSKGIFPAVDPLASSSTIIDPAIVGDEHYRVAQEVIRILQRYKDLQDIIAILGIDEL 394
Mtub DYTDPAPATTFAHLDATTELSRAVFSKGIFPAVDPLASSSTIIDPSVVGDEHYRVAQEVIRILQRYKDLQDIIAILGIDEL 405
Mphl DYTDPAPATTFAHLDATTELSRAVFSKGIFPAVDPLASSSTIIHPSIVGEEHYRVAQEVIRILQRYKDLQDIIAILGIDEL 397
Mlep DYTDPAPATTFAHLDATTELSRSVFAKGIFPAVDPLASSSTIIDPGIVGEEHYRVAQEVIRILQRYKDLQDIIAILGIDEL 404
Msme SEEDKQLVNRARRIERFLSQNMMAAEQFTGQPGSTVPLKETIEAFDKLTKGEFDHLPEQAFFLIGGLDDLAKKAESLGAKL 475
Mtub SEEDKQLVNRARRIERFLSQNMMAAEQFTGQPGSTVPVKETIEAFDRLCKGDFDHVPEQAFFLIGGLDDLAKKAESLGAKL 486
Mphl SEEDKVLVYRARKVERFLSQNMMAAEQFTGQPGSTVPLKETIEAFDKLCKGEFDHLPEQAFFLIGGLDDLAKKAESLGAKL 478
Mlep SEEDKQLVNRARRIERFLSQNMMAAEQFTGQPGSTVPVKETIDAFDRLCKGEFDHVPEQAFFLIGGLDDLTKKAESLGAKL 485

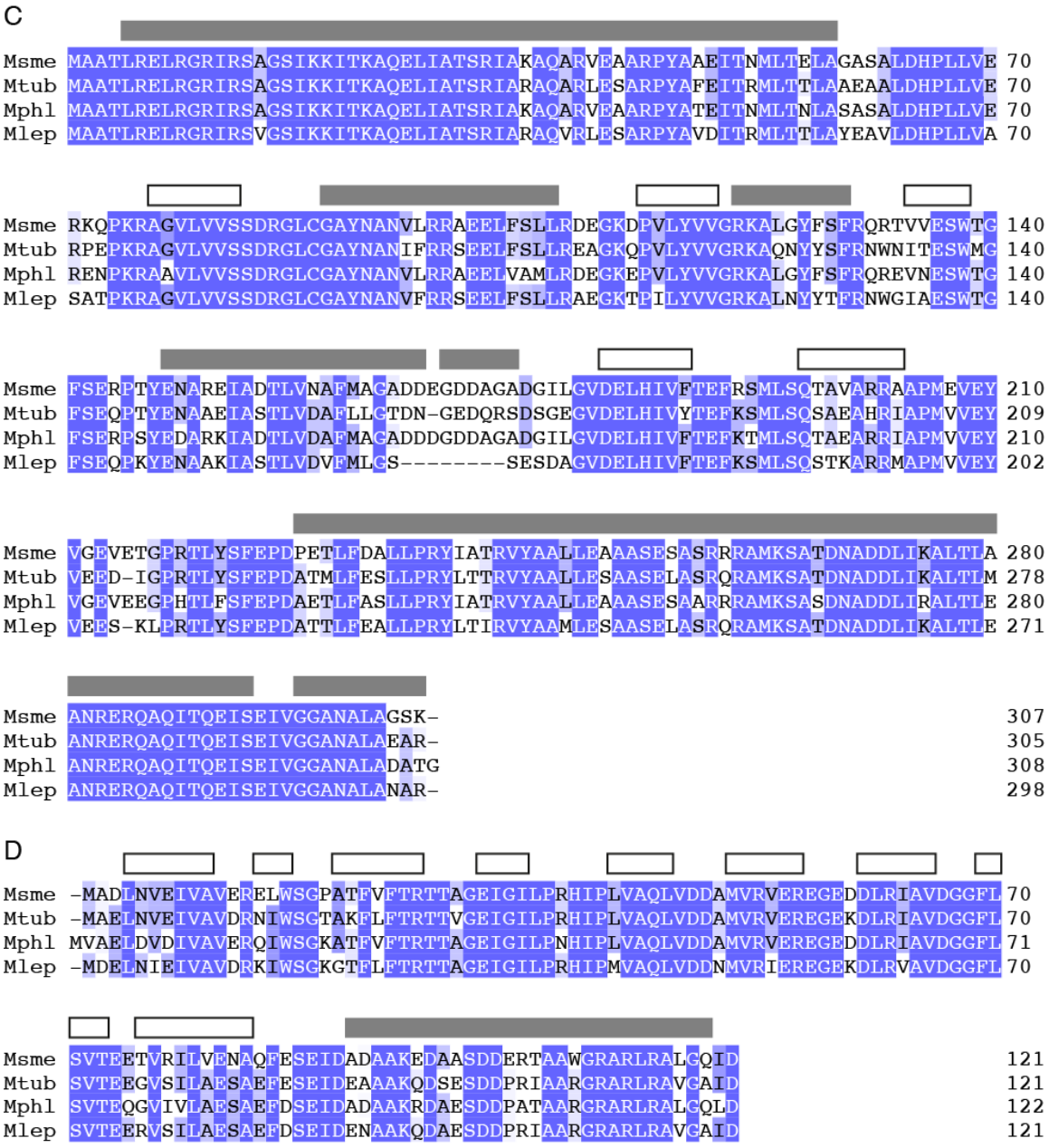


Fig S1. Comparison of the sequences of the subunits of the F_1-domain of ATP synthases from various mycobacterial species. Msme, *M. smegmatis*; Mtub, *M. tuberculosis*; Mphl, *M. phlei*; Mlep, *M. leprae*. Parts (A-D), the α-, β-, γ- and ε-subunits, respectively. Dark blue areas contain identical residues and light blue areas conserved residues. The grey and white boxes indicate regions of α-helix and β-sheet, respectively. In parts (A) and (B), orange and green lines denote the Walker A and Walker B motifs, respectively, and in (A), the red star indicates the catalytically essential "arginine finger" residue.

Table S1. Percentage identities of sequences of subunits of the F_1-domain of ATP synthases from four species of mycobacteria. The full names of the species are given in the legend to Fig. S1.

α-subunit	Msme	Mtub	Mphl	Mlep
Msme	100	84.31	91.06	82.66
Mtub	84.31	100	84.15	89.98
Mphl	91.06	84.15	100	82.33
Mlep	82.66	89.98	82.33	100

β-subunit	Msme	Mtub	Mphl	Mlep
Msme	100	92.39	94.98	90.74
Mtub	92.39	100	91.86	92.37
Mphl	94.98	91.86	100	90.17
Mlep	90.74	92.37	90.17	100

γ-subunit	Msme	Mtub	Mphl	Mlep
Msme	100	78.18	88.60	75.90
Mtub	78.18	100	76.55	83.93
Mphl	88.60	76.55	100	73.94
Mlep	75.90	83.93	73.94	100

ε-subunit	Msme	Mtub	Mphl	Mlep
Msme	100	80.17	81.82	76.03
Mtub	80.17	100	80.17	87.60
Mphl	81.82	80.17	100	77.69
Mlep	76.03	87.60	77.69	100

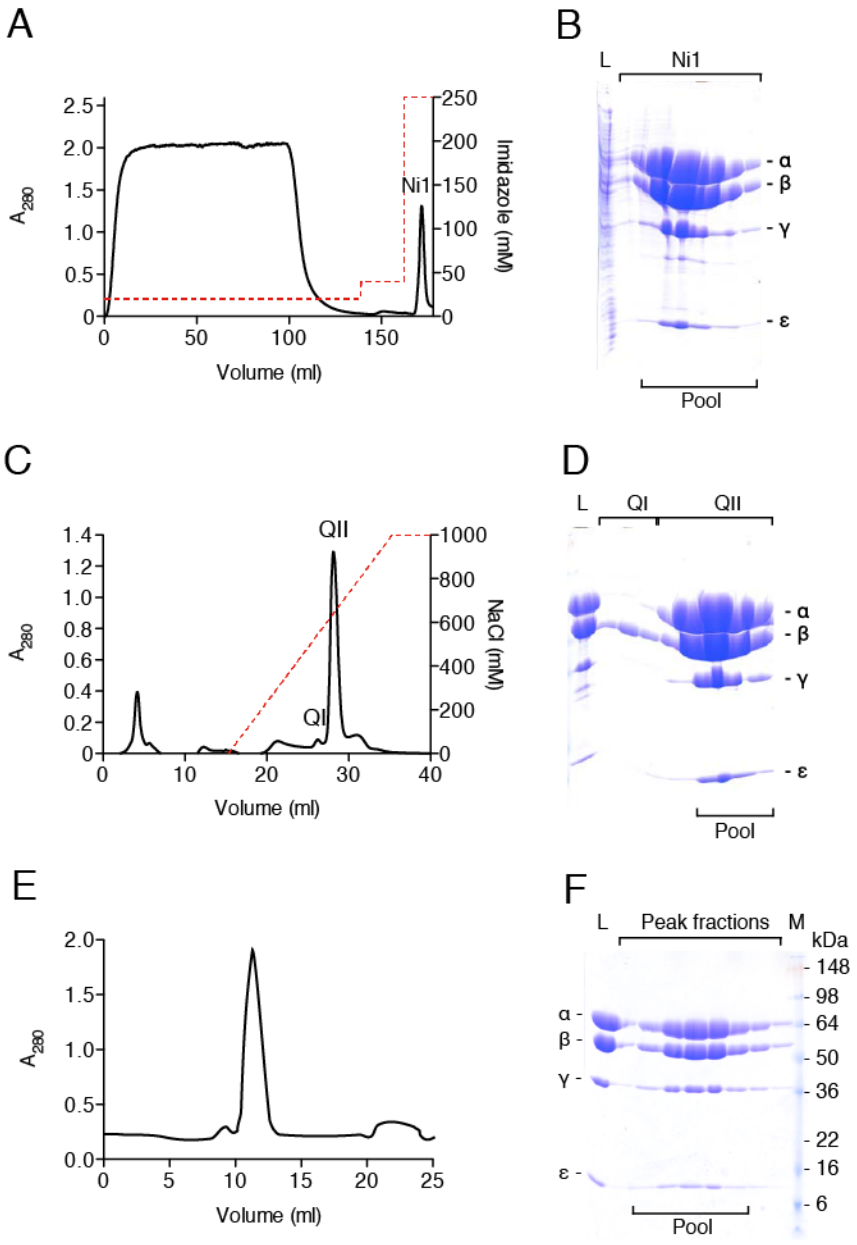


Fig. S2. Purification of F_1-ATPase from *M. smegmatis*. A, Nickel affinity chromatography. Solid line, absorbance of the effluent at 280 nm; red dashed line, gradient of imidazole. B, analysis by SDS-PAGE of fractions 22-30 from A. L, loaded sample. C, Anion exchange chromatography of the pooled fractions from peak Ni1 in A. The red dashed line is the concentration of sodium chloride. D, SDS-PAGE analysis of proteins in peak QI and QII. Fractions from peak QII were pooled as indicated. Lane L is the loaded sample; E, passage of purified F_1-ATPase from *M. smegmatis* through a column of Superdex 200 (10/300); F, analysis by SDS-PAGE of fractions from the main peak in part E. The identities of subunits of the enzyme are indicated on the left, and the positions of molecular weight markers (M) on the right.

Fig. S3. Characterization of F-ATPase purified from *M. smegmatis*. A, subunit composition analyzed by SDS-PAGE and mass-mapping of tryptic peptides. The identities of subunits are indicated on the left; B, analysis by blue native PAGE. The positions of molecular weight markers are shown on the right. In A and B, proteins were detected by staining with Coomassie blue dye.

Table S2. Molecular masses of subunits of the F_1-ATPase from *M. smegmatis*.

	Mass (Da)			
Subunit	Observed (SD[a])	Calculated[b]	Difference	Modification
α	58764.49 (5.4)	58757.6	6.9	- Met[c]
β	51546.25 (2.8)	51542.65	3.6	- Met[c] + Gly[d]
γ	33267.37 (1.8)	33266.71	0.7	- Met[c]
ε	13133.69 (0.9)	13133.62	0.1	- Met[c]

[a] SD, standard deviation; [b] calculated after the modifications given under footnotes [c] and [d]; [c] removal of the N-terminal methionine residue; [d] an additional N-terminal glycine residue, part of the TEV cleavage site.

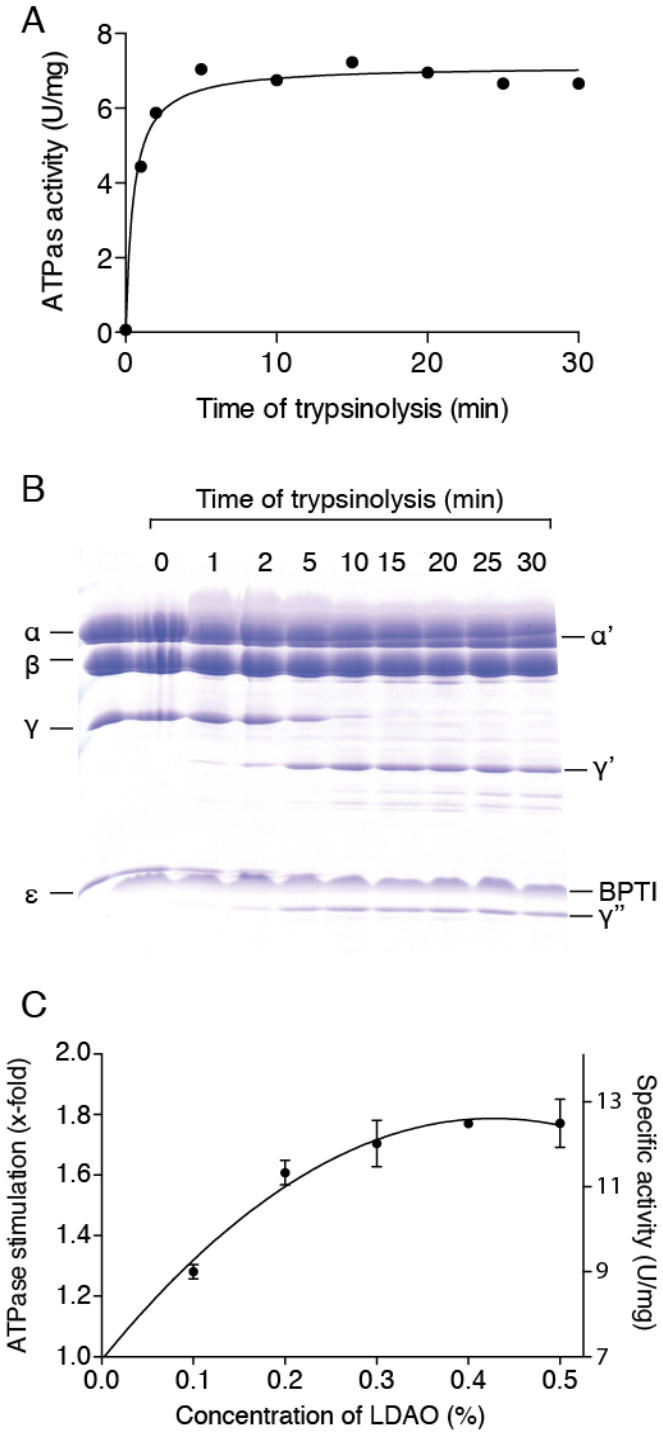


Fig S4. Activation of F_1-ATPase from *M. smegmatis* by trypsinolysis, and further activation of the trypsin treated enzyme with LDAO. The F_1-ATPase: trypsin ratio was 50:1 (w:w). Digestion at each interval was terminated with bovine pancreatic trypsin inhibitor. A, effect of trypsin treatment on the activity; B, analysis of the products of trypsinolysis by SDS-PAGE. Proteins were detected with Coomassie blue dye and

proteolytic fragments were analyzed by mass spectrometry (SI Appendix, Table S3); C, stimulation with LDAO of the activity of the enzyme that had been trypsinized for 10 min. Assays of activation of ATP hydrolysis were performed in triplicate.

Table S3. Characterization of products of trypsinolysis of F_1-ATPase from *M. smegmatis*.

Fragments of the subunits of the enzyme resulting from 10 min digestion with trypsin (see SI Appendix, Fig. S4) were analyzed by electro-spray ionization mass spectrometry. No fragments of the ε-subunit were recovered, and it appears to have been degraded extensively.

	Mass (Da)			
Subunit	Observed (SD[a])	Calculated	Difference	Fragment residues
α’	55160.36 (5.9)	55157.65	2.7	29-541
β	51545.07 (3.0)	51542.65	2.4	1-475
γ’	23814.40 (1.3)	23814.08	0.3	1-219
γ’’	9470.97 (0.4)	9470.64	0.3	220-307

[a] Standard deviations

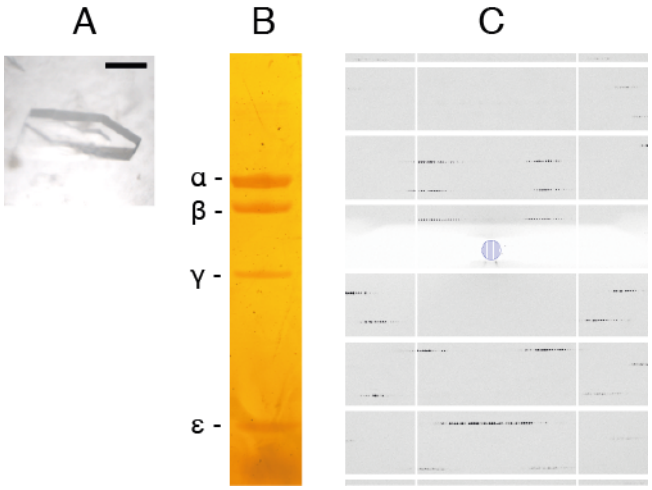


Fig. S5. Crystals of the F_1-ATPase from *M. smegmatis.* (A) a hexagonal crystal. The scale bar represents 100 μm; B, SDS-PAGE analysis of crystals. The protein bands were stained with silver. The positions of subunits are shown on the left; C, central region of X-ray diffraction image of a crystal.

Table S4. Data collection and refinement statistics. Parentheses denote the statistics for the high-resolution bin.

Parameter	
Space group	$P3_121$
Unit cell dimensions a, b, c (Å)	105.2, 105.2, 628.6
Resolution range (Å)	4.0-45.6
High-resolution bin (Å)	4.00-4.20
No. of unique reflections	35040 (4590)
Multiplicity	4.3 (4.0)
Completeness (%)	99.0 (99.0)
R_{merge}[1]	0.125 (0.512)
$\langle I/\sigma (I) \rangle$	7.0 (2.8)
B factor, from Wilson plot $(Å)^2$	108.9
R factor[2] (%)	33.1
Free R factor[3] (%)	36.7
rmsd of bonds (Å)	0.003
rmsd of angles (°)	0.73

[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.

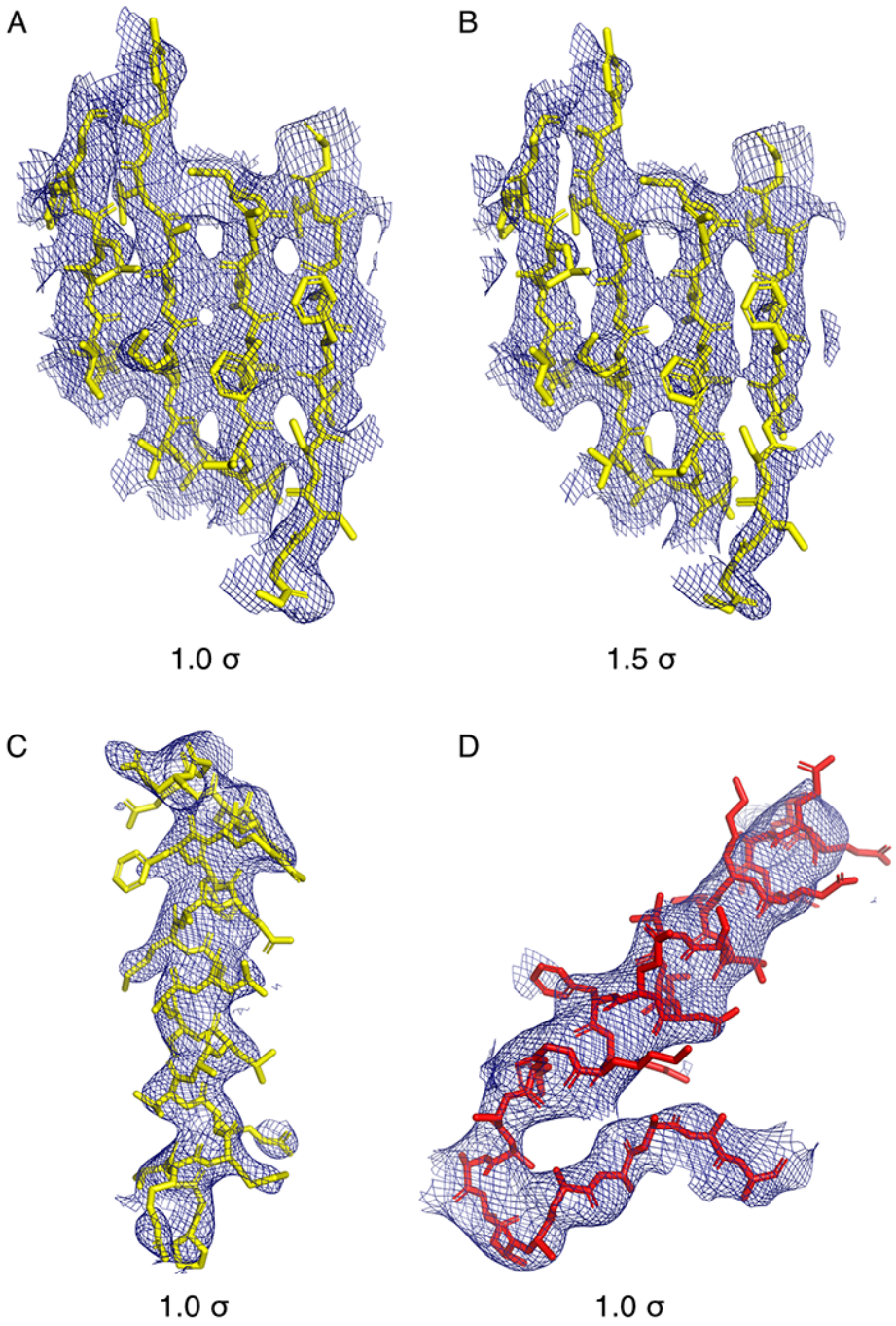


Fig. S6. Examples of regions of the 2Fo-Fc electron density map derived by X-ray analysis of crystals of F_1-ATPase from *M. smegmatis*. A and B, four β-strands consisting, from left to right, of residues155-160, 302-309, 248-255 and 182-191 of the β_{TP}-subunit contoured at 1.0 and 1.5 σ, respectively; C, the α-helix from residues 224-247 in the β_{TP}-subunit; D, residues 492-1522 in the C-terminal region of the α_E-subunit. C and D are contoured at 1.0 σ.

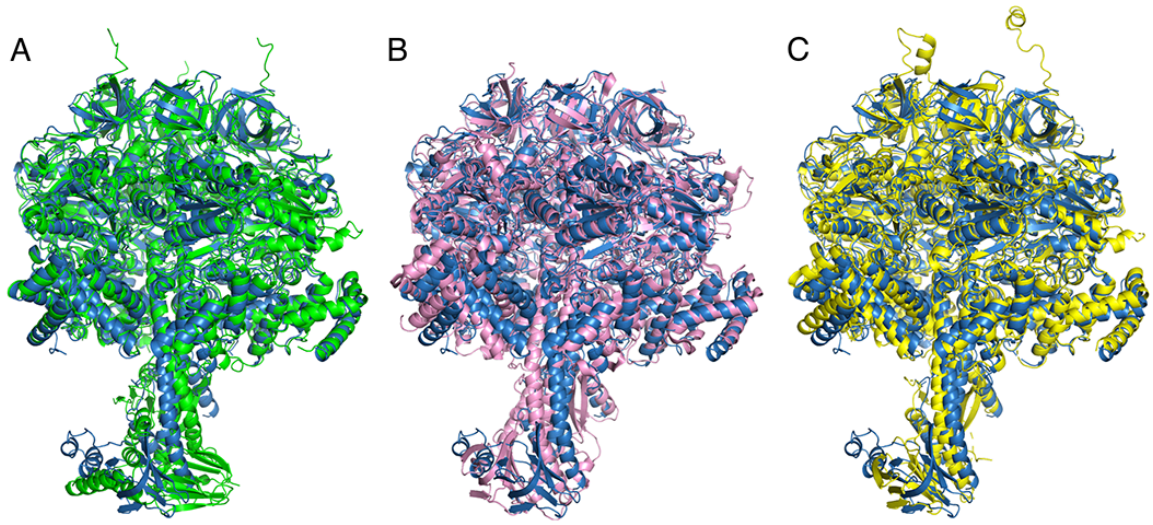


Fig. S7. Comparison of the structure of the F_1-ATPase from *M. smegmatis* with those of other F_1-ATPases. In A, B, and C, the *M. smegmatis* complex (blue) is compared with F_1-ATPases from bovine mitochondria (17) (green, 2JDI), *E. coli* (19) (pink, 3OAA molecule 1), and the F_1-domain from the structure of the intact ATP synthase from *P. denitrificans* (20) (yellow, 5DN6), respectively. For the rmsd values, see Table S5.

Table S5. Comparison of the structure of the F_1-ATPase from *M. smegmatis* with those of other F_1-ATPases.

Species	rmsd[a] (Å) all subunits [residues matched]	rmsd (Å) $\alpha_3\beta_3$ only [residues matched]
Bovine (2jdi)	2.8 [2950]	1.3 [2378]
Bovine (1e79)	4.0 [3034]	1.3 [2747]
Bovine (4xyw)	3.1 [3020]	1.2 [2739]
Bovine (4asu)	2.4 [2980]	1.6 [2734]
Bovine (2jj2 - molecule 1)[b]	1.4 [2900]	1.2 [2744]
C. thermarum (5hkk)	1.9 [3048]	0.9 [2752]
E. coli (3oaa - molecule 1)	4.4 [2998]	3.1 [2729]
P. denitrificans (5DN6)	1.8 [2919]	1.7 [2730]

[a] calculated with α-carbons only; [b] This low rmsd value arises because only the α-helical coiled-coil (α-helices αH1 and αH5 in Fig. 3) of the γ-subunit was resolved. The entire δ- and ε-subunits were not resolved in this structure.

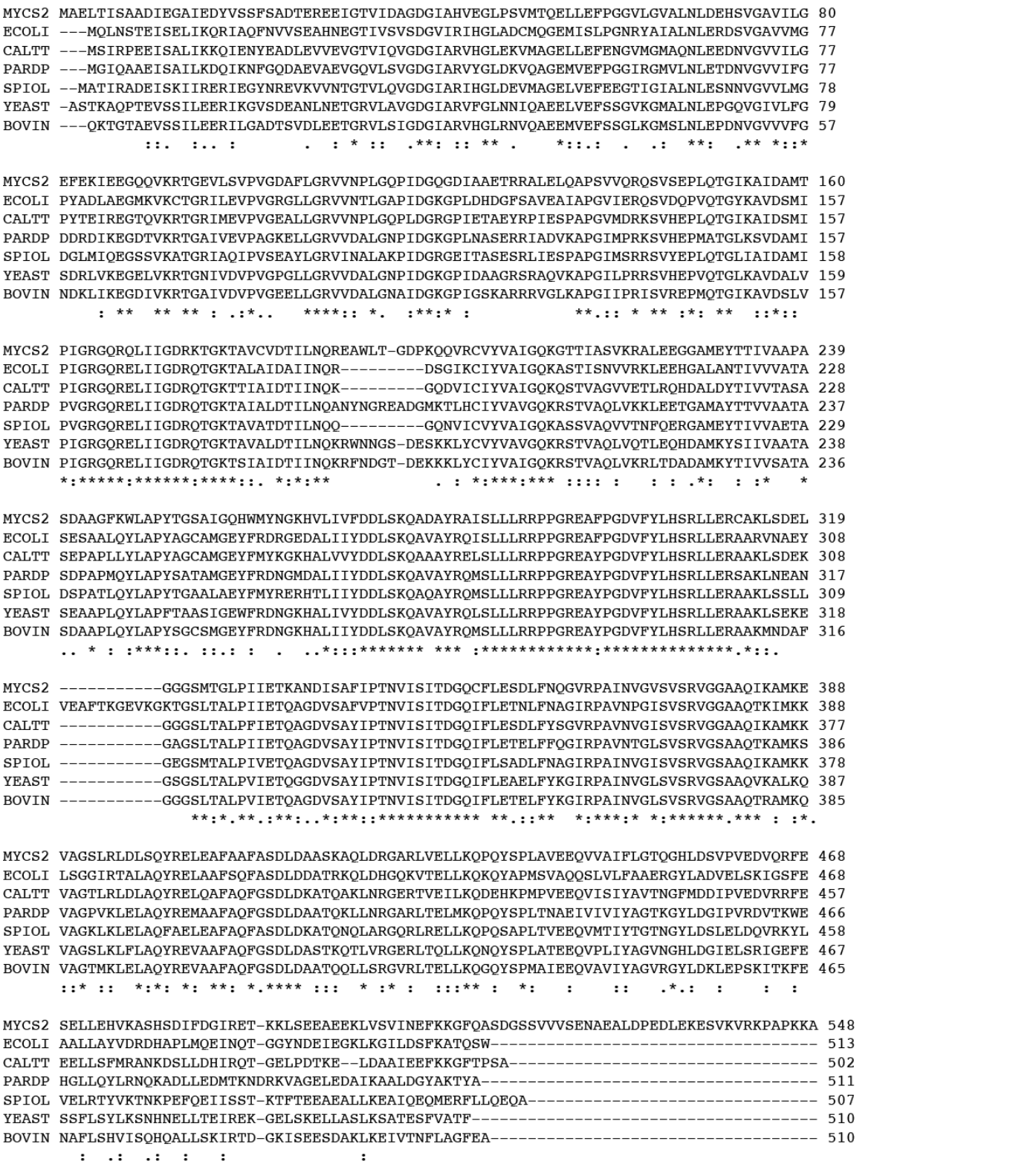


Fig. S8. Comparison of sequences of the α-subunits of ATP synthases from various sources. MYCS2, *M. smegmatis*; ECOLI, *E. coli;* CALTT, *C thermarum*; PARDP, *P.*

denitrificans; SPIOL, *Spinacia oleracea*; YEAST, *S. cerevisiae*; BOVIN, bovine mitochondria. Asterisks denote identities; colons strong conservation, and full stops, weak conservation.

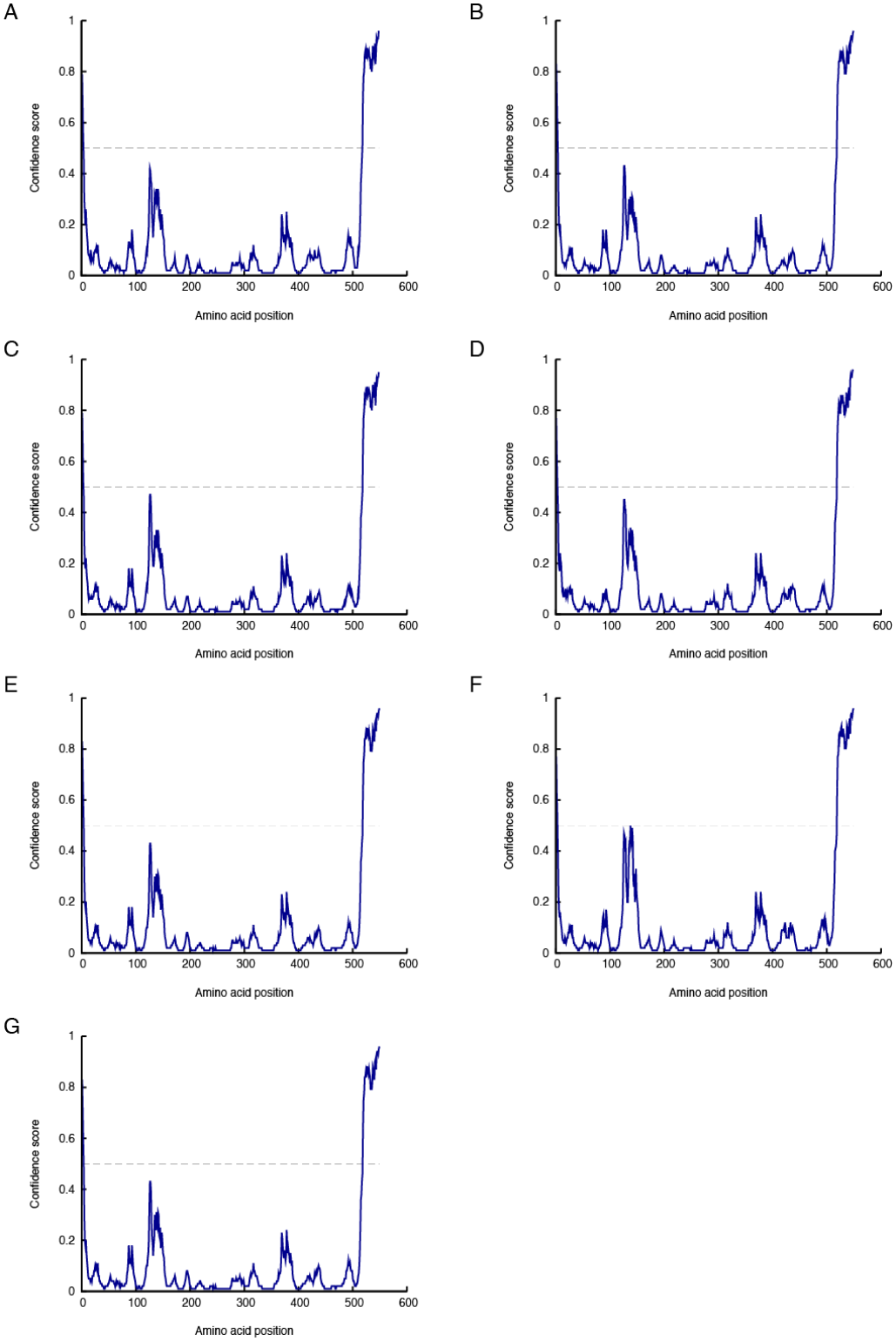


Fig. S9. Predicted regions of intrinsic disorder in the α-subunits of the F_1-ATPases from various mycobacteria. (A), *M. smegmatis*; (B), *M. tuberculosis*; (C) *M. sp JS62*; (D), *M. phlei*; (E), *M. africanum*; (F), *M. ulcerans*; (G), *M. bovis*. The dashed line at a probability of 50% crosses the subunit at approximately residue 510. The predictions were made with PSIPred, and they agree with independent predictions made with PrDOS.

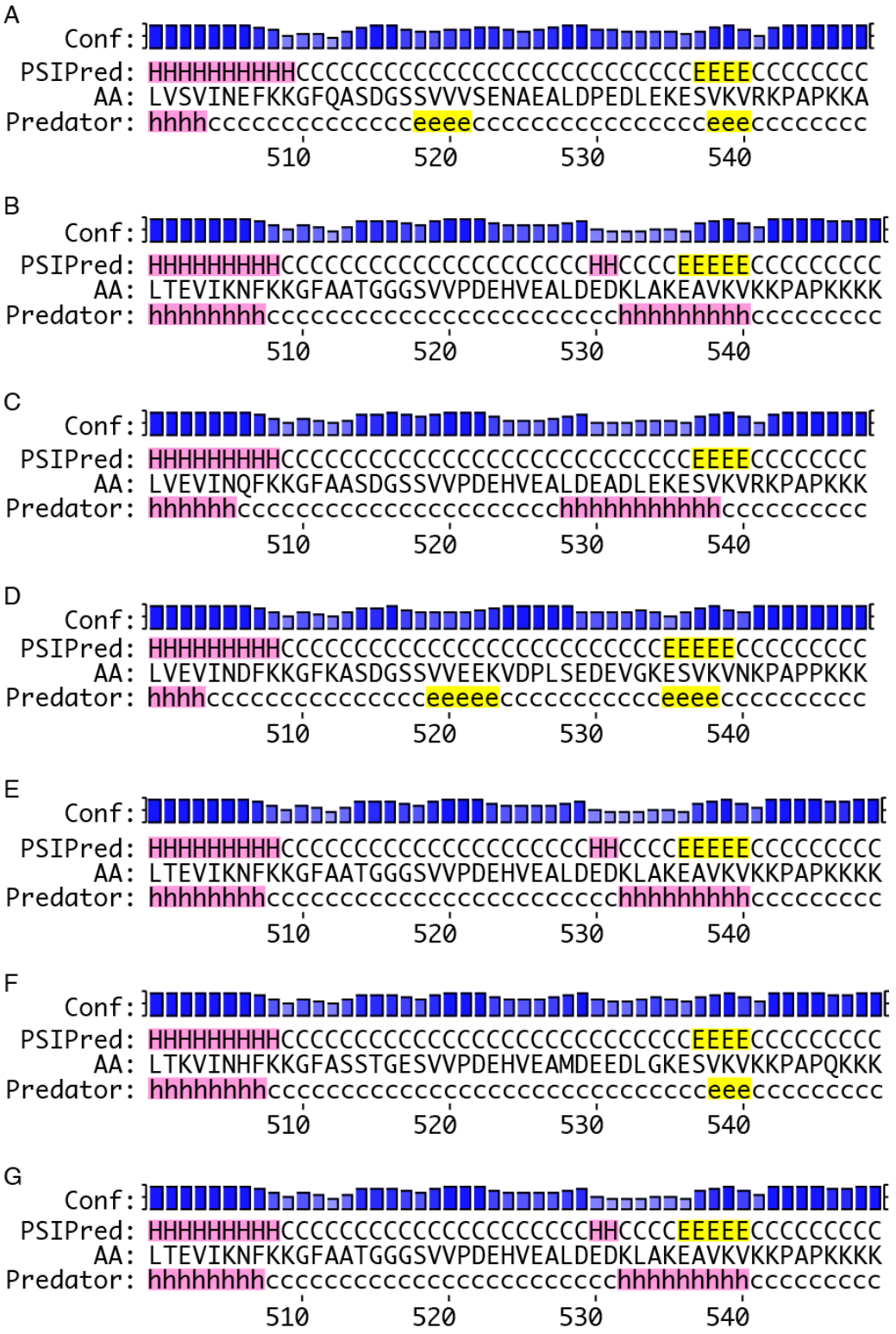


Fig. S10. Predicted secondary structures of the C-terminal regions of the α-subunits of the F-ATPases from various mycobacteria (A), *M. smegmatis*; (B), *M. tuberculosis*; (C) *M. sp*

JS623; (D), *M. phlei*; (E), *M. africanum*; (F), *M. ulcerans*; (G), *M. bovis*. Predictions with PSIPred and Predator are shown. The confidence score applies to PSIPred only.

```
MYCS2   MAATLRELRGRIRSAGSIKKITKAQELIATSRIAKAQARVEAARPYAAEITNMLTELAGA 60
ECOLI   -MAGAKEIRSKIASVQNTQKITKAMEMVAASKMRKSQDRMAASRPYAETMRKVIGHLAHG 59
CALTT   -MQGMREIKRRIRSVKNTRQITKAMKMVAAAKLRRAQETAENARPYADKIKEVISSIAAG 59
PARDP   -MPSLKDLKNRIGSVKNTRKITKAMQMVAAAKLRRAQEAAEAARPYADRMAAVMAGLTAA 59
SPIOL   --ANLRELRDRIGSVKNTQKITEAMKLVAAAKVRRAQEAVVNGRPFSETLVEVLYNMNEQ 58
SYNP6   -MANLKAIRDRIKSVRNTRKITEAMRLVAAAKVRRAQEQVLSTRPFADRLAQVLAGLQQR 59
YEAST   --ATLKEVEMRLKSIKNIEKITKTMKIVASTRLSKAEKAKISAKKMDEAEQLFYKNAE-- 56
BOVIN   --ATLKDITRRLKSIKNIQKITKSMKMVAAAKYARAERELKPARVYGVGSLALYEKAD-- 56
             : :  :: *  . .:**:: .::*:::  :::      :        .

MYCS2   ---SALDHPLLVERKQPKRAGVLVVSSDRGLCGAYNANVLRRAEELFSLLRDEGKDPVLY 117
ECOLI   NL--EYKHPYL-EDRDVKRVGYLVVSTDRGLCGGLNINLFKKLLAEMKTWTDKGVQCDLA 116
CALTT   TK--DFSHPML-EARPVKKTGYMVITSDRGLAGPYNANILRLVSKTIEERHQSKDEYVIF 116
PARDP   AAGSDMAPRLLAGTGEDRRHLLVVMTSERGLAGGFNSSIVKLARLRLQELQAQGKQVSIL 119
SPIOL   LQTEDVDVPLT-KIRTVKKVALMVVTGDRGLCGGFNNMLLKKAESRIAELKKLGVDYTII 117
SYNP6   LQFENVDLPLL-QRREVKTVALLVVSGDRGLCGGYNSNVIRRAEQRARELSAQGLDYKFV 118
YEAST   --TKNLDVEA--TETGAPKELIVAITSDKGLCGSIHSQLAKAVRRHLND----QPNADIV 108
BOVIN   --IK--------TPEDKKKHLIIGVSSDRGLCGAIHSSVAKQMKSEAANLAAAGKEVKII 106
                           : :: ::**.*  :  : :              :  :

MYCS2   VVGRKALGYFSFRQRTVVESWTGFSER--PTYENAREIADTLVNAFMAGADDEGDDAGAD 175
ECOLI   MIGSKGVSFFNSVGGNVVAQVTGMGDN--PSLSELIGPVKVMLQAYDEGR---------- 164
CALTT   AVGRKGRDFFKKRGYPVVEEVTGISDT--PSLTEIQDIAQSAIGMFADET---------- 164
PARDP   TVGKKGREQLKREYGDLFVNHVDLSEVKRIGYDNARAIADEILDRFDNGE---------- 169
SPIOL   SIGKKGNTYFIRRPEIPVDRYFDGTNL--PTAKEAQAIADDVFSLFVSEE---------- 165
SYNP6   IVGRKAGQYFQRREQPIEATYSGLEQI--PTAQEANDIADQLLSLFLSGT---------- 166
YEAST   TIGDKIKMQLLRTHPNNIKLSINGIGKDAPTFQESALIADKLLSVMKAGT---------- 158
BOVIN   GVGDKIRSILHRTHSDQFLVTFKEVGRRPPTFGDASVIALELLNS--GYE---------- 154
         :* *    :                       :    .   .

MYCS2   GILGVDELHIVFTEFRSMLSQTAVARRAAPMEVEYV------------------------- 211
ECOLI   ----LDKLYIVSNKFINTMSQVPTISQLLPLPASDD------------------------ 196
CALTT   ----FDKLTIFYNEFVSPIVQRPVEKQLLPLTSEEV------------------------ 196
PARDP   ----FDVATLFYNRFESVISQVPTARQVIPAVIEEG------------------------ 201
SPIOL   ----VDKVEMLYTKFVSLVKSDPVIHTLLPLSPKGEICDINGKCVDAAEDELFRLTTKEG 221
SYNP6   ----VDRVELVYTKFLSLVASNPVVQTLLPLDPQGL---------ASSDDEIFRLTTRGG 213
YEAST   ----YPKISIFYNDPVSSLSFEPSEKPIFNAKTIEQ------------------------ 190
BOVIN   ----FDEGSIIFNRFRSVISYKTEEKPIFSLDTISS------------------------ 186
                 :. .   . :

MYCS2   ---------GEVETGPRTLYSFEPD-PETLFDALLPRYIATRVYAALLEAAASESASRRR 261
ECOLI   ----------DDLKHKSWDYLYEPD-PKALLDTLLRRYVESQVYQGVVENLASEQAARMV 245
CALTT   ------------LDGPVSAYEYEPD-SESVLEVLLPKYAETLIYSALLDAKASEFGARMT 243
PARDP   ------------EAGASSLYDYEPD-ENAILNDLLPRSVATQVFAALLENAASEQGARMT 248
SPIOL   KLTVERDMIKTETPAFSPILEFEQD-PAQILDALLPLYLNSQILRALQESLASELAARMT 280
SYNP6   SFTVEREKLTSEVAPLPRDMIFEQD-PAQILSALLPLYLSNQLLRALQEAAASELAARMT 272
YEAST   ---------------SPSFGKFEIDTDANVPRDLFEYTLANQMLTAMAQGYAAEISARRN 235
BOVIN   ---------------AESMSIYD-DIDADVLRNYQEYSLANIIYYSLKESTTSEQSARMT 230
                           :: *    :          . :  .: :  ::* .:*

MYCS2   AMKSATDNADDLIKALTLAANRERQAQITQEISEIVGGANALAGSK 307
ECOLI   AMKAATDNGGSLIKELQLVYNKARQASITQELTEIVSGAAAV---- 287
CALTT   AMGNATDNATEMLETLTLQFNRARQAAITQEIAEIVAGANALR--- 286
PARDP   AMDNATRNAGDMIDRLTTVYNRSRQAAITKELIEIISGAEAL---- 290
SPIOL   AMSNATDNANELKKTLSINYNRARQAKITGEILEIVAGANACV--- 323
SYNP6   AMNSASDNANALVGQLTLVYNKARQAAITQELLEVVAGAEALNG-- 316
YEAST   AMDNASKNAGDMINRYSILYNRTRQAVITNELVDIITGASSLG--- 278
BOVIN   AMDNASKNASEMIDKLTLTFNRTRQAVITKELIEIISGAAALD--- 273
        **  *: *.  :        *: *** ** *: ::: ** :
```

Fig. S11. Alignment of sequences of the γ-subunits of ATP synthases from various sources. For definition of species and definition of symbols, see legend to SI Appendix, Fig. S8. The additional sequence SYNP6 relative to SI Appendix, Fig. S8 is from *Synechcococcus* 6301.

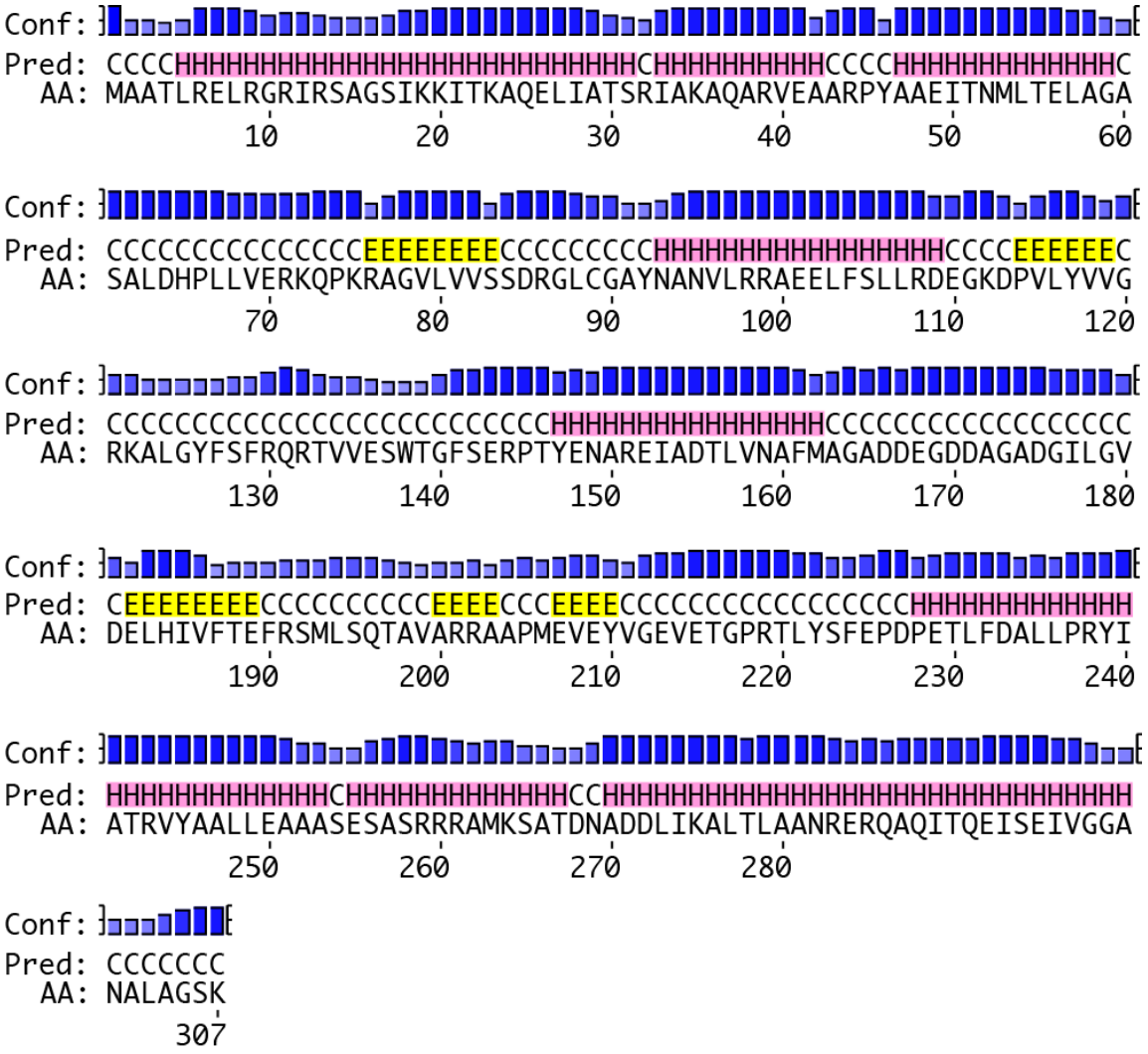


Fig. S12. Predicted secondary structure of the γ-subunit from *M. smegmatis* with PSIPred. The insert (residues 166-179), predicted to be random coil, is in the spring domain.

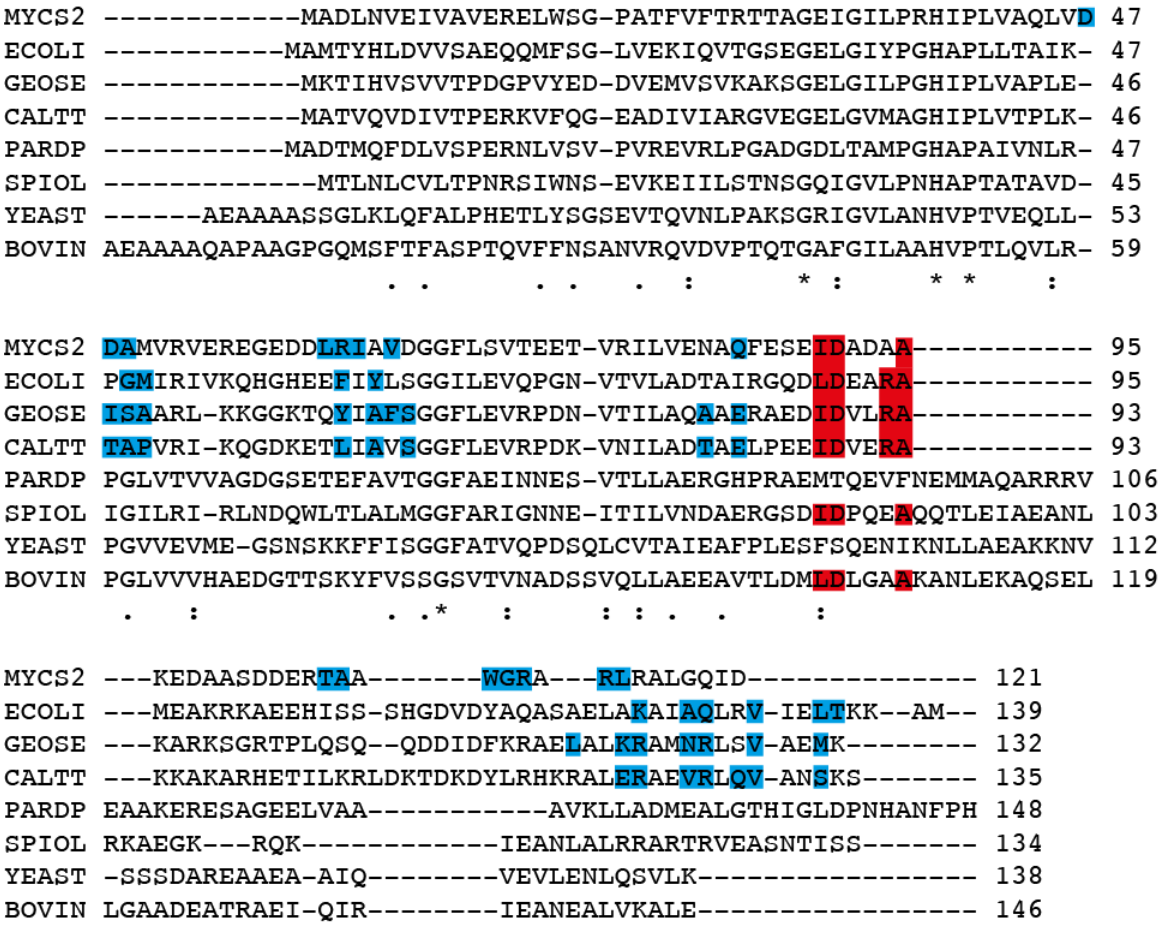


Fig. S13. Alignment of sequences of the ε-subunits of ATP synthases from various species. For definition of species and definition of symbols, see legend to SI Appendix, Fig. S8; GEOSE, *G. stearothermophilus*. The residues marked in red are required for binding an ATP molecule to the ε-subunits in *E. coli*, *C. thermarum, and G. stearothermophilus*. The residues marked with blue preceding the red residues are in the N-terminal domains of the structures of the ε-subunits from *M. smegmatis*, *E. coli* and *G. stearothermophilus*. They are involved in forming interactions within 4 Å with residues in blue that follow the red residues which are found in the C-terminal domains of the respective subunits.

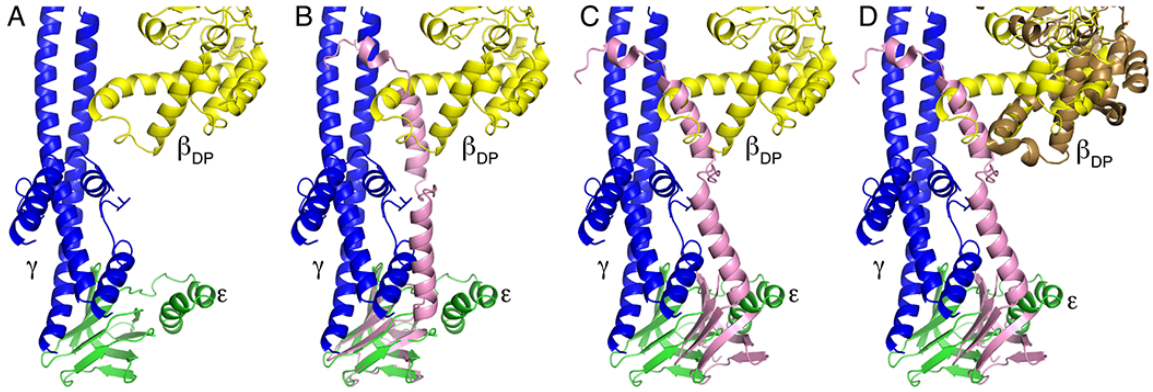


Fig. S14. Superimposition of the ε-subunit from *E. coli* F_1-ATPase (19) in the "up" position upon the structure of the F_1-ATPase from *M. smegmatis*. A, side view of the lower part of F_1-ATPase from *M. smegmatis* showing parts of the β_{DP}- and γ- subunits (yellow and blue respectively), and the entire ε-subunit (green) in the "down" position; B, as A, but with the *E. coli* ε-subunit (pink) in the "up" position superimposed by alignment of their N-terminal domains; C, as B, with the alignment via the $\alpha_3\beta_3$ domains of the F_1-ATPases from *M. smegmatis* and *E. coli*; D, as C, but additionally with the β_{DP}-subunit from *E. coli* (brown). Its position is displaced relative to the *M. smegmatis* β_{DP}-subunit by the rotation of the E. coli γ-subunit to the right compared with the *M. smegmatis* γ-subunit, which also changes the position of the C-terminal domain of the *E. coli* ε-subunit. The C-terminal α-helix of the *E. coli* ε-subunit clashes with the DELSEED region of the *M. smegmatis* β_{DP}-subunit. The shorter C-terminal α-helices of the *M. smegmatis* ε-subunit could theoretically reach inside its F_1-domain to about the position of the end of turn 2 of helix 2 in *E. coli*, but the lack of displacement in the *M. smegmatis* structure of the β_{DP}- and γ-subunits similar to the observed displacement in the *E. coli* complex (part D), prevents this theoretical structure of the *M. smegmatis* ε-subunit from being accommodated.

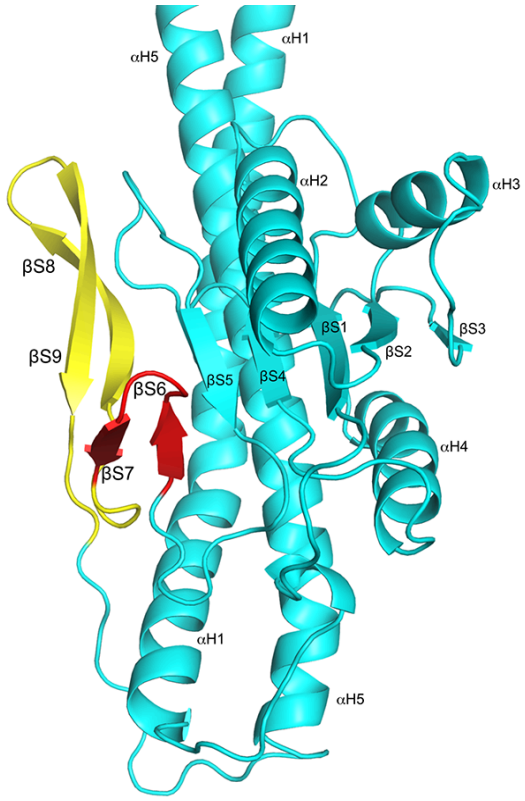


Fig. S15. Detail of the lower part of the γ-subunit of the ATP synthase from spinach chloroplasts (21) (PDB code: 6fkf). αH1 and αH5 are the coiled-coil region of the subunit. The associated Rossmann fold is on the right of the coiled-coil. The yellow and red regions represent the structure of the additional region characteristic of chloroplast ATP synthases that is involved in the inhibition of ATP hydrolysis during the hours of darkness by the formation of a disulfide linkage between cysteine residues at each end of the nine residue red structure. The yellow segments of sequence, but not the red ones, are conserved in cyanobacteria.

Supplementary References

1. Wang F *et al.* (2010) Mycobacterium tuberculosis dihydrofolate reductase is not a target relevant to the antitubercular activity of isoniazid. *Antimicrob Agents Chemother* 54(9):3776–3782.
2. Kapust RB *et al.* (2001) Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng* 14(12):993–1000.
3. Pullman ME, Penefsky H, Datta A, Racker E (1960) Partial resolution of the enzymes catalysing oxidative phosphorylation. Purification and properties of soluble, dinitrophenol-stimulated adenosine triphosphatase. *J Biol Chem* 235:3322–3329.
4. Winn MD *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235–242.

5. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AGW (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):271–281.
6. Evans PR, Murshudov GN (2013) How good are my data and what is the resolution. *Acta Crystallogr D Biol Crystallogr* 69(Pt 7):1204–1214.
7. McCoy AJ *et al.* (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658–674.
8. 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.
9. Murshudov GN *et al.* (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):355–367.
10. Nicholls RA, Fischer M, McNicholas S, Murshudov GN (2014) Conformation-independent structural comparison of macromolecules with ProSMART. *Acta Crystallogr D Biol Crystallogr* 70(Pt 9):2487–2499.
11. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
12. Chen VB *et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12–21.
13. Schrodinger, LLC (2018) The PyMOL Molecular Graphics System, Version 2.2.2.
14. UniProt C (2015) UniProt: a hub for protein information. *Nucleic Acids Res* 43(Database issue):D204–12.
15. 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.
16. Ishida T, Kinoshita K (2007) PrDOS: prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Res* 35(Web Server issue):W460–4.
17. 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.
18. 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.
19. 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.
20. 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.
21. Sobti M *et al.* (2016) Cryo-EM structures of the autoinhibited E. coli ATP synthase in three rotational states. *Elife* 5