Crystal structure of the ε subunit of the proton-translocating ATP synthase from *Escherichia coli*

Ulla Uhlin[1†], Graeme B Cox[2] and J Mitchell Guss[1*]

Background: Proton-translocating ATP synthases convert the energy generated from photosynthesis or respiration into ATP. These enzymes, termed F_0F_1-ATPases, are structurally highly conserved. In *Escherichia coli*, F_0F_1-ATPase consists of a membrane portion, F_0, made up of three different polypeptides (*a*, *b* and *c*) and an F_1 portion comprising five different polypeptides in the stoichiometry $\alpha_3\beta_3\gamma\delta\varepsilon$. The minor subunits γ, δ and ε are required for the coupling of proton translocation with ATP synthesis; the ε subunit is in close contact with the α, β, γ and c subunits. The structure of the ε subunit provides clues to its essential role in this complex enzyme.

Results: The structure of the *E. coli* F_0F_1-ATPase ε subunit has been solved at 2.3 Å resolution by multiple isomorphous replacement. The structure, comprising residues 2–136 of the polypeptide chain and 14 water molecules, refined to an R value of 0.214 (R_{free} = 0.288). The molecule has a novel fold with two domains. The N-terminal domain is a β sandwich with two five-stranded sheets. The C-terminal domain is formed from two α helices arranged in an antiparallel coiled-coil. A series of alanine residues from each helix form the central contacting residues in the helical domain and can be described as an 'alanine zipper'. There is an extensive hydrophobic contact region between the two domains providing a stable interface. The individual domains of the crystal structure closely resemble the structures determined in solution by NMR spectroscopy.

Conclusions: Sequence alignments of a number of ε subunits from diverse sources suggest that the C-terminal domain, which is absent in some species, is not essential for function. In the crystal the N-terminal domains of two ε subunits make a close hydrophobic interaction across a crystallographic twofold axis. This region has previously been proposed as the contact surface between the ε and γ subunits in the complete F_1-ATPase complex. In the crystal structure, we observe what is apparently a stable interface between the two domains of the ε subunit, consistent with the fact that the crystal and solution structures are quite similar despite close crystal packing. This suggests that a gross conformational change in the ε subunit, to transmit the effect of proton translocation to the catalytic domain, is unlikely, but cannot be ruled out.

Addresses: [1]Department of Biochemistry, University of Sydney, NSW 2006, Australia and [2]John Curtin School of Medical Research, Australian National University, Canberra, ACT 0200, Australia.

[†]Present address: Department of Molecular Biology, Swedish University of Agricultural Sciences, PO Box 590, S-751 24 Uppsala, Sweden.

*Corresponding author.
E-mail: M.Guss@biochem.usyd.edu.au

Key words: ATPase, ATP synthase, crystal structure, ε subunit

Received: **27 June 1997**
Revisions requested: **22 July 1997**
Revisions received: **7 August 1997**
Accepted: **11 August 1997**

Structure 15 September 1997, **5**:1219–1230
http://biomednet.com/elecref/0969212600501219

Introduction

The proton-translocating F_0F_1-ATP synthases, located in mitochondrial, chloroplast and bacterial membranes, catalyse the final step in oxidative or photophosphorylation. There are a number of excellent reviews on the structure and mechanism of the ATP synthase [1–3]. The same enzymes can also catalyse ATP driven proton translocation and in this role are termed ATPases. The structure of F_0F_1-ATP synthase is highly conserved among species and the enzyme can be readily separated into two portions: the water-soluble F_1-ATPase and the membrane-bound F_0 portion which acts as a proton pore. The F_0 portion of *Escherichia coli* ATP synthase is made up of three different polypeptides with the approximate composition, ab_2c_{9-12}. The F_1 portion of *E. coli* ATP synthase contains five different polypeptides with the composition, $\alpha_3\beta_3\gamma\delta\varepsilon$. The major subunits of mitochondrial ATPases have sufficient sequence identity to suggest similar functions and are described by the same names [4]. The ε subunit of *E. coli* ATP synthase, however, only has a low sequence identity with the supposed equivalent δ subunit found in mitochondrial systems. We present evidence from the structure that δ from mitochondrial F_1-ATPase, in part at least, serves the same function as ε from the *E. coli* ATPase. The F_1-ATPase has been crystallised and partial structures obtained for the rat liver [5], beef heart mitochondrial [6] and *Bacillus* PS3 [7] enzymes. While the structures of the α and β subunits were unambiguous in the bovine and bacterial enzymes, the structures of the small subunits δ and ε were not

resolved in any of the structures. The γ subunit was partially resolved only in the beef heart F_1-ATPase. The structure of the bovine mitochondrial enzyme reveals alternating α and β subunits arranged in a ring with the visible part of the γ subunit like an axle in the centre of a wheel. The three catalytic sites, which lie in the β subunits close to the interface with a neighbouring α subunit, were each found to be in a different ligand-binding state: the first was empty; the second contained ADP; and the third contained AMP-PNP, a non-hydrolysable ATP analogue. The γ subunit, which has no internal symmetry, interacts differently with each of the catalytic subunits. This is consistent with the binding change mechanism [3,8], a proposed mechanism in which the energy requiring step is not the synthesis of ATP but its release from the active site, and where the binding and release occur simultaneously at separate sites. The binding changes may be explained by a physical rotation of the γ subunit with respect to the α and β subunits [6]. Several experiments, using isolated F_1-ATPase, have now been conducted which support the rotational hypothesis. One of the most convincing experiments reported the cross-linking of the γ subunit to one of the β subunits followed by the replacement of the two remaining β subunits with equivalent radiolabelled subunits. The cross-link was then broken, allowing catalysis to occur, and reformed to see whether the cross-link had moved to one of the radiolabelled subunits [9]. In the absence of catalytic turnover most cross-links reformed to the original unlabelled β subunit but after catalysis the cross-links were randomly distributed among the three β subunits. In addition, a direct observation of the rotation has been made in an experiment in which a fluorescent label, attached to an actin filament coupled to the γ subunit, was seen to rotate with respect to the immobilised $\alpha_3\beta_3$ hexamer in the presence of ATP [10].

What remains to be determined is how the proton translocation is coupled to the rotation of the γ subunit and the synthesis of ATP and in particular the roles of the other minor subunits δ and ε. The ε subunit is known to inhibit ATPase activity in isolated F_1-ATPase [11,12] and as a result ε is sometimes called an inhibitory subunit. In addition, and perhaps more importantly, the ε subunit is essential for the coupling of proton translocation to ATP synthesis. Absence of the ε subunit prevents the binding of F_1 to F_0 [13]. Cross-linking studies have shown that the ε subunit in the *E. coli* enzyme is in close contact with the γ subunit [14–16], the c subunits of F_0 [17], and the α and β subunits [18–20]. It has been proposed that the ε subunit changes conformation or moves in order to make new interactions with different copies of the other subunits during the catalytic cycle [21]. Solution structures of the ε subunit and part of the δ subunit of *E. coli* ATP synthase have been determined using NMR spectroscopy [22,23].

We have determined the structure of the double mutant, Ala1→Gly, Met2→Ser (A1G, M2S), ε subunit of *E. coli* F_0F_1-ATPase using X-ray crystallography at 2.3 Å resolution. We are now able to give a detailed description of the molecular structure including the interactions between the two domains which are relevant to the hypothesis of a conformational change within the ε subunit.

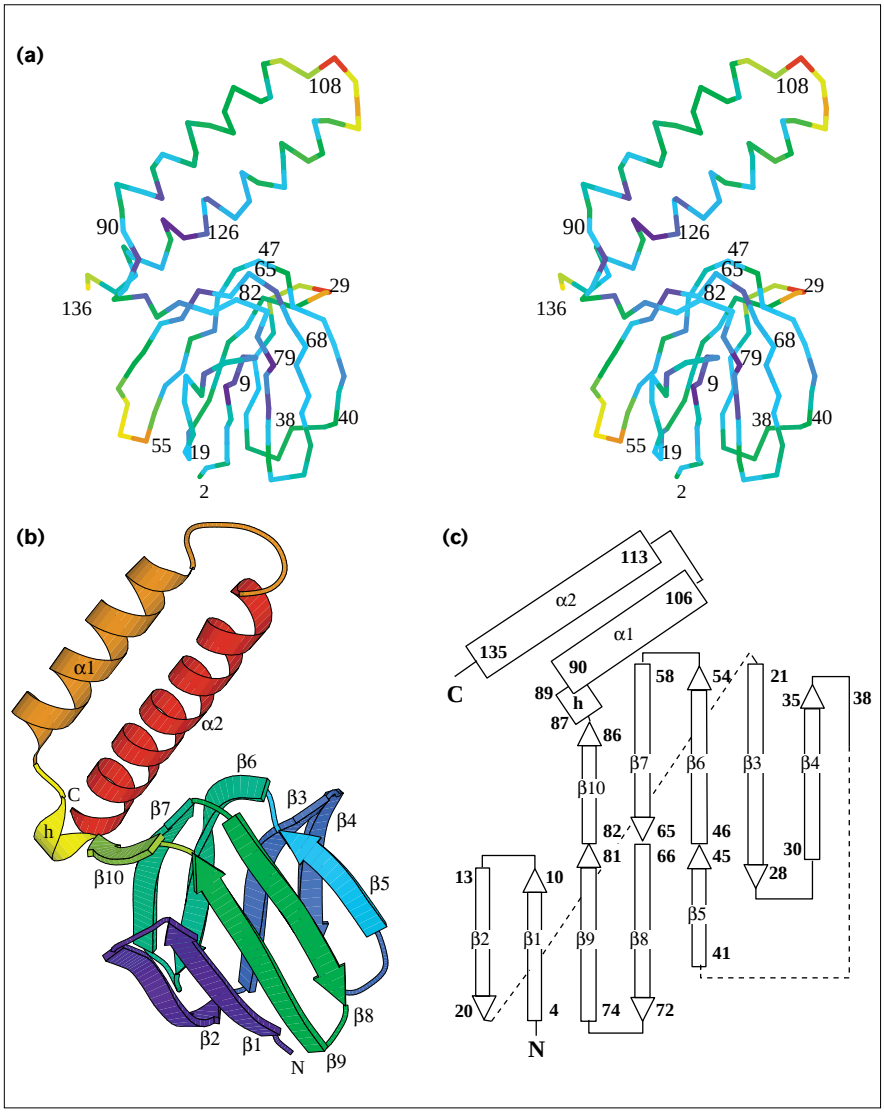
Results

Overall structure of the F_1-ATPase ε subunit

The structure of the *E. coli* F_1ATPase ε subunit forms two distinctive domains as shown in Figure 1. The first 85 residues comprise the β sandwich N-terminal (or β) domain and residues 90–136 form the C-terminal (or α) domain, which comprises an antiparallel coil of two α helices. The two domains are linked by a short loop, residues 86–89, containing a single turn of 3_{10} helix. The α domain folds back over the top of the β domain creating an extensive contact interface that buries a total surface area of 910 Å^2. The overall fold resembles a hand with the three smallest fingers folded against the palm representing the β-sandwich domain and the extended thumb and forefinger representing the two α helices of the C-terminal domain. A search with the program DALI [24] failed to find a close homologue to the entire structure in the data base. Each individual domain in the ε subunit is very similar to the published NMR solution structure [22] (Table 1). There was, however, little NMR data to position one domain accurately with respect to the other, and only a cursory description of the intrasubunit interactions is given in the publication of the structure [22]. The combination of the spin-label data and the few long-range distances derived from nuclear Overhauser effects (NOE) did yield an overall structure for the ε subunit consistent with the description given here.

The N-terminal domain has two predominantly antiparallel β sheets, each of five strands, arranged as a sandwich. No structure in the data base was found with the same sheet topology and architecture. The first sheet, strands β2, β1, β9, β8 and β5, has a single parallel pair of strands, β1 and β9 (Figure 1c). The second, strictly antiparallel sheet, is comprised of strands β10, β7, β6, β3 and β4. An unusual feature of the β sandwich is that the short connecting loops between the pairs of strands β5 and β6, β7 and β8, and β9 and β10, result in an almost continuous sheet of three long bent strands that has been folded nearly in half. Each sheet has two adjacent β hairpins flanked on one end by a short strand, either β5 or β10. The apparent twofold symmetry that relates the two sheets is broken by the reversal of strands β8 and β9. The sidechains pointing to the interior of the β sandwich are hydrophobic, but the only aromatic residues with their sidechains pointing into the core of the sandwich are Tyr4 and Phe16. In addition, one face of the β sandwich has a large hydrophobic patch formed by residues Val9,

Figure 1



The structure of the *E. coli* ATP synthase ε subunit. **(a)** Stereo diagram of the Cα trace with residue numbers highlighting specific points mentioned in the text. The colours indicate average B factor values for each residue with the lowest values in blue and the highest in red. **(b)** Structure of the protein with arrows representing β sheets and coils representing α helices. Colours are ramped from the N terminus in blue to the C terminus in red. (The figure was prepared using the programs O [28] and MOLSCRIPT [42].) **(c)** Topology diagram showing β sheets as arrows and α helices as cylinders. The first and last residues of each secondary structure element are numbered; the N and C termini are marked.

Leu42, Ala44, Ile68 and Leu79 on strands β2, β1, β9, β8 and β5 (Figure 2). This patch may have functional significance for the interaction of the ε subunit with the γ subunit (see below).

The C-terminal helical domain has an antiparallel coiled-coil of two α helices joined by a tight loop which runs from residues Ser107 to Val112. This arrangement of helices occurs as pairs in some four-helix bundle proteins: 46 Cα positions from myohemerythrin [25] can be superposed onto the C-terminal domain with an rmsd of 1.6 Å. The hydrophobic contact surface between the helices is formed by Ala94, Ala97, Ala101 and Ile105 (on the outer helix) and Ala117, Leu121, Ala124 and Leu128 (on the inner helix), forming what might be termed an 'alanine zipper' (Figure 3). These alanine residues actually interdigitate like the teeth of a zipper with the alanines on one helix pointing directly between the two residues on the opposite helix. This is different from the parallel leucine zipper structures, in which the leucines from opposing helices are at the same height when the helix is viewed perpendicular to its axis. The overall spacing of the helices is, however, very similar to that observed in antiparallel helical coiled-coils, such as in myohemerythrin, which do not have alanine residues in the interior, indicating that larger sidechains can be accommodated without altering the distance between the helices. Nevertheless, most of these alanine residues are highly conserved or conservatively substituted in the family of F_1-ATPase ε subunits (Figure 4). The 'outer' face of the longer C-terminal α helix also has a hydrophobic surface formed by residues Ala115, Ala119, Ala122, Ala126, Val130

Table 1

Comparison of the X-ray and NMR structures*.

Secondary structure	Residues included	Number of residues	Rms Δ (Å)
β Strands	4–10, 13–28, 30–35, 41–54, 58–72, 74–86	72	1.8
β Strands and loops†	2–86	85	2.1
α Helices	90–105, 113–135	39	2.1
β Strands, loops† and α helices	2–105, 113–136	128	2.4

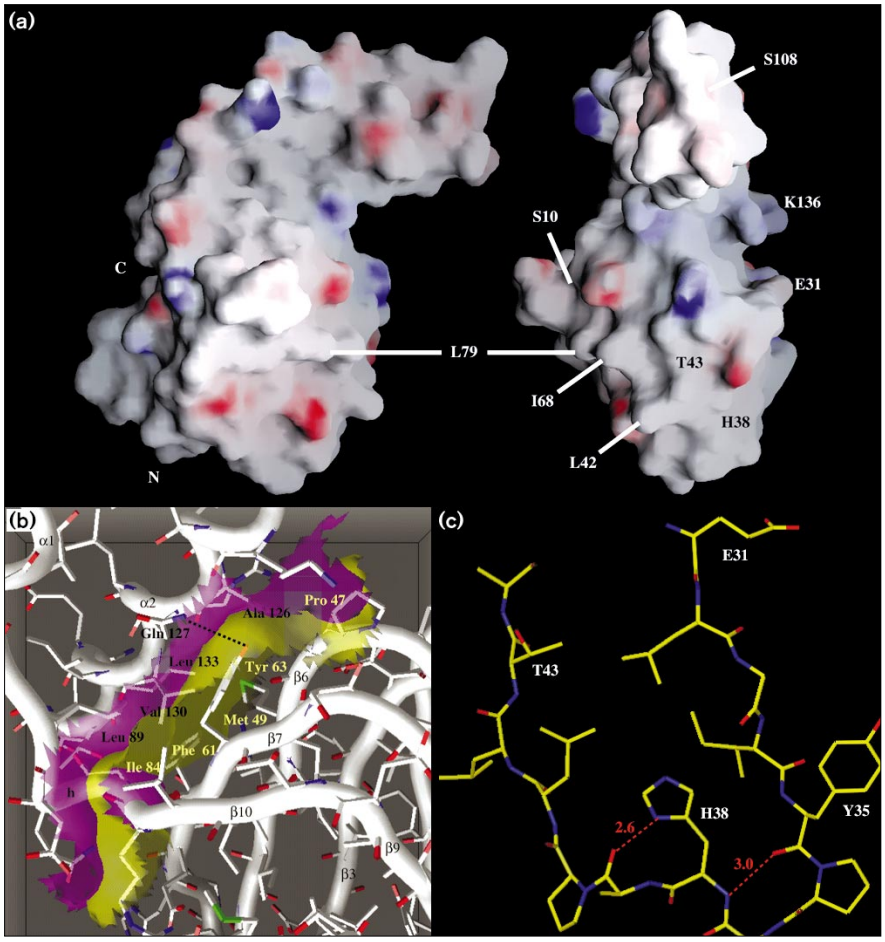
*The X-ray and NMR structures were compared by optimally superimposing equivalent Cα atoms and computing the root mean square (rms) residual. †The loop between the α helices is not included.

and Leu133 (Figure 3). Some of these residues form part of the interdomain interface while others are exposed to the solvent. Few of these residues are conserved.

Interaction between the domains

An outline view of the molecule (Figure 1a) gives the false impression that there is little contact between the two domains resulting in a flexible linker region. On the contrary, the region between the domains together with the linking 3_{10} helix form a closely packed hydrophobic core (Figures 2a,b). Residues contributing to this hydrophobic core include Pro47, Met49 , Phe61, Tyr63 and Ile84 (from the β sandwich) and Leu89, Ala126, Val130, and Leu133 (from the C-terminal α helix). There are no hydrogen bonds or salt bridges further stabilising the contact in this solvent-inaccessible region. Tyr63, which is on strand β7, lies in the crook between the domains with its hydroxyl group pointing towards the solvent (Figure 2b), and forms a hydrogen bond with the sidechain of residue Gln127. In the three proteins which have a tyrosine at position 63, there is always a potential hydrogen-bonding partner at position 127. Other species have a hydrophobic residue at position 63. This apparently stable interface which is preserved in solution and in the tightly packed crystal structure leaves room for a rotational motion of the two domains,

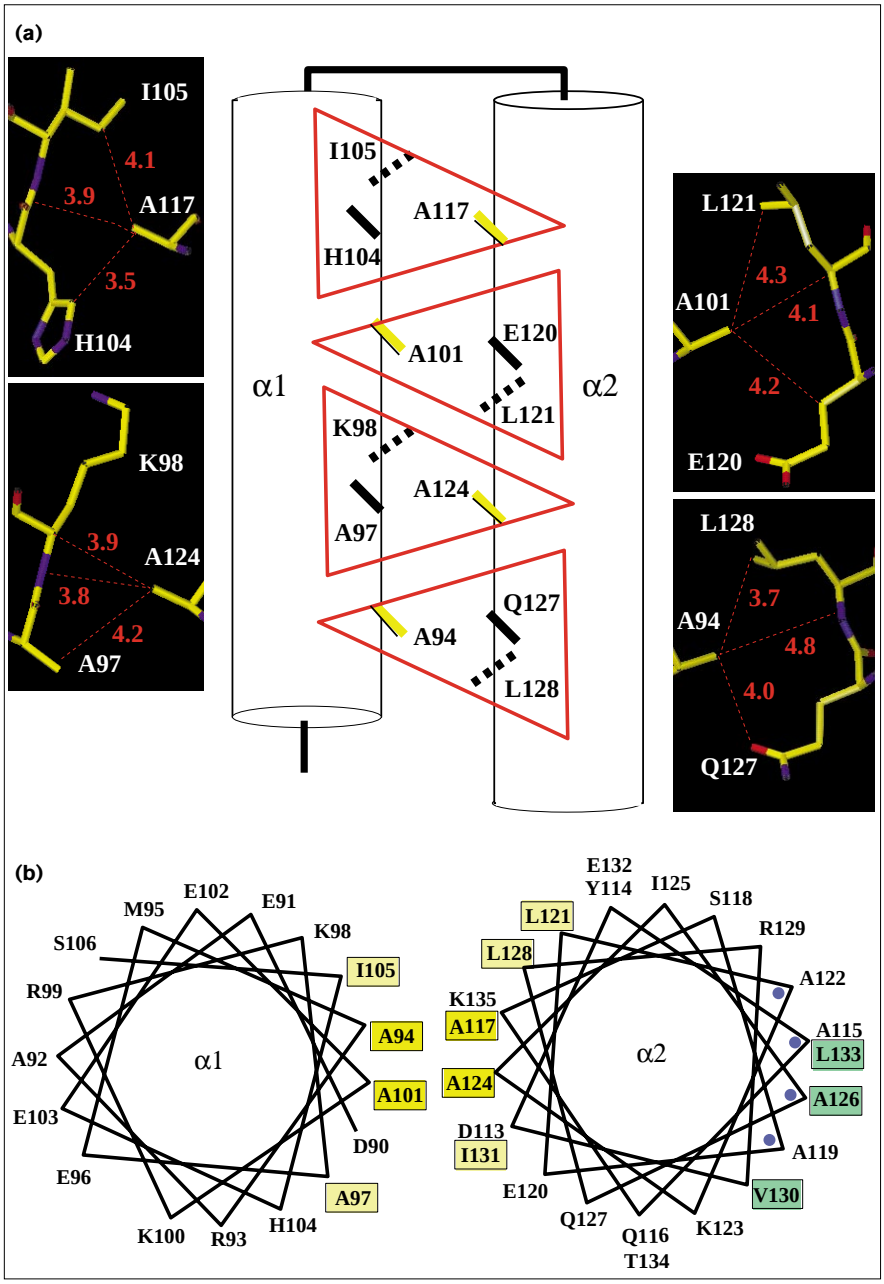
Figure 2



Some important surfaces of the *E. coli* ATP synthase ε subunit. **(a)** Two othogonal views showing the surface charge distribution and the interface between the two domains; positively charged regions are shown in blue, negatively charged regions in red. Leu42, Ile68 and Leu79 form a hydrophobic crystal contact, which is proposed to mimic the interaction with the γ subunit. (The figure was prepared using GRASP [43].) **(b)** The interface between the two domains. The van der Waals surface of the isolated N- and C-terminal domains are shown in yellow and magenta, respectively. **(c)** Interactions involving the highly conserved area around residue His38 at one end of the β sandwich. Hydrogen bonds are shown as red dashed lines; the atoms are shown in standard colours.

Figure 3

Interactions of the 'alanine zipper' in the helical domain. **(a)** The central schematic panel highlights the alanine residues that form the zipper in yellow with black labels; other residues are shown in black. The outer panels show the close hydrophobic contacts (red dotted lines) to each of the alanine methyl groups; distances are given in Å. **(b)** Helical wheels of the coiled-coil. Interacting hydrophobic residues are shown in light yellow with alanines in a darker shade. Hydrophobic residues which interact with the β domain are highlighted in green. Blue circles indicate alanine residues which form a line on the helix, facing the β domain.



relative to each other, resulting in a sliding motion of the two surfaces of the interface.

Crystal packing interactions

This unusually flat molecule packs in the $P6_522$ space group with a single copy in the asymmetric unit, making direct contact (< 3.5 Å) with ten other symmetry-related molecules in the unit cell. These contacts include potential hydrogen bonds (< 2.9 Å): Ser10 O–Arg93 NH1 (y–x,y,3/2–z); Glu70 OE2–Gln72 NE2 (1–y,1–x,7/6+z); Tyr114 OH–Ser118 O (x,1+x–y,5/6–z); and the reverse contacts. In addition, there is a contact across a crystallographic twofold axis that brings Glu70 into close proximity to the same residue in the symmetry-related molecule, Glu70 OE2–Glu70 OE2 (1–y,1–x,7/6+z) 2.5 Å. The orientation of this sidechain was checked in a simulated annealing 'omit' electron-density map. The proximity of these carboxylate groups implies that they share at least one proton. This contact area also contains the hydrophobic residues Leu42, Ile68 and Leu79 which have been

Figure 4

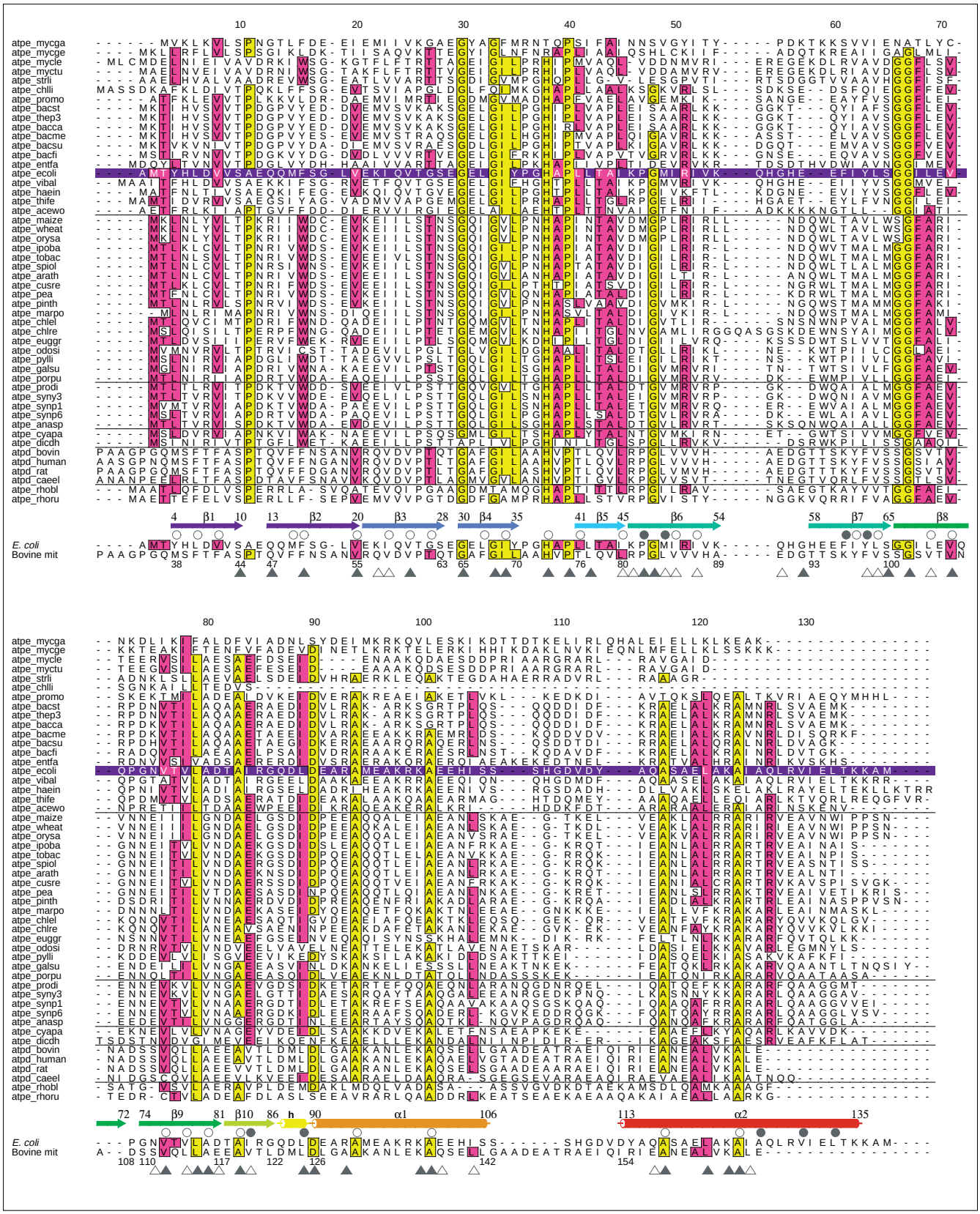

atpe_mycga
atpe_mycge
atpe_mycle
atpe_myctu
atpe_strli
atpe_chlli
atpe_promo
atpe_bacst
atpe_thep3
atpe_bacca
atpe_bacme
atpe_bacsu
atpe_bacfi
atpe_entfa
atpe_ecoli
atpe_vibal
atpe_haein
atpe_thife
atpe_acewo
atpe_maize
atpe_wheat
atpe_orysa
atpe_ipoba
atpe_tobac
atpe_spiol
atpe_arath
atpe_cusre
atpe_pea
atpe_pinth
atpe_marpo
atpe_chlel
atpe_chlre
atpe_euggr
atpe_odosi
atpe_pylli
atpe_galsu
atpe_porpu
atpd_prodi
atpd_syny3
atpe_synp1
atpe_synp6
atpe_anasp
atpe_cyapa
atpe_dicdh
atpd_bovin
atpd_human
atpd_rat
atpd_caeel
atpe_rhobl
atpe_rhoru
E. coli
Bovine mit

Figure 4 continued

Sequence alignment of the ε subunits from bacteria and chloroplasts, and δ subunits from mitochondria. The Swiss Prot [44] accession codes are employed for naming the proteins. Yellow highlights positions with greater than 80% identity and magenta highlights positions with between 50 and 80% identity. The alignment is numbered according to the *E. coli* sequence which is highlighted in purple. In addition to their listing in the main body of the figure, the *E. coli* ε and bovine δ mitochondrial sequences are repeated at the bottom, together with the secondary structure elements identified in the *E. coli* structure. Cylinders represent α helices and arrows β sheets. Amino acids marked with open circles under the β strands have their sidechains pointing into the interior of the β sandwich and similarly open circles under the α helices mark residues buried between the two helices. Filled circles label amino acids which have their sidechains between the two domains. Triangles at the bottom indicate similarities between the *E. coli* and bovine sequences: filled triangles indicate identity and non-filled similarity. (The figure was prepared using the program ALSCRIPT [45].)

proposed to interact with the γ subunit (see below). Residues 68 and 79 are highly conserved in the ε family of proteins (Figure 4). Although the contact in the crystal is not a particularly intimate one, the molecules are mostly close enough to exclude water from the interface. The importance of the close crystal packing involving both domains of the molecule is that the overall structure, relating the domains, is the same as defined by spin labelling and NOE NMR experiments in solution [22], suggesting that the molecule is reasonably rigid.

Discussion

The ε subunit carries out its biological role by interacting with other subunits of the F_0F_1-ATP synthase [13,26]. There is a great deal of evidence from mutation and cross-linking studies, which define both the importance of particular residues in the ε subunit and the importance of residues in other subunits with which the ε subunit interacts. We can now attempt to rationalise some of these observations in terms of the structure.

Sequence alignment and model of the bovine mitochondrial ε subunit

In order to define whether or not the δ subunits of mitochondrial ATPases are equivalent to the ε subunit of the *E. coli* protein, we have first used the program BLAST [27] to align 50 sequences from a wide variety of organisms (Figure 4). There are no fully conserved residues among all ε subunit sequences and the most conserved residues are evenly spread across the alignment. Of these, many are prolines and glycines which lie between and in the heavily bent β strands of the *E. coli* structure suggesting that all the proteins have a similar fold. The most conserved sequence forms part of strand β4 and the connecting loop to strand β5. Residues in this region of the structure have been shown to bind to the F_0 c subunit [17]. Although the more conserved areas of the sequence are in the N-terminal domain, there are a number of conserved alanines in the C-terminal domain, which we have shown form the zipper contact between the helices. The highly conserved residue Asp90 at the beginning of α1 forms the N-terminal cap of the helix.

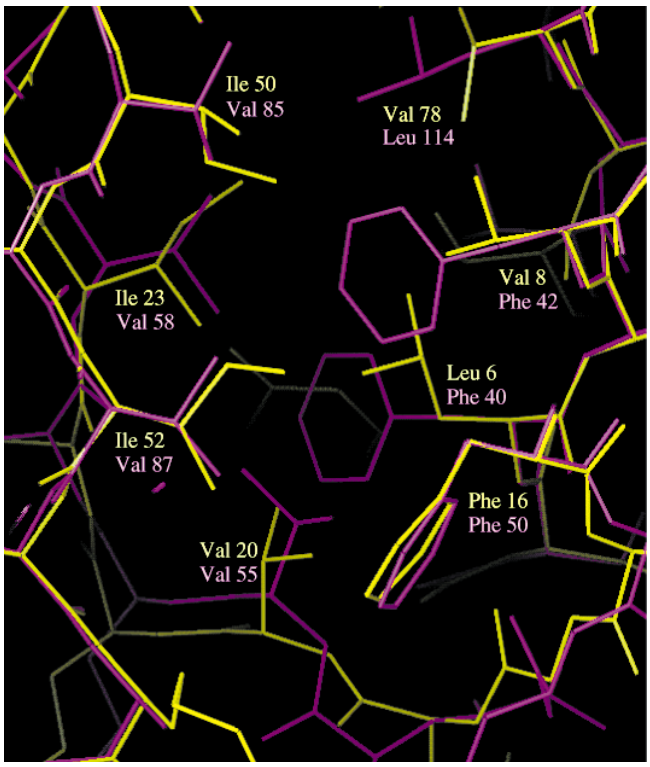
We constructed an homology model of the bovine mitochondrial δ subunit structure, which is equivalent to the *E. coli* ε subunit, using the program O [28] according to the alignment shown at the bottom of Figure 4. This alignment was slightly modified from the BLAST sequence alignment to account for some notable structural features in ε. The multiple sequence alignment left two gaps between strands β6 and β7 in the alignment of the *E. coli* and bovine sequences. These gaps were closed for the model building, giving one additional identical and and one conservative substitution between the two sequences. The internal hydrophobic packing in the β sandwich is preserved by complementary changes in the sizes of amino acids in opposing β sheets. Two insertions at the bottom of the β sandwich, an alanine residue between Gly18 and Leu19 in strand β2 and an alanine at the end of β8, slightly open the barrel allowing for several additional aromatic residues in the bovine structure (Figures 4 and 5).

In the C-terminal domain the overall alignment in the loop between the helices is uncertain. The multiple sequence alignment predicts that the loop in the bovine structure is considerably longer and the last helix much shorter. This difference can be reduced by shuffling the sequences from the right to fill the last gap, however, this results in the loss of the alanine zipper between the helices. Our preferred alignment, shown at the bottom of Figure 4, amalgamates the two gaps in the *E. coli* sequence into one by a shift to fit the secondary structure. The shorter terminal helix in the bovine structure would reduce the tight packing between the domains.

Mutations affecting the ε subunit

Kuki *et al.* [29] constructed a series of C-terminal deletion mutants and found that function was retained with an ε subunit lacking the C-terminal 60 amino acids. This is consistent with the fact that the ε subunit from the green bacterium *Chlorobium limicola* (Swiss Prot. accession number ATPE_CHLLI) naturally lacks the equivalent of these residues (Figure 4). Such a deletion includes all of the C-terminal helical domain and extends into the N-terminal β sandwich domain. The construct containing residues 1–78 forms a functional enzyme, while the construct containing only residues 1–73 does not [29]. The removal of the additional five residues causes the loss of strand β9 in the middle of the molecule, and this together with the loss of strand β10 probably prevents the molecule from folding due to loss of interactions with strands β1 and β8. Using a similar approach, Janouchi *et al.*, [30] found that significant activity could be retained with the deletion of 15 N-terminal amino acids from the ε subunit, but activity

Figure 5



Contacts in the centre of the N-terminal domain showing the conservation of packing in the *E. coli* ε subunit (yellow) and bovine mitochondrial δ subunit (magenta).

was lost with the deletion of 16 residues. In the construct containing residues 16–138, Phe16 could be important for the stabilisation of strand β2 in the absence of strand β1. We would conclude that removal of residues 1–16 also results in improper folding of the ε subunit.

Missense mutations affecting two residues of the ε subunit have been described that significantly affect F_0F_1-ATPase activity. A Gly48→Asp mutation affected assembly of the enzyme, a defect that could be overcome by increasing the gene dosage [31]. The fully assembled enzyme, however, was unable to carry out ATP-dependent proton translocation and ATPase activity was about 30% of the wild-type enzyme. The mutant characteristics were reversed by a second mutation Pro47→Ser or Pro47→Thr (Figure 1a). Gly48 is highly, but not strictly, conserved (Figure 4). It occurs in a bulge in strand β6 which interrupts the regular hydrogen bonding to strand β3. The backbone torsion angles for Gly48 ($\phi = -180°$; $\psi = -140°$) lie in a glycine-only allowed region of the Ramachandran plot, implying that mutation to any other residue would result in a change in the backbone conformation. Furthermore, the hydrogen atom of Gly48, which would be replaced by a sidechain if this residue was mutated, is closely packed against the sidechain of Thr26. A second mutation of Pro47 may be able, by giving the polypeptide backbone greater flexibility, to partially reverse the conformational change when Gly48 is mutated. In those proteins which do not have a glycine residue at position 48 there is never a proline residue at position 47.

Mutations affecting Glu31 and His38 have been investigated by LaRoe and Vik [32]. His38 is highly conserved in all ε subunits and Glu31 interacts with the c subunits of F_0 (see below). As judged by growth tests, 12 amino acid substitutions constructed at position 31 had no significant effect on ATP synthesis. In the same way, 12 amino acid substitutions were constructed at position 38 with the mutation His38→Pro preventing growth on succinate media and causing poor growth on glucose media. Substitution of residue 38 by arginine had a partial effect. Residue 38 makes a mainchain–mainchain hydrogen bond to residue 35 forming a type II β turn. This turn and the following loop are well defined in the X-ray structure but were poorly defined in the NMR structure [22]. The loop contains several of the most conserved residues, including His38 itself (Figures 2 and 4).

Interactions between ε and the F_1 γ and F_0 c subunits

A mutant of the ε subunit, His38→Cys, was constructed and cross-linked to the γ subunit using either heterobifunctional cross-linkers [14,15] or by disulphide formation with a Tyr205→Cys mutant of the γ subunit [16]. Aggeler *et al.* [33] demonstrated that a cysteine residue at position 38 was able to react with maleimides in a purified F_1-ATPase but was unable to react in a purified F_1F_0-ATPase. Furthermore, coupled functions were affected when modified F_1-ATPase was reconstituted with the F_0 portion. Zhang and Fillingame [17] observed disulphide bridge formation between subunit *c* and the ε subunit in double mutant strains carrying the ε mutation Glu31→Cys and any of three c subunit mutants, Ala40→Cys, Gln42→Cys or Pro43→Cys. Both residues 31 and 38 occur at one end of the β-sandwich domain (Figures 2a,c) and the available evidence (outlined above) indicates that this region interacts with the F_0 portion. The N-terminal domain of the ε subunit also clearly interacts with the γ subunit with cross-links obtained between a cysteine residue placed at position 10 of the ε subunit and position 228 of the γ subunit [34]. Cysteines at positions 38 or 43 formed disulphide bridges with a cysteine residue placed at position 205 of the γ subunit [16].

From the NMR structure, Wilkens *et al.* [22] suggested that the hydrophobic patch Val9, Met15, Leu42, Ile68, Thr77 and Leu79 may also interact with the γ subunit (Figures 2 and 6). In the X-ray structure the sidechain of Met15 points in the opposite direction to the other sidechains forming the patch, due to a bulge in the β strand, and is not actually part of the hydrophobic patch.

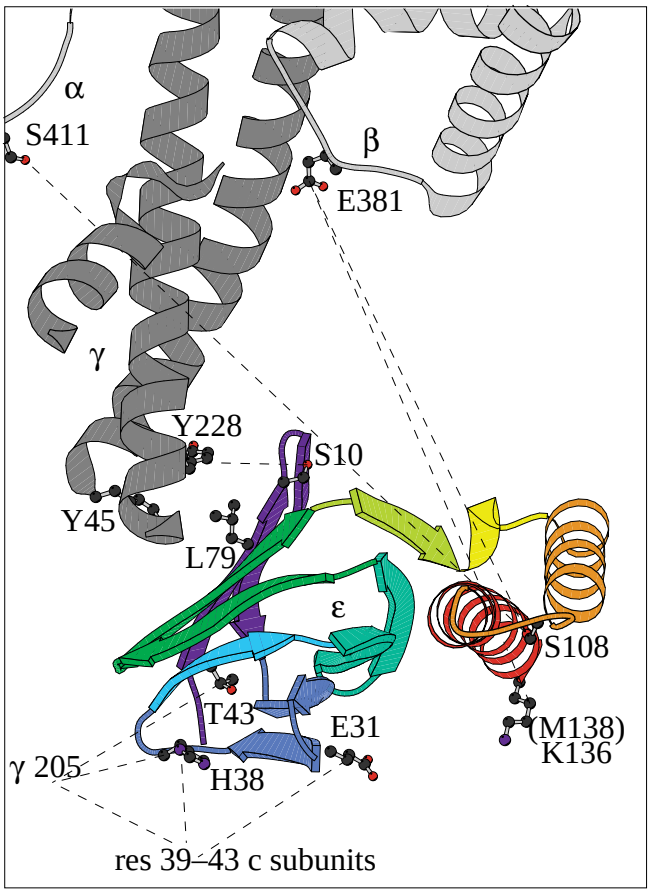
Interaction of the ε subunit with the α and β subunits

The C-terminal domain of the ε subunit has been shown to be located in close proximity to the α and β subunits of the F_1-ATPase. Dallmann *et al.* [18] described a cross-link between the β subunit and the ε subunit that was induced by a water-soluble carbodiimide; an ester linkage was found between a glutamate residue at position 381 in the β subunit and a serine residue at position 108 in the ε subunit. The closeness of these two residues was confirmed by disulphide formation between a cysteine inserted at position 381 of the β subunit and a cysteine inserted at position 108 of the ε subunit [19]. Residue 411 of the α subunit is the structural equivalent of residue 381 of the β subunit. Aggeler *et al.* [20] demonstrated disulphide formation between a cysteine placed at position 411 of the α subunit and a cysteine placed at position 108 of the ε subunit. A cysteine at position 108 of ε formed a disulphide with either cysteine at 411 in the α subunit, in the presence of added ATP, or cysteine 381 in the β subunit in the presence of added ADP. However, α–β cross-links were detected at high levels when the double mutant α Cys411, β Cys381 was constructed. This observation suggests that in these experiments the α and β subunits may move with respect to each other, rather than there being simply a relative movement of the ε subunit. Formation of a β–ε–β cross-linked product has also been reported [21] with one linkage occurring between ε Cys108 and β Cys381 with the second β cross-linked via ε Cys138.

Modelling the binding of the structure to the F_0F_1 complex

We have tried to identify a plausible docking arrangement of the ε subunit within the F_0F_1 complex, which satisfies the above mentioned studies, using the crystal structure of the mitochondrial F_1 complex (MF_1) [6]. Residue 10 in the ε subunit can be cross-linked to Tyr228 in the γ subunit [35]; the corresponding residue in the MF_1 γ subunit is Tyr214. This residue can be seen in the structure of MF_1, but none of the other corresponding residues in the area identified to interact with the ε subunit has been positioned. Residue Tyr205→Cys of the *E. coli* F_1 γ subunit, which can form a disulphide bridge both with residues in the ε subunit (residues His38 and 43) and the c subunit (residues 39, 42 and 43), corresponds to a residue around 190 in the MF_1. This residue cannot be seen in the MF_1 structure. We fitted our structure so as to form a reasonable packing arrangement relative to the observed residues in this area, with residue 10 of the ε subunit at an appropriate distance from Tyr214 of the γ subunit (Figure 6). In this position the highly conserved residue Leu79 of the ε subunit stacks onto the conserved Tyr44 and Tyr214 of the MF_1 γ subunit (Tyr45 and Tyr228 in *E. coli*). There is apparently no way to maintain this interaction between the ε and γ subunits and to simultaneously position the molecule so as to satisfy the binding of the C-terminal helix of the ε subunit to the α and β subunits.

Figure 6



Residues in the *E. coli* F_1-ATPase ε subunit that are known from cross-linking studies to interact with residues on the other F_0F_1-ATPase subunits. The structure of bovine mitochondrial F_1-ATPase was used to model the locations of the interacting residues in the *E. coli* protein. The ε subunit is shown in colour and other subunits are in grey with sidechains in standard colours. The sidechains of residues involved in interactions are shown in ball-and-stick representation; interactions are shown as dashed lines.

Assuming the ε subunit retains the crystal structure in the holoenzyme, if ε-residue 31 cross-links with residues of the c subunit and ε residue 38 is also positioned close to the membrane, then at least in the *E. coli* [21] enzyme the α and β subunits may in fact extend to within about 1 nm of the membrane surface, as dictated by the overall dimensions of the ε subunit. The distance between the α carbon atoms of residues 108 and 136 (the last residue defined by the ε crystal structure) is within a few Ångströms of the largest distance between the α carbon atoms of residues 381 on two β residues, as defined by the crystal structure of MF_1. The formation of β–ε–β cross-links involving residues 108 and 138 of an ε subunit and position 381 on two β subunits is therefore compatible with both crystal structures, providing that a steric clash with the γ subunit is avoided.

Large conformational changes, such as the pH-driven event observed in haemagglutinin and the protease clipping in serpins cannot, however, be ruled out for the ε subunit as it changes its interactions with the other subunits during the catalytic cycle. An essential role for a relative motion of the two domains of the ε subunit in transmitting the force derived from proton translocation is unlikely given the occurrence of at least one species in which the ε subunit has only one domain.

Biological implications

Adenosine triphosphate (ATP), which occurs in all known life forms, is the 'high-energy' intermediate that constitutes the most common cellular energy currency. The enzyme F_0F_1-ATP synthase, which generates ATP from ADP, is energised by a transmembrane proton gradient that is created by photosynthesis or respiration. The proton-translocating ATP synthases in mitochondria, chloroplasts and bacteria have a common architecture comprising two portions, F_1 and F_0. The F_0 portion is an integral membrane protein consisting of three or more different polypeptides. The F_1 portion can be isolated as a soluble fraction which acts as an ATPase, converting ATP to ADP, and has the subunit composition $\alpha_3\beta_3\gamma\delta\varepsilon$. There are three catalytic sites within F_1 which are located in the β subunits, as revealed by the structure of the bovine mitochondrial enzyme.

The coupling of proton translocation through F_0 to ATP synthesis in F_1 has been proposed to proceed via the 'binding change' mechanism, in which each of the three chemically identical active sites is held in a different conformational state by their asymmetric interaction with the single-copy subunits, γ and ε. The crystal structure of the bovine enzyme trapped each of the catalytic sites in a different liganded state, and each of the β subunits made a different interaction with the limited portions of the γ subunit (~45%) visible in the crystal structure. This observation provided support for an earlier hypothesis: that components of the two portions of ATP synthase actually rotate with respect to each other during the catalytic cycle. Recent biophysical experiments have now observed this physical rotation.

What is the role of the ε subunit in ATP synthase? Absence of the ε subunit prevents the binding of F_0 to F_1 and is thus essential for the coupling of proton translocation to ATP synthesis. The ε subunit is also known to inhibit ATPase activity in isolated F_1-ATPase. We report here the crystal structure of the ε subunit of *Escherichia coli* ATP synthase. The molecule is comprised of two domains: the N-terminal β sandwich domain, which has a novel topology with two five-stranded β sheets; and the C-terminal domain, which contains two antiparallel α helices with an interface between them described as an 'alanine zipper'. Each of the individual domains closely resembles an earlier description of the solution structure of the ε subunit determined using NMR spectroscopy. An apparently stable interface between the domains makes it less likely that large conformational changes in the ε subunit accompany the relative rotation of the subunits proposed to occur during ATP synthesis. The still largely unanswered question is how proton translocation through the membrane is coupled to the suggested relative motions of the catalytic and membrane bound subunits.

Materials and methods

Data collection and processing

The expression, purification and crystallisation of the A1G, M2S protein has been described [36]. The crystals were grown by the vapour diffusion method in hanging drops. The reservoir contained 2 M K_2HPO_4/$NaH_2PO_4{\cdot}2H_2O$, 100 mM Hepes buffer, (pH 7.5), and 200 mM NH_2SO_4. The hanging drop (5 to 20 μl) was composed of a 1:1 mixture of a 4 mg/ml protein solution and the reservoir solution. Despite repeated attempts over a wide range of conditions, the original problem which caused crystallisation to take a very long time, typically between nine and 12 months, was never overcome. All diffraction data were recorded on an R-Axis II imaging plate detector mounted on a Rigaku RU-200 rotating-anode generator using mirror optics (Z Otwinowski and G Johnson, Yale University, as marketed by Molecular Structure Corporation, Texas, USA) and a 0.3 mm focal spot. The data were integrated and scaled with DENZO and SCALEPACK [37]. Data for the native protein were recorded from a single crystal to a resolution of 2.3 Å (Table 2).

Phase determination and structure refinement

Initial attempts to solve the structure by molecular replacement (MR) using the NMR structure as a search model were unsuccessful. A number of different search models were tried, including the average structure, a polyalanine and a polyserine structure. Attempts were also made with the individual domains and with both domains using different interdomain angles. Rotation and translation searches were made using both AMORE [38] and X-PLOR [39]. The failure to clearly identify the correct solution probably results from the differences between the NMR structure for the complete molecule and the X-ray structure and the fact that too small a fraction of the molecule was present when searching using only individual domains.

The structure was eventually solved by conventional multiple isomorphous replacement (MIR) using a total of two different heavy-atom compounds which yielded a total of four derivative data sets. Derivatives were prepared by soaking crystals in a stabilising solution (equivalent to the well solution over which the hanging drops had been placed), with the heavy atom compound added over several hours until the final concentration had been reached (Table 2). Given the very small number of crystals that were available the first attempts to prepare derivatives were made using the same crystal that had been used to record the native data. The crystal was subjected to rounds of heavy atom soaking, data recording and back-soaking in heavy atom free buffer. The reagents used were 1 mM K_2PtCl_4, 1.25 mM methylmercury chloride and 1.25 mM phenylmercury acetate. Only the platinum compound yielded useful phasing information. The methylmercury chloride data were used as a second platinum derivative when it became clear that the platinum had not been removed from the crystal by the back-soaking or that the mercury had bound at the same location with lower occupancy. Further heavy-atom data were recorded from a second crystal some years later. This crystal was soaked in 2.5 mM *p*-chloromercurybenzene sulfonic acid (PCMBS). Two data sets were recorded from the same crystal before and after a breakdown of the X-ray diffraction equipment. The data sets were treated independently for the phase calculations. Thus each of the two 'independent' derivatives were each represented by two data sets in the phase calculations. We note that this is a formally incorrect procedure which leads to errors in the phase probability distribution. It was,

Table 2

Data collection and structure solution statistics.

Crystal / derivative*	Native	Pt1	Pt2	Hg1	Hg2
Data					
Cell dimensions a,c (Å)	94.6, 56.9	93.6, 57.3	93.7, 57.1	93.8, 56.9	93.8, 56.9
Resolution (Å)	2.3	3.0	3.0	2.8	3.1
No. of observations	19 833	11 915	10 433	5873	13 331
No. of unique reflections	6626	2768	2679	3343	2825
Completeness (%)					
total / last shell (all data)	94 / 89	86 / 80	83 / 71	85 / 82	96 / 96
R_{merge} †					
overall / last shell (all data)	5.9 / 47.5	6.9 / 35.3	9.6 / 20.4	6.2 / 14.9	8.5 / 17.0
MIR structure solution					
No. of sites	–	2	2	2	2
Phasing power acentric / centric‡	–	1.0 / 0.9	1.0 / 0.8	2.0 / 1.3	2.0 / 1.2
R_{deriv} §	–	26.5	26.5	25.2	24.1

*Pt1 is the potassium tetrachloroplatinum(II), K_2PtCl_4 derivative; Pt2 is the Pt1 derivative after back-soaking in heavy atom free buffer then soaking in methylmercury chloride (see text); Hg1 and Hg2 are both *p*-chloromercurybenzene sulfonic acid (PCMBS) derivatives from the same crystal before and after a breakdown in the X-ray equipment.

†$R_{merge} = \Sigma_h\Sigma_i|I_{(h,i)} - \langle I_{(h)}\rangle|/\Sigma_h\Sigma_i\langle I_{(h)}\rangle$, where $I_{(h,i)}$ is the scaled intensity of the ith observation of the reflection h and $\langle I_{(h)}\rangle$ is the mean value. ‡Phasing power $= \Sigma|F_H|/\Sigma||F_{PHobs}| - |F_{PHcalc}||$. §$R_{deriv} = \Sigma||F_{PH}| - |F_P||/\Sigma F_P|$, where $|F_{PH}|$ and $|F_P|$ are the structure-factor amplitudes of the derivative and the native crystals, respectively.

however, the procedure which produced an interpretable MIR electron-density map. The native and derivative data were merged and scaled using the programs of the CCP4 suite [38]. The platinum sites were located using RSPS and were further refined using MLPHARE [38]. The mercury site in the PCMBS derivative was located in a difference Fourier synthesis with SIR phases from the $K_2Pt(Cl)_4$ derivative. Data collection and scaling for the native and derivative data are summarised in Table 2. The MIR phases were extended and refined by solvent flattening and histogram matching using DM [38]. The correct enantiomorph of the space group $P6_522$ was chosen on the basis of positive refined anomalous occupancies for the heavy atoms and the correct handedness of helices identified in the MIR electron-density map. Model building into electron-density maps was carried out with the program O [28]. Knowledge of the NMR structure was especially helpful in tracing the chain in the β-sheet domain. The structure was extensively checked throughout with omit electron density difference maps. Structure refinement was performed using simulated annealing with X-PLOR [39] and finally with REFMAC [38]. Other computer programs used in the structure solution and analysis included, SIGMAA [38] and MOLEMAN [40].

Table 3

Refinement summary.

Resolution range (Å)	19.0–2.3
No. of reflections*	6134
Number of nonhydrogen protein atoms	1027
Number of ordered water molecules	14
R value†	21.4
Free R value‡	28.8
Deviations from dictionary values§	
bond lengths (Å)	0.01
bond angles (°)	1.5
Thermal parameter restraints#	
$\langle\Delta B\rangle$ mainchain atoms (Å²)	2.1
$\langle\Delta B\rangle$ sidechain atoms (Å²)	5.2
Average thermal parameters	
$\langle B\rangle$ mainchain atoms (Å²)	38
$\langle B\rangle$ sidechain atoms (Å²)	44
Ramachandran plot quality	
residues in core regions (%)	89.7
residues in allowed regions (%)	9.4
residues in generously allowed regions (%)	0.9
residues in disallowed regions (%)	0.0

*Reflections with $I \leq 0$ were excluded from the refinement. †R value $= \Sigma||F_{OBS}| - |F_{CALC}||/\Sigma|F_{OBS}|$, where F_{OBS} and F_{CALC} are the observed and calculated structure factors, respectively. ‡The free R value was calculated using 8% of the data chosen randomly and omitted from the refinement. §The root mean square deviation (rmsd) relates to the Engh and Huber parameters [41]. #The rmsd refers to bonded atoms.

Some areas of the refined structure are in poorly defined electron density and the atoms in the model have correspondingly high B factor values (Figure 1b). There is no or very weak electron density in omit maps for residues 29, 55, 56 and 109–112. Details of the refinement are given in Table 3.

Accession numbers

The atomic coordinates and the structure-factor amplitudes for the structure of the ε subunit of *E. coli* ATPase have been deposited with the Brookhaven Protein Data Bank (accession number 1aqt).

Acknowledgements

Earlier work in this project was support by National Health and Medical Research Council (NHMRC) and Australian Research Council (ARC) grants to JMG and GC. UU is the recipient of a Fellowship from the Swedish Medical Research Council. The coordinates for each of the domains separately and for the overall subunit as defined by NMR were kindly provided by Stephan Wilkens prior to our successful structure determination.

References

1. Senior, A.E. (1990). The proton-translocating ATPase of *Escherichia coli*. *Annu. Rev. Biophys. Biophys. Chem.* **19**, 7–41.
2. Fillingame, R.H. (1990). Molecular mechanics of ATP synthesis for F_1F_0-type H^+-translocating ATP synthases. In *The Bacteria: a Treatise on Structure and Function*. (Krulwich, T.A., ed.), pp. 345–391, Academic Press, Inc, New York, NY, USA.
3. Boyer, P.D. (1993). The binding change mechanism for ATP synthase – some probabilities and possibilities. *Biochim. Biophys. Acta* **1140**, 215–250.

4. Pedersen, P.L. (1996). Frontiers in ATP synthase research: understanding the relationship between subunit movements and ATP synthesis. *J. Bioenerg. Biomembr.* **28**, 389–395.
5. Bianchet, M., Ysern, X., Hullihen, J., Pedersen, P.L. & Amzel, L.M. (1991). Mitochondrial ATP synthase – quaternary structure of the F_1 moiety at 3.6 Å determined by X-ray diffraction analysis. *J. Biol. Chem.* **266**, 21197–21201.
6. Abrahams, J.P., Leslie, A.G.W., Lutter, R. & Walker, J.E. (1994). Structure at 2.8 Å resolution of F_1-ATPase from bovine heart mitochondria. *Nature* **370**, 621–628.
7. Shirakihara, Y., *et al*., & Yoshida, M. (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**, 825–836.
8. Boyer, P.D. (1989). A perspective of the binding change mechanism for ATP synthesis. *FASEB J* **3**, 2164–2178.
9. Duncan, T.M., Bulygin, V.V., Zhou, Y., Hutcheon, M.L. & Cross, R.L. (1995). Rotation of subunits during catalysis by *Escherichia coli* F1-ATPase. *Proc. Natl. Acad. Sci. USA* **92**, 10964–10968.
10. Noji, H., Yasuda, R., Yoshida, M. & Kinosita, K. (1997). Direct observation of the rotation of F_1-ATPase. *Nature* **386**, 299–302.
11. Sternweis, P.C. & Smith, J.B. (1980). Characterization of the inhibitory (ε) subunit of the proton-translocating adenosine triphosphatase from *Escherichia coli*. *Biochemistry* **19**, 526–531.
12. Mendel-Hartvig, J. & Capaldi, R.A. (1991). Catalytic site nucleotide and inorganic phosphate dependence of the conformation of the ε subunit in *Escherichia coli* adenosinetriphosphatase. *Biochemistry* **30**, 1278–1284.
13. Sternweis, P.C. (1978). The ε subunit of *Escherichia coli* coupling factor 1 is required for its binding to the cytoplasmic domain. *J. Biol. Chem.* **253**, 3123–3128.
14. Skakoon, E.N. & Dunn, S.D. (1993). Orientation of the ε subunit in *Escherichia coli* ATP synthase. *Arch. Biochem. Biophys.* **302**, 279–284.
15. Skakoon, E.N. & Dunn, S.D. (1993). Location of conserved residue histidine-38 of the ε subunit of *Escherichia coli* ATP synthase. *Arch. Biochem. Biophys.* **302**, 272–278.
16. Watts, S.D., Tang, C. & Capaldi, R.A. (1996). The stalk region of the *Escherichia coli* ATP synthase. Tyrosine 205 of the γ subunit is in the interface between the F_1 and F_0 parts and can interact with both the ε and c oligomer. *J. Biol. Chem.* **271**, 28341–28347.
17. Zhang, Y. & Fillingame, R.H. (1995). Subunits coupling H^+ transport and ATP synthesis in the *Escherichia coli* ATP synthase. Cys–Cys cross-linking of F_1 subunit epsilon to the polar loop of F_0 subunit c. *J. Biol. Chem.* **270**, 24609–24619.
18. Dallmann, H.G., Flynn, T.G. & Dunn, S.D. (1992). Determination of the 1-ethyl-3-[(3-dimethylamino)propyl]-carbodiimide-induced cross-link between the β and ε subunits of *Escherichia coli* F_1-ATPase. *J. Biol. Chem.* **267**, 18953–18960.
19. Aggeler, R., Haughton, M.A. & Capaldi, R.A. (1995). Disulphide bond formation between the COOH-terminal domain of the β subunits and the γ and ε subunits of the *Escherichia coli* F_1-ATPase. Structural implications and functional consequences. *J. Biol. Chem.* **270**, 9185–9191.
20. Aggeler, R. & Capaldi, R.A. (1996). Nucleotide-dependent movement of the epsilon subunit between α and β subunits in the *Escherichia coli* F_1F_0-type ATPase. *J. Biol. Chem.* **271**, 13888–13891.
21. Capaldi, R.A., Aggeler, R., Wilkens, S. & Gruber, G. (1996). Structural changes in the gamma and epsilon subunits of the *Escherichia coli* F_1F_0-type ATPase during energy coupling. *J. Bioenerg. Biomembr.* **28**, 397–401.
22. Wilkens, S., Dahlquist, F.W., McIntosh, L.P., Donaldson, L.W. & Capaldi, R.A. (1995). Structural features of the ε subunit of the *Escherichia coli* ATP synthase determined by NMR spectroscopy. *Nat. Struct. Biol.* **2**, 961–967.
23. Wilkens, S., Dunn, S.D., Chandler, J., Dahlquist, F.W. & Capaldi, R.A. (1997). Solution structure of the N-terminal domain of the δ subunit of the *E. coli* ATP synthase. *Nat. Struct. Biol.* **4**, 198–201.
24. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matches. *J. Mol. Biol.* **233**, 123–138.
25. Sheriff, S., Hendrickson, W.A. & Smith, J.L. (1987). Structure of myohemerythrin in the azidomet state at 1.7/1.3 Å resolution. *J. Mol. Biol.* **197**, 273–296.
26. Smith, J.B. & Sternweis, P.C. (1977). Purification membrane attachment and inhibitory subunits of the proton translocating adenosine triphosphate from *Escherichia coli*. *Biochemistry* **16**, 306–311.
27. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.
28. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* **47**, 110–119.
29. Kuki, M., Noumi, T., Maeda, M., Amemura, A. & Futai, M. (1988). Functional domains of ε subunit of *Escherichia coli* H^+-ATPase (F_1F_0). *J. Biol. Chem.* **263**, 17437–17422.
30. Janouchi, M., Takeyama, M., Noumi, T., Moriyama, Y., Maeda, M. & Futai, M. (1992). Role of the amino terminal region of the ε subunit of *Escherichia coli* H^+-ATPase (F_0F_1). *Arch. Biochem. Biophys.* **292**, 87–94.
31. Cox, G.B., *et al*., & Gibson, F. (1987). Amino acid substitutions in the ε subunit fo the F_1F_0-ATPase of *Escherichia coli*. *Biochim. Biophys. Acta* **890**, 195–204.
32. LaRoe, D.J. & Vik, S.B. (1992). Mutations at Glu-32 and His-39 in the epsilon subunit of the *Escherichia coli* F_1F_0 ATP synthase affect its inhibitory properties. *J. Bacteriol.* **174**, 633–637.
33. Aggeler, R., Weinreich, F. & Capaldi, R.A. (1995). Arrangement of the ε subunit in the *Escherichia coli* ATP synthase from the reactivity of cysteine residues introduced at different positions in this subunit. *Biochim. Biophys. Acta* **1230**, 62–68.
34. Aggeler, R., Chicas-Cruz, K., Cai, S.X., Keana, J.F. & Capaldi, R.A. (1992). Introduction of reactive cysteine residues in the ε subunit of *Escherichia coli* F_1 ATPase, modification of these sites with tetrafluorophenyl azidemaleimides, and examination of changes in the binding of the ε subunit when different nucleotides are in catalytic sites. *Biochemistry* **31**, 2956–2961.
35. Tang, C. & Capaldi, R.A. (1996). Characterisation of the interface between γ and ε subunits of *Escherichia coli* F_1-ATPase. *J. Biol. Chem.* **271**, 3081–3024.
36. Codd, R., Cox, G.B., Guss, J.M., Solomon, R.G. & Webb, D. (1992). The expression, purification and crystallization of the ε subunit of the F_1 portion of the ATPase of *Escherichia coli*. *J. Mol. Biol.* **228**, 306–309.
37. Otwinowski, Z. (1993). Oscillation data reduction program. In *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*. (Sawyer, L., Isaacs, N. & Bailey, S.S., eds), pp. 56–62, SERC Daresbury Laboratory, Warrington, UK.
38. Collaborative Computational Project no. 4. (1994). *Acta Cryst. D* **50**, 760–763.
39. Brünger, A.T. (1992). *X-PLOR, Version 3.1, A System for Crystallography and NMR*. Yale University Press, New Haven, CT, USA.
40. Kleywegt, G.J. & Jones, T.A. (1996). Phi/Psi-chology: Ramachandran revisited. *Structure* **4**, 1395–1400.
41. Engh, R.A. & Huber, R. (1991). Accurate bond length and angle parameters for X-ray protein structure refinement. *Acta Cryst. A* **47**, 392–400.
42. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structure. *J. Appl. Cryst.* **24**, 946–950.
43. Nicholls, A., Bharadwaj, R. & Honig, B. (1993). GRASP: graphical representation and analysis of surface properties. *Biophys. J.* **64**, 166–170.
44. Bairoch, A. & Boeckmann, B. (1994). The SWISS-PROT protein sequence data bank: current status. *Nucl. Acid Res.* **19**, 3578–3580.
45. Barton, G.J. (1993). ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37–40.