

# The Ancestral Shape of the Access Proton Path of Mitochondrial ATP Synthases Revealed by a Split Subunit- $\alpha$

Jonathan E. Wong,<sup>1,2</sup> Alena Zíková,<sup>1,2</sup> and Ondřej Gahura <sup>\*,1</sup>

<sup>1</sup>Institute of Parasitology, Biology Centre, Czech Academy of Sciences, České Budějovice, Czech Republic

<sup>2</sup>Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

\*Corresponding author: E-mail: gahura@paru.cas.cz.

Associate editor: Banu Ozkan

## Abstract

The passage of protons across membranes through  $F_1F_o$ -ATP synthases spins their rotors and drives the synthesis of ATP. While the principle of torque generation by proton transfer is known, the mechanisms and routes of proton access and release and their evolution are not fully understood. Here, we show that the entry site and path of protons in the luminal half channel of mitochondrial ATP synthases are largely defined by a short N-terminal  $\alpha$ -helix of subunit- $\alpha$ . In *Trypanosoma brucei* and other Euglenozoa, the  $\alpha$ -helix is part of another polypeptide chain that is a product of subunit- $\alpha$  gene fragmentation. This  $\alpha$ -helix and other elements forming the proton pathway are widely conserved across eukaryotes and in Alphaproteobacteria, the closest extant relatives of mitochondria, but not in other bacteria. The  $\alpha$ -helix blocks one of two proton routes found in *Escherichia coli*, resulting in a single proton entry site in mitochondrial and alphaproteobacterial ATP synthases. Thus, the shape of the access half channel predates eukaryotes and originated in the lineage from which mitochondria evolved by endosymbiosis.

**Key words:** subunit- $\alpha$ , proton translocation, gene fragmentation, proton path, *Trypanosoma brucei*, mitochondrial ATP synthase.

## Introduction

$F_1F_o$ -ATP synthases are rotary molecular machines that convert the proton-motive force across energy-transducing membranes of bacteria, chloroplasts, and mitochondria to the chemical energy of ATP. The flow of protons down their electrochemical gradient through the membrane-intrinsic  $F_o$  subcomplex spins the enzyme's asymmetric rotor, whose rotation changes the conformation of three catalytic sites on the membrane-extrinsic  $F_1$ -ATPase subcomplex, resulting in the synthesis of ATP. The transmembrane part of the rotor, the  $c$ -ring, consists of several copies of subunit- $c$ . Proton translocation is based on the sequential protonation of a glutamate residue in each protomer of the  $c$ -ring, rotation of the  $c$ -ring with the neutralized glutamates exposed to the hydrophobic environment of the phospholipid bilayer, and release of protons on the other side of the membrane. The second element crucial for proton passage is the stator subunit- $\alpha$  (also referred to as ATP6 or subunit-6). Canonical subunit- $\alpha$  contains six  $\alpha$ -helices ( $h_1$ – $h_6$ ), five of which are largely embedded in the membrane. While  $h_1$  is a vertical transmembrane helix,  $h_3$ – $h_6$  are horizontal, positioned at an angle of approximately 20° relative to the plane of the membrane. Helices  $h_5$  and  $h_6$  form a hairpin bending around the  $c$ -ring. The sites of proton binding and release from the  $c$ -ring glutamates are separated by a universally conserved arginine of  $h_5$ . The positively charged arginine prevents protons from leaking without

$c$ -ring rotation and induces their dissociation at the release site (reviewed in Walker 2017; Kuhlbrandt 2019; Guo and Rubinstein 2022). The proton-binding glutamates on the  $c$ -ring are accessible from the surface of the complex by two offset solvent-filled half channels. In mitochondrial ATP synthases, the access half channel opens into the lumen of cristae; hence, it is called the luminal half channel. It is shaped by subunit- $\alpha$  and adjacent membrane elements, including lipids (Klusich et al. 2017). Based on the high rate of proton translocation, it was proposed that delivery of protons to the  $c$ -ring is facilitated by a Grotthuss mechanism through a chain of coordinated water molecules and adjacent amino acid residues (Feniouk et al. 2004). Recently, the proton-hopping mechanism was supported by observations of cryoEM densities corresponding to water molecules in the inner part of the channel in mammalian (Spikes et al. 2020) and trypanosomal (Gahura et al. 2022) mitochondrial ATP synthases. Molecular dynamics simulations are also consistent with the role of coordinated water molecules in proton translocation (Ivontsin et al. 2022). Direct evidence for the Grotthuss mechanism of proton transfer in both the inlet and the outlet channels was provided by a combination of mutagenesis and single-molecule rotation studies with the ATP synthase from *Escherichia coli* (Yanagisawa and Frasch 2021).

Here, we explore the luminal half-channel-forming elements across species and show that the N-terminal region

© The Author(s) 2023. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

Open Access

of subunit-a forms a short  $\alpha$ -helix common to alphaproteobacterial and a vast majority of mitochondrial ATP synthases. In Euglenozoa, this helix is a part of a different polypeptide chain encoded by a nuclear gene that is presumably the product of fragmentation of the mitochondrial gene for subunit-a. We document that the helix defines the proton entry site and contributes to the shaping of the inner part of the half channel in modern eukaryotes and their common ancestors.

## Results and Discussion

### The Short N-Terminal $\alpha$ -Helix of Subunit-a is a Widely Conserved Structural Element of Mitochondrial ATP Synthases

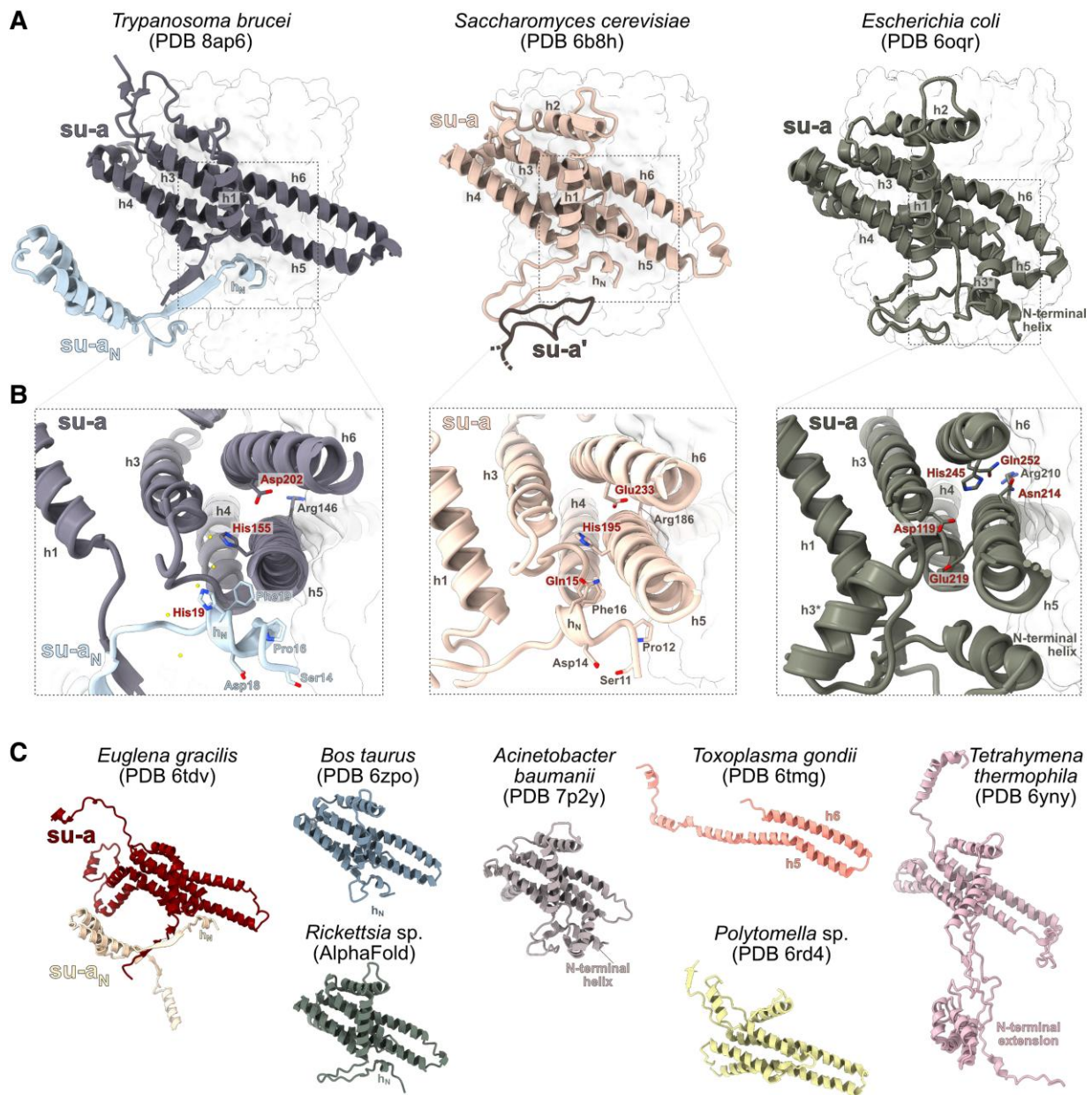
When examining the proton translocation elements of mitochondrial ATP synthases of two divergent eukaryotes, *Saccharomyces cerevisiae* (Guo et al. 2017) and *Trypanosoma brucei* (Gahura et al. 2022), we noticed a remarkable similarity between the N-terminal regions of subunit-a in the yeast and subunit ATPEG3 in the parasite. The N-termini of both proteins form a short single-turn  $\alpha$ -helix, hereafter referred to as  $h_N$ , which is located at the layer corresponding to the luminal surface of the membrane and interacts with the horizontal helix  $h_5$  of subunit-a in the same manner in both structures. Downstream of  $h_N$ , yeast subunit-a, and trypanosomal ATPEG3 follow a similar path; however, after ten residues, the yeast chain turns back toward the first transmembrane helix  $h_1$ , thereby diverging from ATPEG3, which continues to the periphery of the complex. The canonical trypanosomal subunit-a, which contains  $h_1$  and four characteristic horizontal membrane-intrinsic helices  $h_3$ – $h_6$ , begins nine amino acid residues upstream of  $h_1$  in close proximity to the site of divergence between yeast subunit-a and ATPEG3 (fig. 1A). Based on the striking structural and sequence (see below) similarity, we conclude that subunit-a in *T. brucei* has been split and ATPEG3 represents its N-terminal fragment. To acknowledge the origin of ATPEG3, we introduce an alternative name subunit- $a_N$ .

The split of subunit-a in *T. brucei* occurred in the region corresponding to the loop involved in ATP synthase dimerization in *S. cerevisiae* (Guo et al. 2017). In *T. brucei*, the loop is structurally replaced by a two-stranded  $\beta$ -sheet, formed by the N-terminus of subunit-a and a  $\beta$ -strand of subunit- $a_N$  (fig. 1A and B). Unlike in yeast, this structure does not contribute to the ATP synthase dimerization interface, which is provided by subunits g and e in this organism (Gahura et al. 2022). After splitting, subunit- $a_N$  was extended at its C-terminus. The extension interacts exclusively with lineage-specific subunits (ATPTB1, ATPTB6, ATPTB11, ATPTB12) or with lineage-specific extensions of subunits shared with other ATP synthases (subunit-i/j and k; Gahura et al. 2022). It extends to the periphery of the membrane part of the  $F_0$ -moiety, where it co-constitutes a euglenozoan-specific subcomplex, separated from the conserved part by a lipid-filled cavity (Muhleip et al. 2019; Gahura et al. 2022).

Consistently, fragmented subunit-a is also discernible in the structure of ATP synthase from *Euglena gracilis* in contrast to other mitochondrial and bacterial subunit-a exemplified by several selected species (fig. 1C; Muhleip et al. 2019; Gahura et al. 2021).

The sequence of  $h_N$  and flanking residues in yeast is highly similar to the N-terminal sequence of subunit-a in all major groups of Alphaproteobacteria (fig. 2; supplementary table S1, supplementary data S1 and S2, Supplementary Material online), the closest extant relatives to mitochondria (Munoz-Gomez et al. 2022). We predicted the structure of subunit-a from *Rickettsia rickettsii* using AlphaFold (Jumper et al. 2021). The model resembles structures of bacterial and mitochondrial homologs resolved experimentally and includes  $h_N$  located at exactly the same position as the  $h_N$  in *T. brucei* and *S. cerevisiae* (fig. 1C). The similarity of this region in ATP synthases from mitochondria of two deeply branching eukaryotic lineages and the group of bacteria from which mitochondria evolved strongly indicates that the observed architecture of the N-terminal segment of subunit-a is ancestral.

We compared subunit-a sequences from 94 species covering all major eukaryotic lineages, and structures from selected species determined by cryoEM or predicted by AlphaFold. The sequence comparison documented that the split of subunit-a observed in *T. brucei* and *E. gracilis* is restricted to Euglenozoa, because subunit-a from all representatives of the sister lineage Heterolobosea contains a conserved sequence of  $h_N$ . Consistently, the subunit- $a_N$  was found in euglenids, diplomonids, and kinetoplastids, the three lineages of Euglenozoa (Sinha and Wideman 2023; supplementary table S1 and supplementary data S3, Supplementary Material online). Further, the analysis revealed a high-sequence similarity in the N-terminal region, with most sequences showing a degenerate sequence motif SPLEQF (fig. 2; supplementary table S1, Supplementary Material online), supporting the notion that the presence of  $h_N$  represents an ancestral state. The only organisms without the SPLEQF motif were most alveolates, chlorophycean algae *Polytomella* and *Chlamydomonas*, and all metazoans (fig. 2; supplementary table S1, Supplementary Material online). In alveolates and chlorophycean algae, the N-terminal region including  $h_1$  is divergent or absent, as documented by cryoEM structures of ATP synthases from *Tetrahymena* (Flygaard et al. 2020), *Polytomella* (Murphy et al. 2019), and *Toxoplasma* (Muhleip et al. 2021). In *Polytomella*, subunit-a lacks the entire N-terminal region including the transmembrane helix  $h_1$ , which is positionally replaced by the small transmembrane subunit Asa10 (Murphy et al. 2019). However, Asa10 is unlikely to be a fragment of subunit-a, because it crosses the membrane in the opposite direction (Sinha and Wideman 2023). The missing helix  $h_N$  in *Polytomella* is positionally replaced by another conserved component, subunit-8 (Murphy et al. 2019). Subunit-a in *Toxoplasma* is even more reduced and contains horizontal helices  $h_5$  and  $h_6$  only (fig. 1C). On the contrary, subunit-a in *Tetrahymena* features a highly

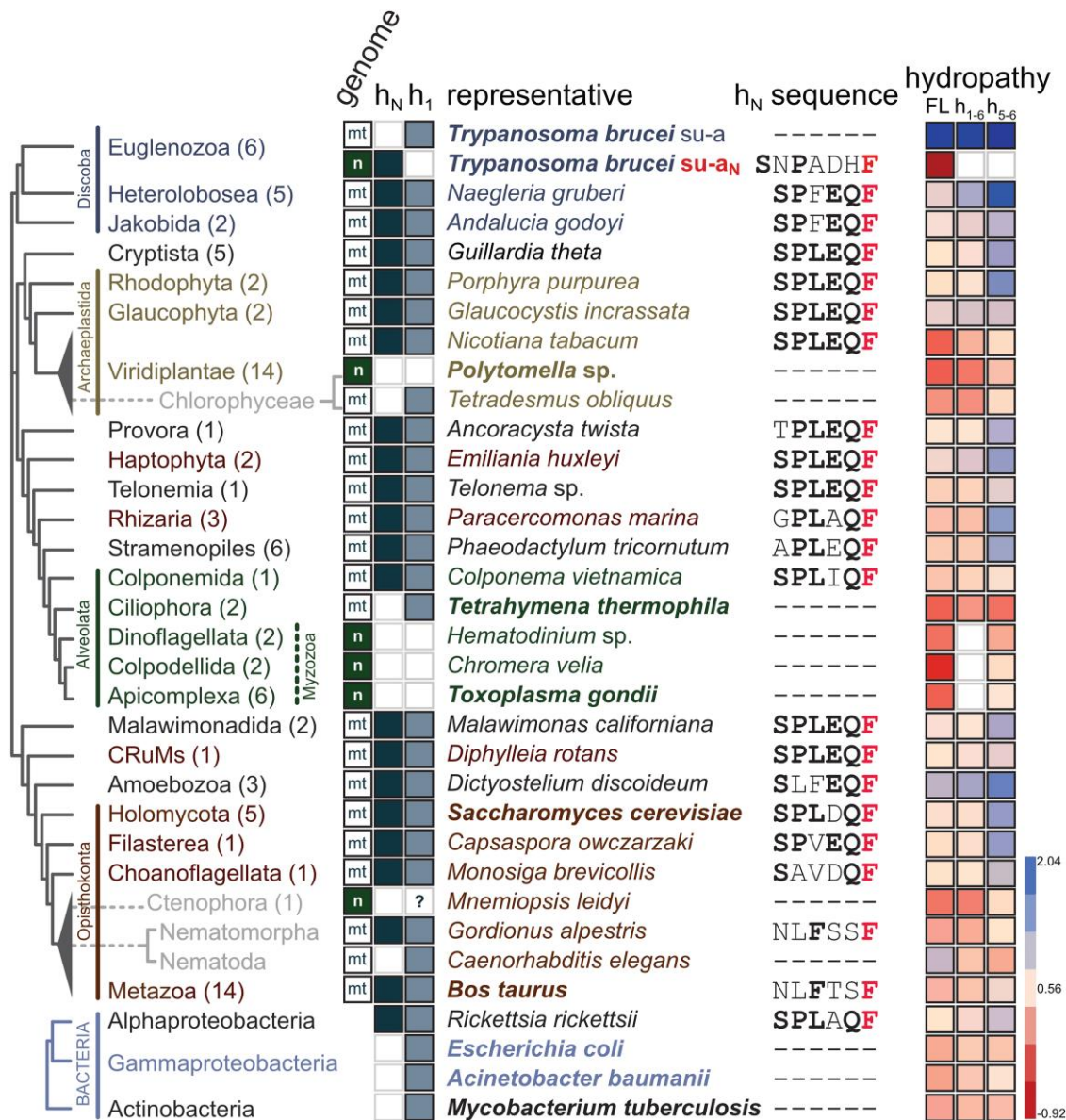


**FIG. 1.** Split subunit-a of mitochondrial ATP synthase in *Trypanosoma brucei* retains ancestral structure and shapes the access proton half channel. (A) Comparison of the structures of subunit-a in *Saccharomyces cerevisiae* and *Escherichia coli* and subunits a and  $a_N$  in *T. brucei*. The c-rings are shown as a transparent surface. In the structure of *S. cerevisiae*, a fragment of the subunit-a from the other ATP synthase protomer is shown (su-a'). (B) The close-up view of the structures in (A). Conserved residues of  $h_N$  and other selected conserved residues are shown as sticks. The residues labeled in red are directly involved in proton translocation or coordinate water molecules in the Grothuss chain. Water molecules inferred from cryoEM density are shown as yellow balls in the *T. brucei* structure. (C) Structures of subunit-a of mitochondrial and bacterial ATP synthases from selected species. The structure of *Rickettsia* was predicted by AlphaFold.

extended and structurally divergent N-terminus, which interacts predominantly with ciliate-specific ATP synthase components. Overall, the regions occupied by N-terminal parts of subunit-a in other ATP synthases are structurally highly divergent in *Toxoplasma* and *Tetrahymena*. The only alveolate with a canonical N-terminus of subunit-a is *Colponema* (fig. 2), which supports the view that it represents a lineage sister to all other Alveolata (Tikhonenkov et al. 2020).

In Metazoa, the N-terminal sequence of subunit-a differs from  $h_N$  in other groups, but invariably contains the motif

FxxF (fig. 2; supplementary table S1 and supplementary data S1, Supplementary Material online). In experimentally determined and predicted structures of mammalian (Pinke et al. 2020; Spikes et al. 2020) and *Drosophila* subunit-a, respectively, this motif adopts the canonical fold of  $h_N$  (fig. 1C), indicating that the ancestral architecture is retained despite sequence divergence. The only two exceptions are ctenophores and nematodes, represented by *Mnemiopsis leidyi* and *Caenorhabditis elegans*, respectively, whose subunit-a begins with helix  $h_1$ , like subunit-a in trypanosomes. In *C. elegans*, this observation is in line with a



**FIG. 2.** Helix h<sub>N</sub> is conserved in most mitochondrial ATP synthases and Alphaproteobacteria. In individual lineages mapped on the phylogenetic tree based on Burki et al. (2020) and Tikhonenkov et al. (2022), the following features are shown: number of species included in the analyses (in parentheses); genome, in which subunit-a/a<sub>N</sub> is encoded (“mt”—mitochondrial, “n”—nuclear); presence or absence of helices h<sub>N</sub> and h<sub>1</sub> of subunit-a/a<sub>N</sub>; selected representative species used in the analyses presented in the following columns (for species shown in bold a cryoEM structure of ATP synthase has been reported); sequence of h<sub>N</sub> and flanking residues (universally conserved phenylalanine is shown in red, other conserved residues of the SPLEQF motif, or the FXXF motif in Metazoa, are shown in bold); average hydrophobicity of the entire subunit-a (FL), region h<sub>1</sub>–h<sub>6</sub>, and region h<sub>5</sub>–h<sub>6</sub> (blue—more hydrophobic, red—less hydrophobic).

recent cryoelectron tomography study that showed that the nematode ATP synthase diverged architecturally from canonical ATP synthases from yeasts and mammals and might lack subunits 8, k, and i/j (Buzzard E, McLaren M, Bragoszewski P, Brancaccio A, Ford H, Daum B, Kuwabara P, Collinson I, Gold VAM, unpublished data), which are otherwise widely conserved in eukaryotes (Sinha and Wideman 2023). Due to high conservancy of h<sub>N</sub> across eukaryotes, we wondered whether an equivalent of h<sub>N</sub> could be retained in nematodes, possibly as a product of subunit-a gene split analogous to that observed in *T. brucei*. Such evolutionary independent event of fission of a

mitochondrial gene would not be unprecedented. The gene encoding the mitoribosomal subunit uS3m has been fragmented in Ciliata (Swart et al. 2012; Tobiasson and Amunts 2020) and some Heterolobosea (Fu et al. 2014), and the gene encoding Cox2, a subunit of cytochrome oxidase, has been independently split at least three times in evolution, specifically in alveolates, chlorophycean algae (Perez-Martinez et al. 2001; Waller and Keeling 2006; Rodriguez-Salinas et al. 2012), and the hymenopteran insect *Campsomeris* (Szafranski 2017a). Therefore, we searched the proteomes of Nematoda using several approaches (blast, motif search, structure similarity search

in AlphaFold database) to identify the protein constituting the missing region including  $h_N$ , but we found no candidate, presumably due to the short length of the region and limited sequence conservancy.

In summary, the vast majority of eukaryotes with the exception of most Alveolata, a group of chlorophycean algae, and possibly Nematoda and Ctenophora have retained the ancestral architecture of the region corresponding to the N-terminus of subunit-a.

### Loss of Transmembrane Helix $h_1$ Predetermines the Transfer of Subunit-a From the Mitochondrial to the Nuclear Genome

While the gene encoding subunit- $a_N$  in Euglenozoa was transferred to the nuclear genome, the subunit-a gene has been retained in the mitochondrial genome. Likewise, subunit-a in nematodes, also starting with the first transmembrane helix  $h_1$ , is encoded mitochondrially. In contrast, the gene of subunit-a in *Polytomella* and *Chlamydomonas*, which lacks  $h_1$ , has been transferred to the nuclear genome. This suggests that while the loss of  $h_1$  allows the transfer of the gene from the mitochondrial to the nuclear genome, the removal of the N-terminal fragment upstream of  $h_1$  is not sufficient for the relocation. Consistently, the reduced subunit-a in *Toxoplasma* and related species are also nuclear encoded (Huet et al. 2018; Muhleip et al. 2021; fig. 2). The last known organism with subunit-a encoded in the nuclear genome is the ctenophore *M. leidy*. The transmembrane helix prediction tools DeepTMHMM (Hallgren J, Tsirigos KD, Pedersen MD, Armenteros JJA, Marcatili P, Nielsen H, Krogh A, Winther O, unpublished data), MEMSAT (Jones et al. 1994), and HMMTOP (Tusnady and Simon 2001) do not predict  $h_1$  in the *M. leidy* subunit-a. AlphaFold does model  $h_1$  in *M. leidy*, but with low confidence scores. Thus, it is unclear whether canonical  $h_1$  is present in the *M. leidy* subunit-a or not. Together, these observations indicate that barriers associated with the insertion of  $h_1$ -containing subunit-a to the inner mitochondrial membrane from the intermembrane space (or with translocation of the protein into mitochondria prior its insertion from inside) may be a factor limiting the transfer of the subunit-a gene from mitochondrial DNA to the nuclear genome. We expanded a previous analysis of hydrophobicity of subunit-a (Muhleip et al. 2021) to all species included in our data set. It revealed that some subunit-as (e.g., from *Nicotiana* or *Tetrahymena*) retained in mitochondrial genomes show relatively low hydrophobicity similar to those transferred to the nuclear genome (fig. 2; supplementary table S1, Supplementary Material online), suggesting that the reduced hydrophobicity is not the only prerequisite for the intracellular subunit-a gene transfer. Regardless of the factor that allowed the subunit-a gene to be transferred, this is a very rare event that has most likely occurred only three times in the evolution of eukaryotes described to date, namely in the ancestors of the present-day myxozoans, ctenophores, and Chlamydomonadales, an order of chlorophycean algae including *Polytomella* and *Chlamydomonas* (fig. 2).

### $h_N$ of Subunit-a Contributes to the Coordination of the Chain of Proton Translocating Water Molecules

The wide conservancy of  $h_N$  and its flanking residues (fig. 2), whether within subunit-a or as part of a fission product, implies the functional importance of this region. The  $h_N$  contributes to coordination of the water chain in the luminal proton half channel in structurally characterized ATP synthases from *Trypanosoma* (Gahura et al. 2022), *S. cerevisiae* (Guo et al. 2017), and mammals (Pinke et al. 2020; Spikes et al. 2020; fig. 1B). One of the water molecules in the Grotthuss water chain is coordinated by positionally equivalent His19 of  $h_N$  of *T. brucei* subunit- $a_N$  and Ser6 of  $h_N$  bovine subunit-a (Pinke et al. 2020; Spikes et al. 2020; Gahura et al. 2022). In the vast majority of mitochondrial subunit-as, this position is occupied by another hydrogen bonding residue, glutamine (Gln15 in *S. cerevisiae*; figs. 1B and 2; supplementary table S1, Supplementary Material online). Thus,  $h_N$  most likely contributes to coordination of water molecules in most eukaryotic lineages and Alphaproteobacteria.

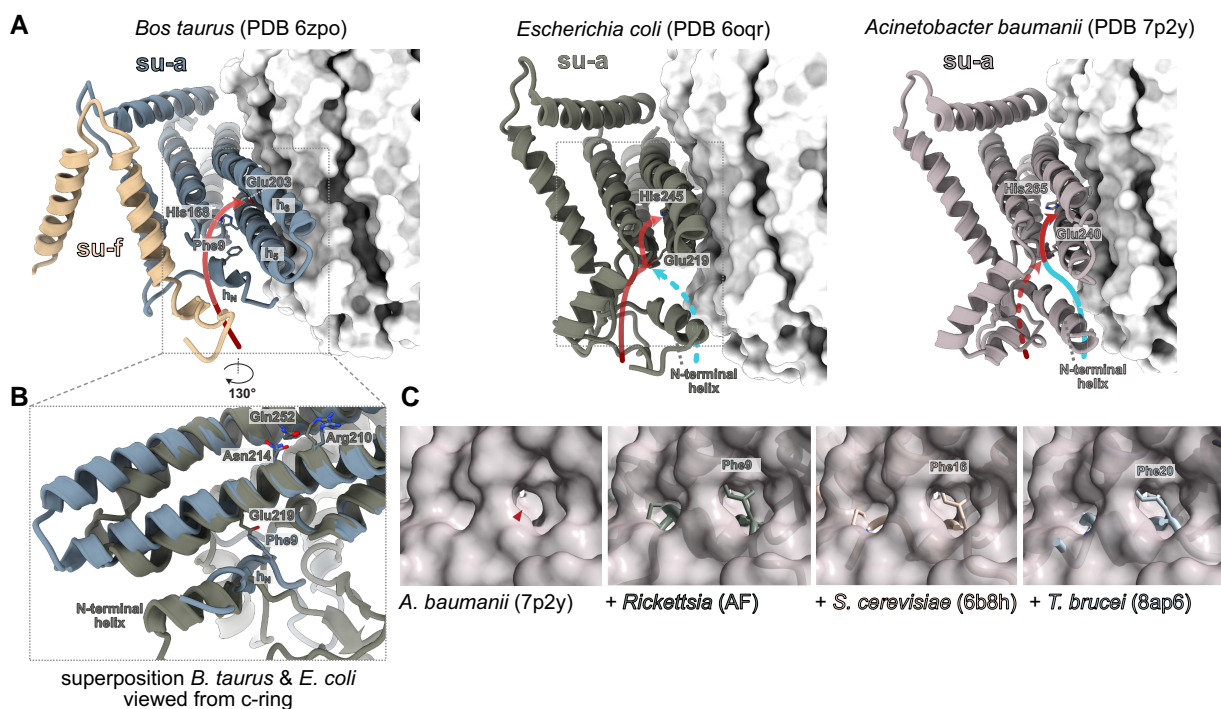
The shape of the inner access half channel in bacteria resembles that of mitochondrial ATP synthase (Montgomery et al. 2021), and a density consistent with the presence of a water molecule in the inlet half channel was observed in the cryoEM map of the *E. coli* ATP synthase (Sobti et al. 2020). Early mutagenesis studies revealed several subunit-a residues crucial for enzymatic function of ATP synthase in *E. coli* (Cain and Simoni 1988; Lightowers et al. 1988; Vik et al. 1988; Cain and Simoni 1989; Howitt et al. 1990; Eya et al. 1991; Hartzog and Cain 1994; Hatch et al. 1995). Most of these residues are localized in the proton half channels or at the interface between subunit-a and c-ring, implying their role in proton translocation. To assess to what extent the inner constricted portion of the luminal proton half channel differs between bacterial and diverse eukaryotic ATP synthases, we examined the conservancy of these residues and compared their positions with the residues proposed to be involved directly in the proton transfer or shown to coordinate ordered proton-transporting water molecules in mitochondrial ATP synthases (Spikes et al. 2020; Gahura et al. 2022; fig. 1B and D; supplementary figure S1, Supplementary Material online). The only two invariable residues of subunit-a are the arginine of  $h_5$  separating the two half channels and a glutamine of  $h_6$  (Arg210 and Gln245 in *E. coli* numbering), juxtaposing with each other. Asn214 and Glu196 lining the proton channel in *E. coli* are widely, yet not universally, conserved in mitochondrial ATP synthases. Functionally important His245 in *E. coli* is replaced by an acidic residue, in most cases glutamate (Glu203, Glu233, and Asp202 in bovine, *S. cerevisiae* and *T. brucei*, respectively), in the vast majority of mitochondrial subunit-as. This residue is assumed to pass protons from the chain of water molecules to the c-ring either directly or by binding other, as yet unidentified, water molecules (Spikes et al. 2020). Other functional residue, Glu219 in *E. coli*, is in most mitochondrial ATP synthases replaced

by histidine (His168/155 in bovine/*T. brucei*), which binds one of the molecules in the water chain in the mammalian and trypanosomal enzymes (Spikes et al. 2020; Gahura et al. 2022). Thus, the positions of these two key histidine and glutamate residues are swapped between *E. coli* and mitochondrial ATP synthases. Nevertheless, due to the properties of these residues, both configurations are compatible with the Grotthuss mechanism of proton translocation. Thus, the segment of the half-channel proximal to the c-ring most likely does not differ between bacterial and mitochondrial ATP synthases, supporting the view that the Grotthuss mechanism represents an ancestral and conserved mode of proton transfer in ATP synthases. In vast majority of mitochondrial ATP synthases and their alpha-proteobacterial ancestors,  $h_N$  is a key component of the network of interactions involved in proton translocation.

### The Entry Point of the Proton Access Half-Channel in Mitochondrial and Alphaproteobacterial ATP Synthases is Defined by $h_N$ of Subunit-a

Notably, the *E. coli* ATP synthase (Sobti et al. 2019) and ATP synthases in bacteria from various groups outside Alphaproteobacteria (*Mycobacterium*; Guo et al. 2021; Montgomery et al. 2021, and *Acinetobacter*; Demmer

et al. 2022) contain an  $\alpha$ -helix at the N-terminus of the subunit-a, which is longer and sequentially dissimilar to  $h_N$  in mitochondrial ATP synthases and positionally shifted further from the loop connecting helices  $h_3$  and  $h_4$  (fig. 1C). As proposed based on the structure of the *Acinetobacter* ATP synthase, the position of the N-terminal helix determines whether protons enter the half channel using the “front” or “rear” entry site as viewed from the c-ring (Demmer et al. 2022). In *Acinetobacter*, the protons have been proposed to enter from the side facing the c-ring through the front entry located between the N-terminal helix and  $h_5$  of subunit-a. In *E. coli*, the protons are funneled from the periplasm to an entry point located between the N-terminal helix of subunit-a and the N-terminus of subunit-b (Yanagisawa and Frasch 2021), termed the rear entry (Demmer et al. 2022). The aqueous funnel is lined with polar and charged residues, which are contributed by the N-terminal helix (Asp10, His14, His15) or  $h_1$  (Glu131, His132) and are not conserved in mitochondrial ATP synthases, and narrows to Glu219, presumably the first residue involved in proton hopping. Notably, molecular dynamics simulations strongly suggest that protons can be delivered to Glu219 in *E. coli* also using another aqueous cavity (Ivontsin et al. 2022), corresponding to the front entry in *Acinetobacter baumannii*. Because *A. baumannii* and



**Fig. 3.** The proton entry point and path in the access half channel are co-determined by  $h_N$ . (A) Comparison of proton paths in the access half channel of the bovine mitochondrial ATP synthases with *Escherichia coli* and *Acinetobacter baumannii* counterparts. The c-ring is shown as white surface. Blue and red arrows indicate the path through the front and rear entry, respectively. (B) Superposition of the structures from *E. coli* and bovine mitochondria shows that  $h_N$  is positionally shifted compared with the N-terminal helix of subunit-a in *E. coli* and that the conserved phenylalanine occupies the position of Glu219 in *E. coli*. Selected functionally important residues are shown as sticks in panels (A) and (B). (C) The position of the proposed front proton entry to the access half channel in *A. baumannii* ATP synthase (Demmer et al. 2022) is blocked by conserved phenylalanine in  $h_N$  of mitochondrial and alphaproteobacterial ATP synthases. A view of the front entry (red arrowhead) to the half channel in the *A. baumannii* subunit-a (surface representation) from the c-ring shown individually and superposed with subunit-a from other ATP synthases (cartoon representation): *Saccharomyces cerevisiae*, the AlphaFold model of *Rickettsia*, and subunits a and  $a_N$  from *Trypanosoma brucei*. The conserved phenylalanine is shown as sticks.

*E. coli* are closely related species and their subunit- $\alpha$ s and adjacent regions are highly similar in both sequence and structure (fig. 1; supplementary data S1, Supplementary Material online), we argue that both proton entry sites are most likely present also in *A. baumannii* (fig. 3A).

Unlike in *E. coli* and *A. baumannii*, in yeast and bovine mitochondrial ATP synthases protons use exclusively the rear opening, formed by a narrow cleft between  $h_N$  and the C-terminal tail of subunit- $f$  (Spikes et al. 2020; Demmer et al. 2022; fig. 3A and B). In the yeast and bovine structures, as well as in the structure of *T. brucei* and the predicted structure of *Rickettsia*, the position of the front entry is blocked by  $h_N$  (fig. 3B and C), specifically by the phenylalanine residue of the SPLEQF motif that is nearly invariably conserved in all  $h_N$  sequences included in our analysis (fig. 2; supplementary table S1, Supplementary Material online). Therefore, we reason that the rear entry is common to all  $h_N$ -containing ATP synthases. In addition, the conserved phenylalanine clashes with the side chain of Glu219 in *E. coli* (fig. 3B), which is the residue, at which the front and rear paths converge. Localization of coordinated water molecules in the part of the trypanosomal inlet channel that is more distal from the c-ring than Glu219 in *E. coli* (fig. 1B) suggests that mitochondrial ATP synthases might use an extended water chain compared with bacteria.

Taken together,  $h_N$  contributes to the shaping of both the surface-exposed and the inner portions of the access proton half channel, underlying its almost universal occurrence in mitochondrial ATP synthases.

## Concluding Remarks

By combining analyses of available ATP synthase structures with sequence comparison and structure prediction across species, we documented that the proton access path and proton transfer mechanism in the access half channel are conserved in the majority of mitochondrial ATP synthases. The key structural determinant, so far largely overlooked, is  $h_N$  of subunit- $\alpha$ , which defines the entry to the access half channel and contributes to hydrogen bonding of water molecules in the Grotthuss chain (fig. 1C). Unlike ATP synthase from *E. coli* and *A. baumannii*, which feature two possible proton entry sites, mitochondrial and alphaproteobacterial counterparts have only one access to the proton inlet half channel, co-defined by  $h_N$  of subunit- $\alpha$ . The functional consequences of this difference remain to be elucidated.

Uniquely,  $F_1F_0$ -ATP synthase in *T. brucei* mitochondria contains two functionally central components, the subunits  $\alpha$  and  $a$ , that are both constituted of two separate polypeptide chains. While the subunit- $\alpha$  is cleaved into two fragments posttranslationally (Gahura et al. 2018; Montgomery et al. 2018), the subunit- $a$  is split on the gene level. Our finding of a split subunit- $a$  in *T. brucei* complements recently reported cases of mitochondrial gene fragmentation revealed by structures of mitochondrial ribosomes (uS4m and uS7m in *Polytomella*; Tobiasson et al. 2022) and respiratory chain supercomplexes (Cox3 and Nad5 in *Tetrahymena*; Zhou et al. 2022; Muhleip et al.

2023), and several occurrences reported earlier (reviewed in Szafranski 2017b). The transfer of the same gene from the mitochondrial to the nuclear genome independently in several lineages facilitates the inference of properties of encoded proteins that are incompatible with such events. We inferred that the transmembrane helix  $h_1$  might represent such an obstacle in the case of subunit- $a$ .

## Materials and Methods

Subunit- $a$  sequences were searched by hmmsearch using the alignment of subunit- $a$  sequences from *Bos taurus*, *S. cerevisiae*, *R. rickettsii*, and *Colponema vietnamica* as a query. The sequences from Ciliata and Apicomplexa were searched by phmmer using sequences from *Tetrahymena* and *Toxoplasma*, respectively, as queries. Both searches were performed on the HMMER web server (Potter et al. 2018) in the Uniprot database. Amino acid sequence alignments were constructed by MUSCLE (Edgar 2004) and corrected manually using available structures as references. Some sequences of subunits  $a$  and  $a_N$  from Euglenozoa were obtained from Sinha and Wideman (2023). Structures were predicted by AlphaFold (Jumper et al. 2021), and superposed by the Matchmaker tool and visualized in ChimeraX (Goddard et al. 2018). Hydrophobicity was calculated as the grand average of hydrophathy using hydrophathy indices from Kyte and Doolittle (1982).

## Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

## Acknowledgments

The authors thank Anzhelika Butenko for the initial search of subunit- $a$  homologs. This work was supported by Czech Science Foundation grant 20-04150Y to O.G., by European Regional Development Fund (ERDF) and Ministry of Education, Youth and Sport (MEYS) project CZ.02.1.01/0.0/0.0/16\_019/0000759, and by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program (grant agreement no. 101044951) to A.Z.

## Author Contributions

Conceptualization: O.G. and A.Z.; investigation and data analyses: J.E.W. and O.G.; writing—original draft: O.G.; writing—review and editing: J.E.W., A.Z., O.G.; funding acquisition: O.G. and A.Z.

## Data availability

All sequences used in this study are from publicly available resources.

**Conflict of interest statement.** They authors declare they have no competing interests.

## References

- Burki F, Roger AJ, Brown MW, Simpson AGB. 2020. The new tree of eukaryotes. *Trends Ecol Evol.* **35**:43–55.
- Cain BD, Simoni RD. 1988. Interaction between Glu-219 and His-245 within the a subunit of F<sub>1</sub>F<sub>o</sub>-ATPase in *Escherichia coli*. *J Biol Chem.* **263**:6606–6612.
- Cain BD, Simoni RD. 1989. Proton translocation by the F1FOATPase of *Escherichia coli*. *J Biol Chem.* **264**:3292–3300.
- Demmer JK, Phillips BP, Uhrig OL, Filloux A, Allsopp LP, Bublitz M, Meier T. 2022. Structure of ATP synthase from ESKAPE pathogen *Acinetobacter baumannii*. *Sci Adv.* **8**:eabl5966.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**:1792–1797.
- Eya S, Maeda M, Futai M. 1991. Role of the carboxyl terminal region of H<sup>+</sup>-ATPase F<sub>o</sub>F<sub>1</sub> a subunit from *Escherichia coli*. *Arch Biochem Biophys.* **284**:71–77.
- Feniouk BA, Kozlova MA, Knorre DA, Cherepanov DA, Mulikidjanian AY, Junge W. 2004. The proton-driven rotor of ATP synthase: ohmic conductance (10 fS), and absence of voltage gating. *Biophys J.* **86**:4094–4109.
- Flygaard RK, Muhleip A, Tobiasson V, Amunts A. 2020. Type III ATP synthase is a symmetry-deviated dimer that induces membrane curvature through tetramerization. *Nat Commun.* **11**:5342.
- Fu CJ, Sheikh S, Miao W, Andersson SG, Baldauf SL. 2014. Missing genes, multiple ORFs, and C-to-U type RNA editing in *Acrasis kona* (Heterolobosea, *Excavata*) mitochondrial DNA. *Genome Biol Evol.* **6**:2240–2257.
- Gahura O, Hierro-Yap C, Zikova A. 2021. Redesignated and reversed: architectural and functional oddities of the trypanosomal ATP synthase. *Parasitology* **148**:1151–1160.
- Gahura O, Muhleip A, Hierro-Yap C, Panicucci B, Jain M, Hollaus D, Slapnickova M, Zikova A, Amunts A. 2022. An ancestral interaction module promotes oligomerization in divergent mitochondrial ATP synthases. *Nat Commun.* **13**:5989.
- Gahura O, Subrtova K, Vachova H, Panicucci B, Fearnley IM, Harbour ME, Walker JE, Zikova A. 2018. The F<sub>1</sub>-ATPase from *Trypanosoma brucei* is elaborated by three copies of an additional p18-subunit. *FEBS J.* **285**:614–628.
- Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. 2018. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**:14–25.
- Guo H, Bueler SA, Rubinstein JL. 2017. Atomic model for the dimeric F<sub>o</sub> region of mitochondrial ATP synthase. *Science* **358**:936–940.
- Guo H, Courbon GM, Bueler SA, Mai J, Liu J, Rubinstein JL. 2021. Structure of mycobacterial ATP synthase bound to the tuberculosis drug bedaquiline. *Nature* **589**:143–147.
- Guo H, Rubinstein JL. 2022. Structure of ATP synthase under strain during catalysis. *Nat Commun.* **13**:2232.
- Hartzog PE, Cain BD. 1994. Second-site suppressor mutations at glycine 218 and histidine 245 in the alpha subunit of F<sub>1</sub>F<sub>o</sub> ATP synthase in *Escherichia coli*. *J Biol Chem.* **269**:32313–32317.
- Hatch LP, Cox GB, Howitt SM. 1995. The essential arginine residue at position 210 in the alpha subunit of the *Escherichia coli* ATP synthase can be transferred to position 252 with partial retention of activity. *J Biol Chem.* **270**:29407–29412.
- Howitt SM, Lightowers RN, Gibson F, Cox GB. 1990. Mutational analysis of the function of the a-subunit of the F<sub>o</sub>F<sub>1</sub>-ATPase of *Escherichia coli*. *Biochim Biophys Acta.* **1015**:264–268.
- Huet D, Rajendran E, van Dooren GG, Lourido S. 2018. Identification of cryptic subunits from an apicomplexan ATP synthase. *Elife* **7**:e38097.
- Ivontsin L, Mashkovtseva E, Nartsissov Y. 2022. Insights on the proton translocation pathways in F<sub>o</sub>F<sub>1</sub>-ATP synthase using molecular dynamics simulations. *Arch Biochem Biophys.* **717**:109135.
- Jones DT, Taylor WR, Thornton JM. 1994. A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry* **33**:3038–3049.
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Zidek A, Potapenko A, et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**:583–589.
- Klusck N, Murphy BJ, Mills DJ, Yildiz O, Kuhlbrandt W. 2017. Structural basis of proton translocation and force generation in mitochondrial ATP synthase. *Elife* **6**:e33274.
- Kuhlbrandt W. 2019. Structure and mechanisms of F-type ATP synthases. *Annu Rev Biochem.* **88**:515–549.
- Kyte J, Doolittle RF. 1982. A simple method for displaying the hydrophobic character of a protein. *J Mol Biol.* **157**:105–132.
- Lightowers RN, Howitt SM, Hatch L, Gibson F, Cox G. 1988. The proton pore in the *Escherichia coli* F<sub>o</sub>F<sub>1</sub>-ATPase: substitution of glutamate by glutamine at position 219 of the alpha-subunit prevents F<sub>o</sub>-mediated proton permeability. *Biochim Biophys Acta.* **933**:241–248.
- Montgomery MG, Gahura O, Leslie AGW, Zikova A, Walker JE. 2018. ATP Synthase from *Trypanosoma brucei* has an elaborated canonical F<sub>1</sub>-domain and conventional catalytic sites. *Proc Natl Acad Sci U S A.* **115**:2102–2107.
- Montgomery MG, Petri J, Spikes TE, Walker JE. 2021. Structure of the ATP synthase from *Mycobacterium smegmatis* provides targets for treating tuberculosis. *Proc Natl Acad Sci U S A.* **118**:e2111899118.
- Muhleip A, Flygaard RK, Baradaran R, Haapanen O, Gruhl T, Tobiasson V, Marechal A, Sharma V, Amunts A. 2023. Structural basis of mitochondrial membrane bending by the I-II-III<sub>2</sub>-IV<sub>2</sub> supercomplex. *Nature* **615**:934–938.
- Muhleip A, Kock Flygaard R, Ovcariakova J, Lacombe A, Fernandes P, Sheiner L, Amunts A. 2021. ATP synthase hexamer assemblies shape cristae of *Toxoplasma* mitochondria. *Nat Commun.* **12**:120.
- Muhleip A, McComas SE, Amunts A. 2019. Structure of a mitochondrial ATP synthase with bound native cardiolipin. *Elife* **8**:e51179.
- Munoz-Gomez SA, Susko E, Williamson K, Eme L, Slamovits CH, Moreira D, Lopez-Garcia P, Roger AJ. 2022. Site-and-branch-heterogeneous analyses of an expanded dataset favour mitochondria as sister to known Alphaproteobacteria. *Nat Ecol Evol.* **6**:253–262.
- Murphy BJ, Klusck N, Langer J, Mills DJ, Yildiz O, Kuhlbrandt W. 2019. Rotary substates of mitochondrial ATP synthase reveal the basis of flexible F<sub>1</sub>-F<sub>o</sub> coupling. *Science* **364**:eaaw9128.
- Perez-Martinez X, Antaramian A, Vazquez-Acevedo M, Funes S, Tolkunova E, d'Alayer J, Claros MG, Davidson E, King MP, Gonzalez-Halphen D. 2001. Subunit II of cytochrome c oxidase in Chlamydomonas algae is a heterodimer encoded by two independent nuclear genes. *J Biol Chem.* **276**:11302–11309.
- Pinke G, Zhou L, Sazanov LA. 2020. Cryo-EM structure of the entire mammalian F-type ATP synthase. *Nat Struct Mol Biol.* **27**:1077–1085.
- Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. 2018. HMMER Web server: 2018 update. *Nucleic Acids Res.* **46**:W200–W204.
- Rodriguez-Salinas E, Riveros-Rosas H, Li Z, Fucikova K, Brand JJ, Lewis LA, Gonzalez-Halphen D. 2012. Lineage-specific fragmentation and nuclear relocation of the mitochondrial cox2 gene in chlorophyte green algae (Chlorophyta). *Mol Phylogenet Evol.* **64**:166–176.
- Sinha SD, Wideman JG. 2023. The persistent homology of mitochondrial ATP synthases. *iScience* **26**:106700.
- Sobti M, Ishmukhametov R, Bouwer JC, Ayer A, Suarna C, Smith NJ, Christie M, Stocker R, Duncan TM, Stewart AG. 2019. Cryo-EM reveals distinct conformations of *E. coli* ATP synthase on exposure to ATP. *Elife* **8**:e43864.
- Sobti M, Walshe JL, Wu D, Ishmukhametov R, Zeng YC, Robinson CV, Berry RM, Stewart AG. 2020. Cryo-EM structures provide insight into how *E. coli* F<sub>1</sub>F<sub>o</sub> ATP synthase accommodates symmetry mismatch. *Nat Commun.* **11**:2615.

- Spikes TE, Montgomery MG, Walker JE. 2020. Structure of the dimeric ATP synthase from bovine mitochondria. *Proc Natl Acad Sci U S A*. **117**:23519–23526.
- Swart EC, Nowacki M, Shum J, Stiles H, Higgins BP, Doak TG, Schotanus K, Magrini VJ, Minx P, Mardis ER, et al. 2012. The *Oxytricha trifallax* mitochondrial genome. *Genome Biol Evol*. **4**:136–154.
- Szafranski P. 2017a. Evolutionarily recent, insertional fission of mitochondrial *cox2* into complementary genes in bilaterian Metazoa. *BMC Genomics* **18**:269.
- Szafranski P. 2017b. Intercompartmental piecewise gene transfer. *Genes (Basel)* **8**:260.
- Tikhonenkov DV, Mikhailov KV, Gawryluk RMR, Belyaev AO, Mathur V, Karpov SA, Zagumyonnyi DG, Borodina AS, Prokina KI, Mylnikov AP, et al. 2022. Microbial predators form a new supergroup of eukaryotes. *Nature* **612**:714–719.
- Tikhonenkov DV, Strassert JFH, Janouskovec J, Mylnikov AP, Aleoshin VV, Burki F, Keeling PJ. 2020. Predatory colponemids are the sister group to all other alveolates. *Mol Phylogenet Evol*. **149**:106839.
- Tobiasson V, Amunts A. 2020. Ciliate mitoribosome illuminates evolutionary steps of mitochondrial translation. *Elife* **9**:e59264.
- Tobiasson V, Berzina I, Amunts A. 2022. Structure of a mitochondrial ribosome with fragmented rRNA in complex with membrane-targeting elements. *Nat Commun*. **13**:6132.
- Tusnady GE, Simon I. 2001. The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**:849–850.
- Vik SB, Cain BD, Chun KT, Simoni RD. 1988. Mutagenesis of the alpha subunit of the  $F_1F_0$ -ATPase from *Escherichia coli*. Mutations at Glu-196, Pro-190, and Ser-199. *J Biol Chem*. **263**:6599–6605.
- Walker JE. 2017. Structure, mechanism and regulation of ATP synthases. In: Wikstrom M, editor. *Mechanisms of primary energy transduction in biology*. Cambridge (UK): The Royal Society of Chemistry. p. 338–373.
- Waller RF, Keeling PJ. 2006. Alveolate and chlorophycean mitochondrial *cox2* genes split twice independently. *Gene* **383**: 33–37.
- Yanagisawa S, Frasch WD. 2021. pH-dependent 11 degrees  $F_1F_0$  ATP synthase sub-steps reveal insight into the  $F_0$  torque generating mechanism. *Elife* **10**:e70016.
- Zhou L, Maldonado M, Padavannil A, Guo F, Letts JA. 2022. Structures of *Tetrahymena*'s respiratory chain reveal the diversity of eukaryotic core metabolism. *Science* **376**:831–839.