The ancestral shape of the access proton path of mitochondrial ATP synthases revealed by a split subunit-a

Jonathan E. Wong[1,2], Alena Zíková[1,2], Ondřej Gahura[1,3*]

[1]Institute of Parasitology, Biology Centre, Czech Academy of Sciences, 37005 České Budějovice, Czech Republic

[2]Faculty of Science, University of South Bohemia, 37005 České Budějovice, Czech Republic

[*]Corresponding author: gahura@paru.cas.cz

Abstract

The passage of protons across membranes through F_1F_o-ATP synthases spins their rotors and drives 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 lumenal half-channel of mitochondrial ATP synthases are largely defined by a short N-terminal α-helix of subunit-a. In *Trypanosoma brucei* and other Euglenozoa, the α-helix is part of another polypeptide chain that is a product of subunit-a gene fragmentation. This α-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 α-helix blocks one of two proton routes found in *Escherichia coli*, resulting in the 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.

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 synthesis of ATP. The transmembrane part of the rotor, the c-ring, consists of several copies of subunit-c. Proton translocation is based on sequential protonation of a glutamate residue of 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-a (also referred to as ATP6 or subunit-6). Canonical subunits-a contain six α-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 ~20° relative to the plane of the membrane. Helices h_5-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 lumenal half-channel. It is shaped by subunit-a and adjacent membrane elements, including lipids (Klusch, 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,

26 the proton-hopping mechanism was supported by observations of cryoEM densities
27 corresponding to water molecules in the inner part of the channel in mammalian (Spikes, et al.
28 2020) and trypanosomal (Gahura, et al. 2022) mitochondrial ATP synthases. Molecular
29 dynamics simulations are also consistent with the role of coordinated water molecules in proton
30 translocation (Ivontsin, et al. 2022). Direct evidence for the Grotthuss mechanism of proton
31 transfer in both the inlet and outlet channels was provided by a combination of mutagenesis and
32 single-molecule rotation studies with the ATP synthase from *Escherichia coli* (Yanagisawa and
33 Frasch 2021).

34 Here, we explore the lumenal half-channel forming elements across species and show that
35 the N-terminal region of subunits-a form a short α-helix common to alphaproteobacterial and
36 vast majority of mitochondrial ATP synthases. In Euglenozoa, this helix is a part of a different
37 polypeptide chain encoded by a nuclear gene that is presumably the product of fragmentation
38 of the mitochondrial gene for subunit-a. We document that the helix defines the proton entry
39 site and contributes to the shaping of the inner part of the half-channel in modern eukaryotes
40 and their common ancestors.

41

42 **Results and discussion**

43 **The short N-terminal α-helix of subunit-a is a widely conserved structural element of**
44 **mitochondrial ATP synthases**

45 When examining the proton translocation elements of mitochondrial ATP synthases of two
46 divergent eukaryotes, *Saccharomyces cerevisiae* (Guo, et al. 2017) and *Trypanosoma brucei*
47 (Gahura, et al. 2022), we noticed a remarkable similarity between the N-terminal regions of
48 subunit-a in the yeast and subunit ATPEG3 in the parasite. The N-termini of both proteins form
49 a short single-turn α-helix, hereafter referred to as h_N, which is located at the layer
50 corresponding to the lumenal surface of the membrane and interacts with the horizontal helix

51 h_5 of subunit-a in the same manner in both structures. Downstream of h_N, yeast subunit-a and
52 trypanosomal ATPEG3 follow a similar path; however, after ten residues the yeast chain turns
53 back towards the first transmembrane helix h_1, thereby diverging from ATPEG3, which
54 continues to the periphery of the complex. The canonical trypanosomal subunit-a, which
55 contains h_1 and four characteristic horizontal membrane-intrinsic helices h_3-h_6, begins nine
56 amino acid residues upstream of h_1 in close proximity to the site of divergence between yeast
57 subunit-a and ATPEG3 (Fig. 1A). Based on the striking structural and sequence (see below)
58 similarity, we conclude that subunit-a in *T. brucei* has been split and ATPEG3 represents its N-
59 terminal fragment. To acknowledge the origin of ATPEG3, we introduce an alternative name
60 subunit-a_N.

61 The split of subunit-a in *T. brucei* occurred in the region corresponding to the loop
62 involved in ATP synthase dimerization in *S. cerevisiae* (Guo, et al. 2017). In *T. brucei*, the loop
63 is structurally replaced by a two-stranded β-sheet, formed by the N-terminus of subunit-a and
64 a β-strand of subunit-a_N (Fig. 1A,B). Unlike in yeast, this structure does not contribute to the
65 ATP synthase dimerization interface, which is provided by subunits g and e in this organism
66 (Gahura, et al. 2022). After splitting, subunit-a_N was extended at its C-terminus. The extension
67 interacts exclusively with lineage-specific subunits (ATPTB1, ATPTB6, ATPTB11,
68 ATPTB12) or with lineage-specific extensions of subunits shared with other ATP synthases
69 (subunit-i/j and k; Gahura, et al. 2022). It extends to the periphery of the membrane part of the
70 F_o-moiety, where it co-constitutes a euglenozoan-specific subcomplex, separated from the
71 conserved part by a lipid-filled cavity (Muhleip, et al. 2019; Gahura, et al. 2022). Consistently,
72 fragmented subunit-a is also discernible in the structure of ATP synthase from *Euglena gracilis*
73 in contrast to other mitochondrial and bacterial subunit-a exemplified by several selected
74 species (Fig. 1C; Muhleip, et al. 2019; Gahura, et al. 2021).

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

85 We compared subunit-a sequences from 94 species covering all major eukaryotic
86 lineages, and structures from selected species determined by cryoEM or predicted by
87 AlphaFold. The sequence comparison documented that the split of subunit-a observed in *T.*
88 *brucei* and *E. gracilis* is restricted to Euglenozoa, because subunits-a from all representatives
89 of the sister lineage Heterolobosea contain a conserved sequence of h_N. Consistently, the
90 subunit-a_N was found in euglenids, diplonemids and kinetoplastids, the three lineages of
91 Euglenozoa (Sinha and Wideman 2023; Supplementary table S1, Supplementary data 3).
92 Further, the analysis revealed high sequence similarity in the N-terminal region, with most
93 sequences showing a degenerate sequence motif SPLEQF (Fig. 2, Supplementary table S1),
94 supporting the notion that presence of h_N represents an ancestral state. The only organisms
95 without the SPLEQF motif were most alveolates, chlorophycean algae *Polytomella* and
96 *Chlamydomonas*, and all metazoans (Fig. 2, Supplementary table S1). In alveolates and
97 chlorophycean algae, the N-terminal region including h_I is divergent or absent, as documented
98 by cryoEM structures of ATP synthases from *Tetrahymena* (Flygaard, et al. 2020), *Polytomella*
99 (Murphy, et al. 2019) and *Toxoplasma* (Muhleip, et al. 2021). In *Polytomella*, subunit-a lacks

100 the entire N-terminal region including the transmembrane helix h_1, which is positionally
101 replaced by the small transmembrane subunit Asa10 (Murphy, et al. 2019). However, Asa10 is
102 unlikely to be a fragment of subunit-a, because it crosses the membrane in the opposite direction
103 (Sinha and Wideman 2023). The missing helix h_N in *Polytomella* is positionally replaced by
104 another conserved component, subunit-8 (Murphy, et al. 2019). Subunit-a in *Toxoplasma* is
105 even more reduced and contains horizontal helices h_5 and h_6 only (Fig. 1C). On the contrary,
106 subunit-a in *Tetrahymena* features a highly extended and structurally divergent N-terminus,
107 which interacts predominantly with ciliate-specific ATP synthase components. Overall, the
108 regions occupied by N-terminal parts of subunit-a in other ATP synthases are structurally highly
109 divergent in *Toxoplasma* and *Tetrahymena*. The only alveolate with a canonical N-terminus of
110 subunit-a is *Colponema* (Fig. 2), which supports the view that it represents a lineage sister to
111 all other Alveolata (Tikhonenkov, et al. 2020).

112 In Metazoa, the N-terminal sequence of subunit-a differs from h_N in other groups, but
113 invariably contains the motif FxxF (Fig. 2 and Supplementary table S1 and Supplementary Data
114 1). In experimentally determined and predicted structures of mammalian (Pinke, et al. 2020;
115 Spikes, et al. 2020) and *Drosophila* subunit-a, respectively, this motif adopts the canonical fold
116 of h_N (Fig. 1C), indicating that the ancestral architecture is retained despite sequence
117 divergence. The only two exceptions are ctenophores and nematodes, represented by
118 *Mnemiopsis leidyi* and *Caenorhabditis elegans*, respectively, whose subunit-a begins with helix
119 h_1, like subunit-a in trypanosomes. In *C. elegans*, this observation is in line with a recent
120 cryoelectron tomography study that showed that the nematode ATP synthase diverged
121 architecturally from canonical ATP synthases from yeasts and mammals and might lack
122 subunits 8, k and i/j (Buzzard, et al. 2023), which are otherwise widely conserved in eukaryotes
123 (Sinha and Wideman 2023). Due to high conservancy of h_N across eukaryotes, we wondered
124 whether an equivalent of h_N could be retained in nematodes, possibly as a product of subunit-a

125 gene split analogous to that observed in *T. brucei*. Such evolutionary independent event of
126 fission of a mitochondrial gene would not be unprecedented. The gene encoding the
127 mitoribosomal subunit uS3m has been fragmented in Ciliata (Swart, et al. 2012; Tobiasson and
128 Amunts 2020) and some Heterolobosea (Fu, et al. 2014), and the gene encoding Cox2, a subunit
129 of cytochrome oxidase, has been independently split at least three times in evolution,
130 specifically in alveolates, chlorophycean algae (Perez-Martinez, et al. 2001; Waller and Keeling
131 2006; Rodriguez-Salinas, et al. 2012), and the hymenopteran insect *Campsomeris* (Szafranski
132 2017a). Therefore, we searched the proteomes of Nematoda using several approaches (blast,
133 motif search, structure similarity search in AlphaFold database) to identify the protein
134 constituting the missing region including h_N, but we found no candidate, presumably due to the
135 short length of the region and limited sequence conservancy.

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

139

140 **Loss of transmembrane helix h_1 predetermines the transfer of subunit-a from the**
141 **mitochondrial to the nuclear genome**

142 While the gene encoding subunit-a_N in Euglenozoa was transferred to the nuclear genome, the
143 subunit-a gene has been retained in the mitochondrial genome. Likewise, subunit-a in
144 nematodes, also starting with the first transmembrane helix h_1, is encoded mitochondrially. In
145 contrast, the gene of subunit-a in *Polytomella* and *Chlamydomonas*, which lacks h_1, has been
146 transferred to the nuclear genome. This suggests that while loss of h_1 allows the transfer of the
147 gene from the mitochondrial to the nuclear genome, the removal of the N-terminal fragment
148 upstream of h_1 is not sufficient for the relocation. Consistently, the reduced subunit-a in
149 *Toxoplasma* and related species are also nuclear encoded (Huet, et al. 2018; Muhleip, et al.

150 2021; Fig. 2). The last known organism with subunit-a encoded in the nuclear genome is the
151 ctenophore *M. leidyi*. The transmembrane helix prediction tools DeepTMHMM (Hallgren et
152 al., bioRxiv), MEMSAT (Jones, et al. 1994), and HMMTOP (Tusnady and Simon 2001) do not
153 predict h1 in the *M. leidyi* subunit-a. Alphafold does model h_1 in *M. leidyi*, but with low
154 confidence scores. Thus, it is unclear whether canonical h_1 is present in the *M. leidyi* subunit-a
155 or not. Together, these observations indicate that barriers associated with the insertion of h_1-
156 containing subunit-a to the inner mitochondrial membrane from the intermembrane space (or
157 with translocation of the protein into mitochondria prior its insertion from inside) may be a
158 factor limiting the transfer of the subunit-a gene from mitochondrial DNA to the nuclear
159 genome. We expanded a previous analysis of hydropathy of subunit-a (Muhleip, et al. 2021) to
160 all species included in our dataset. It revealed that some subunits-a (e.g., from *Nicotiana* or
161 *Tetrahymena*) retained in mitochondrial genomes show relatively low hydrophobicity similar
162 to those transferred to the nuclear genome (Fig. 2, Supplementary table S1), suggesting that the
163 reduced hydrophobicity is not the only prerequisite for the intracellular subunit-a gene transfer.
164 Regardless of the factor that allowed the subunit-a gene to be transferred, this is a very rare
165 event that has most likely occurred only three times in the evolution of eukaryotes described to
166 date, namely in the ancestors of the present-day myzozoans, ctenophores and
167 Chlamydomonadales, an order of chlorophycean algae including *Polytomella* and
168 *Chlamydomonas* (Fig. 2).
169

170 **h_N of subunit-a contributes to the coordination of the chain of proton translocating water**
171 **molecules**

172 The wide conservancy of h_N and its flanking residues (Fig. 2), whether within subunit-a or as
173 part of a fission product, implies the functional importance of this region. The h_N contributes to
174 coordination of the water chain in the lumenal proton half-channel in structurally characterized

175 ATP synthases from *Trypanosoma* (Gahura, et al. 2022), *S. cerevisiae* (Guo, et al. 2017) and
176 mammals (Pinke, et al. 2020; Spikes, et al. 2020) (Fig. 1B). One of the water molecules in the
177 Grotthuss water chain is coordinated by positionally equivalent His19 of h_N of *T. brucei*
178 subunit-a_N and Ser6 of h_N bovine subunit-a (Pinke, et al. 2020; Spikes, et al. 2020; Gahura, et
179 al. 2022). In the vast majority of mitochondrial subunits-a, this position is occupied by another
180 hydrogen bonding residue, glutamine, (Gln15 in *S. cerevisiae*; Figs. 1B&2 and Supplementary
181 table S1). Thus, h_N most likely contributes to coordination of water molecules in most
182 eukaryotic lineages and Alphaproteobacteria.

183 The shape of the inner access half-channel in bacteria resembles that of mitochondrial
184 ATP synthase (Montgomery, et al. 2021), and a density consistent with presence of a water
185 molecule in the inlet half-channel was observed in the cryoEM map of the *E. coli* ATP synthase
186 (Sobti, et al. 2020). Early mutagenesis studies revealed several subunit-a residues crucial for
187 enzymatic function of ATP synthase in *E. coli* (Cain and Simoni 1988; Lightowlers, et al. 1988;
188 Vik, et al. 1988; Cain and Simoni 1989; Howitt, et al. 1990; Eya, et al. 1991; Hartzog and Cain
189 1994; Hatch, et al. 1995). Most of these residues are localized in the proton half-channels or at
190 the interface between the subunit-a and the c-ring, implying their role in proton translocation.
191 To assess to what extent the inner constricted portion of the lumenal proton half-channel differs
192 between bacterial and diverse eukaryotic ATP synthases, we examined the conservancy of these
193 residues and compared their positions with the residues proposed to be involved directly in the
194 proton transfer or shown to coordinate ordered proton-transporting water molecules in
195 mitochondrial ATP synthases (Spikes, et al. 2020; Gahura, et al. 2022); Fig. 1B&D,
196 Supplementary figure 1). The only two invariable residues of subunit-a are the arginine of h5
197 separating the two half-channels and a glutamine of h6 (Arg210 and Gln245 in *E. coli*
198 numbering), juxtaposing with each other. Asn214 and Glu196 lining the proton channel in *E.*
199 *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 vast majority of mitochondrial subunits-a. 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 Grotthus 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 alphaproteobacterial 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 α-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 to the half-channel using the “front” or “rear”

225 entry site as viewed from the c-ring (Demmer, et al. 2022). In *Acinetobacter*, the protons have
226 been proposed to enter from the side facing the c-ring through the front entry located between
227 the N-terminal helix and h_5 of subunit-a. In *E. coli*, the protons are funneled from the periplasm
228 to an entry point located between the N-terminal helix of subunit-a and the N-terminus of
229 subunit-b (Yanagisawa and Frasch 2021), termed the rear entry (Demmer, et al. 2022). The
230 aqueous funnel is lined with polar and charged residues, which are contributed by the N-
231 terminal helix (Asp10, His14, His15) or h_1 (Glu131, His132) and are not conserved in
232 mitochondrial ATP synthases, and narrows to Glu219, presumably the first residue involved in
233 the proton hopping. Notably, molecular dynamics simulations strongly suggest that protons can
234 be delivered to Glu219 in *E. coli* also using another aqueous cavity (Ivontsin, et al. 2022),
235 corresponding to the front entry in *A. baumanii*. Because *A. baumanii* and *E. coli* are closely
236 related species and their subunits-a and adjacent regions are highly similar in both sequence
237 and structure (Fig. 1, Supplementary data 1), we argue that both proton entry sites are most
238 likely present also in *A. baumanii* (Fig. 3A).

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

250 (Fig. 1B) suggests that mitochondrial ATP synthases might use an extended water chain
251 compared to bacteria.

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

255

256 **Concluding remarks**

257 By combination of analyses of available ATP synthase structures with sequence comparison
258 and structure prediction across species, we documented that the proton access path and proton
259 transfer mechanism in the access half-channel is conserved in majority of mitochondrial ATP
260 synthases. The key structural determinant, so far largely overlooked, is h_N of subunit-a, which
261 defines the entry to the access half channel and contributes to hydrogen bonding of water
262 molecules in the Grotthuss chain (Fig. 1C). Unlike ATP synthase from *E. coli* and *A. baumanii*,
263 which feature two possible proton entry sites, mitochondrial and alphaproteobacterial
264 counterparts have only one access to the proton inlet half-channel, co-defined by h_N of subunit-
265 a. Functional consequences of this difference remain to be elucidated.

266 Uniquely, F_1F_o-ATP synthase in *T. brucei* mitochondria contains two functionally central
267 components, the subunits α and a, that are both constituted of two separate polypeptide chains.
268 While the subunit-α is cleaved into two fragments posttranslationally (Gahura, et al. 2018;
269 Montgomery, et al. 2018), the subunit-a is split on the gene level. Our finding of a split subunit-
270 a in *T. brucei* complements recently reported cases of mitochondrial gene fragmentation
271 revealed by structures of mitochondrial ribosomes (uS4m and uS7m in *Polytomella*; Tobiasson,
272 et al. 2022) and respiratory chain supercomplexes (Cox3 and Nad5 in *Tetrahymena*; Mühleip,
273 et al. 2022; Zhou, et al. 2022), and several occurrences reported earlier (reviewed in Szafranski
274 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.

Material and Methods

Subunit-a sequences were searched by hmmsearch using the alignment of subunit-a sequences from *Bos taurus*, *Saccharomyces cerevisiae*, *Rickettsia 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 grand average of hydropathy using hydropathy indices from Kyte and Doolittle 1982.

Acknowledgements

We 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. and 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.

300 **Author contributions**

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

304

305 **Declaration of interests**

306 They authors declare they have no competing interests.

307

308 **References**

309 Burki F, Roger AJ, Brown MW, Simpson AGB. 2020. The New Tree of Eukaryotes. Trends
310 Ecol Evol 35:43-55.

311 Buzzard E, McLaren M, Bragoszewski P, Brancaccio A, Ford H, Daum B, Kuwabara P,
312 Collinson I, Gold VAM, unpublished data,
313 https://www.biorxiv.org/content/10.1101/2023.02.02.526626v2

314 Cain BD, Simoni RD. 1988. Interaction between Glu-219 and His-245 within the a subunit of
315 F_1F_o-ATPase in *Escherichia coli*. Journal of Biological Chemistry 263:6606-6612.

316 Cain BD, Simoni RD. 1989. Proton Translocation by the F1FOATPase of Escherichia coli.
317 Journal of Biological Chemistry 264:3292-3300.

318 Demmer JK, Phillips BP, Uhrig OL, Filloux A, Allsopp LP, Bublitz M, Meier T. 2022.
319 Structure of ATP synthase from ESKAPE pathogen *Acinetobacter baumannii*. Sci Adv
320 8:eabl5966.

321 Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high
322 throughput. Nucleic Acids Res 32:1792-1797.

323 Eya S, Maeda M, Futai M. 1991. Role of the carboxyl terminal region of H^+-ATPase F_oF_1 a
324 subunit from *Escherichia coli*. Arch Biochem Biophys 284:71-77.

325 Feniouk BA, Kozlova MA, Knorre DA, Cherepanov DA, Mulkidjanian AY, Junge W. 2004.
326 The proton-driven rotor of ATP synthase: ohmic conductance (10 fS), and absence of voltage
327 gating. Biophys J 86:4094-4109.
328 Flygaard RK, Muhleip A, Tobiasson V, Amunts A. 2020. Type III ATP synthase is a symmetry-
329 deviated dimer that induces membrane curvature through tetramerization. Nat Commun
330 11:5342.
331 Fu CJ, Sheikh S, Miao W, Andersson SG, Baldauf SL. 2014. Missing genes, multiple ORFs,
332 and C-to-U type RNA editing in *Acrasis kona* (Heterolobosea, Excavata) mitochondrial DNA.
333 Genome Biol Evol 6:2240-2257.
334 Gahura O, Hierro-Yap C, Zikova A. 2021. Redesigned and reversed: architectural and
335 functional oddities of the trypanosomal ATP synthase. Parasitology 148:1151-1160.
336 Gahura O, Muhleip A, Hierro-Yap C, Panicucci B, Jain M, Hollaus D, Slapnickova M, Zikova
337 A, Amunts A. 2022. An ancestral interaction module promotes oligomerization in divergent
338 mitochondrial ATP synthases. Nat Commun 13:5989.
339 Gahura O, Subrtova K, Vachova H, Panicucci B, Fearnley IM, Harbour ME, Walker JE, Zikova
340 A. 2018. The F_1-ATPase from *Trypanosoma brucei* is elaborated by three copies of an
341 additional p18-subunit. FEBS J 285:614-628.
342 Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. 2018.
343 UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci 27:14-
344 25.
345 Guo H, Bueler SA, Rubinstein JL. 2017. Atomic model for the dimeric F_o region of
346 mitochondrial ATP synthase. Science 358:936-940.
347 Guo H, Courbon GM, Bueler SA, Mai J, Liu J, Rubinstein JL. 2021. Structure of mycobacterial
348 ATP synthase bound to the tuberculosis drug bedaquiline. Nature 589:143-147.

349 Guo H, Rubinstein JL. 2022. Structure of ATP synthase under strain during catalysis. Nat
350 Commun 13:2232.
351 Hartzog PE, Cain BD. 1994. Second-site suppressor mutations at glycine 218 and histidine 245
352 in the alpha subunit of F_1F_o ATP synthase in *Escherichia coli*. Journal of Biological Chemistry
353 269:32313-32317.
354 Hatch LP, Cox GB, Howitt SM. 1995. The essential arginine residue at position 210 in the alpha
355 subunit of the *Escherichia coli* ATP synthase can be transferred to position 252 with partial
356 retention of activity. J Biol Chem 270:29407-29412.
357 Howitt SM, Lightowlers RN, Gibson F, Cox GB. 1990. Mutational analysis of the function of
358 the a-subunit of the F_oF_1-ATPase of *Escherichia coli*. Biochim Biophys Acta 1015:264-268.
359 Huet D, Rajendran E, van Dooren GG, Lourido S. 2018. Identification of cryptic subunits from
360 an apicomplexan ATP synthase. Elife 7.
361 Ivontsin L, Mashkovtseva E, Nartsissov Y. 2022. Insights on the proton translocation pathways
362 in F_oF_1-ATP synthase using molecular dynamics simulations. Arch Biochem Biophys
363 717:109135.
364 Jones DT, Taylor WR, Thornton JM. 1994. A model recognition approach to the prediction of
365 all-helical membrane protein structure and topology. Biochemistry 33:3038-3049.
366 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates
367 R, Zidek A, Potapenko A, et al. 2021. Highly accurate protein structure prediction with
368 AlphaFold. Nature 596:583-589.
369 Klusch N, Murphy BJ, Mills DJ, Yildiz O, Kuhlbrandt W. 2017. Structural basis of proton
370 translocation and force generation in mitochondrial ATP synthase. Elife 6.
371 Kuhlbrandt W. 2019. Structure and Mechanisms of F-Type ATP Synthases. Annu Rev Biochem
372 88:515-549.

373 Kyte J, Doolittle RF. 1982. A simple method for displaying the hydropathic character of a
374 protein. J Mol Biol 157:105-132.
375 Lightowlers RN, Howitt SM, Hatch L, Gibson F, Cox G. 1988. The proton pore in the
376 *Escherichia coli* F_oF_1-ATPase: substitution of glutamate by glutamine at position 219 of the
377 alpha-subunit prevents F_o-mediated proton permeability. Biochim Biophys Acta 933:241-248.
378 Montgomery MG, Gahura O, Leslie AGW, Zikova A, Walker JE. 2018. ATP synthase from
379 *Trypanosoma brucei* has an elaborated canonical F_1-domain and conventional catalytic sites.
380 Proc Natl Acad Sci U S A 115:2102-2107.
381 Montgomery MG, Petri J, Spikes TE, Walker JE. 2021. Structure of the ATP synthase from
382 *Mycobacterium smegmatis* provides targets for treating tuberculosis. Proc Natl Acad Sci U S A
383 118.
384 uhleip A, Flygaard RK, Baradaran R, Haapanen O, Gruhl T, Tobiasson V, Marechal A, Sharma
385 V, Amunts A. 2023. Structural basis of mitochondrial membrane bending by the I-II-III_2-IV_2
386 supercomplex. Nature 615:934-938.
387 Muhleip A, Kock Flygaard R, Ovciarikova J, Lacombe A, Fernandes P, Sheiner L, Amunts A.
388 2021. ATP synthase hexamer assemblies shape cristae of *Toxoplasma* mitochondria. Nat
389 Commun 12:120.
390 Muhleip A, McComas SE, Amunts A. 2019. Structure of a mitochondrial ATP synthase with
391 bound native cardiolipin. Elife 8.
392 Munoz-Gomez SA, Susko E, Williamson K, Eme L, Slamovits CH, Moreira D, Lopez-Garcia
393 P, Roger AJ. 2022. Site-and-branch-heterogeneous analyses of an expanded dataset favour
394 mitochondria as sister to known Alphaproteobacteria. Nat Ecol Evol 6:253-262.
395 Murphy BJ, Klusch N, Langer J, Mills DJ, Yildiz O, Kuhlbrandt W. 2019. Rotary substates of
396 mitochondrial ATP synthase reveal the basis of flexible F_1-F_o coupling. Science 364.

397 Perez-Martinez X, Antaramian A, Vazquez-Acevedo M, Funes S, Tolkunova E, d'Alayer J,
398 Claros MG, Davidson E, King MP, Gonzalez-Halphen D. 2001. Subunit II of cytochrome c
399 oxidase in Chlamydomonad algae is a heterodimer encoded by two independent nuclear genes.
400 J Biol Chem 276:11302-11309.
401 Pinke G, Zhou L, Sazanov LA. 2020. Cryo-EM structure of the entire mammalian F-type ATP
402 synthase. Nat Struct Mol Biol 27:1077-1085.
403 Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. 2018. HMMER web server: 2018
404 update. Nucleic Acids Res 46:W200-W204.
405 Rodriguez-Salinas E, Riveros-Rosas H, Li Z, Fucikova K, Brand JJ, Lewis LA, Gonzalez-
406 Halphen D. 2012. Lineage-specific fragmentation and nuclear relocation of the mitochondrial
407 cox2 gene in chlorophycean green algae (Chlorophyta). Mol Phylogenet Evol 64:166-176.
408 Sinha SD, Wideman JG. 2023. The persistent homology of mitochondrial ATP synthases.
409 iScience 26.
410 Sobti M, Ishmukhametov R, Bouwer JC, Ayer A, Suarna C, Smith NJ, Christie M, Stocker R,
411 Duncan TM, Stewart AG. 2019. Cryo-EM reveals distinct conformations of *E. coli* ATP
412 synthase on exposure to ATP. Elife 8.
413 Sobti M, Walshe JL, Wu D, Ishmukhametov R, Zeng YC, Robinson CV, Berry RM, Stewart
414 AG. 2020. Cryo-EM structures provide insight into how *E. coli* F_1F_o ATP synthase
415 accommodates symmetry mismatch. Nat Commun 11:2615.
416 Spikes TE, Montgomery MG, Walker JE. 2020. Structure of the dimeric ATP synthase from
417 bovine mitochondria. Proc Natl Acad Sci U S A 117:23519-23526.
418 Swart EC, Nowacki M, Shum J, Stiles H, Higgins BP, Doak TG, Schotanus K, Magrini VJ,
419 Minx P, Mardis ER, et al. 2012. The *Oxytricha trifallax* mitochondrial genome. Genome Biol
420 Evol 4:136-154.

421 Szafranski P. 2017a. Evolutionarily recent, insertional fission of mitochondrial cox2 into
422 complementary genes in bilaterian Metazoa. BMC Genomics 18:269.
423 Szafranski P. 2017b. Intercompartmental Piecewise Gene Transfer. Genes (Basel) 8.
424 Tikhonenkov DV, Mikhailov KV, Gawryluk RMR, Belyaev AO, Mathur V, Karpov SA,
425 Zagumyonnyi DG, Borodina AS, Prokina KI, Mylnikov AP, et al. 2022. Microbial predators
426 form a new supergroup of eukaryotes. Nature 612:714-719.
427 Tikhonenkov DV, Strassert JFH, Janouskovec J, Mylnikov AP, Aleoshin VV, Burki F, Keeling
428 PJ. 2020. Predatory colponemids are the sister group to all other alveolates. Mol Phylogenet
429 Evol 149:106839.
430 Tobiasson V, Amunts A. 2020. Ciliate mitoribosome illuminates evolutionary steps of
431 mitochondrial translation. Elife 9.
432 Tobiasson V, Berzina I, Amunts A. 2022. Structure of a mitochondrial ribosome with
433 fragmented rRNA in complex with membrane-targeting elements. Nat Commun 13:6132.
434 Tusnady GE, Simon I. 2001. The HMMTOP transmembrane topology prediction server.
435 Bioinformatics 17:849-850.
436 Vik SB, Cain BD, Chun KT, Simoni RD. 1988. Mutagenesis of the alpha subunit of the F_1F_o-
437 ATPase from *Escherichia coli*. Mutations at Glu-196, Pro-190, and Ser-199. Journal of
438 Biological Chemistry 263:6599-6605.
439 Walker JE. 2017. Structure, Mechanism and Regulation of ATP Synthases. In: Wikstrom M,
440 editor: The Royal Society of Chemistry.
441 Waller RF, Keeling PJ. 2006. Alveolate and chlorophycean mitochondrial cox2 genes split
442 twice independently. Gene 383:33-37.
443 Yanagisawa S, Frasch WD. 2021. pH-dependent 11 degrees F_1Fo ATP synthase sub-steps
444 reveal insight into the F_o torque generating mechanism. Elife 10.

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

Figures and legends

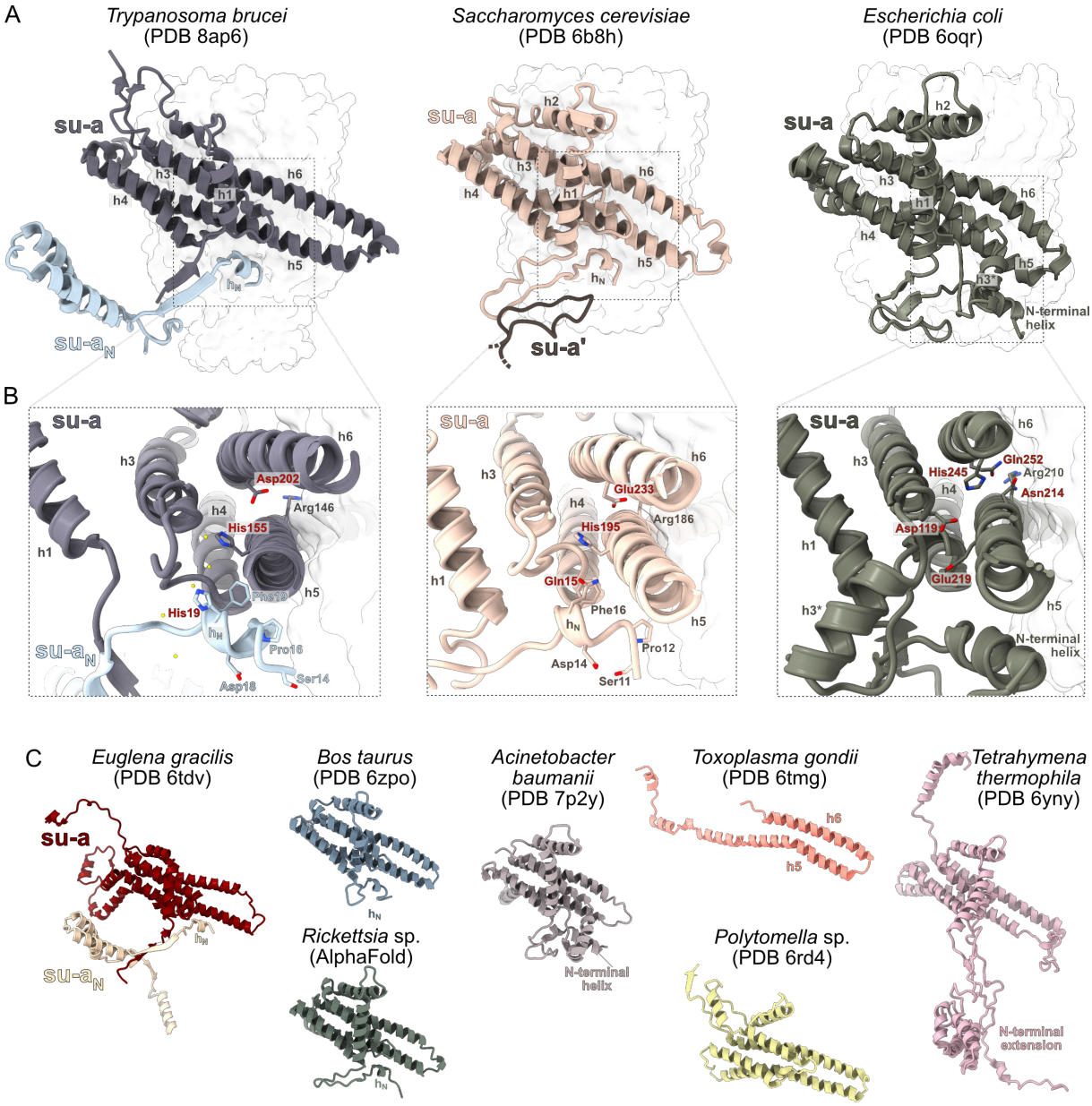
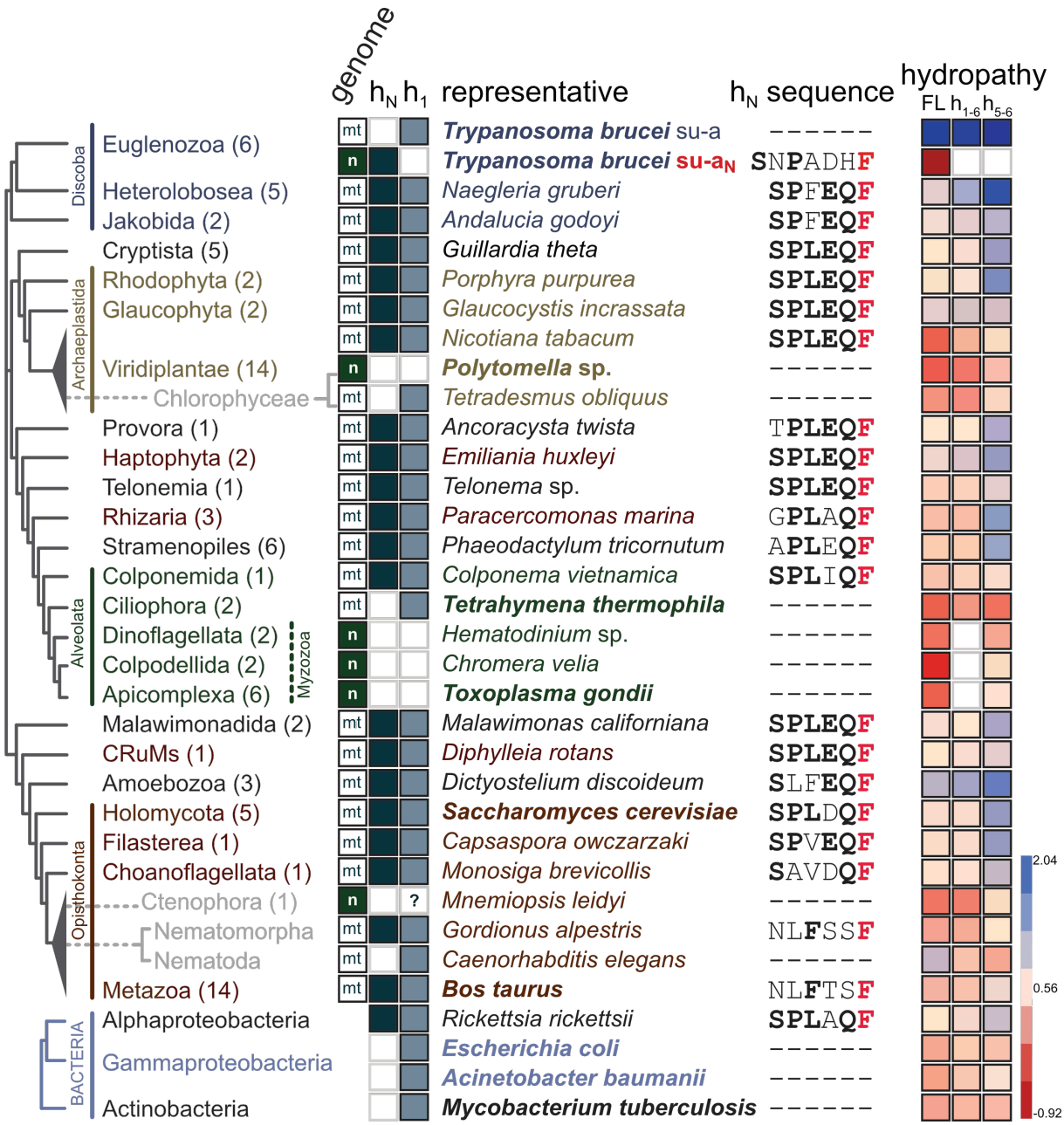


Figure 1. Split subunit-a of mitochondrial ATP synthase in *Trypanosoma brucei* retains ancestral structure and shapes the access proton half-channel (A) Comparison of the structures of subunit-a in *S. cerevisiae* and *E. coli* and subunits a and a_N in *T. brucei*. The c-rings are shown as transparent surface. In the structure from *S. cerevisiae*, a fragment of the

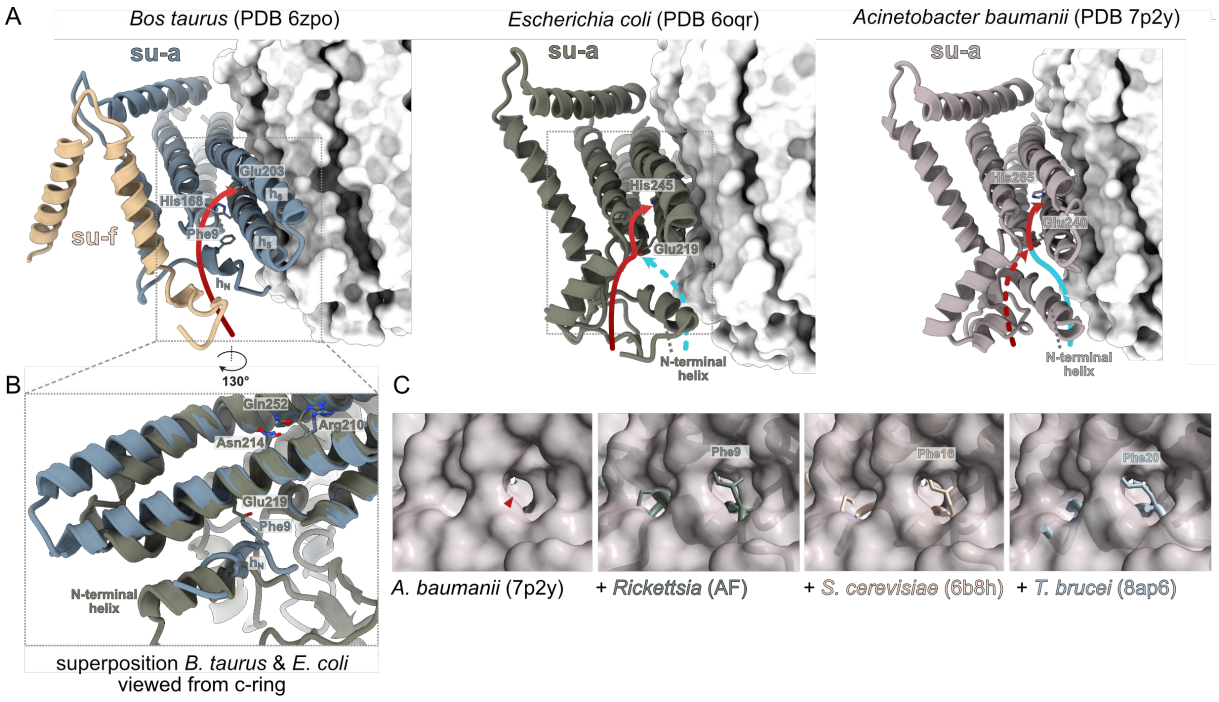
454 subunit-a from the other ATP synthase protomer is shown (su-a'). **(B)** The close-up view of the
455 structures in (A). Conserved residues of h_N and other selected conserved residues are shown as
456 sticks. The residues labeled in red are directly involved in proton translocation or coordinate
457 water molecules in the Grotthuss chain. Water molecules inferred from cryoEM density are
458 shown as yellow balls in the *T. brucei* structure. **(C)** Structures of subunit-a of mitochondrial
459 and bacterial ATP synthases from selected species. The structure from *Rickettsia* was predicted
460 by AlphaFold.



461

462 **Figure 2. Helix h_N is conserved in most mitochondrial ATP synthases and**
463 **Alphaproteobacteria.** For individual lineages mapped on the phylogenetic tree based on
464 (Burki, et al. 2020; Tikhonenkov, et al. 2022) following features are shown: number of species
465 included in the analyses (in parentheses); genome, in which subunit-a/a_N is encoded ("mt" –
466 mitochondrial, "n" – nuclear); presence or absence of helices h_N and h_1 of subunit-a/a_N; selected
467 representative species used in the analyses presented in the following columns (for species
468 shown in bold a cryoEM structure of ATP synthase has been reported); sequence of h_N and
469 flanking residues (universally conserved phenylalanine is shown in red, other conserved
470 residues of the SPLEQF motif, or the FXXF motif in Metazoa, are shown in bold); average
471 hydrophobicity of the entire subunit-a (FL), region h_1-h_6, and region h_5-h_6 (blue – more
472 hydrophobic, red – less hydrophobic).

473



474

475 **Figure 3. The proton entry point and path in the access half-channel is co-determined by**
476 **h_N. (A)** Comparison of proton paths in the access half-channel of the bovine mitochondrial ATP
477 synthases with *Escherichia coli* and *Acinetobacter baumanii* 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 to 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. baumanii* 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. baumanii* subunit-a (surface representation) from the c-ring shown individually and superposed with subunit-a from other ATP synthases (cartoon representation): *S. cerevisiae*, the AlphaFold model of *Rickettsia*, and subunits a and a_N from *T. brucei*. The conserved phenylalanine is shown as sticks.