ABCF ATPases involved in protein synthesis, ribosome assembly and antibiotic resistance: structural and functional diversification across the tree of life

Victoriia Murina*[1,2], Marje Kasari*[1], Hiraku Takada[1,2], Mariliis Hinnu[3], Chayan Kumar Saha[1], James W. Grimshaw[4], Takahiro Seki[5], Michael Reith[1], Marta Putrinš[3], Tanel Tenson[3], Henrik Strahl[4], Vasili Hauryliuk[1,2,3] and Gemma Catherine Atkinson[†1]

[1]Department of Molecular Biology, Umeå University, 901 87, Umeå, Sweden

[2]Laboratory for Molecular Infection Medicine Sweden (MIMS), Umeå University, 901 87, Umeå, Sweden

[3]University of Tartu, Institute of Technology, Nooruse 1, 50411 Tartu, Estonia

[4]Centre for Bacterial Cell Biology, Institute for Cell and Molecular Biosciences Newcastle University, Richardson Road, Newcastle upon Tyne, NE2 4AX, United Kingdom

[5]Department of Applied Chemistry and Biotechnology, Faculty of Engineering, Chiba University, 263-8522, Chiba, Japan

* These authors contributed equally to this work

[†] To whom correspondence should be addressed

Abstract

Within the larger ABC superfamily of ATPases, ABCF family members eEF3 in *Saccharomyces cerevisiae* and EttA in *Escherichia coli* have been found to function as ribosomal translation factors. Several other ABCFs including biochemically characterised VgaA, LsaA and MsrE confer resistance to antibiotics that target the peptidyl transferase centre and exit tunnel of the ribosome. However, the diversity of ABCF subfamilies, the relationships among subfamilies and the evolution of antibiotic resistance factors from other ABCFs have not been explored. To address this, we analysed the presence of ABCFs and their domain architectures in 4505 genomes across the tree of life. We find 45 distinct subfamilies of ABCFs that are widespread across bacterial and eukaryotic phyla, suggesting they were present in the last common ancestor of both. Surprisingly, currently known antibiotic resistance (ARE) ABCFs are not confined to a distinct lineage of the ABCF family

tree. This suggests that either antibiotic resistance is a pervasive feature of bacterial ABCFs, or it is relatively easy to evolve antibiotic resistance from other ABCF functions. Our data suggest there are a number of previously unidentified ARE ABCFs in antibiotic producers and important human pathogens. We also find that ATPase-deficient mutants of all four *E. coli* ABCFs (EttA, YbiT, YheS and Uup) inhibit protein synthesis, indicative of their ribosomal function, and demonstrate a genetic interaction of ABCFs Uup and YheS with translational GTPase BipA involved in assembly of the 50S ribosome subunit. Finally, we show that *Bacillus subtilis* VmlR is a ribosome-binding resistance factor localised to the cytoplasm.

Author summary

Isolated members of the ABCF protein family of ATP-hydrolysing enzymes have been found to have important roles in protein synthesis and antibiotic resistance. However, their full diversity across the tree of life, and their evolutionary histories have never been examined. Therefore, we analysed the presence of ABCFs and their constituent domains in genomes across the tree of life, discovering 45 distinct subfamilies of ABCFs that are widespread across bacterial and eukaryotic phyla. This includes several subfamilies that we predict comprise novel antibiotic resistance (ARE) ABCFs, present in antibiotic producers and important human pathogens. There are significant gaps in our knowledge about the functional capabilities of different ABCF families. To address this, we have made ATPase domain mutants of all four *Escherichia coli* ABCFs, showing that they inhibit protein synthesis and indicating a role on the ribosome. Furthermore, we demonstrate a genetic interaction of two *E. coli* ABCFs with the GTPase BipA, involved in ribosome assembly. Finally, we show that *Bacillus subtilis* VmlR in the ARE2 subfamily is a ribosome-binding resistance factor localised to the cytoplasm. As more is discovered about the function of individual ABCFs, the more it will be possible to predict functions of uncharacterised members, using the ABCF family tree as a framework.

54 **Introduction**

55 Protein biosynthesis – translation – is the reading and deciphering of information coded in genes
56 to produce proteins. It is one of the most ancient and central cellular processes, and control of the
57 various stages of translation is achieved via an intricate interplay of multiple molecular
58 interactions. For many years, enzymatic control of the ribosomal cycle was thought to be mainly
59 orchestrated mainly by translational GTPases (trGTPases). That view of translation has been
60 nuanced by the identification of multiple ATPases in the ABC superfamily that have important
61 roles in translational regulation on the ribosome. The ABC protein eEF3 (eukaryotic Elongation
62 Factor 3) is an essential factor for polypeptide elongation in *Saccharomyces cerevisiae* [1] with
63 proposed roles in E-site tRNA release and ribosome recycling [2-4]. This fungi-specific
64 translational ABC ATPase appeared to be an exception to the tenet that trGTPases are the
65 enzymatic rulers of the ribosome, until ABCE1 (also known as Rli1), a highly conserved protein in
66 eukaryotes and archaea, was identified as another ribosome recycling factor [5-7].

67

68 The ABC ATPases together comprise one of the most ancient superfamilies of proteins, evolving
69 well before the last common ancestor of life [8]. The superfamily contains members with wide
70 varieties of functions, but is best known for its membrane transporters [9]. Families of proteins
71 within the ABC superfamily are named alphabetically ABCA to ABCH, following the nomenclature
72 of the human proteins [10]. While most ABCs carry membrane-spanning domains (MSDs), these
73 are lacking in ABCE and ABCF families [11]. ABCF proteins of eukaryotes include eEF3, and also
74 other ribosome-associated proteins: Gcn20 is involved in sensing starvation by the presence of
75 uncharged tRNAs on the eukaryotic ribosome [12]; ABC50 (ABCF1) promotes translation initiation
76 in eukaryotes [13]; and both Arb1 (ABCF2) and New1 have been proposed to be involved in
77 biogenesis of the eukaryotic ribosome [14, 15]. Ribosome-binding by ABCF proteins seemed to be
78 limited to eukaryotes until characterisation of ABCF member EttA (energy-dependent translational
79 throttle A) found in diverse bacteria. *Escherichia coli* EttA binds to the E-site of the ribosome where
80 it is proposed to 'throttle' the elongation stage of translation in response to change in the

intercellular ATP/ADP ratio [16, 17]. EttA is one of four *E. coli* ABCFs, the others being YheS, Uup and YbiT, none of which have yet been shown to operate on the ribosome [16].

Bacterial ABCF family members have been found to confer resistance to ribosome-inhibiting antibiotics widely used in clinical practise, such as ketolides [18], lincosamides [19-22], macrolides [23, 24], oxazolidinones [25], phenicols [25], pleuromutilins [22] and streptogramins A [22, 26] and B [24]. These antibiotic resistance (AREs) ABCFs have been identified in antibiotic-resistant clinical isolates of *Staphylococcus*, *Streptomyces* and *Enterococcus* among others [27]. This includes the so-called ESKAPE pathogens *Enterococcus faecium* and *Staphylococcus aureus* that contribute to a substantial proportion of hospital-acquired multidrug-resistant infections [28]. As some efflux pumps carry the ABC ATPase domain, it was originally thought that ARE ABCFs similarly confer resistance by expelling antibiotics. However, as they do not carry the necessary transmembrane domains, this is unlikely ~~[29, 30]~~. In support of this, it was recently shown that *Staphylococcus aureus* ARE VgaA protects protein synthesis activity in cell lysates from antibiotic inhibition and that *Enterococcus faecalis* ARE LsaA displaces radioactive lincomycin from *S. aureus* ribosomes [31]. Using a reconstituted biochemical system, we have shown that VgaA and LsaA directly protect the ribosome peptidyl transferase centre (PTC) from antibiotics in an ATP-dependent manner [32]. Recent cryo-EM structures of AREs *Pseudomonas aeruginosa* MsrE and *Bacillus subtilis* VmlR on the ribosome show that like EttA, these ABCFs bind to the E-site of the ribosome, with extended inter-ABC domain linkers protruding into the PTC [33, 34]. The ARE ABCFs therefore appear to either physically interact with the drug to displace it from the ribosome, or allosterically induce a change in conformation of the ribosome that ultimately leads to drug drop-off [35-37].

Here, we carry out an in-depth survey of the diversity of ABCFs across all species with sequenced genomes. We find 45 groups (15 in eukaryotes and 30 in bacteria), including 7 groups of AREs. So-called EQ_2 mutations, double glutamic acid to glutamine substitutions in the two ATPase active sites of EttA lock the enzyme on ribosome in an ATP-bound conformation, inhibiting protein

synthesis and cellular growth [16, 17]. We have tested the effect of equivalent mutations in the other three *E. coli* ABCFs – YbiT, YheS and Uup – as well as the *Bacillus subtilis* ARE VmlR. We establish genetic associations of *E. coli* ABCFs YheS and Uup with the translational GTPase BipA (also known as TypA), and through microscopy and polysome profile analyses, confirm that VmlR does not confer lincomycin resistance through acting as a membrane-bound pump, but via direct interaction with cytoplasmic ribosomes.

Results

ABCFs are widespread among bacteria and eukaryotes

To identify candidate subfamilies of ABCFs and refine the classifications an iterative bioinformatic protocol of sequence searching and phylogenetic analysis was applied. Sequence searching was carried out against a local database of 4505 genomes from across the tree of life. To first get an overview of the breadth of diversity of ABCFs across life, sequence searching began with a local BlastP search against a translated coding sequence database limited by taxonomy to one representative per class, or order if there was no information on class for that species in the NCBI taxonomy database. From phylogenetic analysis of the hits, preliminary groups were identified and extracted to make Hidden Markov Models (HMMs) for further sequence searching and classification. Additional sequences from known ABCF AREs were included in phylogenetic analyses in order to identify groups of ARE-like ABCFs. HMM searching was carried out at the genus level followed by phylogenetic analysis to refine subfamily identification, with final predictions made at the species level. The resulting classification of 16848 homologous sequences comprises 45 subfamilies, 15 in eukaryotes and 30 in bacteria. Phylogenetic analysis of representatives across the diversity of ABCFs shows a roughly bipartite structure, with most eukaryotic sequences being excluded from those of bacteria with strong support (fig. 1). Five eukaryotic groups that fall in the bacterial part of the tree are likely to be endosymbiotic in origin (see the section *Bacteria-like eukaryotic ABCFs*, below).

137 ABCFs are widespread among bacteria and eukaryotes; there are on average four ABCFs per
138 bacterial genome, and five per eukaryotic genome. However, there is considerable variation in how
139 widespread each subfamily is (table 1). The presence of all subfamilies in each genome considered
140 here is shown in table S1, with the full set of sequence IDs and domain composition recorded in
141 table S2. Domain coordinates by amino acid position can be found in table S3. In bacteria,
142 Actinobacteria, and Firmicutes are the phyla with the largest numbers of ABCFs (up to 11 per
143 genome), due to expansions in ARE and potential novel ARE ABCF subfamilies. Among eukaryotes
144 plants and algae encode the most subfamilies, probably due in part to gene acquisition from
145 endosymbiosis events. The diatom *Fragilariopsis cylindrus* has 30 ABCFs and the Haptophyte
146 *Emiliania huxleyi* has 26. Bacterial contamination can sometimes inflate the number of genes in
147 eukaryotic genomes as noted previously for trGTPases [35]. However, as all the *Fragilariopsis* and
148 *Emiliania* sequences belong to typically eukaryotic subgroups, they do not appear to be the result
149 of bacterial contamination. The Tibetan antelope *Pantholops hodgsonii*, on the other hand has 25
150 ABCFs, 20 of which belong to bacterial subgroups. The genome is known to be contaminated by
151 *Bradyrhizobium* sequences [38], and thus the bacteria-like hits from *P. hodgsonii* are most likely
152 artifacts.
153

154 **Table 1 | The subfamilies of the ABCF family, and the numbers (N) of phyla and**
155 **species in which they are encoded**

Subfamily	N Phyla	N Species	Notes on function, relationships and taxonomic distribution
YdiF	42	1852	Broad distribution in bacteria; polyphyletic
Uup [a]	24	3104	Broad distribution in bacteria; paraphyletic to EttA. Inc. P resistance TaeA [39]
EttA	18	2337	Broad distribution in bacteria; translation factor
YbiT	15	1874	Broad distribution in bacteria; potential translation factor
BAF2	7	305	Proteobacteria, Planctomycetes, Spirochaetes, Actinobacteria, Bacteroidetes, Gemmatimonadetes, Cyanobacteria
ARE1[a]	6	269	M,L,S,P,K resistance, inc. VgaA [21], MsrA [18], MsrE [33]; Firmicutes, Actinobacteria, Spirochaetes, Bacteroidetes, Proteobacteria, Tenericutes
ARE3[a]	5	261	L resistance inc LsaA [20]; Firmicutes, Spirochaetes, Proteobacteria, Fusobacteria, Actinobacteria
DAF1	5	58	Spirochaetes, Proteobacteria, Deferribacteres, Fibrobacteres, Chlamydiae
YfmM	5	587	Firmicutes, Tenericutes, Proteobacteria, Fusobacteria, Bacteroidetes
YheS	5	1234	Proteobacteria, Bacteroidetes, Cyanobacteria, Arthropoda, Elusimicrobia
BAF3	4	111	Bacteroidetes, Proteobacteria, Cyanobacteria, Elusimicrobia
DAF2	4	24	Proteobacteria, Planctomycetes, Spirochaetes, Omnitrophica
ARE2[a]	2	54	Antibiotic resistance inc. VmlR[19]; Firmicutes, Tenericutes
ARE4[a]	2	173	M,M16 resistance inc. CarA [40], srmB [40], tlrC [40]; Actinobacteria, Chloroflexi
ARE5[a]	2	408	L,S resistance inc. VarM [26], LmrC [41]; Actinobacteria, Proteobacteria
BAF1	2	20	Firmicutes, Actinobacteria
PAF1	1	138	Proteobacteria
AAF1	1	313	Actinobacteria
AAF2	1	301	Actinobacteria
AAF4	1	219	Actinobacteria
AAF6	1	649	Actinobacteria
AAF3	1	8	Actinobacteria
AAF5	1	25	Actinobacteria
ARE6[a]	1	8	L,S resistance SalA [22]; Firmicutes
ARE7[a]	1	35	Oxazolidinone resistance OptrA [25], Firmicutes
BdAF1	1	176	Bacteroidetes
DAF3	1	36	Proteobacteria
FAF1	1	37	Firmicutes
FAF2	1	42	Firmicutes
SAF1	1	66	Firmicutes
ABCF2	27	560	Arb1 ribosome biogenesis factor; broadly distributed but lacking in Apicomplexa and Microsporidia
ABCF1	23	376	ABC50 translation initiation factor; found in plants, diverse algae, and opisthokonts excluding fungi
ABCF7	16	131	Found in plants, diverse algae, Alveolata, Excavata and Microsporidia
ABCF3	13	382	Gcn20 starvation response; Opisthokonts
eEF3L	9	83	Diverse algae, Chytridiomycota and choanoflagellates
ABCF4	7	37	Diverse algae and Filozoa
ABCF5	7	105	Diverse algae and fungi
cpYdiF	7	85	Diverse algae
algAF1	6	25	Diverse algae
cpEttA	6	18	Chloroplast targeting peptides predicted; plants and diverse algae
algUup	5	17	Found in diverse algae
ABCF6	5	17	Found in diverse algae and Amoebozoa
mEttA	3	14	mitochondrial targeting peptides predicted; diverse algae, Amoebozoa
New1	3	127	Fungi
eEF3	2	138	Translation factor; fungi

Notes - [a] Subfamilies containing known AREs; resistance to antibiotic class in notes column is as follows: M, 14- and 15-membered ring macrolides; M16, 16-membered ring macrolides; L, lincosamides; S, streptogramins; K, ketolides; P, pleuromutilins

156
157

158 The wide distribution and multi-copy nature of ABCFs suggests an importance of these proteins.
159 However they are not completely universal, and are absent in almost all Archaea. The
160 euryarcheaotes *Candidatus Methanomassiliicoccus intestinalis*, *Methanomethylophilus alvus*,
161 *Methanomassiliicoccus luminyensis*, and *Thermoplasmatales archaeon BRNA1* are the only archaea
162 found to encode ABCFs, in each case YdiF. ABCFs are also lacking in 214 bacterial species from
163 various phyla, including many endosymbionts. However, the only phylum that is totally lacking
164 ABCFs is Aquificae. ABCFs are almost universal in Eukaryotes; the only genomes where they were
165 not detected were those of Basidomycete *Postia placenta*, Microsporidium *Enterocytozoon bieneusi*,
166 and apicomplexan genera *Theileria, Babesia* and *Eimeria*.
167
168 Many of the phylogenetic relationships among subfamilies of ABCFs are poorly resolved (fig. 1).
169 This is not surprising, since these represent very deep bacterial relationships that predate the
170 diversification of major phyla and include gene duplication, and differential diversification and
171 loss, and likely combine both vertical inheritance and horizontal gene transfer. Although
172 relationships can not be resolved among all subfamilies, some deep relationships do have strong
173 support (i.e. maximum likelihood bootstrap percentage (MLB) of more than 85% and Bayesian
174 inference posterior probability (BIPP) of 1.0); EttA and Uup share a common ancestor to the
175 exclusion of other ABCFs with full support (100% MLB, 100% UFB and 1.0 BIPP fig. 1, fig. S1) and
176 YheS is the closest bacterial group to the eukaryotic ABCFs with strong support (94% MLB, 98%
177 UFB, and 1.0 BIPP; (fig. 1, fig. S1, text S1). This latter observation suggests that eukaryotic-like
178 ABCFs evolved from within the diversity of bacterial ABCFs. However, this depends on the root of
179 the ABCF family tree. To address this, phylogenetic analysis was carried out of all ABCFs from *E.*
180 *coli*, *Homo sapiens*, *S. cerevisiae* and *B. subtilis*, along with ABCE family sequences from the UniProt
181 database [42]. Rooting with ABCE does not provide statistical support for a particular group at the
182 base of the tree, but does support the eukaryotic subgroups being nested within bacteria, with
183 YheS as the closest bacterial group to the eukaryotic types (fig. 2). It also shows that eEF3 and
184 New1 are nested within the rest of the ABCF family, thus confirming their identity as ABCFs,
185 despite their unusual domain structure. To address the possibility that due to recombination the

186 two ABC domains of the ABCF family may have had different evolutionary histories, we repeated
187 our phylogenetic analysis of representative sequences with the ABC domains uncoupled and
188 aligned to each other (fig. 1, Text S1). With this very short alignment (204 positions) containing a
189 larger proportion of almost invariant active site residues, there is even less statistical support for
190 relationships among subgroups. Nevertheless, we still retain the branches that are well supported
191 in our full-length analyses (fig. 1), and thus there is no evidence for recombination.
192
193 **Domain architectures in the ABCF family are variable**
194
195 To assess the conservation of domains across the family tree of ABCFs, we extracted the domain
196 regions from subfamily alignments, made HMMs representing each domain region and scanned
197 every sequence in our database. We find the most common domain and subdomain structure is an
198 N-terminal ABC1 nucleotide binding domain (NBD) containing an internal Arm subdomain,
199 followed by the Linker region joining to the ABC2 NBD (fig. 3A). Variations on this basic structure
200 include deletions in the Arm and Linker regions, insertion of a Chromo subdomain in the ABC2
201 NBD, and extension of N and C termini by sequence extensions (fig. 3A). The domain structures of
202 eukaryotic ABCFs are more diverse than those of bacteria, with greater capacity for extensions of
203 the N-terminal regions to create new domains (fig. 3A). An increased propensity to evolve
204 extensions, especially at the N terminus is also seen with eukaryotic members of the trGTPase
205 family [35]. In bacteria, terminal extensions of ABCFs tend to be at the C terminus (fig. 3A-B).
206
207 Cryo-electron microscopy structures show that bacterial ABCFs bind to the same site of the
208 ribosome, which is different from that of eEF3 [2, 17, 33, 34], and this is reflected in their domain
209 architectures. eEF3 carries additional N-terminal domains and does not have the Arm subdomain
210 that in EttA binds the L1 stalk and ribosomal protein L1. Instead it carries a Chromo (Chromatin
211 Organization Modifier) subdomain in ABC2 that – from a different orientation to the Arm domain
212 of EttA - interacts with, and stabilises the conformation of the L1 stalk [2]. The Arm subdomain is
213 also missing in eEF3-like close relatives New1 and eEF3L. Although the Arm is widespread in

214 bacterial ABCFs, it is not universal; it is greatly reduced in a number of subfamilies, with the most
215 drastic loss seen in ARE1 and 2 (figs. 3A, and see *Putative AREs* section, below). Subdomain
216 composition can even vary within subfamilies; Uup has lost its arm independently in multiple
217 lineages (fig. 1, fig. S1).
218
219 Arms, linkers and CTD extensions are poorly conserved at the primary sequence level, but are
220 similar in terms of composition, all being rich in charged amino acids, particularly arginine and
221 lysine. Their variable presence and length suggests they can be readily extended or reduced during
222 evolution. The CTD of Uup forms a coiled coil structure that is capable of binding DNA [43].
223 However, whether DNA binding is its primary function is unclear. The CTD of YheS has significant
224 sequence similarity (E value 3.58e-03) to the tRNA binding CTD of Valine-tRNA synthetase (NCBI
225 conserved domains database accession cl11104). *In silico* coiled coil prediction suggests there is a
226 propensity of all these regions to form coiled coil structures (fig. 3B). Extensions and truncations of
227 the arms, linkers and CTD extensions possibly modulate the length of coiled coil protrusions that
228 extend from the globular mass of the protein.
229
230 **Eukaryotic ABCFs comprise 15 subfamilies**
231
232 *eEF3, New1 and eEF3L*
233 eEF3, eEF3L and New1 group together and are particularly distinct members of the ABCF family.
234 The eEF3 subfamily represents the classical fungal proteins, while eEF3L is a more divergent group
235 found mainly in protists (see below). eEF3 has a recent paralogue in *S. cerevisiae* (Hef3 (fig. 2),
236 YEF3B) that apparently arose as a result of the whole genome duplication in yeast [44, 45]. The
237 conservation of domain structure in eEF3/New1/eEF3L suggests they bind the ribosome similarly
238 (fig. 3A). Ribosome binding by eEF3 involves the Chromo subdomain, and the HEAT domain [2],
239 which in addition to New1 and eEF3L, is also found in the protein Gcn1, a binding partner of Gcn20
240 ABCF [46] (see the section *ABCF1-7*, below). eEF3 has a distinct C-terminal extension (fig. 3A),
241 through which it interacts with eEF1A [47]. It has been suggested that this leads to the recruitment

242 of eEF1A to *S. cerevisiae* ribosomes [47, 48]. The New1 CTD contains a region of sequence
243 similarity to the eEF3 CTD; both contain a polylysine/arginine-rich tract of 20-25 amino acids (fig.
244 S2). In *S. cerevisiae* eEF3, this is at positions 1009-1031, which falls within the eEF1A-binding site.
245 Thus eEF1A binding may be a common feature of eEF3, New1, and also eEF3L, which commonly
246 includes an eEF3-like C-terminal extension (table S2).
247
248 We find eEF3-like (eEF3L) factors in a range of eukaryotes including choanoflagellates,
249 haptophytes, heterokonts, dinoflagellates, cryptophytes and red and green algae. This suggests the
250 progenitor of eEF3 was an ancient protein within eukaryotes, and has been lost in a number of
251 taxonomic lineages. Alternatively, eEF3L may have been horizontally transferred in eukaryotes.
252 eEF3 is found in *Chlorella* viruses [49] and *Phaeocystis* viruses (fig. S3), suggesting this may be a
253 medium of transfer. There is no "smoking gun" for viral-mediated transfer in the phylogenies, in
254 that eukaryotic eEF3L sequences do not nest within viral sequence clades (fig. S3). However, given
255 the close association of viral and protist eEF3L, this still remains a possibility. Curiously, the
256 taxonomic distribution of eEF3L in diverse and distantly related protists is similar to that of the
257 unusual elongation factor 1 (eEF1A) paralogue EFL [50] (table S4). The propensity for
258 eEF3/eEF3L/New1 to be present in EFL-encoding organisms and absent in eEF1A-encoding
259 organisms is significant at the level of $P<0.0001$ with Fisher's exact test. Like eEF3L, EFL can be
260 found in viruses such as *Aureococcus anophagefferens* virus (NCBI protein accession
261 YP_009052194.1). As eEF3 interacts with eEF1A [48], the equivalents eEF3L and EFL may also
262 interact in the organisms that encode them.
263
264 Like eEF3, New1 is found across the fungal tree of life (table S1, fig. S3). *S. cerevisiae* New1 has
265 previously been reported to carry a prion-like Y/N/Q/G repetitive region in the N-terminal region
266 before the HEAT domain [51]. We find that this region is limited in taxonomic distribution to
267 Saccharomycetale yeast (table S2, fig. 3A and fig. S2). Thus it is not found in the N-terminal region
268 of *Schizosaccharomyces pombe* New1 (also known as Elf1).
269

270 *ABCF1-7*

271 ABCF1-7 comprise the "ancestral-type" eukaryotic ABCFs, in that they have the typical ABC domain
272 structure that is seen in bacterial ABCFs, and they lack the Chromo and HEAT domains that are
273 found in eEF3, New1 and eEF3L (fig. 3A). All of the terminal extensions found in ABCF subfamilies
274 are biased towards charged amino acids, often present as repeated motifs. The ABCF1 (ABC50)
275 NTD (which interacts with eIF2 [52]) is, due to its length and number of repeats, one of the most
276 striking (fig. S2). It contains multiple tracts of poly-lysine/arginine, poly–glutamic acid/aspartic
277 acid and – in animals - poly-glutamine. ABCF1 and ABCF2 have moderate support as sister groups
278 (fig. 1), and both have representatives in all eukaryotic superphyla, but are not universal. Notable
279 absences of ABCF1 are fungi, amoebozoa and most Aves (birds) (table 1, table S1).

280

281 ABCF2 (Arb1) is essential in yeast and its disruption leads to abnormal ribosome assembly [14],
282 and human ABCF2 can complement an Arb1 deletion, suggesting conservation of function [53]. The
283 protein is broadly distributed across eukaryotes, with the notable exception of the *Alveolata*
284 superphylum (Table S1). Like other ABCF terminal extensions, the ABCF2 N-terminal domain is
285 rich in lysine, in this case lysine and alanine repeats.

286

287 In yeast, where it is known as Gcn20, ABCF3 is a component of the general amino acid control
288 (GAAC) response to amino acid starvation, acting in a complex with Gcn1 [54]. Gcn20 binds to
289 Gcn1 via the latter's HEAT-containing N-terminal domain [46], an interaction that is conserved in
290 ABCF3 and Gcn1 of *Caenorhabditis elegans* [55]. As the HEAT domain is also found in eEF3, this
291 raises the possibility that Gcn20/ABCF3 and eEF3 interact in encoding organisms (fig. 3A).
292 Possible support for this comes from the observation that eEF3 overexpression impairs Gcn2
293 activation [56].

294

295 ABCF3 is widespread in eukaryotes, but absent in heterkont algae and archaeplastida. However,
296 these taxa encode ABCF4 and ABCF7 of unknown function, which have N-terminal domains
297 homologous to ABCF3. Thus, ABCF4 and ABCF7 may be the functional equivalents of ABCF3 in

298 these taxa, potentially interacting with HEAT domain-containing eEF3L in organisms that encode
299 the latter (fig. 3A). ABCF3 from four fungi (*Setosphaeria turcica,* NCBI protein accession number
300 XP_008030281.1; *Cochliobolus sativus,* XP_007703000.1; *Bipolaris oryzae,* XP_007692076.1; and
301 *Pyrenophora teres,* XP_003306113.1) are fused to a protein with sequence similarity to WHI2, an
302 activator of the yeast general stress response [57].
303
304 ABCF5 is a monophyletic group limited to fungi and green algae (*Volvox* and *Chlamydomonas*), with
305 a specific NTD and CTD (fig. 1 and 2A). ABCF5 is found in a variety of Ascomycete and
306 Basidiomycete fungi, including the yeast *Debaryomyces hansenii*, but is absent in yeasts *S. pombe*
307 and *S. cerevisiae*. ABCF4 is a polyphyletic group of various algal and amoeba protists that can not
308 be assigned to the ABCF5, or eEF3/eEF3L/New1 clades. In eukaryotic ABCF-specific phylogenetic
309 analysis, ABCF4/ABCF5/eEF3/eEF3L/New1 are separated from all other eukaryotic ABCFs with
310 moderate support (85% MLB fig. S3). ABCF6 represents a collection of algal and amoebal
311 sequences that associate with the ABCF1+ABCF2 clade with mixed support in phylogenetic
312 analysis (fig. 1).
313
314 *Bacterial-like eukaryotic ABCFs*
315 Five eukaryotic subfamilies are found in the bacteria-like subtree: algAF1, algUup, mEttA, algEttA
316 and cpYdiF (fig. 1). Given their affiliation with bacterial groups, they may have entered the cell with
317 an endosymbiotic ancestor. Indeed, chloroplast-targeting peptides are predicted at the N termini of
318 the majority of cpYdiF sequences, and mitochondrial localization peptides at most mEttA N
319 termini. The situation is less clear for the three remaining groups, with a mix of signal peptides, or
320 none at all being predicted across the group members (table S5).
321
322 **Bacterial ABCFs comprise 30 subfamilies, most of which have unknown function**
323
324 There are 30 groups of bacterial ABCFs, the most broadly distributed being YdiF (the subfamily is
325 given the name of the *Bacillus subtilis* protein as it is not present in *E. coli*) (table 1). This subfamily

326 is a paraphyletic grouping comprising ABCF sequences that can not confidently be classified into
327 any of the other subgroups (fig. 1). The next most broadly distributed group is Uup (*B. subtilis*
328 protein name YfmR), which itself is paraphyletic to EttA. *B. subtilis* does not encode EttA, but does
329 encode YbiT (*B. subtilis* name YkpA). It also encodes two ABCFs not present in *E. coli*: YfmM and
330 VmlR (also known as ExpZ). VmlR is in the ARE2 subfamily, and confers resistance to
331 virginiamycin M1 and lincomycin [19]. Insertional disruptants of all the chromosomal ABCF genes
332 in *B subtilis* strain 168 have been examined for resistance to a panel of nine MLS class antibiotics,
333 and only VmlR showed any hypersensitivity [19]. With the exception of EttA [16, 17], and the
334 seven antibiotic resistance ARE ABCFs (table 1), the biological roles of the other 22 bacterial
335 ABCFs are largely obscure.
336
337 B. subtilis *ARE VmlR is a cytoplasmic protein that directly protects the ribosome from antibiotics*
338 *B. subtilis* virginiamycin M and lincomycin resistance factor ABCF VmlR was originally annotated as
339 an ABC efflux transporter, i.e. a membrane protein [19, 58]. To probe VmlR's interaction with the
340 ribosome we took advantage of ATPase-deficient VmlR mutants generated by simultaneous
341 mutation of both glutamate residues for glutamine (EQ_2) [59] that lock ABC enzymes in an ATP-
342 bound active conformation [17, 60]. In the case of EttA, expression of the EQ_2 mutant results in a
343 dominant-negative phenotype as EttA incapable of ATP hydrolysis acts as a potent inhibitor of
344 protein synthesis and, consequently, bacterial growth [16]. We constructed C-terminally tagged
345 His_6-TEV-3xFLAG-tagged (HTF-tagged) wild type and EQ_2 (*vmlR*-HTF and *vmlR*EQ_2-HTF) under the
346 control of an IPTG-inducible $P_{hy\text{-}spank}$ promotor [61]. To probe the intracellular localization of VmlR,
347 we C-terminally tagged VmlR with the mNeonGreen fluorescent protein [62] under the control of
348 xylose inducible promoter protein P_{xyl} [63]. We have validated the functionality of the fusion
349 constructs by lincomycin resistance assays using a Δ*vmlR* knock-out strain as a negative control
350 and a Δ*vmlR* knock-out strain expressing untagged VmlR under the control of an IPTG-inducible
351 $P_{hy\text{-}spank}$ promotor as a positive control (fig. S4A-C). While C-terminal tagging with either HTF or
352 mNeonGreen does not abolish VmlR's activity, the EQ_2 versions of the tagged proteins are unable
353 to protect from lincomycin (fig. S4B).

354

355 After establishing the functionality of the tagged VmlR constructs, we tested the effects of
356 expression of either wild type or EQ_2 VmlR-HTF on *B. subtilis* growth in rich LB media (fig. 4A).
357 Expression of the wild type protein in the Δ*vmlR* background has no detectable effect. In contrast,
358 the EQ_2 version inhibits growth: while exponential growth is unaffected, the cells enter the
359 stationary phase at lower cell densities, abruptly stopping growth instead of slowing down
360 gradually (fig. 4A). Two factors are likely to cause the growth-phase specificity of the inhibitory
361 effect. First, during the exponential growth cells efficiently dilute the toxic protein via cell division,
362 and when the growth slows down, VmlR-EQ_2-HTF accumulates. Second, upon entering the early
363 stationary phase, *B. subtilis* sequesters 70S ribosomes into inactive 100S dimers [64], and this
364 decrease of active ribosome concentration could conceivably render the cells more vulnerable to
365 the inhibitory effects of VmlR-EQ_2. We probed the interaction of wild type and EQ_2 VmlR-HTF with
366 ribosomes using polysome analysis in sucrose gradients in combination with Western blotting (fig.
367 4B). While the wild type protein barely enters the gradient (most likely dissociating from the
368 ribosomes during centrifugation), the EQ_2 version almost exclusively co-localises with the 70S
369 peak fraction and is absent from the polysomal fractions (fig. 4B), suggesting co-sedimentation of a
370 tight 70S:VmlR-EQ_2 complex.

371

372 Finally, having ascertained the functionality of VmlR-mNeonGreen (fig. S4C), we imaged *B. subtilis*
373 cells expressing VmlR-mNeonGreen in the presence and absence of lincomycin (fig. 4C) and
374 quantified the intensity of the fluorescent signal across the cell (fig. 4D). As a positive control for
375 membrane localization we used WALP23-GFP [65] – an artificial model transmembrane helix
376 WALP23 [66] fused with an N-terminal GFP label. We observe no evidence for association of VmlR
377 with the membrane: the protein is clearly cytoplasmic, with a slight exclusion from the nucleoid in
378 the presence of 5 μg/mL lincomycin. A likely explanation for this effect is the nucleoid compaction
379 caused by inhibition of translation resulting in protein exclusion from the nucleoid-occupied space
380 [67-69]. However, we cannot rule out that this general effect potentiated by specific interaction of
381 VmlR with strongly nucleoid-excluded ribosomes.

382

383 E. coli *ABCFs EttA, YbiT, YheS and Uup interact genetically and functionally with protein synthesis*
384 *and ribosome assembly*

385 *E. coli* encodes four ABCFs: EttA, YbiT, YheS and Uup. An array of structural, biochemical and
386 microbiological methods has been used to establish that EttA operates on the ribosome [16, 17].
387 Ribosomal association of the other *E. coli* ABCFs has not been shown. However, a recent PhD thesis
388 by Dr. Katharyn L. Cochrane suggests that Uup genetically interacts with an enigmatic ribosome-
389 associated factor, the translational GTPase BipA (TypA) [70]. The *E. coli bipA* knock-out strain is
390 characterised by a decreased level of 50S subunits accompanied by an accumulation of pre-50S
391 particles [71]. In the presence of its native substrate, GTP, BipA associates with mature 70S
392 ribosomes [72], occupying the ribosomal A-site [73]. However, in the presence of stress alarmone
393 (p)ppGpp – a molecular mediator of the stringent response [74] – BipA binds the 30S subunit [75].

394

395 We have set out to systematically probe the involvement of *E. coli* ABCFs in protein synthesis. We
396 used two experimental systems. The first is geared towards low level constitutive expression of
397 native, untagged wild type and EQ_2 proteins in a clinically relevant uropathogenic *E. coli* strain
398 CFT073 [76]. For this, we cloned ABCF genes into a low copy pSC101 vector under control of a
399 constitutive tet-promoter (P_{tet}) that in the original plasmid drives expression of the tetracycline
400 efflux pump TcR [77]. Using the λRed-mediated gene disruption method [78] we generated a set of
401 mutants of lacking each of the four ABCF genes, as well as a Δ*bipA* and Δ*bipA*Δ*uup* knock-out
402 strains. The second system allows inducible high-level expression of tagged proteins in the
403 avirulent BW25113 *E. coli* strain [79, 80]. We used wild type and EQ_2 mutants of EttA, YbiT, YheS
404 and Uup with N-terminal FLAG-TEV-His_6 (FTH)-tags expressed from a low copy pBAD18 plasmid
405 under an arabinose-inducible araBAD (P_{BAD}) promoter. The BW25113 strain can not metabolise
406 arabinose (Δ*(araD-araB)567* genotype) [79], and therefore the inducer is not metabolised during
407 the experiment.

408

409 First we tested the genetic interactions between *bipA* and all *E. coli* ABCFs in CFT073 background.
410 At 37°C the *bipA* CFT073 knock-out strain has no growth defect (fig. S5A), however at 18°C, the
411 Δ*bipA* strain displays a pronounced growth defect characteristic of strains defective in ribosome
412 assembly (fig. 5A) [81]. Ectopic expression of Uup efficiently suppresses the growth defect, while
413 deletion of *uup* in the *bipA* background exacerbates it (fig. 5A). Expression of EttA and Ybit have no
414 effect, but expression of YheS leads to a dramatic growth defect. Importantly, in the wild type
415 background, the expression of YheS has no effect on growth at 18°C (fig. S5B), indicating that the
416 genetic interaction between *bipA* and *yheS* is specific. As reported previously, disruption of *bipA*
417 leads to a dramatic ribosome assembly defect at low (18°C) temperature [71] (fig. 5B). The levels
418 of mature 70S ribosomes as well as 50S subunits are dramatically decreased, accompanied by an
419 accumulation of 50S assembly precursors (the peak marked with an asterisk) and free 30S
420 subunits. Ectopic expression of Uup partially suppresses these defects, and in the Δ*bipA*Δ*uup* strain
421 the defects are exacerbated. All of the effects described above are conditional on disruption of *bipA*
422 since neither disruption of individual ABCF genes nor simultaneous disruption of *uup* and *ettA* –
423 the only well-characterised ribosome-associated *E. coli* ABCF to date – causes cold-sensitivity (fig.
424 S5C) or affects polysome profiles (fig. S5D).
425
426 Next we set out to test the effects of EQ_2 versions of ABCFs on translation. We validated the
427 expression of the FTH-tagged ABCFs using Western blotting (fig. S6A). As was observed for
428 untagged Uup (fig. 5A), the expression of FTH-tagged Uup suppresses the cold sensitivity caused
429 by *bipA* deletion while YheS expression exacerbates the growth defect (fig. S6B-C). Expression of
430 the EQ_2 versions universally causes growth inhibition, both at 18°C in *ΔbipA* CFT073 (fig. S6C) and
431 at 37°C in the wild type BW25113 background (fig. 6). Overexpression of none of the wild type
432 ABCFs results in a growth defect (fig. 6A-D). Next we used a ^{35}S-methionine pulse-labeling assay as
433 a readout of translational inhibition. For all the ABCF-EQ_2s, the methionine incorporation
434 decreases, showing protein synthesis is clearly inhibited. The strongest effect is observed for EttA-
435 EQ_2 (fig. 6A) and YbiT-EQ_2 (fig. 6C), and the weakest is seen in for Uup-EQ_2 (fig. 6B).
436

437 *Phylogenetic analysis reveals putative AREs*

438 Through phylogenetic analysis of predicted proteins and previously documented AREs with
439 sequences available in UniProt [42] and the Comprehensive Antibiotic Resistance Database (CARD)
440 [82], we have identified seven groups of AREs (fig. 1, table 1). Surprisingly, these can be quite
441 variable in their subdomain architecture (fig. 7A). Some AREs (ARE1-5) have experienced
442 extension of the Linker by on average around 30 amino acids compared to EttA, which is in line
443 with the observation that the extended Linker is in close contact with the bound antibiotic in the
444 case of ARE1 MsrE [33]. However, linker extension is not the rule for AREs; ARE7 (OptrA) has a
445 linker of comparable length to the *E. coli* ABCFs (fig. 7A). This supports the notion based on the
446 VmlR –ribosome co-structure that antibiotic protection by ABCFs can involve allosteric changes in
447 the ribosome as well as through direct interaction [34]. The Arm subdomain that in EttA interacts
448 with the L1 ribosomal protein and the L1 stalk rRNA (fig. 7A) varies in length among AREs, and the
449 CTD extension may or not be present (fig. 7B). Surprisingly, the Uup protein from cave bacterium
450 *Paenibacillus sp. LC231*, a sequence that is unremarkable among Uups, confers resistance to the
451 pleuromutilin antibiotic tiamulin [39] (fig. S1).

452

453 Actinobacteria are the source of many ribosome-targeting antibiotics [83], and they have evolved
454 measures to protect their own ribosomes, including ARE4 (OleB) and ARE5 (VarM). In addition to
455 these known AREs, Actinobacteria encode a number of other ABCFs specific to this phylum (AAF1-
456 6; table 1). It is possible that some – if not all – of these groups are in fact AREs. Other subfamilies
457 may also be unidentified AREs, but two particularly strong candidates are BAF2 and BAF3 that
458 have strong support for association with ARE5 (fig. 1) and are found in a wide range of bacteria
459 (table 1). PAF1 is also worthy of investigation; it is found in the genomes of several pathogens in
460 the proteobacterial genera *Vibrio*, *Enterobacter*, *Klebsiella*, *Serratia* and *Citrobacter*, but not
461 *Escherichia* (table S1). With the exception of antibiotic producers, known antibiotic resistance
462 ABCFs tend to have a variable presence within genera (fig. S7), probably because they are
463 frequently transferred by mobile elements such as plasmids and transposons. Therefore,
464 variability in the presence of a subfamily across species in a genus can be an indication that an

465 ABCF is an ARE. Taking into account phylogenetic relationships (fig. 1, fig. S1) and disjunction
466 within genera (fig. S7), we predict the following novel AREs: AAF1-5 (which tend to be found in
467 antibiotic producers), BAF1-3, FAF1-2, and PAF1.
468
469 *ABCFs are polyproline-rich proteins*
470 Curiously, we find that ABCFs from both bacteria and eukaryotes are often rich in polyproline
471 sequences, which are known to cause ribosome slow-down or stalling during translation [84]. This
472 stalling is alleviated by the elongation factor EF-P in bacteria [85, 86], and indeed EF-P is required
473 for full expression of EttA, which contains two XPPX motifs [87]. 56% of all the sequences in our
474 ABCF database contain at least two consecutive prolines, compared to an overall 37% of all the
475 proteins in the predicted proteomes considered here. There is a particularly proline-rich hotspot
476 in the C-terminal part of the linker (fig. 7B), which can be up to nine consecutive prolines long in
477 the case of Uup from *Novosphingobium aromaticivorans* (NCBI protein accession number
478 WP_028641352.1). Arms and CTD extensions are also polyproline hotspots; PPP is a common
479 motif in the Arm of EttA proteins, and polyprolines are frequent in YdiF, Uup and YheS CTD
480 extensions. Prolines are rigid amino acids, and conceivably their presence may support the tertiary
481 structures of ABCFs, particularly the orientation of the subdomain coiled coils [88].
482
483 **Discussion**
484
485 *Towards a general model for non-eEF3 ABCF function*
486 In the case of ABCFs that act on the assembled ribosome during translation, the E-site should be
487 vacant (i.e. not filled by an E-site tRNA) for the protein to bind. Specifically, this would be when the
488 E-site has not yet received a tRNA (during initiation, where ABCF50 (ABCF1) functions [13]) or
489 when an empty tRNA has dissociated and not been replaced (during slow or stalled translation
490 such as in the presence of an antibiotic, as in the case of AREs and EttA [16, 17, 34]). EttA has been
491 proposed to promote the first peptide bond after initiation through modulation of the PTC
492 conformation [16, 17]; similarly, allosteric effects acting on the PTC have been observed for the

493 ARE VmlR [34]. This structural modulation or stabilisation could conceivably be a general function
494 of ABCFs, with the specific ribosomal substrate differing depending on the stage of translation,
495 assembly, or cellular conditions. Differences in subdomains would determine both what is sensed
496 and the resulting signal. For instance the presence of the Arm would affect signal transmission
497 between the PTC and the L1 protein and/or the L1 stalk.
498
499 *Evolution of AREs*
500 In order to determine antibiotic resistance capabilities and track the transfer routes of resistance,
501 it is critical to be able to annotate antibiotic resistance genes in genomes. This requires
502 discrimination of antibiotic genes from homologous genes from which resistance functions have
503 evolved. At present this is not straightforward for ABCF AREs, as the distinction between potential
504 translation and antibiotic resistance factors is ambiguous. ABCF AREs have been compared to the
505 Tet family of antibiotic resistance proteins that evolved from trGTPase EF-G to remove tetracycline
506 from the ribosome [29]. However the distinction between Tet proteins and EF-G is much more
507 clear-cut, with Tet comprising a distinct lineage in the evolutionary history of trGTPases [35]. The
508 surprising lack of a clear sequence signature for antibiotic resistance in the ARE ABCFs suggests
509 that antibiotic resistance functions may evolve in multiple ways in ABCFs, and that ABCFs closely
510 related to AREs may have similar functions to the AREs, while not conferring resistance. For
511 example, there are multiple small molecules that bind the PTC and exit tunnel [89], and
512 conceivably ABCFs may be involved in sensing such cases, removing the small molecule, or
513 allowing translation of a subset of mRNAs to continue in its presence. Even macrolide antibiotics
514 that target the exit tunnel do not abrogate protein synthesis entirely, but rather reshape the
515 translational landscape [90, 91].
516
517 **Conclusion**
518
519 ABCFs are stepping into the limelight as important translation, ribosome assembly and antibiotic
520 resistance factors. We have found hydrolysis-incompetent EQ_2 mutants of all four *E. coli* ABCFs

521 inhibit protein synthesis, suggesting they all function on the ribosome. Overexpression of Uup
522 suppresses both the cold-sensitivity and the 50S ribosome assembly defect caused the loss of
523 translational GTPase BipA, suggesting that Uup is involved in the 50S ribosome subunit assembly,
524 either directly or indirectly, for example by fine-tuning expression of ribosomal proteins.
525 Additionally, we have added ARE2 VmlR to the repertoire of AREs confirmed to act on the
526 ribosome, joining the ranks of LsaA, VgaA and MsrE, all ARE1s. Considering the well-established
527 ribosome association of eukaryotic ABCFs, our results suggest that ribosome-binding is a general -
528 perhaps ancestral - feature of ABCFs. However, the ABCF family is diverse, and even within
529 subfamilies, there can be differences in subdomain architecture. We have identified clusters of
530 antibiotic resistance ARE ABCFs, and predicted likely new AREs. Strikingly, the AREs do not form a
531 clear monophyletic group, and either antibiotic resistance has evolved multiple times
532 independently from the ABCF diversity, or this is an innate ability of ABCFs, raising the possibility
533 of a general role of ABCFs in ribosome-binding small molecule sensing and signaling.
534
535 **Methods**
536 **Sequence searching and classification**
537
538 Predicted proteomes were downloaded from the NCBI genome FTP site (2nd December 2014). One
539 representative was downloaded per species of bacteria (i.e. not every strain). Seven additional
540 proteomes were downloaded from JGI (*Aplanochytrium kerguelense, Aurantiochytrium limacinum,*
541 *Fragilariopsis cylindrus, Phytophthora capsici, Phytophthora cinnamomi, Pseudo-nitzschia*
542 *multiseries,* and *Schizochytrium aggregatum*). Previously documented AREs were retrieved from
543 UniProt [42] and the Comprehensive Antibiotic Resistance Database (CARD) [82]. Taxonomy was
544 retrieved from NCBI, and curated manually where some ranks were not available.
545
546 An initial local BlastP search was carried out locally with BLAST+ v 2.2.3 [92] against a proteome
547 database limited by taxonomy to one representative per class or (order if there was no information
548 on class from the NCBI taxonomy database), using EttA as the query. Subsequent sequence

549 searching against the proteome collections used hmmsearch from HMMER 3.1b1, with HMMs
550 made from multiple sequence alignments of subfamilies, as identified below. The E value threshold
551 for hmmsearch was set to $1e^{-70}$, a value at which the subfamily models hit outside of the
552 eukaryotic-like or bacterial-like ABCF bipartitions, ensuring complete coverage while not picking
553 up sequences outside of the ABCF family.
554
555 Sequences were aligned with MAFFT v7.164b (default settings) and Maximum Likelihood
556 phylogenetic analyses were carried out with RAxML-HPC v.8 [93] on the CIPRES Science Gateway
557 v3 [94] using the LG model of substitution, after removing positions containing >50% gaps.
558 Additional phylogenetic analyses of representative sequences were carried out as described in the
559 section "phylogenetic analysis of representatives", below. Trees were visualised with FigTree v.
560 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/) and visually inspected to identify putative
561 subfamilies that preferably satisfied the criteria of 1) containing mostly orthologues, and 2) had at
562 least moderate (>60% bootstrap support). These subfamilies were then isolated, aligned
563 separately and used to make HMM models. Models were refined with subsequent rounds of
564 searching and classification into subfamilies first by comparisons of the E values of HMM hits, then
565 by curating with phylogenetic analysis. After the final classification of all ABCF types in all
566 predicted protein sequences using HMMER, some manual correction was still required. For
567 example, cyanobacterial sequences always hit the chloroplast HMM with a more significant E value
568 than the bacterial model, and eEF3/New1/eEF3L sequences could not be reliably discriminated
569 between using E value comparisons. Therefore, the final classification is a manually curated
570 version of that generated from automatic predictions (table S2). All sequence handling was carried
571 out with bespoke Python scripts, and data was stored in a MySQL database (exported to Excel files
572 for the supplementary material tables). Identification of EFL and eEF1A in predicted proteomes
573 was carried out with HMMER, using the HMMs previously published for these trGTPases [35], The
574 E value cut-of was set to e-200, lenient enough to match both eEF1A or EFL, with assignment to
575 either protein subfamily made by E value comparisons as above.
576

577 **Domain prediction**

578 Domain HMMs were made from subalignments extracted from subfamily alignments. Partial and
579 poorly aligned sequences were excluded from the alignments. All significant domain hits (<E value
580 $1e^{-3}$) for each ABCF sequence were stored in the MySQL database. Sequence logos of domains were
581 created with Skylign [95]. Putative transit peptides for mitochondrial and plastid subcellular
582 localization were predicted with the TargetP web server hosted at the Technical University of
583 Denmark [96].

584

585 **Phylogenetic analysis of representatives**

586

587 For the representative tree of the ABCF family, taxa were selected from the ABCF database to
588 sample broadly across the tree of life, including eukaryotic protistan phyla, while also covering all
589 subfamilies of ABCFs. Sequences were aligned with MAFFT with the L-ins-i strategy [97] and
590 positions with >50% gaps, and several ambiguously aligned positions at the termini were
591 removed. The resulting 249 sequences, and 533 positions were subject to RAxML and IQ-TREE
592 Maximum Likelihood phylogenetic analysis, both run on the CIPRES Science Gateway v3 [94].
593 RAxML was run with the LG substitution matrix, as favoured by ProtTest 2.4 [98] and 100
594 bootstrap replicates. Bootstrapping with RaxML yields a value (maximum likelihood bootstrap
595 percentage, MLB) for how much of the input alignment supports a particular branch in the tree
596 topology, and therefore the reliability of that branch. These support values are indicated on
597 branches in the tree figures. In the case of IQ-TREE, the most appropriate model was selected by
598 the program during the run, which also favoured the LG substitution matrix. IQ-TREE was run with
599 its ultrafast bootstrapping approximation method to ascertain support values (UFB) for branches
600 out of 1000 replicates [99]. To test whether our trees made with ABC domains separately are
601 incompatible (as might indicate recombination), the RaxML and IQ-TREE analyses were repeated
602 with a dataset containing the ABC domains uncoupled from each other and aligned together.
603 Alignments were prepared as above, to make a data set of 204 alignment positions from 525 taxa
604 (Text S1).

605

606 Bayesian inference phylogenetic analysis was carried out with MrBayes v3.2.6, also on the CIPRES
607 gateway. The analysis was run for 1 million generations, after which the standard deviation of split
608 frequencies (SDSF) was 0.08. The mixed model setting was used for determining the amino acid
609 substitution matrix, which converged on WAG. RAxML analysis with the WAG model showed no
610 difference in topology for well-supported branches compared to the RAxML tree with the LG
611 model. Branch support values in this case are posterior probabilities, shown on the tree figure as
612 Bayesian Inference posterior probabilities (BIPP), on a scale of 0 to 1, with increasing probability.

613

614 For the rooted tree with ABCE as the outgroup, all ABCFs were selected from *Arabidopsis thaliana,*
615 *Homo sapiens, Escherichia coli, Bacillus subtilis, Saccharomyces cerevisae, Schizosaccharomyces*
616 *pombe* and *Methanococcus maripaludis*. Ambiguously aligned sites were identified and removed
617 manually. RAxML, IQ-TREE and MrBayes were carried out as above on the resulting 344 positions
618 from 35 sequences. The MrBayes analysis stopped automatically when the SDSF dropped to the
619 0.009 threshold, which was at 235,000 generations.

620

621 For the tree of eukaryotic and viral ABCFs rooted with YheS, sequences were extracted from the
622 ABCF database, and viral sequences were found in the NCBI protein database using BlastP. The
623 resulting 658 sequences were aligned with MAFFT and a RAxML analysis of 645 positions was
624 carried out as above. The alignments used to build the phylogenies presented in the main text as
625 supplementary information are available in text S1, along with trees that are used for ascertaining
626 branch support but not included as figures.

627

628 **Structural analyses**

629 Homology modeling was carried out using Swiss Model [100] with EttA (PDB ID 3J5S) as the
630 template structure. Because the linkers were lacking secondary structure, QUARK [101] was used
631 for *ab initio* structure modeling of these regions, and the resulting coils were aligned back to the
632 homology model using the structural alignment method of MacPyMOL [102]. The presence of

633 coiled coil regions was predicted with the COILS program hosted at the ExPASy Bioinformatics
634 Research Portal [103].
635
636 **Construction of plasmids and bacterial strains**
637 All bacterial strains and plasmids used in this study are described in **Supplementary Methods**
638 and listed in **Table 2**.
639
640 **Table 2. Strains and plasmids used in the study.**

Strain or Plasmid	Description	Reference
Strains: E.coli		
BW	BW25113 *E. coli*	[79]
CFT073	Uropathogenic *E. coli* O6:K2H1	[76]
CFTuup	Δ*uup* (locus tag c1085)	This study
CFTettA	Δ*yjjK* (locus tag c5478)	This study
CFTyheS	Δ*yheS* (locus tag c4127)	This study
CFTybiT	Δ*ybiT* (locus tag c0906)	This study
CFTbipA	Δ*yihK* (locus tag c4820)	This study
CFTbipA_pUup	Δ*yihK* with pSC-uup	This study
Strains: B. subtilis		
B. subtilis 168	*trpC2*	[104]
VHB5	*trpC2* Δ*vmlR*	This study
VHB38	*trpC2* Δ*vmlR amyE*::P_{xyl}-*vmlR-mNeoGreen* Spc^r	This study
VHB44	*trpC2* Δ*vmlR thrC*::$P_{hy\text{-}spnak}$-*vmlR* Kan^r	This study
VHB45	*trpC2* Δ*vmlR thrC*::$P_{hy\text{-}spnak}$-*vmlREQ2* Kan^r	This study
VHB91	*trpC2* Δ*vmlR thrC*::$P_{hy\text{-}spnak}$-*vmlR-HTF* Kan^r	This study
VHB92	*trpC2* Δ*vmlR thrC*::$P_{hy\text{-}spnak}$-*vmlREQ*$_2$*-HTF* Kan^r	This study
HS64	*trpC2 amyE*::P_{xyl}-*WALP23-gfp* Spc^r	[65]
Plasmids		
pKD4	λRed PCR template plasmid with Kan resistance cassette; Kan^r Amp^r	[78]
pKD13	λRed PCR template plasmid with Kan resistance cassette; Kan^r Amp^r	[78]
pKD46	λRed recombinase helper plasmid, temperature sensitive; Amp^r	[78]
pCP20	FLP recombinase encoding plasmid, temperature sensitive; Amp^r	[78]
pSC101	pSC101 empty vector; Kan^r	This study
pSC-ettA	Constitutive *ettA* overexpression plasmid; Kan^r	This study
pSC-uup	Constitutive *uup* overexpression plasmid; Kan^r	This study
pSC-yheS	Constitutive *yheS* overexpression plasmid; Kan^r	This study

pSC-ybiT	Constitutive *ybiT* overexpression plasmid; Kan^r	This study
pSC-bipA	Constitutive *bipA* overexpression plasmid; Kan^r	This study
pHT009	Integration plasmid ; Kan^r $Amp^r Kan^r$	This study
pSG1154	Integration plasmid; Spc^r Amp^r	[63]
pSHP2	Integration plasmid; Spc^r Amp^r	This study
VHp62	pAPNC with vmlR-HTF in SalI/BamHI sites; Spc^r Amp^r	Laboratory stock
VHp66	pAPNC with vmlREQ_2-HTF in SalI/BamHI sites; Spc^r Amp^r	Laboratory stock
pHT009-vmlR	pHT009 with vmlREQ2; Kan^r Amp^r	This study
pHT009-vmlR-HTF	pHT009 with vmlR-HTF in HindIII/SphI site; Kan^r Amp^r	This study
pHT009-vmlREQ2-HTF	pHT009 with vmlREQ2-HTF in HindIII/SphI site; Kan^r Amp^r	This study
pSHP2-vmlR	pSHP2 with vmlR in ApaI/EcoRI site ; Spc^r Amp^r	This study

641

642 **Growth assays**

643 Bacterial growth (OD_{600}) was monitored using a Bioscreen C (Oy Growth Curves Ab Ltd)
644 microplate reader in Honeycomb plates (150 µL culture per well) with continuous shaking (speed:
645 fast, amplitude: normal). All experiments with CFT073 were performed at 18°C unless stated
646 otherwise. Three biological replicates were averaged for each growth curve and the data presented
647 as geometric means ± standard deviation.

648

649 E. coli *transformed with pSC101-based expression plasmids:*

650 Overnight (16 h) cultures were pre-grown in LB medium supplemented with 50 µg/mL
651 kanamycin, diluted to OD_{600} of 0.03 in filtered LB and grown in Bioscreen C microplate reader as
652 described above.

653

654 E. coli *transformed with pBAD-based expression plasmids:*

655 Overnight (16 h) cultures were pre-grown in Neidhardt MOPS medium [105] supplemented with
656 0.1% of casamino acids, 0.4% glucose as a carbon source and 100 µg/mL carbenicillin, diluted to
657 OD_{600} of 0.03 in the same media but containing and 0.5% glycerol instead of 0.4% glucose as well
658 as supplemented with 0.5% arabinose and grown in Bioscreen C microplate reader as described
659 above.

660

661 *Antibiotic resistance testing of tagged* B. subtilis *VmlR*:

662 *B. subtilis* strains VHB38, VHB91 and VHB92 were pre-grown on LB plates overnight at 30˚C. Fresh
663 individual colonies were used to inoculate filtered LB medium, either in the presence and absence
664 of 1 mM IPTG (for VHB91 and VHB92) or in the presence and absence of 0.3% xylose (for VHB38),
665 and OD_{600} adjusted to 0.01. The cultures were seeded on Honeycomb plates, and plates incubated
666 in a Bioscreen C at 37 °C with continuous shaking as described above for *E. coli* cultures. After 90
667 minute incubation ($OD_{600} \approx 0.1$) increasing concentrations of lincomycin (final concentration 0 - 5
668 µg/ml) were added and growth was monitored for additional 6 hours. Three biological replicates
669 were averaged for each growth curve and the data presented as geometric means ± standard
670 deviation.

671

672 **Fluorescence microscopy**

673 Fluorescence microscopy was carried out with cell grown to early-mid logarithmic growth phase
674 (OD_{600} of 0.2-0.5) in LB medium at 37 °C in the presence or absence of inducers. The used inducer
675 concentrations were 0.3% for VmlR-mNG and 1% for WALP23-GFP. If indicated, the cells were
676 incubated with 5µg/ml lincomycin upon shaking at 37 °C prior to the microscopy. The cells were
677 immobilised on microscopy slides covered with a thin film of 1.2% (w/v) agarose in H_2O as
678 described in detail elsewhere [106]. The microscopy was carried out with Nikon Eclipse Ti
679 equipped with Nikon Plan Apo 100x/1.40 Oil Ph3 objective, Sutter Instrument Company Lambda
680 LS xenon arc light source, and Photometrics Prime sCMOS camera. The images were captured
681 using Metamorph 7.7 (Molecular Devices) and analysed using Fiji [107].

682

683 **Western blot analysis of FTH-tagged wt and EQ_2 E. coli ABCF proteins**

684 *Preparation of bacterial samples:*

685 Bacteria were grown either at 37 °C (*E. coli* BW25113 derivatives) or 18°C (*E. coli ΔbipA* CFT073
686 derivatives) up to OD_{600} of 0.5 in 50 mL of Neidhardt MOPS minimal medium [105] supplemented
687 with 0.1% casamino acids (w/v), 0.5% glycerol (w/v) and 100 µg/mL carbenicillin and L-

688 arabinose was added to a final concentration of 0.2% (w/v). Cultures grown at 37°C were
689 harvested 10 minutes after induction by pouring them into precooled centrifuge bottles containing
690 100 g of crushed ice and centrifuged at 10,000 rpm for 10 minutes at 4 °C (Beckman JLA16.250
691 rotor). Cultures grown at 18°C were harvested 5 hours after induction by collecting into precooled
692 centrifuge bottles and pelleting at 10,000 rpm for 10 minutes at 4 °C (Beckman JA25.50 rotor).
693 Lysates were prepared the same as for polysome profiling of *E. coli* (see below).

694 *Western blotting:*

695 3 μg of total protein as determined by Bradford assay of each sample was resolved on 10% SDS-
696 PAGE gel and transferred to 0.2 μm nitrocellulose membrane (Trans-Blot® Turbo™ Transfer Pack,
697 Bio-Rad) using Turbo MIXED MW protocol in Trans-Blot® Turbo™ Transfer System (Bio-Rad). The
698 membrane was blocked in PBS-T (1x PBS 0.05% Tween-20) with 5% w/v nonfat dry milk at room
699 temperature for one hour. Antibody incubations were performed for one hour in 1% nonfat dry
700 milk in PBS-T with five 5-minute washes in fresh PBS-T between and after antibody incubations.
701 FTH-tagged ABCFs were detected using anti-Flag M2 primary (Sigma-Aldrich, F1804; 1:10,000
702 dilution) antibodies combined with anti-mouse-HRP secondary (Rockland; 610-103-040; 1:10,000
703 dilution) antibodies. AECL detection was performed on ImageQuant LAS 4000 (GE Healthcare)
704 imaging system using Pierce® ECL Western blotting substrate (Thermo Scientific).

705

706 **Polysome profiling analysis of E. coli strains**

707 *Preparation of bacterial samples:*

708 Overnight (16 h) cultures were pre-grown at 37°C in LB medium supplemented with 50 μg/mL
709 kanamycin in the case of strains transformed with pSC101-based expression plasmids. Overnight
710 cultures were diluted in filtered LB (33 mL cultures) and after 24 h growth at 18 °C harvested by
711 pouring into precooled centrifuge bottles and pelleting at 10,000 rpm for 10 minutes at 4 °C
712 (Beckman JA25.50 rotor). For the sake of convenience, cultures were diluted to different starting
713 densities (Table S3) to ensure that all of them reach OD_{600} ≈0.5 simultaneously.

714 *Preparation of clarified lysates:*

715 Cell pellets were resuspended in 0.4 mL of Polymix buffer [108] (20 mM HEPES:KOH pH 7.5, 95
716 mM KCl, 5 mM NH_4Cl, 5 mM $Mg(OAc)_2$, 0.5 mM $CaCl_2$, 8 mM putrescine, 1 mM spermidine, 1 mM
717 DTT) and 200 µL of pre-chilled zirconium beads (0.1 mm) were added to each sample. Cellular
718 lysates were prepared by a FastPrep homogeniser (MP Biomedicals) (three 20 seconds pulses at
719 speed 6.0 mp/sec with chilling on ice for 1 minutes between the cycles) and clarified by
720 centrifugation at 21,000 g for 10 minutes at 4 °C. The supernatant was carefully collected avoiding
721 the lipid layer and cellular pellet, aliquoted, frozen in liquid nitrogen and stored at -80 °C until
722 further processing.

723 *Sucrose gradient centrifugation:*

724 After melting the frozen samples on ice, 2 A_{260} units of each extract was loaded onto 5–25% (w/v)
725 sucrose density gradients in Polymix buffer, 5 mM Mg^{2+} [108]. Gradients were resolved at 35,000
726 rpm for 2.5 hours at 4 °C in SW41 rotor (Beckman) and analysed using Biocomp Gradient Station
727 (BioComp Instruments) with A_{260} as a readout. The ribosome profiles presented were normalised
728 to the total area under the curve and are representative of at least three independent experiments
729 for each strain.

730

731 **Polysome profiling and Western blot analysis of B. subtilis strains**

732 Experiments were performed as described above for *E. coli* strains, with minor modifications.

733 *Preparation of bacterial samples and preparation of clarified lysates:*

734 VHB90 and VHB91 strains were pre-grown on LB plates overnight at 30 °C. Fresh individual
735 colonies were used to inoculate 200 mL LB cultures. The cultures were grown at 37 °C until OD_{600}
736 of 0.3 and IPTG was added to final concentration of 30 µM. After 30 min cells were collected by
737 centrifugation (8,000 rmp, 10 minutes), dissolved in 0.5 mL of Polymix buffer [108] (20 mM
738 HEPES:KOH pH 7.5, 95 mM KCl, 5 mM NH_4Cl, 10 mM $Mg(OAc)_2$, 0.5 mM $CaCl_2$, 8 mM putrescine, 1
739 mM spermidine, 1 mM DTT, 2 mM PMSF), lysed (FastPrep homogeniser (MP Biomedicals): four 20
740 seconds pulses at speed 6.0 mp/sec with chilling on ice for 1 minutes between the cycles), and
741 clarified by ultracentrifugation (14,800 rpm, 20 minutes).

742 *Sucrose gradient centrifugation and Western blotting:*

743 Clarified cell lysates were loaded onto 7–35% sucrose gradients in Polymix buffer [108] (20 mM
744 HEPES:KOH pH 7.5, 95 mM KCl, 5 mM NH_4Cl, 10 mM $Mg(OAc)_2$, 0.5 mM $CaCl_2$, 8 mM putrescine, 1
745 mM spermidine, 1 mM DTT) and subjected to centrifugation (35,000 rpm for 3 hours at 4 °C). C-
746 terminally HTF-tagged VmlR (wild type and EQ_2 mutant) and ribosomal protein L3 of the 50S
747 ribosomal subunit were detected using either anti-Flag M2 primary combined with anti-mouse-
748 HRP secondary antibodies or anti-L3 primary (a gift from Fujio Kawamura) combined with goat
749 anti-rabbit IgG-HRP secondary antibodies, respectively. All antibodies were used at 1:10,000
750 dilution.
751
752 **L-[^{35}S]-methionine pulse-labelling**
753 *Preparation of bacterial samples:*
754 Since glucose specifically inhibits the arabinose promoter, the cultures were grown in defined
755 Neidhardt MOPS medium [105] supplemented with 0.4% glycerol as a carbon source. One colony
756 of freshly transformed *E. coli* BW25113 cells expressing N-terminal FTH-tagged ABCFs (wild type
757 and EQ_2 mutants) from pBad vector was used to inoculate 10 mL of 1x MOPS media supplemented
758 with 0.4% glycerol 100 μg/mL carbenicillin, and the cultures were grown until early stationary
759 phase (about 24 hours). Stationary phase cells were diluted to OD_{600} of 0.04–0.07 in 25 mL of the
760 same media, grown at 37 °C with vigorous shaking (200 rpm) to OD_{600} of 0.15–0.2 and expression
761 of ABCFs was induced by addition of L-arabinose to the final concentration of 0.2%.
762 *L-[^{35}S]-methionine pulse-labelling:*
763 For radioactive pulse labelling, 1 μCi L-[^{35}S]-methionine (500μCi, PerkinElmer) aliquots were
764 prepared in 1.5 mL Eppendorf tubes. As a zero time point, 1 mL of cell culture was taken and mixed
765 with an aliquot of radioactive methionine just before inducing cells with L-arabinose.
766 Simultaneously a 1 mL aliquot was taken for an OD_{600} measurement. All consecutive samples were
767 processed similarly at designated time points after induction. ^{35}S-methionine incorporation was
768 stopped after 5 minutes by chloramphenicol added to the final concentration of 200 μg/mL.
769 Subsequent processing of samples differs in the case of scintillation counting and autoradiography.
770 *Scintillation counting:*

771 1 mL of culture was combined with 200 µL of 50% trichloroacetic acid (TCA), passed through a
772 GF/C filter (Whatman) prewashed with 5% TCA and unincorporated label was removed by
773 washing the filter with 5 mL of ice-cold 5% TCA followed by 5 mL of ice-cold 95% EtOH [109].
774 Filters were dried for at least 2 hours, and counted on a TRI-CARB 4910TR 110 V scintillation
775 counter (PerkinElmer) (5 mL of ScintiSafe 3 scintillation cocktail (FisherScientific) per sample,
776 pre-soaked with shaking for 15 minutes prior to counting).

777 *Autoradiography:*

778 1 mL cultures were pelleted by centrifugation, cell pellet washed with Phosphate Buffered Saline
779 (PBS) to remove unincorporated L-[^{35}S]-methionine and dissolved/lysed in 50 µL 1x SDS-loading
780 buffer. Samples were normalised by OD600 by addition of appropriate volume of 1x SDS-loading
781 buffer (50-80 µL according to OD600), and 10 µL of the sample was loaded onto 10% SDS-PAGE
782 and resolved electrophoretically (BioRad), gels were dried on Whatman paper, exposed on BAS
783 storage phosphor screen (GE Heathcare) overnight, and scanned by Typhoon imaging system (GE
784 Heathcare).

785

786 **Acknowledgements**

787 We are grateful to the anonymous reviewers in their suggestions of additional experiments and
788 analyses that strengthened the manuscript. We gratefully acknowledge Fujio Kawamura for
789 providing anti-L3 primary antibodies and Stijn Hendrik Peeters for constructing the pSHP2
790 plasmid. Thanks to Martin Carr for discussion and advice about eEF3 and EFL co-distribution.

791

792 **Author contributions:**

793 GCA conceived the study and carried out all the bioinformatic analyses except for sequence logo
794 construction and eEF1A/EFL/eEF3 co-distribution analysis, which was carried out by CKS. VH, VM,
795 MK, HS, HT, TT and MP designed experimental analyses. VM, MK, HT, MH, TS, MR, HS and JWG
796 carried out the experiments. VH, GCA, HS, TT, HT, VM and MK analyzed the data. GCA and VH
797 drafted the manuscript with input from VM, MK, HT, MH, CKS, JWG, TS, MP, TT and HS.

798

799 **References**

800

801 1. Skogerson, L. and E. Wakatama, *A ribosome-dependent GTPase from yeast distinct from*
802 *elongation factor 2.* Proc Natl Acad Sci U S A, 1976. **73**(1): p. 73-6.
803 2. Andersen, C.B., et al., *Structure of eEF3 and the mechanism of transfer RNA release from the*
804 *E-site.* Nature, 2006. **443**(7112): p. 663-8.
805 3. Kurata, S., et al., *Ribosome recycling step in yeast cytoplasmic protein synthesis is catalyzed*
806 *by eEF3 and ATP.* Proc Natl Acad Sci U S A, 2010. **107**(24): p. 10854-9.
807 4. Triana-Alonso, F.J., K. Chakraburtty, and K.H. Nierhaus, *The elongation factor 3 unique in*
808 *higher fungi and essential for protein biosynthesis is an E site factor.* J Biol Chem, 1995.
809 **270**(35): p. 20473-8.
810 5. Pisarev, A.V., et al., *The role of ABCE1 in eukaryotic posttermination ribosomal recycling.* Mol
811 Cell, 2010. **37**(2): p. 196-210.
812 6. Becker, T., et al., *Structural basis of highly conserved ribosome recycling in eukaryotes and*
813 *archaea.* Nature, 2012. **482**(7386): p. 501-6.
814 7. Young, D.J., et al., *Rli1/ABCE1 Recycles Terminating Ribosomes and Controls Translation*
815 *Reinitiation in 3'UTRs In Vivo.* Cell, 2015. **162**(4): p. 872-84.
816 8. Caetano-Anolles, D., et al., *Proteome evolution and the metabolic origins of translation and*
817 *cellular life.* J Mol Evol, 2011. **72**(1): p. 14-33.
818 9. Davidson, A.L., et al., *Structure, function, and evolution of bacterial ATP-binding cassette*
819 *systems.* Microbiol Mol Biol Rev, 2008. **72**(2): p. 317-64, table of contents.
820 10. Dean, M., A. Rzhetsky, and R. Allikmets, *The human ATP-binding cassette (ABC) transporter*
821 *superfamily.* Genome Res, 2001. **11**(7): p. 1156-66.
822 11. Kerr, I.D., *Sequence analysis of twin ATP binding cassette proteins involved in translational*
823 *control, antibiotic resistance, and ribonuclease L inhibition.* Biochem Biophys Res Commun,
824 2004. **315**(1): p. 166-73.
825 12. Vazquez de Aldana, C.R., M.J. Marton, and A.G. Hinnebusch, *GCN20, a novel ATP binding*
826 *cassette protein, and GCN1 reside in a complex that mediates activation of the eIF-2 alpha*
827 *kinase GCN2 in amino acid-starved cells.* EMBO J, 1995. **14**(13): p. 3184-99.
828 13. Paytubi, S., et al., *ABC50 promotes translation initiation in mammalian cells.* J Biol Chem,
829 2009. **284**(36): p. 24061-73.
830 14. Dong, J., et al., *The novel ATP-binding cassette protein ARB1 is a shuttling factor that*
831 *stimulates 40S and 60S ribosome biogenesis.* Mol Cell Biol, 2005. **25**(22): p. 9859-73.
832 15. Li, Z., et al., *Rational extension of the ribosome biogenesis pathway using network-guided*
833 *genetics.* PLoS Biol, 2009. **7**(10): p. e1000213.

834 16. Boël, G., et al., *The ABC-F protein EttA gates ribosome entry into the translation elongation*
835 *cycle.* Nat Struct Mol Biol, 2014. **21**(2): p. 143-51.
836 17. Chen, B., et al., *EttA regulates translation by binding the ribosomal E site and restricting*
837 *ribosome-tRNA dynamics.* Nat Struct Mol Biol, 2014. **21**(2): p. 152-9.
838 18. Reynolds, E.D. and J.H. Cove, *Resistance to telithromycin is conferred by msr(A), msrC and*
839 *msr(D) in Staphylococcus aureus.* J Antimicrob Chemother, 2005. **56**(6): p. 1179-80.
840 19. Ohki, R., et al., *Transcriptional termination control of a novel ABC transporter gene involved*
841 *in antibiotic resistance in Bacillus subtilis.* J Bacteriol, 2005. **187**(17): p. 5946-54.
842 20. Singh, K.V., G.M. Weinstock, and B.E. Murray, *An Enterococcus faecalis ABC homologue (Lsa)*
843 *is required for the resistance of this species to clindamycin and quinupristin-dalfopristin.*
844 Antimicrob Agents Chemother, 2002. **46**(6): p. 1845-50.
845 21. Novotna, G. and J. Janata, *A new evolutionary variant of the streptogramin A resistance*
846 *protein, Vga(A)LC, from Staphylococcus haemolyticus with shifted substrate specificity*
847 *towards lincosamides.* Antimicrob Agents Chemother, 2006. **50**(12): p. 4070-6.
848 22. Hot, C., N. Berthet, and O. Chesneau, *Characterization of sal(A), a novel gene responsible for*
849 *lincosamide and streptogramin A resistance in Staphylococcus sciuri.* Antimicrob Agents
850 Chemother, 2014. **58**(6): p. 3335-41.
851 23. Buche, A., C. Mendez, and J.A. Salas, *Interaction between ATP, oleandomycin and the OleB*
852 *ATP-binding cassette transporter of Streptomyces antibioticus involved in oleandomycin*
853 *secretion.* Biochem J, 1997. **321 (Pt 1)**: p. 139-44.
854 24. Ross, J.I., et al., *Inducible erythromycin resistance in staphylococci is encoded by a member of*
855 *the ATP-binding transport super-gene family.* Mol Microbiol, 1990. **4**(7): p. 1207-14.
856 25. Wang, Y., et al., *A novel gene, optrA, that confers transferable resistance to oxazolidinones*
857 *and phenicols and its presence in Enterococcus faecalis and Enterococcus faecium of human*
858 *and animal origin.* J Antimicrob Chemother, 2015. **70**(8): p. 2182-90.
859 26. Kitani, S., et al., *Characterization of varM Encoding Type II ABC Transporter in Streptomyces*
860 *virginiae, a Virginiamycin M1 Producer.* Actinomycetologica, 2010. **24**(2): p. 51-57.
861 27. Kerr, I.D., E.D. Reynolds, and J.H. Cove, *ABC proteins and antibiotic drug resistance: is it all*
862 *about transport?* Biochem Soc Trans, 2005. **33**(Pt 5): p. 1000-2.
863 28. Rice, L.B., *Progress and challenges in implementing the research on ESKAPE pathogens.* Infect
864 Control Hosp Epidemiol, 2010. **31 Suppl 1**: p. S7-10.
865 29. Wilson, D.N., *The ABC of Ribosome-Related Antibiotic Resistance.* MBio, 2016. **7**(3).
866 30. Sharkey, L.K.R. and A.J. O'Neill, *Antibiotic Resistance ABC-F Proteins: Bringing Target*
867 *Protection into the Limelight.* ACS Infect Dis, 2018. **4**(3): p. 239-246.
868 31. Sharkey, L.K., T.A. Edwards, and A.J. O'Neill, *ABC-F Proteins Mediate Antibiotic Resistance*
869 *through Ribosomal Protection.* MBio, 2016. **7**(2).

870 32. Murina, V., et al., *Antibiotic resistance ABCF proteins reset the peptidyl transferase centre of*
871 *the ribosome to counter translational arrest.* Nucleic Acids Res, 2018. **46**(7): p. 3753-3763.
872 33. Su, W., et al., *Ribosome protection by antibiotic resistance ATP-binding cassette protein.* Proc
873 Natl Acad Sci U S A, 2018. **115**(20): p. 5157-5162.
874 34. Crowe-McAuliffe, C., et al., *Structural basis for antibiotic resistance mediated by the Bacillus*
875 *subtilis ABCF ATPase VmlR.* Proc Natl Acad Sci U S A, 2018.
876 35. Atkinson, G.C., *The evolutionary and functional diversity of classical and lesser-known*
877 *cytoplasmic and organellar translational GTPases across the tree of life.* BMC Genomics,
878 2015. **16**: p. 78.
879 36. Donhofer, A., et al., *Structural basis for TetM-mediated tetracycline resistance.* Proc Natl
880 Acad Sci U S A, 2012. **109**(42): p. 16900-5.
881 37. Li, W., et al., *Mechanism of tetracycline resistance by ribosomal protection protein Tet(O).* Nat
882 Commun, 2013. **4**: p. 1477.
883 38. Laurence, M., C. Hatzis, and D.E. Brash, *Common contaminants in next-generation sequencing*
884 *that hinder discovery of low-abundance microbes.* PLoS One, 2014. **9**(5): p. e97876.
885 39. Pawlowski, A.C., et al., *A diverse intrinsic antibiotic resistome from a cave bacterium.* Nat
886 Commun, 2016. **7**: p. 13803.
887 40. Schoner, B., et al., *Sequence similarity between macrolide-resistance determinants and ATP-*
888 *binding transport proteins.* Gene, 1992. **115**(1-2): p. 93-6.
889 41. Peschke, U., et al., *Molecular characterization of the lincomycin-production gene cluster of*
890 *Streptomyces lincolnensis 78-11.* Mol Microbiol, 1995. **16**(6): p. 1137-56.
891 42. The UniProt Consortium, *UniProt: the universal protein knowledgebase.* Nucleic Acids Res,
892 2017. **45**(D1): p. D158-D169.
893 43. Carlier, L., et al., *The C-terminal domain of the Uup protein is a DNA-binding coiled coil motif.*
894 J Struct Biol, 2012. **180**(3): p. 577-84.
895 44. Maurice, T.C., et al., *A highly conserved intraspecies homolog of the Saccharomyces cerevisiae*
896 *elongation factor-3 encoded by the HEF3 gene.* Yeast, 1998. **14**(12): p. 1105-13.
897 45. Sarthy, A.V., et al., *Identification and kinetic analysis of a functional homolog of elongation*
898 *factor 3, YEF3 in Saccharomyces cerevisiae.* Yeast, 1998. **14**(3): p. 239-53.
899 46. Marton, M.J., et al., *Evidence that GCN1 and GCN20, translational regulators of GCN4, function*
900 *on elongating ribosomes in activation of eIF2alpha kinase GCN2.* Mol Cell Biol, 1997. **17**(8):
901 p. 4474-89.
902 47. Anand, M., et al., *Functional interactions between yeast translation eukaryotic elongation*
903 *factor (eEF) 1A and eEF3.* J Biol Chem, 2003. **278**(9): p. 6985-91.
904 48. Anand, M., et al., *Domain and nucleotide dependence of the interaction between*
905 *Saccharomyces cerevisiae translation elongation factors 3 and 1A.* J Biol Chem, 2006.
906 **281**(43): p. 32318-26.

907 49. Yamada, T., et al., *Expression of the gene encoding a translational elongation factor 3*
908 *homolog of Chlorella virus CVK2.* Virology, 1993. **197**(2): p. 742-50.
909 50. Atkinson, G.C., et al., *An evolutionary ratchet leading to loss of elongation factors in*
910 *eukaryotes.* BMC Evol Biol, 2014. **14**(1): p. 35.
911 51. Inoue, Y., et al., *Yeast prion protein New1 can break Sup35 amyloid fibrils into fragments in*
912 *an ATP-dependent manner.* Genes Cells, 2011. **16**(5): p. 545-56.
913 52. Paytubi, S., et al., *The N-terminal region of ABC50 interacts with eukaryotic initiation factor*
914 *eIF2 and is a target for regulatory phosphorylation by CK2.* Biochem J, 2008. **409**(1): p. 223-
915 31.
916 53. Kachroo, A.H., et al., *Evolution. Systematic humanization of yeast genes reveals conserved*
917 *functions and genetic modularity.* Science, 2015. **348**(6237): p. 921-5.
918 54. Castilho, B.A., et al., *Keeping the eIF2 alpha kinase Gcn2 in check.* Biochim Biophys Acta,
919 2014. **1843**(9): p. 1948-68.
920 55. Hirose, T. and H.R. Horvitz, *The translational regulators GCN-1 and ABCF-3 act together to*
921 *promote apoptosis in C. elegans.* PLoS Genet, 2014. **10**(8): p. e1004512.
922 56. Visweswaraiah, J., et al., *Overexpression of eukaryotic translation Elongation Factor 3 (eEF3)*
923 *impairs GCN2 activation.* J Biol Chem, 2012.
924 57. Kaida, D., et al., *Yeast Whi2 and Psr1-phosphatase form a complex and regulate STRE-*
925 *mediated gene expression.* Genes Cells, 2002. **7**(6): p. 543-52.
926 58. Dar, D., et al., *Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria.*
927 Science, 2016. **352**(6282): p. aad9822.
928 59. Orelle, C., et al., *The conserved glutamate residue adjacent to the Walker-B motif is the*
929 *catalytic base for ATP hydrolysis in the ATP-binding cassette transporter BmrA.* J Biol Chem,
930 2003. **278**(47): p. 47002-8.
931 60. Smith, P.C., et al., *ATP binding to the motor domain from an ABC transporter drives formation*
932 *of a nucleotide sandwich dimer.* Mol Cell, 2002. **10**(1): p. 139-49.
933 61. Britton, R.A., et al., *Genome-wide analysis of the stationary-phase sigma factor (sigma-H)*
934 *regulon of Bacillus subtilis.* J Bacteriol, 2002. **184**(17): p. 4881-90.
935 62. Shaner, N.C., et al., *A bright monomeric green fluorescent protein derived from*
936 *Branchiostoma lanceolatum.* Nat Methods, 2013. **10**(5): p. 407-9.
937 63. Lewis, P.J. and A.L. Marston, *GFP vectors for controlled expression and dual labelling of*
938 *protein fusions in Bacillus subtilis.* Gene, 1999. **227**(1): p. 101-10.
939 64. Akanuma, G., et al., *Ribosome dimerization is essential for the efficient regrowth of Bacillus*
940 *subtilis.* Microbiology, 2016. **162**(3): p. 448-458.
941 65. Scheinpflug, K., et al., *Antimicrobial peptide cWFW kills by combining lipid phase separation*
942 *with autolysis.* Sci Rep, 2017. **7**: p. 44332.

943 66. Schafer, L.V., et al., *Lipid packing drives the segregation of transmembrane helices into*
944 *disordered lipid domains in model membranes.* Proc Natl Acad Sci U S A, 2011. **108**(4): p.
945 1343-8.
946 67. Jahn, N., S. Brantl, and H. Strahl, *Against the mainstream: the membrane-associated type I*
947 *toxin BsrG from Bacillus subtilis interferes with cell envelope biosynthesis without increasing*
948 *membrane permeability.* Mol Microbiol, 2015. **98**(4): p. 651-66.
949 68. Lewis, P.J., S.D. Thaker, and J. Errington, *Compartmentalization of transcription and*
950 *translation in Bacillus subtilis.* EMBO J, 2000. **19**(4): p. 710-8.
951 69. Sanamrad, A., et al., *Single-particle tracking reveals that free ribosomal subunits are not*
952 *excluded from the Escherichia coli nucleoid.* Proc Natl Acad Sci U S A, 2014. **111**(31): p.
953 11413-8.
954 70. Cochrane, K.L., *Elucidating Ribosomes-Genetic Studies of the ATPase Uup and the Ribosomal*
955 *Protein L1*. 2015, University of Michigan. p. 147.
956 71. Gibbs, M.R. and K. Fredrick, *Roles of elusive translational GTPases come to light and inform*
957 *on the process of ribosome biogenesis in bacteria.* Mol Microbiol, 2018. **107**(4): p. 445-454.
958 72. deLivron, M.A. and V.L. Robinson, *Salmonella enterica serovar Typhimurium BipA exhibits*
959 *two distinct ribosome binding modes.* J Bacteriol, 2008. **190**(17): p. 5944-52.
960 73. Kumar, V., et al., *Structure of BipA in GTP form bound to the ratcheted ribosome.* Proc Natl
961 Acad Sci U S A, 2015. **112**(35): p. 10944-9.
962 74. Hauryliuk, V., et al., *Recent functional insights into the role of (p)ppGpp in bacterial*
963 *physiology.* Nat Rev Microbiol, 2015. **13**(5): p. 298-309.
964 75. deLivron, M.A., et al., *A novel domain in translational GTPase BipA mediates interaction with*
965 *the 70S ribosome and influences GTP hydrolysis.* Biochemistry, 2009. **48**(44): p. 10533-41.
966 76. Welch, R.A., et al., *Extensive mosaic structure revealed by the complete genome sequence of*
967 *uropathogenic Escherichia coli.* Proc Natl Acad Sci U S A, 2002. **99**(26): p. 17020-4.
968 77. Cohen, S.N. and A.C. Chang, *Recircularization and autonomous replication of a sheared R-*
969 *factor DNA segment in Escherichia coli transformants.* Proc Natl Acad Sci U S A, 1973. **70**(5):
970 p. 1293-7.
971 78. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia*
972 *coli K-12 using PCR products.* Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6640-5.
973 79. Grenier, F., et al., *Complete Genome Sequence of Escherichia coli BW25113.* Genome
974 Announc, 2014. **2**(5).
975 80. Baba, T., et al., *Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants:*
976 *the Keio collection.* Mol Syst Biol, 2006. **2**: p. 2006 0008.
977 81. Guthrie, C., H. Nashimoto, and M. Nomura, *Structure and function of E. coli ribosomes. 8.*
978 *Cold-sensitive mutants defective in ribosome assembly.* Proc Natl Acad Sci U S A, 1969. **63**(2):
979 p. 384-91.

980 82. Jia, B., et al., *CARD 2017: expansion and model-centric curation of the comprehensive*
981 *antibiotic resistance database.* Nucleic Acids Res, 2017. **45**(D1): p. D566-D573.
982 83. Mahajan, G.B. and L. Balachandran, *Antibacterial agents from actinomycetes - a review.*
983 Front Biosci (Elite Ed), 2012. **4**: p. 240-53.
984 84. Lassak, J., D.N. Wilson, and K. Jung, *Stall no more at polyproline stretches with the translation*
985 *elongation factors EF-P and IF-5A.* Mol Microbiol, 2016. **99**(2): p. 219-35.
986 85. Ude, S., et al., *Translation elongation factor EF-P alleviates ribosome stalling at polyproline*
987 *stretches.* Science, 2013. **339**(6115): p. 82-5.
988 86. Peil, L., et al., *Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the*
989 *translation elongation factor EF-P.* Proc Natl Acad Sci U S A, 2013. **110**(38): p. 15265-70.
990 87. Elgamal, S., et al., *EF-P dependent pauses integrate proximal and distal signals during*
991 *translation.* PLoS Genet, 2014. **10**(8): p. e1004553.
992 88. Receveur, V., et al., *Dimension, shape, and conformational flexibility of a two domain fungal*
993 *cellulase in solution probed by small angle X-ray scattering.* J Biol Chem, 2002. **277**(43): p.
994 40887-92.
995 89. Seip, B. and C.A. Innis, *How Widespread is Metabolite Sensing by Ribosome-Arresting Nascent*
996 *Peptides?* J Mol Biol, 2016. **428**(10 Pt B): p. 2217-27.
997 90. Kannan, K., N. Vazquez-Laslop, and A.S. Mankin, *Selective protein synthesis by ribosomes*
998 *with a drug-obstructed exit tunnel.* Cell, 2012. **151**(3): p. 508-20.
999 91. Almutairi, M.M., et al., *Co-produced natural ketolides methymycin and pikromycin inhibit*
1000 *bacterial growth by preventing synthesis of a limited number of proteins.* Nucleic Acids Res,
1001 2017. **45**(16): p. 9573-9582.
1002 92. Camacho, C., et al., *BLAST+: architecture and applications.* BMC Bioinformatics, 2009. **10**: p.
1003 421.
1004 93. Stamatakis, A., *RAxML version 8: a tool for phylogenetic analysis and post-analysis of large*
1005 *phylogenies.* Bioinformatics, 2014. **30**(9): p. 1312-3.
1006 94. Miller, M.A., W. Pfeiffer, and T. Schwartz. *Creating the CIPRES Science Gateway for Inference*
1007 *of Large Phylogenetic Trees*. in *Gateway Computing Environments Workshop (GCE)*. 2010.
1008 New Orleans, LA.
1009 95. Wheeler, T.J., J. Clements, and R.D. Finn, *Skylign: a tool for creating informative, interactive*
1010 *logos representing sequence alignments and profile hidden Markov models.* BMC
1011 Bioinformatics, 2014. **15**: p. 7.
1012 96. Emanuelsson, O., et al., *Locating proteins in the cell using TargetP, SignalP and related tools.*
1013 Nat Protoc, 2007. **2**(4): p. 953-71.
1014 97. Katoh, K. and D.M. Standley, *MAFFT multiple sequence alignment software version 7:*
1015 *improvements in performance and usability.* Mol Biol Evol, 2013. **30**(4): p. 772-80.

1016 98. Abascal, F., R. Zardoya, and D. Posada, *ProtTest: selection of best-fit models of protein*
1017 *evolution.* Bioinformatics, 2005. **21**(9): p. 2104-5.
1018 99. Minh, B.Q., M.A. Nguyen, and A. von Haeseler, *Ultrafast approximation for phylogenetic*
1019 *bootstrap.* Mol Biol Evol, 2013. **30**(5): p. 1188-95.
1020 100. Biasini, M., et al., *SWISS-MODEL: modelling protein tertiary and quaternary structure using*
1021 *evolutionary information.* Nucleic Acids Res, 2014. **42**(Web Server issue): p. W252-8.
1022 101. Xu, D. and Y. Zhang, *Ab initio protein structure assembly using continuous structure*
1023 *fragments and optimized knowledge-based force field.* Proteins, 2012. **80**(7): p. 1715-35.
1024 102. Schrodinger, L., *The PyMOL Molecular Graphics System, Version 1.8.* 2015.
1025 103. Lupas, A., M. Van Dyke, and J. Stock, *Predicting coiled coils from protein sequences.* Science,
1026 1991. **252**(5009): p. 1162-4.
1027 104. Barbe, V., et al., *From a consortium sequence to a unified sequence: the Bacillus subtilis 168*
1028 *reference genome a decade later.* Microbiology, 2009. **155**(Pt 6): p. 1758-75.
1029 105. Neidhardt, F.C., P.L. Bloch, and D.F. Smith, *Culture medium for enterobacteria.* J Bacteriol,
1030 1974. **119**(3): p. 736-47.
1031 106. Te Winkel, J.D., et al., *Analysis of Antimicrobial-Triggered Membrane Depolarization Using*
1032 *Voltage Sensitive Dyes.* Front Cell Dev Biol, 2016. **4**: p. 29.
1033 107. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nat Methods,
1034 2012. **9**(7): p. 676-82.
1035 108. Antoun, A., et al., *Ribosome formation from subunits studied by stopped-flow and Rayleigh*
1036 *light scattering.* Biol Proced Online, 2004. **6**: p. 35-54.
1037 109. Esposito, A.M. and T.G. Kinzy, *In vivo [35S]-methionine incorporation.* Methods Enzymol,
1038 2014. **536**: p. 55-64.
1039

1040 **Supplementary file captions**

1041

1042 **S1 Table. Taxonomy of 4505 species, and their ABCF composition**

1043 All species considered in the analysis are listed, ordered by taxonomy. The number and identity of

1044 ABCF subfamilies are recorded.

1045

1046 **S2 Table. Classification of 16848 ABCF sequences into subfamilies, accompanied by domain**

1047 **assignments**

1048 The unique sequence identifiers are included for retrieval of data from online repositories.

1049

1050 **S3 Table. Domain coordinates**

1051 The domain coordinates for the representative sequences in fig. 3A are listed. In the second tab, all

1052 the coordinates for each identified domain in all ABCFs are given. As these are from HMM hits, the

1053 same domain can have more than one hit in each protein. For example, the ABC1 HMM always hits

1054 the ABC2 domain, and vice versa. Duplicate domain hits were removed when generating fig 3A.

1055

1056 **S4 Table. Presence and absence of EFL, eEF1A and eEF3 in eukaryotes**

1057 Where the distribution is unchanged within a specific taxonomic lineage, those rows are collapsed

1058 down to one, and the highest common taxonomic rank is given. The full lineage data is available in

1059 the second tab.

1060

1061 **S5 Table. Transit peptide predictions**

1062 Predictions were made separately for plastid-containing and non-plastid containing eukaryotes.

1063 The description of the output format is shown below the predictions.

1064

1065 **Table S6. Primers used in the study**

1066

1067 **Table S7. The starting OD_{600} of *E. coli* CFT073 and its derivatives.**

1068

1069 **S1 Figure. Ladderised version of the Figure 1 tree**

1070 All branch support values and taxon names including subfamily identity are shown. Branch
1071 colouring is as per Figure 1. Orange stars show Uup sequences that have truncated Arm
1072 subdomains.

1073

1074 **S2 Figure. Sequence logos of domains show amino acid biases**

1075 Sequence logos of each domain HMM. The height of stacked amino acids at each position show the
1076 information content, in bits, with letters dividing the height according to their estimated
1077 probability. Beneath the stacked amino acids there are three lines showing probabilities; line 1 is
1078 occupancy, the probability of observing a letter –rather than a gap - at that position. Line 2 is the
1079 probably of seeing an insertion at that position, and line 3 is the expected length of an insertion
1080 following that position.

1081

1082 **S4 Figure. Functional testing of inducible *B. subtilis* VmlR-HTF and VmlR-NeonGreen fusion**
1083 **proteins.** All experiments were performed in filtered LB at 37˚C in the presence of increasing
1084 concentrations of lincomycin and presented as the geometric mean ± standard deviation (n = 3).
1085 HTF stands for C-terminal His_6-TEV- 3xFLAG tag.

1086

1087 **S5 Figure. Growth and polysome analysis of *E. coli* CFT073 wild type, Δ*bipA*, Δ*abcf* and**
1088 **Δ*bipA*Δ*uup* strains, and *E. coli* CFT073 wild type overexpressing ABCFs under the control of**
1089 **constitutive P_{tet} promoter.** All experiments were performed in filtered LB at either 18˚C (A) or
1090 37˚C (B-D) and growth data are presented as geometric means ± standard deviation (n = 3).

1091

1092 **S6 Figure. Functional testing of FTH-tagged *E. coli* ABCF wild type and EQ_2 proteins.**

1093 Western blot (A) and growth assays (B and C) N-terminally FTH-tagged *E. coli* ABCF proteins
1094 expressed under the control of arabinose-inducible P_{BAD} promoter. Growth experiments were
1095 performed in MOPS media at 18˚C and presented as geometric means ± standard deviation (n = 3).

1096 Western blotting was performed either at 18°C or 37°C. FTH stands for N-terminal 3xFLAG-TEV-
1097 His_6 tag.
1098
1099 **S7 Figure. AREs show variability in presence across species of the same genus, which can be**
1100 **used to predict novel ARE ABCFs.** Genera were selected that contained more than 20 species in
1101 the ABCF database. For each subfamily in each of those genera, the percent of species in which that
1102 subfamily is found, is plotted. Colder colours are those ABCFs that are more universal, as typical of
1103 housekeeping genes. AREs on the other hand have a more patchy distribution, as shown by their
1104 tendencies for warmer colours. Antibiotic producing genera are the exception, where AREs and
1105 ARE-like subfamilies can be universal within a genus.
1106
1107 **S3 Figure. Maximum likelihood phylogeny of eukaryotic subgroup-type ABCFs from**
1108 **eukaryotes and viruses.**
1109 The tree is rooted with bacterial YheS sequences. Branch support is from 100 bootstrap replicates
1110 and branch length is proportional to the number of amino acid substitutions (see lower scale bar).
1111 Names are coloured by subfamily.
1112
1113 **S1 Text. Supplementary sequence alignments and phylogenetic trees.**
1114 The file contains sequence alignments used to generate Figures 1 (and S1), 2 and S3 in FASTA
1115 format, and Newick format phylogenetic trees that are not shown as figures. Each entity
1116 (alignment or tree) is separated by comments preceded by "#".
1117
1118 **S1 Materials and Methods.**
1119 **Supplementary methods describing the construction of plasmids and bacterial strains.**
1120
1121

1122 **Figure legends**

1123

1124 **Fig. 1 | The family tree of ABCFs has a bipartite structure corresponding to eukaryotic-like**
1125 **and bacterial (and organellar)-like sequences**

1126 The tree is a RaxML maximum likelihood phylogeny of representatives across the ABCF family with
1127 branch support values from 100 bootstrap replicates with RaxML (MLB), 1000 ultrafast bootstrap
1128 replicates with IQ-TREE (UFB) and Bayesian inference posterior probability (BIPP). The inset box
1129 shows the legend for subfamily and intersubfamily support; support values within subfamilies and
1130 that are less that 60% MLB are not shown. Species were chosen that sample broadly across the
1131 tree of ABCF-encoding life, sampling at least one representative from each subfamily. Green
1132 shading shows the eukaryotic type ABCFs; other subgroups are bacterial unless marked with a
1133 green shaded circle to indicate eukaryotic groups with potentially endosymbiotic origin. CpYdif
1134 contains both cyanobacterial and predicted chloroplast sequences. The full tree with taxon names
1135 and sequence IDs is shown in fig. S1. Branch lengths are proportional to amino acid substitutions
1136 as per the scale bar in the lower right. The asterisked branch is not supported by this data set,
1137 however it is supported at 85% MLB in phylogenetic analysis of the eukaryotic subgroup and its
1138 viral relatives, rooted with YheS (fig. S3). Branch lengths are proportional to amino acid
1139 substitutions as per the lower right scale bar.

1140

1141 **Fig. 2 | Rooting with ABCE shows eukaryotic-like ABCFs nesting within bacterial-like ABCFs,**
1142 **with YheS as the sister group to the eukaryotic-like clade**

1143 Maximum likelihood phylogeny of representatives across the ABCF family, and ABCE sequences
1144 from the UniProt database. Branch support from 200 bootstrap replicates with RaxML (MBP),
1145 1000 ultrafast bootstrap replicates with IQ-TREE and Bayesian inference posterior probability is
1146 indicated with the key in the inset box. Branch lengths are proportional to amino acid substitutions
1147 as per the inset scale bar.

1148

1149 **Fig. 3 | Typical domain and subdomain architectures of ABCFs**

1150 (A) Boxes show domains as predicted by HMMs. Full coordinates and sequence data for these
1151 examples are recorded in Table S3. Dotted lines shows possible interactions between the HEAT
1152 domain of eEF3/New1/eEF3L and the N-terminal domain of ABCF3, ABCF4 and ABCF7. (B)
1153 Predicted coiled coil regions of *E. coli* YheS along the protein length. Inset: cartoon representation
1154 of the coiled coil subdomains protruding from the core ABC domains.
1155

1156 **Fig. 4 | *B. subtilis* ARE VmlR is a cytoplasmic protein that directly protects the ribosome from**
1157 **antibiotics**

1158 (A) Growth of wild type *B. subtilis* 168, isogenic Δ*vmlR* knockout as well as Δ*vmlR* knockout
1159 expressing either wild type or EQ_2 version of VmlR under the control of IPTG-inducible $P_{hy\text{-}spank}$
1160 promoter. Six biological replicates were averaged for each growth curve and the data presented as
1161 geometric means ± standard deviation. (B) Polysome analysis and western blotting of Δ*vmlR B.*
1162 *subtilis* expressing C-terminally HTF-tagged wild type and EQ_2 version of VmlR. (C) Phase contrast
1163 and fluorescence images of uninhibited *B. subtilis* cells expressing VmlR-mNeonGreen (VmlR-mNG)
1164 in the presence and absence of lincomycin (40 min incubation with 5 μg/ml), and a model
1165 transmembrane protein WALP23-GFP are shown for comparison. (D) Fluorescence intensity
1166 profiles were measured perpendicular to the cell length axis along a 325 nm wide and 5.8 μm long
1167 line as indicated. Fluorescence intensity profiles of cells expressing WALP23-GFP [65], and cells
1168 expressing VmlR-mNG in the presence and absence of lincomycin. The graph depicts the average
1169 fluorescence intensity profiles and the corresponding standard deviations (n = 30).
1170

1171 **Fig. 5 | Overexpression of *E. coli* ABCF Uup suppresses cold sensitivity and ribosome**
1172 **assembly defects caused by loss of translational GTPase BipA**

1173 Growth (A) and sucrose gradient polysome analysis (B) of CFT073 wild type, isogenic Δ*bipA* and
1174 Δ*bipA*Δ*uup*, as well as CFT073 Δ*bipA* transformed with low-copy pSC vector expressing either BipA
1175 or ABCFs EttA, Uup, UheS and YbiT under control of constitutive promoter P_{tet}. All experiments
1176 were performed in filtered LB at 18°C and data are presented as geometric means ± standard
1177 deviation (n = 3).

1178

1179 **Fig. 6 | Expression of *E. coli* ABCF-EQ$_2$ mutants inhibits growth and protein synthesis**

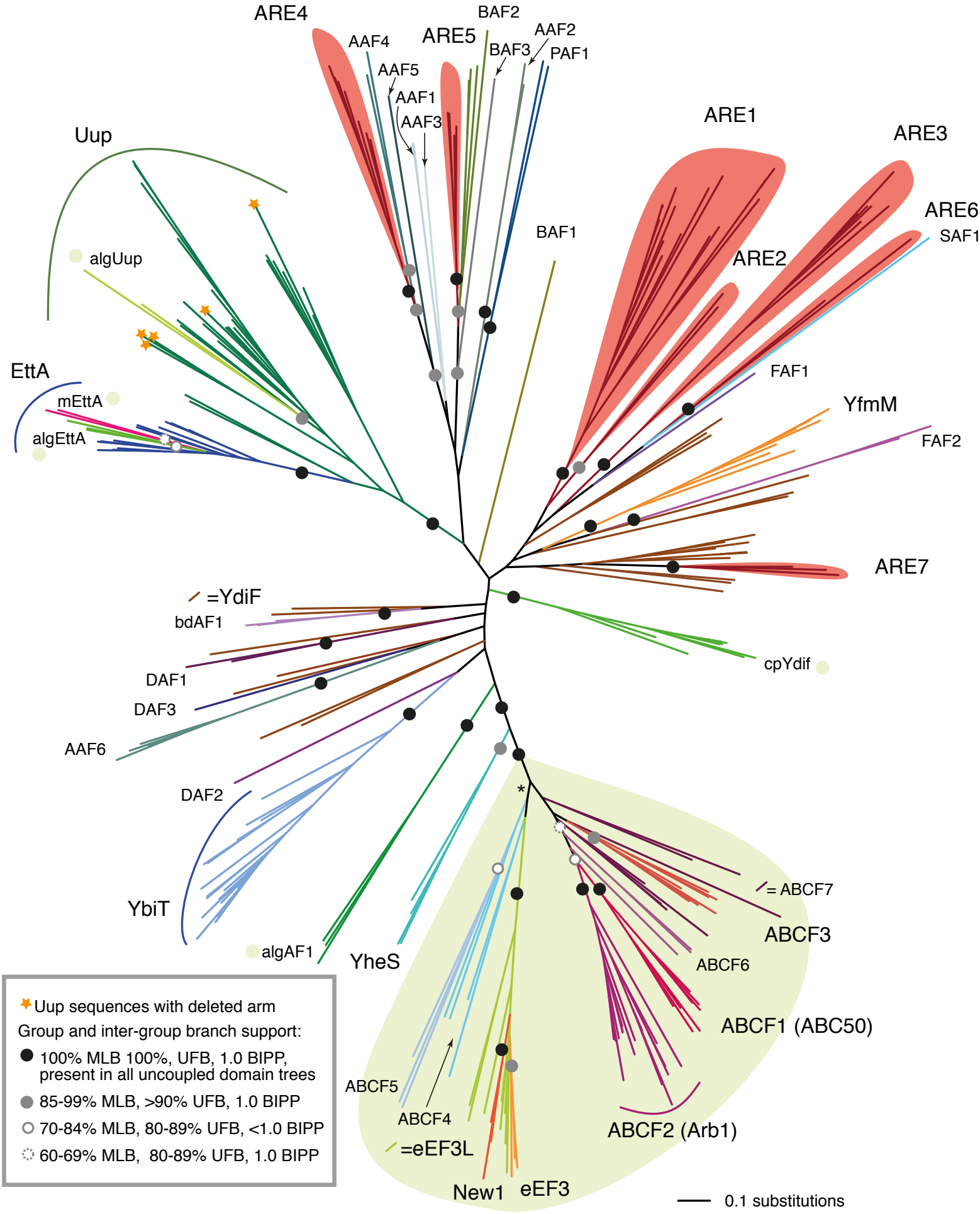
1180 Growth of wildtype *E. coli* BW2513 transformed with pBAD18 vector (grey trace) as well as *E. coli*
1181 BW2513 expressing either wild type (black trace) or EQ$_2$ mutants (red trace) of EttA (A), Uup (B),
1182 YbiT (C), and YheS (D) under the control of arabinose-inducible promoter P_{BAD}. Radiographs show
1183 the effect of wild type and EQ$_2$ ABCF expression on protein synthesis, as probed by pulse labeling
1184 with L-[^{35}S]-methionine. Expression was induced by the addition of L-arabinose to a final
1185 concentration of 0.2% at time point 0, and efficiency of incorporation was quantified by
1186 scintillation counting and visualised by autoradiography at 0 and 20-minute time points.
1187 Scintillation counting data are presented as geometric means ± standard deviation (n = 3). All
1188 experiments were performed at 37˚C in Neidhardt MOPS medium [105] supplemented with 0.4%
1189 glycerol as a carbon source. The inset cartoons are a representation of ABCF domains and sub-
1190 domains, as per the legend in the lower box.

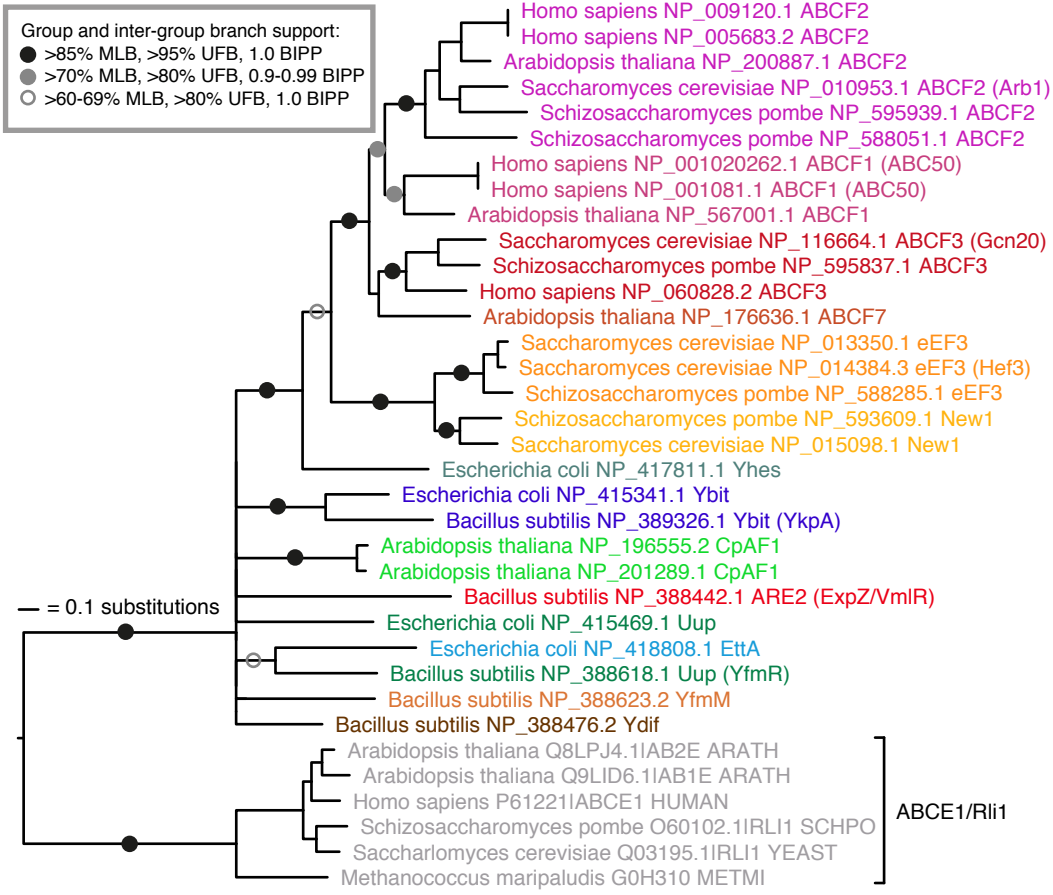
1191

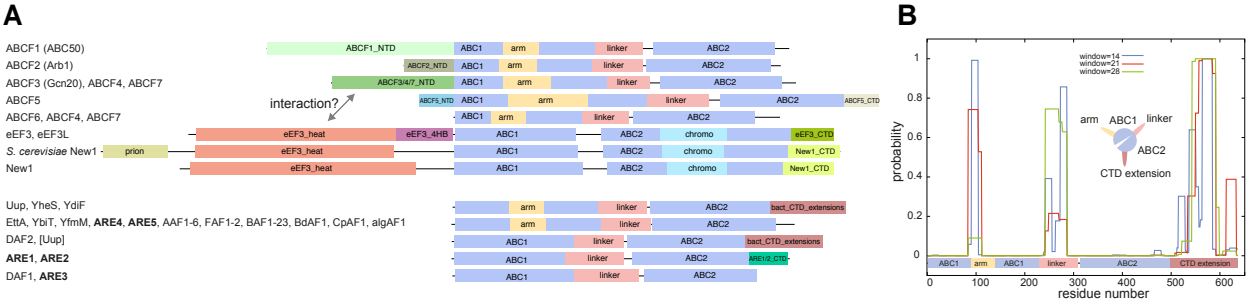
1192 **Fig. 7 | AREs tend to have relatively long linker regions that potentially extend towards the**
1193 **ribosome bound antibiotics**

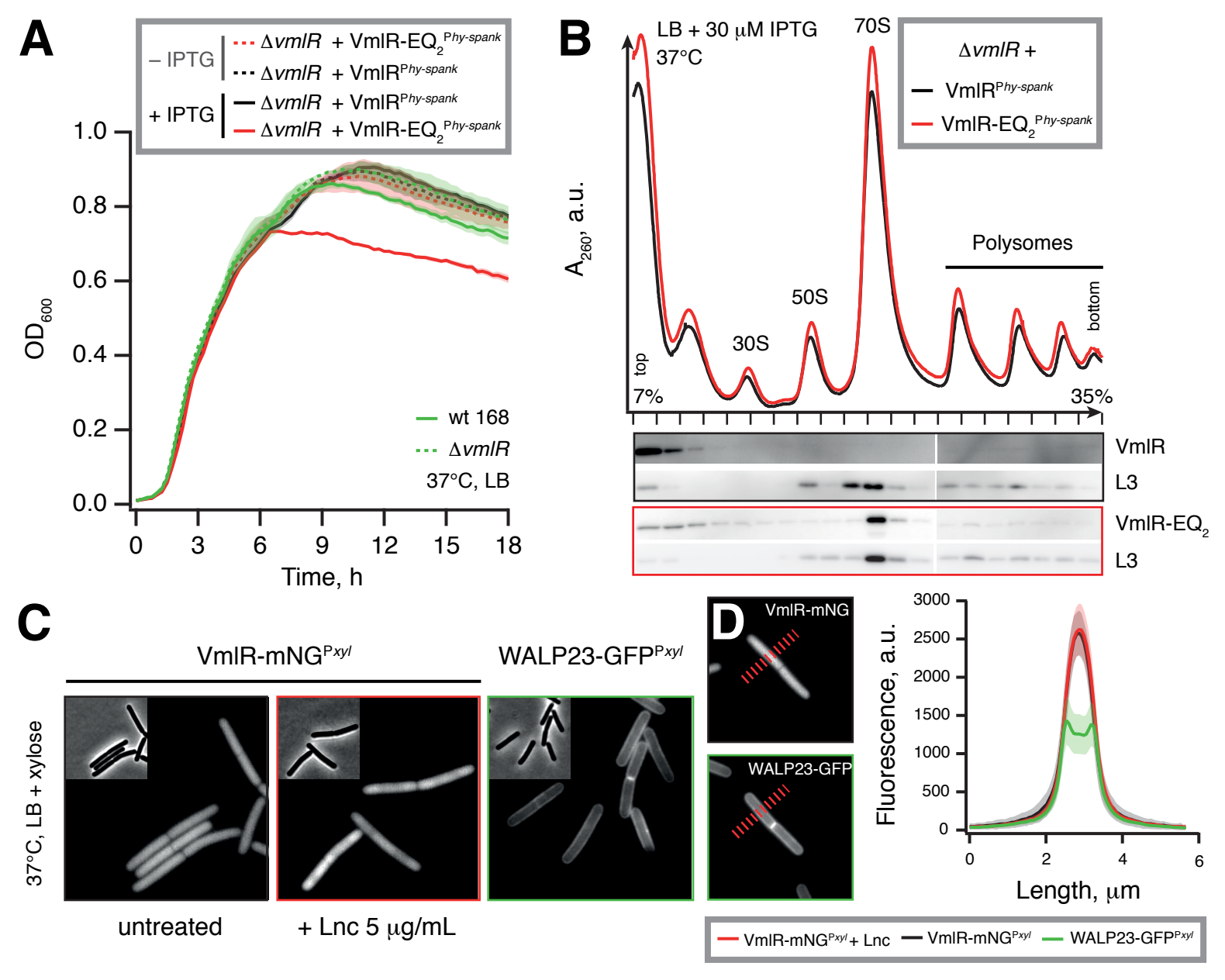
1194 (A) The structure of EttA and its interacting ribosomal components from PDB 3J5S [17] is shown
1195 alongside homology models of *Staphylococcus aureus* VgaA and *Entercoccus faecalis* LsaA, using
1196 3J5S as the template, with *de novo* modeling of the linker regions. The dotted circle shows the
1197 relative location of PTC-inhibiting antibiotics. Arm and linker regions are shaded in yellow and
1198 turquoise respectively. (B) Extracts from the multiple sequence alignment of *E. coli* and *B. subtilis*
1199 ABCFs, and representative AREs, containing the Arm (yellow shading) and Linker (turquoise
1200 shading) subdomains. Alignment numbering is according to the EttA sequence. A boxed region
1201 shows a region that is particularly rich in proline and polyproline in various ABCF family members.

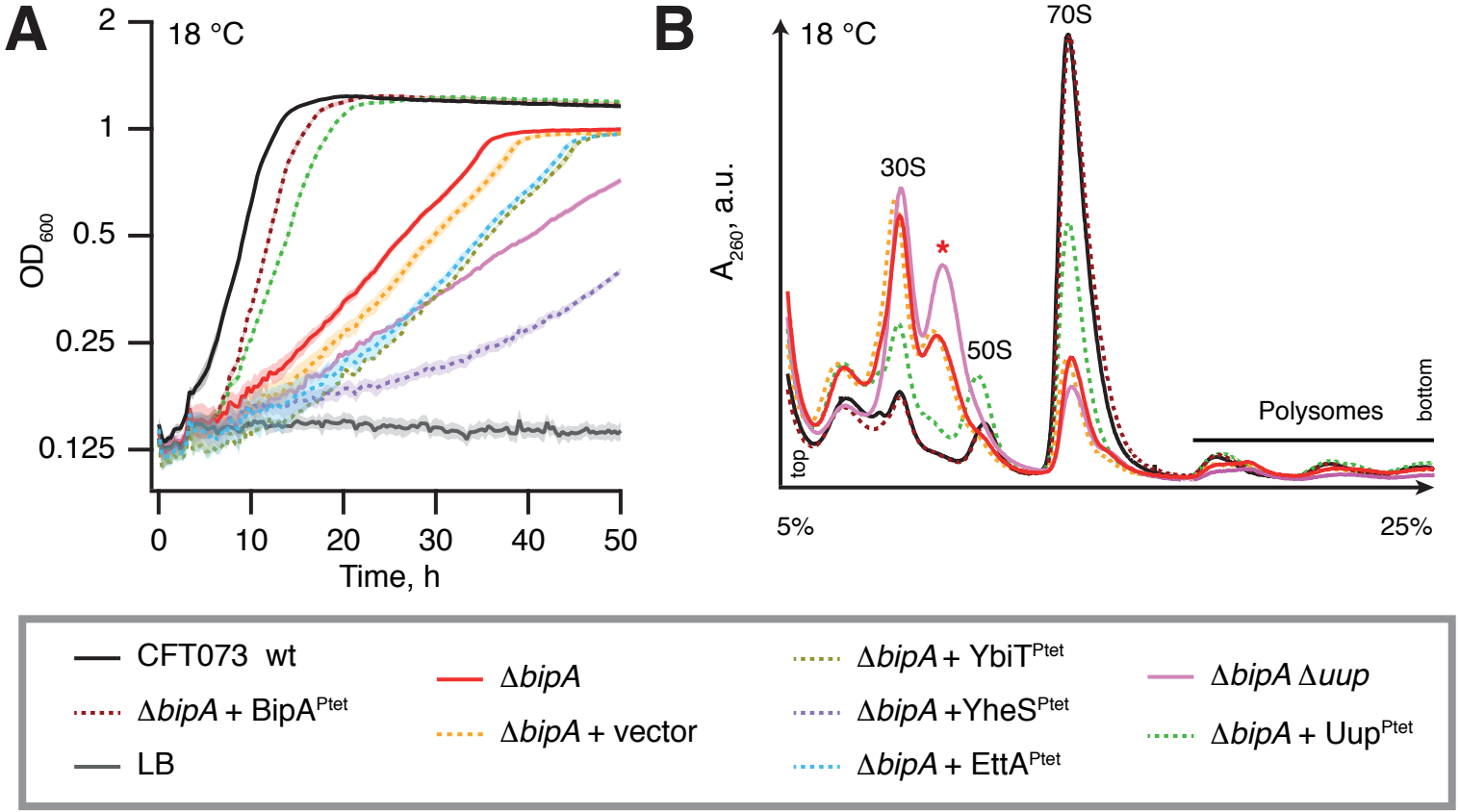
1202

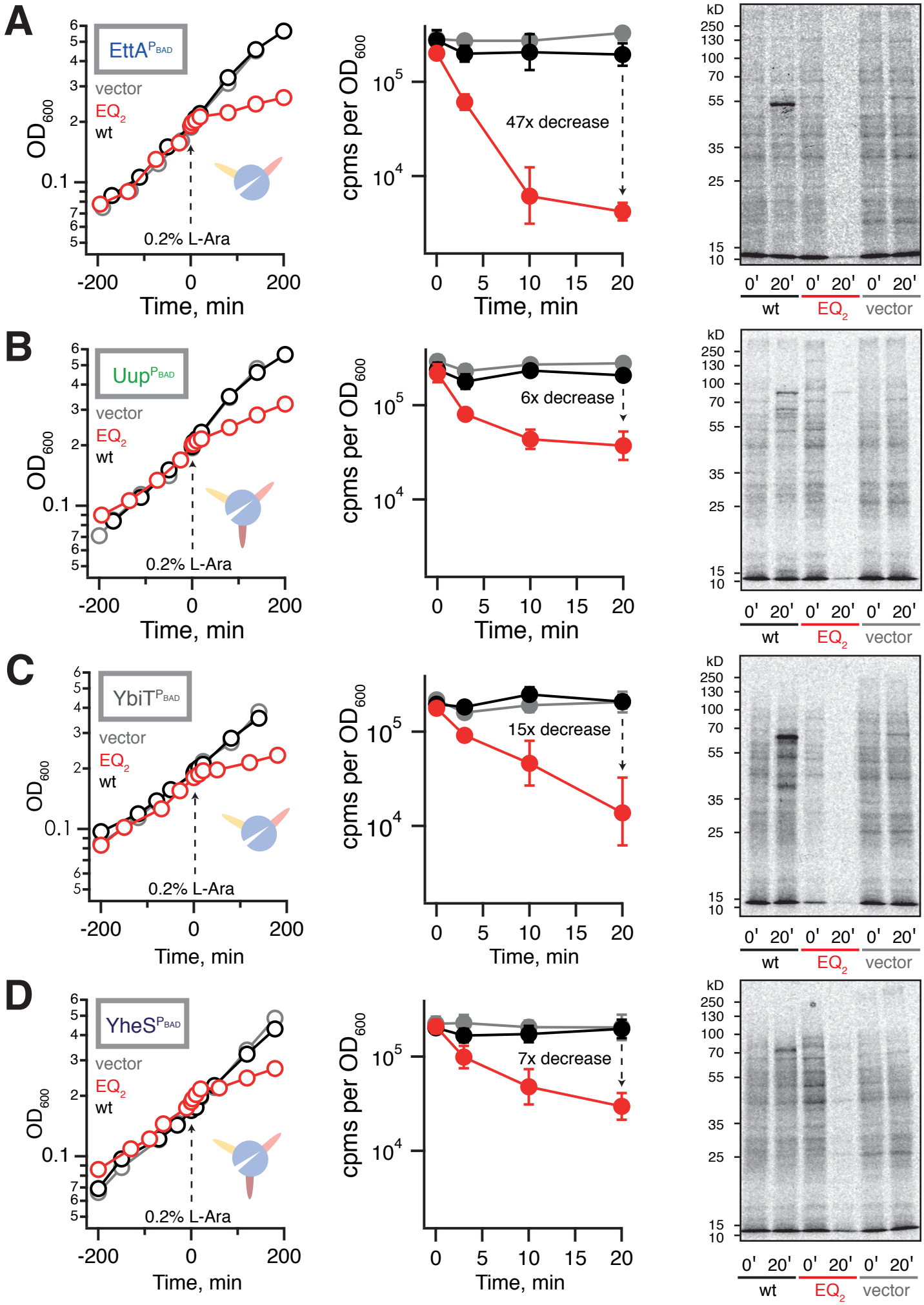

ARE4
AAF4
ARE5
BAF2
AAF2
BAF3
PAF1
AAF5
AAF1
AAF3
ARE1
ARE3
Uup
ARE6
SAF1
BAF1
ARE2
algUup
EttA
FAF1
mEttA
YfmM
algEttA
FAF2
ARE7
=YdiF
bdAF1
cpYdif
DAF1
DAF3
AAF6
DAF2
*
YbiT
= ABCF7
ABCF3
algAF1
YheS
ABCF6
ABCF1 (ABC50)
ABCF5
ABCF4
ABCF2 (Arb1)
=eEF3L
New1
eEF3
0.1 substitutions
Uup sequences with deleted arm
Group and inter-group branch support:
100% MLB 100%, UFB, 1.0 BIPP, present in all uncoupled domain trees
85-99% MLB, >90% UFB, 1.0 BIPP
70-84% MLB, 80-89% UFB, <1.0 BIPP
60-69% MLB, 80-89% UFB, 1.0 BIPP


Group and inter-group branch support:
>85% MLB, >95% UFB, 1.0 BIPP
>70% MLB, >80% UFB, 0.9-0.99 BIPP
>60-69% MLB, >80% UFB, 1.0 BIPP
Homo sapiens NP_009120.1 ABCF2
Homo sapiens NP_005683.2 ABCF2
Arabidopsis thaliana NP_200887.1 ABCF2
Saccharomyces cerevisiae NP_010953.1 ABCF2 (Arb1)
Schizosaccharomyces pombe NP_595939.1 ABCF2
Schizosaccharomyces pombe NP_588051.1 ABCF2
Homo sapiens NP_001020262.1 ABCF1 (ABC50)
Homo sapiens NP_001081.1 ABCF1 (ABC50)
Arabidopsis thaliana NP_567001.1 ABCF1
Saccharomyces cerevisiae NP_116664.1 ABCF3 (Gcn20)
Schizosaccharomyces pombe NP_595837.1 ABCF3
Homo sapiens NP_060828.2 ABCF3
Arabidopsis thaliana NP_176636.1 ABCF7
Saccharomyces cerevisiae NP_013350.1 eEF3
Saccharomyces cerevisiae NP_014384.3 eEF3 (Hef3)
Schizosaccharomyces pombe NP_588285.1 eEF3
Schizosaccharomyces pombe NP_593609.1 New1
Saccharomyces cerevisiae NP_015098.1 New1
Escherichia coli NP_417811.1 Yhes
Escherichia coli NP_415341.1 Ybit
Bacillus subtilis NP_389326.1 Ybit (YkpA)
Arabidopsis thaliana NP_196555.2 CpAF1
Arabidopsis thaliana NP_201289.1 CpAF1
Bacillus subtilis NP_388442.1 ARE2 (ExpZ/VmlR)
Escherichia coli NP_415469.1 Uup
Escherichia coli NP_418808.1 EttA
Bacillus subtilis NP_388618.1 Uup (YfmR)
Bacillus subtilis NP_388623.2 YfmM
Bacillus subtilis NP_388476.2 Ydif
= 0.1 substitutions
Arabidopsis thaliana Q8LPJ4.1IAB2E ARATH
Arabidopsis thaliana Q9LID6.1IAB1E ARATH
Homo sapiens P61221IABCE1 HUMAN
Schizosaccharomyces pombe O60102.1IRLI1 SCHPO
Saccharlomyces cerevisiae Q03195.1IRLI1 YEAST
Methanococcus maripaludis G0H310 METMI
ABCE1/Rli1

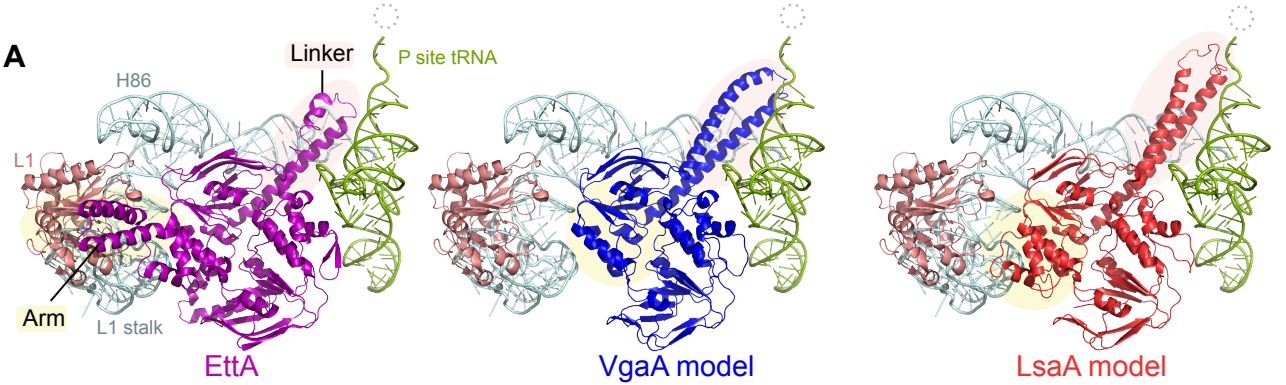

A
ABCF1 (ABC50)
ABCF2 (Arb1)
ABCF3 (Gcn20), ABCF4, ABCF7
ABCF5
ABCF6, ABCF4, ABCF7
eEF3, eEF3L
S. cerevisiae New1
New1
prion
ABCF1_NTD
ABCF2_NTD
ABCF3/4/7_NTD
ABCF5_NTD
ABCF5_CTD
eEF3_heat
eEF3_4HB
eEF3_CTD
New1_CTD
ABC1
arm
linker
ABC2
chromo
interaction?
Uup, YheS, YdiF
EttA, YbiT, YfmM, ARE4, ARE5, AAF1-6, FAF1-2, BAF1-23, BdAF1, CpAF1, algAF1
DAF2, [Uup]
ARE1, ARE2
DAF1, ARE3
bact_CTD_extensions
ARE1/2_CTD
B
window=14
window=21
window=28
probability
residue number
CTD extension


A
– IPTG
ΔvmlR + VmlR-EQ2 Phy-spank
ΔvmlR + VmlR Phy-spank
+ IPTG
ΔvmlR + VmlR Phy-spank
ΔvmlR + VmlR-EQ2 Phy-spank
OD600
Time, h
wt 168
ΔvmlR
37°C, LB
B
LB + 30 μM IPTG
37°C
ΔvmlR +
VmlR Phy-spank
VmlR-EQ2 Phy-spank
A260, a.u.
70S
50S
30S
Polysomes
top
bottom
7%
35%
VmlR
L3
VmlR-EQ2
L3
C
VmlR-mNG Pxyl
WALP23-GFP Pxyl
37°C, LB + xylose
untreated
+ Lnc 5 μg/mL
D
VmlR-mNG
WALP23-GFP
Fluorescence, a.u.
Length, μm
VmlR-mNG Pxyl + Lnc
VmlR-mNG Pxyl
WALP23-GFP Pxyl


A
18 °C
OD600
2
1
0.5
0.25
0.125
0
10
20
30
40
50
Time, h
B
18 °C
70S
30S
50S
*
A260, a.u.
Polysomes
top
bottom
5%
25%
CFT073 wt
ΔbipA + BipAPtet
LB
ΔbipA
ΔbipA + vector
ΔbipA + YbiTPtet
ΔbipA +YheSPtet
ΔbipA + EttAPtet
ΔbipA Δuup
ΔbipA + UupPtet



Predicted domain architectures:


A
Linker
P site tRNA
H86
L1
Arm
L1 stalk
EttA
VgaA model
LsaA model

B
EttA
Uup
YheS
YbiT
YdiF
YfmM
ARE1 VgaA
ARE2 VmlR
ARE3 LsaA
ARE4 OleB
ARE5 VarM
ARE6 SalA
ARE7 OptrA
44
166
242
323
proline hotspot