Deep mutational scanning reveals characteristics important for targeting of the tail-anchored protein Fis1

Abdurrahman Keskin[†], Emel Akdoğan, Cory D. Dunn[‡]

Department of Molecular Biology and Genetics, Koç University, Sarıyer, İstanbul, Turkey

[†] Current affiliation:

Department of Biological Sciences

Columbia University

New York, NY 10027

[‡] Corresponding author

Dr. Cory Dunn

Koç Üniversitesi

Fen Fakultesi

Rumelifeneri Yolu

Sarıyer, İstanbul 34450

Turkey

Email: cdunn@ku.edu.tr

Phone: +90 212 338 1449

Fax: +90 212 338 1559

Running head: Comprehensive analysis of Fis1 TA

Abbreviations: TA, tail anchor; OM, outer membrane; MAD, membrane-anchoring domain; 3-AT, 3-aminotriazole; CHX, cycloheximide

ABSTRACT:

Proteins localized to mitochondria by a carboxyl-terminal tail anchor (TA) play roles in apoptosis, mitochondrial dynamics, and mitochondrial protein import. To reveal characteristics of TAs that may be important for mitochondrial targeting, we focused our attention upon the TA of the *Saccharomyces cerevisiae* Fis1 protein. Specifically, we generated a library of Fis1p TA variants fused to the Gal4 transcription factor, then, using next-generation sequencing, revealed which Fis1p TA mutations inhibited membrane insertion and allowed Gal4p activity in the nucleus. Prompted by our global analysis, we subsequently analyzed the ability of individual Fis1p TA mutants to localize to mitochondria. Our findings suggest that the membrane-associated domain of Fis1p TA may be bipartite in nature, and we encountered evidence that the positively charged patch at the carboxyl-terminus of Fis1p is required for both membrane insertion and organelle specificity. Furthermore, lengthening or shortening the Fis1 TA by up to three amino acids did not inhibit mitochondrial targeting, arguing against a model in which TA length directs insertion of TAs at specific organelles. Most importantly, positively charged residues were more acceptable at several positions within the membrane-associated domain of the Fis1p TA than negatively charged residues. These findings, emerging from the first high-resolution analysis of an organelle targeting sequence by deep mutational scanning, provide strong, *in vivo* evidence that lysine and arginine can "snorkel," or become stably incorporated within a lipid bilayer by placing terminal charges of their side chains at the membrane interface.

INTRODUCTION:

Proteins inserted within the mitochondrial outer membrane (OM) by a carboxyl-terminal tail anchor (TA) are important for programmed cell death, mitochondrial protein import, and the control of mitochondrial shape and number (Wattenberg and Lithgow 2001). While ER-directed tail-anchored proteins can take advantage of a conserved set of soluble proteins and membrane-bound receptors (Denic *et al.* 2013; Johnson *et al.* 2013), the machinery targeting many TAs to mitochondria is yet to be discovered (Lee *et al.* 2014; Neupert 2015), and genetic and biochemical evidence suggest that spontaneous insertion of TAs at mitochondria may occur without the need for a translocation machinery (Setoguchi *et al.* 2006; Kemper *et al.* 2008). Furthermore, some tail-anchored proteins can be dual-localized to mitochondria and other organelles (Borgese and Fasana 2011), but how membrane specificity is controlled is unclear. TA targeting seems to depend, in general, upon incompletely defined structural characteristics of the TA rather than a defined consensus sequence (Beilharz 2003; Rapaport 2003; Borgese *et al.* 2007).

Genetic selection schemes using the organism *Saccharomyces cerevisiae* have been of high value in understanding how proteins reach their proper destination within eukaryotic cells. During such studies, a protein required for survival under selective conditions can be mislocalized, and thereby made inactive, by a targeting sequence utilized by the transport process being studied. Next, mutations that allow return of this mistargeted protein to a region of the cell at which it can perform its function are recovered under selective conditions. *Trans* factors related to protein targeting are identified by standard genetic approaches. Alternatively, *cis* mutations in the targeting sequence are revealed, typically by Sanger sequencing of individual fusion construct clones. Most prominently, this genetic approach to studying protein targeting and transport has been important in understanding protein transit to and through the endomembrane system (Deshaies and Schekman 1987; Robinson *et al.* 1988; Stirling *et al.* 1992). This approach has also been applied to the study of protein import and

3

83 export at the mitochondrial inner membrane (Jensen *et al.* 1992; Maarse *et al.* 1992; He

84 and Fox 1999).

85

86 Even with the availability of these powerful genetic strategies, a fine-grained analysis of

87 any single eukaryotic protein targeting signal has been lacking. However, with the

88 advent of next-generation sequencing, more comprehensive studies of protein

89 targeting sequences are possible. In this study, we successfully coupled genetic

90 selection to next-generation sequence analysis in order to define characteristics

91 important for localization of the tail-anchored Fis1 protein to the mitochondrial outer

92 membrane (OM).

93

RESULTS:

Localization to mitochondria via the Fis1 tail anchor prevents Gal4-mediated transcriptional activation

The TA of Fis1p is necessary (Mozdy *et al.* 2000; Beilharz 2003) and sufficient (Kemper *et al.* 2008; Förtsch *et al.* 2011) for insertion of this polypeptide into the mitochondrial OM. No cellular machinery involved in Fis1p insertion has been identified (Kemper *et al.* 2008). Fis1p has been suggested to reach a final topology in the outer membrane in which the amino-terminal bulk of the protein faces the cytosol, a very short and positively charged carboxyl-terminus protrudes into the mitochondrial intermembrane space, and the two are connected by a membrane-anchoring domain (MAD) passing through the OM (Mozdy *et al.* 2000). In developing our selection for TA mutations that diminish Fis1p targeting, we reasoned that fusion of a transcription factor to Fis1p would lead to insertion within the mitochondrial OM and a lack of nuclear function (Figure 1A). Mutations within the TA of Fis1p that prevent effective membrane insertion would, however, presumably allow the linked transcription factor to enter the nucleus, promote expression of its targets, and allow survival under specific selective conditions, provided that the fusion protein is not degraded, aggregated, or misdirected to another cellular location. Toward this goal, we generated a construct containing the Gal4 transcription factor at the amino terminal end of the polypeptide and full-length Fis1p at the carboxyl terminal end of the protein, since *S. cerevisiae* strains allowing titratable selection based upon nuclear entry and subsequent binding to Gal4p-responsive DNA elements are readily available. Superfolder GFP (Pédelacq *et al.* 2005) was placed between the Gal4 and Fis1 moieties and was visible at mitochondria upon overexpression of this fusion protein (Figure 1B). While Fis1p has been reported to be homogenously distributed on the mitochondrial surface (Mozdy *et al.* 2000), puncta containing Gal4-sfGFP-Fis1p are observed, perhaps due to the formation of heteromeric complexes with nuclear import components attempting to transport Gal4-sfGFP-Fis1p to the nucleus.

124

125   To assess failed Gal4-sfGFP-Fis1p targeting to mitochondria, we specifically took

126   advantage of the Gal4-driven *HIS3* and *URA3* auxotrophic markers in MaV203, a strain

127   commonly used for yeast-two-hybrid assays (Vidal *et al.* 1996a). Similar to cells

128   containing an empty vector, Gal4p fused to Fis1p was unable to allow proliferation on

129   medium lacking histidine and containing 20mM 3-aminotriazole (3-AT) to competitively

130   inhibit any His3p produced independently of Gal4p activation (Durfee *et al.* 1993) or in

131   medium lacking uracil (SC-Ura) (Figure 1C). However, the same Gal4-sfGFP-Fis1

132   polypeptide devoid of its TA [Gal4-sfGFP-Fis1(ΔTA)] permitted ample proliferation on

133   the same two selective media. This result indicated that our fusion protein could

134   translocate to the nucleus upon TA disruption and that any potential lipid binding

135   mediated by the cytosolic domain of Fis1p (Wells and Hill 2011) will not prevent genetic

136   assessment of TA localization.

137

138   We immediately sought mutations in the TA that would block targeting of Gal4-sfGFP-

139   Fis1p to the mitochondrial OM by isolating colonies spontaneously arising on medium

140   lacking uracil. Limited sequencing of the constructs encoding Gal4-sfGFP-Fis1p within

141   these isolates revealed at least seven nonsense and 19 frameshift mutations out of a

142   total of 32 plasmids analyzed. While these findings further validated the link between

143   proliferation under selective conditions and damage to the Fis1p TA, continued

144   encounters with nonsense and frameshift mutations would not be greatly informative

145   regarding sequential or structural determinants important for TA targeting. As

146   expected, we noted that release of the fusion construct from mitochondria following

147   mutation of the Fis1p TA leads to detection of sfGFP in the nucleus (C. Dunn,

148   unpublished observations).

149

150   In the strain used for our selective scheme, a Ura[+] phenotype requires greater Gal4-

151   dependent transcriptional activation than is required for a His[+] phenotype (Vidal *et al.*

152   1996b). Therefore, we reasoned that initial selection of TA mutants based on a His[+]

153   phenotype may provide informative mutations that weaken, but do not totally inhibit

154   membrane association. We used mutagenic PCR to generate altered TAs within the

155   context of a Gal4-sfGFP-Fis1 fusion protein. We then isolated four colonies that

156   proliferated upon SMM-His medium containing 20mM 3-AT, yet exhibited diminished

157   proliferation on SC-Ura medium when compared to cells expressing the Gal4-sfGFP-

158   Fis1(ΔTA) polypeptide. Sanger sequencing of the region encoding the TA of Gal4-

159   sfGFP-Fis1p within these colonies revealed one clone containing a V145E mutation

160   (amino acid numbering provided in this study will correspond to that of the unmodified,

161   full-length Fis1 protein; the necessary and sufficient region for mitochondrial

162   association of Fis1p begins at amino acid L129), two clones containing a L139P

163   mutation, and one clone harboring two mutations: L129P and V138A. Serial dilution

164   assays (Figure 2A) confirmed that V145E and L139P provided a less than maximal, but

165   still apparent Ura⁺ phenotype, with the V145E mutant allowing more rapid proliferation

166   on medium lacking uracil than the L139P mutant. The L129P/V138A mutant provided a

167   His⁺ phenotype, but could not drive uracil prototrophy, suggesting a less severe

168   localization defect than that exhibited by the other two mutant TAs. Interestingly, the

169   V145E mutation falls within the predicted MAD of the Fis1p TA (Figure S1A), consistent

170   with poor accommodation of a negatively charged amino acid within this hydrophobic

171   stretch of amino acids. Moreover, the Fis1p TA is predicted to be mostly alpha-helical

172   in nature, and some evidence suggests that helicity is an important determinant of TA

173   targeting to mitochondria (Wattenberg *et al.* 2007). Therefore, isolation of the

174   potentially helix-disrupting L139P replacement by selection supports the need for TA

175   helicity during mitochondrial targeting.

176

177   To further examine whether isolated TA mutations affect mitochondrial OM targeting,

178   we linked wild-type (WT) or mutant Fis1p TAs to the carboxyl-terminus of mCherry. To

179   assess mitochondrial localization of these fusion proteins, mitochondria were

180   specifically labelled with GFP targeted to mitochondria by the presequence of the

181   Cox4 protein (Sesaki and Jensen 1999). Importantly, the TA of Fis1p lacks information

182   required for mediating mitochondrial fragmentation (Habib *et al.* 2003), indicating that

183 the localization of these constructs should not be influenced by interaction partners of

184 full-length, inserted Fis1p.

185

186 The location of mCherry fused to Fis1p TAs was consistent with our genetic findings.

187 V145E and L139P mutations in the Fis1p TA led to substantial cytosolic and nuclear

188 accumulation of mCherry (Figure 2B). Moreover, the L129P/V138A TA, consistent with

189 its weaker activation of Gal4 targets in our selection system, provided still discernable

190 mitochondrial localization of the mCherry signal, but extraorganellar levels of this

191 mutant fusion protein appeared to be increased compared to mCherry fused to the WT

192 TA. These results suggest that our genetic approach is likely to allow recovery of

193 mutations affecting the ability of the Fis1p TA to moor proteins to the mitochondrial

194 outer membrane.

195

196 *A deep mutational scan for mutations potentially affecting Fis1p TA targeting*

197

198 Buoyed by our initial isolation of TA mutations affecting Fis1p localization, we decided

199 to take a more global approach to the analysis of the Fis1p TA. Using degenerate

200 primers and recombination-based cloning in *S. cerevisiae*, we sought construction of a

201 library consisting of all possible codons at every one of 27 amino acid positions within

202 the TA of Gal4-sfGFP-Fis1p. We then allowed the pool of cells containing mutant TAs

203 to divide four times under six possible culture conditions: no specific selection for *HIS3*

204 or *URA3* reporter activation (SC-Trp, selecting only for plasmid maintenance), selection

205 for *HIS3* activation at different 3-AT concentrations (SMM-Trp-His +0, 5, 10, or 20mM

206 3-AT), or selection for *URA3* activation (SC-Ura). Plasmid DNA was isolated from each

207 pool of mutants, and TA coding sequences were amplified and subjected to next-

208 generation sequencing. We focused our subsequent analysis only on clones within our

209 pools harboring a WT TA sequence or encoding TAs with only a single amino acid

210 change.

211

212   While all potential replacement mutations could not be detected within our starting

213   library (Figure S2), and some biases did exist at each TA position, the vast majority of

214   potential amino acid mutations were represented within our pool. 98.9% of potential

215   amino acid replacements were identified in the starting pool cultured in SC-Trp, and

216   95.9% of TAs with single mutations were represented by at least 10 counts.

217   Quantification of counts from all samples can be found in Table S1. When comparing

218   the mutant pool cultured in SC-Trp with selection in SMM-Trp-His without added 3-AT,

219   there was no appreciable difference in the relative abundance of most mutant TAs,

220   including truncation mutations expected to totally prevent mitochondrial targeting of

221   Gal4-sfGFP-Fis1p (Figure S3A). Such a result is consistent with 'leaky' expression of

222   *HIS3* independent of Gal4-driven activation (Durfee *et al.* 1993). However, upon

223   addition of 3-AT at concentrations of 5mM (Figure S3B), 10mM (Figure S3C), or 20mM

224   (Figure 3) to medium lacking histidine, there were substantial shifts in the composition

225   of the mutant pools toward specific amino acids, prompting further experiments that

226   we describe below. The pool cultured in SC-Ura medium showed very strong selection

227   for nonsense mutations within the TA (Figure S3D), but less prominent biases among

228   amino acids. When considering our initial findings, in which recovery of uracil

229   prototrophs by our genetic scheme led to a high recovery of frameshift and nonsense

230   mutations, assessment of *HIS3* activation seems more informative regarding

231   determinants of Fis1p TA targeting than competition assays performed in the more

232   strongly selective medium lacking uracil. Independently of the primary amino acid

233   sequence, the specific codons used to direct synthesis of a protein can affect that

234   polypeptide's translation rate, folding, or even synthesis of its transcript (Yu *et al.* 2015;

235   Zhou *et al.* 2016). While the data are more 'noisy' due to a lack of representation of

236   certain codons at each position, codons encoding the same amino acid generally

237   acted in concert with one another within our selection scheme (Figure S4). Therefore,

238   our focus remained upon the amino acid sequence of library variants rather than on

239   codon sequence.

240

241

9

Proline disruption of Fis1p tail anchor targeting

Previous analyses of various tail-anchored mitochondrial proteins similar in general structure to Fis1p suggested that no primary consensus sequence is required for TA insertion (Horie *et al.* 2002; Beilharz 2003; Rapaport 2003). While meaningful alignment of Fis1p TAs across species is difficult due to constraints in amino acid choice within hydrophobic domains and as a consequence of the apparently variable Fis1p TA length across species, only G131 (as pertains to the *S. cerevisiae* Fis1p sequence) might be considered highly conserved (Figure S1B). Our comprehensive analysis supports the idea that no consensus sequence within the Fis1p TA is necessary to achieve membrane insertion, since most amino acid replacements within the necessary and sufficient region required for Fis1p targeting, including at position G131, fail to lead to notable selectable reporter activation (Figure 3). Consequently, we focused our subsequent analysis on structural characteristics of the TA that might be most important for mitochondrial OM targeting.

The recovery of the L139P mutation during preliminary selection for Fis1p TA mutations indicated that proline may not be acceptable within the hydrophobic core of the Fis1p TA. Our deep mutational scan of the Fis1p TA in SMM-Trp-His + 20mM 3-AT (Figure 3) also strongly indicated that proline insertion across many positions disrupted mitochondrial TA localization. When focusing specifically upon those mutants that were in the top 75% most commonly counted variants in the starting pool (>126 counts) and enriched at least four-fold in SMM-Trp-His + 20mM 3-AT, 12 of 33 missense mutations within this set were proline replacements (Figure 4), further indicating failure of TA targeting following placement of proline at many TA positions.

Subsequently, we carried out directed experiments to further examine poor accommodation of proline within the Fis1p TA. We further studied the L139P mutant that was initially isolated during selection for Fis1p TA targeting mutants, and we also generated four additional, individual proline replacements within Gal4-sfGFP-Fis1p and

10

272 tested for Gal4-driven reporter activation. Newly constructed V134P, G137P, A140P,

273 and A144P substitutions, consistent with our larger scale analysis (Figure 3 or Figure

274 4), provided ample proliferation on medium selective for *HIS3* activation (Figure 5A).

275 Upon visualization of mCherry fused to these Fis1p TA mutants, V134P, L139P, A140P,

276 and A144P replacements all clearly diminished mCherry localization to mitochondria

277 (Figure 5B). Our results suggest that the secondary structure of the Fis1p TA is

278 important for its function, and that disruption of helicity at many locations may make

279 targeting to the mitochondrial OM unfavorable. We also noted, however, variability in

280 the propensity of proline replacements to disrupt Fis1p TA targeting. Most prominently,

281 the G137P substitution allowed apparently normal targeting to mitochondria, as

282 assessed by Gal4p-driven reporter activation (Figures 3 and 5A) and by microscopic

283 analysis (Figure 5B), potentially suggesting the existence of two separable, helical

284 segments within the Fis1p TA rather than a single, monolithic helix.

285

286 We considered the possibility that targeting of these mutant TAs is not impeded, but

287 rather a lack of stability within the lipid bilayer leads to ejection from the mitochondrial

288 OM by a quality control mechanism. Recently, the yeast Msp1 protein and its human

289 ortholog ATAD1, have been identified as potential 'extractases' that can remove

290 improperly folded or mislocalized proteins from the mitochondrial OM (Okreglak and

291 Walter 2014; Chen *et al.* 2014). We tested whether any of tail-anchored fluorescent

292 proteins containing proline replacements and not strongly localized to mitochondria

293 could recover mitochondrial localization in cells deleted of Msp1p. However, deletion

294 of Msp1p did not lead to relocalization of any tested mutant mCherry-TA fusion protein

295 to mitochondria (Figure S5A), providing no evidence for proper targeting, then

296 subsequent removal, of assayed Fis1p TAs harboring proline replacements.

297

298 In mammalian cells, the ubiquilin protein UBQLN1 can bind to tail-anchored proteins

299 that fail to target mitochondria, expediting their degradation (Itakura *et al.* 2016). We

300 wondered whether the yeast UBQLN1 orthlog Dsk2p (Chuang *et al.* 2016) might bind to

301 mutant Fis1p TAs, sequestering them in the cytosol and preventing mitochondrial

302   insertion. However, deletion of Dsk2p did not change the cellular location of

303   mislocalized mCherry-TA fusion proteins carrying proline substitutions (Figure S5B).

304

305   *The positively charged carboxyl-terminus of the tail anchor allows both efficient and*

306   *specific targeting to mitochondria*

307   Analysis of the data from our deep mutational scan suggested that nonsense

308   mutations throughout much of the TA can allow Gal4-sfGFP-Fis1p to move to the

309   nucleus and activate transcription (Figures 3, 4, and S3). Stop codons placed within

310   the highly charged RNKRR pentapeptide near the carboxyl-terminus of Fis1p, however,

311   seem to permit some membrane localization, as reported by proliferation rates in

312   selective medium. Therefore, we examined the behavior of a R151X mutant, which

313   lacks all charged amino acids following the predicted MAD (X represents mutation to a

314   stop codon). Supporting partial localization to a cellular membrane, the R151X mutant

315   of Gal4-sfGFP-Fis1p did not activate Gal4-controlled expression of *HIS3* to the same

316   extent of a Gal4-sfGFP-Fis1p construct lacking the entire TA (Figure 6A). Consistent

317   with those results, the R151X TA directed mCherry to intracellular organelles (Figure

318   6B). However, along with some apparent mitochondrial association, the R151X TA was

319   also clearly localized to the ER (Figure 6C). Interestingly, ER localization of mCherry

320   fused to the R151X TA was not completely dependent upon Get3p, a receptor for ER

321   tail-anchored proteins (Schuldiner *et al.* 2008), suggesting the existence of an

322   alternative pathway for localization of the R151X TA to the ER (Figure S6A and S6B).

323   Similarly, R151X TA could be targeted to the ER in *get2Δ* mutants, lacking a receptor

324   for several ER-directed tail-anchored proteins (Figure S6C). However, puncta that

325   might indicate clustering with Get3p (Schuldiner *et al.* 2008) were apparent in some

326   *get2Δ* cells, raising the unexplored possibility that Get3p can bind at least some

327   portion of the cytosolic pool of R151X TA.

328   Together, our results demonstrate that the charged amino acids at the carboxyl-

329   terminus of the TA provide organelle specificity, yet are not totally required for

330   membrane localization. These findings are consistent with previous results reporting

331 that positively charged amino acids following the MAD of Fis1p allow targeting of TAs

332 specifically to mitochondria rather than ER (Isenmann *et al.* 1998; Kuroda *et al.* 1998;

333 Borgese *et al.* 2001; Stojanovski *et al.* 2004), perhaps by decreasing total TA

334 hydrophobicity, a factor important for determining the final location of a TA (Beilharz

335 2003; Wattenberg *et al.* 2007).

336

337 Since the R151X variant of Gal4-sfGFP-Fis1p activated Gal4-driven reporters, yet was

338 at least partially localized to the ER, we wondered if ER localization of any protein

339 fused to Gal4 might similarly lead to Gal4-driven transcription due to the physical

340 continuity between the ER and nuclear envelope. Therefore, we examined the TA of the

341 human FIS1 protein (hFIS1), since full-length hFIS1 can localize to ER in *S. cerevisiae*

342 (Stojanovski *et al.* 2004). Indeed, we found that mCherry fused to the hFIS1 TA was

343 poorly localized to mitochondria (Figure S7A) and abundant at the ER (Figure S7B).

344 However, a fusion protein consisting of Gal4 fused to the TA of hFIS1 did not provide

345 *HIS3* activation (Figure S7C), indicating that the hFIS1 TA is quantitatively membrane-

346 targeted in *S. cerevisiae* and that activation of Gal4-dependent reporters upon removal

347 of the positively charged carboxyl-terminus from Gal4-sfGFP-Fis1p is unlikely to be a

348 consequence of ER localization.

349

350 Due to the mislocalization of the hFIS1 TA, we then investigated the possibility that

351 other mitochondrial TA proteins from human would be targeted improperly in *S.*

352 *cerevisiae*. We fused mCherry to the TA of human BAX, a region that is sufficient for

353 insertion at the mammalian mitochondrial OM (Schinzel *et al.* 2004). While mCherry

354 signal was diminished in comparison with other mCherry fusion proteins examined in

355 this study and expressed under the same promoter, mCherry fused to the BAX TA was

356 properly targeted to mitochondria (Figure S7D). Gal4 fused to the BAX TA did not

357 activate selectable reporters (C. Dunn, unpublished results), further indicating effective

358 mitochondrial targeting mediated by the BAX TA.

359

13

360   *Extension or reduction of Fis1p TA length does not affect targeting to mitochondria*

361

362   Targeting of tail-anchored proteins to specific membranes has been suggested to

363   depend, at least in part, upon the specific length of the MAD within the TA (Isenmann

364   *et al.* 1998; Horie *et al.* 2002). We reasoned that the region within the MAD at which

365   prolines do not strongly disrupt mitochondrial targeting, as defined by our analysis,

366   may be amenable to the insertion or deletion of new amino acids, thereby allowing us

367   to test the relationship between Fis1p TA length and mitochondrial targeting. We

368   inserted one (∇1A), two (∇2A), or three (∇3A) additional alanines between A135 and

369   G136 within the TA of Gal4-sfGFP-Fis1p, but none of these mutant constructs led to

370   apparent *HIS3* (Figure 7A) activation. We then analyzed the location of mCherry fused

371   to a Fis1p TA carrying these same insertions. All constructs were localized properly to

372   mitochondria (Figure 7B).

373

374   Next, we deleted one (ΔG136), two (ΔA135-G136), or three (ΔA135-G137) amino acids

375   within the Fis1p MAD and performed similar assays. Like our insertion mutants,

376   deletion mutants were apparently targeted to a membrane, as assessed by Gal4-driven

377   reporter transcription (Figure 7A). Moreover, mCherry remained localized to

378   mitochondria when up to three amino acids were deleted (Figure 7C).

379

380   Disruption of the ER-localized Spf1 protein reduces the contrast in ergosterol content

381   between the ER and the mitochondrial OM (Krumpe *et al.* 2012). Consequently, TAs

382   normally localized to mitochondria are mistargeted to the ER upon Spf1p deletion. The

383   sterol concentration of membranes can determine bilayer thickness (Dufourc 2008),

384   raising the possibility that insertions or deletions may allow mitochondrial TAs to once

385   again prefer mitochondrial OM targeting over ER localization in a background lacking

386   Spf1p. However, mCherry linked to insertion or deletion mutants of the Fis1p TA

387   remained poorly targeted to mitochondria (Figure S8A) and prominently localized to the

388   ER (Figure S8B) in mutants lacking Spf1p.

389

390   Together, our results demonstrate that the Fis1p TA is properly localized to

391   mitochondria even when its length is substantially altered.

392

393   *Positively charged amino acids are more acceptable than negatively charged amino*

394   *acids within the predicted transmembrane domain of the Fis1p tail anchor*

395

396   Our deep mutational scan of the Fis1p TA demonstrated that Gal4-sfGFP-Fis1p was

397   generally able to activate gene expression when aspartate or glutamate was placed

398   within the MAD (Figure 3). In fact, upon examination of those amino acid replacements

399   found within the top three quartiles of counts in the initial library and also enriched at

400   least four-fold upon culture in SMM-Trp-His + 20mM 3-AT, 18 of 33 missense

401   mutations were aspartate or glutamate substitutions (Figure 4). We were surprised to

402   find that placement of positively charged arginine or lysine residues appeared to be

403   much more acceptable within the MAD of Fis1p than aspartate or glutamate; none of

404   the amino acid substitutions within this high-count, high-enrichment set were by lysine

405   or arginine.

406

407   To further pursue the possibility that positively charged amino acids can be

408   accommodated within the Fis1p MAD, we mutated four amino acids within the

409   hydrophobic stretch of the Fis1p TA to aspartate, glutamate, lysine, or arginine.

410   Specifically, we generated amino acid replacements at positions V132, A140, A144, or

411   F148, then retested these mutants under selection for Gal4-sfGFP-Fis1p transcriptional

412   activity. The results from our global analysis were verified, with aspartate and

413   glutamate mutations providing stronger reporter activation than lysine and arginine

414   mutations (Figure 8A). Only the A144D mutation provided sufficient Gal4 activation for

415   proliferation on medium lacking uracil (Figure S9A) after two days of incubation,

416   suggesting a very severe TA localization defect caused by this TA mutation. We noted

417   that these mutant Gal4-sfGFP-Fis1 constructs exhibit altered behavior at different

418   temperatures. For example, lysine and arginine substitutions at positions A140, A144,

419   or F148 led to reduced proliferation at 37°C under conditions selective for *HIS3*

15

420   activation (Figure S9B) when compared to the same TA substitutions assessed at 18°C

421   (Figure S9C) or 30°C (extended incubation, Figure S9D). This outcome is consistent

422   with the idea that altered phospholipid dynamics at different temperatures may lead to

423   consequent changes in TA insertion efficiency (de Mendoza and Cronan 1983).

424

425   We then tested the ability of these Fis1p TAs containing charge substitutions to

426   promote mitochondrial localization of mCherry. At V132 and F148, positions within the

427   MAD nearer to the putative water-lipid bilayer interface, mutation to positively charged

428   amino acids allowed abundant localization to mitochondria (Figure 8B). In contrast,

429   mutation to negatively charged amino acids clearly hindered mitochondrial targeting.

430   We noted that F148D and F148E replacements hampered mitochondrial localization

431   more severely than V132D and V132E replacements, consistent with phenotypic

432   testing of Gal4-sfGFP-Fis1 fusion proteins. At position A144, lying more deeply within

433   the predicted MAD, all charge mutations inhibited mCherry targeting, yet A144K and

434   A144R substitutions allowed some mCherry localization at mitochondria while A144D

435   and A144E substitutions appeared undetectable at mitochondria. Finally, no

436   mitochondrial signal was apparent for any of the charge mutants tested at position

437   A140 of the Fis1p TA. However, A140K and A140R mutants differed from A140D and

438   A140E mutants by localizing to other membranes within the cell, including the plasma

439   membrane, rather than providing a diffuse cytosolic signal. Msp1p removal did not

440   permit relocalization of tail-anchored fluorescent proteins to mitochondria (Figure S5C),

441   supporting the idea that charge replacements within the Fis1p TA lead to a failure of

442   association with the OM rather than enhanced removal from mitochondria. Removal of

443   UBQLN1 ortholog Dsk2p also had no discernable effect on mCherry-TA mutant

444   localization (Figure S5D).

445

446   Negligible Fis1p activity at mitochondria is apparently sufficient to promote

447   mitochondrial fission (Habib *et al.* 2003; Krumpe *et al.* 2012), suggesting that even

448   minimal localization and related functionality at the OM would be detectable by

449   functional assays. Interestingly, Fis1p variants harboring charged residues, positive or

450   negative, at positions V132, A144, or V148, with the exception of Fis1p carrying the

451   A144D mutation, provided at least some Fis1p function, as indicated by microscopic

452   (Figures S10A and S10B) and genetic (Figure S10C) assays. Such a result is consistent

453   with the finding that only the A144D charge mutant provides sufficient *URA3* activation

454   for rapid proliferation on medium lacking uracil (Figure S9A).

455

456   Together, our results demonstrate that positively charged amino acids within the MAD

457   can better promote Fis1p localization than negatively charged amino acids, but that

458   even negatively charged amino acids can be accommodated within the MAD and lead

459   to a detectable level of mitochondrial targeting.

460

DISCUSSION:

Using a deep mutational scanning approach, we explored structural characteristics of the Fis1p TA important for targeting to the mitochondrial outer membrane. To our knowledge, this work is the first application of this technique to the study of a eukaryotic organelle targeting signal. Deep mutational scanning, when coupled to an effective method of screening or selection, is very cost- and time-effective (Araya and Fowler 2011; Fowler and Fields 2014; Boucher *et al.* 2014). Mutant library generation, subsequent pool selection, and next-generation sequencing can be completed in just a few months. Consequently, this approach generates far more useful data over a shorter duration than, for example, alanine scanning mutagenesis or low-throughput genetic selection followed by Sanger sequencing. Deep mutational scanning has recently been applied successfully to other areas of study, such as membrane protein insertion within bacteria (Elazar *et al.* 2016), tumor suppressor structure and function (Starita *et al.* 2015), and the relationship between a polypeptide's evolutionary path and its present fitness (Hietpas *et al.* 2011; Melamed *et al.* 2013). As is true for most genetic screens and selections, further experimentation, as performed here by testing mutant Fis1p TAs in microscopic and functional assays, is required in order to make full use of a larger-scale dataset. We also note that some mutations affecting specific insertion of the Fis1p TA at mitochondria may have been missed during our global analysis: transcription driven by Gal4-sfGFP-Fis1p indirectly reports on TA targeting and may be affected by mutant stability, potential sequestration at other intracellular membranes, or ability of the fusion protein to enter and function within the nucleus.

In vivo evidence of the ability of positively charged amino acids to "snorkel" their side chains to surface of the lipid bilayer

Because there is a high energetic barrier to placing any charged residue into a lipid bilayer (Cymer *et al.* 2015), we were initially surprised to find that positively charged amino acids within the MAD of the Fis1p TA promoted mitochondrial targeting far

491 better than negatively charged amino acids. However, it has been suggested that

492 lysine and arginine within a MAD can "snorkel," or integrate the uncharged portion of

493 their side chain into the hydrophobic milieu of the lipid bilayer while locating positive

494 charges near polar head groups at the interface between the membrane and aqueous

495 environment (Gavel *et al.* 1988; Segrest *et al.* 1990). Some phospholipid head groups

496 carry a net negative charge, potentially providing further favorability to snorkeling lysine

497 and arginine. The shorter hydrophobic portion of aspartate and glutamate, however,

498 may not permit the side chain to easily reach the membrane interface to remove the

499 negative charge from the hydrophobic environment, and if they do reach the lipid

500 bilayer surface, charge repulsion might make such a conformation unfavorable.

501

502 Most experimental support for amino acid snorkeling has been obtained using *in vitro*

503 approaches or has been gathered during structural studies (Monné *et al.* 1998; Long *et*

504 *al.* 2005; Kim *et al.* 2012; Öjemalm *et al.* 2016), and little *in vivo* evidence has been

505 reported in support of this phenomenon. Our comprehensive study of the Fis1p TA, a

506 region dedicated only to the process of membrane integration (Habib *et al.* 2003;

507 Kemper *et al.* 2008), strongly supports the ability of lysine or arginine to be

508 accommodated by snorkeling at numerous positions within the Fis1p MAD. We note

509 that snorkeling, if operative for positive charges within the Fis1p TA, may not be

510 permitted within the context of all mitochondrial TAs: replacement of S184 within the

511 BAX TA by lysine does not seem to permit mitochondrial localization of this protein

512 (Nechushtan *et al.* 1999). We also note that while the Fis1p TA is typically modelled as

513 bitopic, or reaching through the mitochondrial outer membrane from the cytosol to the

514 intermembrane space, the possibility that the Fis1p TA lies side-long in the outer

515 membrane has not been ruled out. However, snorkeling is thought to be possible for a

516 MAD found in either the monotopic or the bitopic configuration (Strandberg and Killian

517 2003).

518

519 Further comprehensive mutational scans of MADs may further substantiate the

520 concept of snorkeling. Interestingly, since those Fis1p TAs mutated to contain

521 positively charged amino acids within the MAD were often not targeted to the OM with

522 full efficiency, one might imagine a scenario in which dual localization of a protein with

523 a single MAD to cytosol and to intracellular membranes can be easily evolved by the

524 appearance of positively charged amino acids at positions previously lacking such

525 charges. Moreover, some prediction methods for MADs might be considered overly

526 conservative, especially for prediction programs emphasizing amino acid charge.

527 Therefore, the accumulating evidence of amino acid snorkeling should prompt the

528 development of improved algorithms that explicitly consider positively charged amino

529 acids to be acceptable within certain positions of a predicted MAD.

530

531 *The membrane-associated domain of the Fis1p tail-anchor may consist of two*

532 *separable segments*

533

534 Computational analyses suggest that the Fis1p TA is mostly alpha-helical in nature

535 (Buchan *et al.* 2013; Drozdetskiy *et al.* 2015). We found that proper localization of the

536 Fis1p TA likely requires its predicted alpha-helicity within the MAD, since proline, which

537 is known to break or kink helices (Senes *et al.* 2004), profoundly disrupts targeting

538 when substituted at many positions throughout this hydrophobic region. However, we

539 found that replacement by proline is more acceptable at a specific location, G137, than

540 proline mutations found at many other locations within this region, potentially indicating

541 that the Fis1p MAD is bipartite in nature. Further supporting a bipartite structure of the

542 Fis1p MAD, insertion of new amino acids between A135 and G136 did not apparently

543 affect mitochondrial TA targeting. Moreover, mutations toward the carboxyl-terminal

544 end of the Fis1p MAD appear to affect mitochondrial targeting more drastically, as

545 reported by deep mutational scanning, than mutations toward the amino terminus of

546 the transmembrane segment. Previous analysis of the rat OMP25 TA also support a

547 bipartite structure of the MAD, with higher sensitivity to mutation nearer to the

548 carboxyl-terminal end of this hydrophobic stretch (Horie *et al.* 2002). In addition,

549 prolines are found within the MAD of the mammalian OMb and OMP25 TAs. These

550 results suggest that those prolines might demarcate the boundary between distinct

20

551 structural regions of the targeting sequence. On the other hand, prolines within a single

552 helical segment may simply be more easily housed within an alpha-helix when buried

553 deep in the lipid bilayer (Li *et al.* 1996; Senes *et al.* 2004) and may not reflect two

554 separable MAD segments. If this is the case, prolines found in mitochondrial TAs might

555 indicate the portion of the TA found at the midpoint of the OM.

556 Glycine is not preferred within alpha-helices (Chou and Fasman 1974; O'Neil and

557 DeGrado 1990) as a consequence of its conformational flexibility. However, our deep

558 mutational scan does not indicate reduced membrane targeting when most amino

559 acids within the Fis1p TA are individually mutated to glycine. This might be surprising

560 in light of the pronounced effects provided by several proline replacement mutations

561 throughout this domain. Yet, glycine may not be as disruptive for alpha-helices found

562 within a lipid bilayer environment when compared with alpha-helices of soluble

563 proteins, due to better intra-helical hydrogen bonding within the hydrophobic

564 environment of the membrane (Dong *et al.* 2012). Indeed, four glycines already exist

565 within the *S. cerevisiae* Fis1p TA, and the TAs of Fis1p orthologs are also littered with

566 glycines (Stojanovski *et al.* 2004) and (Figure S1B), further indicating that glycines are

567 less disruptive of the Fis1p TA than prolines. Interestingly, GXXXG motifs, and other

568 similarly spaced small amino acids like alanine and serine, can promote helix packing

569 within lipid bilayers (Russ and Engelman 2000; Senes *et al.* 2004; Gimpelev *et al.* 2004).

570 However, our findings indicate that the sole Fis1p GXXXG motif and a nearby AXXXA

571 motif do not play a significant role in targeting of the Fis1p TA to membrane.

572 Furthermore, a GXXXXG motif may mediate multimerization of the mitochondrial OM

573 protein Bcl-XL (Ospina *et al.* 2011), and such a motif can be found within the *S.*

574 *cerevisiae* Fis1p and several of its orthologs. However, the GXXXXG motif also seems

575 to have little to no role in insertion, at least as determined by the ability of Gal4-sfGFP-

576 Fis1p to activate gene expression.

577

578   *Does a machinery dedicated to mitochondrial tail anchor insertion exist?*

579

580   So far, no cellular machinery dedicated to the insertion of mitochondrial tail anchored

581   proteins has been revealed. Evidence supporting the existence of such a machinery

582   includes saturable import of mitochondrial TA-containing proteins in mammalian cells

583   (Setoguchi *et al.* 2006), potentially indicating a finite number of binding sites for

584   mitochondrial TAs. Furthermore, the TOM complex has been reported to assist in

585   insertion of full-length BAX into mitochondria (Ott *et al.* 2007; Cartron *et al.* 2008; Colin

586   *et al.* 2009). Consistent with the need for a mitochondrial TA translocation machinery,

587   the hFIS1 TA localizes specifically to mitochondria in human cells (Suzuki *et al.* 2003),

588   but cannot effectively localize mCherry to *S. cerevisiae* mitochondria, possibly

589   suggesting evolutionary divergence and a structural mismatch between the hFIS1 TA

590   and any putative yeast TA translocation apparatus. However, we note that not all

591   human mitochondrial TAs fail to be imported at the proper organelle in yeast, since our

592   genetic and microscopic results indicate that the human BAX TA can be targeted to

593   yeast mitochondria.

594

595   Other, perhaps more abundant evidence supports the idea that mitochondrial tail-

596   anchored proteins similar in structure to Fis1p do not require a translocation machinery

597   and can spontaneously insert into the OM. First, the MAD of the TA is protected from

598   chemical modification upon exposure to lipid vesicles devoid of protein (Kemper *et al.*

599   2008), suggesting that the Fis1p TA can insert into lipid bilayers without assistance.

600   Moreover, blockade or destruction of the general insertion pore for mitochondrial

601   proteins, associated receptors, or other outer membrane protein biogenesis machinery

602   such as the SAM complex or the MIM complex did not prevent Fis1p insertion at yeast

603   mitochondria (Stojanovski *et al.* 2007; Kemper *et al.* 2008; Sinzel *et al.* 2016). Tail-

604   anchored proteins also appear to have the ability to spontaneously and rapidly insert

605   into mammalian mitochondria without the need for the TOM complex or soluble

606   cytosolic chaperones (Setoguchi *et al.* 2006), although cytosolic chaperones potentially

607   play a role in maintaining solubility of tail-anchored proteins while they are *en route* to

608 mitochondria (Itakura *et al.* 2016). Further supporting the absence of a translocation

609 machinery dedicated to TA insertion, a large-scale screen for proteins required for

610 proper localization of mitochondrial tail-anchored proteins uncovered no putative

611 translocon components (Krumpe *et al.* 2012). We note that Fis1p is also localized to

612 peroxisomes (Kuravi *et al.* 2006), suggesting that any machinery that allows Fis1p TA

613 insertion would potentially be shared by both mitochondria and peroxisomes.

614 However, no dual-localized translocation machinery has yet been identified.

615

616 While a fraction of Fis1p can be targeted to the peroxisome, we found that detection of

617 mutant Fis1p TAs in the cytosol and nucleus by microscopic and genetic assays is not

618 driven by a specific failure of peroxisomal insertion. mCherry fused to the wild-type

619 Fis1p TA is not evident in the cytosol or nucleus upon severe disruption of peroxisomal

620 biogenesis (Hettema *et al.* 2000) by deletion of the peroxisomal membrane protein

621 import components Pex3p (Figure S11A) or Pex19p (C. Dunn, unpublished

622 observations). Moreover, reporter activation by Gal4-sfGFP-Fis1p containing an

623 unmutated TA is not increased when *PEX3* is removed (Figure S11B), and the relative

624 level of Gal4p-driven reporter activation among proline substitution mutants (Figure

625 S11C) or charge substitution mutants (Figure S11D) did not differ between *PEX3* and

626 *pex3Δ* cells.

627

628 If a TA insertion machinery does exist at the mitochondrial OM, loss-of-function

629 mutations affecting this machinery would presumably be recovered by the application

630 of our genetic selection scheme. Moreover, we expect that refined analysis of other

631 organelle targeting signals and membrane insertion sequences will be accomplished

632 by applying the deep mutational scanning approach outlined in this study.

633

MATERIALS AND METHODS:

Yeast strains and plasmids

Details of strains used in this study are provided in Table S2. Plasmid acquisition details and associated references, as well as details of plasmid construction, are found in Table S3. Oligonucleotides used in this study are listed in Table S4.

Culture conditions

Synthetic complete (SC) medium contains 0.67% yeast nitrogen base without amino acids, 2% dextrose, 0.1% casamino acids, 50 µg/ml adenine hemisulfate, and either 25 µg/ml uracil (SC-Trp) or 100 µg/ml L-tryptophan (SC-Ura). Supplemented minimal medium (SMM) contains 0.67% yeast nitrogen base without amino acids, 2% dextrose, 20 µg/ml adenine hemisulfate, 20 µg/ml uracil, 20 µg/ml methionine, 30 µg/ml lysine. SMM also contains, depending on selection needs, 20 µg/ml histidine, 100 µg/ml leucine, and/or 20 µg/ml tryptophan, as indicated. SLac medium lacking histidine contains 0.67% yeast nitrogen base without amino acids, 1.2% NaOH, a volume of lactic acid sufficient to subsequently bring the pH to 5.5, 20 µg/ml adenine hemisulfate, 20 µg/ml uracil, 20 µg/ml methionine, 30 µg/ml lysine, 100 µg/ml leucine, and 20 µg/ml tryptophan. Solid media also contain 1.7% bacteriological agar. Cells were incubated at 30℃ unless otherwise indicated. For serial dilution assays, strains in logarithmic proliferation phase were diluted to an OD_{600} of 0.1, and 4 µL of this dilution and three serial five-fold dilutions were spotted to solid medium. Experiments have been carried out at 30℃ unless otherwise noted.

Fis1p TA mutant library construction

Recombination-based cloning (Oldenburg *et al.* 1997) was used to generate constructs expressing Gal4-sfGFP-Fis1p under control of the *FIS1* promoter and mutated at one

24

664 of 27 positions within the Fis1p TA. Two DNA segments generated by PCR were fused

665 in this recombination reaction. The 5' portion was amplified by PCR from template

666 plasmid b100 using primer 698 and the appropriate primer (rvsposX) listed in Table S4.

667 The 3' section was generated from template b100 using primers 517 and the relevant

668 primer (fwdposX) listed in the same table. PCR products were recombined into *Not*I-

669 linearized pKS1 by co-transformation of vector and PCR products into strain MaV203.

670 The sub-library for each Fis1p TA position was generated individually by selection of

671 Trp$^+$ clones in liquid medium, with a portion of each transformation reaction plated to

672 solid SC-Trp medium to confirm recombination and transformation efficiency. To

673 generate the total pool prior to selection for Gal4p-mediated transcription, equal

674 numbers of cells, as determined by OD$_{600}$ measurement, were taken from overnight

675 cultures of each sub-library and combined within the same liquid culture.

676

677 *Deep mutational scanning of the Fis1p TA library*

678

679 The pool of constructs containing Fis1p TA mutations was cultured for four generations

680 in SC-Trp medium, SC-Ura medium, or SMM-Trp-His medium containing 0 mM, 5 mM,

681 10 mM, or 20 mM 3-AT. Plasmids present under each culture condition were then

682 recovered from 10 OD$_{600}$ units of cells. To harvest each plasmid library, cells were

683 pelleted at 4,000 g for 3 min, then washed with 5 ml 0.9 M D-sorbitol and resuspended

684 in 1 ml of 0.9 M D-sorbitol. One "stick-full" of zymolyase 20T (Amsbio, Abingdon,

685 United Kingdom), was added, and cells were incubated at 37°C for 45 min. Cells were

686 again collected at 4,000 g for 3 min and processed using a plasmid purification kit

687 (GeneJET Plasmid Miniprep Kit, Thermo Scientific, Waltham, USA) according to the

688 manufacturer's instructions. Primers 882 and 883 were used to amplify the genomic

689 region encoding the Fis1p TA from each plasmid pool. Using the provided PCR

690 products, next-generation, paired-end sequencing was performed by Microsynth

691 (Balgach, Switzerland) on a MiSeq Nano (2x150v2). The resulting FASTQ output can be

692 found at [to be uploaded to Dryad Digital Repository and link provided - upload

693 requires manuscript acceptance. Data are also available upon reviewer request].

694   FASTQ output from paired ends, stripped of adaptor sequences, was combined into a

695   single segment using the PANDAseq assembler version 2.8 (Masella *et al.* 2012). The

696   TRIM function (trimmer Galaxy tool version 0.0.1) was performed using the resources

697   of the Galaxy Project (Goecks *et al.* 2010) in order to remove sequences not directly

698   encoding the defined Fis1p and stop codon. Further processing in Microsoft Excel

699   (Redmond, USA) subsequently allowed conversion of DNA sequence to amino acid

700   sequence and removal of those TAs with more than one amino acid alteration from

701   further analysis. Enrichment values reflect, at a given amino acid position, the ratio of

702   the fraction of amino acid counts following selection to the fraction of amino acid

703   counts in the starting library. Enrichment values are not derived through comparisons

704   across different amino acid positions. Counts for the native amino acid at each position

705   were set as the total number of TA counts for which all amino acids were WT within a

706   given selected pool. When calculating enrichment values, TA amino acid replacements

707   for which there were zero reads in the SC-Trp sample had their value changed to one

708   in order to allow possible detection of enrichment under selective conditions by

709   preventing division by zero. Heat maps were generated using the Matrix2png utility

710   (Pavlidis and Noble 2003).

711

712   *Microscopy*

713

714   For epifluorescence microscopy, cells in the logarithmic phase of culture proliferation

715   were examined using an Eclipse 80i microscope with a 100X Plan Fluor objective and

716   linked to a DS-Qi1Mc camera (Nikon, Tokyo, Japan). Cells were cultured in SMM

717   medium appropriate for plasmid selection. Exposure times were automatically

718   determined, and images were captured using NIS-Elements version AR 3.2. mCherry

719   fusions are driven by the *ADH1* promoter and universally contain Fis1p amino acids

720   119-128 linking mCherry to the Fis1p TA, the region of Fis1p that is necessary and

721   sufficient for mitochondrial insertion (Mozdy *et al.* 2000; Beilharz 2003; Kemper *et al.*

722   2008; Förtsch *et al.* 2011) or to alternative TAs. All images of mCherry expression were

723   brightness adjusted in Adobe Photoshop CS5 (Adobe, San Jose, California) to an

26

724 equivalent extent, except when the mCherry-BAX(TA) signal was assessed. For

725 presentation of data associated with that specific fusion protein, the 'autolevels'

726 adjustment was used. Scoring of mitochondrial morphology was performed blind to

727 genotype. To promote Fis1p-dependent mitochondrial fragmentation, sodium azide

728 was added at a concentration of 500 µM for 60 min before fluorescence microscopy

729 (Fekkes *et al.* 2000; Klecker *et al.* 2015).

730

731 *Genetic assessment of Fis1p functionality*

732

733 Cells lose mtDNA and the ability to proliferate on non-fermentable medium when

734 mitochondrial fusion is blocked unless mitochondria division is also abrogated (Sesaki

735 and Jensen 1999; Fekkes *et al.* 2000; Mozdy *et al.* 2000; Tieu and Nunnari 2000).

736 Strain CDD688 (Mutlu *et al.* 2014) harbors chromosomal deletions in *FZO1* and *FIS1*

737 and a CHX-counterselectable plasmid expressing *FZO1*. Upon removal of plasmid-

738 expressed *FZO1*, cells will maintain mtDNA and respire unless functional *FIS1* is also

739 present to allow mitochondrial division. To assess the functionality of Fis1p variants

740 containing TA mutations, strain CDD688 was transformed with plasmids expressing

741 WT *FIS1* or variants mutated within the Fis1p TA. Transformants were cultured

742 overnight in SMM-His medium lacking CHX to permit cells to lose the *FZO1*-encoding

743 plasmid. Serial dilutions were then spotted to SLac-His + 3 µg/mL CHX ("lactate / no

744 fusion") and incubated for 5 d to test for maintenance of mtDNA following

745 counterselection for *FZO1*, with cell proliferation indicating a lack of Fis1p function. As

746 a control for cell proliferation under conditions not selective for mtDNA maintenance,

747 an equal number of cells was also spotted to SMM-Trp-His medium ("glucose /

748 fusion") and incubated for 2 d.

749

27

750 **ACKNOWLEDGEMENTS:**

751

756

757 **CONFLICT OF INTEREST STATEMENT:**

758

759 The authors have no known conflict of interest affecting the outcome or interpretation

760 of this study.

761

FIGURE LEGENDS:

Figure 1. A genetic selection based on protein mislocalization allows recovery of mutations blocking Fis1p TA localization to mitochondria. (A) A scheme for selection of mutations preventing mitochondrial targeting of the Fis1p TA. Full-length Fis1p is fused to the transcription factor Gal4p. Upon failure of Fis1p to be localized to the mitochondrial OM, Gal4p may be free to translocate to the nucleus and activate the selectable markers *HIS3* and *URA3*. (B) A Gal4-sfGFP-Fis1 fusion localizes to mitochondria. Strain CDD898 was transformed with plasmid b102, which overexpresses the Gal4-sfGFP-Fis1p construct used in this study. Mitochondria were visualized using mCherry fused to the Cox4 presequence expressed from plasmid pHS12-mCherry. Scale bar, 5 µm. (C) Removal of the Fis1p TA allows proliferation on medium requiring *HIS3* activation or *URA3* activation. Strain MaV203 expressing Gal4-sfGFP-Fis1p from plasmid b100 (WT TA), a variant lacking the Fis1p TA from plasmid b101 (ΔTA), or harboring empty vector pKS1 was cultured in SC-Trp medium, then, following serial dilution, spotted to SC-Trp, SMM-His + 20 mM 3-AT, or SC-Ura and incubated for 2 d.

Figure 2. Selection for reporter activation by Gal4-sfGFP-Fis1p reveals TA mutations inhibiting mitochondrial localization. (A) Missense mutations within the Fis1p TA provide selectable marker activation. Strain MaV203 expressing Gal4-sfGFP-Fis1p from plasmid b100 or variants expressed from plasmids b128 (V145E), b129 (L139P), b130 (L129P, V138A), or b101 (ΔTA) were treated as in Figure 1C. pKS1 (vector) is also provided as a negative control. (B) Missense mutations within the Fis1p TA allow cytosolic and nuclear accumulation of a linked mCherry protein. mCherry fused to variants of the Fis1p TA were expressed in WT strain CDD961 from plasmids b109 (WT), b134 (V145E), b135 (L139P), b136 (L129P,V138A), or b252 (ΔTA) and visualized by fluorescence microscopy. Mitochondria were labelled with a mitochondria-targeted GFP expressed from plasmid pHS1. Scale bar, 5 µm.

29

792 **Figure 3.** Global discovery of mutations within the TA of a Gal4-sfGFP-Fis1 fusion

793 protein that allow Gal4-driven transcription. The $\log_2$ of enrichment values for each

794 amino acid were calculated for each position following selection in SMM-Trp-His

795 medium containing 20 mM 3-AT. Enrichment values are generated for individual amino

796 acid positions within the TA, and not across positions. Black outlines denote the native

797 amino acid for each position. Amino acid replacements not detectable under selective

798 conditions are denoted by black, filled squares. The predicted MAD is indicated by a

799 red line. 'X' represents substitution by a stop codon.

800

801 **Figure 4.** Identification of abundant Gal4-sfGFP-Fis1p clones which are highly

802 enriched upon selection for Gal4-sfGFP-Fis1p nuclear translocation. (A) TA substitution

803 mutations are plotted, with $\log_2$ enrichment values provided on the X-axis and

804 sequence counts recovered from the starting pool (SC-Trp) provided on the Y-axis.

805 Those replacement mutations that are within the top 75th percentile of mutant

806 abundance in the starting pool and enriched at least four-fold following selection in

807 SMM-Trp-His medium containing 20 mM 3-AT are highlighted in a blue box. (B)

808 Expansion of the highlighted region in (A) showing specific TA mutations.

809

810 **Figure 5.** Proline substitution is acceptable at a discrete position within the Fis1p TA.

811 (A) Replacement of specific amino acids within the TA of Gal4-sfGFP-Fis1p with proline

812 can lead to Gal4-mediated selectable marker activation. Strain MaV203 expressing

813 Gal4-sfGFP-Fis1p variants from plasmids b100 (WT), b188 (V134P), b189 (G137P),

814 b129 (L139P), b190 (A140P), b296 (A144P), or b101 (ΔTA) was cultured in SC-Trp

815 medium then spotted to SC-Trp or SMM-His + 20 mM 3-AT medium for 2 d. (B) TAs

816 with specific proline replacements can reduce mitochondrial targeting of a linked

817 fluorescent protein. Variants of the Fis1p TA fused to mCherry were expressed in WT

818 strain CDD961 from plasmids b109 (WT), b208 (V134P), b209 (G137P), b135 (L139P),

819 b210 (A140P), b211 (A144P) and examined, along with mitochondria-targeted GFP, as

820 in Figure 2B. Scale bar, 5 μm.

821

822

823 **Figure 6.** The positively charged carboxyl-terminus of the Fis1p TA is important for

824 specific localization to and insertion at the mitochondrial outer membrane. (A) Deletion

825 of the final five amino acids from the Fis1p TA permits transcriptional activation by

826 Gal4-sfGFP-Fis1p. Strain MaV203 harboring plasmids b100 (WT), b253 (R151X), or

827 b101 (ΔTA) was treated as in Figure 5A. (B) Removal of the last five amino acids from

828 the Fis1p TA allows mislocalization to the secretory system. Strain CDD961 expressing

829 mCherry fused to the WT Fis1p TA from plasmid b109 or expressing mCherry linked to

830 a truncated Fis1p TA (R151X) from plasmid b254 was evaluated as in Figure 2B. (C)

831 Strain CDD961 was cured of plasmid pHS1. The resulting strain was transformed with

832 plasmid pJK59 to label ER by expression of Sec63p-GFP, then transformed with either

833 plasmid b109 or plasmid b254 to localize the WT and R151X TAs and examined by

834 fluorescence microscopy. Scale bar, 5 μm.

835

836 **Figure 7**. Targeting of the Fis1p TA is not dependent upon a specific TA length. (A)

837 Deletion of up to three amino acids or insertion of up to three amino acids does not

838 allow Gal4-sfGFP-Fis1p to activate transcription. MaV203 cells expressing Gal4-

839 sfGFP-Fis1p variants from plasmids b229 (∇1A), b230 (∇2A), b231 (∇3A), b226 (ΔG136),

840 b227 (ΔA135-G136), b228 (ΔA135-G137), or b101 (ΔTA) were treated as in Figure 5A.

841 (B) mCherry fused to a Fis1p TA containing an insertion of up to three amino acids in

842 length localizes properly to mitochondria. Strain CDD961 expressing mCherry-TA

843 fusions from plasmids b109 (WT), b235 (∇1A), b236 (∇2A), or b237 (∇3A) was visualized

844 as in Figure 2B. (C) mCherry fused to a Fis1p TA deleted of up to three amino acids is

845 properly targeted to mitochondria. Strain CDD961 expressing mCherry-TA fusions from

846 plasmids b109 (WT), b232 (ΔG136), b233 (ΔA135-G136), or b234 (ΔA135-G137) was

847 examined as in Figure 2B. Scale bar, 5 μm.

848

849 **Figure 8.** Fis1p TA targeting is hindered to a greater extent by inclusion of negatively

850 charged amino acids in the MAD than by positively charged amino acids. (A) Negative

851 charges allow higher transcriptional activity than positive charges when placed at

31

852   specific positions within the Gal4-sfGFP-Fis1p TA. Strain MaV203 was transformed

853   with plasmids pKS1 (vector), b100 (Gal4-sfGFP-Fis1), b101 [Gal4-sfGFP-Fis1(ΔTA)], or

854   plasmids encoding the indicated charge replacements within the Fis1p TA (plasmids

855   b173-b187 and b295). The resulting transformants were spotted to SC-Trp medium (1

856   d, 30℃) or SMM-His + 20mM 3-AT medium (2 d, 30℃). (B) mCherry-TA localization is

857   disrupted more severely by negatively charged amino acids within the MAD than by

858   positively charged amino acids. Strain CDD961 was transformed with plasmids (b192-

859   b207) expressing mCherry linked to Fis1p TAs harboring the indicated substitutions.

860   Cells were visualized as in Figure 2B. Scale bar, 5 μm.

REFERENCES:

Akdoğan E., Tardu M., Garipler G., Baytek G., Kavakli İ. H., Dunn C. D., 2016 Reduced Glucose Sensation Can Increase the Fitness of Saccharomyces cerevisiae Lacking Mitochondrial DNA (M Polymenis, Ed.). PLoS ONE **11**: e0146511–32.

Araya C. L., Fowler D. M., 2011 Deep mutational scanning: assessingprotein function on a massive scale. Trends in Biotechnology **29**: 435–442.

Beilharz T., 2003 Bipartite Signals Mediate Subcellular Targeting of Tail-anchored Membrane Proteins in Saccharomyces cerevisiae. **278**: 8219–8223.

Borgese N., Fasana E., 2011 Targeting pathways of C-tail-anchored proteins. BBA - Biomembranes **1808**: 937–946.

Borgese N., Brambillasca S., Colombo S., 2007 How tails guide tail-anchored proteins to their destinations. Current opinion in cell biology **19**: 368–375.

Borgese N., Gazzoni I., Barberi M., Colombo S., Pedrazzini E., 2001 Targeting of a tail-anchored protein to endoplasmic reticulum and mitochondrial outer membrane by independent but competing pathways. Mol Biol Cell **12**: 2482–2496.

Boucher J. I., Cote P., Flynn J., Jiang L., Laban A., Mishra P., Roscoe B. P., Bolon D. N. A., 2014 Viewing protein fitness landscapes through a next-gen lens. Genetics **198**: 461–471.

Buchan D. W. A., Minneci F., Nugent T. C. O., Bryson K., Jones D. T., 2013 Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Res **41**: W349–57.

Cartron P.-F., Bellot G., Oliver L., Grandier-Vazeille X., Manon S., Vallette F. M., 2008 Bax inserts into the mitochondrial outer membrane by different mechanisms. FEBS Lett **582**: 3045–3051.

Chen Y.-C., Umanah G. K. E., Dephoure N., Andrabi S. A., Gygi S. P., Dawson T. M., Dawson V. L., Rutter J., 2014 Msp1/ATAD1 maintains mitochondrial function by facilitating the degradation of mislocalized tail-anchored proteins. EMBO J **33**: 1548–1564.

Chou P. Y., Fasman G. D., 1974 Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins. Biochemistry **13**: 211–222.

Chuang K.-H., Liang F., Higgins R., Wang Y., 2016 Ubiquilin/Dsk2 promotes inclusion body formation and vacuole (lysosome)-mediated disposal of mutated huntingtin. Mol Biol Cell **27**: 2025–2036.

895 Colin J., Garibal J., Mignotte B., Guenal I., 2009 The mitochondrial TOM complex
896   modulates bax-induced apoptosis in Drosophila. Biochemical and Biophysical
897   Research Communications **379**: 939–943.

898 Cymer F., Heijne von G., White S. H., 2015 Mechanisms of Integral Membrane Protein
899   Insertion and Folding. J Mol Biol **427**: 999–1022.

900 de Mendoza D., Cronan J. E. Jr., 1983 Thermal regulation of membrane lipid fluidity in
901   bacteria. Trends in Biochemical Sciences **8**: 49–52.

902 Denic V., Dötsch V., Sinning I., 2013 Endoplasmic reticulum targeting and insertion of
903   tail-anchored membrane proteins by the GET pathway. Cold Spring Harbor
904   Perspectives in Biology **5**: a013334–a013334.

905 Deshaies R. J., Schekman R., 1987 A yeast mutant defective at an early stage in import
906   of secretory protein precursors into the endoplasmic reticulum. J Cell Biol **105**:
907   633–645.

908 Dong H., Sharma M., Zhou H.-X., Cross T. A., 2012 Glycines: Role in α-Helical
909   Membrane Protein Structures and a Potential Indicator of Native Conformation.
910   Biochemistry **51**: 4779–4789.

911 Drozdetskiy A., Cole C., Procter J., Barton G. J., 2015 JPred4: a protein secondary
912   structure prediction server. Nucleic Acids Res **43**: W389–94.

913 Dufourc E. J., 2008 Sterols and membrane dynamics. J Chem Biol **1**: 63–77.

914 Dunn C. D., Jensen R. E., 2003 Suppression of a defect in mitochondrial protein import
915   identifies cytosolic proteins required for viability of yeast cells lacking mitochondrial
916   DNA. Genetics **165**: 35–45.

917 Durfee T., Becherer K., Chen P. L., Yeh S. H., Yang Y., Kilburn A. E., Lee W. H., Elledge
918   S. J., 1993 The retinoblastoma protein associates with the protein phosphatase
919   type 1 catalytic subunit. Genes Dev **7**: 555–569.

920 Elazar A., Weinstein J., Biran I., Fridman Y., Bibi E., Fleishman S. J., 2016 Mutational
921   scanning reveals the determinants of protein insertion and association energetics in
922   the plasma membrane. eLife **5**: 1302.

923 Fekkes P., Shepard K. A., Yaffe M. P., 2000 Gag3p, an outer membrane protein
924   required for fission of mitochondrial tubules. J Cell Biol **151**: 333–340.

925 Fowler D. M., Fields S., 2014 Deep mutational scanning: a new style of protein science.
926   Nature Publishing Group **11**: 801–807.

927 Förtsch J., Hummel E., Krist M., Westermann B., 2011 The myosin-related motor
928   protein Myo2 is an essential mediator of bud-directed mitochondrial movement in

929 yeast. J Cell Biol **194**: 473–488.

930 Garipler G., Mutlu N., Lack N. A., Dunn C. D., 2014 Deletion of conserved protein
931 phosphatases reverses defects associated with mitochondrial DNA damage in
932 Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences **111**:
933 1473–1478.

934 Gavel Y., Nilsson L., Heijne von G., 1988 Mitochondrial targeting sequences. Why
935 "non-amphiphilic" peptides may still be amphiphilic. FEBS Lett **235**: 173–177.

936 Gimpelev M., Forrest L. R., Murray D., Honig B., 2004 Helical Packing Patterns in
937 Membrane and Soluble Proteins. Biophysical Journal **87**: 4075–4086.

938 Goecks J., Nekrutenko A., Taylor J., Galaxy Team, 2010 Galaxy: a comprehensive
939 approach for supporting accessible, reproducible, and transparent computational
940 research in the life sciences. Genome Biol **11**: R86.

941 Gross A., Pilcher K., Blachly-Dyson E., Basso E., Jockel J., Bassik M. C., Korsmeyer S.
942 J., Forte M., 2000 Biochemical and Genetic Analysis of the Mitochondrial Response
943 of Yeast to BAX and BCL-XL. Mol Cell Biol **20**: 3125–3136.

944 Habib S. J., Vasiljev A., Neupert W., Rapaport D., 2003 Multiple functions of tail-anchor
945 domains of mitochondrial outer membrane proteins. FEBS Lett **555**: 511–515.

946 He S., Fox T. D., 1999 Mutations affecting a yeast mitochondrial inner membrane
947 protein, pnt1p, block export of a mitochondrially synthesized fusion protein from
948 the matrix. Mol Cell Biol **19**: 6598–6607.

949 Hettema E. H., Girzalsky W., Van den Berg M., Erdmann R., Distel B., 2000
950 Saccharomyces cerevisiae pex3p and pex19p are required for proper localization
951 and stability of peroxisomal membrane proteins. EMBO J **19**: 223–233.

952 Hietpas R. T., Jensen J. D., Bolon D. N. A., 2011 Experimental illumination of a fitness
953 landscape. Proceedings of the National Academy of Sciences **108**: 7896–7901.

954 Horie C., Suzuki H., Sakaguchi M., Mihara K., 2002 Characterization of signal that
955 directs C-tail-anchored proteins to mammalian mitochondrial outer membrane. Mol
956 Biol Cell **13**: 1615–1625.

957 Isenmann S., Khew-Goodall Y., Gamble J., Vadas M., Wattenberg B. W., 1998 A
958 splice-isoform of vesicle-associated membrane protein-1 (VAMP-1) contains a
959 mitochondrial targeting signal. Mol Biol Cell **9**: 1649–1660.

960 Itakura E., Zavodszky E., Shao S., Wohlever M. L., Keenan R. J., Hegde R. S., 2016
961 Ubiquilins Chaperone and Triage Mitochondrial Membrane Proteins for
962 Degradation. Molecular Cell: 1–14.

963 Jensen R. E., Schmidt S., Mark R. J., 1992 Mutations in a 19-amino-acid hydrophobic
964   region of the yeast cytochrome c1 presequence prevent sorting to the
965   mitochondrial intermembrane space. Mol Cell Biol **12**: 4677–4686.

966 Johnson N., Powis K., High S., 2013 Post-translational translocation into the
967   endoplasmic reticulum. BBA-Molecular Cell Research **1833**: 2403–2409.

968 Kemper C., Habib S. J., Engl G., Heckmeyer P., Dimmer K. S., Rapaport D., 2008
969   Integration of tail-anchored proteins into the mitochondrial outer membrane does
970   not require any known import components. Journal of Cell Science **121**: 1990–
971   1998.

972 Kim C., Schmidt T., Cho E.-G., Ye F., Ulmer T. S., Ginsberg M. H., 2012 Basic amino-
973   acid side chains regulate transmembrane integrin signalling. Nature **481**: 209–213.

974 Klecker T., Wemmer M., Haag M., Weig A., Böckler S., Langer T., Nunnari J.,
975   Westermann B., 2015 Interaction of MDM33 with mitochondrial inner membrane
976   homeostasis pathways in yeast. Sci. Rep.: 1–14.

977 Krogh A., Larsson B., Heijne von G., Sonnhammer E. L. L., 2001 Predicting
978   transmembrane protein topology with a hidden markov model: application to
979   complete genomes11Edited by F. Cohen. J Mol Biol **305**: 567–580.

980 Krumpe K., Frumkin I., Herzig Y., Rimon N., Özbalci C., Brügger B., Rapaport D.,
981   Schuldiner M., 2012 Ergosterol content specifies targeting of tail-anchored proteins
982   to mitochondrial outer membranes. Mol Biol Cell **23**: 3927–3935.

983 Kuravi K., Nagotu S., Krikken A. M., Sjollema K., Deckers M., Erdmann R., Veenhuis
984   M., van der Klei I. J., 2006 Dynamin-related proteins Vps1p and Dnm1p control
985   peroxisome abundance in Saccharomyces cerevisiae. Journal of Cell Science **119**:
986   3994–4001.

987 Kuroda R., Ikenoue T., Honsho M., Tsujimoto S., Mitoma J. Y., Ito A., 1998 Charged
988   amino acids at the carboxyl-terminal portions determine the intracellular locations
989   of two isoforms of cytochrome b5. J Biol Chem **273**: 31097–31102.

990 Lee J., Kim D. H., Hwang I., 2014 Specific targeting of proteins to outer envelope
991   membranes of endosymbiotic organelles, chloroplasts, and mitochondria. Front
992   Plant Sci **5**: 173.

993 Lee S., Lim W. A., Thorn K. S., 2013 Improved blue, green, and red fluorescent protein
994   tagging vectors for S. cerevisiae. (AS Gladfelter, Ed.). PLoS ONE **8**: e67902.

995 Li S. C., Goto N. K., Williams K. A., Deber C. M., 1996 Alpha-helical, but not beta-
996   sheet, propensity of proline is determined by peptide environment. Proc Natl Acad
997   Sci USA **93**: 6676–6681.

998   Long S. B., Campbell E. B., MacKinnon R., 2005 Voltage Sensor of Kv1.2: Structural
999   Basis of Electromechanical Coupling. Science **309**: 903–908.

1000 Maarse A. C., Blom J., Grivell L. A., Meijer M., 1992 MPI1, an essential gene encoding
1001   a mitochondrial membrane protein, is possibly involved in protein import into yeast
1002   mitochondria. EMBO J **11**: 3619–3628.

1003 Masella A. P., Bartram A. K., Truszkowski J. M., Brown D. G., Neufeld J. D., 2012
1004   PANDAseq: paired-end assembler for illumina sequences. BMC Bioinformatics **13**:
1005   31.

1006 Melamed D., Young D. L., Gamble C. E., Miller C. R., Fields S., 2013 Deep mutational
1007   scanning of an RRM domain of the Saccharomyces cerevisiae poly(A)-binding
1008   protein. RNA **19**: 1537–1551.

1009 Monné M., Nilsson I., Johansson M., Elmhed N., Heijne von G., 1998 Positively and
1010   negatively charged residues have different effects on the position in the membrane
1011   of a model transmembrane helix. J Mol Biol **284**: 1177–1183.

1012 Mozdy A. D., McCaffery J. M., Shaw J. M., 2000 Dnm1p GTPase-mediated
1013   mitochondrial fission is a multi-step process requiring the novel integral membrane
1014   component Fis1p. J Cell Biol **151**: 367–380.

1015 Mutlu N., Garipler G., Akdoğan E., Dunn C. D., 2014 Activation of the pleiotropic drug
1016   resistance pathway can promote mitochondrial DNA retention by fusion-defective
1017   mitochondria in Saccharomyces cerevisiae. G3 (Bethesda) **4**: 1247–1258.

1018 Nechushtan A., Smith C. L., Hsu Y. T., Youle R. J., 1999 Conformation of the Bax C-
1019   terminus regulates subcellular location and cell death. EMBO J **18**: 2330–2341.

1020 Neupert W., 2015 A Perspective on Transport of Proteins into Mitochondria: A Myriad
1021   of Open Questions. J Mol Biol **427**: 1135–1158.

1022 O'Neil K. T., DeGrado W. F., 1990 A thermodynamic scale for the helix-forming
1023   tendencies of the commonly occurring amino acids. Science **250**: 646–651.

1024 Okreglak V., Walter P., 2014 The conserved AAA-ATPase Msp1 confers organelle
1025   specificity to tail-anchored proteins. Proceedings of the National Academy of
1026   Sciences **111**: 8019–8024.

1027 Oldenburg K. R., Vo K. T., Michaelis S., Paddon C., 1997 Recombination-mediated
1028   PCR-directed plasmid construction in vivo in yeast. Nucleic Acids Res **25**: 451–
1029   452.

1030 Ospina A., Lagunas-Martínez A., Pardo J., Carrodeguas J. A., 2011 Protein
1031   oligomerization mediated by the transmembrane carboxyl terminal domain of Bcl-

37

1032   XL. FEBS Lett **585**: 2935–2942.

1033   Ott M., Norberg E., Walter K. M., Schreiner P., Kemper C., Rapaport D., Zhivotovsky
1034   B., Orrenius S., 2007 The mitochondrial TOM complex is required for tBid/Bax-
1035   induced cytochrome c release. J Biol Chem **282**: 27633–27639.

1036   Öjemalm K., Higuchi T., Lara P., Lindahl E., Suga H., Heijne von G., 2016 Energetics of
1037   side-chain snorkeling in transmembrane helices probed by nonproteinogenic amino
1038   acids. Proceedings of the National Academy of Sciences **113**: 10559–10564.

1039   Pavlidis P., Noble W. S., 2003 Matrix2png: a utility for visualizing matrix data.
1040   Bioinformatics **19**: 295–296.

1041   Pédelacq J.-D., Cabantous S., Tran T., Terwilliger T. C., Waldo G. S., 2005 Engineering
1042   and characterization of a superfolder green fluorescent protein. Nat Biotechnol **24**:
1043   79–88.

1044   Prinz W. A., Grzyb L., Veenhuis M., Kahana J. A., Silver P. A., Rapoport T. A., 2000
1045   Mutants affecting the structure of the cortical endoplasmic reticulum in
1046   Saccharomyces cerevisiae. J Cell Biol **150**: 461–474.

1047   Rapaport D., 2003 Finding the right organelle. Targeting signals in mitochondrial outer-
1048   membrane proteins. EMBO Rep **4**: 948–952.

1049   Robinson J. S., Klionsky D. J., Banta L. M., Emr S. D., 1988 Protein sorting in
1050   Saccharomyces cerevisiae: isolation of mutants defective in the delivery and
1051   processing of multiple vacuolar hydrolases. Mol Cell Biol **8**: 4936–4948.

1052   Russ W. P., Engelman D. M., 2000 The GxxxG motif: A framework for transmembrane
1053   helix-helix association. J Mol Biol **296**: 911–919.

1054   Ryan K. R., Leung R. S., Jensen R. E., 1998 Characterization of the mitochondrial inner
1055   membrane translocase complex: the Tim23p hydrophobic domain interacts with
1056   Tim17p but not with other Tim23p molecules. Mol Cell Biol **18**: 178–187.

1057   Schinzel A., Kaufmann T., Schuler M., Martinalbo J., Grubb D., Borner C., 2004
1058   Conformational control of Bax localization and apoptotic activity by Pro168. J Cell
1059   Biol **164**: 1021–1032.

1060   Schuldiner M., Metz J., Schmid V., Denic V., Rakwalska M., Schmitt H. D.,
1061   Schwappach B., Weissman J. S., 2008 The GET Complex Mediates Insertion of
1062   Tail-Anchored Proteins into the ER Membrane. Cell **134**: 634–645.

1063   Segrest J. P., De Loof H., Dohlman J. G., Brouillette C. G., Anantharamaiah G. M.,
1064   1990 Amphipathic Helix Motif: Classes and Properties. PROTEINS: Structure,
1065   Function, and Genetics **8**: 103–117.

Senes A., Engel D. E., DeGrado W. F., 2004 Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. Current Opinion in Structural Biology **14**: 465–479.

Sesaki H., Jensen R. E., 1999 Division versus fusion: Dnm1p and Fzo1p antagonistically regulate mitochondrial shape. J Cell Biol **147**: 699–706.

Setoguchi K., Otera H., Mihara K., 2006 Cytosolic factor- and TOM-independent import of C-tail-anchored mitochondrial outer membrane proteins. EMBO J **25**: 5635–5647.

Sikorski R. S., Hieter P., 1989 A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics **122**: 19–27.

Sinzel M., Tan T., Wendling P., Kalbacher H., Özbalci C., Chelius X., Westermann B., Brügger B., Rapaport D., Dimmer K. S., 2016 Mcp3 is a novel mitochondrial outer membrane protein that follows a unique IMP-dependent biogenesis pathway. EMBO Rep **17**: 965–981.

Soding J., 2005 Protein homology detection by HMM-HMM comparison. Bioinformatics **21**: 951–960.

Starita L. M., Young D. L., Islam M., Kitzman J. O., Gullingsrud J., Hause R. J., Fowler D. M., Parvin J. D., Shendure J., Fields S., 2015 Massively Parallel Functional Analysis of BRCA1 RING Domain Variants. Genetics **200**: 413–422.

Stirling C. J., Rothblatt J., Hosobuchi M., Deshaies R., Schekman R., 1992 Protein translocation mutants defective in the insertion of integral membrane proteins into the endoplasmic reticulum. Mol Biol Cell **3**: 129–142.

Stojanovski D., Guiard B., Kozjak-Pavlovic V., Pfanner N., Meisinger C., 2007 Alternative function for the mitochondrial SAM complex in biogenesis of alpha-helical TOM proteins. J Cell Biol **179**: 881–893.

Stojanovski D., Koutsopoulos O. S., Okamoto K., Ryan M. T., 2004 Levels of human Fis1 at the mitochondrial outer membrane regulate mitochondrial morphology. Journal of Cell Science **117**: 1201–1210.

Strandberg E., Killian J. A., 2003 Snorkeling of lysine side chains in transmembrane helices: how easy can it get? FEBS Lett **544**: 69–73.

Suzuki M., Jeong S. Y., Karbowski M., Youle R. J., Tjandra N., 2003 The solution structure of human mitochondria fission protein Fis1 reveals a novel TPR-like helix bundle. J Mol Biol **334**: 445–458.

1100   Taxis C., Knop M., 2006 *System of centromeric, episomal, and integrative vectors*
1101   *based on drug resistance markers for Saccharomyces cerevisiae.*

1102   Tieu Q., Nunnari J., 2000 Mdv1p is a WD repeat protein that interacts with the
1103   dynamin-related GTPase, Dnm1p, to trigger mitochondrial division. J Cell Biol **151**:
1104   353–366.

1105   Vidal M., Brachmann R. K., Fattaey A., Harlow E., Boeke J. D., 1996a Reverse two-
1106   hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-
1107   protein interactions. Proc Natl Acad Sci USA **93**: 10315–10320.

1108   Vidal M., Braun P., Chen E., Boeke J. D., Harlow E., 1996b Genetic characterization of
1109   a mammalian protein-protein interaction domain by using a yeast reverse two-
1110   hybrid system. Proc Natl Acad Sci USA **93**: 10321–10326.

1111   Wattenberg B., Lithgow T., 2001 Targeting of C-terminal (tail)-anchored proteins:
1112   understanding how cytoplasmic activities are anchored to intracellular membranes.
1113   Traffic **2**: 66–71.

1114   Wattenberg B. W., Clark D., Brock S., 2007 An artificial mitochondrial tail signal/anchor
1115   sequence confirms a requirement for moderate hydrophobicity for targeting. Biosci
1116   Rep **27**: 385–401.

1117   Wells R. C., Hill R. B., 2011 The cytosolic domain of Fis1 binds and reversibly clusters
1118   lipid vesicles. (P Kursula, Ed.). PLoS ONE **6**: e21384.

1119   Yu C.-H., Dang Y., Zhou Z., Wu C., Zhao F., Sachs M. S., Liu Y., 2015 Codon Usage
1120   Influences the Local Rate of Translation Elongation to Regulate Co-translational
1121   Protein Folding. Molecular Cell **59**: 744–754.

1122   Zhou Z., Dang Y., Zhou M., Li L., Yu C.-H., Fu J., Chen S., Liu Y., 2016 Codon usage is
1123   an important determinant of gene expression levels largely through its effects on
1124   transcription. Proceedings of the National Academy of Sciences **113**: E6117–
1125   E6125.

1126

1127

SUPPLEMENTAL FIGURE AND TABLE LEGENDS:

Figure S1. Structural and sequence characteristics of the Fis1p TA. (A) The Fis1p TA is predicted to be a transmembrane helix. The indicated amino acid sequence was tested using the TMHMM Server version 2.0 (Krogh *et al.* 2001), and individual amino acid scores were plotted. The predicted MAD is indicated. (B) Glycines are abundant among Fis1p orthologs. Fis1 proteins from the listed species (Uniprot access numbers listed) were aligned using Clustal Omega (Soding 2005). Glycines are colored green, positively charged amino acids are colored purple, negatively charged amino acids are colored blue, and proteins are colored red. A red bar over the alignment indicates the predicted MAD, and a black arrow denotes a potentially conserved glycine. The region necessary and sufficient for Fis1p insertion is bracketed in pink.

Figure S2. Amino acid replacement representation within the Gal4-sfGFP-Fis1p TA library. The fraction of counts representing each amino acid replacement in the starting SC-Trp library (fObs) was compared to the fraction that would be expected based on randomized codon recovery (fExp). Native amino acids are represented by a black square with a blue dot. Amino acid replacements with no representation in the library are represented by empty black squares. The predicted MAD is indicated by a red line.

Figure S3. Amino acid replacement frequencies within the Fis1p TA under differing selective pressure. Quantification of replacement in SMM-Trp-His medium without 3-AT (A), containing 5 mM 3-AT (B), containing 10mM 3-AT (C), or in SC-Ura medium (D). In all panels, black outlines indicate the native amino acid at each position within the Fis1p TA. Amino acid replacements not detectable under selective conditions are denoted by black, filled squares. The predicted MAD is indicated by a red line.

Figure S4. Little evidence exists for codon-level control of Gal4-sfGFP-Fis1p TA targeting. The $\log_2$ of enrichment values for each codon following selection of the Fis1p TA library in SMM-Trp-His medium containing 20 mM 3-AT are illustrated. Green

41

1158   outlines denote the native amino acid at each position. Codon replacements with no

1159   representation in the library following selection are represented by empty black

1160   squares. The predicted MAD is indicated by a red line.

1161

1162   **Figure S5.** Deletion of Msp1p or Dsk2p does not allow recovery of mitochondrial

1163   localization by poorly targeted mCherry-TA variants. (A) Deletion of the Msp1p

1164   extractase does not allow tail-anchored proteins mistargeted due to proline

1165   substitutions to return to mitochondria. Cells from *msp1Δ/msp1Δ* strain CDD1044

1166   expressing pHS1 (signal not shown) and plasmids b109 (WT), b208 (V134P), b135

1167   (L139P), b210 (A140P), or b211 (A144P) were examined to determine mCherry-TA

1168   location. (B) Deletion of ubiquilin ortholog Dsk2p does not allow mislocalized, proline-

1169   containing Fis1p TAs to target to mitochondria. Cells from *dsk2Δ/dsk2Δ* strain

1170   CDD1179 were transformed and analyzed as in (A). (C) Deletion of Msp1p does not

1171   allow mitochondrial localization of mistargeted Fis1p TAs carrying charge

1172   replacements. Strain CDD1044, deleted of Msp1p, was transformed with the following

1173   plasmids encoding mCherry fused to the mutant TAs: b192 (V132D), b196 (A140D),

1174   b197 (A140E), b198 (A140K), b199 (A140R), b200 (A144D), b201 (A144E), b134

1175   (V145E), b204 (F148D), b205 (F148E). Transformants were examined as in (A). (D)

1176   Removal of Dsk2p fails to allow mutant TAs containing charge substitutions to re-

1177   localized to mitochondria. Strain CDD1179 was transformed with plasmids and

1178   analyzed as in (C). Scale bar, 5 μm.

1179

1180   **Figure S6.** A mutated Fis1p TA lacking the positively charged carboxyl-terminus

1181   localizes to the ER independently of Get3p or Get2p. (A) Get3p is not required for ER

1182   localization of the R151X TA. *get3Δ/get3Δ* strain CDD1033 expressing mCherry fused

1183   to a WT Fis1p TA from plasmid b109 or expressing mCherry fused to a Fis1p TA

1184   lacking the positively charged carboxyl-terminus (R151X) from plasmid b254 were

1185   examined, and mitochondria were labelled with GFP expressed from pHS1. (B)

1186   *get3Δ/get3Δ* strain CDD1033 carrying plasmids b109 or b254 were examined by

1187   expression of Sec63p-GFP from plasmid pJK59 as in Figure 6C. (C) Get2p is not

1188   necessary for ER localization of the Fis1p TA lacking carboxyl-terminal, positively

1189   charged amino acids. *get2Δ* strain CDD948 was transformed with plasmid pJK59 and

1190   either b109 or b254, then imaged as in Figure 6C. White arrows indicate puncta

1191   containing truncated mCherry-Fis1(TA). Scale bar, 5 μm.

1192

1193   **Figure S7.** The TAs of human proteins do not universally target to mitochondria in

1194   yeast. (A) The hFIS1 TA is mistargeted to the ER in *S. cerevisiae*. Strain CDD961

1195   containing plasmid b257 expressing mCherry fused to the hFIS1 TA was analyzed as in

1196   Figure 2B. (B) Strain CDD961 cured of pHS1 and carrying plasmids pJK59 and b257

1197   was imaged as in Figure 6C. (C) The hFIS1TA does not permit activity of a fused Gal4p

1198   within the nucleus. Strain MaV203 was transformed with plasmid b100 (Gal4-sfGFP-

1199   Fis1), b258 [Gal4-hFIS1(TA)], or plasmid b101 (Gal4-sfGFP-Fis1ΔTA) and assessed as

1200   in Figure 5A. (D) The BAX TA can target to mitochondria in *S. cerevisiae*. Strain

1201   CDD961 transformed with plasmid b255, which expresses mCherry fused to the BAX

1202   TA, was examined as in Figure 2B. Scale bar, 5 μm.

1203

1204   **Figure S8.** mCherry fused to Fis1p TAs of varying length remains targeted to the ER in

1205   cells lacking Spf1p. (A) Cells of *spf1Δ/spf1Δ* strain CDD1031 expressing mCherry-TA

1206   fusions from plasmids b109 (WT), b232 (ΔG136), b233 (ΔA135-G136), b234 (ΔA135-

1207   G137), b235 (∇1A), b236 (∇2A), or b237 (∇3A), were examined as in Figure 2B. (B) The

1208   same mCherry-linked TAs were analyzed in CDD1031 cells expressing Sec63p-GFP

1209   from pJK59. Scale bar, 5 μm.

1210

1211   **Figure S9.** Temperature and selection-dependent proliferation in the presence of

1212   charges within the Fis1p MAD. (A) The A144D replacement within Gal4-sfGFP-Fis1p

1213   allows for strong Gal4-dependent proliferation under stringent selective conditions.

1214   Cells spotted in Figure 8A were also spotted to SC-Ura medium and incubated at 30℃

1215   for 2 d. (B) Elevated temperature further distinguishes the behavior of cells expressing

1216   Gal4-sfGFP-Fis1p carrying positively charged or negatively charged amino acid

1217   replacements. Cells used in Figure 8A were also spotted to SMM-His + 20mM 3-AT

43

1218 and incubated at 37°C for 4 d. (C) Cells were treated as in (B), with incubation at 18°C

1219 for 4 d. (D) Extended incubation of the SMM-His + 20 mM 3-AT plate shown in Figure

1220 8A for 4 d at 30°C.

1221

1222 **Figure S10.** Positive and negative charge replacements within the Fis1p TA can

1223 generally support Fis1p function. (A) Normal mitochondrial morphology can be

1224 maintained even when charges are placed within the hydrophobic MAD of the Fis1p

1225 TA. *fis1Δ* strain CDD741, in which mitochondria were visualized by expression of

1226 mitochondria-targeted GFP from plasmid pHS12, was transformed with vector

1227 pRS313, a plasmid expressing WT Fis1p (b239), or plasmids expressing Fis1p

1228 containing the indicated TA alterations (plasmids b240-251). Cells were treated with

1229 sodium azide to provoke Fis1p-dependent mitochondrial fragmentation, and

1230 representative images are shown. Scale bar, 5 μm. (B) Cells treated as in (A) were

1231 scored for the maintenance of a mitochondrial network (n=200 cells). (C) A genetic

1232 assay of Fis1p function demonstrates that charged residues within the TA allow Fis1p

1233 activity. *fzo1Δ fis1Δ* strain CDD688, carrying a CHX-counterselectable, *FZO1*-

1234 expressing plasmid, was transformed with the *FIS1*-expressing plasmids enumerated

1235 above. After allowing cells to lose the *FZO1*-encoding plasmid, serial dilutions were

1236 spotted to SLac-His medium containing 3 μg/mL CHX to counterselect against *FZO1*

1237 expression ("lactate / no fusion") and incubated for 5 d to test for maintenance of

1238 mtDNA. Cell proliferation indicates a lack of Fis1p activity. To control for cell number

1239 spotted, cells from the same culture were placed on SMM-Trp-His medium ("glucose /

1240 fusion") and incubated for 2 d.

1241

1242 **Figure S11.** A block of peroxisomal membrane biogenesis does not lead to re-

1243 localization of the Fis1p TA to the cytosol or nucleus. (A) *WT* strain BY4742 and *pex3Δ*

1244 strain CDD974 were transformed with plasmid b109, expressing mCherry fused to a

1245 wild-type Fis1p TA, and imaged by fluorescence microscopy. Scale bar, 5 μm. (B)

1246 Deletion of Pex3p does not lead to increased Gal4-sfGFP-Fis1p transcriptional activity.

1247 *PEX3* strain MaV203 and isogenic *pex3Δ* strain CDD1172 were transformed with

44

1248   plasmids b100 (Gal4-sfGFP-Fis1), b101 [Gal4-sfGFP-Fis1(ΔTA)], or empty vector pKS1

1249   and treated as in Figure 8A. (C) Relative activity of Gal4-sfGFP-Fis1p containing proline

1250   substitutions does not change upon Pex3p removal. Strain CDD1172 expressing Gal4-

1251   sfGFP-Fis1p variants from plasmids b100 (WT), b188 (V134P), b189 (G137P), b129

1252   (L139P), b190 (A140P), b296 (A144P), or b101 (ΔTA) was treated as in (B). (D) Relative

1253   differences in Gal4-sfGFP-Fis1p transcriptional activity among charge-substituted TAs

1254   do not change in the absence of Pex3p. Strain CDD1172 transformed with plasmids

1255   encoding the indicated mutations within the Fis1p TA (plasmids b173-b187 and b295)

1256   was treated as in (B).

1257

1258   **Table S1.** Tail anchor sequence counts from individual pools.

1259

1260   **Table S2.** Strains used during this study.

1261

1262   **Table S3.** Plasmids used for experiments during this study.

1263

1264   **Table S4.** Oligonucleotides used in this study.

1265

1266

1267

1268

1269

1270

1271

Figure 1

Figure 2

Figure 3

Figure 4

A

B

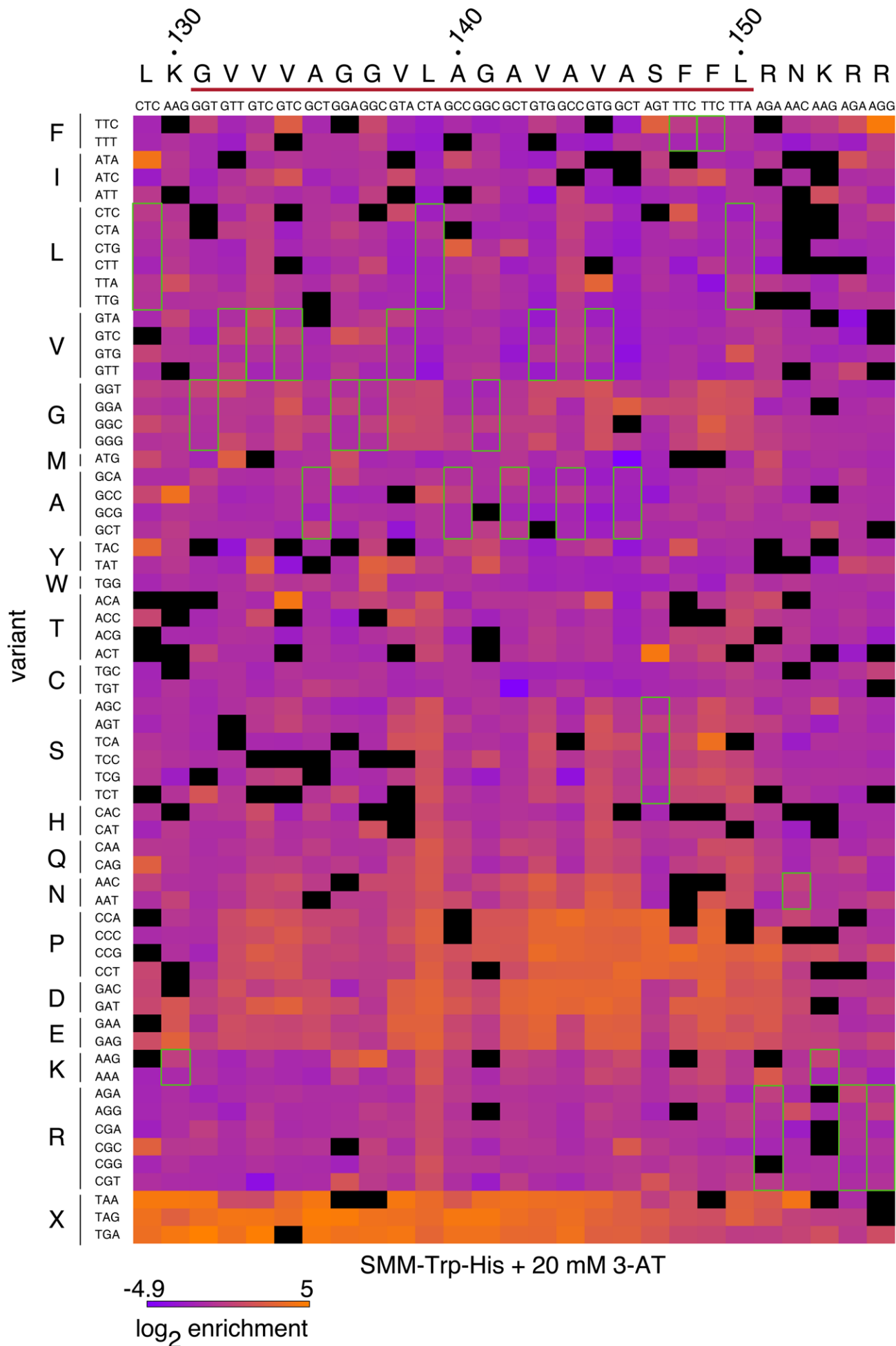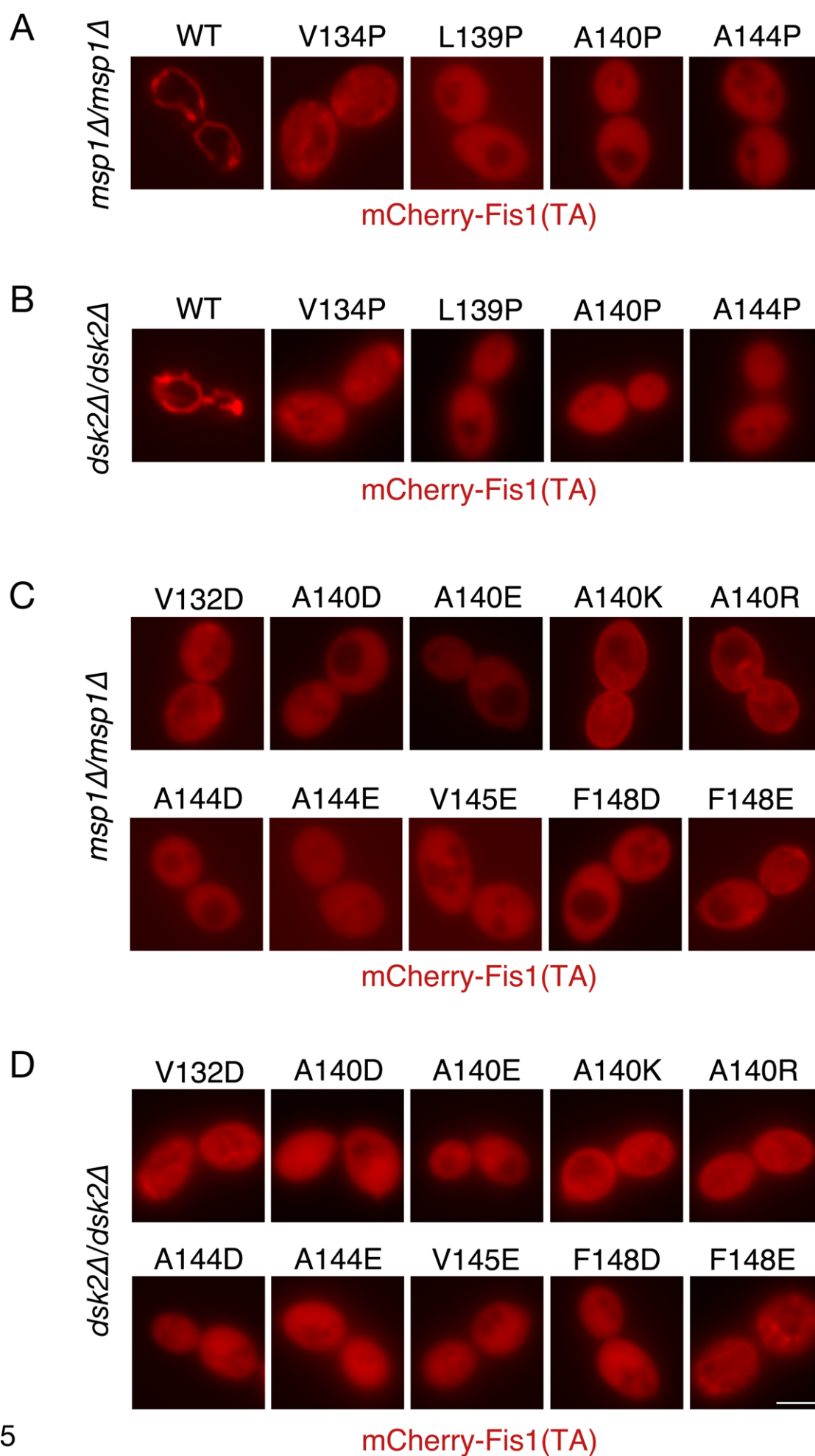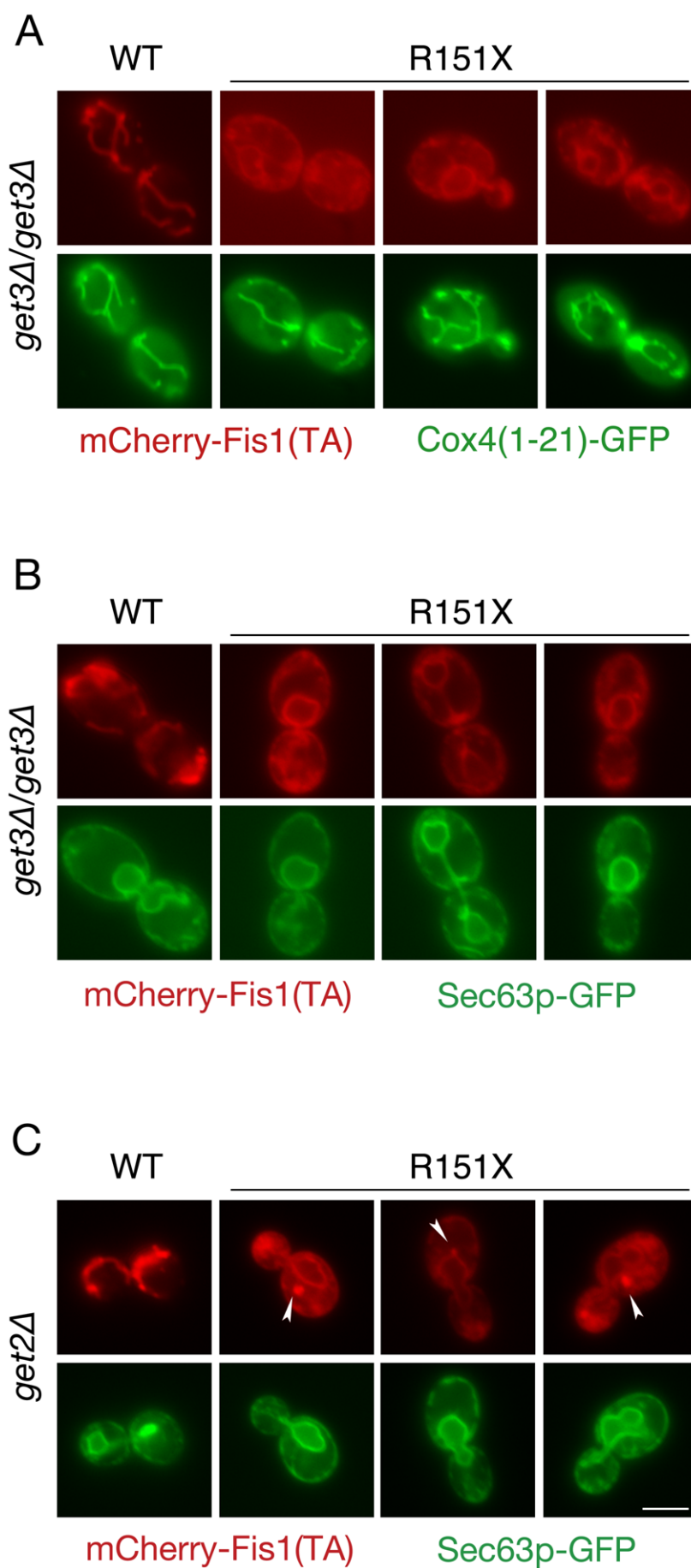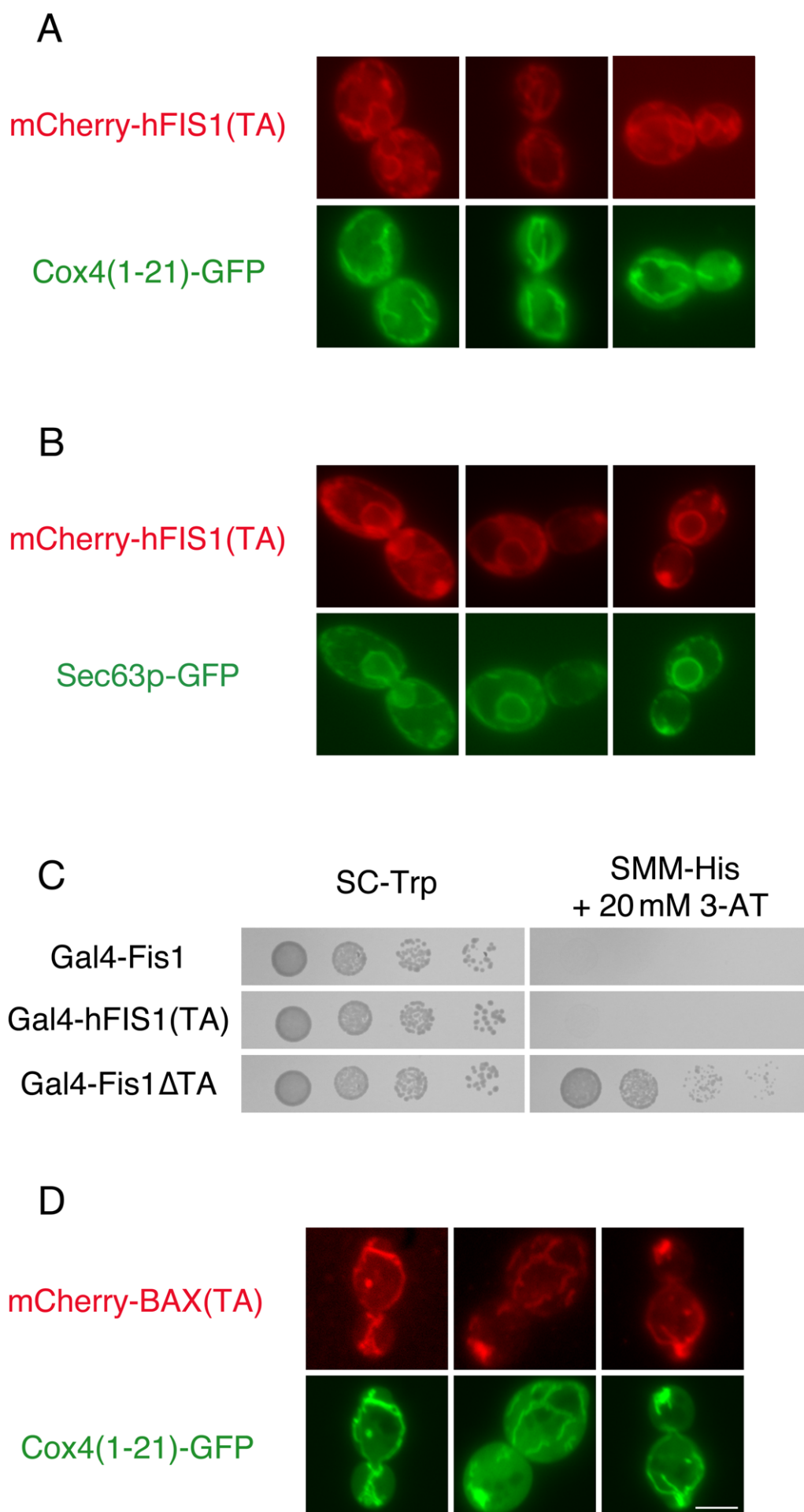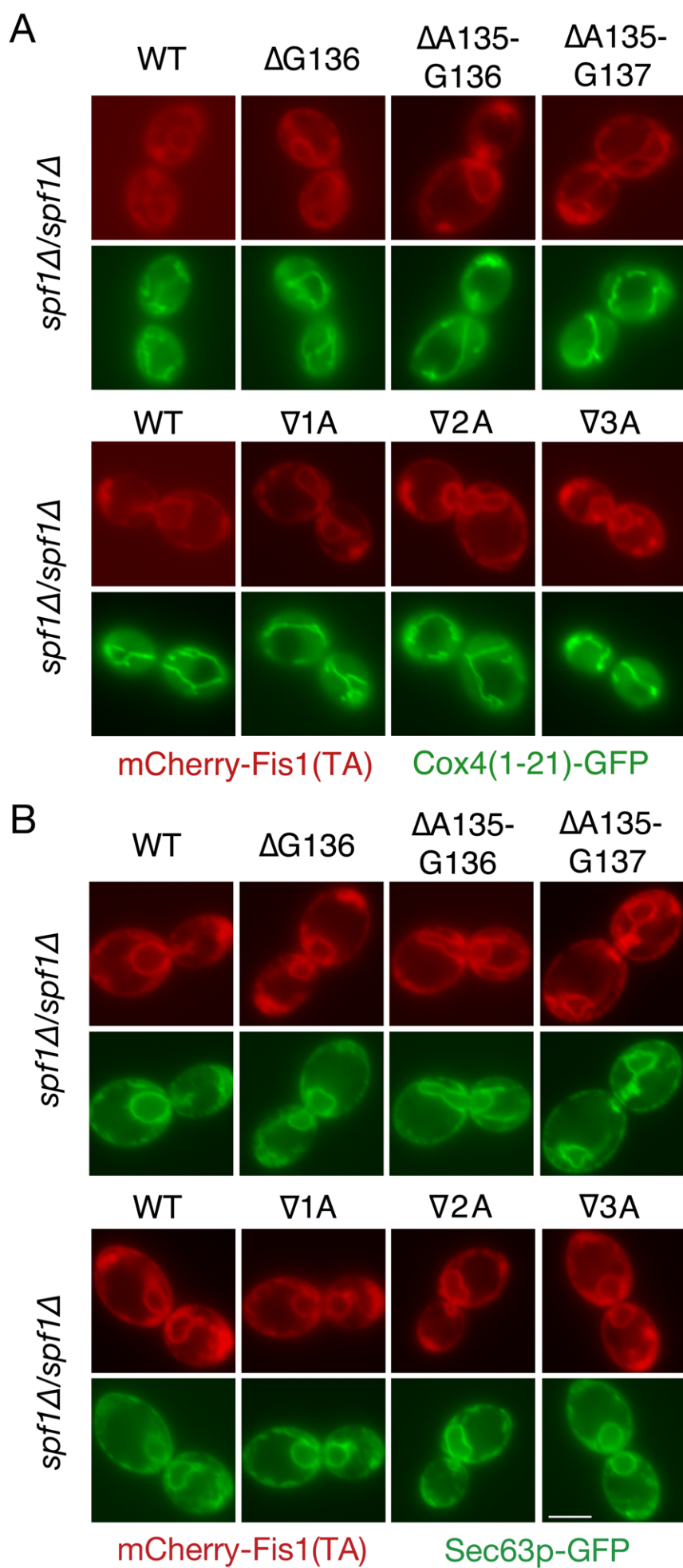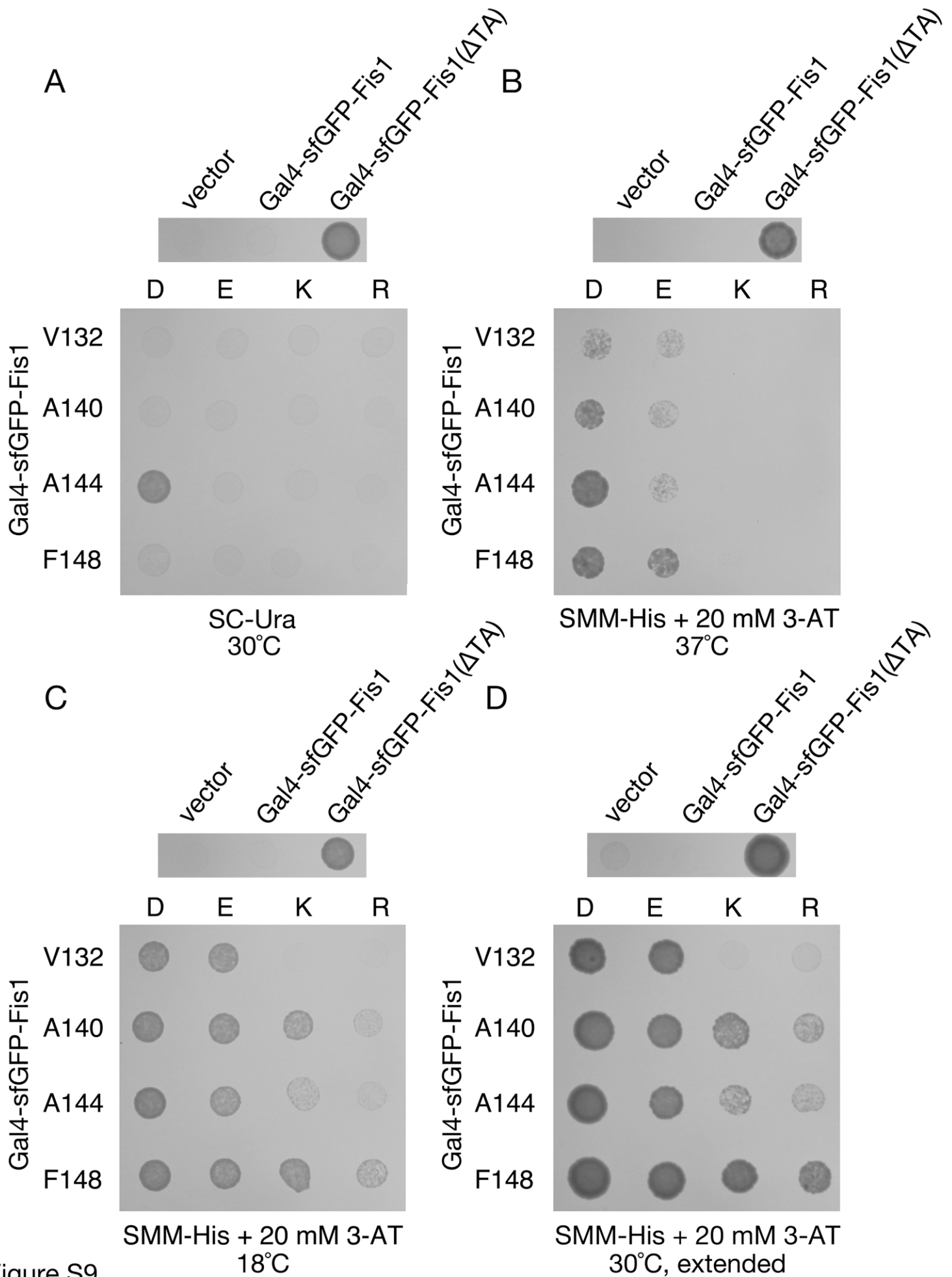mCherry-Fis1(TA)   Cox4(1-21)-GFP
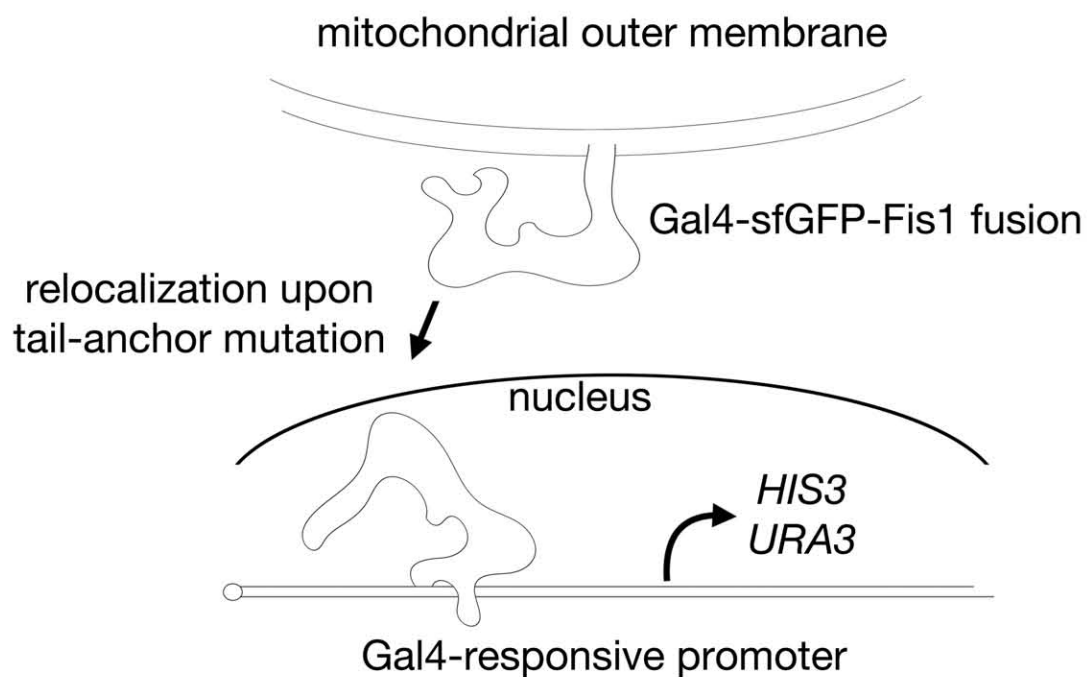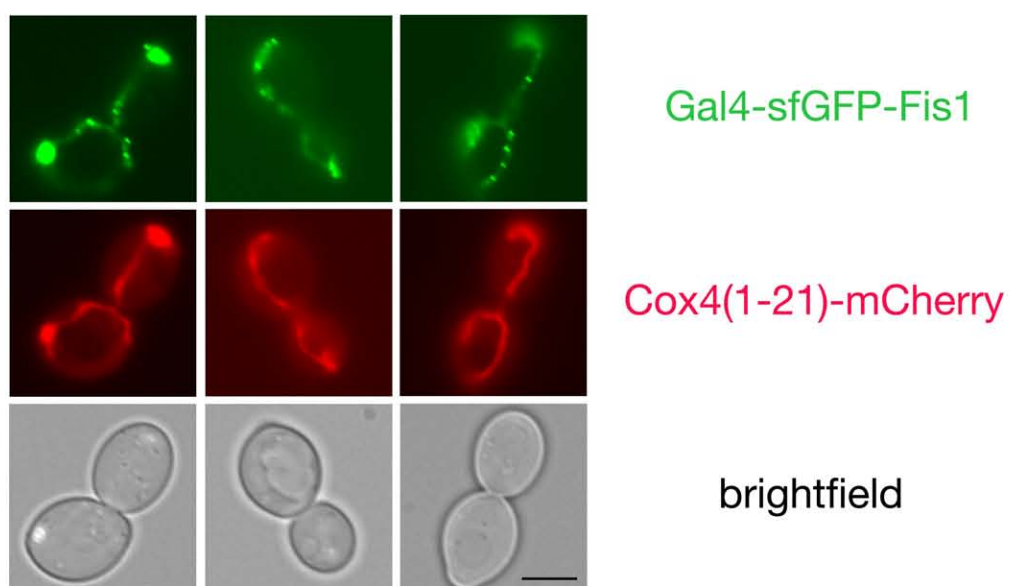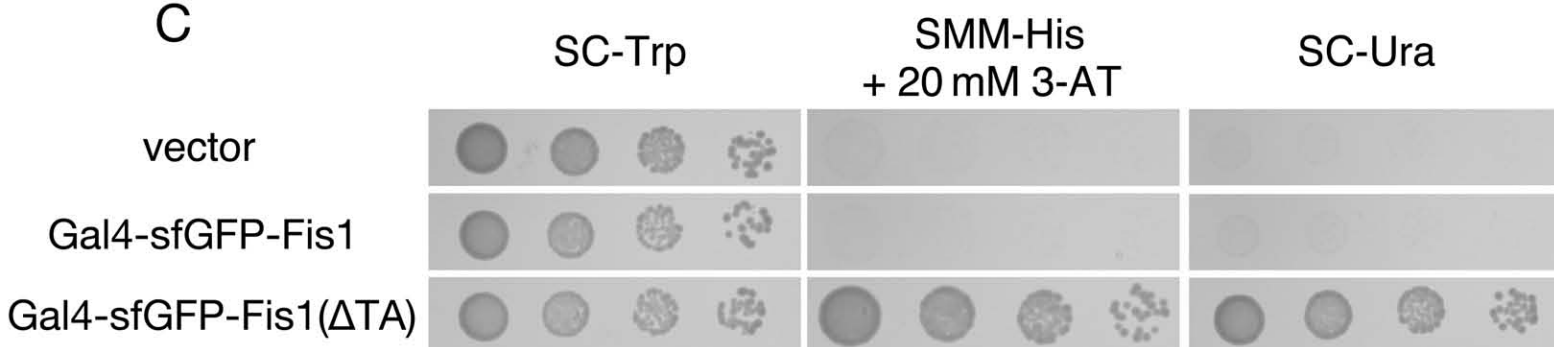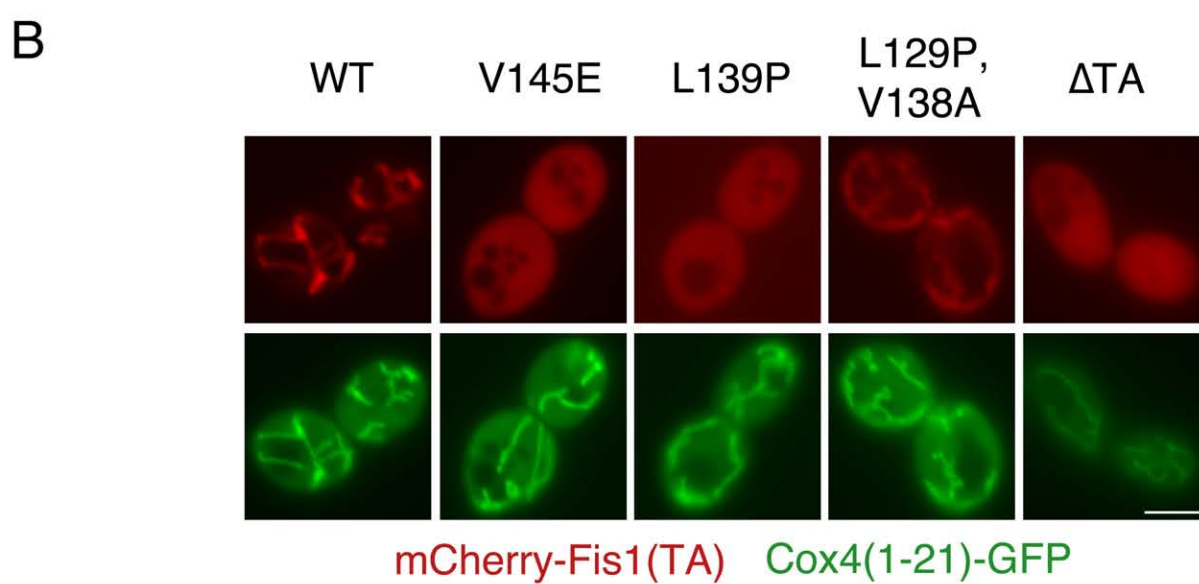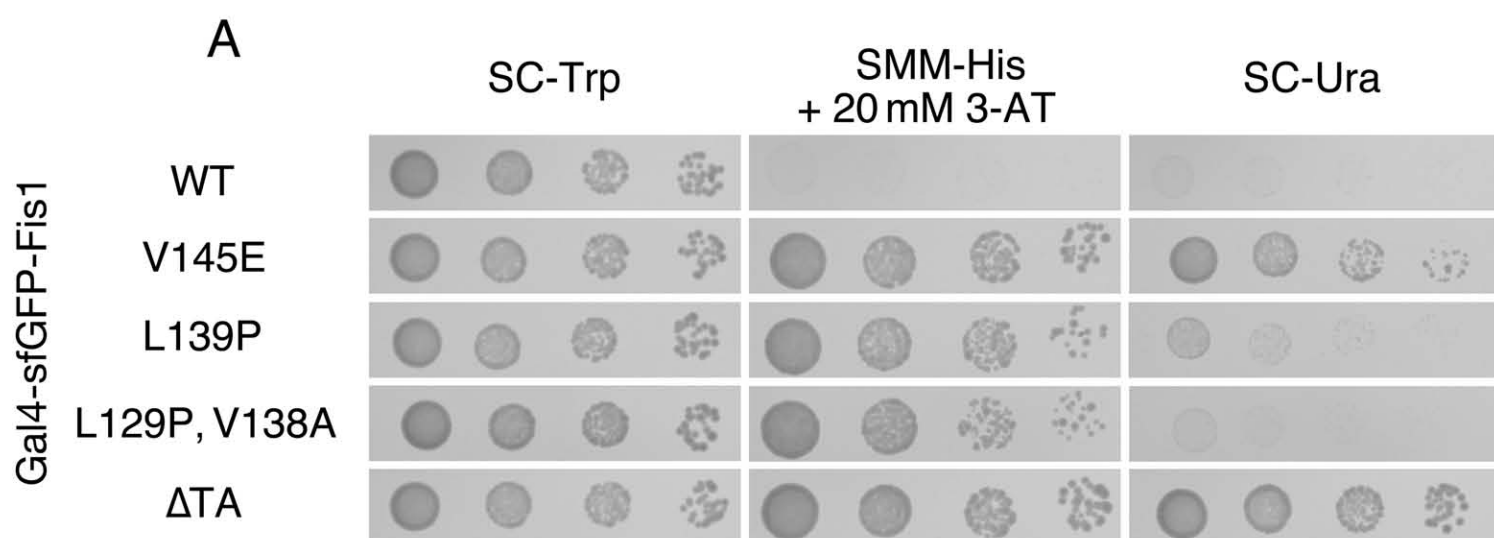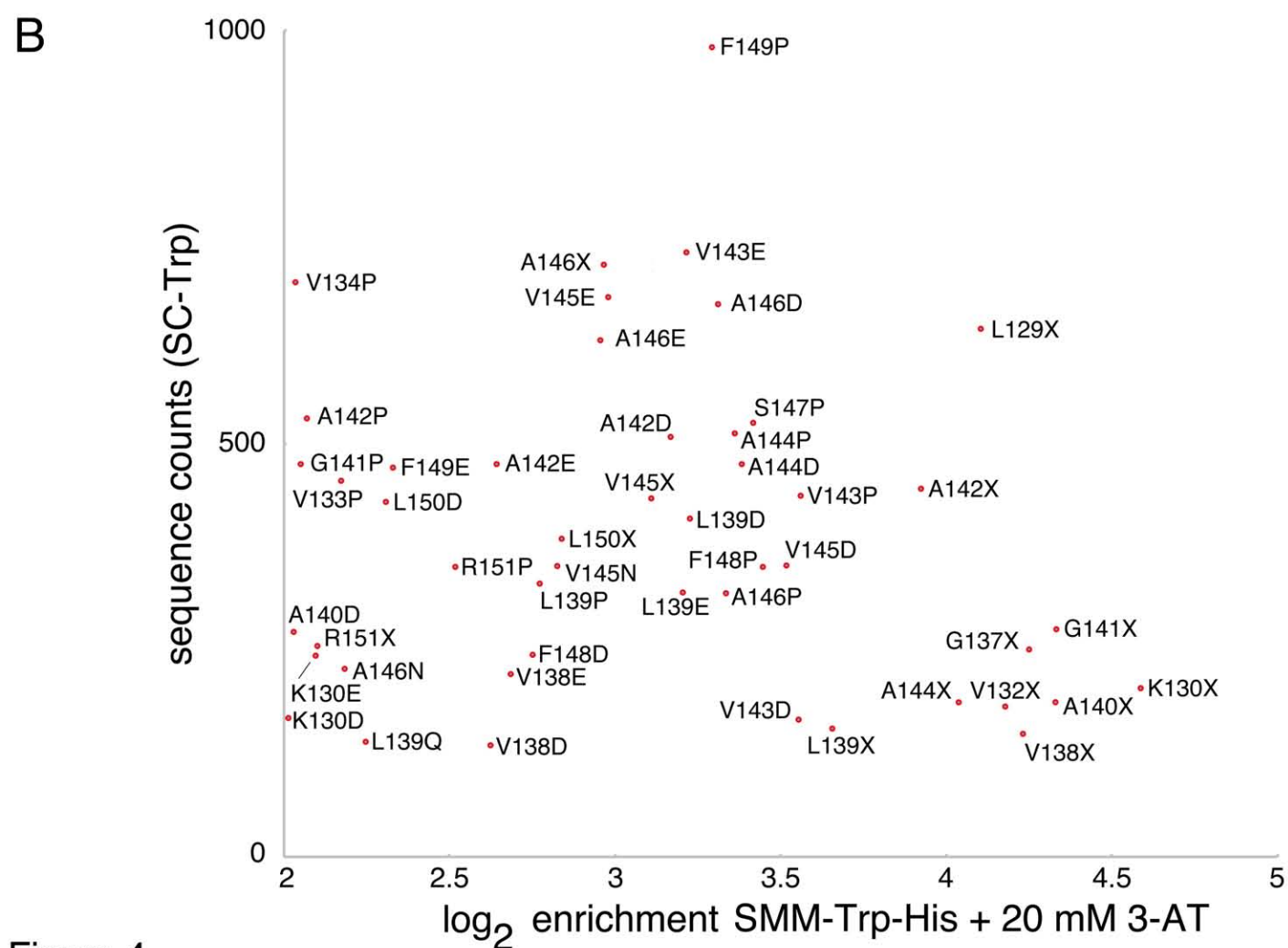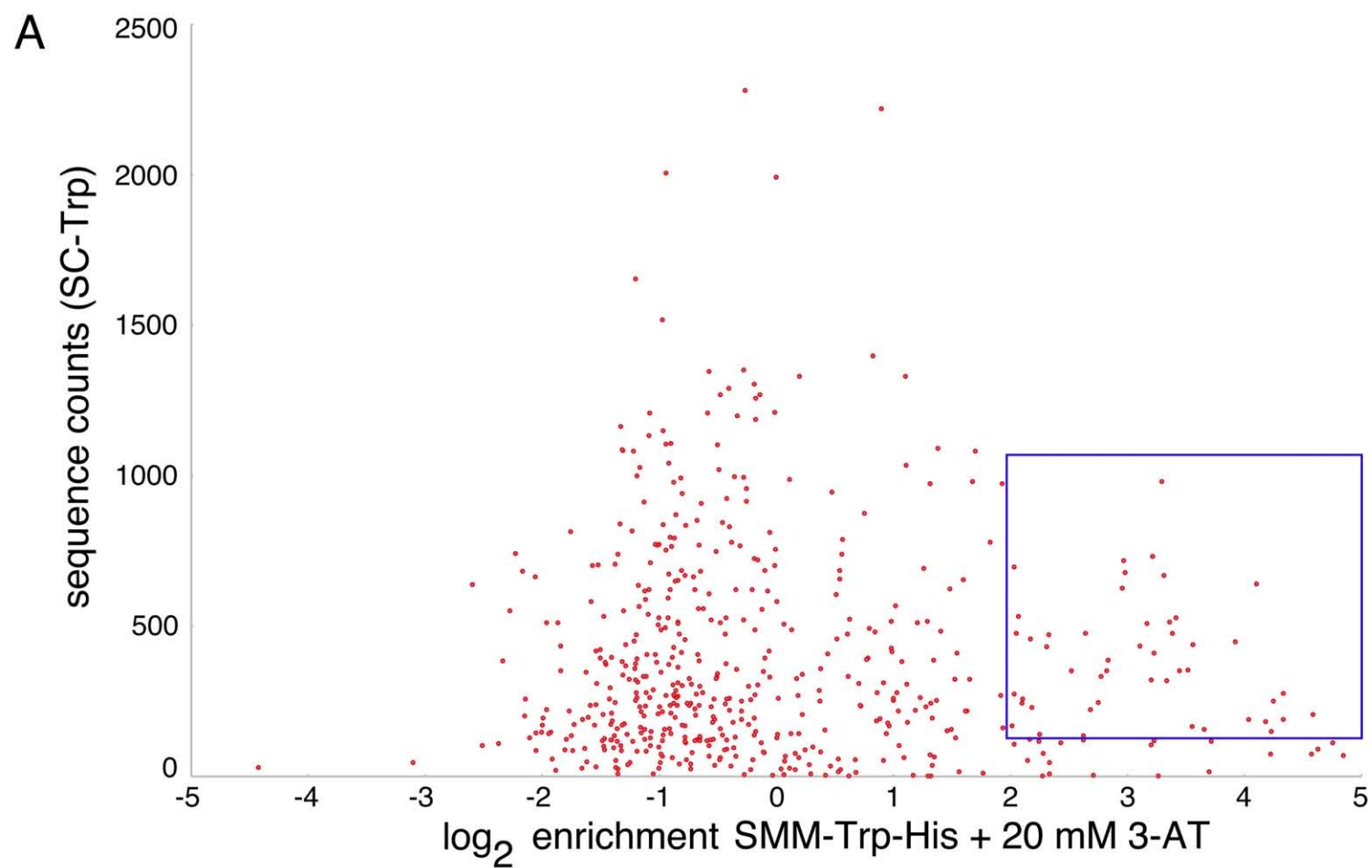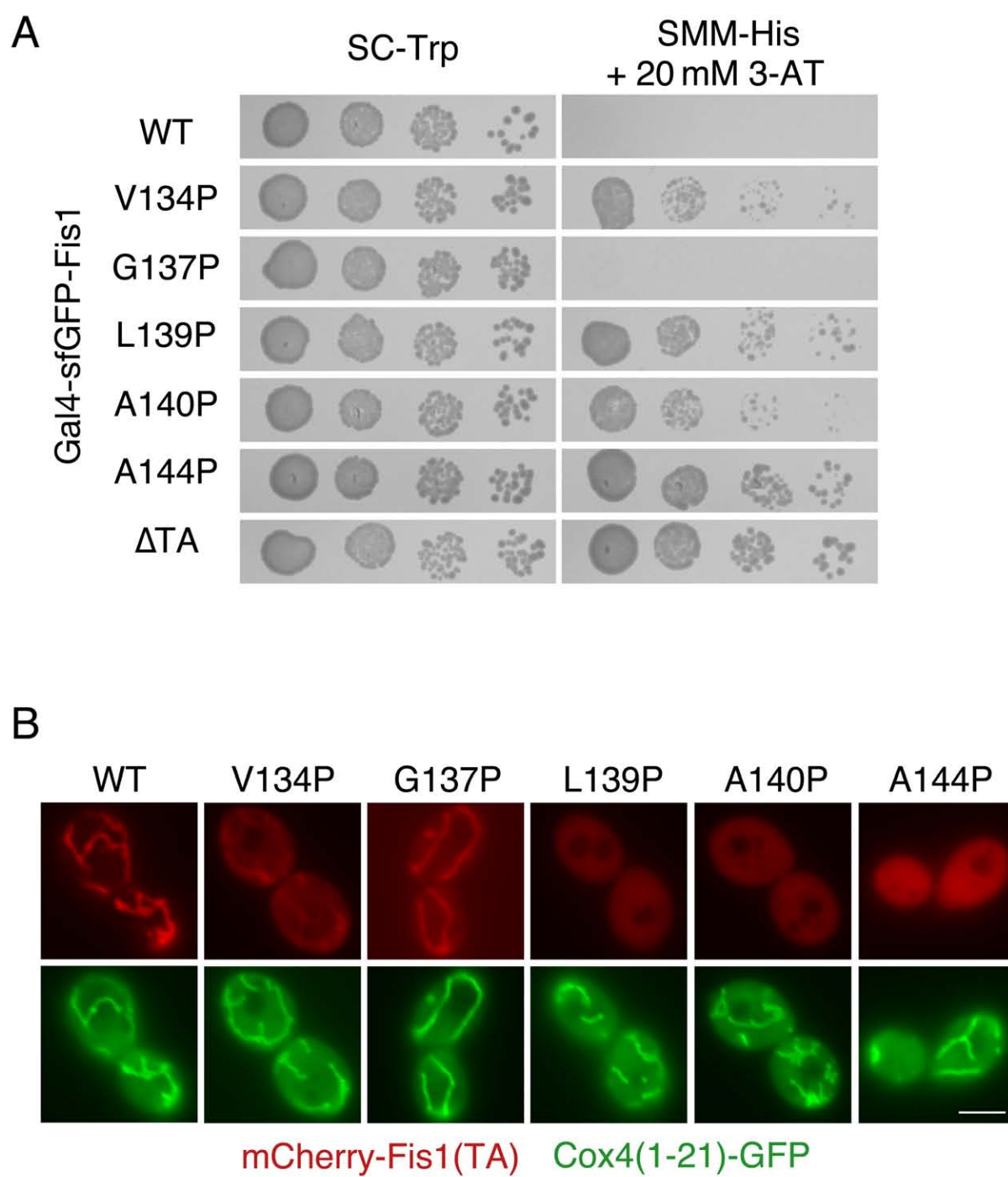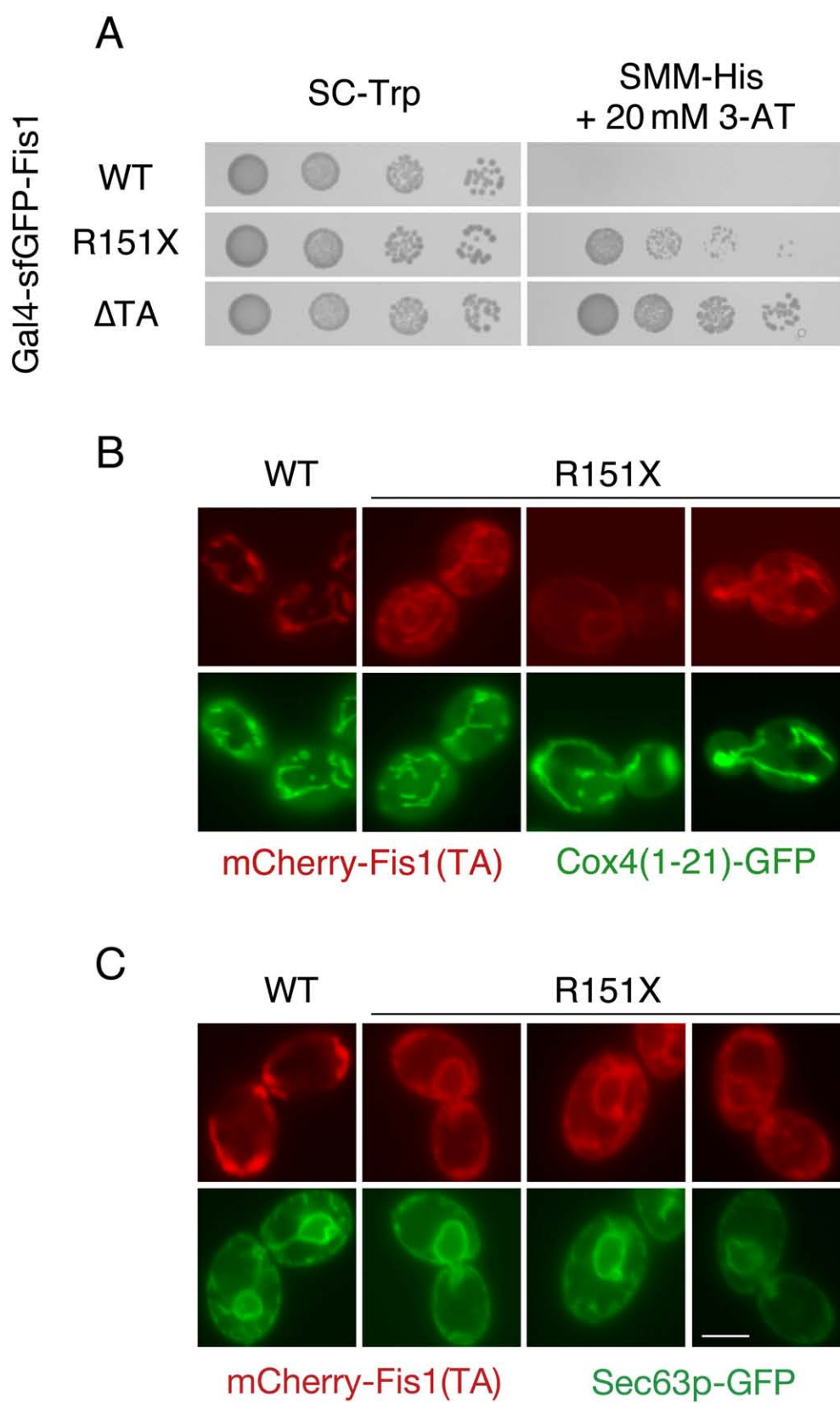
Figure 5

Figure 6

Figure 7

Figure 8

mCherry-Fis1(TA) Cox4(1-21)-GFP
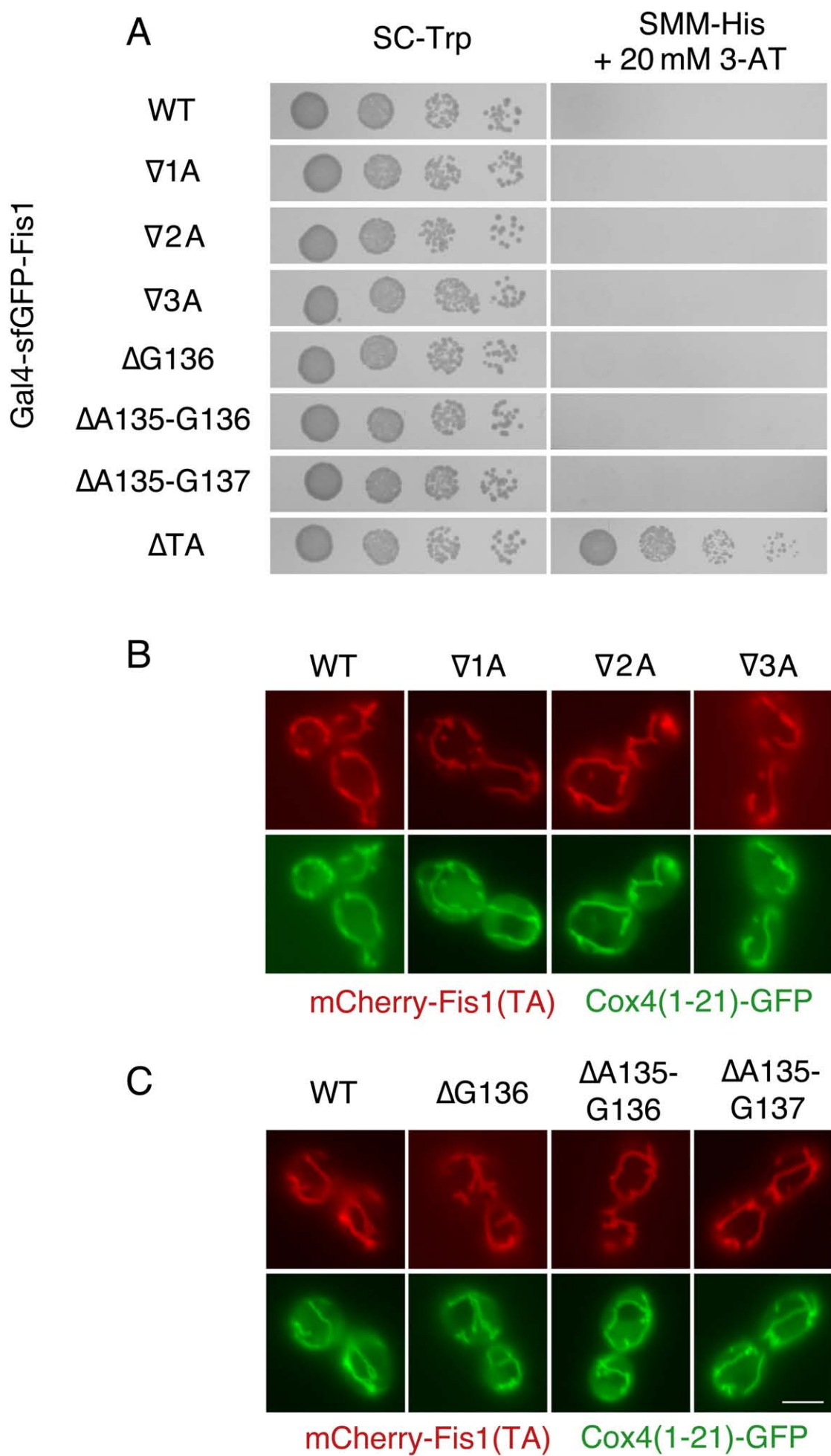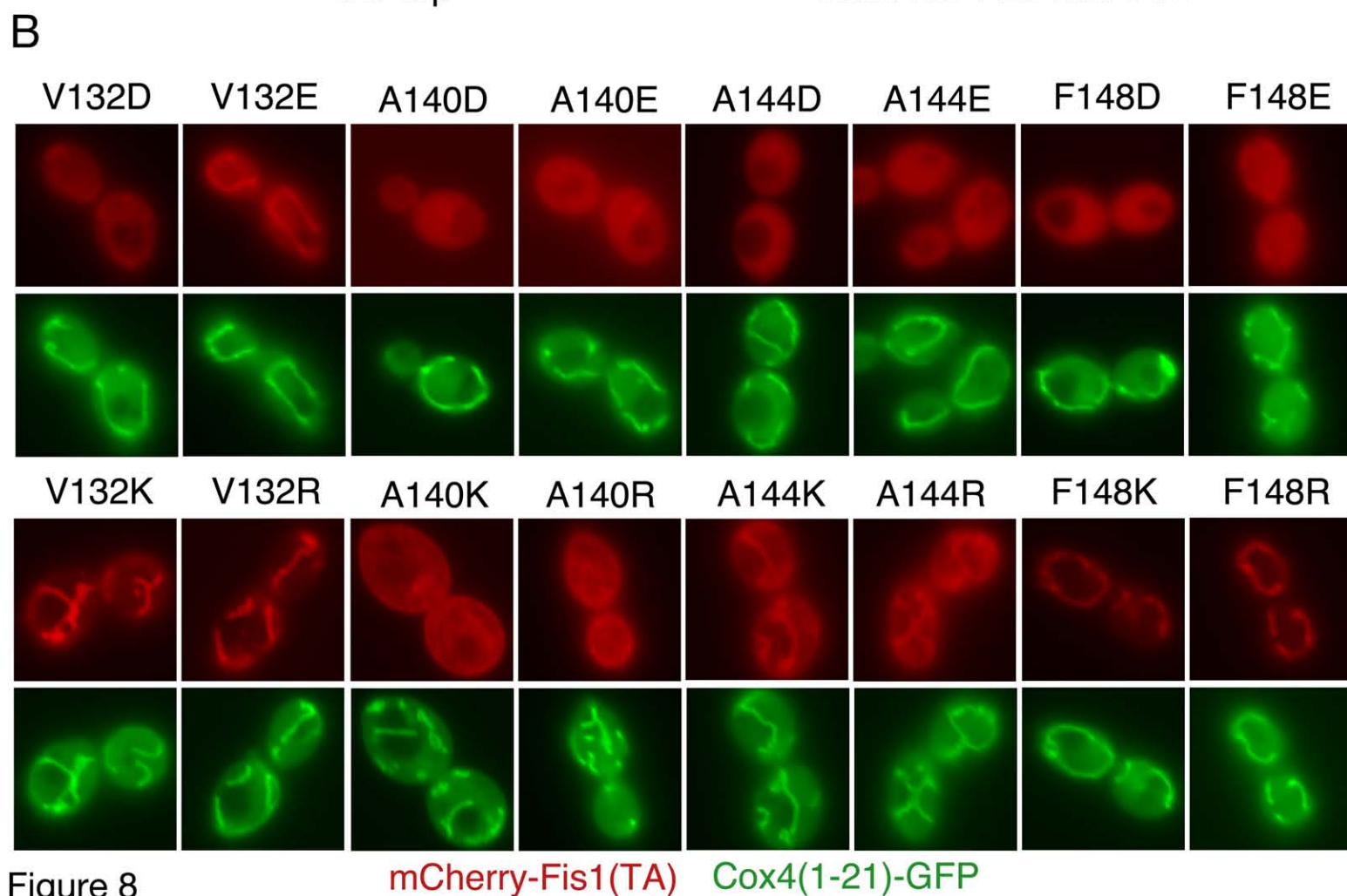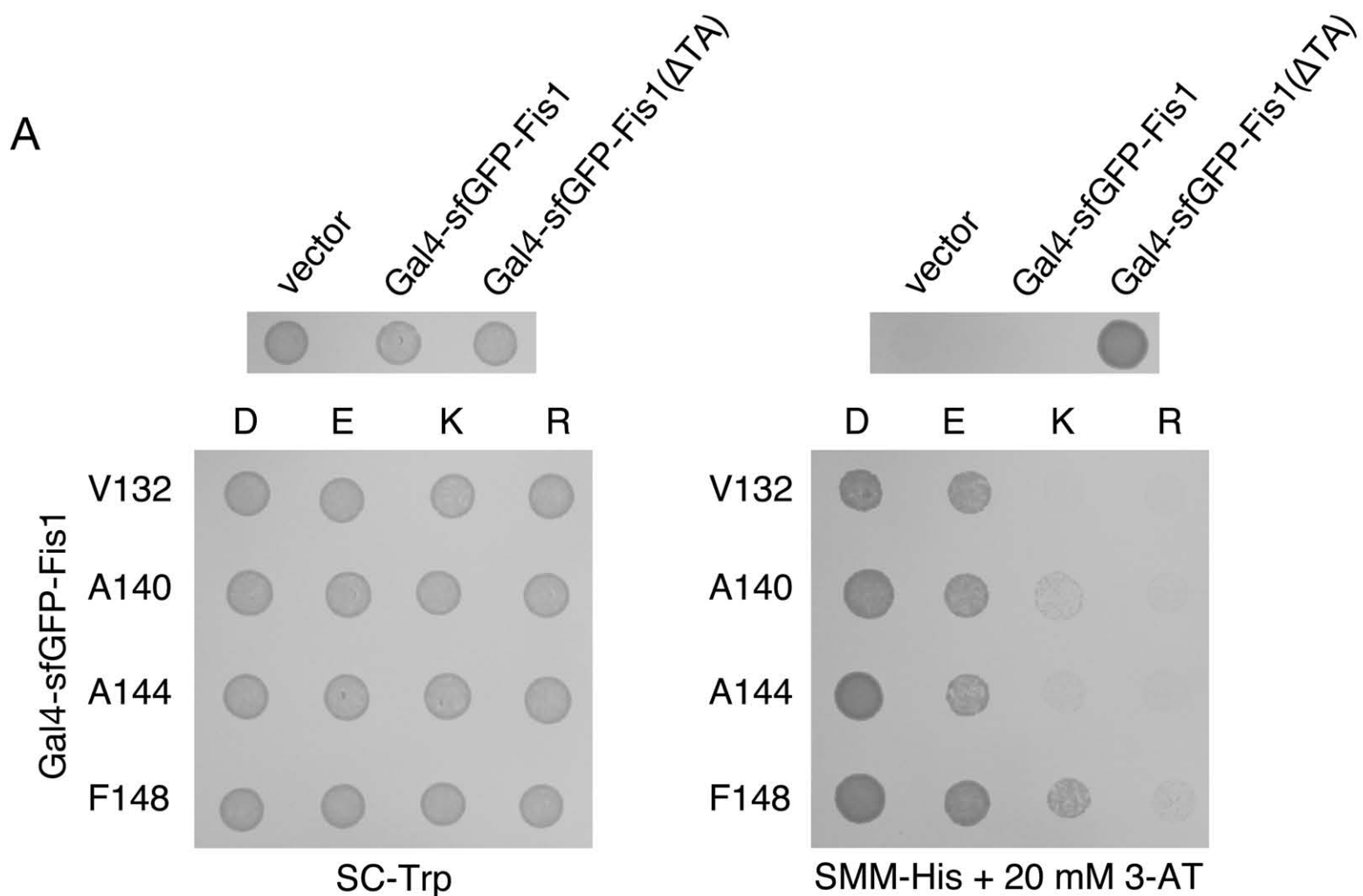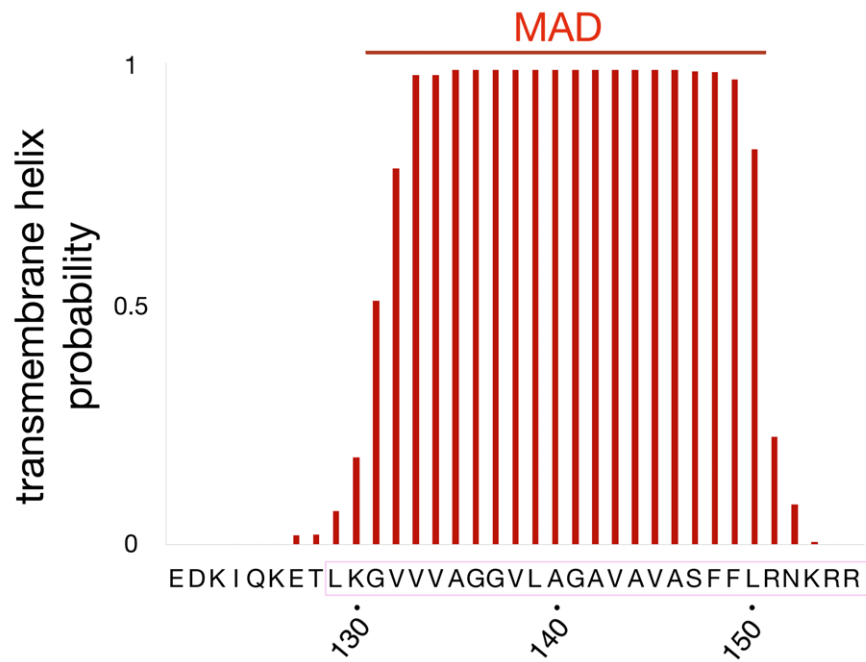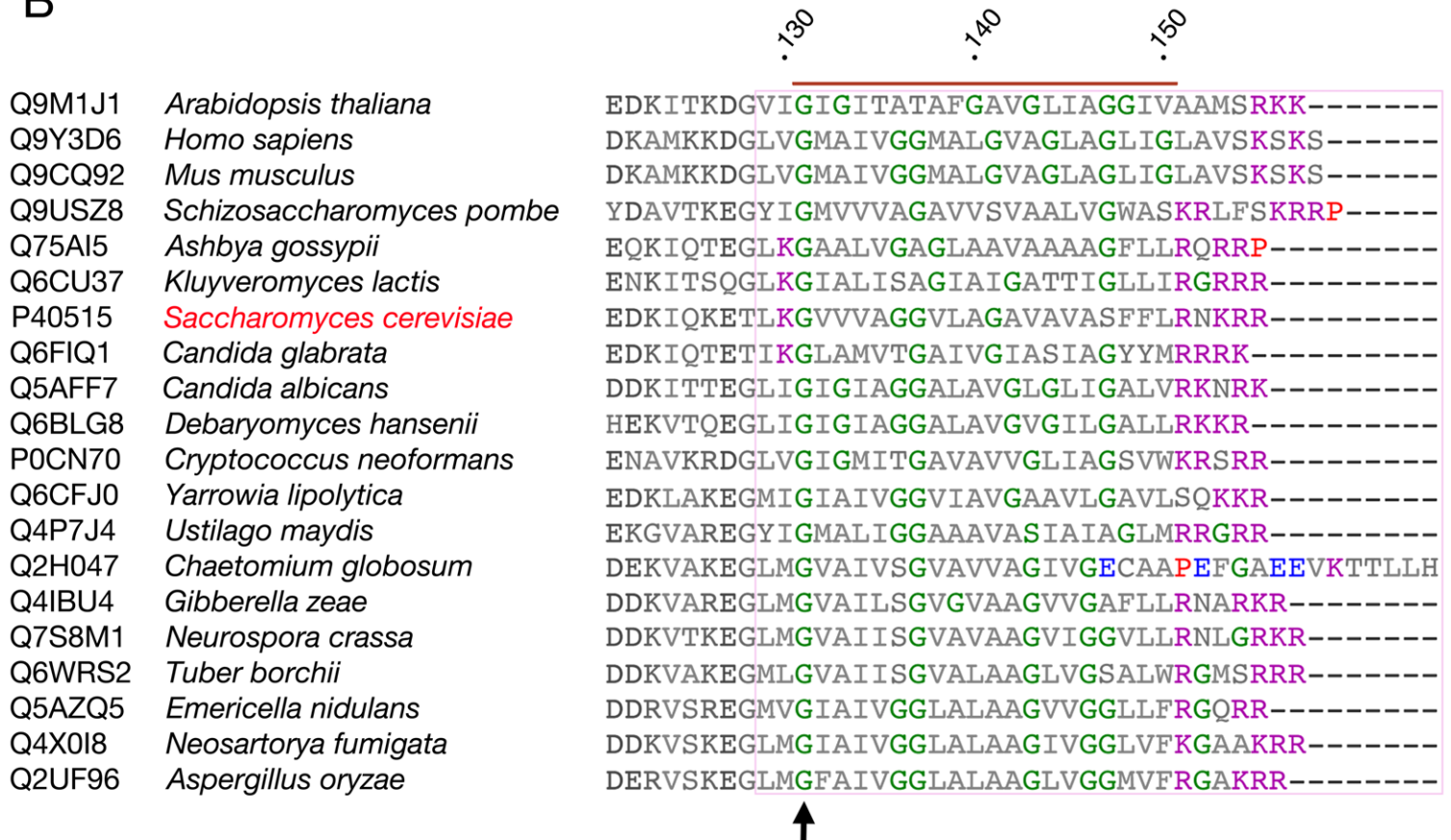
Figure S1

Figure S2

Figure S3

Figure S4

A

msp1Δ/msp1Δ

| WT | V134P | L139P | A140P | A144P |

mCherry-Fis1(TA)

B

dsk2Δ/dsk2Δ

| WT | V134P | L139P | A140P | A144P |

mCherry-Fis1(TA)

C

msp1Δ/msp1Δ

| V132D | A140D | A140E | A140K | A140R |
| A144D | A144E | V145E | F148D | F148E |

mCherry-Fis1(TA)

D

dsk2Δ/dsk2Δ

| V132D | A140D | A140E | A140K | A140R |
| A144D | A144E | V145E | F148D | F148E |

Figure S5

mCherry-Fis1(TA)

Figure S6

A

mCherry-hFIS1(TA)

Cox4(1-21)-GFP

B

mCherry-hFIS1(TA)

Sec63p-GFP

C

	SC-Trp	SMM-His + 20 mM 3-AT
Gal4-Fis1		
Gal4-hFIS1(TA)		
Gal4-Fis1ΔTA		

D

mCherry-BAX(TA)

Cox4(1-21)-GFP
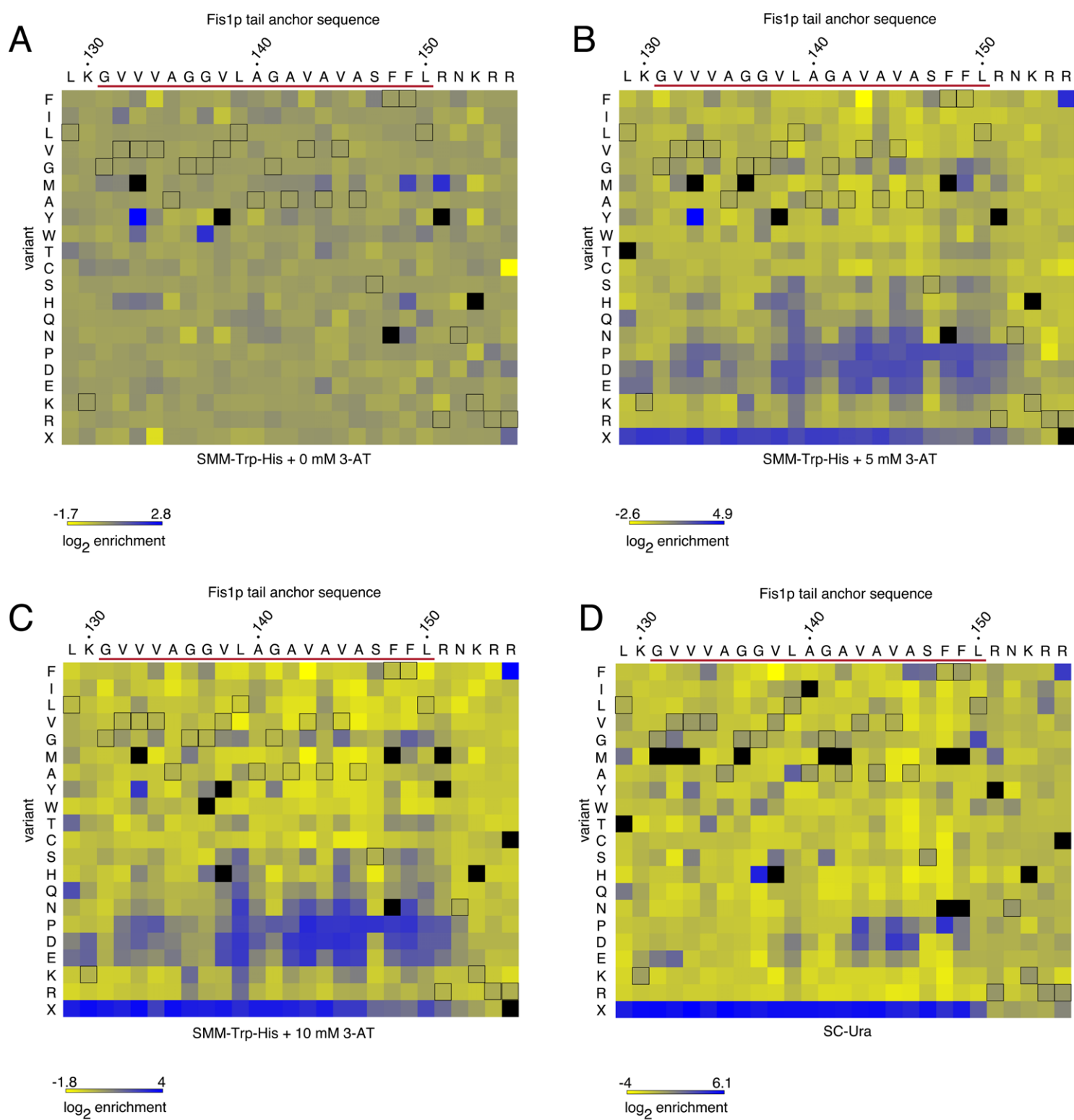
Figure S7

Figure S8

Figure S9

A

WT   vector

(+ azide)

V132D   V132E   V132K   V132R

A144D   A144E   A144K   A144R

F148D   F148E   F148K   F148R

Cox4(1-21)-GFP

B

% cells with networks
undisrupted by azide

C

Figure S10

Figure S11