Plant Cysteine Oxidases are Dioxygenases that Directly Enable Arginyl Transferase-Catalyzed Arginylation of N-End Rule Targets

Mark D. White[1], Maria Klecker[2,3], Richard J. Hopkinson[1], Daan Weits[4], Carolin Mueller[5,6], Christin Naumann[2,3], Rebecca O'Neill[1], James Wickens[1], Jiayu Yang[1], Jonathan C. Brooks-Bartlett[7], Elspeth F. Garman[7],Tom N. Grossmann[5,6], Nico Dissmeyer[2,3,*], Emily Flashman[1,*]

[1] Chemistry Research Laboratory, University of Oxford, 12 Mansfield Road, Oxford, OX1 3TA, United Kingdom
[2] Independent Junior Research Group on Protein Recognition and Degradation, Leibniz Institute of Plant Biochemistry (IPB), Weinberg 3, D-06120 Halle (Saale), Germany
[3] ScienceCampus Halle – Plant-based Bioeconomy, Betty-Heimann-Str. 3, D-06120 Halle (Saale), Germany
[4] Institute of Biology I, RWTH Aachen University, Worringerweg 1, D-52074 Aachen, Germany
[5] Chemical Genomics Centre of the Max Planck Society, Department of Chemistry and Chemical Biology, Technische Universität Dortmund, Otto-Hahn-Str. 15, D-44227 Dortmund, Germany
[6] VU University Amsterdam, De Boelelaan 1083, 1081 HV Amsterdam, The Netherlands
[7] Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, United Kingdom

* Corresponding authors: emily.flashman@chem.ox.ac.uk, +44(0)1865 275920
nico.dissmeyer@ipb-halle.de, ++49 (0) 345 5582 1710

Keywords
cysteine dioxygenase; ethylene response factors; plant hypoxia; N-end rule pathway; arginyl transferase; E3 ubiquitin ligase; proteostasis

Abbreviations
PCO, plant cysteine oxidase; ATE, arginyl tRNA transferase; ERF-VII, group VII ETHYLENE RESPONSE FACTOR; 2OG, 2-oxoglutarate; NMR, nuclear magnetic resonance; Met, methionine; NME, N-terminal Met excision; Nt, N-terminal; NO, nitric oxide; HIF, hypoxia-inducible factor; PHD, prolyl hydroxylase; MALDI-MS, matrix-assisted laser desorption/ionization-mass spectrometry; LC-MS, liquid chromatography-mass spectrometry; HRE, HYPOXIA RESPONSIVE ERF; RAP, RELATED TO APETALA2; EBP, ETHYLENE RESPONSE FACTOR 72; CDO, cysteine dioxygenase; MAP, Met-aminopeptidase.

46 **Abstract**

47

48 **Crop yield loss due to flooding is a threat to food security. Submergence-induced**
49 **hypoxia in plants results in stabilisation of group VII ETHYLENE RESPONSE**
50 **FACTORS (ERF-VIIs), which aid survival under these adverse conditions. ERF-VII**
51 **stability is controlled by the N-end rule pathway, which proposes that ERF-VII N-**
52 **terminal cysteine oxidation in normoxia enables arginylation followed by proteasomal**
53 **degradation. The PLANT CYSTEINE OXIDASEs (PCOs) have been identified as**
54 **catalysts of this oxidation. ERF-VII stabilisation in hypoxia presumably arises from**
55 **reduced PCO activity. We directly demonstrate that PCO dioxygenase activity produces**
56 **Cys-sulfinic acid at the N-terminus of an ERF-VII peptide, which then undergoes**
57 **efficient arginylation by an arginyl transferase (ATE1). This is the first molecular**
58 **evidence showing N-terminal Cys-sulfinic acid formation and arginylation by N-end**
59 **rule pathway components, and the first ATE1 substrate in plants. The PCOs and ATE1**
60 **may be viable intervention targets to stabilise N-end rule substrates, including ERF-**
61 **VIIs to enhance submergence tolerance in agronomy.**

62 **Introduction**

63

64 All aerobic organisms require homeostatic mechanisms to ensure O_2 supply and demand are
65 balanced. When supply is reduced (hypoxia), a hypoxic response is required to decrease
66 demand and/or improve supply. In animals, this well characterized response is mediated by
67 the Hypoxia-Inducible transcription Factor (HIF), which upregulates genes encoding for
68 vascular endothelial growth factor, erythropoietin and glycolytic enzymes amongst many
69 others.[1-3] Hypoxia in plants is typically a consequence of reduced O_2 diffusion under
70 conditions of waterlogging or submergence, or inside of organs such as seeds, embryos, or
71 floral meristems in buds where the various external cell layers act as diffusion barriers.
72 Although plants can survive temporary periods of hypoxia, flooding negatively impacts on
73 plant growth, and if sustained it can result in plant damage or death[4]. This has a major impact
74 on crop yield; for example, flooding resulted in crop loss costing $3 billion in the U.S. in
75 2011.[5] As climate change results in increased severe weather events including flooding[4],
76 strategies to address crop survival under hypoxic stress are needed to meet the needs of a
77 growing worldwide population.

78

79 The response to hypoxia in rice, Arabidopsis, and barley is known to be mediated by the
80 group VII ETHYLENE RESPONSE FACTORs (ERF-VIIs).[6-11] It has been found that these
81 transcription factors promote the expression of core hypoxia-responsive genes, including
82 those encoding alcohol dehydrogenase and pyruvate decarboxylase that facilitate anaerobic
83 metabolism.[12,13] Crucially, it was shown, initially in Arabidopsis, that the stability of the
84 ERF-VIIs is regulated in an O_2-dependent manner via the Arg/Cys branch of the N-end rule
85 pathway, which directs proteins for proteasomal degradation depending on the identity of
86 their N-terminal amino acid.[14-16] Thus, a connection between O_2 availability and the plant

87 hypoxic response was identified.[11,17,18] The Arabidopsis ERF-VIIs are translated with the
88 conserved N-terminal motif MCGGAI/VSDY/F[4] and co-translational N-terminal methionine
89 excision, catalyzed by Met amino peptidases (MAPs)[19,20], leaves an exposed N-terminal Cys
90 which is susceptible to oxidation.[14-16] N-terminally oxidized Cys residues (Cys-sulfinic acid
91 or Cys-sulfonic acid, **Supplementary Figure 1**) are then proposed to render the ERF-VII N-
92 termini substrates for arginyl tRNA transferase (ATE)-catalyzed arginylation. The subsequent
93 Nt-Arg-ERF-VIIs are candidates for ubiquitination by the E3 ligase PROTEOLYSIS6
94 (PRT6)[21] which promotes targeted degradation via the 26S proteasome. It has also been
95 shown that degradation of ERF-VIIs by the N-end rule pathway can be influenced by NO,
96 and that the ERF-VIIs play a role in plant NO-mediated stress responses.[22,23]
97
98 The plant hypoxic response mimics the equivalent well-characterized regulatory system in
99 animals, whereby adaptation to hypoxia is mediated by HIF. In normoxic conditions, HIF is
100 hydroxylated at specific prolyl residues targeting it for binding to the von Hippel-Lindau
101 tumour suppressor protein (pVHL), the recognition component of the E3-ubiquitin ligase
102 complex, which results in HIF ubiquitination and proteasomal degradation.[1,3] Thus, while not
103 substrates for the N-end rule pathway of protein degradation, HIF levels are regulated by
104 post-translational modification resulting in ubiquitination, in a manner that is sensitive to
105 hypoxia. HIF prolyl hydroxylation is catalyzed by O_2-dependent enzymes, the HIF prolyl
106 hydroxylases (PHDs 1-3),[2] which are highly sensitive to O_2 availability.[24,25] These O_2-sensing
107 enzymes are thus the direct link between O_2 availability and the hypoxic response.[26,27]
108
109 Crucially, a family of five enzymes, the PLANT CYSTEINE OXIDASEs (PCO1-5) were
110 identified in Arabidopsis[28] that were reported to catalyze the O_2-dependent reaction in the
111 plant hypoxic response, specifically the oxidation of the conserved Cys residue at the N-

terminus of the Arabidopsis ERF-VIIs, RAP2.2, RAP2.12, RAP2.3, HRE1 and HRE2. It was found that overexpression of PCO1 and 2 *in planta* specifically led to depleted RAP2.12 protein levels and reduced submergence tolerance, while *pco1 pco2* T-DNA insertion mutants accumulated RAP2.12 protein. Isolated recombinant PCO1 and PCO2 were shown to consume O_2 in the presence of pentameric peptides CGGAI corresponding to the N-termini of various ERF-VIIs **(Supplementary Table 1[28]).** The identification of these enzymes indicates that the hypoxic response in plants is enzymatically regulated,[28] potentially in a similar manner to the regulation of the hypoxic response in animals by the HIF hydroxylases. The PCOs may therefore act as plant O_2 sensors.

Validation of the chemical steps in the Arg/Cys branch of the N-end rule pathway is still limited, both in animals and plants. We therefore sought to provide molecular evidence that the PCOs catalyze the oxidation step in ERF-VII proteasomal targeting and to determine if this step is required for further molecular priming by arginylation. Using mass spectrometry and NMR techniques we confirm that PCO1 and also PCO4 – representatives of the 2 different PCO 'subclasses' based on sequence identity and expression behavior[28] – catalyze dioxygenation of the N-terminal Cys of Arabidopsis ERF-VII peptide sequences to Cys-sulfinic acid ($CysO_2$). This oxidation directly incorporates molecular O_2. To our knowledge, these are the first described enzymes that catalyze cysteinyl oxidation, as well as being the first described cysteine dioxygenases in plants. We then verify that the Cys-sulfinic acid product of the PCO-catalyzed reactions is a direct substrate for the arginyl tRNA transferase ATE1, demonstrating that PCO activity is relevant and sufficient for the subsequent step of molecular recognition and modification according to the N-end rule pathway. This provides the first molecular evidence that Nt-Cys-sulfinic acid is a *bona fide* substrate for N-end rule mediated arginylation. Overall, we thus define the PCOs as plant cysteinyl dioxygenases and

137 ATE1 as an active arginyl transferase, establishing for the first time a direct link between
138 molecular O_2, PCO catalysis and ATE1 recognition and modification of N-end rule
139 substrates.
140
141
142 **Results**
143
144 **PCOs catalyze modification of RAP2$_{2\text{-}11}$ in an O_2-dependent manner**
145 N-terminally hexahistidine-tagged recombinant PCO1 and 4 were purified to ~90% purity, as
146 judged by SDS-PAGE (**Supplementary Figure 2a**). Protein identity was confirmed by
147 comparison of observed and predicted mass by LC-MS (PCO1 predicted mass 36,510 Da,
148 observed mass 36,513 Da; PCO4 predicted mass 30,680 Da, observed mass 30,681 Da,
149 **Supplementary Figure 2b**). Both PCO1 and PCO4 were found to be monomeric in solution
150 and to co-purify with substoichiometric levels of Fe(II) (~0.3 atoms Fe(II) per monomer,
151 **Supplementary Figure 2c-e**), in line with the reported parameters of recombinant forms of
152 their distant homologs, the cysteine dioxygenases (CDOs).[28-30] The activity of the purified
153 PCO1 and PCO4 was tested towards a synthetic 10-mer peptide corresponding to the
154 methionine excised N-termini of the ERF-VIIs RAP2.2, RAP2.12, and HRE2 (H_2N-
155 CGGAIISDFI-COOH, hereafter termed RAP2$_{2\text{-}11}$ **Supplementary Table 1**). Assays
156 comprising RAP$_{2\text{-}11}$ at 100 μM in the presence or absence of PCO1 or PCO4 at 0.5 μM
157 underwent aerobic or anaerobic coincubation for 30 minutes at 30°C prior to analysis of the
158 peptide by matrix-assisted laser desorption/ionization-mass spectrometry (MALDI-MS,
159 **Figure 1a,b**;). Only under aerobic conditions and in the presence of PCO1 or PCO4, did the
160 spectra reveal the appearance of two species with mass increases of +32 Da and +48 Da,
161 corresponding to two or three added O atoms, suggesting an O_2-dependent reaction for PCOs

162 1 and 4 (**Figure 1b**), as previously shown for PCOs 1 and 2 (note that supplementation of
163 Fe(II) and/or addition of ascorbate was not required for the end-point PCO1/4 activity assays
164 conducted in this study).[28] These mass shifts were deemed to be consistent with enzymatic
165 formation of Cys-sulfinic ($CysO_2$, +32 Da) and Cys-sulfonic acid ($CysO_3$, +48 Da;
166 **Supplementary Figure 1**). Although homology between the PCOs and CDOs[28,30] leads to
167 the predisposition that they will perform similar chemistry (i.e. catalyse Cys sulfinic acid
168 formation), both Cys-sulfinic and Cys-sulfonic acid are proposed to be Arg transferase
169 substrates in the Arg/Cys branch of N-end rule mediated protein degradation and therefore
170 both were considered as potential products of the PCO-catalysed reaction.[14-16]

171

172 **PCOs are dioxygenases catalyzing the incorporation of both atoms of O_2 into $RAP2_{2-11}$**

173 To ascertain whether the PCOs function as dioxygenases and thus to confirm a direct
174 connection between molecular O_2 and PCO activity, we sought to verify the source of the O
175 atoms in the oxidized $RAP2_{2-11}$ by conducting assays in the presence of $^{18}O_2$ as the
176 cosubstrate or $H_2{}^{18}O$ as the solvent. To probe O_2 as the source of O atoms in the product,
177 anaerobic solutions of $RAP2_{2-11}$ were prepared in sealed vials before addition of PCO4 using
178 a gas-tight syringe. The vials were then purged with $^{16}O_2$ or $^{18}O_2$ and the reactions were
179 allowed to proceed at 30°C for a subsequent 20 minutes. Upon analysis by MALDI-MS, the
180 mass of the products revealed that molecular O_2 was incorporated into the Cys-sulfinic acid
181 product (**Figure 2a**). The Cys-sulfinic acid product had a mass of +32 Da in the presence of
182 $^{16}O_2$ and +36 Da in the presence of $^{18}O_2$, demonstrating addition of two ^{18}O atoms and
183 indicating that O_2 is the source of O atoms in this product. The Cys-sulfonic acid product had
184 a mass of +52 Da in the presence of $^{18}O_2$, indicating a third ^{18}O atom had not been
185 incorporated into this product. To probe whether the source of the additional mass in the
186 apparent Cys-sulfonic acid product was an O atom derived from water, an equivalent reaction

187 was carried out under aerobic conditions in the presence of $H_2{}^{18}O$ ($H_2{}^{18}O$:H_2O in a 3:1 ratio).
188 No additional mass was observed in the peak corresponding to the Cys-sulfonic acid, raising
189 the possibility that the +48 Da species observed by MALDI-MS is not enzymatically formed.
190 Importantly, following incubation in the presence of $H_2{}^{18}O$ no additional mass was observed
191 in the peak corresponding to Cys-sulfinic acid, confirming that this species is a product of a
192 reaction where molecular O_2 is a substrate (**Figure 2b**).
193
194 To further investigate whether the PCO-catalyzed product species observed at +48 Da is
195 enzymatically produced or an artefact of the MALDI-MS analysis method, we turned to
196 liquid chromatography-mass spectrometry (LC-MS) to analyze the products of the PCO-
197 catalyzed reactions. Under these conditions, only peptidic product with a mass increase of
198 +32 Da was observed after incubation with both PCO1 and PCO4, corresponding to the
199 incorporation of two O atoms and the formation of Cys-sulfinic acid (**Figure 2c**), consistent
200 with the products observed using $^{18}O_2$ and $H_2{}^{18}O$ (**Figure 2a,b**). No product was observed
201 with a mass corresponding to Cys-sulfonic acid, which suggested that the +48 Da product
202 detected by MALDI-MS was indeed an artefact. When combined with the observation that
203 significant quantities of Cys-sulfonic acid were not seen in no-enzyme or in anaerobic
204 controls (**Figure 1**), it was hypothesized that the Cys-sulfinic acid product of the PCO-
205 catalyzed reaction is non-enzymatically converted to Cys-sulfonic acid during MALDI-MS
206 analysis, potentially as a result of laser exposure. Upon subjecting the products of PCO1 and
207 4 turnover to MALDI-MS analysis with increasing laser intensity, a direct correlation
208 between laser intensity and the ratio of Cys-sulfonic acid:Cys-sulfinic acid product was
209 observed (**Supplementary Figure 3a**). Of note, significant levels of laser induced formation
210 of +32 and +48 Da species upon analysis of unmodified peptide were not observed
211 (**Supplementary Figure 3b**). Together, these results confirm that the +48 Da species

212 observed following incubation of the PCOs with $RAP2_{2-11}$ are a product of Cys-sulfinic acid
213 exposure to the MALDI-MS laser, and not a product of the PCO-catalyzed reaction. Overall,
214 these data demonstrate that the PCOs are dioxygenase enzymes, similar to the mammalian
215 and bacterial cysteine dioxygenases (CDOs) to which they show sequence homology.[28,30]
216

217 **PCOs catalyze oxidation of N-terminal Cys of $RAP2_{2-11}$ to form Cys-sulfinic acid**

218 Recombinant PCO1 and PCO2 were reported to consume O_2 in the presence of pentameric
219 CGGAI peptides corresponding to the methionine-excised N-terminus of the Arabidopsis
220 ERF-VIIs.[28] To definitively verify that the N-terminal cysteinyl residue of $RAP2_{2-11}$ is indeed
221 the target for the PCO-catalyzed +32 Da modifications, we conducted LC-MS/MS analyses
222 on the reaction products. Fragmentation of $RAP2_{2-11}$ that had been incubated in the presence
223 and absence of PCO1 and PCO4 revealed b- and y-ion series consistent with oxidation of the
224 N-terminal Cys residue (**Figure 3a**), confirming that PCOs 1 and 4 act as cysteinyl
225 dioxygenases.

226 As a final confirmation of the nature of the reaction catalyzed by PCO1 and PCO4, their
227 activity was monitored using ^{1}H-NMR. Reactions were initiated by adding 5 μM enzyme to
228 500 μM $RAP2_{2-11}$ (in the presence of 10% D_2O) and products of the reaction were analysed
229 using a 600 MHz NMR spectrometer. In the presence of both PCO1 and PCO4, modification
230 to the cysteinyl residues was observed, as exemplified by the disappearance of the ^{1}H-
231 resonance corresponding to the β-cysteinyl protons (at δ_H 2.88 ppm) and the emergence of a
232 new ^{1}H-resonance at δ_H 2.67 ppm (**Figure 3b**). The chemical shift of the new resonance is
233 similar to that observed for L-Cys conversion to L-Cys-sulfinic acid by mouse CDO,[31] and
234 also to the chemical shift of an L-Cys-sulfinic acid standard measured under equivalent
235 conditions to the PCO assays (**Supplementary Figure 4**). Therefore, the resonance shift
236 observed upon PCO1/4 reaction was assigned to the β-protons of L-Cys-sulfinic acid. Overall

237 these results provide confirmation at the molecular level that Arabidopsis PCOs 1 and 4 act
238 as plant cysteinyl dioxygenases, catalyzing incorporation of O_2 into N-terminal Cys residues
239 on a RAP2 peptide to form Cys-sulfinic acid.

240

241 **ATE1 arginylates acidic N-termini including Cys-sulfinic acid**

242 We next sought to confirm that the PCO-catalyzed Cys-oxidation to Cys-sulfinic acid renders
243 a RAP2 peptide capable of and sufficient for onward modification by ATE1. Cys-sulfinic
244 acid has been proposed as a substrate for ATE1 on the basis of its structural homology with
245 known ATE1 substrates Asp and Glu, but evidence has only been reported to date for
246 arginylation of Cys-sulfonic acid.[32,33] We further sought to validate the role of a plant ATE1:
247 To date ATE1 has been suggested to be responsible for transfer of ^{3}H-arginine to bovine α-
248 lactalbumin in highly purified plant extracts *in vitro*[34] and RAP2.12 stabilization in *ate1 ate2*
249 double null mutant plant lines implicates ATE1 as an ERF-VII-targeting arginyl transferase
250 *in vivo*.[17,18] To this end, we produced recombinant hexahistidine-tagged Arabidopsis ATE1
251 (**Supplementary Figure 5**) for use in an arginylation assay which detects incorporation of
252 radiolabeled ^{14}C-Arg into biotinylated peptides. C-terminally biotinylated $RAP2_{2-13}$ peptides
253 (H_2N-XGGAIISDFIPP(PEG)K(biotin)-NH_2) where the N-terminal residue, X, constitutes
254 Gly, Asp, Cys or Cys-sulfonic acid were subjected to the arginylation assay in the presence or
255 absence of PCO1/4 (**Figure 4a**). Peptide with an N-terminal Gly did not accept Arg, while an
256 N-terminal Asp did accept Arg, independent of the presence of PCO1 or 4. A peptide
257 comprising an N-terminal Cys-sulfonic acid was also shown to be a substrate for ATE1,
258 again independent of the presence of PCO1 or 4, which is in line with proposed steps of the
259 Arg/Cys N-end rule pathway and has also recently been reported using a similar assay with
260 mouse ATE1.[14-16,35] Crucially, in the absence of PCO1/4, $RAP2_{2-13}$ with an N-terminal Cys

261 was not an acceptor of arginine transfer by ATE1, yet when either PCO1 or PCO4 was
262 incorporated in the reaction, significant ATE1 transferase activity was observed (**Figure 4a**).

263 To confirm that the increased detection of radiolabelled arginine corresponded to arginyl
264 incorporation at the N-termini of the peptides, the experiment was repeated using non-
265 radiolabeled arginine in the presence and absence of PCO4 and ATE1, and peptide products
266 subjected to LC-MS analysis (**Figure 4c**). As with $RAP2_{2-11}$ (**Figure 2c**), the Cys-initiated
267 $RAP2_{2-13}$ peptide displayed a +32 Da increase in mass upon incubation with PCO4 only
268 (**Figure 4c**, red spectrum). Importantly, following incubation of Cys-initiated $RAP2_{2-13}$ with
269 both PCO4 and ATE1, a mass increase equivalent to oxidation coupled to arginylation (+188
270 Da) was observed (**Figure 4c**, blue spectrum). Subsequent tandem MS analysis of these
271 product ions revealed fragmentation species consistent with the assumption that oxidation
272 and sequential arginylation occur at the N-terminus of PCO4- and ATE1-treated peptides
273 (**Figure 4d**, blue spectrum), strongly suggesting that the PCO-oxidized N-termini of ERF-
274 VIIs are rendered N-degrons via additional arginylation (**Figure 4b**).

275 A +12 Da mass increase was observed in products of control assays lacking PCO4 (**Figure**
276 **4c, d**; purple spectra). This appeared to be related to prolonged incubation in the presence of
277 HEPES and DTT as used in the arginylation assay buffer: The +12 Da modification was not
278 observed if the peptide was dissolved in H_2O (**Figure 4c**, black spectrum) or if incubated
279 with HEPES and DTT for just 1 hour, but was observed when the peptide was incubated with
280 HEPES and DTT overnight (**Supplementary Figure 6**). It is proposed that under these
281 conditions, trace levels of contaminating formaldehyde react with free Nt-Cys residues to
282 form thiazolidine N-termini.[36]

283 These results are in line with proposed arginylation requirements for the Arg/Cys branch of
284 the N-end rule pathway[14-16] including the known Cys-initiated arginylation targets from

285 mammals.[32,33,35,37] Importantly, these results demonstrate for the first time Arg transfer
286 mediated by a plant ATE dependent on the N-terminal residue of its substrate, and also that
287 both Cys-sulfinic acid (the product of PCO-catalysis) and Cys-sulfonic acid can act as
288 substrates for ATE1. In particular, the arginylation observed with PCO-catalyzed Cys-sulfinic
289 acid supports the assumption that N-terminal residues sterically and electrostatically
290 resembling Asp or Glu can serve as Arg acceptors in reactions catalyzed by ATEs,[33] and also
291 confirms the importance of the PCOs as a connection between the stability of their ERF-VII
292 substrates and O_2 availability (**Figure 4b**).

293

294 **Discussion**

295

296 The PCOs were identified in *Arabidopsis thaliana* as a set of five enzymes suggested to
297 catalyze oxidation of N-terminal cysteine residues in ERF-VII transcription factors and
298 oxygen consumption was demonstrated for reactions with short peptides corresponding to
299 their N-termini[28]. This putative oxidation was associated with destabilization of the ERF-
300 VIIs, presumably by rendering them substrates of the Arg/Cys branch of the N-end rule
301 pathway.[14,16] Under conditions of sufficient O_2 availability, ERF-VII protein levels are
302 decreased, while under hypoxic conditions, such as those encountered upon plant
303 submergence or in the context of organ development, ERF-VII levels remain high.[17,18]
304 Importantly, the ERF-VII transcription factors are known to upregulate genes which allow
305 plants to cope with or respond to submergence.[13] The PCOs are proposed to act as potential
306 O_2 sensors involved in regulating the plant hypoxic response.[28]

307

308 We sought to biochemically confirm the role of the PCOs in the plant hypoxic response, and
309 present here mass spectrometry and NMR data that clearly demonstrate that two enzymes

310 from different 'subclasses' of this family, PCOs 1 and 4, are dioxygenases that catalyze direct
311 incorporation of O_2 into $RAP2_{2-11}$ peptides to form Cys-sulfinic acid. Their direct use of O_2
312 supports the proposal that these enzymes may act as plant O_2 sensors.[28] A relationship has
313 been demonstrated between O_2 concentration and PCO activity,[28] but it will be of interest to
314 perform detailed kinetic characterization of these enzymes to ascertain their level of
315 sensitivity to O_2 availability, in particular to determine whether their O_2-sensitivity is similar
316 to that of the HIF hydroxylases in animals.[24,25] Although there is functional homology
317 between the PCOs and the HIF hydroxylases, they are apparently mechanistically divergent:
318 The PCOs show sequence homology to the Fe(II)-dependent CDO family of enzymes which
319 do not require an external electron donor for O_2 activation,[28,30] while the HIF hydroxylases
320 are Fe(II)/2OG-dependent oxygenases. They also co-purified with Fe(II) as reported for both
321 the CDOs[29] and PHD2[38]. Of note, the PCOs are the first identified CDOs in plants. Further, in
322 contrast to the reactions of mammalian and bacterial CDOs which oxidize free L-Cys, the
323 PCOs are also, to our knowledge, the first identified cysteinyl (as opposed to free L-Cys)
324 dioxygenases.
325
326 According to the Arg/Cys branch of the N-end rule pathway, N-terminal Cys oxidation is
327 proposed to enable successive arginylation by ATE1 to render proteins as N-degrons. While
328 both Cys-sulfinic and Cys-sulfonic acid are repeatedly reported as potential arginylation
329 substrates[14-16], detailed evidence has only been presented to date for arginylation of Cys-
330 sulfonic acid[32,33] and this only in a mammalian system. We therefore sought to demonstrate
331 that PCO-catalyzed ERF-VII N-terminal Cys oxidation to Cys-sulfinic acid promotes
332 arginylation by ATE1. The arginylation assay and mass spectrometry results we present
333 demonstrate that the PCO-catalyzed dioxygenation reaction is sufficient to trigger N-terminal
334 arginylation of RAP2s by ATE1, thus likely rendering ERF-VIIs (at least those comprising

335 the tested N-terminal sequence) as N-degrons, i.e. recognition by PRT6 and other potential
336 E3 ubiquitin ligases, polyubiquitination and possibly transfer to the 26S proteasome for
337 proteolysis.[14-16] Collectively therefore, we present the comprehensive molecular evidence
338 confirming the Cys-oxidation and subsequent arginylation steps of the Arg/Cys branch of the
339 N-end rule pathway.[32,33,37] We also confirm that ATE1 is able to selectively arginylate, as
340 predicted,[33] acidic N-terminal residues of plant substrates, including Cys-sulfonic acid.
341
342 Arginylation has been known as a posttranslational modification since 1963,[39] to possess a
343 general aminoacyl transferase function in plants (rice and wheat) since 1973 [40] and to have a
344 speculative involvement in the N-end rule pathway since 1988.[41,42] ATE1 is reported as being
345 capable of arginylating proteins at both acidic N-termini and midchain acidic side chains via
346 canonical and non-canonical peptide bonds, respectively.[43] Reports of midchain arginylation
347 highlighted a potentially broad involvement of ATEs in posttranslational protein
348 modifications for various functions[35,43,44,45] but was only very recently brought into question
349 by ^{14}C-Arg incorporation assays using arrays of immobilized synthetic peptides.[35,43,44] To
350 date however only one physiological and two *in vitro* substrates for the Arg/Cys branch of the
351 N-end rule pathway have been characterized, namely mammalian regulator of G protein
352 signaling (RGS) 4, and RGS5 and 10 respectively,[46] where Nt-Cys oxidation was described
353 (to Cys-sulfonic acid) as was Nt-Cys arginylation.[33,37] The first non Cys-branch N-end rule
354 arginylation target was shown to require posttranslational proteolytic cleavage of a (pre-)-
355 proprotein. The C-terminal fragment of proteolytically cleaved mouse BRCA1 is Asp-
356 initiated[47] and gets degraded in an N-end rule-dependent manner. Then, the molecular
357 chaperone BiP (GRP78 and HSPA5, heat shock 70 kDa protein 5) and the oxidoreductase
358 protein disulphide isomerase (PDI), present Glu or Asp after cleavage of their signal peptide,
359 respectively, and were suggested but not shown as putative N-end rule substrates (Hu et al.

360 JBC, 2006). Only very recently, BiP and PDI were identified in mammalian cell culture
361 together with the Glu-initiated calreticulin (CRT) as arginylation targets with a function in
362 autophagy rather than the N-end rule degradation.[48]
363 Similarly, data regarding the molecular requirements of plant ATEs are limited. Already in
364 1973, a general aminoacyl transfer activity was found in rice and wheat cell extracts,
365 however, the nature of enzyme, acceptor position and mechanism remained unclear. It was
366 suggested that the N-terminus could serve as Arg acceptor.[40]
367
368 The first description of a mutant of the single translatable *ATE1* gene in the Arabidopsis
369 accession Wassilewskija (Ws-0) highlighted a role of ATEs in plant development. Ws-0 lacks
370 the second *bona fide* ATE, that is ATE2, due to a single nucleotide polymorphism in *ATE2*
371 causing a premature stop.[49] Developmental functions of the single homolog ATE1 in the
372 moss *Physcomitrella patens* were recently described.[50] Interaction partners of the enzyme
373 were found as well as four arginylated peptides immunologically detected by using antibodies
374 directed against peptides mimicking N-terminal Arg-Asp or Arg-Glu.[51] In one case, that is
375 the acylamino-acid-releasing enzyme PpAARE, which presents for unknown reasons a neo-
376 N-terminal Asp residue which was formerly Asp2 and therefore initiated by Met, an N-
377 terminal arginylation was found with high confidence.Previously, Arg transferase function of
378 Arabidopsis ATE1/2 has been shown using an assay detecting conjugation of ^{3}H-Arg to
379 bovine α-lactalbumin (bearing an N-terminal Glu) in the presence of plant extracts from wild
380 type Arabidopsis, and *ate1* and *ate2* single mutants but not from *ate1 ate2* double mutant
381 seedlings.[34] Therefore, the results we present here demonstrate for the first time Arg
382 transferase activity of a plant ATE towards known plant N-end rule substrates.

383 Interestingly, in combination with O_2, nitric oxide was identified as an RGS oxidizing agent,
384 suggesting a potential role of S-nitrosylation in the Arg/Cys branch of the N-end rule
385 pathway, albeit non-enzymatically controlled.[32] It has also been reported *in planta* that both
386 NO and O_2 are required for ERF-VII degradation, potentially at the Cys oxidation step.[22,23]
387 Although in N-end rule-mediated RGS4/5 degradation it has been proposed that Cys-
388 nitrosylation precedes Cys-oxidation (also currently considered a non-enzymatic process), we
389 find that under the conditions used, the PCO1/4-catalyzed reaction does not require either
390 prior cys-nitrosylation or exogenous NO to proceed efficiently. We cannot rule out that NO
391 plays a role in formation of a Cys-sulfonic acid product, which is also a substrate for ATE1 as
392 shown in our Arg transfer experiments. Alternatively, NO may have a role in ERF-VII
393 degradation *in vivo* via non-enzymatic oxidation or via a secondary mechanism. The manner
394 in which NO contributes to Arg/Cys branch of the N-end rule pathway therefore remains to
395 be elucidated.
396
397 ERF-VII stabilization has been shown to result in improved submergence tolerance, elegantly
398 demonstrated in barley by mutation of the candidate E3-ubiquitin ligase *PRT6,*[11] but also in
399 rice containing the *Sub1A* gene; SUB1A is an apparently stable ERF-VII that confers
400 particular flood tolerance in certain rare varieties of rice.[9,17] Overexpression of *Sub1A* in more
401 commonly grown rice varieties has resulted in a 45% increase in yield relative to *sub1a*
402 mutant lines after exposure to flooding.[52] If ERF-VII stabilization is indeed a proficient
403 mechanism for enhancing flood tolerance, then manipulation of PCO or ATE activity may be
404 an efficient and effective point of intervention. This work presents molecular validation of
405 their function, providing the basis for future targeted chemical/genetic inhibition of their
406 activity. It also highlights genetic strategies for breeding via introgression of variants of N-
407 end rule pathway components or introduction of alleles of enzymatic components of the N-

end rule pathway from non-crop species into crops. . Any of these strategies has the potential to result in stabilized ERF-VII levels and therefore increase stress resistance and may therefore help to address food security challenges.

411 **Online Methods**

412

413 **Peptide Synthesis, Purification and Characterization**

414 All reagents used were purchased from Sigma-Aldrich unless otherwise stated. The 10-mer
415 $RAP2_{2-11}$ peptide (H_2N-CGGAIISDFI-COOH) was purchased from GL Biochem (Shanghai)
416 Ltd, China (**Supplementary Table 1**). The sequence of the 12-mer peptides used in the
417 coupled oxidation-arginylation assay is derived from RAP2.2 and RAP2.12 (H_2N-X-
418 GGAIISDFIPP(PEG)K(biotin)-NH_2) and synthesized by Fmoc-based solid-phase peptide
419 synthesis (SPPS) on NovaSyn®TGR resin (Merck KGaA, **Supplementary Table 2**). Fmoc
420 protected amino acids (Iris Biotech GmbH) were coupled using 4 equivalents of (eq) of the
421 amino acid according to the initial loading of the resin. 4 eq amino acid was mixed with 4 eq
422 O-(6-chlorobenzotriazol-1-yl)-N,N,N′,N′-tetramethyluronium hexafluorophosphate (HCTU)
423 and 8 eq N,N-diisopropylethylamine (DIPEA; Santa Cruz Biotechnology, sc-293894) and
424 added to the resin for 1 h. In a second coupling, the resin was treated with 4 eq of the Fmoc-
425 protected amino acid mixed with 4 eq benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium
426 hexafluoro-phosphate (PyBOP) and 8 eq 4-methylmorpholine (NMM) for 1 h. After double
427 coupling a capping step to block free amines was performed using acetanhydride and DIPEA
428 in N-methyl-2-pyrrolidinone (NMP) (1:1:10) for 5 min. The *C*-terminal Fmoc-Lys(biotin)-
429 OH, the 8-(9-fluorenylmethyloxycarbonyl-amino)-3.6-dioxaoctanoic acid (PEG) linker and
430 the different Fmoc protected N-terminal amino acids were coupled manually. The remaining
431 peptide sequence was assembled using an automated synthesizer (Syro II, MultiSynTech
432 GmbH). Fmoc deprotection was performed using 20 % piperidine in dimethylformamide
433 (DMF) for 5 min, twice. After each step the resin was washed 5 times with DMF, DCM and
434 DMF, respectively. Final cleavage was performed with 94 % trifluoroacetic acid (TFA),
435 2.5 % 1,2-ethanedithiole (EDT) and 1 % triisopropylsilane (TIPS) in aqueous solution for

436 2 h, twice. The cleavage solutions were combined and peptides were precipitated with diethyl
437 ether (Et_2O) at -20°C for 30 min. Peptides were solved in water/acetonitrile (ACN) 7:3 and
438 purified by reversed-phase high performance liquid chromatography (HPLC; Nucleodur C18
439 culumn; 10×125 mm, 110 Å, 5 µm particle size; Macherey-Nagel) using a flow rate of
440 6 $ml \cdot min^{-1}$ (A: ACN with 1 % TFA, B: water with 1 % TFA). Obtained pure fractions were
441 pooled and lyophilized. Peptide characterization was performed by analytical HPLC (1260
442 Infinity, Agilent Technology; flow rate of 1 $ml \cdot min^{-1}$, A: ACN with 1 % TFA, B: water with
443 1 % TFA) coupled with a mass spectrometer (6120 Quadrupole LC/MS, Agilent Technology)
444 using electro spray ionization (Agilent Eclipse XDB-C18 culumn, 4.6×150 mm, 5 µm
445 particle size). Analytical HPLC chromatograms were recorded at 210 nm (**Supplementary**
446 **Figure 7)**. Quantification was performed by HPLC-based comparison (chromatogram at 210
447 nm) with a reference peptide (**Supplementary Table 2**).

448

449 **Preparation of Recombinant Proteins**

450 Arabidopsis PCO1 and PCO4 sequences in pDEST17 bacterial expression vectors
451 (Invitrogen) were kindly provided by F. Licausi and J. van Dongen.[28] Plasmids were
452 transformed into BL21(DE3) *Escherichia coli* cells, and expression of recombinant protein
453 carrying an N-terminal hexahistidine tag was induced with 0.5 mM IPTG and subsequent
454 growth at 18°C for 18 hours. Harvested cells were lysed by sonication and proteins purified
455 using Ni^{++} affinity chromatography, before buffer exchange into 250 mM NaCl/50 mM Tris-
456 HCl (pH 7.5). Analysis by SDS-PAGE and denaturing liquid-chromatography mass
457 spectrometry (LC-MS) showed proteins with more than 90% purity and with the predicted
458 molecular weights.

459 The coding sequence of Arabidopsis ATE1 was cloned according to gene annotations at
460 TAIR (www.arabidopsis.org) from cDNA. The sequence was flanked by an N-terminal

461 tobacco etch virus (TEV) recognition sequence for facilitated downstream purification ("tev":
462 ENLYFQ-X) using the primers ate1_tev_ss (5'-
463 GCTTAGAGAATCTTTATTTTCAGGGGATGTCTTTGAAAAACGATGCGAGT-3') and
464 ate1_as (5'-
465 GGGGACCACTTTGTACAAGAAAGCTGGGTATCAGTTGATTTCATACACCATTCTC
466 TC-3'). A second PCR using the primers adapter (5'-
467 GGGGACAAGTTTGTACAAAAAAGCAGGCTTAGAGAATCTTTATTTTCAGGGG-3')
468 and ate1_as was performed to amplify the construct to use it in a BP reaction for cloning into
469 pDONR201 (Invitrogen) followed by an LR reaction into the vector pDEST17 (Invitrogen).
470 The N-terminal hexahistidine fusion was expressed in BL21-CodonPlus (DE3)-RIL
471 *Escherichia coli* (*E. coli*) cells. The expression culture was induced with 1 mM IPTG at
472 OD=0.6 and grown for 16 hours at 18°C. After resuspension in LEW buffer (50 mM
473 NaH_2PO_4, pH 8; 300 mM NaCl; 1 mM DTT), the cells were lysed by incubation with 1.2
474 mg/ml lysozyme for 30 min and underwent subsequent sonification in the presence of 1 mM
475 PMSF. Recombinant protein was purified by Ni^{++} affinity chromatography and subjected to
476 Amicon Ultra-15 (30K) (Merck Millipore) filtration for buffer exchange to imidazole-free
477 LEW containing 20% glycerol.

478

479 **PCO Activity Assays and MALDI Analysis**

480 PCO activity assays were conducted under the following conditions unless otherwise stated: 1
481 µM PCO1 or 4 was mixed with 100 or 200 µM $RAP2_{2-11}$ peptide in 250 mM NaCl, 1 mM
482 dithiothreitol (DTT), 50 mM Tris-HCl pH 7.5 and incubated at 30°C for 30-60 minutes.
483 Addition of exogenous Fe(II) and/or ascorbate were not required for activity. Assays were
484 stopped by quenching 1 µL sample with 1 µL alpha-cyano-4-hydroxycinnamic acid (CHCA)
485 matrix on a MALDI plate prior to product mass analysis using a Sciex 4800 TOF/TOF mass

486 spectrometer (Applied Biosystems) operated in negative ion reflectron mode. The instrument
487 parameters and data acquisition were controlled by 4000 Series Explorer software and data
488 processing was completed using Data Explorer (Applied Biosystems).
489
490 To test the activity of PCO4 in the presence of $^{18}O_2$, 100 µL of an anaerobic solution of 100
491 µM $RAP2_{2-11}$ in 250 mM NaCl/50 mM Tris-HCl pH 7.5 was prepared in a septum-sealed
492 glass vial by purging with 100% N_2 for 10 minutes at 100 mL/min using a mass flow
493 controller (Brooks Instruments), as used for previous preparation of anaerobic samples to
494 determine enzyme dependence on O_2.[24] PCO4 was then added using a gas-tight Hamilton
495 syringe, followed by purging with a balloon (approx. 0.7 L) of $^{16}O_2$ or $^{18}O_2$ over the course of
496 10 minutes at room temperature. Reaction vials were then transferred to 30°C for a further 20
497 minutes before products were analysed by MALDI-MS as described above.
498
499 PCO4 activity was additionally tested in the presence of $H_2{}^{18}O$ by conducting an assay in
500 75% $H_2{}^{18}O$, 25% H_2O (with all enzyme/substrate/buffer components comprising a portion of
501 the H_2O fraction). Assays were conducted for 10 minutes at room temperature followed by 20
502 minutes at 30°C for comparison with assays conducted with $^{18}O_2$. Products were analysed by
503 MALDI-MS, as described above.
504
505 **UPLC-MS and MS/MS Analysis of PCO Assay Products**
506 Ultra-high performance chromatography (UPLC) mass spectrometry (MS) measurements
507 were obtained using an Acquity UPLC system coupled to a Xevo G2-S Q-ToF mass
508 spectrometer (Waters) operated in positive electrospray mode. Instrument parameters, data
509 acquisition and data processing were controlled by Masslynx 4.1. Source conditions were
510 adjusted to maximize sensitivity and minimize fragmentation while Lockspray was employed

511 during analysis to maintain mass accuracy. 2 µL of each sample was injected on to a
512 Chromolith Performance RP-18e 100-2 mm column (Merck) heated to 40 ^{o}C and eluted using
513 a gradient of 95 % deionized water supplemented with 0.1 % (v/v) formic acid (analytical
514 grade) to 95 % acetonitrile (HPLC grade) and a flow rate of 0.3 mL/min. Fragmentation
515 spectra of substrate and product peptide ions (MS/MS) were obtained using a targeted
516 approach with a typical collision-induced dissociation (CID) energy ramp of 30 to 40 eV.
517 Analysis was carried out with the same source settings, flow rate and column elution
518 conditions as above.
519

520 **^{1}H NMR Assay**

521 Reaction components (5 µM PCO1 or PCO4 and 500 µM RAP2$_{2-11}$) were prepared to 75 µL
522 in 156 mM NaCl, 31 mM Tris-HCl (pH 7.5) and 10% D_2O (enzyme added last), in a 1.5 mL
523 microcentrifuge tube before being transferred to a 2 mm diameter NMR tube. ^{1}H NMR
524 spectra at 310 K were recorded using a Bruker AVIII 600 (with inverse cryoprobe optimized
525 for ^{1}H observation and running topspin 2 software; Bruker) and reported in p.p.m. relative to
526 D_2O (δ_H 4.72). The deuterium signal was also used as internal lock signal and the solvent
527 signal was suppressed by presaturating its resonance.
528
529

530 **Arginylation Assay**

531 The conditions for arginylation of the 12-mer peptide substrates were modified from [43]. In
532 detail, ATE1 was incubated at 10 µM in the reaction mixture containing 50 mM HEPES, pH
533 7.5; 25 mM KCl; 15 mM $MgCl_2$; 1 mM DTT; 2.5 mM ATP; 0.6 mg/ml *E. coli* tRNA (R1753,
534 Sigma); 0.04 mg/ml *E. coli* aminoacyl-tRNA synthetase (A3646, Sigma); 80 µM (4 nCi/µl)
535 ^{14}C-arginine (MC1243, Hartmann Analytic); 50 µM C-terminally biotinylated 12-mer peptide

536 substrate and, where indicated, 1 µM purified recombinant PCO1 or PCO4 in a total reaction
537 volume of 50 µL. The reaction was conducted at 30°C for 16 to 40 hours. After incubation,
538 each 50 µL of avidin agarose bead slurry (20219, Pierce) equilibrated in PBSN (PBS-
539 Nonidet; 100 mM NaH_2PO_4; 150 mM NaCl; 0.1% Nonidet-P40) was added to the samples
540 and mixed with an additional 350 µL of PBSN. After 2 hours of rotation at room temperature,
541 the beads were washed 4 times in PBSN, resuspended in 4 mL of FilterSafe scintillation
542 solution (Zinsser Analytic) and scintillation counting was performed using a Beckmann
543 Coulter LS 6500 Multi-Purpose scintillation counter.

Acknowledgements

Petra Majovsky, Domenika Thieme and Wolfgang Hoehenwarter from the Proteomics Unit of the Leibniz Institute of Plant Biochemistry (IPB), Halle, are acknowledged for mass spectrometry of recombinant ATE1. David Staunton from the Biophysical Facility, Department of Biochemistry, University of Oxford is acknowledged for multi-angle light scatter analysis of PCO1/4. Geoff Grime from the University of Surrey Ion Beam Centre is acknowledged for assistance with the microPIXE data collection. This work was supported by a Biotechnology and Biological Sciences Research Council (U.K.) New Investigator grant (BB/M024458/1) to E.F., a grant for setting up the junior research group of the *ScienceCampus Halle – Plant-based Bioeconomy* to N.D., by a Ph.D. fellowship of the Landesgraduiertenförderung Sachsen-Anhalt awarded to C.N., by an Engineering and Physical Sciences Research Council (U.K.) studentship (EP/G03706X/1) to J.B-B., a Royal Society Dorothy Hodgkin Fellowship to E.F., a William R. Miller Junior Research Fellowship (St. Edmund Hall, Oxford) to R.H. and grant DI 1794/3-1 by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) to N.D. Financial support came from the Leibniz Association, the state of Saxony Anhalt, the Deutsche Forschungsgemeinschaft (DFG) Graduate Training Center GRK1026 "*Conformational Transitions in Macromolecular Interactions*" at Halle, and the Leibniz Institute of Plant Biochemistry (IPB) at Halle, Germany. We thank Prof. J. van Dongen (RWTH Aachen University, Germany) and Prof. F. Licausi (Scuolo Superiore Sant'Anna, Pisa, Italy) for sharing pDEST-PCO plasmids and helpful discussions.

566 **Author Contributions**

567 M.W. performed the PCO1/4 activity assays and MALDI/LC-/MS/MS analyses. M.K.
568 performed and established arginylation reactions on peptides coupled to biotin pulldown and
569 scintillation measurements and purified ATE1 protein. R.H. performed the NMR assays with
570 E.F. D.W. prepared the pDEST17-PCO1 and 4 plasmids. C.M. synthesized the peptides,
571 T.N.G. supervised and designed the synthesis, C.N. cloned and established purification and
572 activity assays for ATE1. R.O. conducted LC-MS to analyse +12 Da mass shifts. J.W.
573 performed LC-MS analysis. J.Y. and J.C.B-B. prepared samples for micro-PIXE analysis and
574 J.C.B-B. and E.F.G. collected and analysed micro-PIXE data. This work was supported by
575 the network of the European Cooperation in Science and Technology (COST) Action
576 BM1307—" European network to integrate research on intracellular proteolysis pathways in
577 health and disease (PROTEOSTASIS). E.F. performed the PCO1 and PCO4 protein
578 purification and selected activity assays. E.F., M.W., M.K. and N.D. designed the study, E.F.
579 and N.D. wrote the manuscript. M.W., M.K., N.D. and E.F. designed the figures. All authors
580 read and approved the final version of this manuscript.

581

582 **References**

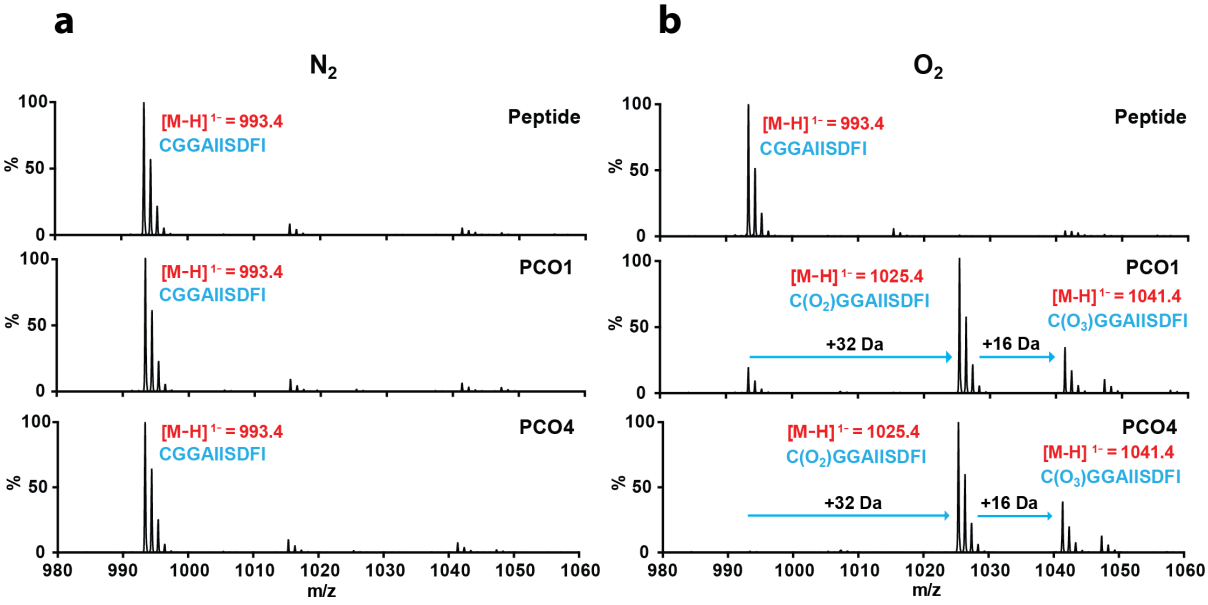
583 1. Kaelin, W.G., Jr. & Ratcliffe, P.J. Oxygen sensing by metazoans: the central role of the HIF
584 hydroxylase pathway. *Mol Cell* **30**, 393-402 (2008).
585 2. Myllyharju, J. Prolyl 4-hydroxylases, master regulators of the hypoxia response. *Acta Physiol*
586 *(Oxf)* **208**, 148-65 (2013).
587 3. Semenza, G.L. Oxygen sensing, homeostasis, and disease. *N Engl J Med* **365**, 537-47 (2011).
588 4. Bailey-Serres, J. et al. Making sense of low oxygen sensing. *Trends Plant Sci* **17**, 129-38
589 (2012).
590 5. Bailey-Serres, J., Lee, S.C. & Brinton, E. Waterproofing crops: effective flooding survival
591 strategies. *Plant Physiol* **160**, 1698-709 (2012).
592 6. Hattori, Y. et al. The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt
593 to deep water. *Nature* **460**, 1026-U116 (2009).
594 7. Hinz, M. et al. Arabidopsis RAP2.2: an ethylene response transcription factor that is
595 important for hypoxia survival. *Plant Physiol* **153**, 757-72 (2010).
596 8. Licausi, F. et al. HRE1 and HRE2, two hypoxia-inducible ethylene response factors, affect
597 anaerobic responses in Arabidopsis thaliana. *Plant J* **62**, 302-15 (2010).
598 9. Xu, K. et al. Sub1A is an ethylene-response-factor-like gene that confers submergence
599 tolerance to rice. *Nature* **442**, 705-8 (2006).
600 10. Papdi, C. et al. Functional identification of Arabidopsis stress regulatory genes using the
601 controlled cDNA overexpression system. *Plant Physiol* **147**, 528-42 (2008).
602 11. Mendiondo, G.M. et al. Enhanced waterlogging tolerance in barley by manipulation of
603 expression of the N-end rule pathway E3 ligase PROTEOLYSIS6. *Plant Biotechnol J* **14**, 40-50
604 (2016).
605 12. Lee, S.C. et al. Molecular characterization of the submergence response of the Arabidopsis
606 thaliana ecotype Columbia. *New Phytol* **190**, 457-71 (2011).
607 13. Mustroph, A. et al. Profiling translatomes of discrete cell populations resolves altered
608 cellular priorities during hypoxia in Arabidopsis. *Proc Natl Acad Sci U S A* **106**, 18843-8
609 (2009).
610 14. Varshavsky, A. The N-end rule pathway and regulation by proteolysis. *Protein Sci* **20**, 1298-
611 1345 (2011).
612 15. Tasaki, T., Sriram, S.M., Park, K.S. & Kwon, Y.T. The N-end rule pathway. *Annu Rev Biochem*
613 **81**, 261-89 (2012).
614 16. Gibbs, D.J., Bacardit, J., Bachmair, A. & Holdsworth, M.J. The eukaryotic N-end rule pathway:
615 conserved mechanisms and diverse functions. *Trends Cell Biol* **24**, 603-611 (2014).
616 17. Gibbs, D.J. et al. Homeostatic response to hypoxia is regulated by the N-end rule pathway in
617 plants. *Nature* **479**, 415-8 (2011).
618 18. Licausi, F. et al. Oxygen sensing in plants is mediated by an N-end rule pathway for protein
619 destabilization. *Nature* **479**, 419-22 (2011).
620 19. Ross, S., Giglione, C., Pierre, M., Espagne, C. & Meinnel, T. Functional and developmental
621 impact of cytosolic protein N-terminal methionine excision in Arabidopsis. *Plant Physiol* **137**,
622 623-37 (2005).
623 20. Giglione, C., Boularot, A. & Meinnel, T. Protein N-terminal methionine excision. *Cell Mol Life*
624 *Sci* **61**, 1455-74 (2004).
625 21. Garzon, M. et al. PRT6/At5g02310 encodes an Arabidopsis ubiquitin ligase of the N-end rule
626 pathway with arginine specificity and is not the CER3 locus. *FEBS Lett* **581**, 3189-96 (2007).
627 22. Gibbs, D.J. et al. Nitric oxide sensing in plants is mediated by proteolytic control of group VII
628 ERF transcription factors. *Mol Cell* **53**, 369-79 (2014).
629 23. Gibbs, D.J. et al. Group VII Ethylene Response Factors Coordinate Oxygen and Nitric Oxide
630 Signal Transduction and Stress Responses in Plants. *Plant Physiol* **169**, 23-31 (2015).

24. Tarhonskaya, H. et al. Investigating the contribution of the active site environment to the slow reaction of hypoxia-inducible factor prolyl hydroxylase domain 2 with oxygen. *Biochem J* (2014).
25. Hirsila, M., Koivunen, P., Gunzler, V., Kivirikko, K.I. & Myllyharju, J. Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor. *J Biol Chem* **278**, 30772-80 (2003).
26. Bruick, R.K. Oxygen sensing in the hypoxic response pathway: regulation of the hypoxia-inducible transcription factor. *Genes Dev* **17**, 2614-23 (2003).
27. Epstein, A.C. et al. C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell* **107**, 43-54 (2001).
28. Weits, D.A. et al. Plant cysteine oxidases control the oxygen-dependent branch of the N-end-rule pathway. *Nat Commun* **5**, 3425 (2014).
29. Imsand, E.M., Njeri, C.W. & Ellis, H.R. Addition of an external electron donor to in vitro assays of cysteine dioxygenase precludes the need for exogenous iron. *Arch Biochem Biophys* **521**, 10-7 (2012).
30. Joseph, C.A. & Maroney, M.J. Cysteine dioxygenase: structure and mechanism. *Chem Commun (Camb)*, 3338-49 (2007).
31. Li, W. & Pierce, B.S. Steady-state substrate specificity and O2-coupling efficiency of mouse cysteine dioxygenase. *Arch Biochem Biophys* **565**, 49-56 (2015).
32. Hu, R.G. et al. The N-end rule pathway as a nitric oxide sensor controlling the levels of multiple regulators. *Nature* **437**, 981-6 (2005).
33. Kwon, Y.T. et al. An essential role of N-terminal arginylation in cardiovascular development. *Science* **297**, 96-9 (2002).
34. Graciet, E. et al. The N-end rule pathway controls multiple functions during Arabidopsis shoot and leaf development. *Proc Natl Acad Sci U S A* **106**, 13618-23 (2009).
35. Wadas, B., Piatkov, K.I., Brower, C.S. & Varshavsky, A. Analyzing N-terminal Arginylation through the Use of Peptide Arrays and Degradation Assays. *J Biol Chem* **291**, 20976-20992 (2016).
36. Kallen, R.G. The mechanism of reactions involving Schiff base intermediates. Thiazolidine formation from L-cysteine and formaldehyde. *J Am Chem Soc* **93**, 6236-48 (1971).
37. Davydov, I.V. & Varshavsky, A. RGS4 is arginylated and degraded by the N-end rule pathway in vitro. *J Biol Chem* **275**, 22931-41 (2000).
38. McNeill, L.A. et al. Hypoxia-inducible factor prolyl hydroxylase 2 has a high affinity for ferrous iron and 2-oxoglutarate. *Mol Biosyst* **1**, 321-4 (2005).
39. Kaji, H., Novelli, G.D. & Kaji, A. A Soluble Amino Acid-Incorporating System from Rat Liver. *Biochim Biophys Acta* **76**, 474-7 (1963).
40. Manahan, C.O. & App, A.A. An arginyl-transfer ribonucleic Acid protein transferase from cereal embryos. *Plant Physiol* **52**, 13-6 (1973).
41. Ciechanover, A. et al. Purification and characterization of arginyl-tRNA-protein transferase from rabbit reticulocytes. Its involvement in post-translational modification and degradation of acidic NH2 termini substrates of the ubiquitin pathway. *J Biol Chem* **263**, 11155-67 (1988).
42. Bohley, P., Kopitz, J. & Adam, G. Surface hydrophobicity, arginylation and degradation of cytosol proteins from rat hepatocytes. *Biol Chem Hoppe Seyler* **369 Suppl**, 307-10 (1988).
43. Wang, J. et al. Arginyltransferase ATE1 catalyzes midchain arginylation of proteins at side chain carboxylates in vivo. *Chem Biol* **21**, 331-7 (2014).
44. Eriste, E. et al. A novel form of neurotensin post-translationally modified by arginylation. *J Biol Chem* **280**, 35089-97 (2005).
45. Wong, C.C. et al. Global analysis of posttranslational protein arginylation. *PLoS Biol* **5**, e258 (2007).
46. Lee, M.J. et al. RGS4 and RGS5 are in vivo substrates of the N-end rule pathway. *Proc Natl Acad Sci U S A* **102**, 15030-5 (2005).

682 47. Piatkov, K.I., Brower, C.S. & Varshavsky, A. The N-end rule pathway counteracts cell death by
683 destroying proapoptotic protein fragments. *Proc Natl Acad Sci U S A* **109**, E1839-47 (2012).
684 48. Cha-Molstad, H. et al. Amino-terminal arginylation targets endoplasmic reticulum chaperone
685 BiP for autophagy through p62 binding. *Nat Cell Biol* **17**, 917-29 (2015).
686 49. Yoshida, S., Ito, M., Callis, J., Nishida, I. & Watanabe, A. A delayed leaf senescence mutant is
687 defective in arginyl-tRNA:protein arginyltransferase, a component of the N-end rule pathway
688 in Arabidopsis. *Plant J* **32**, 129-37 (2002).
689 50. Schuessele, C. et al. Spatio-temporal patterning of arginyl-tRNA protein transferase (ATE)
690 contributes to gametophytic development in a moss. *New Phytol* **209**, 1014-27 (2016).
691 51. Hoernstein, S.N. et al. Identification of Targets and Interaction Partners of Arginyl-tRNA
692 Protein Transferase in the Moss Physcomitrella patens. *Mol Cell Proteomics* **15**, 1808-22
693 (2016).
694 52. Dar, M.H., de Janvry, A., Emerick, K., Raitzer, D. & Sadoulet, E. Flood-tolerant rice reduces
695 yield variability and raises expected yield, differentially benefitting socially disadvantaged
696 groups. *Sci Rep* **3**, 3315 (2013).

697 **Figures**

698



699

700

701 **Figure 1. PCO1, PCO4 and O_2-dependent modification of a RAP2$_{2\text{-}11}$ peptide substrate,**

702 **consistent with Cys-oxidation.** MALDI-MS spectra showing products following PCO1 and

703 PCO4 incubation with RAP2$_{2\text{-}11}$ under anaerobic (**a**) or aerobic (**b**) conditions. Products with

704 mass increases of +32 Da and +48 Da were only observed in the presence of PCO1 or PCO4

705 and O_2.

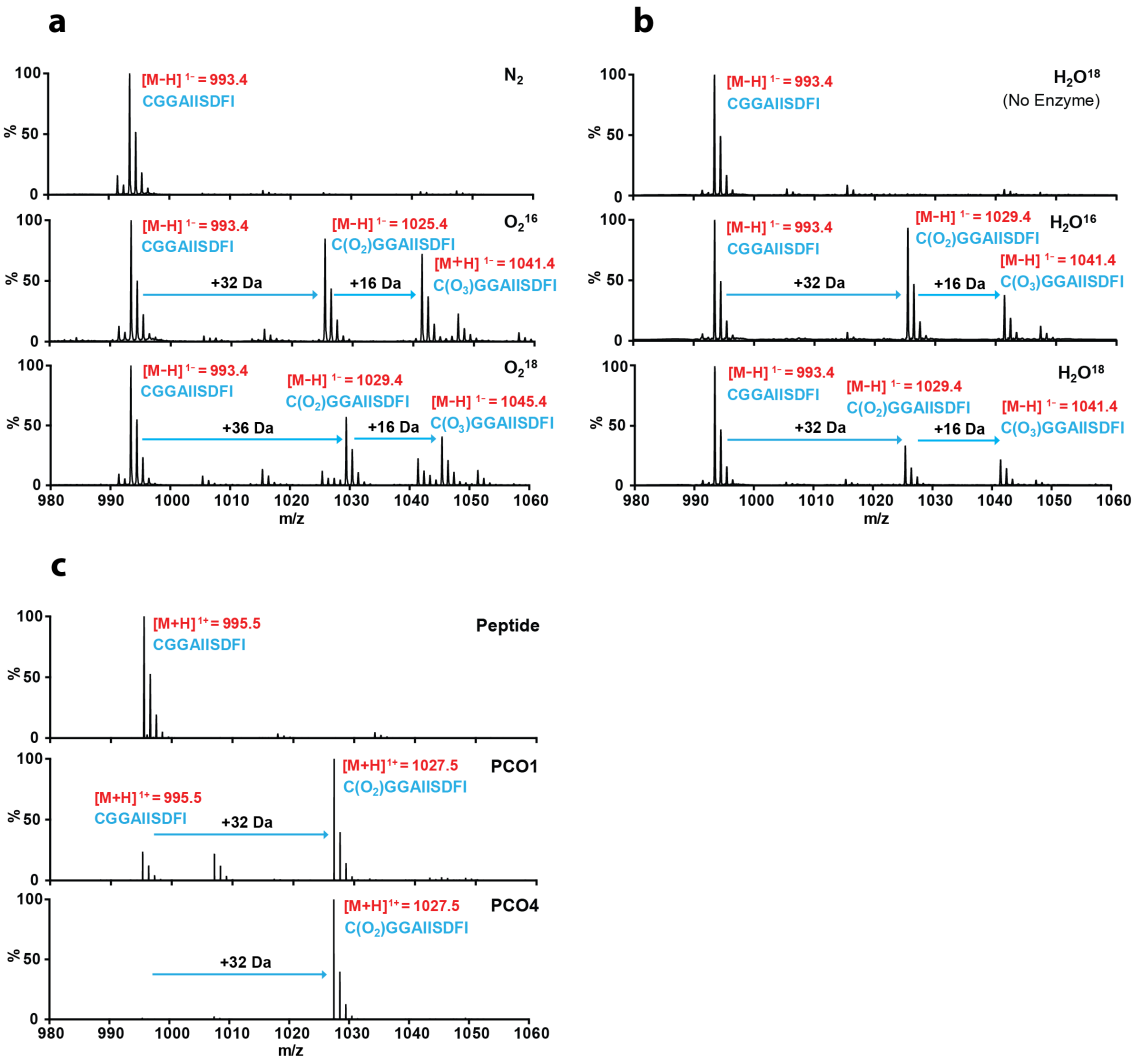


Figure 2. The PCOs are dioxygenases which catalyze incorporation of molecular O_2 into $RAP2_{2-11}$. (a) MALDI-MS spectra showing that PCO4-catalyzed reactions carried out in the presence of $^{18}O_2$ result in a +4 Da increase in the mass of the putative Cys-sulfinic acid product, however a +6 Da increase in the size of the putative Cys-sulfonic acid product is not observed; **(b)** MALDI-MS spectra showing that PCO4-catalyzed reactions carried out in the presence of $H_2{}^{18}O$ show no additional incorporation of mass compared with products of reactions in the presence of $H_2{}^{16}O$; **(c)** LC-MS spectra confirm that the +48 Da reaction product is an artifact of MALDI-MS analysis (**Supplementary Figure 3**) and incubation of PCO1 and PCO4 with $RAP2_{2-11}$ results in formation of a single product with a mass increase of +32 Da, consistent with Cys-sulfinic acid formation.

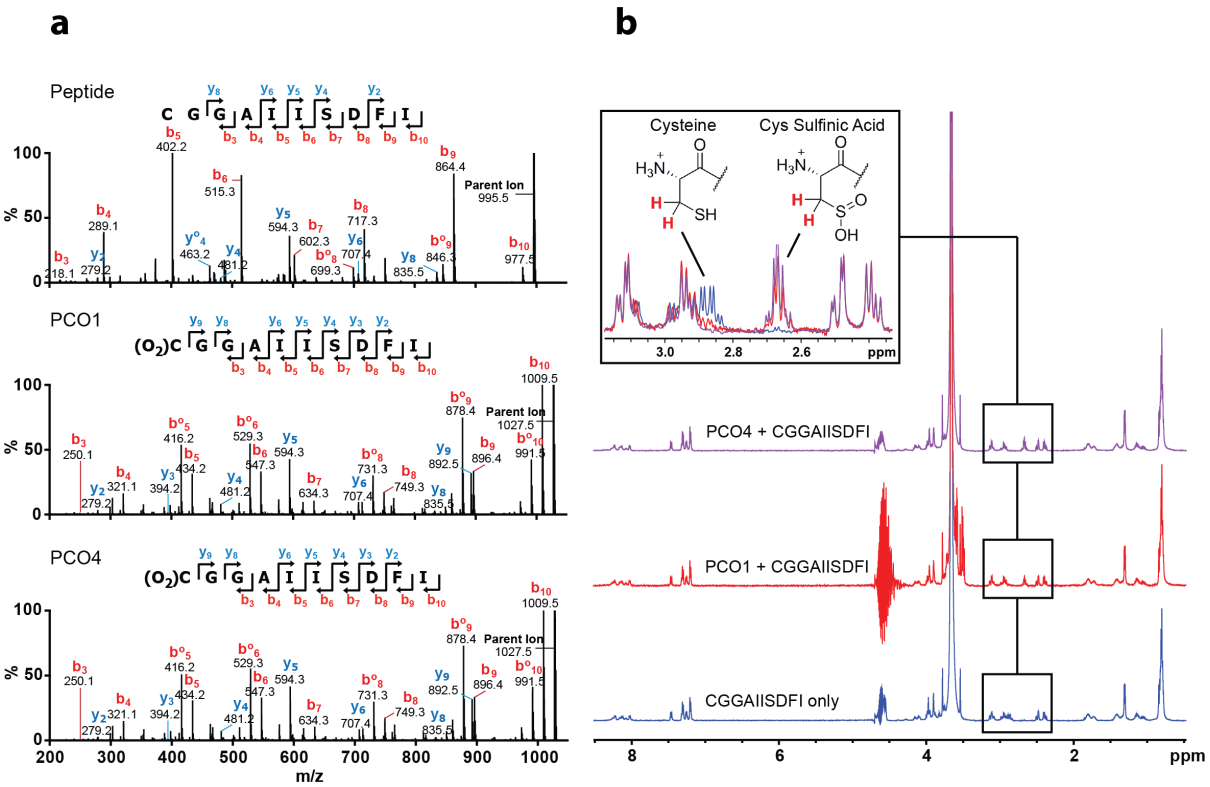


Figure 3. PCO1 and PCO4 oxidize the N-terminal Cys of RAP2$_{2\text{-}11}$ to Cys-sulfinic acid as confirmed by (a) LC-MS/MS and (b) ^{1}H-NMR. (**a**) Peptidic products of PCO-catalyzed reactions were subjected to LC-MS/MS analysis. In the presence of enzyme, fragment assignment was consistent with expected b- and y-series ion masses for RAP2$_{2\text{-}11}$ with N-terminal Cys-sulfinic acid. (**b**) ^{1}H-NMR was used to monitor changes to RAP2$_{2\text{-}11}$ (500 μM) upon incubation with enzyme (5 μM). In the presence of PCO1 (red) and PCO4 (purple), the ^{1}H-resonance at δ_H 2.88 ppm (assigned to the β-cysteinyl protons of RAP2$_{2\text{-}11}$, blue) was observed to decrease in intensity, with concomitant emergence of a resonance at δ_H 2.67 ppm. This new resonance was assigned to the β-protons of Cys-sulfinic based on chemical shift analysis (see **Supplementary Figure 4**).

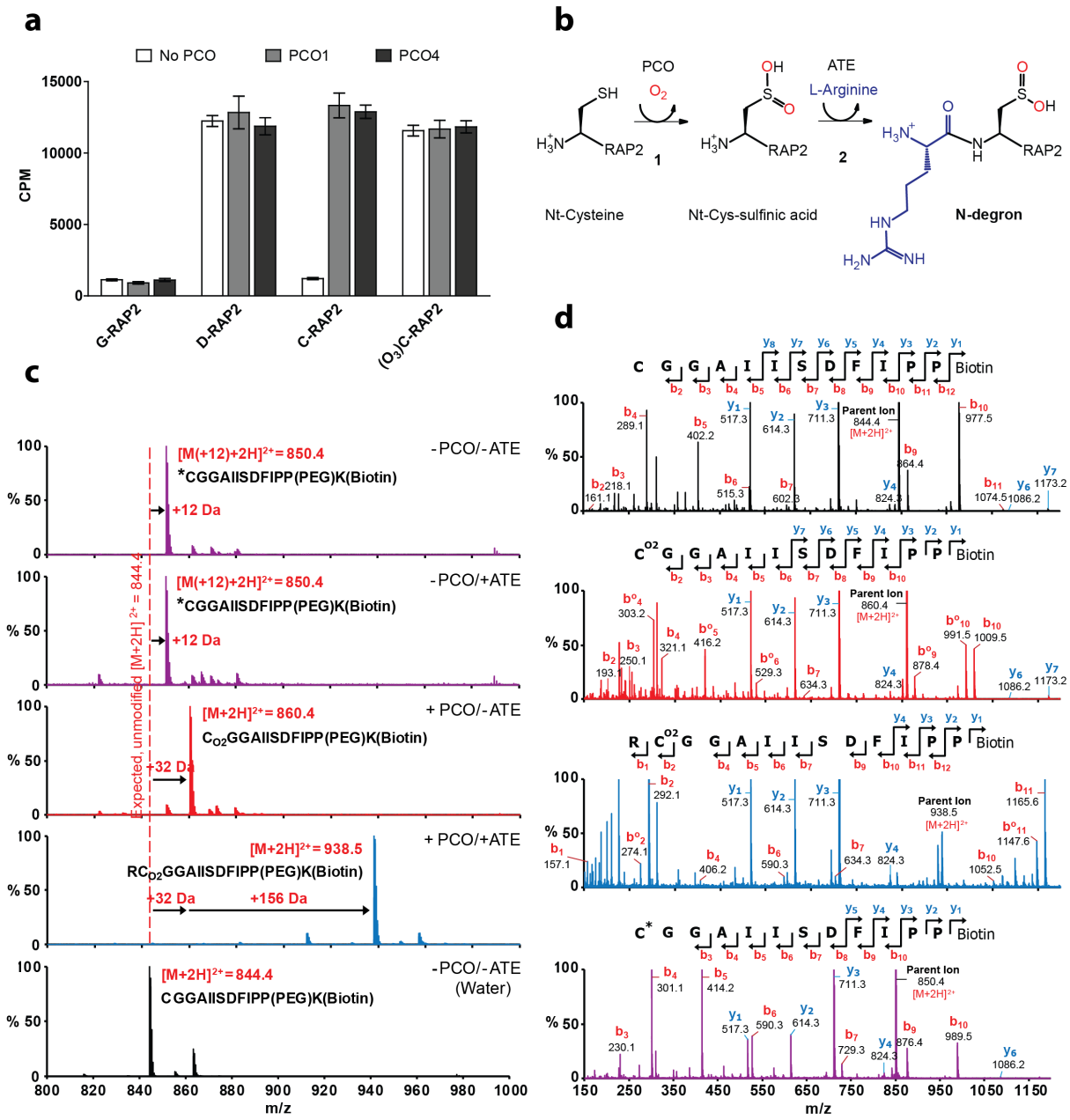


Figure 4. PCO-Catalyzed Cys-sulfinic acid formation renders RAP2$_{2\text{-}13}$ a substrate for ATE1 catalyzed arginylation. (a) ^{14}C-Arg incorporation by ATE1 into the 12-mer N-terminal RAP2 peptide (H_2N-XGGAIISDFIPP(PEG)K(biotin)-NH_2, X = Gly, Asp, Cys or Cys-sulfonic acid (C(O_3))), was assayed by liquid scintillation counting of immobilized biotinylated peptides after the arginylation reaction and removal of unreacted ^{14}C-Arg (n=3). In the case of the Cys-starting peptide (RAP2$_{2\text{-}13}$), ATE1 activity was strongly dependent on the presence of PCO1 or PCO4. **(b)** Scheme showing PCO- and ATE1-catalyzed reactions on

738 Nt-Cys RAP2.2, as validated in this study. **(c)** LC-MS spectra of products of equivalent
739 assays with Cys-initiated $RAP2_{2-13}$ using non-radiolabelled Arg, revealing a sequential mass
740 increase of +32 (corresponding to oxidation) and +156 Da (corresponding to arginylation)
741 only in the presence of PCO and ATE1 (blue spectrum). The red spectrum shows a +32 Da
742 mass increase for Cys-$RAP2_{2-13}$ incubated +PCO/-ATE, demonstrating Cys-sulfonic acid
743 formation as expected. Purple spectra show +12 Da products formed upon incubation of Cys-
744 $RAP2_{2-13}$ in the absence of PCO +/-ATE (for explanation of this mass increase see text and
745 **Supplementary Figure 6**); the black spectrum shows Cys-$RAP2_{2-13}$ dissolved in H_2O. **(d)**
746 b- and y-ion series spectra generated by MS/MS analysis of Cys-$RAP2_{2-13}$ only (no
747 incubation; black), Cys-$RAP2_{2-13}$ incubated +PCO/-ATE (red), Cys-$RAP2_{2-13}$ incubated with
748 PCO and ATE1 (blue) and Cys-$RAP2_{2-13}$ incubated without PCO or ATE1 (purple),
749 confirming arginylation only at the N-terminus of PCO-modified $RAP2_{2-13}$.