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

Mark D. White¹, Maria Klecker²,³, Richard J. Hopkinson¹, Daan Weits⁴, Carolin Mueller⁵,⁶, Christin Naumann²,³, Rebecca O’Neill¹, James Wickens¹, Tom N. Grossmann⁵,⁶, Nico Dissmeyer²,³,*⁷, Emily Flashman¹,*

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

* 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.
Abstract

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

All aerobic organisms require homeostatic mechanisms to ensure O₂ supply and demand are balanced. When supply is reduced (hypoxia), a hypoxic response is required to decrease demand and/or improve supply. In animals, this well-characterized response is mediated by the Hypoxia-Inducible transcription Factor (HIF), which upregulates genes encoding for vascular endothelial growth factor, erythropoietin and glycolytic enzymes amongst many others.¹⁻³ Hypoxia in plants is typically a consequence of reduced O₂ diffusion under conditions of waterlogging or submergence, or inside of organs such as seeds, embryos, or floral meristems in buds where the various external cell layers act as diffusion barriers.

Although plants can survive temporary periods of hypoxia, flooding negatively impacts on plant growth, and if sustained it can result in plant damage or death⁴. This has a major impact on crop yield, for example, flooding resulted in crop loss costing $3 billion in the U.S. in 2011.⁵ As climate change results in increased severe weather events including flooding⁴, strategies to address crop survival under hypoxic stress are needed to meet the needs of a growing worldwide population.

The response to hypoxia in rice, Arabidopsis, and barley is known to be mediated by the group VII ETHYLENE RESPONSE FACTORs (ERF-VIIs).⁶⁻¹¹ It has been found that these transcription factors promote the expression of core hypoxia-responsive genes, including those encoding alcohol dehydrogenase and pyruvate decarboxylase that facilitate anaerobic metabolism.¹²,¹³ Crucially, it was shown, initially in Arabidopsis, that the stability of the ERF-VIls is regulated in an O₂-dependent manner via the Arg/Cys branch of the N-end rule pathway, which directs proteins for proteasomal degradation depending on the identity of their N-terminal amino acid.¹⁴⁻¹⁶ Thus, a connection between O₂ availability and the plant
hypoxic response was identified. The Arabidopsis ERF-VIIIs are translated with the conserved N-terminal motif MCGGAI/VSDY/F and co-translational N-terminal methionine excision, catalyzed by Met amino peptidases (MAPs), leaves an exposed N-terminal Cys which is susceptible to oxidation. N-terminal oxidized Cys residues are then proposed to render the ERF-VII N-termini substrates for arginyl tRNA transferase (ATE)-catalyzed arginylation. The subsequent Nt-Arg-ERF-VIIIs are candidates for ubiquitination by the E3 ligase PROTEOLYSIS6 (PRT6) which promotes targeted degradation via the 26S proteasome. It has also been shown that degradation of ERF-VIIIs by the N-end rule pathway is influenced by NO, and that the ERF-VIIIs play a key role in plant NO-mediated stress responses. The relationship between O2- and NO-mediated ERF-VII modification remains to be fully elucidated.

The plant hypoxic response mimics the equivalent well-characterized regulatory system in animals, whereby adaptation to hypoxia is mediated by HIF. In normoxic conditions, HIF is hydroxylated at specific prolyl residues targeting it for binding to the von Hippel-Lindau tumour suppressor protein (pVHL), the recognition component of the E3-ubiquitin ligase complex, which results in HIF ubiquitination and proteasomal degradation. Thus, while not substrates for the N-end rule pathway of protein degradation, HIF levels are regulated by post-translational modification resulting in ubiquitination, in a manner that is sensitive to hypoxia. HIF prolyl hydroxylation is catalyzed by O2-dependent enzymes, the HIF prolyl hydroxylases (PHDs 1-3), which are highly sensitive to O2 availability. These O2-sensing enzymes are thus the direct link between O2 availability and the hypoxic response.

Crucially, a family of five enzymes, the PLANT CYSTEINE OXIDASEs (PCO1-5) were identified in Arabidopsis that were reported to catalyze the O2-dependent reaction in the plant hypoxic response, specifically the oxidation of the conserved Cys residue at the N-
terminus of the Arabidopsis ERF-VII, 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 O2 in the presence of pentameric peptides CGGAI corresponding to the N-termini of various ERF-VIIs (Supplementary Table 1). The identification of these enzymes indicates that the hypoxic response in plants is enzymatically regulated, 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 O2 sensors.

We 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 catalyze dioxygenation of the N-terminal Cys of Arabidopsis ERF-VII peptide sequences to Cys-sulfenic acid (CysO2), rather than the fully oxidized Cys-sulfonic acid (CysO3) (Supplementary Figure 1). This oxidation directly incorporates molecular O2. To our knowledge, these are the first described enzymes that catalyze cysteinylation, as well as being the first described cysteine dioxygenases in plants. We then verify that the product of the PCO-catalyzed reactions is a 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. Overall, we thus define the PCOs as plant cysteinyldioxygenases and ATE1 as an active arginyl transferase, establishing for the first time a direct link between molecular O2, PCO catalysis and ATE1 recognition and modification of N-end rule substrates.
Results

PCOs catalyze modification of RAP2_{2-11} in an O_2-dependent manner

N-terminally hexahistidine-tagged recombinant PCO1 and 4 were purified to ~90% purity, as judged by SDS-PAGE (Supplementary Figure 2a). Protein identity was confirmed by comparison of observed and predicted mass by LC-MS (PCO1 predicted mass 36,510 Da, observed mass 36,513 Da; PCO4 predicted mass 30,680 Da, observed mass 30,681 Da, Supplementary Figure 2b). The activity of the purified PCO1 and PCO4 was tested towards a synthetic 10-mer peptide corresponding to the methionine excised N-termin of the ERF-VII s RAP2.2, RAP2.12, and HRE2 (H_{2}N-CGGAIISDFI-COOH, hereafter termed RAP2_{2-11} Supplementary Table 1). Assays comprising RAP2_{2-11} at 100 µM in the presence or absence of PCO1 or PCO4 at 0.5 µM underwent aerobic or anaerobic coinubcation for 30 minutes at 30°C prior to analysis of the peptide by matrix-assisted laser desorption/ionization-mass spectrometry (MALDI-MS, Figure 1a,b). Only under aerobic conditions and in the presence of PCO1 or PCO4, the spectra revealed the appearance of two species with mass increases of +32 Da and +48 Da, corresponding to two or three added O atoms, highlighting an O_2-dependent reaction for PCOs 1 and 4 (Figure 1b), as previously shown for PCOs 1 and 2. These mass shifts were deemed to be consistent with enzymatic formation of Cys-sulfinic (CysO_2, +32 Da) and Cys-sulfonic acid (CysO_3, +48 Da; Supplementary Figure 1), both proposed to be Arg transferase substrates in the Arg/Cys branch of N-end rule mediated protein degradation. 

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

To ascertain whether the PCOs function as dioxygenases, we sought to verify the source of the O atoms in the oxidized RAP2_{2-11} by conducting assays in the presence of ^{18}O_2 as the
cosubstrate or $H_2^{18}O$ as the solvent. To probe $O_2$ as the source of $O$ atoms in the product, anaerobic solutions of RAP2$_{2,11}$ were prepared in sealed vials before addition of PCO4 using a gas-tight syringe. The vials were then purged with $^{16}O_2$ or $^{18}O_2$ and the reactions were allowed to proceed at 30°C for a subsequent 20 minutes. Upon analysis by MALDI-MS, the mass of the products revealed that molecular $O_2$ was incorporated into the Cys-sulfinic acid product, but not the Cys-sulfonic acid product (Figure 2a). The Cys-sulfinic acid product has a mass of +32 Da in the presence of $^{16}O_2$ and +36 Da in the presence of $^{18}O_2$, demonstrating addition of two $^{18}O$ atoms and indicating that $O_2$ is the source of $O$ atoms in this product. The Cys-sulfonic acid product has a mass of +52 Da in the presence of $^{18}O_2$, indicating a third $^{18}O$ atom has not been incorporated into this product. To probe whether the source of the additional mass in the apparent Cys-sulfonic acid product was an $O$ atom derived from water, an equivalent reaction 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). No additional mass was observed in the peak corresponding to the Cys-sulfonic acid, raising the possibility that the +48 Da species observed by MALDI-MS is not enzymatically formed. Importantly, no additional mass was observed in the peak corresponding to Cys-sulfonic acid, confirming that this species is a product of a reaction where molecular $O_2$ is a substrate (Figure 2b).

To further investigate whether the PCO-catalyzed product species observed at +48 Da is enzymatically produced or an artifact of the MALDI-MS analysis method, we turned to liquid chromatography-mass spectrometry (LC-MS) to analyze the products of the PCO-catalyzed reactions. Under these conditions, only peptidic product with a mass increase of +32 Da was observed after incubation with both PCO1 and PCO4, corresponding to the incorporation of two $O$ atoms and the formation of Cys-sulfinic acid (Figure 2c), consistent with the products observed using $^{18}O_2$ and $H_2^{18}O$ (Figure 2a,b). No product was observed with a mass
corresponding to Cys-sulfonic acid, which suggested that the +48 Da product detected by MALDI-MS was indeed an artifact. When combined with the observation that significant quantities of Cys-sulfonic acid were not seen in no-enzyme or anaerobic controls (Figure 1), it was hypothesized that the Cys-sulfinic acid product of the PCO-catalyzed reaction is non-enzymatically converted to Cys-sulfonic acid during MALDI-MS analysis, potentially as a result of laser exposure. Upon subjecting the products of PCO1 and 4 turnover to MALDI-MS analysis with increasing laser intensity, a direct correlation between laser intensity and the ratio of Cys-sulfonic acid:Cys-sulfinic acid product was observed (Supplementary Figure 3a). Of note, significant levels of laser induced formation of +32 and +48 Da species upon analysis of unmodified peptide were not observed (Supplementary Figure 3b). Together, these results confirm that the +48 Da species observed following incubation of the PCOs with RAP2,11 are a product of Cys-sulfinic acid exposure to the MALDI-MS laser, and not a product of the PCO-catalyzed reaction. Overall, these data demonstrate that the PCOs are dioxygenase enzymes, similar to the mammalian and bacterial cysteine dioxygenases (CDOs) to which they show sequence homology.28

**PCOs catalyze oxidation of N-terminal Cys of RAP2,11 to form Cys-sulfinic acid**

Recombinant PCO1 and PCO2 were reported to consume O2 in the presence of pentameric CGGAI peptides corresponding to the methionine-excised N-terminus of the Arabidopsis ERF-VIIs.28 To definitively verify that the N-terminal cysteiny1 residue of RAP2,11 is indeed the target for the PCO-catalyzed +32 Da modifications, we conducted LC-MS/MS analyses on the reaction products. Fragmentation of RAP2,11 that had been incubated in the presence and absence of PCO1 and PCO4 revealed b- and y-ion series consistent with oxidation of the N-terminal Cys residue (Figure 3a), confirming that the PCOs 1 and 4 act as cysteiny1 dioxygenases.
As a final confirmation of the nature of the reaction catalyzed by PCO1 and PCO4, their activity was monitored using $^1$H-NMR. Reactions were initiated by adding 5 µM enzyme to 500 µM RAP2$_{2,11}$ (in the presence of 10% D$_2$O) and products of the reaction were analysed using a 600 MHz NMR spectrometer. In the presence of both PCO1 and PCO4, modification to the cysteinyln residues was observed, as exemplified by the disappearance of the $^1$H-resonance corresponding to the β-cysteinyln protons (at $\delta_H$ 2.88 ppm) and the emergence of a new $^1$H-resonance at $\delta_H$ 2.67 ppm (Figure 3b). The chemical shift of the new resonance is similar to that observed for L-Cys conversion to L-Cys-sulfinic acid by mouse CDO,$^{29}$ and also to the chemical shift of an L-Cys-sulfinic acid standard measured under equivalent conditions to the PCO assays (Supplementary Figure 4). Therefore, the resonance shift observed upon PCO1/4 reaction was assigned to the β-protons of L-Cys-sulfinic acid. Overall these results provide confirmation at the molecular level that Arabidopsis PCOs 1 and 4 act as plant cysteinyln dioxygenases, catalyzing incorporation of O$_2$ into N-terminal Cys residues on a RAP2 peptide to form Cys-sulfinic acid.

ATE1 arginylates acidic N-termini including Cys-sulfinic acid

We next sought to confirm that the PCO-catalyzed Cys-oxidation to Cys-sulfinic acid renders a RAP2 peptide capable of and sufficient for onward modification by ATE1, as proposed by the N-end rule pathway. ATE1 has been suggested to be responsible for transfer of $^3$H-arginine to bovine α-lactalbumine in highly purified plant extracts in vitro$^{30}$ and RAP2.12 stabilization in $ate1 ate2$ double null mutant plant lines implicates ATE1 as an ERF-VII-targeting arginyl transferase in vivo.$^{17,18}$ To this end, we produced recombinant hexahistidine-tagged Arabidopsis ATE1 (Supplementary Figure 5) for use in an arginylation assay which detects incorporation of radiolabeled $^{14}$C-Arg into biotinylated peptides. C-terminally biotinylated RAP2$_{2,13}$ peptides (H$_2$N-XGGAIISDFIPP(PEG)K(biotin)-NH$_2$) where the N-
terminal residue, X, constitutes Gly, Asp, Cys or Cys-sulfonic acid were subjected to the arginylation assay in the presence or absence of PCO1/4 (Figure 4a). Peptide with an N-terminal Gly did not accept Arg, while an N-terminal Asp did accept Arg, independent of the presence of PCO1 or 4. A peptide comprising an N-terminal Cys-sulfonic acid was also shown to be a substrate for ATE1, again independent of the presence of PCO1 or 4.

Crucially, RAP22-13 with an N-terminal Cys in the absence of PCO1/4 was not an acceptor of arginine transfer by ATE1, yet when either PCO1 or PCO4 was incorporated in the reaction, significant ATE1 transferase activity was observed (Figure 4a).

To confirm that the increased detection of radiolabel corresponded to arginyl incorporation at the N-termini of the peptides, the experiment was repeated using non-radiolabeled arginine in the presence and absence of PCO4 and ATE1, and peptide products subjected to LC-MS analysis (Figure 4c). As with RAP22-11 (Figure 2c), the Cys-initiated RAP22-13 peptide displayed a +32 Da increase in mass upon incubation with PCO4 only (Figure 4c, red spectrum). Importantly, following incubation of Cys-initiated RAP22-13 with both PCO4 and ATE1, a mass increase equivalent to oxidation coupled to arginylation (+188 Da) was observed (Figure 4c, blue spectrum). Subsequent tandem MS analysis of these product ions revealed fragmentation species consistent with the assumption that oxidation and sequential arginylation occur at the N-terminus of PCO4- and ATE1-treated peptides (Figure 4d, blue spectrum), strongly suggesting that the PCO-oxidized N-termini of ERF-VIIs are rendered N-degrons via additional arginylation (Figure 4b).

A +12 Da mass increase was observed in products of control assays lacking PCO4 (Figure 4c, d; purple spectra). This appeared to be related to prolonged incubation in the presence of HEPES and DTT as used in the arginylation assay buffer: The +12 Da modification was not observed if the peptide was dissolved in H2O (Figure 4c, black spectrum) or if incubated in
White M et al.

HEPES and DTT for just 1 hour, but was observed when the peptide was incubated in HEPES and DTT overnight (Supplementary Figure 6). It is proposed that trace levels of contaminating formaldehyde react with free Nt-Cys residues to form thiazolidine N-termini. These results are in line with proposed arginylation requirements for the Arg/Cys branch of the N-end rule pathway including the known Cys-initiated arginylation targets from mammals. Importantly, these results demonstrate for the first time Arg transfer mediated by ATE1 dependent on the N-terminal residue of its substrate, and also that both Cys-sulfenic acid (the product of PCO-catalysis) and Cys-sulfonic acid can act as substrates for ATE1. The arginylation observed with PCO-catalyzed Cys-sulfinic acid in particular supports the assumption that N-terminal residues sterically and electrostatically resembling Asp or Glu can serve as Arg acceptors in reactions catalyzed by ATEs, and also confirms the importance of the PCOs as a connection between the stability of their ERF-VII substrates and O₂ availability (Figure 4b).

Discussion

The PCOs were identified in Arabidopsis thaliana as a set of five enzymes suggested to catalyze oxidation of N-terminal cysteine residues in ERF-VII transcription factors and oxidation was demonstrated in short peptides corresponding to their N-termini. This oxidation was associated with destabilization of the ERF-VIIs, presumably by rendering them substrates of the Arg/Cys branch of the N-end rule pathway. Under conditions of sufficient O₂ availability, ERF-VII protein levels are decreased, while under hypoxic conditions, such as those encountered upon plant submergence or in the context of organ development, ERF-VII levels remain high. Importantly, the ERF-VII transcription factors are known to upregulate genes which allow plants to cope with or respond to submergence.
The PCOs are proposed to act as potential O\textsubscript{2} sensors involved in regulating the plant hypoxic response.\textsuperscript{28}

We sought to biochemically confirm the role of the PCOs in the plant hypoxic response, and present here mass spectrometry and NMR data that clearly demonstrate that two enzymes from this family, PCOs 1 and 4, are dioxygenases that catalyze direct incorporation of O\textsubscript{2} into RAP2\textsubscript{2-11} peptides to form Cys-sulfinic acid. Their direct use of O\textsubscript{2} supports the proposal that these enzymes are plant O\textsubscript{2} sensors.\textsuperscript{28} A relationship has been demonstrated between O\textsubscript{2} concentration and PCO activity,\textsuperscript{28} but it will be of interest to perform detailed kinetic characterization of these enzymes to ascertain their level of sensitivity to O\textsubscript{2} availability, in particular whether their O\textsubscript{2}-sensitivity is similar to that of the HIF hydroxylases in animals.\textsuperscript{24,25} Although there is functional homology between the PCOs and the HIF hydroxylases, they are apparently mechanistically divergent: The PCOs show sequence homology to the Fe(II)-dependent CDO family of enzymes which do not require an external electron donor for O\textsubscript{2} activation,\textsuperscript{28,35} while the HIF hydroxylases are Fe(II)/2OG-dependent oxygenases. Of note, the PCOs are the first identified CDOs in plants. Further, in contrast to the reactions of mammalian and bacterial CDOs which oxidize free L-Cys, the PCOs are also, to our knowledge, the first identified cysteinyI (as opposed to free L-Cys) dioxygenases.

According to the Arg/Cys branch of the N-end rule pathway, N-terminal Cys oxidation is proposed to enable successive arginylation by ATE1 to render proteins as N-degrons. We therefore sought to demonstrate that PCO-catalyzed ERF-VII N-terminal Cys oxidation to Cys-sulfinic acid promotes arginylation by ATE1. The arginylation assay and mass spectrometry results we present demonstrate that the PCO-catalyzed dioxygenation reaction is sufficient to trigger N-terminal arginylation of RAP2s by ATE1, thus likely rendering
ERF-VIIIs (at least those comprising the tested N-terminal sequence) as N-degrons, i.e. recognition by PRT6 and other potential E3 ubiquitin ligases, polyubiquitination and possibly transfer to the 26S proteasome for proteolysis.\textsuperscript{14-16} Collectively therefore, we present the first molecular evidence confirming the Cys-oxidation and subsequent arginylation steps of the Arg/Cys branch of the N-end rule pathway.\textsuperscript{32-34} We also confirm that ATE1 is able to selectively arginylate, as predicted,\textsuperscript{34} acidic N-terminal residues of plant substrates, including Cys-sulfonic acid.

ATE1 is capable of arginylation at both acidic N-termini and midchains, though N-terminal acidic residues are reported to be arginylated via canonical peptide bonds.\textsuperscript{36} While reports of midchain arginylation are relatively common,\textsuperscript{36,37,38} to date only one physiological and two \textit{in vitro} substrates for the Arg/Cys branch of the N-end rule pathway have been characterized, namely mammalian regulator of G protein signaling (RGS) 4, and RGS5 and 10 respectively,\textsuperscript{39} where Nt-Cys oxidation was described (to either Cys-sulfinic or Cys-sulfonic acid) as was Nt-Cys arginylation.\textsuperscript{32,34} The first three non Cys-branch N-end rule arginylation targets were only very recently identified in mammalian cell culture:\textsuperscript{40} the molecular chaperone BiP (GRP78 and HSPA5, heat shock 70 kDa protein 5; Glu-initiated), calreticulin (CRT, Glu-initiated), and the oxidoreductase protein disulphide isomerase (PDI, Asp-initiated). Similarly, data regarding the molecular requirements of plant ATEs is limited. Developmental functions of the single homolog ATE1 in the moss \textit{Physcomitrella patens} were recently described\textsuperscript{41} and interaction partners of the enzyme found and arginylated peptides immunologically-detected.\textsuperscript{42} Previously, Arg transferase function of Arabidopsis ATE1/2 has been shown using an assay detecting conjugation of \textsuperscript{3}H-Arg to bovine α-lactalbumin (bearing an N-terminal Glu) in the presence of plant extracts from wild type Arabidopsis, \textit{atel} and \textit{ate2} single mutants but not from \textit{atel ate2} double mutant seedlings.\textsuperscript{30}
Therefore, the results we present here demonstrate for the first time Arg transferase activity of a plant ATE towards known plant N-end rule substrates.

Interestingly, as well as O₂, nitric oxide was identified as an RGS oxidizing agent, suggesting a potential role of S-nitrosylation in the Arg/Cys branch of the N-end rule pathway, albeit non-enzymatically controlled. It has also been reported in planta that both NO and O₂ are required for ERF-VII degradation, potentially at the Cys oxidation step. Although in N-end rule-mediated RGS4/5 degradation it has been proposed that Cys-nitrosylation precedes Cys-oxidation (also currently considered a non-enzymatic process), we find that under the conditions used, the PCO1/4-catalyzed reaction does not require either prior cys-nitrosylation or exogenous NO to proceed efficiently. We cannot rule out that NO plays a role in formation of a Cys-sulfonic acid product, which is also a substrate for ATE1 as shown in our Arg transfer experiments. Alternatively, NO may have a role in ERF-VII degradation in vivo via non-enzymatic oxidation or a secondary mechanism. The manner in which NO contributes to Arg/Cys branch of the N-end rule pathway therefore remains to be elucidated.

ERF-VII stabilization has been shown to result in improved submergence tolerance, elegantly demonstrated in barley by mutation of the candidate E3-ubiquitin ligase PRT6, but also in rice containing the Sub1A gene; SUB1A is an apparently stable ERF-VII that confers particular flood tolerance in certain rare varieties of rice. Overexpression of Sub1A in more commonly grown rice varieties has resulted in a 45% increase in yield relative to sub1a mutant lines after exposure to flooding. If ERF-VII stabilization is indeed a proficient mechanism for enhancing flood tolerance, then manipulation of PCO or ATE activity may be an efficient and effective point of intervention. This work presents molecular validation of their function, providing the basis for future targeted chemical/genetic inhibition of their
activity, but also highlighting genetic strategies for breeding via introgression of variants of N-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.

Online Methods

Peptide Synthesis, Purification and Characterization

All reagents used were purchased from Sigma-Aldrich unless otherwise stated. The 10-mer RAP2_{2.11} peptide (H_{2}N-CGGAIIISDFI-COOH) was purchased from GL Biochem (Shanghai) Ltd, China (Supplementary Table 1). The sequence of the 12-mer peptides used in the coupled oxidation-arginylation assay is derived from RAP2.2 and RAP2.12 (H_{2}N-X-GGAIISDFIPP(PEG)K(biotin)-NH_{2}) and synthesized by Fmoc-based solid-phase peptide synthesis (SPPS) on NovaSyn®TGR resin (Merck KGaA, Supplementary Table 2). Fmoc protected amino acids (Iris Biotech GmbH) were coupled using 4 equivalents of (eq) of the amino acid according relative to initial resin loading. 4 eq amino acid was mixed with 4 eq O-(6-chlorobenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HCTU) and 8 eq N,N-diisopropylethylamine (DIPEA; Santa Cruz Biotechnology, sc-293894) and added to the resin for 1 h. In a second coupling, the resin was treated with 4 eq of the Fmoc-protected amino acid mixed with 4 eq benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluoro-phosphate (PyBOP) and 8 eq 4-methylmorpholine (NMM) for 1 h. After double coupling a capping step to block free amines was performed using acetyanhydride and DIPEA in N-methyl-2-pyrrolidinone (NMP) (1:1:10) for 5 min. The C-terminal Fmoc-Lys(biotin)
OH, the 8-(9-fluorenylmethyloxycarbonyl-amino)-3.6-dioxaoctanoic acid (PEG) linker and the different Fmoc protected N-terminal amino acids were coupled manually. The remaining peptide sequence was assembled using an automated synthesizer (Syro II, MultiSynTech GmbH). Fmoc deprotection was performed using 20 % piperidine in dimethylformamide (DMF) for 5 min, twice. After each step the resin was washed 5 times with DMF, DCM and DMF, respectively. Final cleavage was performed with 94 % trifluoroacetic acid (TFA), 2.5 % 1,2-ethanedithiole (EDT) and 1 % triisopropylsilane (TIPS) in aqueous solution for 2 h, twice. The cleavage solutions were combined and peptides were precipitated with diethyl ether (Et₂O) at -20°C for 30 min. Peptides were solved in water/acetonitrile (ACN) 7:3 and purified by reversed-phase high performance liquid chromatography (HPLC; Nucleodur C18 column; 10×125 mm, 110 Å, 5 µm particle size; Macherey-Nagel) using a flow rate of 6 ml·min⁻¹ (A: ACN with 1 % TFA, B: water with 1 % TFA). Obtained pure fractions were pooled and lyophilized. Peptide characterization was performed by analytical HPLC (1260 Infinity, Agilent Technology; flow rate of 1 ml·min⁻¹, A: ACN with 1 % TFA, B: water with 1 % TFA) coupled with a mass spectrometer (6120 Quadrupole LC/MS, Agilent Technology) using electro spray ionization (Agilent Eclipse XDB-C18 column, 4.6×150 mm, 5 µm particle size). Analytical HPLC chromatograms were recorded at 210 nm (Supplementary Figure 7). Quantification was performed by HPLC-based comparison (chromatogram at 210 nm) with a reference peptide (Supplementary Table 2).

Preparation of Recombinant Proteins

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

The coding sequence of Arabidopsis ATE1 was cloned according to gene annotations at TAIR (www.arabidopsis.org) from cDNA. The sequence was flanked by an N-terminal tobacco etch virus (TEV) recognition sequence for facilitated downstream purification (“tev”: ENLYFQ-X) using the primers ate1_tev_ss (5’-GCTTAGAGAAATCTTTATTTTCAGGGGATGTCTTTGAAAAACGATGCGAGT-3’) and ate1_as (5’-GGGGACCACTTTGTACAAGAAAGCTGGGTATCAGTTGATTTTCATACACCATTTC-3’). A second PCR using the primers adapter (5’-GGGGACAAGTTTGTACAAAAAAGCAGGCTTAGAGAATCTTTATTTTCAGGG-3’) and ate1_as was performed to amplify the construct to use it in a BP reaction for cloning into pDONR201 (Invitrogen) followed by an LR reaction into the vector pDEST17 (Invitrogen).

The N-terminal hexahistidine fusion was expressed in BL21-CodonPlus (DE3)-RIL Escherichia coli (E. coli) cells. The expression culture was induced with 1 mM IPTG at OD=0.6 and grown for 16 hours at 18°C. After resuspension in LEW buffer (50 mM NaH₂PO₄, pH 8; 300 mM NaCl; 1 mM DTT), the cells were lysed by incubation with 1.2 mg/ml lysozyme for 30 min and subsequent sonification in the presence of 1 mM PMSF. Recombinant protein was purified by Ni⁺⁺ affinity chromatography and subjected to Amicon Ultra-15 (30K) (Merck Millipore) filtration for buffer exchange to imidazole-free LEW containing 20% glycerol.
PCO Activity Assays and MALDI Analysis

PCO activity assays were conducted under the following conditions unless otherwise stated: 1 µM PCO1 or 4 was mixed with 100 or 200 µM RAP2_{2-11} peptide in 250 mM NaCl, 1 mM dithiothreitol (DTT), 50 mM Tris-HCl pH 7.5 and incubated at 30°C for 30-60 minutes. Assays were stopped by quenching 1 µL sample with 1 µL alpha-cyano-4-hydroxycinnamic acid (CHCA) matrix on a MALDI plate prior to product mass analysis using a Sciex 4800 TOF/TOF mass spectrometer (Applied Biosystems) operated in negative ion reflectron mode. The instrument parameters and data acquisition were controlled by 4000 Series Explorer software and data processing was completed using Data Explorer (Applied Biosystems).

To test the activity of PCO4 in the presence of ^{18}O_2, 100 µL of an anaerobic solution of 100 µM RAP2_{2-11} in 250 mM NaCl/50 mM Tris-HCl pH 7.5 was prepared in a septum-sealed glass vial by purging with 100% N_2 for 10 minutes at 100 mL/min using a mass flow controller (Brooks Instruments), as used for previous preparation of anaerobic samples to determine enzyme dependence on O_2. PCO4 was then added using a gas-tight Hamilton syringe, followed by purging with a balloon (approx. 0.7 L) of ^{16}O_2 or ^{18}O_2 over the course of 10 minutes at room temperature. Reaction vials were then transferred to 30°C for a further 20 minutes before products were analysed by MALDI-MS as described above.

PCO4 activity was additionally tested in the presence of H_2^{18}O by conducting an assay in 75% H_2^{18}O, 25% H_2O (with all enzyme/substrate/buffer components comprising a portion of the H_2O fraction). Assays were conducted for 10 minutes at room temperature followed by 20 minutes at 30°C for comparison with assays conducted with ^{18}O_2. Products were analysed by MALDI-MS, as described above.
UPLC-MS and MS/MS Analysis of PCO Assay Products

Ultra-high performance chromatography (UPLC) mass spectrometry (MS) measurements were obtained using an Acquity UPLC system coupled to a Xevo G2-S Q-ToF mass spectrometer (Waters) operated in positive electrospray mode. Instrument parameters, data acquisition and data processing were controlled by Masslynx 4.1. Source conditions were adjusted to maximize sensitivity and minimize fragmentation while Lockspray was employed during analysis to maintain mass accuracy. 2 µL of each sample was injected on to a Chromolith Performance RP-18e 100-2 mm column (Merck) heated to 40 °C and eluted using a gradient of 95 % deionized water supplemented with 0.1 % (v/v) formic acid (analytical grade) to 95 % acetonitrile (HPLC grade) and a flow rate of 0.3 mL/min. Fragmentation spectra of substrate and product peptide ions (MS/MS) were obtained using a targeted approach with a typical collision-induced dissociation (CID) energy ramp of 30 to 40 eV. Analysis was carried out with the same source settings, flow rate and column elution conditions as above.

1H NMR Assay

Reaction components (5 µM PCO1 or PCO4 and 500 µM RAP22-11) were prepared to 75 µL in 156 mM NaCl, 31 mM Tris-HCl (pH 7.5) and 10% D2O (enzyme added last), in a 1.5 mL microcentrifuge tube before being transferred to a 2 mm diameter NMR tube. 1H NMR spectra at 310 K were recorded using a Bruker AVIII 600 (with inverse cryoprobe optimized for 1H observation and running topspin 2 software; Bruker) and reported in p.p.m. relative to D2O (δH 4.72). The deuterium signal was also used as internal lock signal and the solvent signal was suppressed by presaturating its resonance.
Arginylation Assay

The conditions for arginylation of the 12-mer peptide substrates were modified from 36. In detail, ATE1 was incubated at 10 µM in the reaction mixture containing 50 mM HEPES, pH 7.5; 25 mM KCl; 15 mM MgCl₂; 1 mM DTT; 2.5 mM ATP; 0.6 mg/ml *E. coli* tRNA (R1753, Sigma); 0.04 mg/ml *E. coli* aminoacyl-tRNA synthetase (A3646, Sigma); 80 µM (4 nCi/µl) ¹⁴C-arginine (MC1243, Hartmann Analytic); 50 µM C-terminally biotinylated 12-mer peptide substrate and, where indicated, 1 µM purified recombinant PCO1 or PCO4 in a total reaction volume of 50 µL. The reaction was conducted at 30°C for 16 to 40 hours. After incubation, each 50 µL of avidin agarose bead slurry (20219, Pierce) equilibrated in PBSN (PBS-Nonidet; 100 mM NaH₂PO₄; 150 mM NaCl; 0.1% Nonidet-P40) were added to the samples and mixed with an additional 350 µL of PBSN. After 2 hours of rotation at room temperature, the beads were washed 4 times in PBSN, resuspended in 4 mL of FilterSafe scintillation solution (Zinsser Analytic) and scintillation counting was performed using a Beckmann 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. 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., 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.

Author Contributions

M.W. performed the PCO1/4 activity assays and MALDI/LC-/MS/MS analyses. M.K. performed and established arginylation reactions on peptides coupled to biotin pulldown and scintillation measurements and purified ATE1 protein. R.H. performed the NMR assays with E.F. D.W. prepared the pDEST17-PCO1 and 4 plasmids. C.M. synthesized the peptides, T.N.G. supervised and designed the synthesis, C.N. cloned and established purification and activity assays for ATE1. R.O. conducted LC-MS to analyse +12 Da mass shifts. J.W.
performed LC-MS analysis. E.F. performed the PCO1 and PCO4 protein purification and selected activity assays. E.F., M.W., M.K. and N.D. designed the study, E.F. and N.D. wrote the manuscript. M.W., M.K., N.D. and E.F. designed the figures. All authors read and approved the final version of this manuscript.
References


Figure 1. PCO1, PCO4 and O₂-dependent modification of a RAP2<sub>2-11</sub> peptide substrate, consistent with Cys-oxidation. MALDI-MS spectra showing products following PCO1 and PCO4 incubation with RAP2<sub>2-11</sub> under anaerobic (a) or aerobic (b) conditions. Products with mass increases of +32 Da and +48 Da were only observed in the presence of PCO1 or PCO4 and O₂.
Figure 2. The PCOs are dioxygenases which catalyze incorporation of molecular O\textsubscript{2} into RAP2\textsubscript{2-11}. (a) MALDI-MS spectra showing that PCO4-catalyzed reactions carried out in the presence of \textsuperscript{18}O\textsubscript{2} result in a +4 Da increase in the mass of the putative Cys-sulfenic 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\textsubscript{2}\textsuperscript{18}O show no additional incorporation of mass compared with products of reactions in the presence of H\textsubscript{2}\textsuperscript{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\textsubscript{2-11} results in formation of a single product with a mass increase of +32 Da, consistent with Cys-sulfenic acid formation.
Figure 3. PCO1 and PCO4 oxidize the N-terminal Cys of RAP2-11 to Cys-sulfenic 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 RAP2-11 with N-terminal Cys-sulfenic acid. (b) $^1$H-NMR was used to monitor changes to RAP2-11 (500 µM) upon incubation with enzyme (5 µM). In the presence of PCO1 (red) and PCO4 (purple), the $^1$H-resonance at δH 2.88 ppm (assigned to the β-cysteinyl protons of RAP2-11, blue) was observed to decrease in intensity, with concomitant emergence of a resonance at δH 2.67 ppm. This new resonance was assigned to the β-protons of Cys-sulfenic based on chemical shift analysis (see Supplementary Figure 4).
Figure 4. PCO-Catalyzed Cys-sulfinic acid formation renders RAP2\textsubscript{2-13} a substrate for ATE1 catalyzed arginylation. (a) \textsuperscript{14}C-Arg incorporation by ATE1 into the 12-mer N-terminal RAP2 peptide (H\textsubscript{2}N-XGGAISDFIPP(PEG)K(biotin))-NH\textsubscript{2}, X = Gly, Asp, Cys or Cys-sulfonic acid (C(O\textsubscript{3})), was assayed by liquid scintillation counting of immobilized biotinylated peptides after the arginylation reaction and removal of unreacted \textsuperscript{14}C-Arg (n=3). In the case of the Cys-starting peptide (RAP2\textsubscript{2-13}), ATE1 activity was strongly dependent on the presence of PCO1 or PCO4. (b) Scheme showing PCO- and ATE1-catalyzed reactions on...
Nt-Cys RAP2.2, as validated in this study. (c) LC-MS spectra of products of equivalent assays with Cys-initiated RAP2_{2,13} using non-radiolabelled Arg, revealing a sequential mass increase of +32 (corresponding to oxidation) and +156 Da (corresponding to arginylation) only in the presence of PCO and ATE1 (blue spectrum). The red spectrum shows a +32 Da mass increase for Cys-RAP2_{2,13} incubated +PCO/-ATE, demonstrating Cys-sulfonic acid formation as expected. Purple spectra show +12 Da products formed upon incubation of Cys-RAP2_{2,13} in the absence of PCO +/-ATE (for explanation of this mass increase see text and Supplementary Figure 6); the black spectrum shows Cys-RAP2_{2,13} dissolved in H_2O. (d) b- and y-ion series spectra generated by MS/MS analysis of Cys-RAP2_{2,13} only (no incubation; black), Cys-RAP2_{2,13} incubated +PCO/-ATE (red), Cys-RAP2_{2,13} incubated with PCO and ATE1 (blue) and Cys-RAP2_{2,13} incubated without PCO or ATE1 (purple), confirming arginylation only at the N-terminus of PCO-modified RAP2_{2,13}. 

White M et al.