


## STATE-OF-THE-ART REVIEW

**Emerging roles for thiol dioxygenases as oxygen sensors**Dona M. Gunawardana, Karen C. Heathcote and Emily Flashman 

Department of Chemistry, University of Oxford, UK

**Keywords**

cysteine dioxygenase; hypoxia; oxygen-sensing; thiol dioxygenase

**Correspondence**

E. Flashman, Department of Chemistry,  
 University of Oxford, 12 Mansfield Road,  
 Oxford, OX1 3TA, UK  
 Tel. +44 1865 275920  
 E-mail: emily.flashman@chem.ox.ac.uk

Dona M. Gunawardana and Karen C.  
 Heathcote contributed equally to this study

(Received 19 May 2021, revised 21 July  
 2021, accepted 3 August 2021)

doi:10.1111/febs.16147

Cysteine dioxygenases, 3-mercaptopropionate dioxygenases and mercaptosuccinate dioxygenases are all thiol dioxygenases (TDOs) that catalyse oxidation of thiol molecules to sulphinates. They are Fe(II)-dependent dioxygenases with a cupin fold that supports a 3xHis metal-coordinating triad at the active site. They also have other, broadly common features including arginine residues involved in substrate carboxylate binding and a conserved trio of residues at the active site featuring a tyrosine important in substrate binding catalysis. Recently, N-terminal cysteinyl dioxygenase enzymes (NCOs) have been identified in plants (plant cysteine oxidases, PCOs), while human 2-aminoethanethiol dioxygenase (ADO) has been shown to act as both an NCO and a small molecule TDO. Although the cupin fold and 3xHis Fe(II)-binding triad seen in the small molecule TDOs are conserved in NCOs, other active site features and aspects of the overall protein architecture are quite different. Furthermore, the PCOs and ADO appear to act as biological O<sub>2</sub> sensors, as shown by kinetic analyses and hypoxic regulation of the stability of their biological targets (N-terminal cysteine oxidation triggers protein degradation via the N-degron pathway). Here, we discuss the emergence of these two subclasses of TDO including structural features that could dictate their ability to bind small molecule or polypeptide substrates. These structural features may also underpin the O<sub>2</sub>-sensing capability of the NCOs. Understanding how these enzymes interact with their substrates, including O<sub>2</sub>, could reveal strategies to manipulate their activity, relevant to hypoxic disease states and plant adaptive responses to flooding.

**Introduction**

Thiol dioxygenases (TDOs) have been known for many decades as having important roles in cysteine catabolism, taurine metabolism and other small molecule thiol turnovers. These enzymes are Fe(II)-dependent oxygenases from the cupin superfamily of proteins, and a large body of work has been dedicated to understanding the structural and mechanistic features that underpin their catalytic properties [1]. Until 2014, there was

no reason to link their dioxygenase function with a role in O<sub>2</sub> sensing. However, the identification of plant cysteine oxidases (PCOs) with an O<sub>2</sub>-sensing role revealed a new subset of TDOs, enzymes that oxidise cysteine residues at the N termini of proteins, triggering their degradation in an O<sub>2</sub>-dependent manner via the Cys/Arg branch of the N-degron pathway [2,3]. We will refer to this subset of TDOs as NCOs (N-terminal

**Abbreviations**

3MDO, 3-mercaptopropionate dioxygenase; ADO, 2-aminoethanethiol dioxygenase; CDO, cysteine dioxygenase; ERF-VII, group VII ethylene response factor; HIF, hypoxia-inducible factor; IL32, interleukin 32; MSDO, mercaptosuccinate dioxygenase; NCO, N-terminal cysteinyl dioxygenase; PCO, plant cysteine oxidase; PHD, prolyl hydroxylase domain; RGS, regulator of G-protein signalling; TDO, Thiol dioxygenase.

cysteinyl dioxygenases). Subsequently, cysteamine dioxygenase (also known as 2-aminoethanethiol dioxygenases, ADO), until then understood to oxidise cysteamine, was also found to act as an NCO, regulating the stability of certain Cys-initiating proteins in an O<sub>2</sub>-sensitive manner [4]. Emerging structural information and sequence analysis revealed that these two subfamilies of TDO appear to have divergent features, which may rationalise their different functions and could help to reveal their O<sub>2</sub>-sensing mechanisms [5,6]. In this review, we summarise what we know to date about the two subfamilies and what is still to be uncovered. As TDOs are conserved across eukaryotic and prokaryotic kingdoms, such knowledge could help identify novel O<sub>2</sub>-sensitive pathways.

## Small molecule thiol dioxygenases

### Cysteine dioxygenase

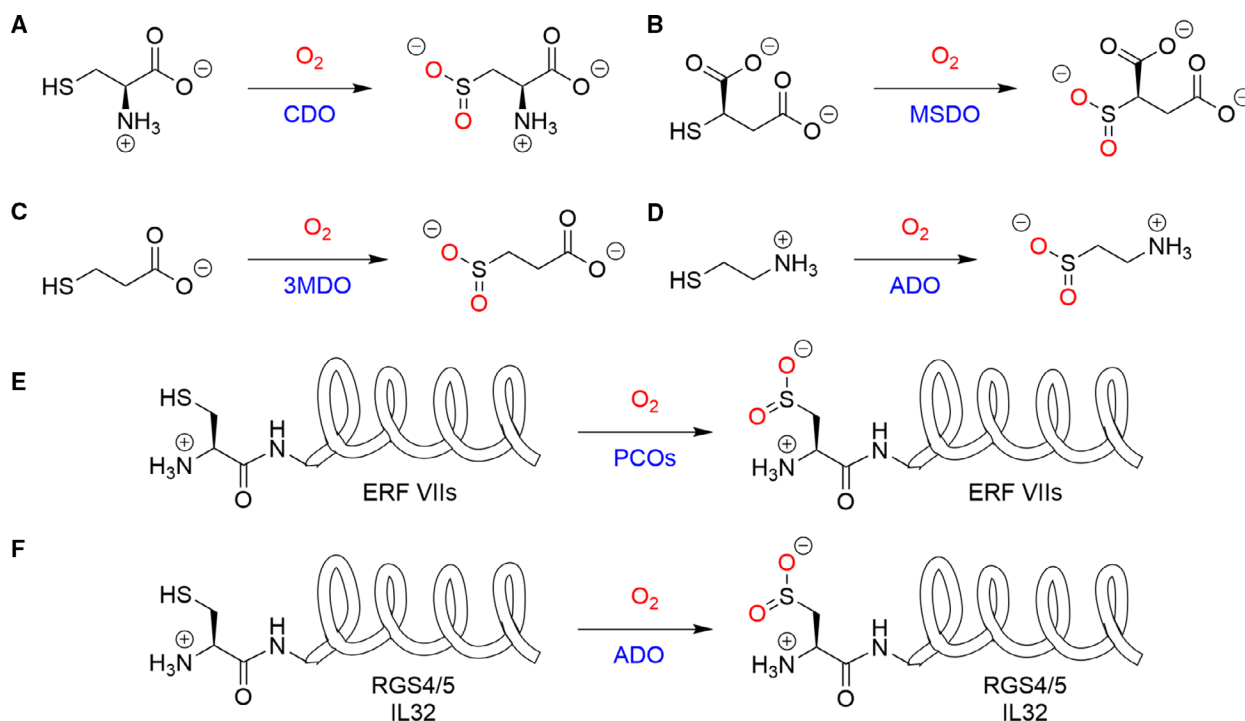
Cysteine dioxygenase (CDO) was the first characterised TDO and has been known for several decades to regulate cysteine levels in animals [7,8]. It catalyses oxidation of L-cysteine to L-cysteine sulphinic acid in an O<sub>2</sub>-dependent oxidation reaction, incorporating both oxygen atoms into the product (Scheme 1) [1]. Homeostasis of cysteine levels in the body is crucial as high cysteine levels are toxic and have been linked to neurological disorders such as Parkinson's disease and Alzheimer's disease [9]. CDO is expressed in a tissue-specific manner, particularly in liver, adipose tissues and lungs [10,11].

Sequence and structural analyses have identified that CDO belongs to the cupin superfamily of enzymes, which are non-haem metalloenzymes that catalyse a wide range of reactions [12]. They share common structural features such as the two conserved motifs (named cupin motif 1 and 2) and double  $\beta$  sheet helix (DBSH), and the active site of the enzymes is formed within this cupin fold. While CDO shares these conserved structural features (Fig. 1AI), its active site Fe(II) coordination is different to that of other cupin superfamily enzymes. Typically, 3-His-1-Glu or 2-His-1-Glu/Asp motifs coordinate the metalcenter; [12] however, in CDOs just three conserved histidine residues coordinate the active site iron [13]. Uncommon to the rest of the cupin enzymes, this motif is only seen in two other enzymes which are gentisate 1,2-dioxygenase (GDO) and diketone dioxygenase (Dke1) [1]. Sequence analysis shows that the metal-coordinating glutamate residue seen in other non-haem metalloenzymes is replaced by a cysteine residue in mammalian CDOs (Cys<sub>93</sub> in RnCDO).

A number of high-resolution crystal structures of mammalian CDOs have been published, including from mouse (PDB 2ATF) [13], rat (PDB 2B5H, 3ELN) [14,15] and humans (PDB 2IC1) [16], which reveal features of the enzymes' active sites (Fig. 1A). A substrate-free structure of rat CDO (PDB 2B5H) showed the active site iron cofactor in an unusual tetrahedral coordination with the three His residues as well as a water molecule, which was also coordinated with Tyr<sub>157</sub> [14]. The substrate-free mouse CDO structure (PDB 2ATF) instead suggested a distorted octahedral geometry at the active site (in which Fe(II) was substituted with Ni). Other conserved features of CDO active sites, at least in eukaryotes, are a Ser<sub>153</sub>-His<sub>155</sub>-Tyr<sub>157</sub> motif (RnCDO numbering) which form a hydrogen-bonding network, a Cys<sub>93</sub>-Tyr<sub>157</sub> thioether bond and Arg<sub>60</sub> involved in binding the substrate carboxylate group (Fig. 1AIII; Fig. 2B). CDOs are also present in certain phyla of bacteria, where their active sites differ slightly from their eukaryotic homologues with a Gly residue replacing Cys<sub>93</sub>, meaning a Cys-Tyr thioether bond is not observed [13,14].

Human CDO was the first structure (PDB 2IC1) reported with substrate present in the active site, although subsequent analysis of electron density of this structure disputes the presence of L-cysteine [16]. The structure of a rat CDO with L-Cys bound and a proposed intermediate (persulfenate) trapped in the active site have since been reported (PDB 3ELN), along with similar structures at a range of pH values [15,17]. These structures (e.g. PDB 4IEV at pH 8.0; Fig. 1AIV) show that the thiolate and amine groups of L-cysteine bind directly to the active site, coordinating the iron in a bidentate manner [15,17]. CDO Arg<sub>60</sub> is positioned to interact with the carboxylate group of the substrate via its guanidinium group. Tyr<sub>157</sub> is also in close proximity (2.2 Å) to the bound substrate, and Tyr<sub>157</sub>Phe variants show significantly reduced activity compared with the wild-type, suggesting an important role of the Tyr in substrate interaction and the catalytic mechanism (likely including formation of the Cys-Tyr cross-link) [16,18–20].

The post-translational covalent cross-link observed between residues Cys<sub>93</sub> and Tyr<sub>157</sub> is a modification which has only otherwise been seen to date in galactose oxidase [21]. The role of the arising Cys-Tyr thioether bond has been examined in some detail. It is not necessary for CDO activity, forming over the course of multiple turnovers, but enzymatic activity is enhanced in its presence. Cys<sub>93</sub> variants of CDO, as well as the Tyr<sub>157</sub>Phe variant, have been shown to be functional but with decreased activity [22–24].

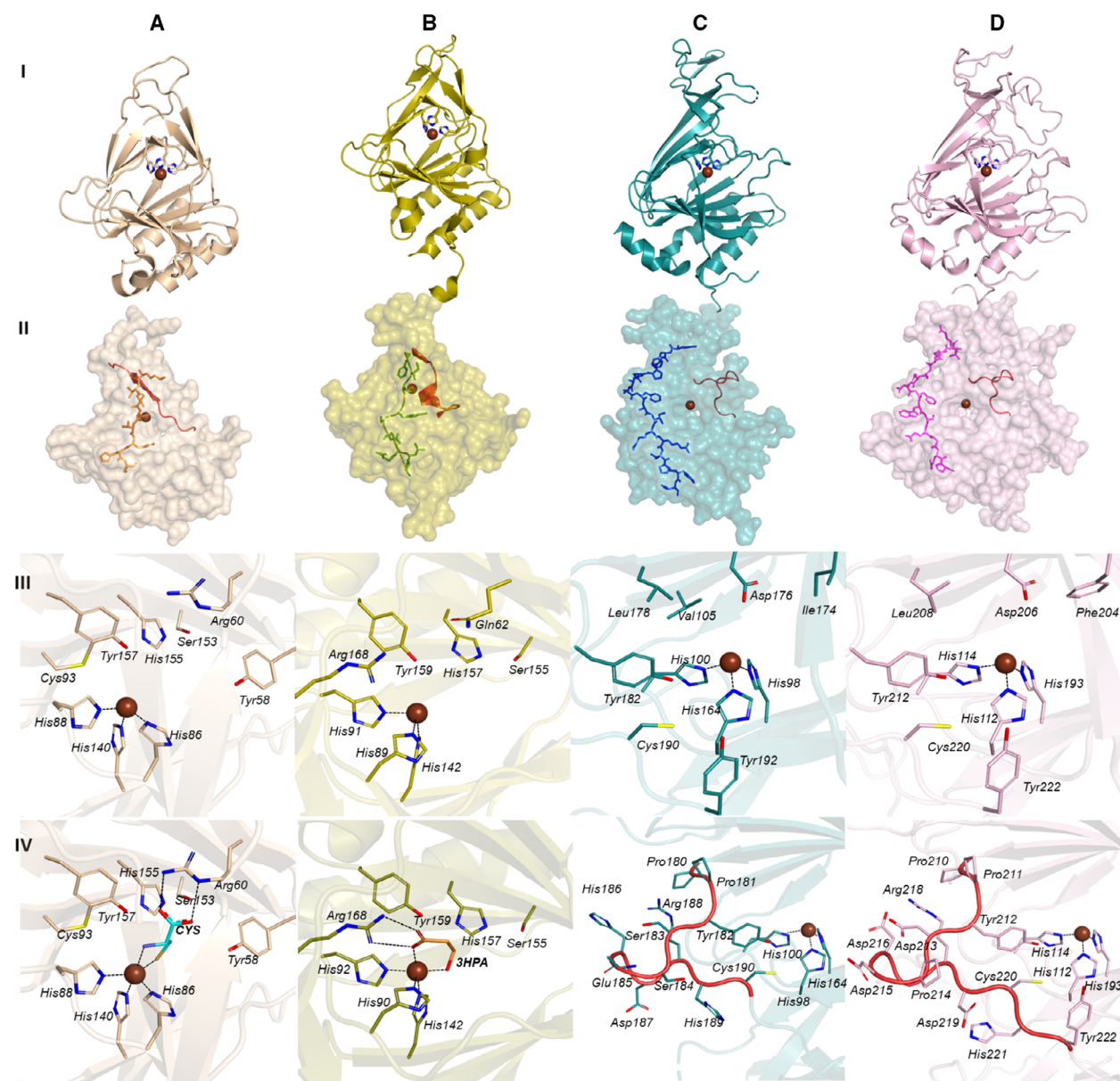


**Scheme 1.** Reaction mechanism of small molecule TDOs and NCOs. Thiol dioxygenases (TDOs) oxidise cysteine and analogous substrates by incorporating both atoms of molecular  $O_2$  (red). (A) Cysteine dioxygenase (CDO) catalyses conversion of L-cysteine to cysteine sulphinic acid while homologs of CDO found in bacteria, (B) mercaptosuccinate dioxygenase (MSDO) and (C) 3-mercaptopropionic acid dioxygenase (3MDO) catalyse conversion of cysteine analogues, mercaptosuccinate and 3-mercaptopropionic acid to sulfinosuccinate and 3-sulfiniopropionic acid, respectively. Human CDO homolog (D) HsADO (2-aminoethanethiol dioxygenase) catalyses oxidation of cysteamine to hypotaurine. (E) Plant cysteine oxidases (PCOs) and (F) ADO catalyse the oxidation of conserved cysteine residues at the N termini of protein/peptides including group VII ethylene response factors (ERF-VIIIs), regulators of G-protein signalling (RGS) 4/5 and interleukin 32 (IL32).

Cross-link formation is proposed to be initiated by  $O_2$  activation at the active site resulting in oxidation of Cys<sub>93</sub> instead of L-Cys (substrate) to generate a thiyl radical, which subsequently oxidises Tyr<sub>157</sub> to form the cross-link [25]. Although the mechanism by which the Cys<sub>93</sub>-Tyr<sub>157</sub> cross-link enhances enzyme kinetics is not definitive, current evidence indicates that it is through stabilisation of the Tyr<sub>157</sub>-OH group, which forms part of a hydrogen-bonding network across the Ser<sub>153</sub>-His<sub>155</sub>-Tyr<sub>157</sub> motif [13,14,18], such that it is at a favourable distance and position to interact with the substrate as well as  $O_2$ , and may play a role in catalysis of substrate oxidation [26,27].

Despite reports of structural, spectroscopic and enzymatic characterisation, the CDO catalytic mechanism is not fully resolved. An electron paramagnetic resonance study of CDO with iron cofactor, NO (substituted for  $O_2$ ) and L-cysteine showed that the cysteine coordinates the active site iron first, followed by molecular  $O_2$  [27]. This study also showed that in the substrate-free form of CDO, Fe(II) is stabilised in

the reduced state in aerobic conditions even in the absence of a reductant [27]. Hence, the high selectivity of substrate binding to the enzyme prior to  $O_2$  binding may be due to the high reduction potential of the substrate-free Fe(II) centre. This order of binding means CDO is likely to act via activation of  $O_2$  rather than activation of substrate. There are currently two proposed mechanisms for thiol oxidation (Scheme 2). The first is based on a crystallographically observed persulfenate species and predicts that, once substrate and  $O_2$  are bound, the thiolate attacks the proximal O coordinating the iron cofactor, resulting in thiadioxirane formation and heterolytic cleavage of the O-O bond [15,17]; this mechanism proposes a role for Tyr<sub>157</sub> (RnCDO numbering) in H-bonding with the proximal O (Scheme 2). The second mechanism is based on theoretical calculations, which suggest that the thiolate is attacked by the distal O, resulting in O-O bond cleavage and a reactive Fe(IV)-oxo species from which the second O is transferred to the substrate (Scheme 2) [28,29]. Kinetic experiments



**Fig. 1.** Structural comparison of a mammalian CDO, a bacterial 3MDO, Arabidopsis PCO4 and human ADO (model). Structures of small molecule TDOs (A) RnCDO (wheat, PDB 2B5H) [60] (B) Pa3MDO (olive, panel I-III, PDB 4TLF) [61] and Av3MDO (deep olive, panel IV, PDB 6XB9) [33] in comparison with NCOs (C) AtPCO4 (deep teal, PDB 6S7E) [5] and (D) HsADO model based on AtPCO4 (PDB 6S7E) generated with MODELLER [56]. Each structure is shown in 4 representations. Panel (I) shows cartoon representation of whole protein showing the cupin fold and 3xHis residues (sticks) coordinating active site metal (brown sphere). Panel (II) shows surface representation highlighting the active site accessibility; note potential substrate binding loops for AtPCO4/HsADO and homologous region in CDO/3MDO (highlighted in red) and a C-terminal chain (orange, green, blue and magenta sticks) which occludes the active site of RnCDO and Pa3MDO but in AtPCO4 and the HsADO model, sits adjacent to the active site creating a large cavity which leads to the active site. In panel (III), active site residues (sticks) and active site metal (brown sphere) are shown. Panel (IV) shows CDO substrate, CYS (cyan sticks) bound to RnCDO (PDB 4IEV) [17], inhibitor 3HPA (orange sticks) bound to Av3MDO (deep olive, PDB 6XB9) [33] as well as the potential substrate binding loop residues (highlighted in red) for AtPCO4 (PDB 6S7E) [5] and the HsADO model. (All four structures in each panel are structurally aligned using PyMol to show a near equivalent view).

attempting to trap and identify catalytic intermediates to resolve the mechanism have not yet been successful: although a catalytically relevant spectroscopically

visible intermediate has been observed, it was short-lived (< 20 ms) and its identity has not yet been experimentally resolved [29].

## Bacterial thiol dioxygenases

TDOs are also found in certain phyla of bacteria, where they are divided into subclasses, 'Arg-type' and 'Gln-type' TDOs [23]. CDOs fall into the Arg-type category, where an active site Arg residue is retained in an equivalent position to Arg<sub>60</sub> in RnCDO to stabilise substrate binding (Fig. 1AIII) [23]. 'Gln-type' TDOs replace this conserved Arg with a Gln residue which is not involved in substrate binding (Fig. 2B). These TDOs in fact appear to preferentially act as 3-mercaptopropionate dioxygenases (3MDOs, Scheme 1), another class of TDO found in bacteria [23,30]. 3MDOs have a conserved Arg elsewhere in the active site (e.g. Av/Pa3MDO Arg<sub>168</sub>; Fig. 1BIII and IV; Fig. 2B) which has been shown to stabilise substrate binding via interaction with its carboxylate group, fulfilling a similar role to RnCDO Arg<sub>60</sub> [18]. Mercaptosuccinate dioxygenase (MSDO, Scheme 1) also has both Gln and Arg residues in the active site; while *Variovorax paradoxus* B4 MSDO Gln<sub>64</sub> characterises this enzyme as a Gln-type TDO, Arg<sub>66</sub> is important in catalysis likely via substrate binding (Fig. 2B) [31]. Substitution of the Tyr<sub>157</sub> equivalent residue to Phe in MSDO did not impact activity, suggesting its OH group is nonessential for mercaptosuccinate oxidation [31]. In line with this observation, in bacterial TDOs (including CDOs) the equivalent Cys is replaced with a Gly, and the Cys-Tyr cross-link is not observed (Fig. 2B) [30]. Neither mercaptosuccinate nor 3-mercaptopropionic acid (3MPA) has an amine group as is seen for the CDO substrate, L-cysteine. Therefore, substrate binding to the active site of these TDOs cannot involve amine as well as thiol coordination to the active site metal. Kinetic and spectroscopic data have indicated that substrate coordination is instead monodentate [30]; however, recent modelling and structural reports, including Av3MDO with a 3-hydroxypropionic acid (3HPA) substrate analogue in the active site, indicate bidentate coordination of substrate via its carboxylate and hydroxyl groups is possible (Fig. 1BIV) [32,33].

## Cysteamine dioxygenase

A second mammalian TDO, capable of oxidising cysteamine to hypotaurine, has been postulated since the 1960s, [34] when an enzyme extracted from animal tissue was found to be specific for cysteamine over cysteine and other cysteine derivatives [34]. This reaction is part of an alternative pathway in taurine biosynthesis and also regulates levels of cysteamine, which is released during coenzyme A degradation. The enzyme,

cysteamine (2-aminoethanethiol) dioxygenase (ADO), is more widely expressed than CDO, with higher levels of expression in the brain and muscles [35] and was first biochemically characterised in 2007 by Dominy *et al.* [35]. They reported that the recombinant ADO protein, like the extracted enzyme, was very specific for the dioxygenation of cysteamine to hypotaurine (Scheme 1) and could not use L-cysteine as a substrate. L-Cysteine was also found to be a very poor inhibitor of ADO's cysteamine dioxygenase activity, as was 2-mercaptoethanol, indicating neither of these alternative thiols bound readily to the active site.

There is no published structure of ADO; its sequence suggests that, like CDO, it belongs to the cupin superfamily of enzymes but, also like CDO, the common glutamate residue usually found in cupin motif 1 is not present, replaced by glycine or valine (Cys<sub>93</sub> in RnCDO) [36]. ADO is therefore expected to have a DBSH, with the active site located within the cupin fold. Sequence alignment of ADO across different organisms and with CDO shows 3 conserved histidine residues (HsADO His<sub>112</sub>, His<sub>114</sub> and His<sub>193</sub>), which are expected to be involved in binding iron via a facial triad. Indeed as for CDO, iron coordination is needed for ADO activity and when iron binding was ablated using a His to Ala variant cysteamine dioxygenase activity was lost [35]. It is also possible that ADO shares another structural feature with CDO: in 2018, Wang *et al.* published evidence that a Cys-Tyr cross-link can occur in ADO between Cys<sub>220</sub> and Tyr<sub>222</sub> [37]. Interestingly, recent spectroscopic investigations indicate that metal-substrate coordination in ADO is distinctive from that of CDO: While crystallographic evidence has shown these enzymes to bind their substrates in a bidentate manner, spectroscopic evidence has indicated that ADO binds cysteamine in a monodentate manner via the thiol but not via its amine group [38,39].

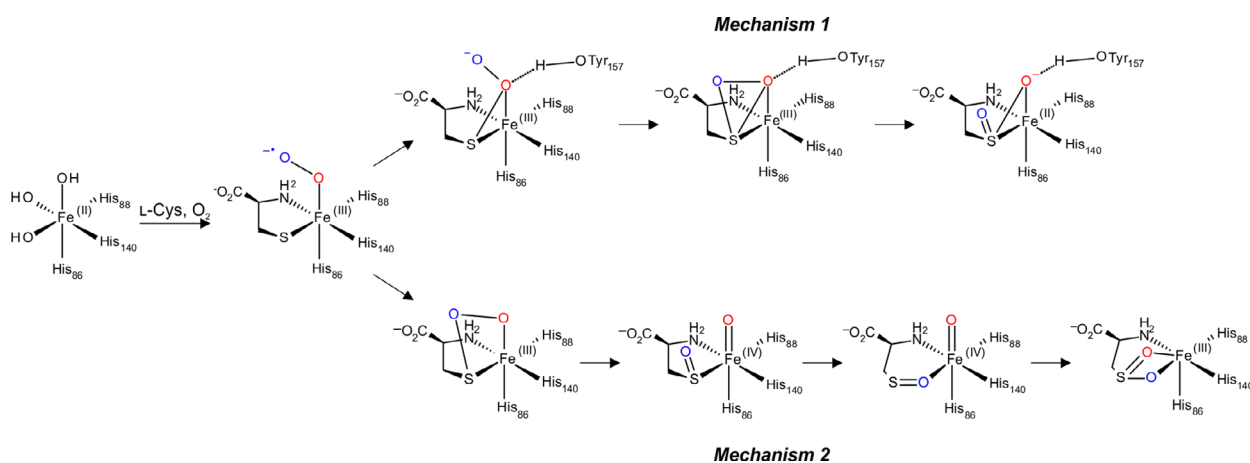
## Thiol dioxygenases as N-terminal cysteine dioxygenases

### Plant cysteine oxidases

In 2011, two reports identified that in plants, cysteine oxidation plays a major role in adaptive responses to the reduced oxygen availability (hypoxia) associated with flooding [40,41]. Group VII ethylene response factors (ERF-VIIs) are transcription factors which upregulate genes that trigger either a metabolic quiescence response or an escape response involving rapid growth above the level of the floodwater. While the ERF-VIIs are stable in these hypoxic conditions, they are



**Fig. 2.** Phylogenetic analysis and alignment of known and predicted secondary structures of TDOs. Thiol dioxygenase sequences from various organisms were analysed using NGPhylogeny.fr one click workflow [62–67]. Sequences (accession, *organism*): AvMDO (WP\_012701646, *Azotobacter vinelandii*); ReCDO (4QMA\_A, *Cupriavidus pinatubonensis*); Pa3MDO (WP\_003113364, *Pseudomonas aeruginosa*); VPMSDO (WP\_062365821.1, *Variovorax paradoxus*); BsCDO (WP\_003243534, *Bacillus subtilis*); HsCDO (AAH24241, *Homo sapiens*), RnCDO (AAH70509, *Rattus norvegicus*); MmCDO (NP\_001005419, *Mus musculus*); ZmPCO2 (PWZ57039, *Zea mays* (maize)); HsADO (NP\_116193, *Homo sapiens*); MmADO (NP\_001005419, *Mus musculus*); AtPCO2 (NP\_198805, *Arabidopsis thaliana*); AtPCO1 (NP\_197016, *Arabidopsis thaliana*); CsPCO1 (XP\_010492336.1, *Camelina sativa*, *camelina*); AtPCO3 (NP\_564055, *Arabidopsis thaliana*); MpPCO (hypothetical protein PTQ35572.1, *Marchantia polymorpha*); OsPCO4 (XP\_015633582, *Oryza sativa* (Japonica group, rice)); SiPCO4 (XP\_004243282, *Solanum lycopersicum* (tomato)); AtPCO4 (NP\_565980, *Arabidopsis thaliana*); AtPCO5 (NP\_001078310, *Arabidopsis thaliana*). A selection of these sequences was aligned using Clustal Omega Multiple Sequence Alignment and subsequently processed using ESPrnt 3. [68] Cartoon representations of secondary structure are overlaid in place of amino acid sequence, showing  $\alpha$  helices in red and  $\beta$  sheets in light blue. Solid black lines represent remaining secondary structure, with grey lines indicated less conservation across sequences. Residues of note are highlighted in dark blue. (A): Phylogenetic tree of all chosen TDO sequences. (B): Alignment of small molecule thiol dioxygenases. Secondary structures were assigned using published structures: RnCDO (PDB 2B5H); Pa3MDO (PDB 4TLF); BsCDO (PDB 4QM9); HsCDO (PDB 2IC1). Secondary structure for VpMSDO was determined using predicted structure produced with MODELLER [56] using RnCDO (PDB 2B5H) as a template (C): alignment of N-terminal cysteinyl dioxygenases. Secondary structure for AtPCO4 was determined using published structure (PDB 6S7E). Secondary structure for remaining NCOs was determined using predicted structures produced with MODELLER [56] using AtPCO4 (PDB 6S7E) as a template.



**Scheme 2.** Proposed catalytic mechanisms for CDO. Coordination of L-cysteine and  $O_2$  to free enzyme-Fe(II) leads to formation of Fe(III)-superoxo intermediate. **Mechanism 1** (top) based on the persulfenate intermediate formation observed in X-ray crystallography [15,17]. Addition of sulfur to the proximal oxygen leads to persulfenate intermediate formation, followed by formation of a thiadioxirane intermediate. Heterolytic O-O cleavage facilitated by H-bonding between the Tyr<sub>157</sub> and the proximal oxygen (red) forms a sulfinate. **Mechanism 2** (bottom) based on theoretical studies [28,29]: the distal oxygen (blue) attacks the sulfur resulting in cleavage of O-O bond and formation of an Fe(IV)-oxo intermediate and a sulfoxide intermediate. Rotation of the sulfoxide bond allows the proximal oxygen to transfer to the sulfoxide to generate the Fe(II)-sulfinate complex.

degraded in normoxia via the Cys/Arg branch of the N-degron pathway. This degradation is initiated by oxidation of the N-terminal Cys of the ERF-VIIs, connecting  $O_2$  availability with the hypoxic response [42]. This is an analogous mechanism to the hypoxic regulation of the hypoxia-inducible factor (HIF), a transcription factor which performs a similar role in triggering hypoxic adaptation in animals [43].

Knowing that HIF stability is controlled by  $O_2$ -sensing HIF hydroxylase enzymes, equivalent enzymatic control of ERF-VII N-terminal Cys oxidation was suspected. By

searching plant genomes for homologues of the small molecule TDOs described above, Weits *et al* [2] reported in 2014 that they had identified 5 enzymes in *Arabidopsis thaliana* that could oxidise ERF-VII N termini and regulate ERF-VII stability; these were termed plant cysteine oxidases (PCOs). Subsequent analyses confirmed that the PCOs catalyse oxidation of the Nt-Cys of ERF-VII to Cys-sulfinic acid (Scheme 1) [40,41] and that these sulfinylated N termini are then arginylated by arginyl-tRNA-protein transferase 1 (ATE1) to generate N-degrons [44–46]. The N-degron pathway then recognises

and ubiquitinates N-degron proteins or fragments (degradation signals) directing them for proteolysis by the 26S proteasome [47]. The PCOs therefore represent a form of TDO that can catalyse oxidation of N-terminal cysteine residues in the context of a protein (N-terminal cysteinyl dioxygenases, NCOs) in contrast to the role described for small molecule TDOs. Since this discovery, PCOs have also been shown to catalyse oxidation of Nt-Cys residues of arabidopsis proteins polycomb group protein vernalisation 2 (VRN2) [48] and LITTLE ZIPPER 2 (ZPR2) [49], expanding the role of these enzymes beyond flood tolerance to developmental signalling processes. Indeed, as plants have hundreds of Nt-Cys-initiating proteins, this suggests the potential for multiple PCO substrates.

Kinetic studies have been performed with AtPCO enzymes towards a peptide representing the Met-excised N terminus of two of the ERF-VII substrates in Arabidopsis (RAP2.2 and RAP2.12). These studies revealed that the catalytic activity among the five AtPCOs varied significantly [3], with one enzyme, AtPCO4, having a significantly higher turnover rate than the other enzymes. AtPCO4 and AtPCO5 have  $k_{\text{cat}}/K_m$  values of 44927 and 49068  $\text{M}^{-1} \text{s}^{-1}$ , respectively, which are at least twice the rates of AtPCOs1-3, indicating these are the most catalytically efficient enzymes [3]. AtPCO4 and AtPCO5 are both constitutively expressed, along with AtPCO3, but in contrast to AtPCO1 and AtPCO2 which are hypoxically upregulated [2,50], potentially indicating a particular biological relevance of AtPCOs 4 and 5. Importantly, given their role in triggering adaptive hypoxic responses, all five AtPCOs showed sensitivity to  $\text{O}_2$  availability as determined by  $K_{\text{M}\text{O}_2}$  values [3]. Both AtPCO1 and 4 have a lower affinity for  $\text{O}_2$  ( $K_{\text{M}\text{O}_2}$  values of 15.7 and 17.3%, respectively) compared with AtPCO2 and AtPCO5 ( $K_{\text{M}\text{O}_2}$  values of 6.57 and 4.83%, respectively) [3]. Although these values are lower than atmospheric  $\text{O}_2$  levels in solution, when compared to reported  $\text{O}_2$  concentrations across plant tissues (< 1%  $\text{O}_2$  to > 15%  $\text{O}_2$ ) [51], these  $K_{\text{M}\text{O}_2}$  values suggest the PCOs have the potential to act as physiologically relevant  $\text{O}_2$  sensors [3]. Combined with apparent differing substrate preferences of AtPCO1/2 and AtPCO4/5 towards ERF-VII substrates and the nature of their expression in plant cells (hypoxia-induced vs. constitutively expressed) [2], the kinetic data suggest the possibility of distinct physiological roles for the five enzymes, such as stabilisation and destabilisation of ERF-VIIs on submergence/desubmergence. Further investigation of the dynamics of PCO activity across  $\text{O}_2$  gradients *in planta* will be of great interest.

To date, crystal structures of three AtPCOs have been reported: PCO4 (PDB 6S0P[5], 6S7E[5] and 7CHJ)[6], PCO5 (PDB 6SBP[5] and 7CHI)[6] and

PCO2 (PDB 7CXZ) [6]. These structures reveal that, similar to other TDOs, the PCOs are members of the cupin superfamily of enzymes, with DBSH core motifs flanked by an N-terminal  $\alpha$ -helical and a C-terminal loop region (Fig. 1CI). Their active sites are formed within the cupin fold where the active site iron is coordinated by 3xHis residues (His<sub>98</sub>, His<sub>100</sub> and His<sub>164</sub>, AtPCO4/5 numbering here and throughout) [5]. In the absence of substrate, the octahedral iron coordination at the iron centre is completed with three water molecules bound in the vacant sites [5]. Beyond the iron-binding residues, the active sites of the PCOs appear quite different to those of the small molecule TDOs (Fig. 1CIII). The Ser-His-Tyr catalytic triad is replaced with Asp-Ile-Leu, while the equivalent residues to the CDO Cys<sub>93</sub>-Tyr<sub>157</sub> cross-link are not present in the PCOs (Fig. 1CIII; Fig. 2). Although there are other potential cross-link sites (Cys<sub>190</sub>, Tyr<sub>182</sub>, Tyr<sub>192</sub>), thioether bond formation has not been identified in PCOs to date [5,37]. Mutagenesis studies have revealed roles for Asp<sub>176</sub> and Tyr<sub>182</sub> in substrate binding and/or catalysis, where conservative substitutions to Asn and Phe (respectively) saw decreased activity compared with the wild-type [5,6]. Curiously, given that 2-His-Asp metal coordination is a common feature in cupin superfamily enzymes [52], an AtPCO4 His164Asp variant lost all activity.

A structure of AtPCO2 with Cys bound in the active site has been modelled based on a structure of AtPCO2 with Tris bound in the active site (PDB 7CXZ) [6]. This model suggested quite different substrate-interacting residues to those seen in CDO, for example, van der Waals interactions between L-Cys and residues Phe<sub>123</sub>, Ile<sub>131</sub> and Phe<sub>199</sub> (AtPCO2 numbering) [6]. Although these residues are conserved in all five isoforms, to date there are no mutagenesis studies reported to validate their role in substrate recognition and/or binding in the active site.

### ADO functions as an N-terminal cysteinyl dioxygenase

Having identified the ability of the AtPCOs to oxidise N-terminal cysteinyl residues of target proteins, examination of human TDO sequences suggested that ADO showed a greater homology towards PCO sequences than towards CDO sequences, particularly with respect to secondary active site residues (Fig. 2). Regulators of G-protein signalling, RGS4 and RGS5, were already known targets of the Cys/Arg branch of the N-degron pathway in humans [45]; ADO was therefore investigated as a potential NCO-type enzyme as ADO-catalysed oxidation of RGS4/5 N-terminal Cys

residues seemed a distinct possibility. In 2019, Masson *et al.* showed that ADO was indeed required for O<sub>2</sub>-dependent degradation of RGS4 and RGS5 [4]. As for the PCOs, ADO was shown to catalyse dioxygenation of the N-terminal cysteine residue of peptides representing the N termini of each of these proteins using both oxygen atoms from molecular oxygen. This activity was not replicated with human CDO. The stability of the pro-inflammatory cytokine Interleukin 32 (IL32) was also found to be regulated by ADO [4], although other Nt-Cys-initiating mammalian proteins were found to be poor ADO substrates, suggesting that ADO may only regulate O<sub>2</sub>-dependent protein degradation in a select group of Nt-Cys-initiating proteins.

Similar to the PCOs, the Nt-Cys oxidation activity catalysed by ADO was shown to be sensitive to O<sub>2</sub> availability, with an apparent  $K_M O_2$  of around 50% O<sub>2</sub> with respect to RGS4 and RGS5 oxidation, a degree of O<sub>2</sub> sensitivity only observed to date for the HIF hydroxylase enzyme prolyl hydroxylase 2 (PHD2) [53]. Kinetic studies indicated that RGS4, RGS5 and IL32 were significantly better substrates for ADO than its small molecule substrate, cysteamine: ADO had a catalytic efficiency ( $k_{cat}/K_M$ ) of 420 M<sup>-1</sup> s<sup>-1</sup> for cysteamine [35], compared to  $k_{cat}/K_M$  values of 164 000, 237 000 and 307 000 M<sup>-1</sup> s<sup>-1</sup> with RGS4, RGS5 and IL32 peptides, respectively. The ability of cysteamine and L-cysteine to compete with turnover of the RGS5 peptide was determined and IC<sub>50</sub> values reported as 37.65 mM and 13.63 mM, respectively. Although these values are high, they were determined using a high RGS5<sub>2-15</sub> substrate concentration (320 μM) to give maximal ADO activity ( $K_M^{app} = 71.5 \mu M$ ). Assuming cysteine and cysteamine act as competitive inhibitors, it is therefore possible that they could compete with the NCO activity of ADO under physiological conditions with lower concentrations of Nt-Cys-initiating substrates (the normal concentration of cysteine in human cells is 80–200 μM [54] while cysteamine is likely less than this [55]) [38,39]. Although the kinetic data therefore suggest that the primary biological role of ADO is hypoxic regulation of Nt-Cys-initiating protein stability, the role of ADO in catalysing cysteamine oxidation likely remains biologically relevant, dependent on the relative concentrations of cysteamine and Nt-Cys-initiating substrates in cells.

Sequence similarity between ADO and PCOs enabled us to predict a structure of ADO using MODELLER [56] with AtPCO4 (PDB 6S7E) [5] as a template (Fig. 1D). The structure predicted had 93% of residues in the Ramachandran favoured region, with just 1.49% Ramachandran outliers and 2.65% rotamer outliers. This predicted structure featured the expected

distinctive DBSH and active site iron cofactor coordination via a facial triad of histidine residues (Fig. 1DI). Other aspects of the AtPCO4 active site are also predicted to be conserved in the ADO model, including a motif comprising Phe<sub>202</sub>, Asp<sub>204</sub> and Leu<sub>206</sub>, similar to the Ile<sub>174</sub>-Asp<sub>176</sub>-Leu<sub>178</sub> triad seen in the PCOs (Fig. 1DIII). The predicted structure of ADO also featured a hairpin loop near the active site opening (residues 212–220), similar to a hairpin loop observed in the PCO structures (Fig. 1DIV) [5]. Intriguingly, a recent spectroscopic study suggests that ADO binds RGS5 peptide in a monodentate manner through coordination to the Fe(II) by the thiol of the Nt-Cys, similar to the proposed binding mode for cysteamine described above [39].

## Two subclasses of thiol dioxygenase

The structural and functional analyses of PCOs and ADO have revealed clear differences to the related TDOs, CDOs, MSDO and MPDO. These suggest that TDOs can be divided into two subclasses depending on the thiol substrate: protein or small molecule. This is seen upon phylogenetic analysis of their sequences (Fig. 2A), and available crystal structures of these two subclasses of TDO also appear to show different intrinsic structural features which may fulfil their requirements for substrate binding (Fig. 1).

The Ser-His-Tyr motif seen in small molecule TDOs (Fig. 1AIII and BIII; Fig. 2B), known as the ‘catalytic triad’, forms a H-bond network adjacent to the active site iron. This motif is conserved in all small molecule thiol dioxygenases and facilitates substrate binding by the Tyr-OH group [14,15,33]. This motif is not conserved in AtPCOs or ADO and is replaced by Ile-Asp-Leu residues at equivalent positions, or Phe-Asp-Leu in ADO [5]. This alternative motif is conserved in all AtPCOs as well as PCO sequences from a wide range of other plants, although Leu is a common substitution for Ile (Fig. 2C) [5]. The precise role of the alternative triad in NCOs is not yet clear, but an Asp176Asn variant in AtPCO4 resulted in near-ablated activity. If, as reported for ADO, substrate coordination of NCOs is monodentate to the active site metal [38], it is possible that the conserved Asp in the NCO form of this triad could play a role in binding the N-terminal amine of their polypeptide substrates. Tyr<sub>157</sub> plays a particularly important role in CDO substrate binding, so although this role could be fulfilled by other Tyr residues in the active sites of NCOs, distinct modes of primary substrate (and/or O<sub>2</sub>) binding between the two classes of TDO seem likely, especially as there is no equivalent in NCOs to

the interaction between Arg residues and substrate carboxylate groups seen in CDOs, 3MDOs and MSDOs. The PCOs do, however, appear to share a common feature with CDOs, which are a *cis*-peptide bond following the catalytic triad: this comprises Ala<sub>179</sub>-Pro<sub>180</sub> and Ser<sub>179</sub>-Pro<sub>180</sub> in AtPCO4 and AtPCO5, respectively, and Ser<sub>158</sub>-Pro<sub>159</sub> in RnCDO (Fig. 2). This may prove to be important in aligning substrate for oxidation; NCO:substrate structures will be invaluable in understanding these and other active site interactions.

As PCO:substrate interactions are likely to occur across a greater surface area of the enzyme than is the case for small molecule TDOs, it was interesting to find that the position of a C-terminal loop in the AtPCO substrate-free structures creates an open tunnel-like cavity leading to the active site (Fig. 1CII) [5]. Our modelled ADO structure also indicates similar accessibility to the active site (Fig. 1DII). In contrast, the equivalent C-terminal loop in small molecule TDO enzymes sits across this active site face, rendering access to the active site more hindered (Fig. 1AII, BII). It could be that the more accessible active sites seen for the NCOs may enable binding of a protein substrate, whereas the more restricted enzymes are better suited for small molecule thiol substrate binding. Situated at the entrance to this proposed active site cavity in the AtPCO4 structure (PDB 6S0P) is a hairpin loop ( $\beta$ 9- $\beta$ 10 loop, residues 182-190, AtPCO4 numbering) consisting of charged (Glu<sub>185</sub>, Asp<sub>187</sub> and Arg<sub>188</sub>) and polar (Tyr<sub>182</sub>, Ser<sub>183</sub>, Ser<sub>184</sub> and Cys<sub>190</sub>) residues [5] (Fig. 1CIV). The predicted structure of ADO also features a hairpin loop near the active site opening (residues 212-220; Fig. 1DIV, Fig. 2C), similar to the hairpin loops observed in the PCO structures and rich in aspartate residues [5]. This region is noticeably variable among NCO sequences (Fig. 2C) and, given its location, may play a role in substrate recognition and/or binding [5].

### An emerging role for N-terminal cysteine dioxygenases in oxygen sensing

A particularly interesting feature of the NCOs identified to date is their capacity to act as O<sub>2</sub> sensors. For enzymes to act as O<sub>2</sub> sensors, their rate of activity must be sensitive to variations in O<sub>2</sub> that are physiologically relevant. Certainly, AtPCO  $K_M$ O<sub>2</sub> values ranging from 5 to 17% are in line with this [3], while the high  $K_M$ O<sub>2</sub> value of ADO towards RGS4/5 (50%) shows a similar O<sub>2</sub> sensitivity to the well-established O<sub>2</sub>-sensing enzyme in humans, PHD2 [4]. While the biological consequences of ADO O<sub>2</sub> sensitivity are

unknown and do not appear to regulate a transcription factor cascade [57], this nevertheless implies that there will be a rapid stabilisation of its substrates in hypoxic conditions that may be important in relevant disease conditions such as ischaemia or in hypoxic tumour cores. The sequence homology between ADO from the animal kingdom and the PCOs from the plant kingdom implies that these O<sub>2</sub> sensors could have arisen from a common evolutionary ancestor, earlier than the PHD/HIF mediated O<sub>2</sub>-sensing system, which is only present in metazoans [58]. Our current understanding of NCO/PHD-mediated O<sub>2</sub> sensing is centred on their ability to regulate transcription factors that evolved in complex eukaryotes, as is the case for both HIF and ERF-VIIs. It will be of interest to uncover substrates for these enzymes in simpler organisms which may enlighten our understanding of how O<sub>2</sub>-sensing capacity has driven adaptation to different environments [57].

It is important to note that, although there has been no connection identified between CDO, 3MDO or MSDO activity and O<sub>2</sub>-sensitive responses, the kinetics of their activity specifically with respect to O<sub>2</sub> availability has not been determined. Such knowledge could help elucidate whether all TDOs, or just the NCO subset, are responsive to physiologically relevant levels of hypoxia. Whether or not this is the case, the O<sub>2</sub>-sensitive regulation of protein stability via NCO activity is an important new pathway in hypoxia biology and common structural features could underpin this property. Although the mechanistic rationale can only be speculated at present, O<sub>2</sub> sensitivity may be conferred by active site features regulating the rate at which O<sub>2</sub> binds and is activated by iron or through a restricted rate of O<sub>2</sub> delivery to the active site, which in turn could be impacted by the nature of the enzyme:substrate interaction and accessibility. Further mechanistic and kinetic assays will reveal how the NCOs control their rate of reaction with O<sub>2</sub>, with such knowledge potentially enabling rational manipulation of plant PCOs to alter their O<sub>2</sub> sensitivity and with it, their capacity to respond to flood-induced hypoxia.

### Conclusions

We have described the recent characterisation of a new class of thiol dioxygenase, the N-terminal cysteinyl dioxygenases (NCOs), which have an O<sub>2</sub>-sensing role in regulating protein stability via the N-degron pathway. While there are many common features, there are distinct structural differences between the NCOs and small molecule thiol dioxygenases characterised to date, particularly in the active site. Some of these

structural differences are likely due to their requirement to bind protein versus small molecule substrates.

The common features of the NCOs, particularly at the active site, in combination with their O<sub>2</sub>-sensing function, raise the possibility of being able to predict whether other organisms may have such O<sub>2</sub>-sensing enzymes. The existence of such sequences in early land plants (Fig. 2C) and even algae is intriguing [57], particularly as known substrates of the PCOs have evolved with flowering and seeding plants. Investigating NCOs from distantly related organisms could reveal new O<sub>2</sub>-sensing functions in a range of biological niches.

Although there is much work to be done to correlate the structural and functional features of the NCOs, particularly with respect to O<sub>2</sub>-sensing, so doing could reveal opportunities for targeted manipulation of their activity. Such manipulation, either via pharmaceutical, agrochemical or genetic intervention, could impact on biological responses to hypoxia. Ultimately, this may be of use from a therapeutic perspective in humans or from an agronomic perspective where reduced PCO activity could stabilise ERF-VII levels and enhance adaptive responses to flooding [59].

## Acknowledgements

DMG and EF were supported by the European Research Council (European Union's Horizon 2020 Research and Innovation Programme Grant 864888), and KCH was supported by the Oxford-GSK-Crick Doctoral Programme in Chemical Biology via the Systems Approaches to Biomedical Science CDT EPSRC Grant Reference: EP/S513866/1 and GlaxoSmithKline. Graphical abstract created with BioRender.com

## Conflict of interest

The authors declare no conflict of interest.

## Author contributions

All authors contributed to writing and editing the manuscript.

## References

- 1 Stipanuk MH, Simmons CR, Andrew Karplus P & Dominy JE (2011) Thiol dioxygenases: unique families of cupin proteins. *Amino Acids* **41**, 91–102.
- 2 Weits DA, Giuntoli B, Kosmacz M, Parlanti S, Hubberten H-M, Riegler H, Hoefgen R, Perata P, van Dongen JT & Licausi F (2014) Plant cysteine oxidases control the oxygen-dependent branch of the N-end-rule pathway. *Nat Commun* **5**, 3425.
- 3 White MD, Kamps JJAG, East S, Taylor Kearney LJ & Flashman E (2018) The plant cysteine oxidases from *Arabidopsis thaliana* are kinetically tailored to act as oxygen sensors. *J Biol Chem* **293**, 11786–11795.
- 4 Masson N, Keeley TP, Giuntoli B, White MD, Puerta ML, Perata P, Hopkinson RJ, Flashman E, Licausi F & Ratcliffe PJ (2019) Conserved N-terminal cysteine dioxygenases transduce responses to hypoxia in animals and plants. *Science* **365**, 65–69.
- 5 White MD, Dalle Carbonare L, Lavilla Puerta M, Iacopino S, Edwards M, Dunne K, Pires E, Levy C, McDonough MA, Licausi F *et al.* (2020) Structures of *Arabidopsis thaliana* oxygen-sensing plant cysteine oxidases 4 and 5 enable targeted manipulation of their activity. *Proc Natl Acad Sci* **117**, 23140–23147.
- 6 Chen Z, Guo Q, Wu G, Wen J, Liao S & Xu C (2021) Molecular basis for cysteine oxidation by plant cysteine oxidases from *Arabidopsis thaliana*. *J Struct Biol* **213**, 107663.
- 7 Lombardini JB, Singer TP & Boyer PD (1969) Cysteinase oxygenase. II. Studies on the mechanism of the reaction with 18oxygen. *J Biol Chem* **244**, 1172–1175.
- 8 Sakakibara S, Yamaguchi K, Hosokawa Y, Kohashi N, Ueda I & Sakamoto Y (1976) Purification and some properties of rat liver cysteine oxidase (cysteine dioxygenase). *Biochim Biophys Acta - Enzymol* **422**, 273–279.
- 9 Heafield MT, Fearn S, Steventon GB, Waring RH, Williams AC & Sturman SG (1990) Plasma cysteine and sulphate levels in patients with motor neurone, Parkinson's and Alzheimer's disease. *Neurosci Lett* **110**, 216–220.
- 10 Stipanuk MH, Ueki I, Dominy JE Jr, Simmons CR & Hirschberger LL (2009) Cysteine dioxygenase: a robust system for regulation of cellular cysteine levels. *Amino Acids* **37**, 55–63.
- 11 Ueki I, Roman HB, Valli A, Fieselmann K, Lam J, Peters R, Hirschberger LL & Stipanuk MH (2011) Knockout of the murine cysteine dioxygenase gene results in severe impairment in ability to synthesize taurine and an increased catabolism of cysteine to hydrogen sulfide. *Am J Physiol - Endocrinol Metab* **301**, E668–E684.
- 12 Clifton IJ, McDonough MA, Ehrismann D, Kershaw NJ, Granatino N & Schofield CJ (2006) Structural studies on 2-oxoglutarate oxygenases and related double-stranded  $\beta$ -helix fold proteins. *J Inorg Biochem* **100**, 644–669.
- 13 McCoy JG, Bailey LJ, Bitto E, Bingman CA, Aceti DJ, Fox BG & Phillips GN Jr (2006) Structure and mechanism of mouse cysteine dioxygenase. *Proc Natl Acad Sci USA* **103**, 3084–3089.
- 14 Simmons CR, Liu Q, Huang Q, Hao Q, Begley TP, Karplus PA & Stipanuk MH (2006) Crystal structure of

- mammalian cysteine dioxygenase: a novel mononuclear iron center for cysteine thiol oxidation. *J Biol Chem* **281**:18723–18733.
- 15 Simmons CR, Krishnamoorthy K, Granett SL, Schuller DJ, Dominy JE, Begley TP, Stipanuk MH & Karplus PA (2008) A putative Fe<sup>2+</sup> bound persulfenate intermediate in cysteine dioxygenase. *Biochemistry* **47**, 11390–11392.
- 16 Ye S, Wu X, Wei L, Tang D, Sun P, Bartlam M & Rao Z (2007) An insight into the mechanism of human cysteine dioxygenase: Key roles of the thioether-bonded tyrosine-cysteine cofactor. *J Biol Chem* **282**, 3391–3402.
- 17 Driggers CM, Cooley RB, Sankaran B, Hirschberger LL, Stipanuk MH & Karplus PA (2013) Cysteine dioxygenase structures from pH4 to 9: consistent cys-persulfenate formation at intermediate pH and a cys-bound enzyme at higher pH. *J Mol Biol* **425**, 3121–3136.
- 18 Li W, Blaesi EJ, Pecore MD, Crowell JK & Pierce BS (2013) Second-sphere interactions between the C93–Y157 cross-link and the substrate-bound Fe site influence the O<sub>2</sub> coupling efficiency in mouse cysteine dioxygenase. *Biochemistry* **52**, 9104–9119.
- 19 Dominy JE Jr, Hwang J, Guo S, Hirschberger LL, Zhang S & Stipanuk MH (2008) Synthesis of amino acid cofactor in cysteine dioxygenase is regulated by substrate and represents a novel post-translational regulation of activity. *J Biol Chem* **283**, 12188–12201.
- 20 Njeri CW & Ellis HR (2014) Shifting redox states of the iron center partitions CDO between crosslink formation or cysteine oxidation. *Arch Biochem Biophys* **558**, 61–69.
- 21 Ito N, Phillips SEV, Stevens C, Ogel ZB, McPherson MJ, Keen JN, Yadav KDS & Knowles PF (1991) Novel thioether bond revealed by a 1.7 Å crystal structure of galactose oxidase. *Nature* **350**, 87–90.
- 22 Dominy JE, Simmons CR, Karplus PA, Gehring AM & Stipanuk MH (2006) Identification and characterization of bacterial cysteine dioxygenases: A new route of cysteine degradation for eubacteria. *J Bacteriol* **188**, 5561–5569.
- 23 Driggers CM, Hartman SJ & Karplus PA (2015) Structures of Arg- and Gln-type bacterial cysteine dioxygenase homologs. *Protein Sci* **24**, 154–161.
- 24 Forbes DL, Meneely KM, Chilton AS, Lamb AL & Ellis HR (2020) The 3-His metal coordination site promotes the coupling of oxygen activation to cysteine oxidation in cysteine dioxygenase. *Biochemistry* **59**, 2022–2031.
- 25 Li J, Koto T, Davis I & Liu A (2019) Probing the Cys-Tyr cofactor biogenesis in cysteine dioxygenase by the genetic incorporation of fluorotyrosine. *Biochemistry* **58**, 2218–2227.
- 26 Driggers CM, Kean KM, Hirschberger LL, Cooley RB, Stipanuk MH & Karplus PA (2016) Structure-based insights into the role of the Cys-Tyr crosslink and inhibitor recognition by mammalian cysteine dioxygenase. *J Mol Biol* **428**, 3999–4012.
- 27 Pierce BS, Gardner JD, Bailey LJ, Brunold TC & Fox BG (2007) Characterization of the nitrosyl adduct of substrate-bound mouse cysteine dioxygenase by electron paramagnetic resonance: electronic structure of the active site and mechanistic implications. *Biochemistry* **46**, 8569–8578.
- 28 Kumar D, Thiel W & de Visser SP (2011) Theoretical study on the mechanism of the oxygen activation process in cysteine dioxygenase enzymes. *J Am Chem Soc* **133**, 3869–3882.
- 29 Tchesnokov EP, Faponle AS, Davies CG, Quesne MG, Turner R, Fellner M, Souness RJ, Wilbanks SM, de Visser SP & Jameson GNL (2016) An iron–oxygen intermediate formed during the catalytic cycle of cysteine dioxygenase. *Chem Commun* **52**, 8814–8817.
- 30 Pierce BS, Subedi BP, Sardar S & Crowell JK (2015) The “Gln-Type” thiol dioxygenase from *Azotobacter vinelandii* is a 3-mercaptopropionic acid dioxygenase. *Biochemistry* **54**, 7477–7490.
- 31 Brandt U, Galant G, Meinert-Berning C & Steinbüchel A (2019) Functional analysis of active amino acid residues of the mercaptosuccinate dioxygenase of *Variovorax paradoxus* B4. *Enzyme Microb Technol* **120**, 61–68.
- 32 Aloï S, Davies CG, Karplus PA, Wilbanks SM & Jameson GNL (2019) Substrate specificity in thiol dioxygenases. *Biochemistry* **58**, 2398–2407.
- 33 York NJ, Lockart MM, Sardar S, Khadka N, Shi W, Stenkamp RE, Zhang J, Kiser PD & Pierce BS (2021) Structure of 3-mercaptopropionic acid dioxygenase with a substrate analog reveals bidentate substrate binding at the iron center. *J Biol Chem* **296**, 100492.
- 34 Cavallini D, Scandurra R & Demarco C (1963) The enzymatic oxidation of cysteamine to hypotaurine in the presence of sulfide. *J Biol Chem* **238**, 2999–3005.
- 35 Dominy JE, Simmons CR, Hirschberger LL, Hwang J, Coloso RM & Stipanuk MH (2007) Discovery and characterization of a second mammalian thiol dioxygenase, cysteamine dioxygenase. *J Biol Chem* **282**, 25189–25198.
- 36 Sarkar B, Kulharia M & Mantha AK (2017) Understanding human thiol dioxygenase enzymes: structure to function, and biology to pathology. *Int J Exp Pathol* **98**, 52–66.
- 37 Wang Y, Griffith WP, Li J, Koto T, Wherritt DJ, Fritz E & Liu A (2018) Cofactor biogenesis in cysteamine dioxygenase: C–F bond cleavage with genetically incorporated unnatural tyrosine. *Angew Chemie Int Ed* **57**, 8149–8153.
- 38 Fernandez RL, Dillon SL, Stipanuk MH, Fox BG & Brunold TC (2020) Spectroscopic Investigation of Cysteamine Dioxygenase. *Biochemistry* **59**, 2450–2458.

- 39 Wang Y, Davis I, Chan Y, Naik SG, Griffith WP & Liu A (2020) Characterization of the nonheme iron center of cysteamine dioxygenase and its interaction with substrates. *J Biol Chem* **295**, 11789–11802.
- 40 Licausi F, Kosmacz M, Weits DA, Giuntoli B, Giorgi FM, Voeselek LACJ, Perata P & van Dongen JT (2011) Oxygen sensing in plants is mediated by an N-end rule pathway for protein destabilization. *Nature* **479**, 419–422.
- 41 Gibbs DJ, Lee SC, Md Isa N, Gramuglia S, Fukao T, Bassel GW, Correia CS, Corbineau F, Theodoulou FL, Bailey-Serres J *et al.* (2011) Homeostatic response to hypoxia is regulated by the N-end rule pathway in plants. *Nature* **479**, 415–418.
- 42 Bailey-Serres J, Fukao T, Gibbs DJ, Holdsworth MJ, Lee SC, Licausi F, Perata P, Voeselek LACJ & van Dongen JT (2012) Making sense of low oxygen sensing. *Trends Plant Sci* **17**, 129–138.
- 43 Schofield CJ & Ratcliffe PJ (2004) Oxygen sensing by HIF hydroxylases. *Nat Rev Mol Cell Biol* **5**, 343–354.
- 44 White MD, Klecker M, Hopkinson RJ, Weits DA, Mueller C, Naumann C, O'Neill R, Wickens J, Yang J, Brooks-Bartlett JC *et al.* (2017) Plant cysteine oxidases are dioxygenases that directly enable arginyl transferase-catalysed arginylation of N-end rule targets. *Nat Commun* **8**, 14690.
- 45 Lee MJ, Tasaki T, Moroi K, An JY, Kimura S, Davydov IV & Kwon YT (2005) RGS4 and RGS5 are in vivo substrates of the N-end rule pathway. *Proc Natl Acad Sci* **102**, 15030–15035.
- 46 Graciet E, Mesiti F & Wellmer F (2010) Structure and evolutionary conservation of the plant N-end rule pathway. *Plant J* **61**, 741–751.
- 47 Varshavsky A. (2019) N-degron and C-degron pathways of protein degradation. *Proc Natl Acad Sci* **116**, 358–366.
- 48 Gibbs DJ, Tedds HM, Labandera A-M, Bailey M, White MD, Hartman S, Sprigg C, Mogg SL, Osborne R & Dambire C (2018) Oxygen-dependent proteolysis regulates the stability of angiosperm polycomb repressive complex 2 subunit VERNALIZATION 2. *Nat Commun* **9**, 1–11.
- 49 Weits DA, Kunkowska AB, Kamps NCW, Portz KMS, Packbier NK, Nemeck VENZA Z, Gaillochet C, Lohmann JU, Pedersen O, van Dongen JT *et al.* (2019) An apical hypoxic niche sets the pace of shoot meristem activity. *Nature* **569**, 714–717.
- 50 Weits DA, Zhou L, Giuntoli B, Carbonare LD, Iacopino S, Piccinini L, Shukla V, Bui LT, Novi G, van Dongen JT *et al.* (2020) Acquisition of hypoxia inducibility by oxygen sensing N-terminal cysteine oxidase in spermatophytes. *bioRxiv* **24**, 169417.
- 51 Van Dongen JT & Licausi F (2015) Oxygen sensing and signaling. *Annu Rev Plant Biol* **66**, 345–367.
- 52 Stipanuk MH, Simmons CR, Karplus PA & Dominy JE Jr (2011) Thiol dioxygenases: unique families of cupin proteins. *Amino Acids* **41**, 91–102.
- 53 Tarhonskaya H, Chowdhury R, Leung IKH, Loik ND, McCullagh JSO, Claridge TDW, Schofield CJ & Flashman E (2014) Investigating the contribution of the active site environment to the slow reaction of hypoxia-inducible factor prolyl hydroxylase domain 2 with oxygen. *Biochem J* **463**, 363–372.
- 54 Stipanuk MH, Dominy JE, Lee J-I & Coloso RM (2006) Mammalian cysteine metabolism: new insights into regulation of cysteine metabolism. *J Nutr* **136**, 1652S–1659S.
- 55 Dominy JE Jr, Hirschberger LL, Coloso RM & Stipanuk MH (2006) Regulation of cysteine dioxygenase degradation is mediated by intracellular cysteine levels and the ubiquitin–26 S proteasome system in the living rat. *Biochem J* **394**, 267–273.
- 56 Šali A & Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779–815.
- 57 Hammarlund EU, Flashman E, Mohlin S & Licausi F (2020) Oxygen-sensing mechanisms across eukaryotic kingdoms and their roles in complex multicellularity. *Science* **370**, eaba3512.
- 58 Holdsworth MJ & Gibbs DJ (2020) Comparative biology of oxygen sensing in plants and animals. *Curr Biol* **30**, R362–R369.
- 59 Pan J, Sharif R, Xu X & Chen X (2021) Mechanisms of waterlogging tolerance in plants: research progress and prospects. *Front Plant Sci* **11**, 2319.
- 60 Simmons CR, Hao Q & Stipanuk MH (2005) Preparation, crystallization and X-ray diffraction analysis to 1.5 Å resolution of rat cysteine dioxygenase, a mononuclear iron enzyme responsible for cysteine thiol oxidation. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **61**, 1013–1016.
- 61 Tchesnokov EP, Fellner M, Siakkou E, Kleffmann T, Martin LW, Aloï S, Lamont IL, Wilbanks SM & Jameson GNL (2015) The cysteine dioxygenase homologue from *Pseudomonas aeruginosa* is a 3-mercaptopyruvate dioxygenase. *J Biol Chem* **290**, 24424–24437.
- 62 Lemoine F, Correia D, Lefort V, Doppelt-Azeroual O, Mareuil F, Cohen-Boulakia S & Gascuel O (2019) NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Res* **47**, W260–W265.
- 63 Criscuolo A & Gribaldo S (2010) BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol Biol* **10**, 210.
- 64 Lemoine F, Domelevo Entfellner J-B, Wilkinson E, Correia D, Dávila Felipe M, De Oliveira T & Gascuel

- O (2018) Renewing Felsenstein's phylogenetic bootstrap in the era of big data. *Nature* **556**, 452–456.
- 65 Katoh K & Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* **30**, 772–780.
- 66 Junier T & Zdobnov EM (2010) The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* **26**, 1669–1670.
- 67 Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W & Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307–321.
- 68 Robert X & Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320–W324.