








ARTICLE

<https://doi.org/10.1038/s41467-019-12711-7>

OPEN

Aspartate/asparagine- β -hydroxylase crystal structures reveal an unexpected epidermal growth factor-like domain substrate disulfide pattern

Inga Pfeffer¹, Lennart Brewitz¹, Tobias Krojer ², Sacha A. Jensen ³, Grazyna T. Kochan², Nadia J. Kershaw ¹, Kirsty S. Hewitson¹, Luke A. McNeill¹, Holger Kramer⁴, Martin Münzel¹, Richard J. Hopkinson ¹, Udo Oppermann ^{2,5}, Penny A. Handford³, Michael A. McDonough ¹ & Christopher J. Schofield ^{1*}

AspH is an endoplasmic reticulum (ER) membrane-anchored 2-oxoglutarate oxygenase whose C-terminal oxygenase and tetratricopeptide repeat (TPR) domains present in the ER lumen. AspH catalyses hydroxylation of asparaginyl- and aspartyl-residues in epidermal growth factor-like domains (EGFDs). Here we report crystal structures of human AspH, with and without substrate, that reveal substantial conformational changes of the oxygenase and TPR domains during substrate binding. Fe(II)-binding by AspH is unusual, employing only two Fe(II)-binding ligands (His679/His725). Most EGFD structures adopt an established fold with a conserved Cys1-3, 2-4, 5-6 disulfide bonding pattern; an unexpected Cys3-4 disulfide bonding pattern is observed in AspH-EGFD substrate complexes, the catalytic relevance of which is supported by studies involving stable cyclic peptide substrate analogues and by effects of Ca(II) ions on activity. The results have implications for EGFD disulfide pattern processing in the ER and will enable medicinal chemistry efforts targeting human 2OG oxygenases.

¹Chemistry Research Laboratory, University of Oxford, Mansfield Road, Oxford OX1 3TA, UK. ²Structural Genomics Consortium, University of Oxford, Old Road Campus, Roosevelt Drive, Headington OX3 7DQ, UK. ³Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK. ⁴Department of Physiology, Anatomy and Genetics, University of Oxford, South Parks Road, Oxford OX1 3QX, UK. ⁵NDORMS, Botnar Research Centre, University of Oxford, Old Road, Oxford OX3 7LD, UK. *email: christopher.schofield@chem.ox.ac.uk

The discovery that specific asparagine- and aspartate-residues in epidermal growth factor-like domains (EGFDs) undergo hydroxylation was a landmark in post-translational modification (PTM) research, because it demonstrated proteins other than collagen type domains are directly modified by O₂ (Fig. 1a)^{1,2}. EGFD hydroxylation is catalysed by the aspartate/asparagine- β -hydroxylase (AspH, BAH), a non-haem ferrous iron and 2-oxoglutarate (2OG) oxygenase^{3–5}, which produces succinate and carbon dioxide as co-products. Human ASPH (NCBI ID: 444 [<https://www.ncbi.nlm.nih.gov/gene/?term=444>], A β H-J-J locus) has two promoters and undergoes extensive alternative splicing^{6,7} resulting in >10 AspH isoforms (Supplementary Fig. 1). An ASPH mutation resulting in the R735W substitution correlates with Traboulsi syndrome (OMIM 601552/refSNP rs374385878), which manifests as facial dysmorphism and lens dislocation⁸; another mutation (G434V) affects human kidney function resulting in vesicoureteral reflux⁹. AspH and its truncated, likely non-catalytic, isoform Humbug are overexpressed in cancers^{10–14}; AspH is strongly upregulated by hypoxia¹⁵, consistent with its clinical use as a tumour biomarker^{16,17}. Loss of murine EGFD Asp/Asn β -hydroxylation correlates with increased tumour incidence and developmental defects similar to those caused by disrupted Notch signalling¹⁸. AspH localises to the endoplasmic reticulum (ER) in normal cells, but (at least part of AspH) localises to the surface of tumour cells^{16,17} and its hydroxylase activity was reported to enhance cell migration^{19,20}.

EGFDs are ~40 residue domains often present in membrane protein extracellular regions; their structures are primarily comprised of β -strands stabilised by 3 conserved internal disulfides, which, as revealed by multiple crystal and NMR structures, normally adopt a thermodynamically preferred canonical Cys1–3, 2–4, 5–6 disulfide pattern^{21–23}. This disulfide pattern is a functional requirement for some EGFDs²⁴ and for Ca(II) ion binding^{25–27}. AspH catalyses (3R)-hydroxylation of specific Asp/Asn residues in EGFDs, some of which are involved in binding Ca(II) in medically relevant proteins, e.g. as in coagulation factors, Notch, Jagged, and Fibrillins^{28,29}.

There are about 60 human 2OG oxygenases with a variety of functional roles including in collagen biosynthesis, fatty acid metabolism/carnitine biosynthesis, nucleic acid repair/modification, O₂ sensing, and transcriptional/translational regulation³⁰. Some 2OG oxygenases are therapeutic targets, including those involved in carnitine biosynthesis and the hypoxia-inducible factor hydroxylases³¹. Structurally informed sequence alignments reveal most 2OG oxygenases possess a conserved His-Xxx-Asp/Glu...His Fe(II)-binding motif located on β -strands II and VII of the double-stranded β -helix (DSBH) core fold³². By contrast, sequences of AspH indicate a His-Xxx-Gly...His Fe(II)-binding motif³³, though it has been unclear if an additional residue from another part of AspH could be involved in Fe(II) binding. Despite its importance in healthy biology and cancer, few biochemical studies and no efficient substrates of isolated recombinant AspH have been reported to date.

Here, we report crystallographic and biochemical results on the structure and substrate selectivity of AspH. The results reveal a key role for the TPR domain of AspH in substrate binding and, unexpectedly, that the preferred EGFD-substrates of AspH have a non-canonical (Cys3–4) disulfide pattern.

Results

AspH oxygenase domain structure. As preceded in work with other multi-domain 2OG oxygenases^{34,35}, we initially prepared the isolated 2OG oxygenase domain of AspH for studies on catalysis. Recombinant N-terminally hexa-His-tagged (His₆)

AspH_{562–758} (AspH-Ox) was prepared in *Escherichia coli* (Supplementary Fig. 2) and tested for catalytic activity using the N-terminal EGFD of human coagulation factor X (hFX EGF1_{39mer}; aa 86–124, Fig. 1b), a known cellular AspH substrate^{2,36}, which was prepared by solid-phase peptide synthesis followed by thiol oxidation in air-saturated buffer to give disulfides. However, no hydroxylation of hFX EGF1_{39mer} was detected by mass spectrometry (MS). Using a [1-¹⁴C]-2OG based assay, we observed AspH-Ox catalysed turnover of 2OG to CO₂ in the absence of substrate, demonstrating oxygenase activity, albeit uncoupled from substrate hydroxylation as preceded with some other 2OG oxygenases under certain conditions (Supplementary Fig. 3).

To investigate the predicted unusual Fe(II)-binding motif of AspH, we initiated crystallographic studies. A structure of the AspH oxygenase domain (AspH-Ox) was determined with Ni(II), substituting for Fe(II), and with L-malate (from the crystallisation buffer) bound in the predicted 2OG binding pocket (Supplementary Fig. 4). As expected, the AspH-Ox structure is comprised of 8 β -strands forming a double-stranded beta-helix (DSBH) fold, with its active site located at the more open end of the β -sandwich. In contrast to most 2OG oxygenases, the active site metal is only coordinated by two histidine residues (His679, His725), located on DSBH strands II and VII. Two water molecules and a malate molecule complete octahedral coordination of the metal ion (Supplementary Fig. 4). Overall, the AspH-Ox structure validates the expected DSBH fold with intact Fe(II) and 2OG-binding sites, consistent with the observed 2OG turnover. To further investigate the lack of EGFD substrate hydroxylation by the isolated AspH-Ox domain, we produced a longer AspH construct containing the AspH-Ox and a predicted N-terminal tetratricopeptide repeat (TPR) domain.

AspH oxygenase and TPR domain structure and activity. In contrast to AspH-Ox, the longer AspH_{315–758} construct containing both 2OG oxygenase and TPR domains (AspH-TPR-Ox, Supplementary Fig. 5) catalysed hydroxylation of hFX EGF1_{39mer} at the anticipated Asp-residue (Asp103_{hFX}, Fig. 1c–e). However, incomplete substrate turnover (40%) was observed even with prolonged incubation times and at higher enzyme-to-substrate ratios using hFX EGF1_{39mer} synthesized by thiol oxidation in air-saturated buffer. We next set out to obtain crystal structures of AspH-TPR-Ox, in part aiming to rationalise the observed incomplete substrate hydroxylation.

The longer AspH-TPR-Ox construct was crystallised in the presence of Mn(II), substituting for the Fe(II) co-factor, and the 2OG analogue, N-oxalylglycine (NOG); L-malate was not included in the crystallization buffer (Fig. 2 and Supplementary Fig. 6). The oxygenase domain fold conformation in the AspH-TPR-Ox structure is near identical to that of the AspH-Ox structure (main chain RMSD 0.20–0.22 Å), with the largest conformational differences occurring in an acidic loop (aa 614–620) which forms part of the active site (Supplementary Figs. 7 and 8).

The solenoid-like TPR domain (aa 330–555) comprises 6 tandem repeating pairs of anti-parallel helices (12 helices in total, α 1– α 12) and is connected to the AspH-Ox domain by a hinge-loop region (aa 556–577) (Fig. 2a–c). Like that of AspH-Ox, the AspH-TPR-Ox structure shows His679, His725, and two waters coordinating the metal, but instead of malate, NOG is bound in the 2OG binding pocket. The NOG C-5 carboxylate is positioned to salt bridge with Arg735 and to hydrogen-bond with the Ser668 hydroxyl (which form a “RXS motif” as observed in some other 2OG oxygenases)³¹. The NOG oxalate is bound to the active site via bidentate metal coordination and forms hydrogen bonds with the side chains of His690 and Arg688 (Fig. 2d, e). This structure

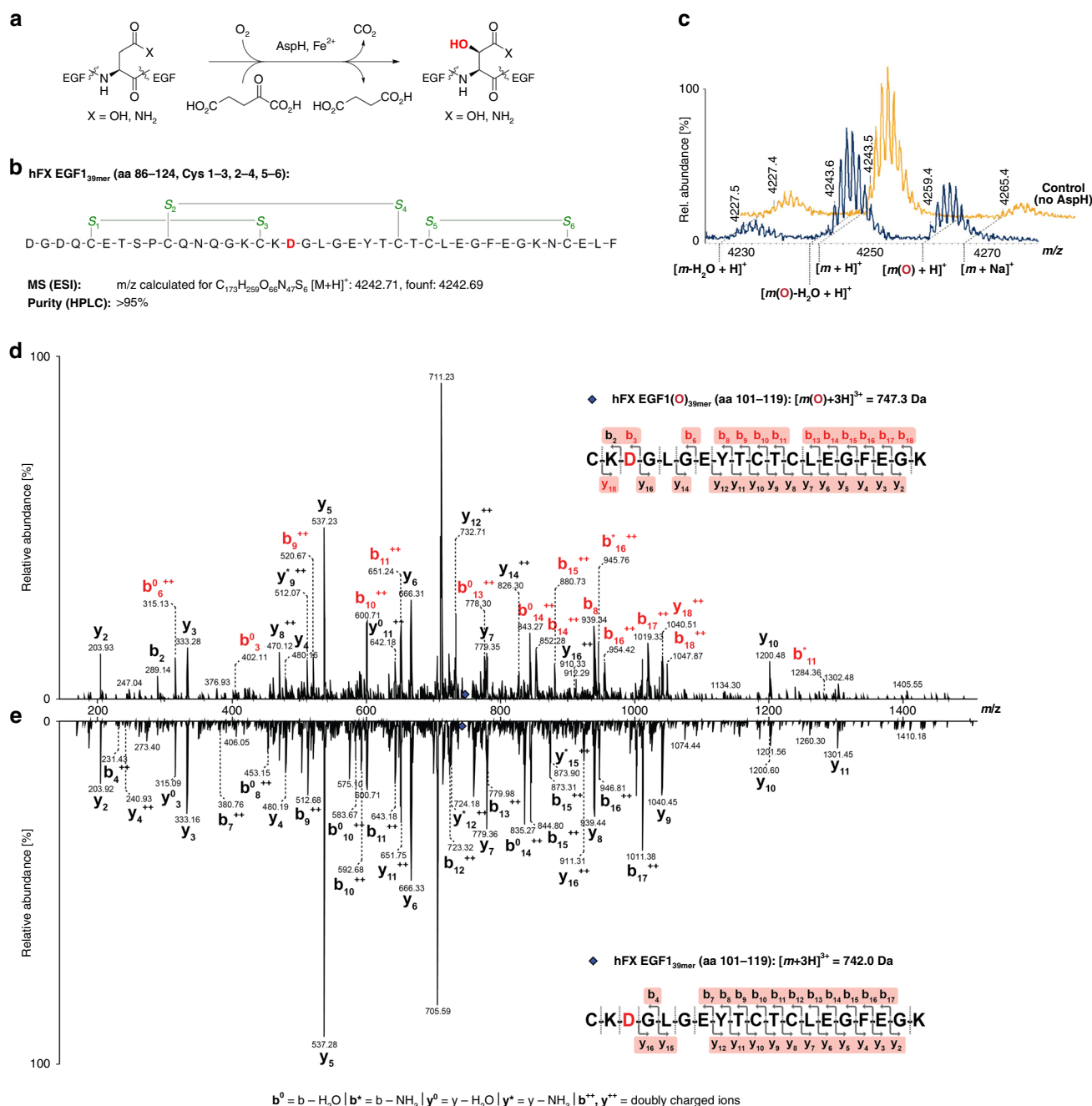


Fig. 1 EGF1 of human coagulation factor X undergoes AspH-catalysed hydroxylation at Asp103_{hFX}. MS/MS-analysis was performed using a Bruker Daltonics amaZon™ Ion Trap LC-MS/MS system equipped with a Dionex™ UltiMate® 3000 HPLC machine. Endpoint turnover assays were performed under standard (non-redox) conditions (for details see Methods section). For MS/MS-analysis, after quenching, cystine disulfides were reduced (dithiothreitol) and cysteine thiols derivatized with iodoacetamide. **a** Reaction scheme for diastereospecific AspH-catalysed hydroxylation of Asp/Asn-residues in EGFs. **b** Schematic structure and calculated mass of the expected disulfide isomer of hFX EGF1_{39mer} (aa 86–124) featuring a canonical disulfide connectivity pattern (Cys1–3, 2–4, 5–6; green); the hydroxylation site (Asp103_{hFX}) is in red. **c** His₆-AspH_{315–758} catalysed hydroxylation of hFX EGF1_{39mer} proceeds with ~40% conversion under standard (non-redox) conditions; the light orange graph represents a control in which AspH was replaced by buffer. **d** LC-MS/MS analysis of the relevant digestion fragment of hFX EGF1_{39mer} (aa 101–119), which is partially hydroxylated (~40%) after treatment with His₆-AspH_{315–758}, confirms AspH-catalysed hydroxylation takes place at Asp103_{hFX} as anticipated based on the AspH-substrate consensus sequence (fragments bearing β-hydroxy-Asp are in red). **e** LC-MS/MS analysis of the relevant digestion fragment of hFX EGF1_{39mer} (aa 101–119) before exposure to His₆-AspH_{315–758}

indicates that the R735W substitution, present in a Traboulsi syndrome patient⁸, likely interferes with 2OG binding.

The major β-sheet of the DSBH core is flanked by seven helices (α13–α19, Fig. 2a, b); the minor β-sheet is partially flanked by two helices derived from the C-terminal repeats (repeat 6) of the TPR

domain. The 5 N-terminal TPR repeats form a right handed superhelix, the concave face of which partially aligns to the shorter linking loop end of repeats 2–4 that point towards the oxygenase domain and help to form the active site (Fig. 2b, c). The hinged oxygenase-TPR domain arrangement forms a large

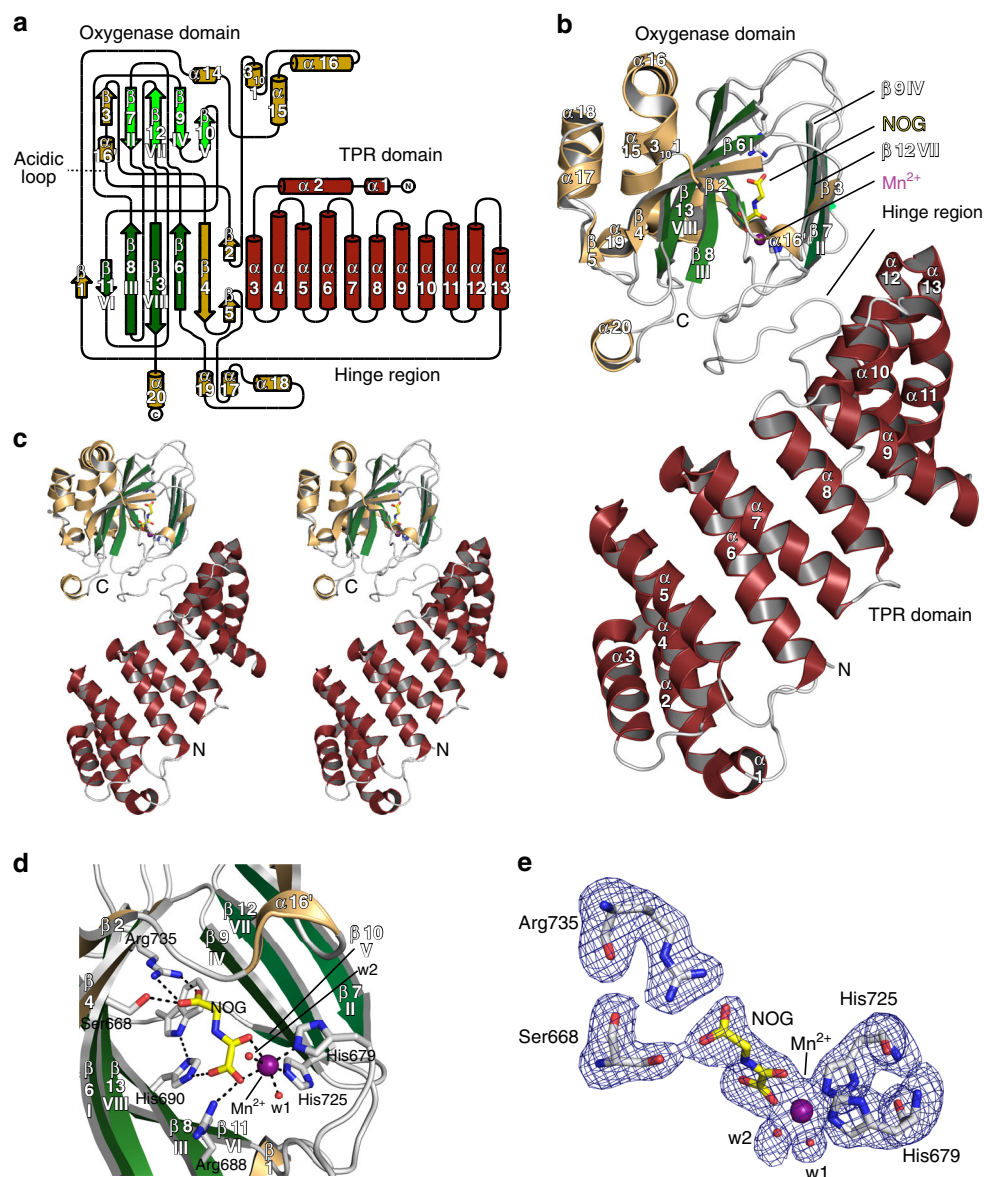


Fig. 2 Crystal structure of AspH-TPR-Ox reveals Fe(II)-binding involving only two residues. Color code: grey: His₆-AspH_{315–758}; yellow: carbon-backbone of *N*-oxalylglycine (NOG); violet: Mn; red: oxygen; blue: nitrogen. w: water. **a** Topology diagram of human AspH_{315–758}. **b** Overview of the AspH-TPR-Ox crystal structure. **c** Stereoview of the AspH-TPR-Ox crystal structure. **d** Close-up of the AspH-active site: Arg735 forms a salt bridge with the distal (C-5) carboxylate oxygens of NOG (2.4 and 3.1 Å) while Ser668 interacts with one oxygen lone pair through a hydrogen bond (2.7 Å). His690 is positioned to hydrogen bond to the C-1 carboxylate of NOG (2.8 Å). Mn(II) is bound to His679 (2.3 Å) and His725 (2.2 Å) of His₆-AspH_{315–758} and coordinates two water molecules (both 2.2 Å) as well as the C-1 (2.2 Å) and C-2 (2.5 Å) carbonyl oxygens of NOG. **e** Representative OMIT electron density map (mFo-DFc) contoured to 3σ around NOG, Mn, water, and relevant AspH-active site residues are shown

and open cavity, in which we envisaged EGFD substrates may fit. Comparison of the AspH TPR domain with other TPR domain structures identified by a DALI search reveals variations in the extent of curvature, likely due to the differences in intra-repeat interactions and loop lengths. The AspH TPR repeats 1–5 align particularly well with 5 repeats of the *Candidatus magnetobacterium bavaricum* magnetosome-associated (MAMA) TPR-containing protein (RMSD over 117 Cα atoms: 3.7 Å)³⁷. The AspH-TPR-Ox structure suggested that the TPR domain may be directly involved in substrate binding. To investigate, we crystallized AspH-TPR-Ox in the presence of the hFX EGF1_{39mer} substrate.

AspH-TPR-Ox was successfully co-crystallized with a 39-residue fragment of hFX EGF1 (AspH-TPR-Ox:hFX, Fig. 3 and Supplementary Fig. 9). The AspH-TPR-Ox structure reveals

electron density at the active site corresponding to 18 of the 39 hFX EGF1 substrate residues (aa 99–116: -GKC₃KDGLGEYTC₄TC₅LEGF-) and for a disulfide linkage between Cys101_{hFX} and Cys110_{hFX} (Fig. 3e). Highly specific interactions, including multiple protein–protein/peptide interactions with both AspH-Ox and TPR domains are apparent (Fig. 3b–d). The N-terminal region of hFX (aa 100–105) interacts with the AspH-Ox domain and the C-terminal region of hFX (aa 106–116) interacts with the TPR domain. Comparison of AspH structures with and without substrate implies that induced fit involving substantial conformational changes during formation of the AspH-TPR-Ox:hFX protein:protein complex occurs; the distance between Leu433 on repeat 3 of the TPR domain to Pro756 of the oxygenase domain near the C-terminus is reduced from 20 to ~14 Å on substrate binding. This movement is

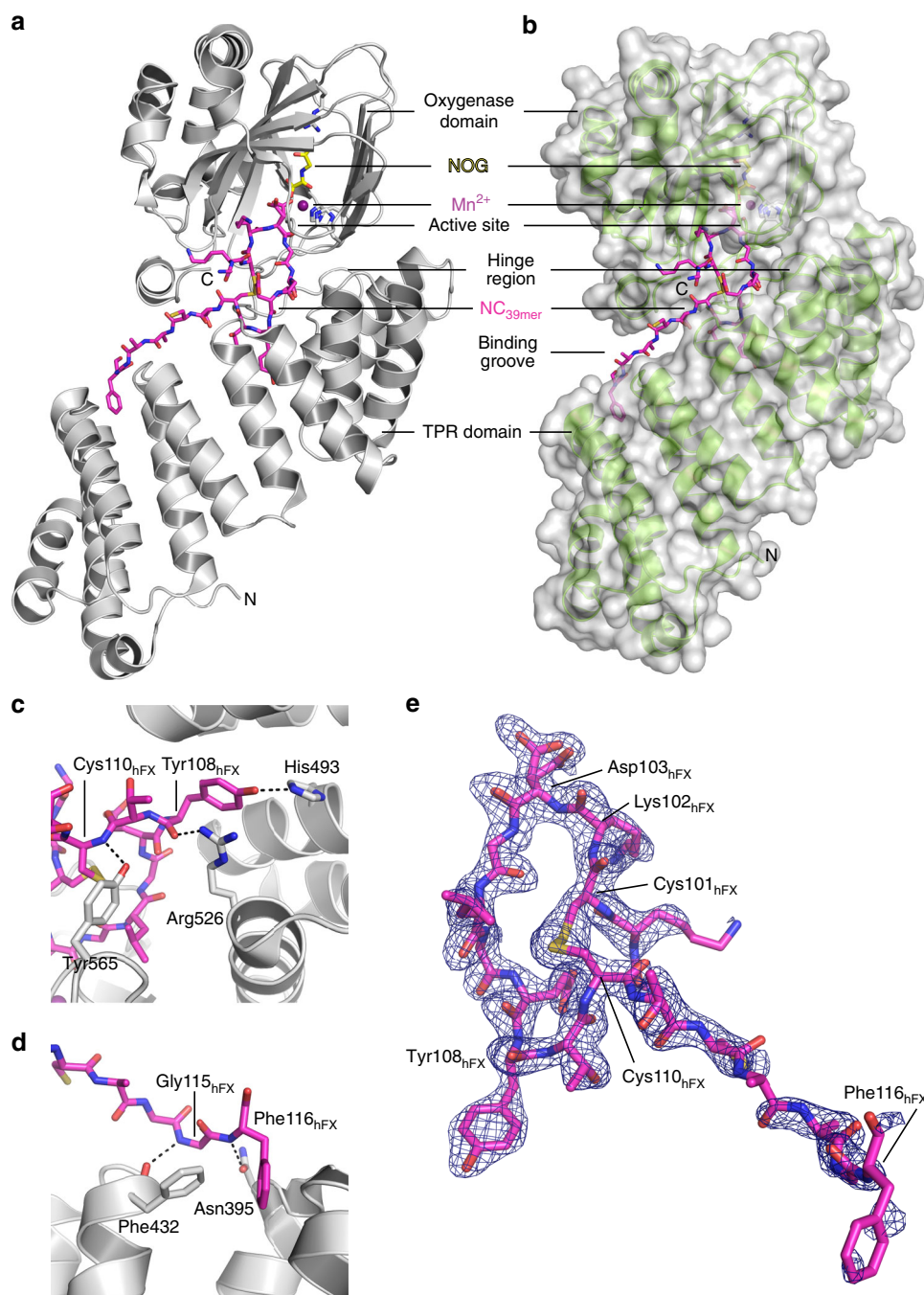


Fig. 3 Structure of AspH-TPR-Ox:hFX features a non-canonical EGFD disulfide connectivity. Color code: magenta: carbon-backbone of NC_{39mer} peptide; yellow: carbon-backbone of N-oxalylglycine (NOG); violet: Mn; red: oxygen; blue: nitrogen; pale yellow: sulfur. **a** Overview of the AspH-TPR-Ox:hFX structure; His₆-AspH₃₁₅₋₇₅₈ in grey. **b** Surface representation (grey) of the AspH-TPR-Ox:hFX structure showing the substrate-binding groove, His₆-AspH₃₁₅₋₇₅₈ in green. **c** Interactions of the TPR domain residues His493 and Arg526 as well as the hinge region residue Tyr565 with the NC_{39mer} peptide. **d** Additional interactions of the TPR domain with the NC_{39mer} peptide: Asn395 forms a hydrogen bond with Phe116_{hFX} (3.0 Å) and Phe432 forms a hydrogen bond with Gly115_{hFX} (3.2 Å). **e** OMIT electron density map (mFo-DFc) contoured to 3σ around the hFX derived NC_{39mer} peptide supports the presence of a non-canonical disulfide bridge between Cys101_{hFX} and Cys110_{hFX} as a substrate requirement for AspH (canonical EGFD disulfide isomer: Cys1-3, 2-4, 5-6; non-canonical EGFD disulfide isomer: Cys1-2, 3-4, 5-6; see Fig. 5a). Note that electron density for two alternative conformation of the Asp103_{hFX} side chain is observed

apparently enabled by the hinge region (aa 556–577) linking the TPR and AspH-Ox domains (Fig. 4a, b).

The side chain of Asp103_{hFX}, the substrate asparagine undergoing hydroxylation, is at the i+2 position of a Type I β-turn adjacent to the active site metal and was modelled in two conformations (A,B; see Supplementary Fig. 10 for details), both

of which are positioned to form salt bridges with the side chains of Arg688 and Lys666 (Figs. 4c and 3e). The Asp103_{hFX} carboxylate side chain in conformation A is positioned to form a hydrogen bond (3.2 Å) with the side chain of Gln627, whilst in conformation B the carboxylate is positioned to interact with the active site metal (2.6 Å) *trans* to His725. Asp103_{hFX} conformation

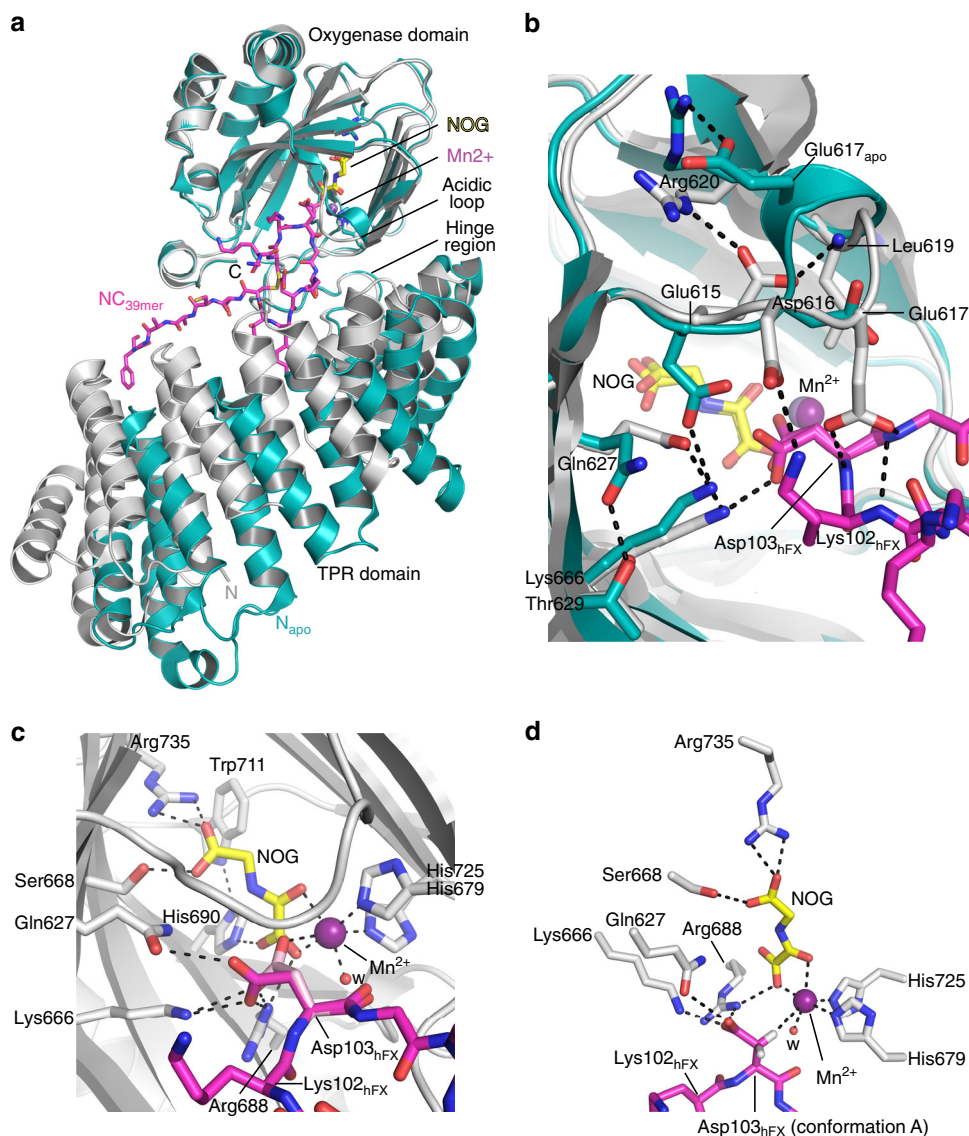


Fig. 4 Significant conformational changes occur in AspH upon substrate binding. Color code: grey: His₆-AspH_{315–758} (AspH-TPR-Ox:hFX); turquoise: His₆-AspH_{315–758} (AspH-TPR-Ox); magenta: carbon-backbone of NC_{39mer} peptide; yellow: carbon-backbone of N-oxalylglycine (NOG); violet: Mn; red: oxygen; blue: nitrogen; pale yellow: sulfur. w: water. **a** Superimposition of the AspH-TPR-Ox:hFX (grey) and AspH-TPR-Ox (turquoise) structures indicate conformational changes in the TPR domain, the hinge region, and the oxygenase acidic loop on substrate binding. **b** Superimposition of the AspH-active sites of the AspH-TPR-Ox:hFX (grey) and AspH-TPR-Ox (turquoise) structures: The interaction between Glu617_{apo} and Arg620 (2.5 Å) in the AspH-TPR-Ox structure is lost on substrate binding; both Asp616 and Glu617 interact with the substrate in the AspH-TPR-Ox:hFX structure. Gln627 (3.2 Å) and Lys666 (2.7 Å) bind to the Asp103_{hFX} carboxylate of the active conformer of the AspH-substrate. On substrate binding, the side chain of Glu615 rotates by ~90° to interact with the side chain of Arg620 (2.8 Å) and the main chain of Leu619 (2.8 Å), rather than Lys666 (2.7 Å) as in the substrate unbound state. **c** The side chain of the Asp103_{hFX} residue undergoing hydroxylation is observed in two conformations (A: magenta and B: pink; see Supplementary Fig. 10 for details). The Asp103_{hFX} side chain carboxylate of conformation B (pink) is positioned (2.6 Å) to interact with the Mn. **d** Close-up of the AspH-active site: The *pro-R* hydrogen at the Asp103_{hFX} β-position of the likely productive NC_{39mer} conformation A (magenta) is positioned to interact with the Mn (distance Cβ-Mn: 4.2 Å), consistent with hydroxylation at this position

A is poised for C-3 hydroxylation such that its *pro-3R*-hydrogen projects towards the metal (Cβ-metal distance ~4.5 Å) (Fig. 4d), consistent with reported (3R)-stereoselectivity for AspH catalysis^{1,36}. The Asp103_{hFX} conformation B displaces one of the two metal coordinating waters that are observed in the AspH-TPR-Ox structure without bound substrate.

Substrate binding to AspH also involves major changes within an acidic loop which forms part of the active site (aa 614–620), with the side chain carboxylates of Glu615, Asp616 and Glu617 all being shifted towards the substrate (Fig. 4b). The Glu617 side chain is positioned to form a hydrogen bond (2.8 Å) with the

main chain amide nitrogen of Asp103_{hFX}; the Asp616 side chain is positioned to form a salt bridge with the Lys102_{hFX} side chain; and Glu615 folds into the inner apex of the acidic loop forming a salt bridge with Arg620 and a hydrogen bond with the main chain amide nitrogen of Leu619. In the AspH-Ox structure, the side chain of Glu617 appears to mimic the Asp103_{hFX} side chain interactions with Lys666 and Arg688, possibly reflecting a function for this residue in protecting the active site from oxidation in the absence of substrate. In AspH-TPR-Ox:hFX structure, the Arg686 side chain is positioned to form a hydrogen bond with the main chain carbonyl oxygen of Lys102_{hFX} and a

salt bridge with the Asp721 side chain. The position of the side chains of Arg686, Arg688 and Asp721 in AspH-TPR-Ox:hFX does not change in comparison with AspH-TPR-Ox.

Role of the TPR domain. Notably, the highly-conserved hFX tyrosine, Tyr108_{hFX}, that is part of the consensus EGFD sequence for AspH hydroxylation³⁸, binds in a mostly hydrophobic pocket located between TPR repeats 5 and 6 located on the concave surface of the TPR domain. The Tyr108_{hFX} side chain hydroxyl is positioned to form a hydrogen bond to the His493 side chain Nε2 (3.0 Å) and its main chain carbonyl to form a hydrogen bond (2.8 Å) with the Arg526 side chain. The Tyr108_{hFX} side chain is apparently positioned to form skewed π - π -stacking interactions with the side chains of Phe496 and Arg526. The main chain nitrogen of Cys110_{hFX} (2.9 Å) interacts with the Tyr565 side chain (Fig. 3c).

At the hFX C-terminus, the phenyl ring of Phe116_{hFX} is buried in a hydrophobic pocket on the TPR domain formed by Ala389, Leu398, and Phe432, which are located between TPR repeats 2 and 3. The phenyl ring of Phe116_{hFX} is positioned to π -stack with the Arg393 guanidine side chain and the peptide bonds linking Arg393 with Ser394 and Asn395 (Fig. 3d). The Phe116_{hFX} main chain amide NH forms a hydrogen bond with the Asn395 side chain Oδ1 (2.6 Å). Sequence alignment of known hydroxylated EGFDs shows that the aromatic residues at the Phe116_{hFX} position are highly conserved (Supplementary Fig. 11).

Unexpectedly, in the AspH-TPR-Ox substrate complex we observed a non-canonical disulfide bridge, i.e. between Cys101_{hFX} and Cys110_{hFX} for the hFX substrate, in effect forming a ten-residue ring (Fig. 3e). The substrate disulfide is positioned in a hydrophobic pocket formed by Tyr565, Pro682, and Ile758. This pocket is shared with Leu105_{hFX} and expanded by Phe529 and Leu564 of AspH.

The observed hFX disulfide corresponds to a non-canonical Cys3–4 EGFD disulfide link, suggesting this pattern is required for AspH catalysis rather than the established (Cys1–3, 2–4, 5–6) disulfide pattern. The canonical EGFD disulfide pattern (Cys1–3, 2–4, 5–6) has been observed in all reported crystal and NMR structures of EGFDs bearing the AspH-substrate consensus sequence, including for hFX^{39,40}. A non-canonical disulfide pattern (Cys1–2, 3–4, 5–6) has been reported in crystalline and solution state studies on the EGFD-5 of human thrombomodulin^{41–43}; this, however, does not have the AspH substrate consensus sequence. MS⁴⁴ and computational⁴⁵ work also suggests the possibility of a non-canonical disulfide connectivity pattern for EGFDs bearing the consensus sequence for AspH-catalysed hydroxylation. It was, therefore, important to investigate AspH activity towards disulfide patterns with the non-canonical 3–4 link (i.e. Cys1–2, 3–4, 5–6; 1–5, 3–4, 2–6; or 1–6, 2–5, 3–4).

AspH accepts an unexpected EGFD disulfide pattern. MS analyses imply that the disulfide pattern arising from thiol oxidation of the linearly synthesised hFX EGF1_{39mer} in air-saturated buffer predominantly comprises a mixture of the canonical (Cys1–3, 2–4, 5–6) and non-canonical (Cys1–2, 3–4, 5–6) disulfide patterns (Fig. 5a); other additional disulfide isomers may be present in the mixture but were not detected. The canonical and non-canonical disulfide isomers co-elute as a single peak in MS coupled HPLC (Supplementary Figs. 12 and 13). Similar results have been reported for the synthesis of EGFD5 of human thrombomodulin⁴⁶. The hFX EGF1_{39mer} mixture underwent partial hydroxylation (~40%, Fig. 1c) by AspH-TPR-Ox under normal assay conditions. However, it underwent complete hydroxylation (>95%) when incubated with enzyme in a redox

buffer containing a mixture of reduced and oxidized glutathione (GSH and GSSG). The conditions were employed to enable reversible thiol-disulfide interchange of the canonical and non-canonical hFX forms (Fig. 5b).

The complexities of the reversible nature of thiol-disulfide interchange in solution render definitive identification of the preferred disulfide pattern of the AspH EGFD substrate difficult. We thus synthesised an irreversibly-linked cyclic analogue of the Cys3–4 EGFD disulfide substrate; A 10mer cyclic peptide (CP_{101–110}) comprising the core residues (aa 101–110) of hFX EGF1_{39mer}, wherein the Cys3–4 disulfide was replaced by a stable thioether, was prepared using the method of Suga et al.^{47,48}. CP_{101–110} was efficiently Asp- β -hydroxylated by AspH-TPR-Ox (~85%, Supplementary Figs. 14a and 15), whereas cyclic peptides with variations in consensus sequence residues or a linear 10mer peptide were not (Supplementary Fig. 14), further supporting the crystallographically observed Cys3–4 disulfide as the preferred substrate form for AspH catalysis. An extended 19mer CP (CP_{101–119}, Fig. 6a) was designed based on the enzyme-substrate contacts observed in the AspH-TPR-Ox:hFX structure and found to be a better substrate than CP_{101–110} (>95%, Fig. 6b). A structure of AspH co-crystallised with the CP_{101–119} (AspH-TPR-Ox:CP_{101–119}) reveals this substrate bound to one of the two AspH molecules in the asymmetric unit (chain A). Conformational differences between the bound (chain A) and unbound (chain B) (C α RMSD 3.82 Å) are consistent with afore-described observations implying an AspH induced fit mechanism with hFX (Fig. 6c and Supplementary Fig. 16).

Substrate variant analysis. Next, we assayed AspH for its ability to hydroxylate truncated 26mer hFX EGF1 substrates (aa 86–111) lacking the Cys5–6 disulfide-forming residues, synthesized using orthogonal cysteine protection to yield selectively both the pure canonical (C_{26mer}, Fig. 7a) or non-canonical (NC_{26mer}, Fig. 7d) disulfide isomers. Only the NC_{26mer} peptide was hydroxylated under normal assay conditions (Fig. 7b, e), whereas both 26mer peptides were hydroxylated using redox buffer conditions (Fig. 7c, f). A structure of AspH in complex with NC_{26mer} was obtained (AspH-TPR-Ox:NC_{26mer}, Supplementary Fig. 17). Comparison of AspH-TPR-Ox:NC_{26mer} and AspH-TPR-Ox:hFX structures reveals little overall difference (C α RMSD 0.13 Å). The presence of a Cys3–4 disulfide bond indicates that the additional residues present in the longer 39mer substrate, which were not observed in electron density maps of AspH-TPR-Ox:NC_{26mer} due to disorder, are likely not essential for substrate recognition. The efficient hydroxylation of the synthetic hFX EGF1_{39mer} with cysteines 1, 2, 5 and 6 (aa 90, 95, 112 and 121, respectively) all being replaced by serines (NC-4Ser_{39mer}) further supports the proposal that only the substrate Cys3–4 disulfide is required for AspH-catalysed hydroxylation (Fig. 8a–c). A structure of AspH in complex with the NC-4Ser_{39mer} substrate was obtained (AspH-TPR-Ox:NC-Ser_{39mer}, Supplementary Fig. 18). Comparison of the AspH-TPR-Ox:NC-Ser_{39mer} and AspH-TPR-Ox:hFX structures reveals little overall difference (C α RMSD 0.08 Å, Fig. 8d), including for the Ser112_{NC-Ser39mer} and Cys112_{hFX} residues, suggesting that Cys112_{hFX} may not necessarily form a disulfide during productive AspH-catalysis.

AspH with other EGFD protein substrates. To investigate the generality of the hFX results, we tested recombinant Ca(II)-binding EGFDs (cbEGFDs) (~10–15 kDa) from human Fibrillin-1 (hFib1); a two EGFD containing construct, cbEGF32–33⁴⁹, a three EGFD construct, cbEGF41–43⁵⁰, and the single EGFD construct, TB4cbEGF23 (with a cbEGFD preceded by a transforming growth factor β -binding protein-like (TB) domain)⁵¹, and a three EGFD

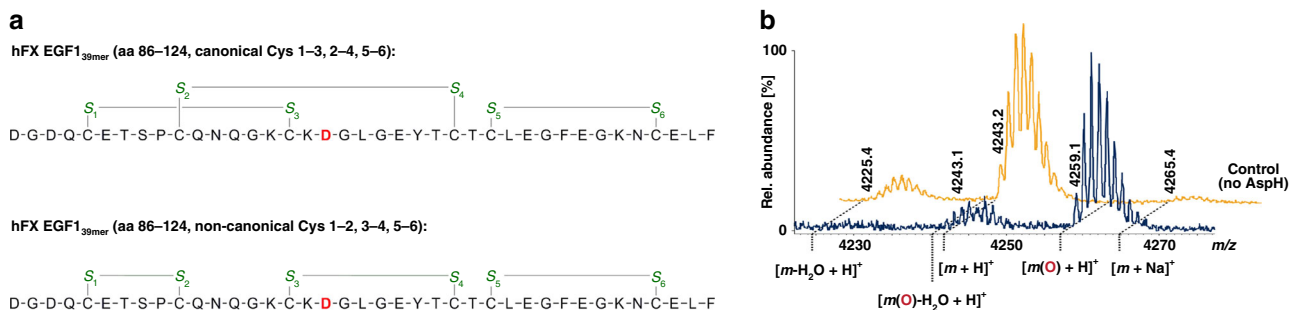


Fig. 5 AspH fully hydroxylates a mixture of canonical and non-canonical EGFD disulfides under redox conditions. End-point turnover reactions were performed as in the Methods Section. **a** Schematic structures of the two major hFX EGF1_{39mer} disulfide isomers identified in a single batch of hFX EGF1_{39mer} obtained by thiol oxidation in air-saturated buffer (see Supplementary Information); disulfides are in green (canonical isomer, Cys1-3, 2-4, 5-6, top; non-canonical isomer, Cys1-2, 3-4, 5-6, bottom), the hydroxylation sites (Asp103_{hFX}) are in red. **b** >95% Hydroxylation of hFX EGF1_{39mer} was observed under redox conditions as opposed to partial hydroxylation under standard (non-redox) conditions (Fig. 1c), indicating that a 'non-canonical' EGFD-disulfide pattern (Cys 1-2, 3-4, 5-6) is the actual AspH-substrate; the light orange graph represents a control in which AspH was replaced by buffer

construct, cbEGF11-13, of Notch-1 (hNotch1)⁵². All these were hydroxylated by AspH-TPR-Ox (30–85%) under redox buffer conditions as determined by MS after proteolytic digestion, but were much poorer substrates under non-redox conditions (0–30%) (Supplementary Fig. 19).

Effect of calcium ions on EFGD stability and AspH activity. As most AspH-substrate EFGDs bear a Ca(II)-binding site nearby the AspH-hydroxylation site²⁸, we next investigated if Ca(II) ions inhibit EFGD hydroxylation, i.e. by stabilising the canonical EFGD fold which is apparently not an AspH-substrate. No effect on AspH-TPR-Ox activity was observed on addition of 3 mM Ca(II) ions with any of the tested substrates in the absence of redox buffer. By contrast, in the presence of redox buffer, hydroxylase activity was not significantly affected for the recombinant multi-domain proteins hFib1 (cbEGF32–33 and cbEGF41–43) and hNotch 1 (cbEGF11–13). However, AspH hydroxylase activity was substantially inhibited for the recombinant hFib1 fragment, TB4cbEGF23 (aa 1527–1647, Fig. 9a), which binds Ca(II) tightly ($K_d = 16 \pm 1$ nM)⁵¹ and which is a poor AspH substrate under non-redox hydroxylation conditions (<5% hydroxylation, Fig. 9b), but an efficient one in redox buffer (~85% hydroxylation, Fig. 9c). The addition of Ca(II) ions at 1 mM final concentration in redox buffer reduced the extent of TB4cbEGF23 hydroxylation from ~85 to ~25%; in the presence of 3 mM Ca(II) ions in redox buffer TB4cbEGF23 hydroxylation was further reduced to ~10% (Fig. 9d, e). No inhibitory effect was observed when 3 mM Mg(II) was added instead of Ca(II) (Supplementary Fig. 20a). Ca(II) ions do not directly inhibit AspH, as the irreversibly formed cyclic thioether, CP2_{101–110}, was fully hydroxylated in the presence of 10 mM Ca(II) (Supplementary Fig. 20b).

Since Ca(II) ions preferably bind to and stabilise the canonical EFGD fold²⁷, these results further support assignment of the non-canonical 3–4 EGFD disulfide form as a preferred AspH substrate. The effects of the neighbouring domains and the different intrinsic Ca(II) affinities of EFGD proteins complicates the understanding of the influence of Ca(II) ions on AspH-catalysis. Nonetheless, the results clearly reveal the potential for regulation of AspH activity by Ca(II) ions by stabilisation of cbEGFDs in the canonical disulfide pattern fold, so making them unavailable for disulfide shuffling to the non-canonical Cys3–4 pattern.

Discussion

There is an unmet need for further investigations into the biological functions of human AspH and its role in genetic

diseases and cancer. Such work is challenging given the membrane bound nature of AspH⁵, the presence of multiple different isoforms/splice variants (>10 in humans)^{6,7}, its location in the ER⁵, and the redox sensitive nature of its disulfide substrates⁵³. Our work defining the first substrates for isolated AspH and the roles for the TPR domain in its catalysis provides a structural basis for future work on the cellular and physiological roles of AspH.

Whilst the oxygenase domain fold of AspH is typical for the 2OG oxygenase superfamily (Fig. 2), details of its active site and the role of the TPR domain in EFGD-substrate binding are remarkable. Key insights arising from our structures include: the presence of an active site metal bound by only two protein ligands, the role of the TPR domain in substrate binding, and the fact that AspH only accepts EFGD-substrates with the non-canonical Cys3–4, instead of the canonical Cys1–3, 2–4, 5–6, disulfide pattern. The latter finding was unexpected given the prevalence of the canonical EGFD disulfide pattern (at least outside of the ER) reported in the literature⁵⁴. The lack of in vitro hydroxylation assays for AspH has hindered progress on its characterisation; our results will enable further detailed kinetic, mechanistic, and inhibition studies on AspH.

Comparison of the AspH structures with and without substrates implies substantial induced fit during substrate binding (Fig. 4a, b); experience with NMR studies on single oxygenase domain 2OG oxygenases^{55–57} implies that in the case of AspH, even more substantial changes in conformation may occur in solution, likely involving both oxygenase and TPR domains, as well as AspH-substrates. Whatever the precise extent of conformational changes during catalysis, the results reveal a central role for the TPR domain in enabling AspH catalysis—without it, 2OG turnover occurs, but EFGDs are not hydroxylated (Fig. 4 and Supplementary Fig. 3).

TPR domains are involved in protein-protein interactions and substrate recognition in other proteins/enzymes⁵⁸ including human 2OG oxygenases, i.e. the procollagen C-4 (CP4Hs) and C-3 (Leprecans) prolyl- and C-5 (PLODs) lysyl-hydroxylases, which contain (or are predicted to contain) TPR domains *N*-terminal to the oxygenase domain⁵⁹. As yet, there are no structures reported for the oxygenase domains of eukaryotic members of these TPR containing enzymes, which like AspH are ER localised, though the TPR domain of CPH in complex with a collagen-like peptide has been structurally characterised⁶⁰. Our work reveals how the AspH oxygenase and TPR domains recognise and bind a non-canonical 3–4 disulfide EFGD and, through conformational changes to the enzyme, guide and position the Asp/Asn residue to be hydroxylated to the active site in a catalytically productive

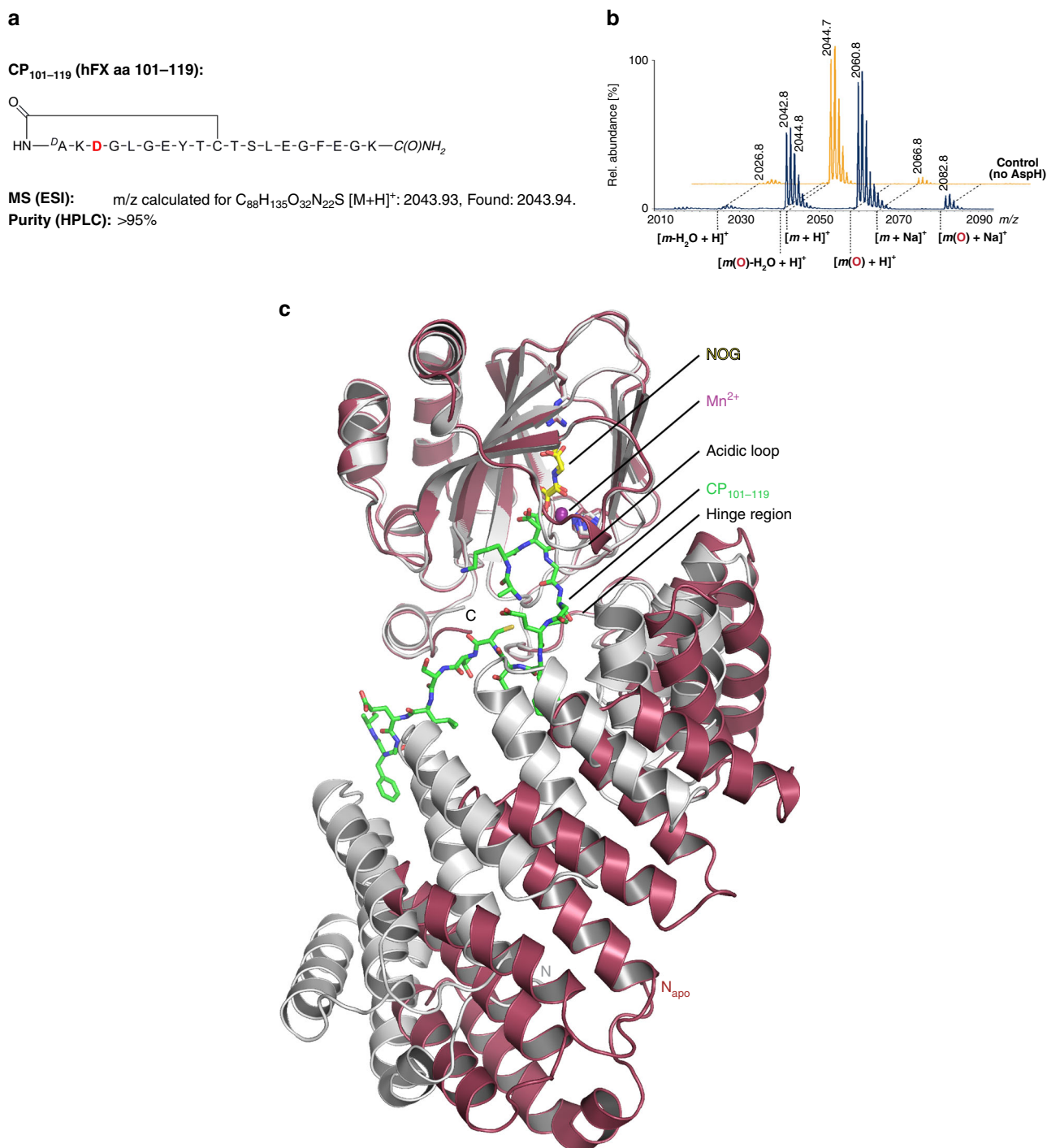


Fig. 6 A cyclic peptide mimicking the non-canonical Cys3-4 EGFD is an excellent AspH-substrate. End-point turnover reactions were performed as in the Methods Section. **a** Schematic structure and calculated mass of the cyclic peptide CP_{101–119} mimicking the central macrocyclic disulfide of the non-canonical isomer of hFX EGF1_{39mer} (Cys3-4); the hydroxylation site (Asp103_{hFX}) is in red. **b** >95% Hydroxylation of the CP_{101–119} peptide was observed under standard (non-redox) conditions; the light orange graph represents a control in which AspH was replaced by buffer. **c** Two His₆-AspH_{315–758}-molecules are present in the asymmetric unit of the AspH-TPR-Ox:CP_{101–119} crystal structure, only one binds the CP_{101–119} peptide (details in the Supplementary Information): The superimposition of the two independent AspH-molecules of this crystal structure (colour code: unbound AspH: raspberry; AspH bound to CP_{101–119}: grey; CP_{101–119}: green) highlights an induced fit mechanism of AspH upon substrate binding: Major conformational changes occur in the TPR domain, the hinge region and the acidic loop whilst the overall oxygenase domain conformation is hardly affected. The thioether linker of the cyclic peptide could not be accurately modelled into the electron density; however, complementary analytical methods imply the presence of a cyclic thioether linkage (see Supplementary Figure 16 for further details)

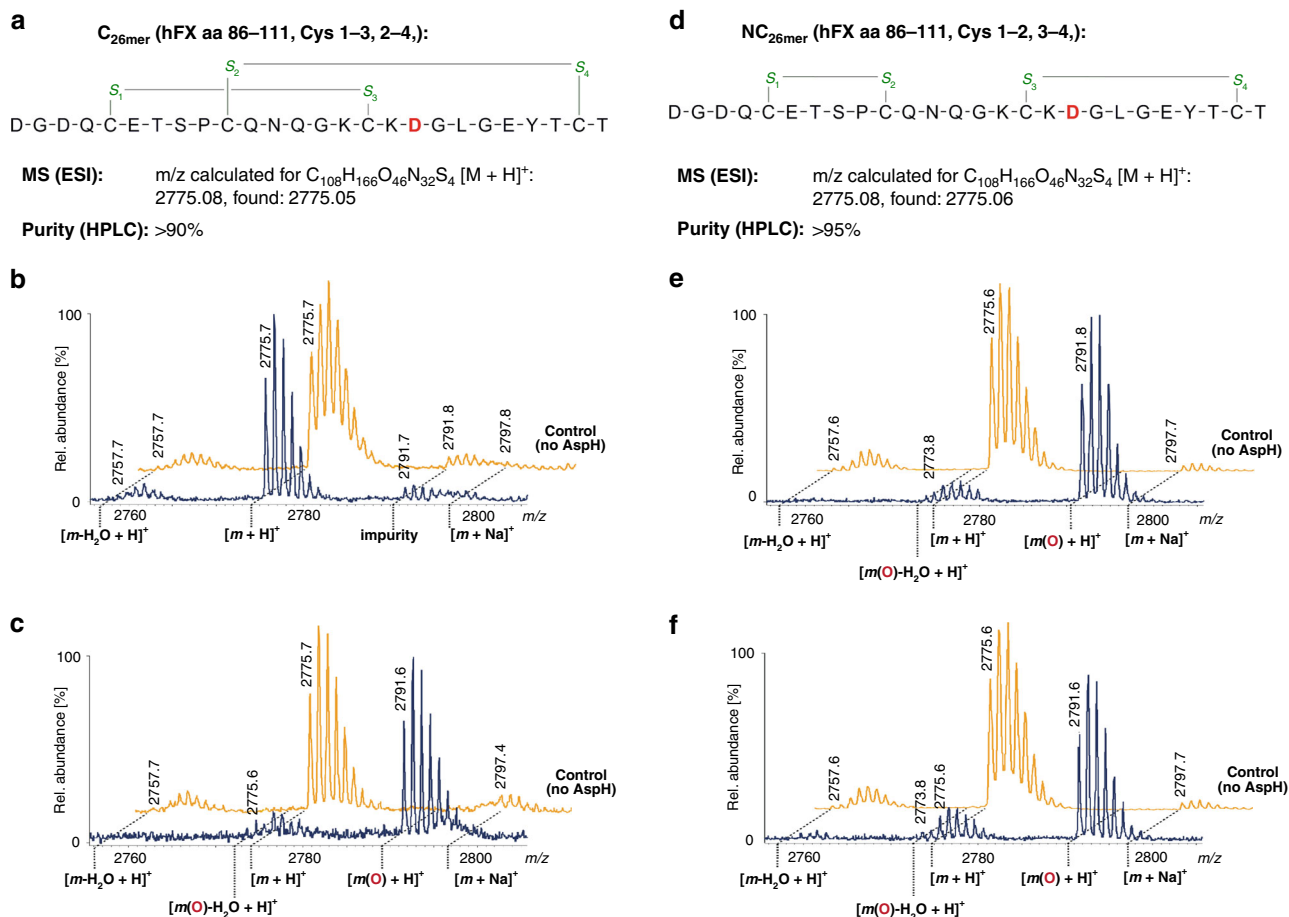


Fig. 7 The EGFD disulfide connectivity determines the degree of AspH-catalysed hydroxylation under non-redox conditions. End-point turnover reactions were performed as in the Methods Section. Light orange graphs represent controls in which AspH was replaced by buffer. **a** Schematic structure and calculated mass of the C_{26mer} peptide, bearing a canonical disulfide arrangement (Cys1–3, 2–4); disulfides are shown in green, the hydroxylation site (Asp103_{hFX}) is indicated in red. **b** No hydroxylation of the C_{26mer} peptide was observed under standard (non-redox) conditions; ~7% impurity (might correspond to an oxidized byproduct) was observed in this sample, including the control samples. **c** >95% Hydroxylation of the C_{26mer} peptide was observed under redox conditions. **d** Schematic structure and calculated mass of the NC_{26mer} peptide, bearing a non-canonical disulfide arrangement (Cys1–2, 3–4); disulfides are in green, the hydroxylation site (Asp103_{hFX}) is in red. **e** >95% Hydroxylation of the NC_{26mer} peptide was observed under standard (non-redox) conditions. **f** ~83% Hydroxylation of the NC_{26mer} peptide was observed under redox conditions

manner. It would seem probable that some of the TPR domains in other ER-localised 2OG oxygenases have analogous roles.

The role of the TPR domain in AspH catalysis is preceded in the roles of some non-catalytic domains in substrate targeting by the 2OG dependent JmjC N^{ϵ} -methyl lysine histone demethylases (KDMs), though this is only understood in detail from a structural perspective for two KDM7 subfamily members (KDM7A/B)⁶¹. Binding of a highly-conserved EGFD tyrosine (e.g. Tyr108_{hFX}) to a hydrophobic pocket on the TPR domain surface is reminiscent of H3K4me3 binding to the plant homeobox domain (PHD) of KDM7A⁶¹. Analogous recognition domains to that of KDM7A/B occur in other JmjC KDMs, involving different domain types (e.g. Tudor domains), though to our knowledge TPR domains are not present in the chromatin modifying 2OG oxygenases; they thus appear to be characteristic of the ER localised 2OG oxygenases.

The combined results reveal that AspH accepts EGFD substrates with a non-canonical Cys3–4 disulfide pattern, rather than the canonical Cys1–3, 2–4, 5–6 pattern. Evidence for this comes from structural data, MS turnover assays under different redox conditions, inhibition by Ca(II) ions (which presumably bind and stabilise the canonical but not the non-canonical disulfide pattern), and the use of a stable disulfide cyclic peptide analogue.

These observations are interesting because they raise the possibility of EGFDs manifesting dynamic disulfide patterns in the redox active environment of the ER, rather than being limited to the canonical form normally observed by in vitro biophysics and which current evidence implies would likely be dominating in the oxidising extracellular environment. Indeed, the combined crystallographic evidence suggests the canonical form is the most thermodynamically stable^{21–23}. AspH could act as a chaperone regulating EGFD folding or recognising misfolded EGFDs, possibly as part of the redox sensitive unfolded protein response (UPR) pathway⁶². This is consistent with the identification of various degrees of β -hydroxylation in purified native proteins containing EGFDs, most notably human coagulation proteins hFIX (30%) and hFX (100%)⁶³, which implies they have been through a folding cycle prior to secretion into the bloodstream. Interestingly, purified hFVII does not show any β -hydroxylation, despite having the AspH-substrate consensus sequence, which suggests a possible role of AspH in recognising misfolded rather than functionally active folded EGFDs. One role of AspH could be to regulate EGFD folding with hydroxylation being an oxygen sensitive element of this process rather than an end in itself; this proposal is consistent with the lack of observed effect of Asn/Asp hydroxylation on the canonical EGFD fold conformation and on

a NC-4Ser_{39mer} (hFX aa 86–124, Cys 3–4):

MS (ESI): m/z calculated for C₁₇₃H₂₆₃O₇₀N₄₇S₂ [M + H]⁺: 4182.81, Found: 4182.79

Purity (HPLC): >95%

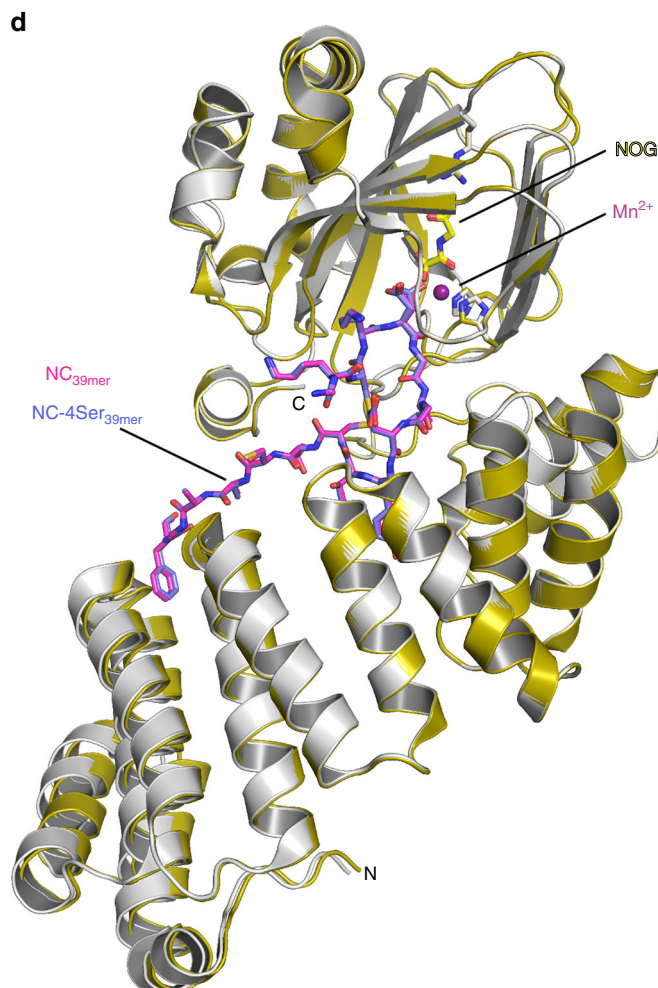
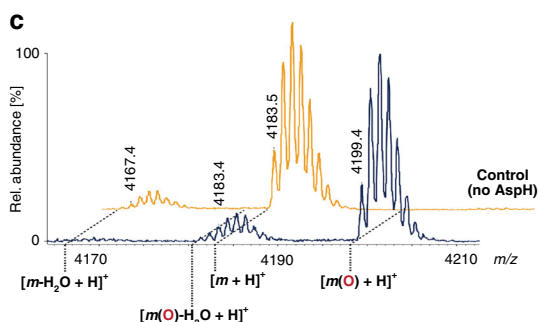
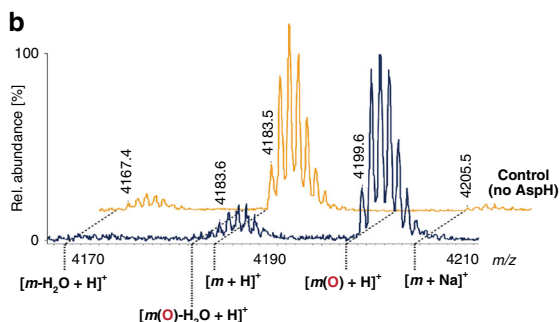


Fig. 8 The AspH-substrate requirement is a ring composed of 10 amino acid residues. End-point turnover reactions were performed as in the Methods Section. Light orange graphs represent controls in which AspH was replaced by buffer. **a** Schematic structure and calculated mass of the NC-4Ser_{39mer} peptide featuring a single disulfide between Cys3–4 (green); its sequence is identical to the one of hFX EGF1_{39mer} (see Fig. 1a), except that Cys90_{hFX}, 95, 112, 121 are substituted for Ser (light blue) to avoid disulfide scrambling; the hydroxylation site (Asp103_{hFX}) is in red. **b** >95% Hydroxylation was observed under standard (non-redox) conditions. **c** >95% Hydroxylation was observed under redox conditions. **d** Superimposition of the AspH-TPR-Ox:hFX crystal structure (color code: AspH: grey; NC_{39mer} peptide: magenta) with the AspH-TPR-Ox:NC-Ser_{39mer} crystal structure (color code: AspH: gold; NC-4Ser_{39mer} peptide: slate blue) shows a high conservation of the conformations of both enzyme and ligands

Ca(II) ion binding⁶³. Extracellular AspH could play a role in recognising and labelling mis-folded EGFDs on the cell surface, possibly in a manner relating to over-production of AspH in cancer cells^{10,11,13,16,17}, since the latter can have dysregulated EGFD proteins involved in signalling (e.g. Notch)^{64,65}. The structures reported here may also be of interest with respect to defining factors involved in AspH localisation, in particular to the surfaces of some tumour cells^{16,17}. They raise the possibility of AspH localisation being regulated by EGFD binding in a manner dependent on the EGFD disulfide pattern and redox environment. The results also suggest that AspH activity may be regulated by Ca(II) ions; in this regard it is interesting that there are predicted Ca(II) binding EF-hand motifs in human AspH itself.

The results clearly reveal that, at least under the analysed conditions, AspH has an unusual Fe(II) coordination chemistry for a 2OG-dependent hydroxylase (Fig. 2d, e). It should be noted that all the AspH crystal structures presented here are in complex with Ni(II) or Mn(II) and a 2OG mimetic, i.e. NOG or L-malate. We considered the possibility that in the presence of Fe(II) and 2OG a third protein-derived ligand interacts with the metal. However, this scenario seems unlikely given that the metal coordination geometries and substrate positions observed in many other 2OG-dependent oxygenase crystal structures in complex with Fe(II) and 2OG compared to Ni(II)/Mn(II) and NOG are very similar, with no evidence for additional enzyme derived ligands being observed³¹.

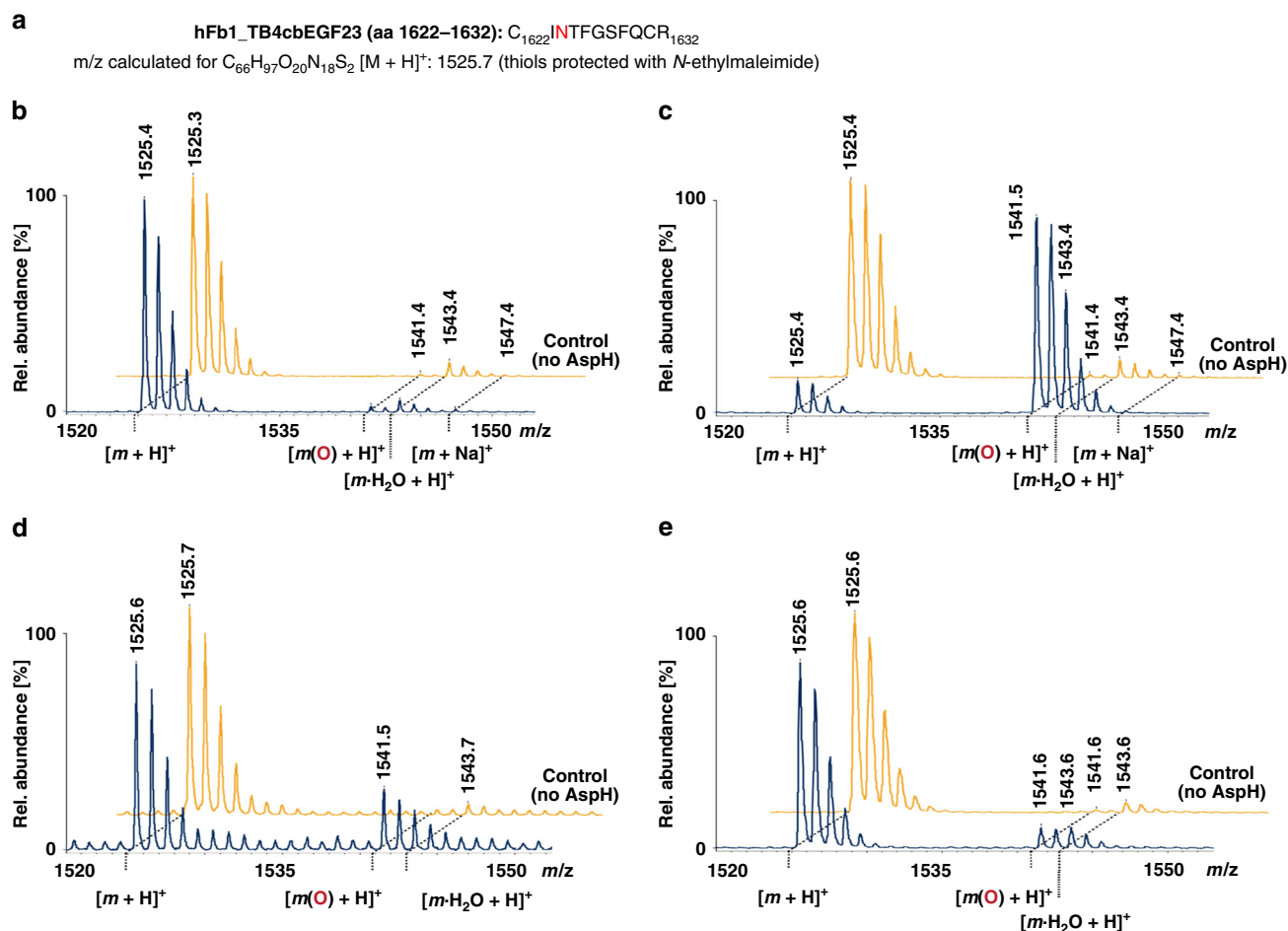


Fig. 9 Ca(II) ions inhibit AspH-catalysed hydroxylation of a human fibrillin-1 fragment. End-point turnover reactions were performed with an incubation time of 180 min (See Methods Section). Prior to analysis, the substrate was reduced and cysteine thiols were derivatized with *N*-ethylmaleimide, then digested using trypsin, and analysed using MALDI-ToF-MS. Light orange graphs represent controls in which AspH was replaced by buffer. **a** Sequence of the fragment of human fibrillin-1, TB4cbEGF23 (hFb1_TB4cbEGF23, aa 1527–1647; Uniprot Database entry: FBN1_HUMAN), bearing the hydroxylation site (Asn1624_{hFb1}, red). **b** <5% Hydroxylation was observed under standard (non-redox) conditions. **c** ~85% Hydroxylation was observed under redox conditions. **d** ~25% Hydroxylation was observed under redox conditions in the presence of 1 mM Ca(II). **e** ~10% Hydroxylation was observed under redox conditions in the presence of 3 mM Ca(II)

The reason why AspH employs two- rather than three-protein derived metal ligands is unclear, but sequence comparisons suggest this is a highly-conserved feature (Supplementary Fig. 6). It also appears to be linked to other characteristic active site features involving 2OG and substrate binding, which may compensate for the lack of a third protein ligand. It is of interest that in the structure of AspH-TPR-Ox complexed with a 39-residue fragment of hFX EGF1 (AspH-TPR-Ox:hFX, Fig. 4) the Asp residue positioned for AspH-catalysed hydroxylation is observed in two conformations. One conformer is apparently positioned for stereoselective hydroxylation in a manner preceded for other 2OG oxygenases (Supplementary Fig. 21). By contrast, the Asp side chain carboxylate of the other conformer is positioned to coordinate with the active site metal (Fig. 4). Whilst the latter conformation may be involved in catalysis, it may be that within the endoplasmic reticulum, AspH substrates are involved in maintaining Fe(II) at its active site.

The unusual coordination chemistry of AspH may relate to an as yet undetermined physiological or disease role as it is the case for the 2OG dependent hypoxia-inducible factor (HIF) prolyl hydroxylases (PHDs), where it is proposed that the unusual kinetic properties of PHDs, i.e. their slow reaction with O₂, both reflect their active site chemistry and roles as hypoxia sensors³¹.

Given the unusual coordination chemistry of AspH, detailed kinetic and solution biophysical studies to investigate dioxygen binding and water ligand displacement are of considerable interest and are the subject of ongoing work.

We compared the AspH structures with the human HIF Asn-hydroxylase, factor inhibiting HIF (FIH, PDB: 1H2K)⁶⁶, and a prokaryotic asparagine hydroxylase, AsnO (PDB: 2OG7)⁶⁷ (Supplementary Fig. 21). Analysis of enzyme-substrate complex structures reveals similar substrate side chain orientations at the active sites of AsnO and FIH, consistent with knowledge that they produce the same (3*S*)-stereochemistry products^{67,68}, which contrasts with the (3*R*)-product of AspH^{1,36}. AsnO has an active site arginine (Arg305_{AsnO}), which interacts with the Ca carboxylate of the asparagine substrate, similarly to the interaction of AspH Arg688 with the side chain of Asp103_{hFX} (Supplementary Fig. 20). However, Arg305_{AsnO} is located on β-strand VIII and AspH Arg688 is located on β-strand III of their respective DSBHs, possibly reflecting convergent evolution within the 2OG oxygenase superfamily. Interestingly, FIH also interacts with the side chain of its Asn (and Asp) substrates using an arginine, Arg238_{FIH}, but one which is located on an insert linking DSBH β-strands βIV-V (Supplementary Fig. 21). Biochemical and structural studies on FIH metal binding carboxylate variants D201A

and D201G reveal coordination by only two His residues and that the D201G variant is catalytically viable⁶⁹. Comparison of these structures with AspH shows the main chain Ca of the carboxylate ligand positions are very similar, but that the hFX substrate residue side chain likely causes changes in metal geometry. It is probable that second sphere residues, including the substrate side chain, contribute to catalysis in the two-His metal binding 2OG oxygenases. 2OG dependent halogenases (SyrB2/CytC3/WelO5/AmbO5) also have two His metal binding ligands (His-Xxx-Ala... His motif), an arrangement enabling halide coordination to the active site metal⁷⁰. We have not observed halogenase activity with AspH in the presence of chloride ions, implying other factors in addition to the number of metal ligands control substrate hydroxylation versus halogenation.

The ability to carry out assays with isolated AspH and knowledge of its active site and substrate binding mode provide a platform from which to develop small molecule AspH inhibitors. These will be useful in dissecting its biological functions and will help to enable new cancer treatments and the development of selective inhibitors for other human 2OG oxygenases, some of which are current medicinal chemistry targets⁷¹.

Methods

Recombinant protein production and purification. A pET-28a(+) vector encoding for N-terminally His₆-tagged AspH_{315–758} was transformed into *E. coli* BL21 (DE3) cells. Cells were grown in 2TY media supplemented with kanamycin (0.05 mM; 30 mg/ml) at 37 °C with shaking (180 rpm). AspH production was induced by adding isopropyl β-D-thiogalactopyranoside to a final concentration of 0.1 mM when the OD₆₀₀ reached 1.2 at 18 °C. Cells were shaken for 16 h at 18 °C, then harvested by centrifugation (10,967 × g, 8 min, 4 °C); the resultant cell pellets were stored at –80 °C. The frozen cell pellets were resuspended in ice-cold (30 g/100 mL) 50 mM HEPES buffer (pH 7.5, 500 mM NaCl, 5 mM imidazole) containing EDTA-free protease inhibitor cocktail tablets (1 tablet/50 mL, Roche Diagnostics) and DNase I (bovine pancreas, grade II, Roche Diagnostics). Cells were lysed by sonication on ice (8 × 30 s bursts; Sonics Vibra-Cell VCX500, amplitude: 60%) and the lysate was centrifuged (48,384 × g, 30 min, 4 °C). The supernatant containing AspH was purified at 4 °C using Ni(II)-affinity chromatography (HisTrap HP column, GE Healthcare) on an AKTA Pure machine (GE Healthcare) with a step elution gradient (to 50 mM HEPES, pH 7.5, 500 mM NaCl, 40 mM imidazole) and elution buffers (50 mM HEPES, pH 7.5, 500 mM NaCl, 500 mM imidazole; AspH typically eluted at around 150–200 mM imidazole). Eluted fractions containing AspH were pooled, concentrated using Amicon Ultra centrifugal filters (3082 × g, 4 °C), and further purified by size-exclusion chromatography using a HiLoad 26/60 Superdex 75 pg 300 mL column with a flow rate of 1 mL/min and 50 mM HEPES (pH 7.5, 500 mM NaCl) as elution buffer. The protein was subsequently concentrated and buffer exchanged into 50 mM HEPES, pH 7.5 for storage at –80 °C until further use. Purity of the proteins was assessed by SDS-PAGE and MS (Supplementary Fig. 5).

AspH-substrates. A 39mer peptide based on the amino acid sequence of the first EGFD of hFX (hFX EGF1_{39mer}, aa 86–124) was synthesized by solid phase peptide synthesis (SPPS) and purified by Peptide Synthetics (Peptide Protein Research Ltd, UK). Regioselective disulfide formation to yield exclusively (within our analytical limits) the canonical disulfide pattern (Cys1–3, 2–4, 5–6) was achieved using a sequence of orthogonal cysteine protection and selective cysteine oxidation⁷². However, no hydroxylation of this peptide was observed when exposed to His₆-AspH_{315–758} and cofactors. Therefore, the 39mer peptide was resynthesized with the disulfide bridges being formed by thiol-oxidation in air-saturated buffer. This method has previously been applied to synthesize bovine FX⁴ and FIX²⁴ peptides that were successfully hydroxylated by bovine AspH. Using this protocol, a mixture of the canonical (Cys1–3, 2–4, 5–6; C_{39mer}) and non-canonical (Cys1–2, 3–4, 5–6; NC_{39mer}) disulfide-isomers was obtained (Supplementary Fig. 12 and 13), which was used in all subsequent hydroxylation assays. Derivatives of hFX EGF1_{39mer} containing either canonical (C_{26mer}) or non-canonical (NC_{26mer}; NC-4Ser_{39mer}) disulfides were synthesized as single disulfide isomers by GL Biochem (Shanghai) Ltd using orthogonal cysteine protection. The Ca(II)-binding EGFD-containing multi-domain AspH-substrates (hFIB1₁-cbEGFD32–33⁴⁹, hFIB1₁-cbEGFD41–43⁵⁰, hFIB1₁-TB4cbEGF23⁵¹, hNotch1₁-cbEGFD11–13⁵²) were produced recombinantly in *E. coli* and purified as reported in the cited literature.

Design and synthesis of cyclic peptides. Based on the AspH-TPR-Ox:hFX crystal structure, stable cyclic peptides comprising the core ring residues (aa 101–110) of the hFX EGF1_{39mer}-substrate with a thioether replacing the Cys3–4 disulfide were synthesized. The D-stereochemistry of the N-terminal amino acid (aa 101) aligns the peptide side chain with the main chain of the original substrate.

Commercial Fmoc-protected amino acids (AGTC Bioproducts; Alfa Aesar; CSBio; Iris Biotech; Novabiochem; Sigma-Aldrich; TCI) were used as received. SPPS was performed using a CSBio CS336X automated peptide synthesizer following standard Fmoc-strategy: linear peptides were synthesized from the C- to N-termini on a 0.1 mmol scale using a Rink amide linker using N,N-diisopropylcarbodiimide/1-hydroxybenzotriazole for coupling.

To obtain cyclic peptides, the N-terminal Fmoc-protecting group was cleaved on the resin, and the resin suspended in DMF (4 mL) containing N-chloroacetyl succinimide⁷³ (150 mg). The mixture was gently shaken for 3 h, filtered, and the dried resin treated with a solution of trifluoroacetic acid, triisopropylsilane, and water (4 mL; 95/2.5/2.5). After 3 h, the mixture was filtered and the solution diluted with 45 mL cold diethyl ether. The suspension was centrifuged (4000 rpm, 10 min, 4 °C), decanted, and then taken up in 1.5 mL aqueous triethylammonium acetate buffer (1 M, pH 8.5, pH readjusted with trimethylamine) and heated for 10 min at 100 °C in a microwave reactor (Biotage Initiator). Directly afterwards, the crude cyclic peptides were purified by semi-preparative HPLC (DionexTM UltiMate[®] 3000, Thermo Scientific) using a reverse phase column (Grace Vydac[®] 218TP101522) and a gradient of acetonitrile in milliQ water (each containing 0.1%_{v/v} trifluoroacetic acid) specified in Supplementary Fig. 14.

AspH activity assays. All reagents were obtained from commercial sources (Alfa Aesar; Sigma-Aldrich; TCI). For all turn-over experiments, His₆-tagged AspH-constructs were used, as no effect upon cleavage of the His₆-tag was observed in initial activity assays. Hydroxylation assays were initialised by mixing a solution containing AspH (10 μM), L-ascorbic acid (400 μM), and (NH₄)₂Fe(SO₄)₂·7H₂O (50 μM) in the appropriate buffer with a solution containing substrate (100 μM), disodium 2-oxoglutarate (300 μM), and, if desired, additional reagents (CaCl₂, EGTA) in the appropriate buffer. The buffers used comprised: Standard, non-redox, Buffer (50 mM HEPES pH 7.5, 150 mM NaCl); Redox Buffer (50 mM Tris, pH 8.5, 3.0 mM reduced L-glutathione, 0.3 mM oxidized L-glutathione, 150 mM NaCl); Refolding Buffer (50 mM Tris, pH 8.5, 3.0 mM L-cysteine, 0.3 mM L-cystine, 10 mM CaCl₂). Reactions were performed at 37 °C for 60 min unless otherwise noted, and then quenched by the addition of an equal volume of 1%_{v/v} aqueous formic acid. All assays were performed in triplicate alongside an additional no-enzyme control (His₆-AspH substituted by buffer).

The hydroxylase activity of AspH was assayed by MALDI-ToF MS in the positive ion reflection mode (Bruker Daltonics Ultraflex I machine with 32–60% laser energy; Waters Micromass MALDI micro MX with flight tube voltage: 12 kV, reflection voltage: 5.2 kV, laser fire rate: 10 Hz, pulse voltage: 1950 V, detector voltage: 2750 V); the relative quantities of substrate and hydroxylated product were analysed with respect to no-enzyme controls. A 1:4 volume ratio of sample to MALDI-matrix was used for sample preparation (20 mg/mL 2,5-dihydroxybenzoic acid, 10 mg/mL α-cyano-4-hydroxycinnamic acid or 10 mg/mL sinapinic acid in aqueous 50% CH₃CN containing 0.1%_{v/v} trifluoroacetic acid). For the disulfide containing substrates a matrix composed of a 10:1 mixture of 10 mg/mL α-cyano-4-hydroxycinnamic acid:10 mg/mL 2-(4'-hydroxybenzenazo)benzoic acid in aqueous 50%_{v/v} CH₃CN containing 0.1%_{v/v} trifluoroacetic acid was used. Mass spectra were analysed using MassLynx software (Version 4.1). Apart from the AspH-substrate peaks, peaks corresponding to the desired AspH-substrate mass reduced by 18 Da are frequently observed in the MALDI-MS spectra; these peaks result from the ionization induced liberation of H₂O upon intramolecular cyclization of aspartates to form aspartyl succinimides (–18 Da)⁷⁴ and are taken into account when integrating the peaks to determine conversions.

Crystallography. Crystals of N-terminally His₆-tagged AspH_{562–758} were grown by the vapour diffusion method at 20 °C in 150 nL sitting drops. The drops were prepared by mixing 100 nL of protein solution (9.5 mg/mL containing 1 mM N-oxalylglycine (NOG) and 10 mM NiSO₄) and 50 nL of precipitant (50 mM disodium malate, 30%_{v/v} PEG-3350, 30% dextran sulfate sodium salt) with a protein to well ratio of 1:2. Crystals were cryo-cooled in liquid N₂ with 25% glycerol as cryoprotectant and data were collected at 100 K using synchrotron radiation at the Diamond Light Source (DLS).

For the remaining structures, high-throughput crystallization screens were set up using a Phoenix RE liquid dispensing robot (Art Robbins Instruments) on 96-well, 3-subwell low profile Intelliplates (Art Robbins Instruments) using Hampton Research Molecular Dimensions crystallization screens (JCSG plus, PACT, INDEX, PEG/Ion, Structure, Salt, MIDAS) or optimization screens. N-Terminally His₆-tagged AspH_{315–758} (18 mg/mL in 50 mM HEPES buffer, pH 7.5) was mixed with 1 mM MnCl₂ (pH 7.5), 2 mM N-oxalylglycine (pH 7.5), and, when appropriate, an AspH-substrate (see Supplementary Table 1 for details). Crystals were grown by the vapour diffusion method at 4 °C in 200 or 300 nL sitting drops using 2:1, 1:1 or 1:2 sample:well solution ratios. Crystals were cryoprotected using mother liquor supplemented with 25%_{v/v} glycerol before cryo-cooling in liquid N₂. Data were collected at 100 K at DLS. Data were indexed, integrated, and scaled using HKL-3000⁷⁵, XDS⁷⁶, SCALA⁷⁷, Xia2⁷⁸ or CrystalClear (Rigaku). Detailed crystallization conditions for all crystal structures and a summary of data collection parameters are given in Supplementary Table 1.

Structure solution and refinements. The structure of AspH-Ox was determined by single isomorphous replacement with anomalous scattering (SIRAS). Briefly, AspH crystals were derivatised by incubating in a reservoir solution supplemented with 10 mM K_2PtCl_4 for 5 h. Crystals were transferred to and backsoaked for several seconds in a reservoir solution supplemented with 25% ethylene glycol instead of Pt(II) and crystals were immediately cryo-cooled in liquid N_2 .

The AspH-Ox (native and Pt-derivatised) datasets were integrated (XDS⁷⁶) and subsequently scaled (SCALA). SHELXD⁷⁹ was used to locate three platinum positions in the Pt-dataset. Heavy atom positions were refined and initial phases calculated with SHARP⁸⁰ using the SIRAS method. The final electron density map after phase extension and solvent flattening with SOLOMON⁸¹ was of excellent quality and automated model building with ARP/wARP⁸² resulted in a 95% complete model. Refinement with PHENIX⁸³ and several rounds of manual rebuilding in COOT⁸⁴ resulted in a model with a final $R_{\text{factor}}/R_{\text{free}}$ of 15.3/18.6%, respectively.

The AspH-TPR-Ox structure was determined by molecular replacement (MR) using the AutoMR (PHASER⁸⁵) subroutine in PHENIX. The coordinates of the AspH oxygenase domain (His₆-tagged AspH_{562–758} as a search model, PDB: 5APA) were first used to successfully identify a rotation and translation solution for the oxygenase domain. Initial attempts to find a rotation and translation solution for the TPR using the C-terminal TPR domain of peroxisomal targeting signal 1 receptor (PEX5)⁸⁶ were complicated by the repeating nature of the TPR fold. In order to overcome problems associated with the repeating fold the search model was trimmed to include three sequential repeats (PEX5 Gln300–Tyr395, PDB: 1FCH). The 3 C-terminal TPR repeats of AspH were identified in a search after the oxygenase domain was fixed. This allowed manual fitting of 3 further TPR repeats to the N-terminus of AspH. Rigid body refinement of the three separate fragments followed by simulated annealing refinement (both cartesian and torsion) led to a model with R_{free} 42.15. Iterative rounds of model building and fitting in COOT⁸⁴ and refinement in PHENIX⁸³ were performed until the decreasing R and R_{free} no longer converged and a final model was obtained. Due to domain movements between different crystal forms and complexes, all other structures were determined by MR using PHASER with the oxygenase and TPR domains searched for independently. The oxygenase domain was always searched first followed by the TPR domain. For data from crystals grown in the presence of substrate, up to 18 residues of bound substrate were modelled. Refinement statistics are given in Supplementary Table 2. Representative electron density (with contour level and map types defined in legends) is shown in Figs. 2 and 3 and in Supplementary Figs. 4, 7, 16–18.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Crystal structure data for N-terminal His₆-tagged AspH_{562–758} and N-terminal His₆-tagged AspH_{315–758} (apo and in complex with different AspH-substrates) are deposited in the protein databank with PDB accession codes: 5APA (AspH-Ox), 5JZA (AspH-TPR-Ox), 5JZ6 (AspH-TPR-Ox:malate), 5JZ8 (AspH-TPR-Ox:hFX), 5JQY (AspH-TPR-Ox:NC-Ser_{39mer}), 5JZU (AspH-TPR-Ox:NC_{26mer}), 6RK9 (AspH-TPR-Ox:CP_{101–119}). Other data are available from the corresponding author upon reasonable request.

Received: 17 May 2019; Accepted: 26 September 2019;

Published online: 28 October 2019

References

- Drakenberg, T., Fernlund, P., Roepstorff, P. & Stenflo, J. β -Hydroxyaspartic acid in vitamin K-dependent protein C. *Proc. Natl Acad. Sci. USA* **80**, 1802–1806 (1983).
- McMullen, B. A. et al. Complete amino acid sequence of the light chain of human blood coagulation factor X: evidence for identification of residue 63 as β -hydroxyaspartic acid. *Biochemistry* **22**, 2875–2884 (1983).
- Stenflo, J. et al. Hydroxylation of aspartic acid in domains homologous to the epidermal growth factor precursor is catalyzed by a 2-oxoglutarate-dependent dioxygenase. *Proc. Natl. Acad. Sci. USA* **86**, 444–447 (1989).
- Gronke, R. S. et al. Partial purification and characterization of bovine liver aspartyl β -hydroxylase. *J. Biol. Chem.* **265**, 8558–8565 (1990).
- Korioth, F., Gieffers, C. & Frey, J. Cloning and characterization of the human gene encoding aspartyl β -hydroxylase. *Gene* **150**, 395–399 (1994).
- Feriotto, G. et al. Myocyte enhancer factor 2 activates promoter sequences of the human $\alpha\beta$ H-J-J locus, encoding aspartyl β -hydroxylase, juncin, and juncate. *Mol. Cell. Biol.* **25**, 3261–3275 (2005).
- Feriotto, G. et al. Multiple levels of control of the expression of the human $\alpha\beta$ H-J-J locus encoding aspartyl β -hydroxylase, juncin, and juncate. *Ann. N. Y. Acad. Sci.* **1091**, 184–190 (2006).
- Patel, N. et al. Mutations in ASPH cause facial dysmorphism, lens dislocation, anterior-segment abnormalities, and spontaneous filtering blebs, or Traboulsi Syndrome. *Am. J. Hum. Genet.* **94**, 755–759 (2014).
- Vivante, A. et al. Exome sequencing discerns syndromes in patients from consanguineous families with congenital anomalies of the kidneys and urinary tract. *Clin. J. Am. Soc. Nephrol.* **28**, 69–75 (2017).
- Yang, H. et al. The distribution and expression profiles of human aspartyl/asparaginyl β -hydroxylase in tumor cell lines and human tissues. *Oncol. Rep.* **24**, 1257–1264 (2010).
- Lavaissiere, L. et al. Overexpression of human aspartyl(asparaginyl) β -hydroxylase in hepatocellular carcinoma and cholangiocarcinoma. *J. Clin. Invest.* **98**, 1313–1323 (1996).
- Wang, J. et al. Prognostic value of humbug gene overexpression in stage II colon cancer. *Hum. Pathol.* **38**, 17–25 (2007).
- Luu, M. et al. Prognostic value of aspartyl (asparaginyl)- β -hydroxylase/humbug expression in non-small cell lung carcinoma. *Hum. Pathol.* **40**, 639–644 (2009).
- Shimoda, M. et al. Endocrine sensitivity of estrogen receptor-positive breast cancer is negatively correlated with aspartate- β -hydroxylase expression. *Cancer Sci.* **108**, 2454–2461 (2017).
- Elvidge, G. P. et al. Concordant regulation of gene expression by hypoxia and 2-oxoglutarate-dependent dioxygenase inhibition: the role of HIF-1 α , HIF-2 α , and other pathways. *J. Biol. Chem.* **281**, 15215–15226 (2006).
- Maeda, T. et al. Clinicopathological correlates of aspartyl (asparaginyl) β -hydroxylase over-expression in cholangiocarcinoma. *Cancer Detect. Prev.* **28**, 313–318 (2004).
- Wang, K. et al. Overexpression of aspartyl-(asparaginyl)- β -hydroxylase in hepatocellular carcinoma is associated with worse surgical outcome. *Hepatology* **52**, 164–173 (2010).
- Dinchuk, J. E. et al. Absence of post-translational aspartyl β -hydroxylation of epidermal growth factor domains in mice leads to developmental defects and an increased incidence of intestinal neoplasia. *J. Biol. Chem.* **277**, 12970–12977 (2002).
- Dong, X. et al. Aspartate β -hydroxylase expression promotes a malignant pancreatic cellular phenotype. *Oncotarget* **6**, 1231–1248 (2015).
- Zou, Q. et al. Hydroxylase activity of ASPH promotes hepatocellular carcinoma metastasis through epithelial-to-mesenchymal transition pathway. *EBioMedicine* **31**, 287–298 (2018).
- Savage, C. R., Hash, J. H. & Cohen, S. Epidermal growth factor: location of disulfide bonds. *J. Biol. Chem.* **248**, 7669–7672 (1973).
- Cooke, R. M. et al. The solution structure of human epidermal growth factor. *Nature* **327**, 339 (1987).
- Ogiso, H. et al. Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell* **110**, 775–787 (2002).
- Gronke, R. S. et al. Aspartyl β -hydroxylase: in vitro hydroxylation of a synthetic peptide based on the structure of the first growth factor-like domain of human factor IX. *Proc. Natl. Acad. Sci. USA* **86**, 3609–3613 (1989).
- Ohlin, A. K., Linse, S. & Stenflo, J. Calcium binding to the epidermal growth factor homology region of bovine protein C. *J. Biol. Chem.* **263**, 7411–7417 (1988).
- Rees, D. J. G. et al. The role of β -hydroxyaspartate and adjacent carboxylate residues in the first EGF domain of human factor IX. *EMBO J.* **7**, 2053–2061 (1988).
- Tolkachev, D. & Ni, F. Calcium binding properties of an epidermal growth factor-like domain from human thrombomodulin. *Biochemistry* **37**, 9091–9100 (1998).
- Stenflo, J., Stenberg, Y. & Muranyi, A. Calcium-binding EGF-like modules in coagulation proteinases: function of the calcium ion in module interactions. *Biochim. Biophys. Acta Protein Struct. Molec. Enzym.* **1477**, 51–63 (2000).
- Rand, M. D., Lindblom, A., Carlson, J., Villoutreix, B. O. & Stenflo, J. Calcium binding to tandem repeats of EGF-like modules. Expression and characterization of the EGF-like modules of human Notch-1 implicated in receptor-ligand interactions. *Protein Sci.* **6**, 2059–2071 (1997).
- Loenarz, C. & Schofield, C. J. Expanding chemical biology of 2-oxoglutarate oxygenases. *Nat. Chem. Biol.* **4**, 152–156 (2008).
- 2-Oxoglutarate-Dependent Oxygenases* (eds Schofield, C. J. & Hausinger, R. P.), (The Royal Society of Chemistry, Cambridge, UK, 2015).
- McDonough, M. A., Loenarz, C., Chowdhury, R., Clifton, I. J. & Schofield, C. J. Structural studies on human 2-oxoglutarate dependent oxygenases. *Curr. Opin. Struct. Biol.* **20**, 659–672 (2010).
- Lancaster, D. E., McDonough, M. A. & Schofield, C. J. Factor inhibiting hypoxia-inducible factor (FIH) and other asparaginyl hydroxylases. *Biochem. Soc. Trans.* **32**, 943–945 (2004).
- Ng, S. S. et al. Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. *Nature* **448**, 87 (2007).
- Yue, W. W. et al. Crystal structure of the PHF8 Jumonji domain, an N ϵ -methyl lysine demethylase. *FEBS Lett.* **584**, 825–830 (2010).

36. Fernlund, P. & Stenflo, J. β -Hydroxyaspartic acid in vitamin K-dependent proteins. *J. Biol. Chem.* **258**, 12509–12512 (1983).
37. Zeytuni, N., Baran, D., Davidov, G. & Zarivach, R. Inter-phylum structural conservation of the magnetosome-associated TPR-containing protein, MamA. *J. Struct. Biol.* **180**, 479–487 (2012).
38. Stenflo, J., Lundwall, Å. & Dahlbäck, B. β -Hydroxyasparagine in domains homologous to the epidermal growth factor precursor in vitamin K-dependent protein S. *Proc. Natl. Acad. Sci. USA* **84**, 368–372 (1987).
39. Padmanabhan, K. et al. Structure of human des(1-45) factor Xa at 2.2 Å resolution. *J. Mol. Biol.* **232**, 947–966 (1993).
40. Selander-Sunnerhagen, M. et al. How an epidermal growth factor (EGF)-like domain binds calcium. High resolution NMR structure of the calcium form of the NH2-terminal EGF-like domain in coagulation factor X. *J. Biol. Chem.* **267**, 19642–19649 (1992).
41. White, C. E., Hunter, M. J., Meininger, D. P., Garrod, S. & Komives, E. A. The fifth epidermal growth factor-like domain of thrombomodulin does not have an epidermal growth factor-like disulfide bonding pattern. *Proc. Natl. Acad. Sci. USA* **93**, 10177–10182 (1996).
42. Fuentes-Prior, P. et al. Structural basis for the anticoagulant activity of the thrombin–thrombomodulin complex. *Nature* **404**, 518 (2000).
43. Sampoli Benitez, B. A., Hunter, M. J., Meininger, D. P. & Komives, E. A. Structure of the fifth EGF-like domain of thrombomodulin: an EGF-like domain with a novel disulfide-bonding pattern. *J. Mol. Biol.* **273**, 913–926 (1997).
44. Klapoetke, S. & Xie, M. H. Disulfide bond characterization of human factor Xa by mass spectrometry through protein-level partial reduction. *J. Pharm. Biomed.* **132**, 238–246 (2017).
45. Sampoli Benitez, B. A. & Komives, E. A. Disulfide bond plasticity in epidermal growth factor. *Proteins: Struct., Funct., Bioinf.* **40**, 168–174 (2000).
46. Hunter, M. J. & Komives, E. A. Thrombin-binding affinities of different disulfide-bonded isomers of the fifth EGF-like domain of thrombomodulin. *Protein Sci.* **4**, 2129–2137 (1995).
47. Goto, Y. et al. Reprogramming the translation initiation for the synthesis of physiologically stable cyclic peptides. *ACS Chem. Biol.* **3**, 120–129 (2008).
48. Sako, Y., Goto, Y., Murakami, H. & Suga, H. Ribosomal synthesis of peptidase-resistant peptides closed by a nonreducible inter-side-chain bond. *ACS Chem. Biol.* **3**, 241–249 (2008).
49. Knott, V., Downing, A. K., Cardy, C. M. & Handford, P. Calcium binding properties of an epidermal growth factor-like domain pair from human fibrillin-1. *J. Mol. Biol.* **255**, 22–27 (1996).
50. Jensen, S. A., Aspinall, G. & Handford, P. A. C-terminal propeptide is required for fibrillin-1 secretion and blocks premature assembly through linkage to domains cbEGF41–43. *Proc. Natl. Acad. Sci. USA* **111**, 10155–10160 (2014).
51. Jensen, S. A., Corbett, A. R., Knott, V., Redfield, C. & Handford, P. A. Ca²⁺-dependent interface formation in fibrillin-1. *J. Biol. Chem.* **280**, 14076–14084 (2005).
52. Hambleton, S. et al. Structural and functional properties of the human notch-1 ligand binding region. *Structure* **12**, 2173–2183 (2004).
53. Wouters, M. A., Fan, S. W. & Haworth, N. L. Disulfides as redox switches: from molecular mechanisms to functional significance. *Antioxid. Redox Signal.* **12**, 53–91 (2010).
54. Campbell, I. D. & Bork, P. Epidermal growth factor-like modules. *Curr. Opin. Struct. Biol.* **3**, 385–392 (1993).
55. Bleijlevens, B. et al. Dynamic states of the DNA repair enzyme AlkB regulate product release. *EMBO Rep.* **9**, 872–877 (2008).
56. Ergel, B. et al. Protein dynamics control the progression and efficiency of the catalytic reaction cycle of the *Escherichia coli* DNA-repair enzyme AlkB. *J. Biol. Chem.* **289**, 29584–29601 (2014).
57. Chowdhury, R. et al. Structural basis for oxygen degradation domain selectivity of the HIF prolyl hydroxylases. *Nat. Commun.* **7**, 12673 (2016).
58. Blatch, G. L. & Lässle, M. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *BioEssays* **21**, 932–939 (1999).
59. Karpenahalli, M. R., Lupas, A. N. & Söding, J. TPRpred: a tool for prediction of TPR-, PPR- and SEL1-like repeats from protein sequences. *BMC Bioinforma.* **8**, 2 (2007).
60. Pekkala, M. et al. The peptide-substrate-binding domain of collagen prolyl 4-hydroxylases is a tetratricopeptide repeat domain with functional aromatic residues. *J. Biol. Chem.* **279**, 52255–52261 (2004).
61. Horton, J. R. et al. Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases. *Nat. Struct. Mol. Biol.* **17**, 38 (2009).
62. Ron, D. & Walter, P. Signal integration in the endoplasmic reticulum unfolded protein response. *Nat. Rev. Mol. Cell Biol.* **8**, 519 (2007).
63. Sunnerhagen, M. S. et al. The effect of aspartate hydroxylation on calcium binding to epidermal growth factor-like modules in coagulation factors IX and X. *J. Biol. Chem.* **268**, 23339–23344 (1993).
64. Wang, N. J. et al. Loss-of-function mutations in Notch receptors in cutaneous and lung squamous cell carcinoma. *Proc. Natl. Acad. Sci. USA* **108**, 17761–17766 (2011).
65. Liu, Y.-F. et al. Somatic mutations and genetic variants of NOTCH1 in head and neck squamous cell carcinoma occurrence and development. *Sci. Rep.* **6**, 24014 (2016).
66. Elkins, J. M. et al. Structure of factor-inhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF-1 α . *J. Biol. Chem.* **278**, 1802–1806 (2003).
67. Strieker, M., Kopp, F., Mahler, C., Essen, L.-O. & Marahiel, M. A. Mechanistic and structural basis of stereospecific C β -hydroxylation in calcium-dependent antibiotic, a daptomycin-type lipopeptide. *ACS Chem. Biol.* **2**, 187–196 (2007).
68. McNeill, L. A. et al. Hypoxia-inducible factor asparaginyl hydroxylase (FIH-1) catalyses hydroxylation at the β -carbon of asparagine-803. *Biochem. J.* **367**, 571–575 (2002).
69. Hewitson, K. S. et al. Evidence that two enzyme-derived histidine ligands are sufficient for iron binding and catalysis by factor inhibiting HIF (FIH). *J. Biol. Chem.* **283**, 25971–25978 (2008).
70. Blasiak, L. C., Vaillancourt, F. H., Walsh, C. T. & Drennan, C. L. Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis. *Nature* **440**, 368–371 (2006).
71. Rose, N. R., McDonough, M. A., King, O. N. F., Kawamura, A. & Schofield, C. J. Inhibition of 2-oxoglutarate dependent oxygenases. *Chem. Soc. Rev.* **40**, 4364–4397 (2011).
72. Gali, H. et al. Chemical synthesis of *Escherichia coli* STh analogues by regioselective disulfide bond formation: Biological evaluation of an 111In-DOTA-Phe19-STh analogue for specific targeting of human colon cancers. *Bioconjug. Chem.* **13**, 224–231 (2002).
73. Kawamura, A. et al. Highly selective inhibition of histone demethylases by de novo macrocyclic peptides. *Nat. Commun.* **8**, 14773 (2017).
74. Yang, H. & Zubarev, R. A. Mass spectrometric analysis of asparagine deamidation and aspartate isomerization in polypeptides. *Electrophoresis* **31**, 1764–1772 (2010).
75. Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: the integration of data reduction and structure solution - from diffraction images to an initial model in minutes. *Acta Cryst.* **D62**, 859–866 (2006).
76. Kabsch, W. XDS. *Acta Cryst.* **D66**, 125–132 (2010).
77. Evans, P. Scaling and assessment of data quality. *Acta Cryst.* **D62**, 72–82 (2006).
78. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Cryst.* **43**, 186–190 (2010).
79. Uson, I. & Sheldrick, G. M. An introduction to experimental phasing of macromolecules illustrated by SHELX; new autotracing features. *Acta Cryst.* **D74**, 106–116 (2018).
80. Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Cryst.* **D59**, 2023–2030 (2003).
81. Abrahams, J. P. & Leslie, A. G. W. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Cryst.* **D52**, 30–42 (1996).
82. Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3**, 1171 (2008).
83. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst.* **D66**, 213–221 (2010).
84. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Cryst.* **D66**, 486–501 (2010).
85. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
86. Gatto, G. J., Geisbrecht, B. V., Gould, S. J. & Berg, J. M. Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5. *Nat. Struct. Mol. Biol.* **7**, 1091–1095 (2000).

Acknowledgements

We thank the Wellcome Trust, Cancer Research UK, and the Biotechnological and Biological Research Council for funding. M.M. thanks the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007–2013) under REA grant agreement number 298603 (Marie Curie IEF) for a Fellowship and R.J.H. acknowledges a William R. Miller Junior Research Fellowship from St Edmund Hall Oxford. L.B. thanks the Deutsche Forschungsgemeinschaft for a fellowship (BR 5486/2-1). S.A.J. and P.A.H. thank Arthritis Research UK for support (20785). U. O. thanks Arthritis Research UK for support (20522). We thank the Diamond Light Source and staff for allocation of beam time and support.

Author contributions

I.P. purified proteins, performed most of the assays, crystallizations, and MS studies (with help of H.K.). L.B. performed assays; T.K. and G.T.K. purified, crystallized, and refined the structure of AspH-Ox under supervision of U.O.; S.A.J. expressed and purified multidomain AspH-substrates under supervision of P.A.H.; L.A.M., K.S.H., and N.J.K. carried out initial biochemical studies on AspH-catalysis; M.M. and I.P. synthesized peptides; R.J.H. performed NMR experiments; M.A.M. performed crystallography for AspH-TPR-Ox structures. C.J.S. designed experiments and wrote the manuscript with help from L.B., I.P. and M.A.M. All authors analysed the data.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-019-12711-7>.

Correspondence and requests for materials should be addressed to C.J.S.

Peer review information *Nature Communications* thanks Stefan Kalkhof, Johanna Myllyharju and the other, anonymous, reviewer for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019