Ribosomal oxygenases are structurally conserved from prokaryotes to humans

Rasheduzzaman Chowdhury, Rok Sekirnik, Nigel C. Brissett, Tobias Krojer, Chia-hua Ho, Stanley S. Ng, Ian J. Clifton, Wei Ge, Nadia J. Kershaw, Gavin C. Fox, João R. C. Muniz, Melanie Vollmar, Claire Phillips, Ewa S. Pilka, Kathryn L. Kavanagh, Frank von Delft, Udo Oppermann, Michael A. McDonough, Aidan J. Doherty & Christopher J. Schofield

2-Oxoglutarate (2OG)-dependent oxygenases have important roles in the regulation of gene expression via demethylation of N-methylated chromatin components and in the hydroxylation of transcription factors and splicing factor proteins. Recently, 2OG-dependent oxygenases that catalyse hydroxylation of transfer RNA and ribosomal proteins have been shown to be important in translation relating to cellular growth, T cell differentiation and translational accuracy. The finding that ribosomal oxygenases (ROXs) occur in organisms ranging from prokaryotes to humans raises questions as to their structural and evolutionary relationships. In Escherichia coli, YcfD catalyses arginine hydroxylation in the ribosomal protein L16; in humans, MYC-induced nuclear antigen (MINA53; also known as MINA) and nucleolar protein 66 (NO66) catalyse histidine hydroxylation in the ribosomal proteins RPL27A and RPL8, respectively. The functional assignments of ROXs open therapeutic possibilities via either ROX inhibition or targeting of differentially modified ribosomes. Despite differences in the residue and protein selectivities of prokaryotic and eukaryotic ROXs, comparison of the crystal structures of E. coli YcfD and Rhodothermus marinus YcfD with those of human MINA53 and NO66 reveals highly conserved folds and novel dimerization modes defining a new structural subfamily of 2OG-dependent oxygenases. ROX structures with and without their substrates support their functional assignments as hydroxylases but not demethylases, and reveal how the subfamily has evolved to catalyse the hydroxylation of different residue side chains of ribosomal proteins. Comparison of ROX crystal structures with those of other JmjC-domain-containing hydroxylases, including the hypoxia-inducible factor asparaginyl hydroxylase FIH and histone \(N\)-methyl lysine demethylases, identifies branch points in 2OG-dependent oxygenase evolution and distinguishes between JmjC-containing hydroxylases and demethylases catalysing modifications of translational and transcriptional machinery. The structures reveal that new protein hydroxylation activities can evolve by changing the coordination position from which the iron-bound substrate-oxidizing species reacts. This coordination flexibility has probably contributed to the evolution of the wide range of reactions catalysed by oxygenases.

To investigate the structural basis of catalytic differences within the ROX subfamily of JmjC-domain-containing hydroxylases and their relationship with the JmjC-containing histone \(N\)-methyl lysine demethylases (KDMs, Fig. 1a), we conducted structural analyses on both prokaryotic (initially YcfD from E. coli (EcYcfD) and subsequently that from the thermophile \(R\. \) marinus (RmYcfD)) and human ROXs (MINA53 \(26\text{--}465\) and NO66 \(183\text{--}641\)). We used RmYcfD to obtain a YcfD substrate structure. All four ROXs showed marked similarities in their folds: the JmjC domain is followed by helical dimerization and carboxy-terminal ‘winged helix’ (WH) domains (Fig. 1b). The ROX JmjC domains consist of 11–12 \(\beta\)-strands, 8 of which (I–VIII) form a double-stranded \(\beta\)-helix (DSBH), which is stereotypical of 2OG-dependent oxygenases (Fig. 1c and Extended Data Fig. 1).

The dimerization domains have a two-fold symmetry and comprise a bundle of three \(\alpha\)-helices (Extended Data Fig. 2); the dimers are stabilized by electrostatic and hydrogen bonding as well as hydrophobic interactions. Consistent with a catalytic role for this domain, dimerization blocking substitutions, EcYcfD(1211R) and MINA53(313E) decrease activity. Hydrogen-bonding and electrostatic interactions are substantially more important in RmYcfD dimerization than for the other ROXs, consistent with the increased occurrence of electrostatic interactions in thermophiles. The ROX C-terminal domains, which are required for activity (Extended Data Fig. 3), are reminiscent of WH domains involved in protein–protein and protein–nucleic-acid interactions; however, their overall negative charge suggests that they may not directly bind nucleic acids. In contrast to ROXs, other JmjC-containing hydroxylases—FIH, \(\text{rRNA}\) yW-synthesizing protein 5 (TYW5), JmjC-domain-containing protein 4 (JMJD4), JMJD5 (ref. 17) and JMJD6 (ref. 4)—do not contain a WH domain (Fig. 2). The combined structures led to the proposal that the ROX fold evolved into those of JmjC-containing hydroxylases and KDMs partly via loss of the WH domain, which enabled the C-terminal helical bundle to take on other roles as in KDMs or the dimerization mode as observed in FIH.

ROX structures were determined in complex with Mn(II) and 2OG or \(\text{N-oxalylglycine (NOG)}\), replacing Fe(II) and 2OG. As for most 2OG-dependent oxygenases, the metal is octahedrally coordinated by a 2-His-1-carboxylate triad from DSBH \(\beta\)-III and \(\beta\)-VI (Fig. 3); two coordination sites are occupied by the 2OG/NOG oxalyl group, leaving one for \(\text{H}_2\text{O}/\text{O}_2\) binding (Fig. 4 and Extended Data Fig. 4). With the YcfDs, the NOG C5 carboxylate is positioned to salt bridge with EcYcfD Arg 140 or RmYcfD Arg 148 on DSBH \(\beta\)-IV (Extended Data Fig. 4). This arrangement is notable because with other 2OG-dependent oxygenases in which the 2OG C5 carboxylate interacts with an Arg residue, it is located on \(\beta\)-VIII. In human ROXs, the 2OG C5-carboxylate-interacting residue is a lysine (MINA53 Lys 194, NO66 Lys 355) from \(\beta\)-IV, in which the most JmjC-containing hydroxylases and KDMs. These observations suggest that eukaryotic JmjC-containing hydroxylases and KDMs evolved from prokaryotic ROXs.

Initial attempts to obtain substrate complexes by co-crystallization or soaking crystals were unsuccessful. We therefore pursued alternatives, one involving using a thermostable YcfD homologue, which we considered may have a relatively low substrate dissociation constant \((K_d)\), enabling complex crystallization. RmYcfD (which has 31% identity with EcYcfD) catalyses L16 fragment (20-residue peptide, amino acids Lys 72–Glu 91) Arg 82 hydroxylation with an approximately sevenfold lower Michaelis constant \((K_m)\) than EcYcfD (268 \(\mu\)M and 1.9 mM, respectively). A RmYcfD–L16–91 structure, obtained by co-crystallization, was solved
The overall folds of the ribosomal oxygenases. a, Reactions catalysed by ROX and related oxygenases. ARD, ankyrin repeat domain; CAD, C-terminal transactivation domain of HIF-α. b, Ribbon representations of EcYcfD, RmYcfD, MINA53 and NO66 homodimers. The monomers contain a JmjC domain with the DSBH core present in all 2OG-dependent oxygenases (blue) followed by dimerization (yellow) and C-terminal WH domains (red). Domain architecture and a schematic representation of the DSBH core β-strands (βI–VIII) that form major (blue, βI, βIII, III and V) and minor sheets (grey, βII, VII, IV and V) is shown boxed. The insert between βIV and βV (purple) is involved in substrate binding. The three Fe-coordinating residues are on the βI and βVII strands (black circles). 2OG is in green sticks; the 2OG C5-carboxylate-binding residue, Arg (YcfDs) or Lys (human ROXs) from βIV is a black circle.

by molecular replacement using the apo EcYcfD structure (Protein Data Bank (PDB) accession 4CCL). The overall EcYcfD and RmYcfD structures are similar (Cα root mean squared deviation (r.m.s.d.) 1.58 Å); L16 residues Lys 77–Lys 85 are visible in the electron density map (Fig. 3c).

For the human ROXs, we used electrospray ionization–mass spectrometry guided disulphide crosslinking19,20 to obtain substrate complexes (Extended Data Fig. 5). Structures were obtained for wild-type NO66–RPL8(G220C) (complex 1), NO66(L299C/C300S)–RPL8(G220C) (complex 2) and NO66(S373C)–RPL8(G214C) (complex 3) pairs. Electron density corresponding to RPL8 residues 215–223 (complex 1), 213–223 (complex 2) and 212–223 (complex 3) was observed at the active site (Fig. 3b and Extended Data Fig. 5). The RPL8 residues (215–219)—including the hydroxylated His 216—adopt near identical conformations (Cα r.m.s.d., 0.29–0.36 Å), implying that all three structures represent catalytically functional complexes (Extended Data Fig. 5). In the light of the NO66–RPL8 structures, we identified a MINA53 residue (Tyr 209) suitable for crosslinking: MINA53(Y209C) crystallized in complex with RPL27A(G37C) with electron density observed for RPL27A residues 36–44 (Fig. 3a). Further validation of the functional relevance of the crosslinked structures comes from comparisons with the wild-type RmYcfD–L16 structure and kinetic studies demonstrating activities with most variants (Extended Data Fig. 6).

MINA53 and NO66 bind their RPL27A and RPL8 substrates in a conserved manner (Cα r.m.s.d., RPL27A38–43, RPL8215–220, 0.8 Å). Comparison of human ROX and RmYcfD complexes reveals similarities in substrate binding, particularly for the hydroxylated residue and for substrate residues to the amino-terminal side of the hydroxylated residue. In all ROX complexes, substrates bind with the same N/C directionality, as observed for FIH17 and for one KDM—plant homeodomain finger 8 (PHF8)23 (and probably other KDM2/7 subfamily members) —but differing from that for most KDMs (KDM4A24, KDM6B25 and KDM6A26) (Fig. 2). The substrates bind in shallow channels on the ROX surfaces and form multiple interactions/hydrogen bonds with residues from DSBH (βI, βII and βVIII, and the extended βIV–βV loop. Although the N-terminal regions of RPL27A (amino acids 36–39), RPL8 (213–216) and L16 (78–81) bind similarly, the C-terminal regions of RPL27A (40–44) and RPL8 (217–223) form more extensive interactions with human ROXs than does L16 (83–85) with RmYcfD (Fig. 3). Notably, both RPL27A and RPL8 substrates make hydrophobic contacts with the WH domains in MINA53 and NO66 (Extended Data Fig. 3). In addition, MINA53 forms a catalytically important salt-bridge interaction between RPL27A Arg 42 and MINA53 Asp 333 located on the α-helix connecting the dimerization and WH domains (Extended Data Figs 6 and 7).

The general binding mode of the hydroxylated residues is conserved between prokaryotic and human ROXs, that is, they bind in deep pockets and the positions of the hydroxylated β-methylene nearly superimpose (Fig. 4). There are, however, clear differences in the way human ROXs and RmYcfD bind their target residue side chains (Fig. 3). With human ROXs, the binding of RPL27A His 39/RPL8 His 216 involves a series of hydrogen bonds to backbone amides or side chains of human ROX residues: MINA53 Gln 136/NO66 Arg 297; MINA53 Asn 165/NO66 Asn 326; MINA53 Tyr 167/NO66 Tyr 328; and MINA53 Ser 257/NO66 Ser 421 (Fig. 3a, b). With RmYcfD, the Arg 82 ‘slots’ into a hydrophobic cleft defined by RmYcfD Tyr 137 and Met 120 side chains and hydrogen bonds to RmYcfD Asp 118 and Ser 208 (Fig. 3c). Mutagenesis studies on ROXs support the observed binding modes of the substrate residues (Extended Data Figs 6 and 8).

There are conflicting reports as to the catalytic activities of some JmjC-containing hydroxylases, including NO66, which has been classified as both a hydroxylase8 and a KDM25. Comparison of ROXs with KDMs and FIH (Figs 2 and 4a) identifies distinctive structural features characteristic of JmjC-containing hydroxylases and KDMs, in addition to the roles of the WH domains. This is important because it supports the assignment of hydroxylase (but not demethylase) activities for ROXs and other human JmjC-containing hydroxylases; for example, FIH17 and JMJD6 (ref. 4). In our assays with isolated MINA53 and NO66 we have consistently not observed enzyme-catalysed demethylation under conditions in which JmjC-containing KDMs are active4. Although we cannot rule out the possibility that some of the JmjC-containing hydroxylases may have KDM activities under different conditions or in cells, the multiple structures reported here suggest that for this to occur, substantial active-site rearrangements would be required on substrate binding.
Figure 2 | Comparison of the substrate structures for ROXs and JmjC-containing enzymes. a–f, Ribbon representations of ROX and related 2OG-dependent oxygenase–substrate complexes. a, MINA53–Mn–2OG–RPL27A(32–50) (PDB accession 2OX0). b, NO66–Mn–NOG–RPL8(205–224) (C2, 2.35 Å). c, RmYcfD–Mn–NOG–L16(72–91) (P22, 2.3, 3.0 Å). d, FIH–Fe–NOG–HIF–H3K9me2(7–14) (PDB accession 1H2K). e, PHF8–Fe–NOG–H3K4me3/K9me2(2–25) (PDB accession 3KV4). PHD, plant homeodomain. f, KDM4A–Ni–NOG–H3K9me2(7–14) (PDB accession 2OX0). For comparison, the DSBH core of each structure is in a similar orientation. Note the directionality of substrate binding in the JmjC domains. The active-site metals (Fe/surrogate) are colour-coded spheres. Analyses of the structures reveal that the ROX overall folds (a–c), oligomerization states and active-site architectures are evolutionarily conserved.

Figure 3 | Features of ROX–substrate binding. a–c, Ribbon representations of MINA53 (a), NO66 (b) and RmYcfD (c) monomers showing difference electron density (Fo – Fc, omit map) for substrates contoured to 3σ (right panels). Left panels depict active-site surface representations, showing key hydrogen bonds and polar interactions (dotted lines) with substrates. a, With MINA53, the RPL27A His 39 imidazole nitrogens form hydrogen bonds with Tyr 167/Ser 257 (NδHis 39—OH-Tyr 167, 2.9 Å; NεHis 39—OεSer 257, 3.1 Å). b, In NO66, RPL8 His 216 is similarly bound in a deep pocket; the RPL8 His 216 imidazole nitrogens form hydrogen bonds with Tyr 328/Ser 421 (NδHis 216—OH-Tyr 328, 3.2 Å; NεHis 216—OεSer 421, 2.7 Å) and hydrophobic interactions with Ile 244 that project its pro-S hydrogen towards the metal (metal—δ-Cδ, 4.4 Å). a. Although MINA53 (a) uses four primary amides—Asn 101, Gln 136, Gln 139 and Asn 165—to interact with RPL27A backbone amides, NO66 (b) uses two arginines (272, 297) to hydrogen bond with the RPL8 Asn 215 side chain and RPL8 His 216 backbone. c, In the RmYcfD–L16 complex, the L16 Arg 82 binds in a pocket defined by the Tyr 137/Met 120 side chains, which form π–cation and hydrophobic interactions with the L16 Arg 82 side chain. The Arg 82 guanidino group makes electrostatic interactions with the RmYcfD Asp 118 carboxylate (O–NH, 2.8–3.1 Å) and hydrogen bonds to RmYcfD Ser 208 (NδArg 82—OHSer 208, 3.5 Å; NδArg 82—COESer 208, 3.2 Å). Although MINA53 Tyr 167 and NO66 Tyr 328 are not positionally related to RmYcfD Tyr 137, the role of the serine (MINA53 Ser 257, NO66 Ser 421, RmYcfD Ser 208, βVIII) in binding the hydroxylated His/Arg is conserved in ROXs. Substitutions of these residues cause marked loss of activity (see Extended Data Fig. 6).
Like ROXs, FIH catalyses β-hydroxylation of an Asn residue in its HIF-α transcription factor substrate and of other residues, including histidines in ankyrins. Superimposition of human ROX and FIH–substrate structures is interesting from catalytic and evolutionary perspectives. Although both FIH and human ROXs catalyse histidine 3S hydroxylation, the positions of their substrate imidazoles is markedly different (Fig. 4 and Extended Data Fig. 9). The positioning of hydroxylated methylenes relative to the metal differs substantially: in the overlaid structures, the angle between the metal and the Cβ atoms of the RPL27A His 39/RPL8 His 216 (human ROX substrate) and HIF-1α Asn 803 (FIH substrate) is ~50° (Fig. 4 and Extended Data Fig. 9), demonstrating that the reactive oxidizing intermediates (Fe(IV) = O) react from different coordination positions in different oxygenases. Studies with 2OG-dependent halogenases have lead to the proposal that iron-bound reactive intermediates abstract a hydrogen from the substrate and deliver a halogen or hydroxyl from different coordination positions to form products. In contrast, our work implies flexibility in the coordination positions with respect to the metal and oxygenase structure of different JmjC-containing hydroxylases from which the ferryl-oxo reacts. Together with other structural considerations, this observation has consequences for the evolution of JmjC-containing oxygenases.

RPL8 (NO66 substrate) has an Asn at the Cβ position relative to the hydroxylated His 216 (YcfD/NO66 substrate) and HIF-1α Asn 803 (FIH substrate) is ~50° in RPL8 His 216. Revealing the extreme sensitivity of oxygenase catalysis to geometric positioning. There is a notable correlation in the binding of RPL8 Asn 215 and HIF-1α Asn 803 to NO66 and FIH, respectively, even though one residue is
hydroxylated and one is not; the primary amides of both RPL8 Asn 215 and HIF-1α Asn 803 hydrogen bond with primary amides, that is, NO66 Asn 376 and FIH Gln 239. Collectively these observations reveal that 2OG-dependent oxygenases can evolve new activities not only by 'directly' altering the nature of enzyme–substrate interactions (including by altering the directionality of substrate binding), but also by changing the coordination position from which the ferryl intermediate reacts. The combined structures reveal that the observed modes of ROX hydroxylations have probably evolved into those of other JmjC-containing hydroxylases and the KDMs, both by altering the coordination position from which the ferryl-oxo react and by engineering the depth of substrate penetration. Structurally informed phylogenetic analyses (Extended Data Fig. 10), coupled to the observation that NO66 is more widely distributed than FIH and MINA53, reveal that prokaryotic YcfDs evolved into NO66, which is a branch point leading to the eukaryotic JmjC-containing hydroxylases and demethylases. 2OG-dependent oxygenases are among the most catalytically flexible of all enzyme families. Recent work has revealed that ROXs react with substrates through a different but evolutionarily related binding mode to FIH. The catalytic capabilities of 2OG-dependent oxygenases for protein oxidations thus probably extend beyond those presently identified.

**METHODS SUMMARY**

Recombinant human MINA53 and NO66 and bacterial EcYcd and RmYcd were produced in *E. coli* and purified by metal affinity/cation exchange and size-exclusion chromatography. Assays comprised incubation with Fe(II), 2OG and substrate followed by mass spectrometry and/or 2OG turnover assays. Crystals were grown by vapour diffusion (Supplementary Table 1) and cryo-cooled in liquid nitrogen. Data were collected on Swiss Light Source X10SA, European Synchrotron Radiation Facility BM16 and Diamond Light Source MX beamlines. MINA53 and EcYcd structures were solved by single-wavelength anomalous diffraction or by single isomorphous replacement with anomalous scattering using SeMet derivatives. The NO66 structure was solved by molecular replacement (MR) using the MINA53 JmjC domain. Phases of the substrate complex structures were solved by MR using apo structures (MINA53 (PDB accession 4BU2), NO66 (PDB accession 4DlQ); Supplementary Tables 2–4).

**Online Content** Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 30 September 2013; accepted 14 March 2014.**

**Published online 11 May 2014.**

1. Klose, R. J. & Zhang, Y. Regulation of histone methylation by demethylimators and demethylasators. *Nature Rev. Mol. Cell Biol.* **8**, 307–318 (2007).
2. Waiprost, L. J., Hopkinson, R. J. & Schofield, C. J. Mechanisms of human histone and nucleic acid demethylases. *Curr. Opin. Chem. Biol.* **16**, 525–534 (2012).
3. Kaelin, W. G. Jr & Ratcliffe, P. J. Oxygen sensing by metazoans: the central role of 2OG-dependent oxygenases for protein oxidations thus probably extend beyond those presently identified. 2OG-dependent oxygenases are among the most catalytically flexible of all enzyme families. Review.
4. Walport, L. J., Hopkinson, R. J. & Schofield, C. J. Mechanisms of human histone and nucleic acid demethylases. *Curr. Opin. Chem. Biol.* **16**, 525–534 (2012).
5. Sinha, K. M., Yasuda, H., Coombes, M. M., Dent, S. Y. & de Crombrugghe, B. Regulation of the osteoblast-specific transcription factor Osterix by NO66, a family of jumonji histone lysine demethylases. *Nature Struct. Mol. Biol.* **17**, 38–43 (2010).
6. Ng, S. S. et al. Crystal structures of histone demethylase JMD2A reveal basis for substrate specificity. *Nature* **448**, 87–91 (2007).
7. Kruidenier, L. et al. A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. *Nature* **488**, 404–408 (2012).
8. Long, T. K. & Yokoyama, S. Structural basis for histone H3 Lys 27 demethylation by UTX/KDM6A. *Genes Dev.* **25**, 2266–2277 (2011).
9. Iyer, L. M., Abhiman, S., de Souza, R. F. & Aravind, L. Origin and evolution of peptide-modifying dioxygenases and identification of the wybutosine hydroxylase/hydroperoxidase. *Nucleic Acids Res.* **38**, 5261–5279 (2010).
10. Henri, J. et al. Elucidation of the Fe(IV)/O2 intermediate in the catalytic cycle of the halogenase SyrB2. *Angew. Chem. Int. Ed.* **50**, 1086–1097 (2011).
11. Hoffart, L. M., Barr, E. W., Guyer, R. B., Bollinger, J. M. Jr & Krebs, C. Direct spectroscopic detection of a C-H/cleaving high-spin Fe(V) complex in a prolyl-4-hydroxylase. *Proc. Natl Acad. Sci. USA* **103**, 14738–14743 (2006).
12. Wong, S. D. et al. Rapid turnover of the Fe(V)/O2 intermediate in the catalytic cycle of the halogenase SyrB2. *Nature* **499**, 320–323 (2013).
13. Yang, M. et al. Substrate selectivity analyses of factor inhibiting hypoxia-inducible factor (FIH) catalyses the post-translational hydroxylation of histidyl residues within ankyrin repeat domains. *FEBS J.* **278**, 1086–1097 (2011).
14. Hausinger, R. P. FeII/2OG-dependent hydroxylase/hydroperoxidase. *Crit. Rev. Biochem. Mol. Biol.* **38**, 29–43 (2003).
15. Elkins, J. M. Structural basis for histone H3 Lys 27 demethylation by UTX/KDM6A. *EMBO J.* **29**, 68–79 (2010).
16. Sengoku, T. & Yokoyama, S. Structural basis for histone H3 Lys 27 demethylation by UTX/KDM6A. *Genes Dev.* **25**, 2266–2277 (2011).
17. Hoffart, L. M., Barr, E. W., Guyer, R. B., Bollinger, J. M. Jr & Krebs, C. Direct spectroscopic detection of a C-H/cleaving high-spin Fe(V) complex in a prolyl-4-hydroxylase. *Proc. Natl Acad. Sci. USA* **103**, 14738–14743 (2006).
18. Wong, S. D. et al. Rapid turnover of the Fe(V)/O2 intermediate in the catalytic cycle of the halogenase SyrB2. *Nature* **499**, 320–323 (2013).
19. Yang, M. et al. Substrate selectivity analyses of factor inhibiting hypoxia-inducible factor (FIH) catalyses the post-translational hydroxylation of histidyl residues within ankyrin repeat domains. *FEBS J.* **278**, 1086–1097 (2011).
20. Iyer, L. M., Abhiman, S., de Souza, R. F. & Aravind, L. Origin and evolution of peptide-modifying dioxygenases and identification of the wybutosine hydroxylase/hydroperoxidase. *Nucleic Acids Res.* **38**, 5261–5279 (2010).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank the Biotechnology and Biological Sciences Research Council, the Wellcome Trust, European Research Council, Medical Research Council, Oxford NIHR Biomedical Research Unit, Cancer Research UK, Arthritis Research UK, Bayer Healthcare, the Rosetree Foundation and the Slovenian Academy of Sciences and Arts for funding; we thank the scientists of the beamlines X10SA (Swiss Light Source) and I02, I03, I04, I04-1 (Diamond Light Source) for assistance. The Structural Genomics Consortium is a registered charity (number 1097737) funded by Abbvie, Boehringer Ingelheim, the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, Elly Lilly, Genome Canada, GlaxoSmithKline, the Ontario Ministry of Economic Development and Innovation, Janssen, Novartis Research Foundation, Pfizer, Takeda and the Wellcome Trust. **Author Contributions** R.C., R.S., N.C.B., C.-H.H., W.G., N.J.K., C.P., S.S.N. and E.S.P. cloned the constructs and purified proteins; R.C. and R.S. performed assays; R.C., R.S., N.C.B., S.S.N., C.-H.H. and E.S.P. crystallized the protein–ligand/substrate complexes; R.C., N.O.B., T.K., S.S.N., I.J.C., G.F.C., K.L.K., F.v.D. and M.A.M. collected and processed X-ray data; T.K.B., T.K., J.R., C.M., M.V. and M.A.M. solved and refined complex structures; R.C., R.S., M.A.M. and C.J.S. analysed data; R.C., U.O., A.J.D. and C.J.S. designed the studies; R.C. and C.J.S. wrote the paper with the help of others. See Supplementary Table 5 for further details.

**Author Information** Atomic coordinates and structure factors for the crystal structures have been deposited in the PDB under accession numbers 2XOV, 4DlQ, 4BU2, 4BFX, 4CJC, 4CCX, 4CEL, 4CMN, 4CCQ, 4UL4, 4Ut, 4UV, 4CSW and 4CUG. Requests and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the comprehensive version of the paper. Correspondence and requests for materials should be addressed to C.J.S. (christopher.schofield@chem.ox.ac.uk) or R.C. (r.chadwick@chem.ox.ac.uk).

©2014 Macmillan Publishers Limited. All rights reserved
METHODS

Recombinant protein production and enzyme assays. Complementary DNA sequences encoding N-terminally truncated MINA53 (amino acids 26–465) and NO66 (amino acids 183–641) were PCR amplified from the Mammalian Gene Collection (MGCC; accession numbers BC014928 and BC011350, respectively) and cloned into pNCIC28-Bsa4 vector. Full-length EcYcDd was cloned into pET-28a (+) vector (Novagen) as previously described5. The RnyfC gene (NCBI gene accession number 8566662) was amplified by PCR from genomic DNA of R. marinus, and was cloned into pGEM-T Easy Vector and then into pET-28a (+). Stratagene’s QuickChange site-directed mutagenesis kit was used to make all ROX mutations using the above constructs as templates.

Wild-type ROX enzymes/variants were produced as native His-tagged proteins in E. coli BL21(DE3) as described49. For crystallization experiments, selenomethionine (SeMet)-derivatized enzymes, SeMet-MINA53 and SeMet-EcYcDd, were produced in E. coli BL21(DE3)-R3-pRARE2 and BL21(DE3) strains, respectively. In general, cells were grown in Le Master media51 (alternatively in Selenomethionine Medium Base plus Nutrient Mix) supplemented with SeMet (40–50 mg ml\(^{-1}\)) and kanamycin (30 mg ml\(^{-1}\)) at 37 °C (while shaking at 200 r.p.m.) until an optical density at 600 nm (OD\(_{600\text{nm}}\)) of 1.2 (SeMet-MINA53) or 0.6 (SeMet-EcYcDd) was reached. Protein expression was then induced with 0.2 mM (SeMet-MINA53) or 1.0 mM (SeMet-EcYcDd) isopropyl-\(\beta\)-D-thiogalactoside (IPTG) and allowed to continue for 18 h at 18 °C. Native/SeMet-derivatized proteins were purified from cell lysates using immobilized Ni\(^{2+}\) affinity chromatography with gradient elution using imidazole and/or ion-exchange chromatography. For YcDd, imidazole was removed by buffer exchange to 50 mM HEPES-Na pH 7.5 using a PD10 desalting column followed by a further purification using Q-Sepharose HP (EcYcDd) or Source Q16 (RmYcDd) anion exchange chromatography. For MINA53 and NO66, the His tag was removed by incubation with TEV protease followed by a final-step purification using size-exclusion chromatography in 50 mM HEPES-Na pH 7.5, 500 mM NaCl, 5% (v/v) glycerol–0.5 mM tris-(2-carboxyethyl)phosphine (TCP). Proteins were concentrated to 10–30 mg ml\(^{-1}\) and were of \(>95\%\) purity, as determined by SDS–PAGE. All columns were supplied by GE Healthcare. Assays were performed as described.

Crystallization, data collection and processing. Crystals of MINA53, NO66, EcYcDd and ROX complexes were grown as described in Supplementary Table 1. In general, crystals were cryoprotected by transferring to a solution of mother liquor supplemented with 20% (v/v) ethylene glycol (MINA53/NO66) or 25% (v/v) glycerol (YcDd) before being cryo-cooled in liquid nitrogen.

As described in Supplementary Tables 2–4, data on native and SeMet-derivatized crystals were collected at 100 K using synchrotron radiation at the Swiss Light Source (SLS) beamline X10SA, European Synchrotron Radiation Facility (ESRF) beamline BM16 and Diamond Light Source (DLS) beamlines. The data were processed as outlined in Supplementary Tables 2–4.

Structure solution and refinement

MINA53 structures. SHAKE-AND-BAKE52 was used to identify five Se positions in the SeMet-MINA53–NOG data set (PDB space group), refinement using a common-atom parameters and phasing was carried out with SHARP53 using the single isomorphous replacement with anomalous scattering (SIRAS) method with MINA53–NOG (native) and SeMet-MINA53-NOG as the derivative data set (Supplementary Table 2). The electron density map after density modification with SOLOMON48 was of good quality; automated model building with ARP/wARP resulted in a \(>80\%\) complete model with one MINA53 molecule per asymmetric unit, which corresponds to an unusually high solvent content of \(\sim75\%\). Refinement was carried out with BUSTER69 and after several cycles of manual rebuilding with COOT63, the model converged to 19.7% R\(_{\text{free}}\) and 22.9% R\(_{\text{free}}\). Atomic coordinates and structure factors for this structure are deposited in the PDB database under the accession number 2DXV.

SeMet-MINA53–2OG structure was solved by using phases from a highly redundant single-wavelength anomalous dispersion (SAD) data set collected around the Se absorption edge. Using Patterson seeding and dual-space direct methods, SHELXD (SHELX pipeline)70 (CCP4 suite)71 located six out of eight possible Se sites. Refinement of substructure solution followed by density modification with SHELXE resulted in good-quality initial phases to 2.8 Å resolution. Automated model building with Buccaneer72 resulted in a model where core regions including the JmJC and dimerization domains were built. Iterative refinement using CNS 1.3 (ref. 40) and model building using COOT63 continued until R\(_{\text{free}}\) was around 30%. Final rounds of manual fitting using COOT63 and refinement using a combination of CNS 1.3 (ref. 40) and PHENIX63 continued until R\(_{\text{free}}\) reached 27.9%, no longer improved (Supplementary Table 2). This structure (deposited in the PDB under accession number 4BU2) was then used as a search model to solve the structure of MINA53(Y209C) in complex with RPL27A(3G37) by molecular replacement (MR) with PHASER62 (P2\(_1\)2\(_1\)2\(_1\), space group, resolution 2.05 Å). The quality of all MINA53 structures was validated using MOLPROBITY14 with \(>95\%\) of the residues in the favoured region of the Ramachandran plot.

NO66 structures. An N-terminally truncated form of MINA53 (amino acids 30–260), comprising the JmJC domain, was used as a search model for MR using PHASER62. The two molecules in the asymmetric unit of NO66 were readily located, but the electron density away from the JmJC core of NO66 was ambiguous. Density modification with RESOLVE66, as implemented in PHENIX63, which took advantage of the two-fold non-crystallographic symmetry (in a P2\(_1\)2\(_1\)2\(_1\) space group), led to a marked map improvement and allowed automated model building with Buccaneer72. Refinement was carried out with REFMAC564 after several cycles of manual rebuilding with COOT63, the model converged to 18.5% R\(_{\text{free}}\) and 23.1% R\(_{\text{free}}\). Atomic coordinates and structure factors for this structure are deposited in the PDB under accession number 4DIQ. The remaining NO66 structures, including those in complex with substrate RPL8, were solved in P2\(_1\) or C2 space groups (resolution 2.15–2.50 Å) with 2–4 molecules per asymmetric unit (Supplementary Table 3) using the NO66/P2\(_1\)/2\(_1\), structure (PDB accession 4DIQ) as a search model. Iterative rounds of model building using COOT63 and refinement using PHENIX63 and/or CNS 1.3 (ref. 40) were performed until the decrease in R\(_{\text{free}}\) and R\(_{\text{free}}\) no longer converged (Supplementary Table 3). All residues were in acceptable regions of Ramachandran plots as calculated by MOLPROBITY14.

YcDd structures. SOLVE65, a comprehensive Python-based system for macromolecular structure determination, solves a search model to locate 17 out of 22 possible Se sites using the SeMet-EcYcDd data set. Eight pairs of sites were related by non-crystallographic symmetry. The initial electron density map after solvent flattening density modification with RESOLVE66 was of good quality and automated model building resulted in a model where core regions (68% of residues in the crystallized protein’s sequence) of both molecules in the asymmetric unit were built. Refinement and fitting cycles were performed using PHENIX63 and COOT63 that converged to a final 19.5% R\(_{\text{free}}\) and R\(_{\text{free}}\) 25.0%. Phasing and refinement statistics are summarized in Supplementary Table 4. Structures of RmYcDd in complex with 1-chloro-4-hydroxyisoquinoline-3-carbonylhydrazine (IOX3) (ref. 46) or substrate 1,6 were solved by MR using the EcYcDd structure as the search model. The structural refinement was carried out with PHENIX with iterative rebuilding of the models using COOT until R\(_{\text{free}}\) converged to final values (Supplementary Table 4).

31. Guerrero, S. A., Hecht, H. J., Hofmann, B., Bleibl, H. & Singh, M. Production of selenomethionine-labelled proteins using simplified culture conditions and generally applicable host/vector systems. Appl. Microbiol. Biotechnol. 56, 718–723 (2001).
32. Miller, R. et al. On the application of the minimal principle to solve unknown structures. Science 259, 1430–1433 (1993).
33. 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 Crystallogr. D 59, 2023–2030 (2003).
34. Abrahams, J. P. & Leslie, A. G. Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta Crystallogr. D 52, 30–42 (1996).
35. Smart, O. S. et al. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. Acta Crystallogr. D 68, 368–380 (2012).
36. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr. D 60, 2126–2132 (2004).
37. Tserka, O., Tserkou, N., Markaki, E., Theodorou, G. & Tomi, A. Structural insights into histone demethylase NO66 in interaction with oestobead specific transcription factor Ostexin and gene repression. J. Biol. Chem. 288, 16430–16437 (2013).
50. Lancaster, D. E. et al. Disruption of dimerization and substrate phosphorylation inhibit factor inhibiting hypoxia-inducible factor (FIH) activity. Biochem. J. 383, 429–437 (2004).

51. Holm, L. & Rosenstrom, P. Dali server: conservation mapping in 3D. Nucleic Acids Res. 38, W545–W549 (2010).

52. Regni, C. A. et al. How the MccB bacterial ancestor of ubiquitin E1 initiates biosynthesis of the microcin C7 antibiotic. EMBO J. 28, 1953–1964 (2009).

53. Schmeing, T. M. et al. The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. Science 326, 688–694 (2009).

54. Wang, S., Engohang-Ndong, J. & Smith, I. Structure of the DNA-binding domain of the response regulator PhoP from Mycobacterium tuberculosis. Biochemistry 46, 14751–14761 (2007).

55. Groisman, E. A. The pleiotropic two-component regulatory system PhoP-PhoQ. J. Bacteriol. 183, 1835–1842 (2001).

56. Aik, W., McDonough, M. A., Thalhammer, A., Chowdhury, R. & Schofield, C. J. Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. Curr. Opin. Struct. Biol. 22, 691–700 (2012).

57. 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).

58. Han, M. V. & Zmasek, C. M. phyloXML: XML for evolutionary biology and comparative genomics. BMC Bioinformatics 10, 356 (2009).

59. Larkin, M. A. et al. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948 (2007).
Extended Data Figure 1 | Schematic protein topologies of ROXs and related 2OG-dependent oxygenases. a–f, Protein topologies of MINA53–Mn–2OG–RPL27A_{32–50} (a), NO66–Mn–NOG–RPL8_{205–224} (b), RmYcfD–Mn–NOG–L16_{72–91} (c), FIH–Fe–NOG–HIF-1α_{786–826} (PDB accession 1H2K) (d), PHF8–Fe–NOG–H3K4me3K9me2_{2–25} (PDB accession 3KV4) (e) and KDM4A–Ni–NOG–H3K9me2_{7–14} (PDB accession 2OX0) (f) (substrates are not shown). DSBH core elements, labelled βI–βVIII, are in green, helices in cyan, additional β-strands in red, random coils in black and the insert between the fourth and fifth β-strands in blue. Note that not all the DSBH oxygenases maintain antiparallel hydrogen-bond pairing between βII and βVII even though the ψ/φ angles (βII) are within the β-region of the Ramachandran plot. Figures were generated using TopDraw47.
Extended Data Figure 2 | ROX dimerization domains. a, Comparison of the dimerization domains in ROXs and FIH. b, Intermolecular interactions observed at dimerization interfaces (monomer A, grey; monomer B, yellow). Validation of the functional relevance of the ROX dimers comes from biochemical and kinetic studies demonstrating loss of activities with most variants. The dimer interfaces in the ROXs are related to that of FIH; we propose that the FIH dimerization fold evolved from that of ROXs. The large buried surface area (>3,000 Å²) within all ROX dimerization domains is sufficient for dimerization in solution, as reported for NO66 (ref. 49). The interactions observed in dimerization include both hydrogen bonds/electrostatic interactions and hydrophobic interactions. In the EcYcfD/RmYcfD dimerization domains, residues involved in hydrophobic interactions are mainly from α2 and are well conserved (RmYcfD residues in parentheses): Phe 214 (Met 223), Val 242 (Ile 250), Met 247 (Leu 255), Leu 250 (Ile 258). Hydrogen bonding/electrostatic interactions are more important in RmYcfD Met 253 (Leu 261), Met 254 (Leu 262), Leu 257 (Leu 265), Ile 258 (Ile 257). 

RmYcfD dimerization domains, residues involved in hydrophobic interactions are mainly from α2 and are well conserved (RmYcfD residues in parentheses): Phe 214 (Met 223), Val 242 (Ile 250), Met 247 (Leu 255), Leu 250 (Ile 258). Hydrogen bonding/electrostatic interactions are more important in RmYcfD Met 253 (Leu 261), Met 254 (Leu 262), Leu 257 (Leu 265), Ile 258 (Ile 257). Hydrogen bonding/electrostatic interactions are more important in RmYcfD Met 253 (Leu 261), Met 254 (Leu 262), Leu 257 (Leu 265), Ile 258 (Ile 257). Hydrogen bonding/electrostatic interactions are more important in RmYcfD Met 253 (Leu 261), Met 254 (Leu 262), Leu 257 (Leu 265), Ile 258 (Ile 257). Hydrogen bonding/electrostatic interactions are more important in RmYcfD Met 253 (Leu 261), Met 254 (Leu 262), Leu 257 (Leu 265), Ile 258 (Ile 257). Hydrogen bonding/electrostatic interactions are more important in RmYcfD Met 253 (Leu 261), Met 254 (Leu 262), Leu 257 (Leu 265), Ile 258 (Ile 257). Hydrogen bonding/electrostatic interactions are more important in RmYcfD Met 253 (Leu 261), Met 254 (Leu 262), Leu 257 (Leu 265), Ile 258 (Ile 257). Hydrogen bonding/electrostatic interactions are more important in RmYcfD Met 253 (Leu 261), Met 254 (Leu 262), Leu 257 (Leu 265), Ile 258 (Ile 257). Hydrogen bonding/electrostatic interactions are more important in RmYcfD Met 253 (Leu 261), Met 254 (Leu 262), Leu 257 (Leu 265), Ile 258 (Ile 257).
Extended Data Figure 3 | Interaction of the ROX C-terminal WH domains with their respective ribosomal protein substrates. a–e. The figure shows how ROX C-terminal domains interact with their substrates. A DALI search indicates that a close structural homologue of the ROX C-terminal domain is the ‘peptide clamp’ (WH) domain of MccB, an enzyme involved in the biosynthesis of the microcin C7 antibiotic. WH domains, a subtype of the helix-turn-helix (HTH) family, are nucleic acid/protein-interacting domains and occur in different cellular pathways, from transcriptional regulation to RNA processing. Although the overall negative charge of ROX WH domains suggests that they may not directly interact with nucleic acids, it is notable that the prokaryotic ribosomal proteins L6, which is located proximal to L16 in intact ribosomes, and the transcriptional regulator PhoP contain WH folds. The latter is interesting because in the *E. coli* K12 genome the *ycfD* gene is located adjacent to those for the PhoP/PhoQ two component signalling system, which is involved in stress responses. a. General topology of the C-terminal WH domain showing two distinct binding sites for L16 (yellow) and RPL27A (magenta)/RPL8 (orange) involving residues either from an N-terminal loop connecting the WH and dimerization domains (as in RmYcfD) or from an extended loop between WH β3–β4 (as in human ROXs). b–e. Comparisons between the WH domains in MccB (b), MINA53 (c), NO66 (d) and EcYcfD (e) showing the interactions observed between this domain and the substrate(s). Note that although both the RPL27A and RPL8 substrates make hydrophobic contacts with the WH domains in MINA53 (Met 405 and Met 406) (c), and NO66 (Val 576 and Tyr 577) (d), RmYcfD uses Arg 285 to form a hydrogen bond with the L16 Met 83 (RmYcfD Arg 285 NH2–L16 Met 83 O, 2.5 Å) (e). Right panels show the partial loss of activity with mutations of MINA53 (M405A), NO66 (Y577A) and EcYcfD (H277C) residues from WH domains. Data show mean and s.e.m. (n = 3).
Extended Data Figure 4 | Comparison of 2OG/co-substrate binding in ROXs and representative 2OG-dependent oxygenases. The identity of the basic residue (Arg or Lys) that binds the 2OG C5 carboxylate via electrostatic interactions is indicated along with which of the eight DSBH (I–VIII) strands it is located on. The occurrence and positioning of the basic Arg/Lys is characteristic of each subfamily. 2OG binding also involves other polar residues including alcohols, that is, a Ser (βVIII, part of the RXS motif as present in, for example, DAOCS, ANS, FTO and algal P4H) or Thr (βII, for example, as in some KDMs: JMJD3, JMJD6, PHF8 and UTX) or Tyr (non-DSBH β-strand, for example, as in FIH, KDM4A, ABH2 and PHD2) and sometimes, water molecule(s) (reviewed in refs 15,56,57). In an analogous position to the serine of the RXS motif (βVIII), human ROXs have histidine residues, MINA53 His 255/NO66 His 417 (βVIII), which form part of a hydrogen-bond network involving MINA53 Thr 255/NO66 Thr 417 (βVIII), a water molecule, and the 2OG carboxylates. Although EcYcfD/RmYcfD has Asn 197/Thr 206 at this position (βVIII), it is the conserved serine from βI (114 in EcYcfD and 122 in RmYcfD) that is positioned to hydrogen bond with the 2OG C5 carboxylate.
Extended Data Figure 5 | Human ROX–substrate complexes showing disulphide crosslinking sites and difference electron density for the substrate residues. a, Strategy adopted to obtain the crosslinked structures (the same strategy can be used for other protein hydroxylases/KDMs). b–d, Different disulphide crosslinking sites (red arrows) that form NO66–RPL8 cysteine–disulphide pairs under equilibrating conditions. Analyses of the 2OG-oxygenase–substrate complexes reveal that substrate residues at ±2 positions relative to the hydroxylated residues make interactions with enzyme residues within a ∼12 Å radius of the metal. To obtain stable NO66–RPL8 complexes, we engineered NO66 variants substituting Cys residues within ∼12 Å radius of the metal at positions considered likely to be involved in substrate binding based on the analyses of other 2OG-oxygenase–substrate structures21,22,26 and the evolutionary/phylogenetic analyses of NO66/NO66-like proteins in eukaryotes. We also substituted Cys residues at ±2 positions on the peptide substrate sequence, relative to the hydroxylated residue. Electrospray ionization–mass spectrometry (ESI–MS) assays were used to identify the best crosslinking yields for the NO66–RPL8 pairs under equilibrating conditions. The following crosslinked pairs were used for crystallization: wild-type NO66 with RPL8(G220C), a double NO66 variant L299C/C300S with RPL8(G220C), and a single NO66 variant S373C with RPL8(G214C). Structures were obtained for wild-type NO66–RPL8(G220C) (complex 1; b), NO66(L299C/C300S)–RPL8(G220C) (complex 2; c), and NO66(S373C)–RPL8(G214C) (complex 3; d) in combination with NOG/Mn(II) in C2 space group, 2.25–2.50 Å resolution with two molecules per asymmetric unit; RPL8 residues 215–223 (complex 1), 213–223 (complex 2) and 212–223 (complex 3) were observed bound to the NO66 active site. e, Superimposition of the three complex structures. Note that the key RPL8 residues (215–219), including the hydroxylated His 216, are observed in near identical conformations (r.m.s.d. 0.29–0.36 Å for Cα atoms); the similarity of the substrate positions in all the three NO66 structures suggests that they all probably represent functional complexes. On the basis of the NO66–RPL8 structures, we identified a MINA53 residue, Y209C, suitable for crosslinking, which we crystallized in complex with RPL27A(G37C) (g). Fobs – Fcalc electron-density maps contoured at 3σ are shown as green (RPL8) and grey (RPL27A) meshes around the substrate residues. To test whether the wild-type/mutant enzymes and altered substrates still function catalytically we carried out endpoint and time-course assays using variable enzyme-to-substrate ratios. f, h, The biochemical data show that for both wild-type NO66 (f) and MINA53 (h) (wild type and Y209C), all the Cys-substituted peptides function as substrates. In the case of MINA53, the Y209C variant with which we obtained the MINA53–RPL27A complex structure is approximately fourfold more active than wild-type MINA53. Data are mean and s.e.m. (n = 3). We also tested wild-type NO66 for reaction between enzyme cysteines and the cysteines of modified substrate peptides by ESI–MS. Despite testing multiple combinations, we only observed disulphide formation in cases where we were also able to obtain crystal structures for substrate complexes. All possible combinations of human ROX wild type or variants and the peptides containing Cys at variable positions were used for the cross-reactivity tests: NO66: wild type, R297C, L299C/C300S, S373C, S421C; RPL8: wild type, G214C, H218C and G220C; MINA53: wild type and Y209C; RPL27A: wild type and G37C. The combined activity and MS analyses suggest that in order to form stable/crystallizable cross-linked complexes, the substrates need to be recognized by the enzyme active sites in a catalytically relevant manner (a).
Extended Data Figure 6 | Mutagenesis analyses of the substrate-binding residues located on the JmjC catalytic domains of MINA53, NO66 and RmYcfD. a–c, MINA53 (a), NO66 (b) and RmYcfD (c) are shown in colour-coded sticks. Left panels show views from the active sites of ROX–substrate complexes and the right panels show the effects of mutations on ROX catalysis. Data are mean and s.e.m. (n = 3). Analyses of ROX–substrate complexes reveal important interactions between ROX and their ribosomal protein substrates. With human ROXs, the binding of ribosomal RPL27A His 39 (light blue)/RPL8 His 216 (orange) involves a series of hydrogen bonds to backbone amides and the side chains of MINA53/NO66 residues: MINA53 Gln 136/NO66 Arg 297, MINA53 Asn 165/NO66 Asn 326, MINA53 Tyr 167/NO66 Tyr 328 and MINA53 Ser 257/NO66 Ser 421. In addition, in the MINA53–RPL27A complex, Leu 38 and Arg 42 of RPL27A make hydrophobic contacts with MINA53 Leu 176 and a salt-bridge interaction with MINA53 Asp 333, respectively. We produced variants of all these residues to investigate their roles on substrate binding. The results of the endpoint assays as well as kinetic studies on the variants (right panels) show that substitution of these residues causes substantial losses of activity. c, In the case of RmYcfD, the hydroxylated residue L16 Arg 82 binds in a hydrophobic cleft lined by RmYcfD Tyr 129 and RmYcfD Ser 208. To test the crystallographically observed binding mode, variants of RmYcfD residues (Asp 118, Met 112, Tyr 129 and Ser 208, highlighted) were prepared in EcYcfD (corresponding to Asp 110, Met 112, Tyr 129 and Ser 199, respectively). Mutagenesis studies on all ROXs support the crystallographically observed binding modes of the substrate residues. The combined biochemical and structural data also provide insights into the substrate selectivity of ROXs over other oxygenases.
Extended Data Figure 7 | Conformational changes on substrate binding in ROX. a–c, Conformational changes at the domain and residue levels in MINA53 (dark salmon and red with/without RPL27A, light blue) (a), NO66 (slate and cyan with/without RPL8, orange) (b) and RmYcfD (grey and split pea with/without L16, yellow) (c). Although the overall movement observed for the C-terminal WH domain on substrate binding is more significant in MINA53 as compared to other ROXs, the RmYcfD structures with and without substrate show marked local changes in the side chains of substrate-binding residues (see below). a, The inset highlights local changes to the active-site region in MINA53 in the presence (green sticks) or absence (yellow sticks) of substrate; MINA53 uses an acidic residue, Asp 333, located on an α-helix connecting the dimerization and WH domains, to form a catalytically important salt-bridge interaction with RPL27A Arg 42. Support for this statement comes from activity analyses on variants of both RPL27A and MINA53. We have previously reported that a mutation of Arg 42 in RPL27A to Ala results in <5% hydroxylation. The D333A variant of MINA53 ablates hydroxylation (almost completely) of native RPL27A in all tested substrate:enzyme ratios (Extended Data Fig. 6). In the substrate-unbound form, MINA53 Asp 333 has two alternative conformations, indicating flexibility. The NO66 substrate RPL8 has an Ile 219 at the analogous position to Arg 42 of RPL27A that makes hydrophobic contacts with the Tyr 577 side chains from the WH domain of NO66 (b). In the case of RmYcfD, the substrate-interacting residues located on the βII–βIII loop (Tyr 137), the βIV–βV insert (Arg 169), the dimerization domain (Arg 212 and Glu 218) and on the loop connecting the dimerization and WH domains (Arg 284) are observed in different conformations in the structures with and without substrate, probably reflecting induced fit on substrate binding (c). Substitutions of these residues have variable effects on ROX catalyses (Extended Data Fig. 6).
Extended Data Figure 8 | Comparison of YcfDs from E. coli and R. marinus.

a–d, Differences between YcfDs from E. coli (green) and R. marinus (grey) are shown. a, Superimposition of EcYcfD and RmYcfD–L16 complex structures showing crystallographically observed differences, particularly in the dimerization and βIV–βV loop regions. The βIV–βV insert is highlighted in crimson red and pink in EcYcfD and RmYcfD, respectively. b, Residue numbering is according to RmYcfD, with the EcYcfD numbering shown in brackets. Note that all of the directly identified substrate-binding residues are strictly conserved between EcYcfD and RmYcfD. However, some residues, particularly those located on the βIV–βV insert including Asp 118, Tyr 137 and Arg 212 in RmYcfD (Asp 110, Tyr 129 and Arg 203 in EcYcfD), are observed in different conformations, suggesting potential roles for these residues in catalysis. c, d, Predicted binding mode of L16 (yellow) to EcYcfD (green). A model complex of EcYcfD with Mn(II), NOG and L16 (residues Pro 77–Lys 84) was generated using EcYcfD-SeMet as the template and by comparison with RmYcfD–L16 and MINA53–RPL27A(32–50) structures. d, Surface representations of the EcYcfD–Mn–NOG–L16(77–84) complex, predicting key hydrogen-bond/polar interactions (dotted lines) with L16. The hydroxylated L16 Arg 81 is predicted to bind in a pocket defined by the Tyr 129 and Met 112 sidechains, which probably form π–cation and hydrophobic interactions with the L16 Arg 81 side chain, as observed in the RmYcfD–L16 crystal structure. The Arg 81 guanidino group is predicted to make electrostatic interactions with the EcYcfD Asp 110 carboxylate and hydrogen bonds to EcYcfD Ser 199. EcYcfD residues Asp 110, Met 112, Tyr 129 and Ser 199 were substituted to test the predicted mode of binding; the assay results are given in Extended Data Fig. 6c.
Extended Data Figure 9 | Comparison of active-site chemistry of ROXs and related enzymes. The figure compares active-site chemistry in representative 2OG-dependent oxygenases and directionality of the peptide substrate binding through the active site. Red/blue arrows indicate hydroxylation/demethylation sites. The active-site metals (Fe/Fe surrogates, Mn or Ni) are in colour-coded spheres.
Extended Data Figure 10 | Phylogenetic relationships of human JmjC 2OG-dependent oxygenases. The figure shows a parsimony tree constructed using Archaeopteryx v.0.9812 (ref. 58) from ClustalW™ aligned protein sequences of human JmjC-containing 2OG-dependent oxygenases showing that distinct branches of JmjC-containing oxygenases exist for hydroxylases (red), demethylases/hydroxylases (light green) and demethylases (blue).