Ribosomal Oxygenases are Structurally Conserved from Prokaryotes to Humans

Rasheduzzaman Chowdhury[1], **Rok Sekirnik**[#1], **Nigel C. Brissett**[#2], **Tobias Krojer**[3], **Chia-hua Ho**[1], **Stanley S. Ng**[3], **Ian J. Clifton**[1], **Wei Ge**[1], **Nadia J. Kershaw**[1], **Gavin C. Fox**[4], **Joao R. C. Muniz**[3], **Melanie Vollmar**[3], **Claire Phillips**[3], **Ewa S. Pilka**[3], **Kathryn L. Kavanagh**[3], **Frank von Delft**[3], **Udo Oppermann**[3,5], **Michael A. McDonough**[1], **Aiden J. Doherty**[2], and **Christopher J. Schofield**[1]

[1]The Department of Chemistry and Oxford Centre for Integrative Systems Biology, University of Oxford, Mansfield Road, Oxford OX1 3TA, U.K.

[2]Genome Damage and Stability Centre, University of Sussex, Brighton BN1 9RQ, U.K.

[3]Structural Genomics Consortium, University of Oxford, Headington, Oxford OX3 7DQ, U.K.

[4]Synchrotron SOLEIL, Saint Aubin, 91192 Gif-sur-Yvette Cedex, France

[5]NIHR Oxford Biomedical Research Unit, Botnar Research Centre, Oxford OX3 7LD, UK

[#] These authors contributed equally to this work.

Abstract

2-Oxoglutarate (2OG)-dependent oxygenases play important roles in the regulation of gene expression via demethylation of *N*-methylated chromatin components[1,2], hydroxylation of transcription factors[3], and of splicing factor proteins[4]. Recently, 2OG-oxygenases that catalyze hydroxylation of tRNA[5-7] and ribosomal proteins[8], have been shown to play roles in translation relating to cellular growth, T_H17-cell differentiation and translational accuracy[9-12]. The finding that the ribosomal oxygenases (ROX) occur in organisms ranging from prokaryotes to humans[8] raises questions as to their structural and evolutionary relationships. In *Escherichia coli*, ycfD catalyzes arginine-hydroxylation in the ribosomal protein L16; in humans, Mina53 (MYC-induced nuclear antigen) and NO66 (Nucleolar protein 66) catalyze histidine-hydroxylation in ribosomal proteins rpL27a and rpL8, respectively. The functional assignments of the ROX open therapeutic

possibilities via either ROX inhibition or targeting of differentially modified ribosomes. Despite differences in residue- and protein-selectivities of prokaryotic and eukaryotic ROX, crystal structures of ycfD and ycfD$_{RM}$ from *E. coli* and *Rhodothermus marinus* with those of human Mina53 and NO66 (hROX) reveal highly conserved folds and novel dimerization modes defining a new structural subfamily of 2OG-oxygenases. ROX structures in complex with/without their substrates, support their functional assignments as hydroxylases, but not demethylases and reveal how the subfamily has evolved to catalyze the hydroxylation of different residue sidechains of ribosomal proteins. Comparison of ROX crystal structures with those of other JmjC-hydroxylases including the hypoxia-inducible factor asparaginyl-hydroxylase (FIH) and histone N^ε-methyl lysine demethylases (KDMs) identifies branchpoints in 2OG-oxygenase evolution and distinguishes between JmjC-hydroxylases and -demethylases catalyzing 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 likely contributed to the evolution of the wide range of reactions catalyzed by iron-oxygenases.

To investigate the structural basis of catalytic differences within the ROX subfamily of JmjC-enzymes and their relationship with JmjC-KDMs, we conducted structural analyses on both prokaryotic (initially ycfD from *E. coli* and subsequently from the thermophile *R. marinus*) and human ROX (Mina53$_{26-465}$ and NO66$_{183-641}$). We used the *R. marinus* ycfD$_{RM}$ to obtain an ycfD-substrate structure. All 4 ROX display striking similarities in their folds: the JmjC domain is followed by helical dimerization and *C*-terminal 'winged-helix' (WH)-domains[13] (Fig. 1b). The ROX JmjC domains consist of 11-12 β-strands, 8 of which (I-VIII) form a double-stranded-β-helix (DSBH), which is stereotypical of 2OG-oxygenases (Fig. 1c and Extended Data Fig. 1)[14,15].

The dimerization domains have a 2-fold symmetry and comprise a bundle of 3 α-helices (Extended Data Fig. 2); dimers are stabilized by electrostatic/hydrogen-bonding and hydrophobic interactions. Consistent with a catalytic role for this domain, dimerization blocking substitutions, I211R$_{ycfD}$ and R313E$_{Mina53}$ decrease activity. Hydrogen-bonding/ electrostatic interactions are substantially more important in ycfD$_{RM}$ dimerization than for the other ROX, consistent with the increased occurrence of electrostatic interactions in thermophiles[16]. 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[13]; however their overall negative charge suggests they do not directly bind nucleic acids. In contrast to ROX, other JmjC-hydroxylases (FIH[17], tRNA wybutosine-synthesizing enzyme 5/TYW5[6], JmjC-domain containing protein 4/JMJD4[9], JMJD5[18], JMJD6[4]) and KDMs do not contain a WH domain (Fig. 2). The combined structures lead to the proposal that ROX fold evolved into those of JmjC-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[17].

ROX structures were determined in complex with Mn(II) and 2OG/*N*-oxalylglycine (NOG), replacing Fe(II) and 2OG. As for most 2OG-oxygenases, the metal is octahedrally coordinated by a 2-His-1-carboxylate triad from DSBH-βII and - βVII[14,15] (Fig. 3); 2

coordination sites are occupied by the 2OG/NOG oxalyl group leaving one for H_2O/O_2 binding (Fig. 4 and Extended Data Fig. 4). With the ycfDs the NOG C5-carboxylate is positioned to salt bridge with $Arg140_{ycfD}/Arg148_{ycfDRM}$ on DSBH-βIV (Extended Data Fig. 4). This arrangement is notable because with other 2OG-oxygenases where the 2OG C5-carboxylate interacts with an Arg-residue, it is located on βVIII[14,15]. In hROX, the 2OG C5-carboxylate interacting residue is a lysine ($Lys194_{Mina53}/Lys355_{NO66}$) from βIV, as in most JmjC-hydroxylases and KDMs. These observations lead to the proposal that the eukaryotic JmjC-hydroxylases/KDMs evolved from prokaryotic ycfDs/ROX.

Initial attempts to obtain substrate complexes by co-crystallization/soaking crystals were unsuccessful. We therefore pursued alternatives, one involving using a thermostable ycfD homolog, which we considered may have a relatively low substrate K_d, enabling complex crystallization. The *R. marinus* $ycfD_{RM}$ (31% identity with ycfD) catalyzes L16 fragment (20-mer, aa Lys72-Glu91) Arg82-hydroxylation with ~7 fold lower K_m relative to ycfD (268 μM and 1.9 mM, respectively). An $ycfD_{RM}·L16_{72-91}$ structure, obtained by co-crystallization was solved by molecular replacement using the apo-ycfD structure (PDB: 4CCL). The $ycfD/ycfD_{RM}$ structures are similar (Cα rmsd: 1.58 Å); L16 residues Lys77-Lys85 are visible in the electron density map (Fig. 3c).

For the hROX, we employed electrospray-ionization mass spectrometry guided disulfide cross-linking[19,20] to obtain substrate complexes (Extended Data Fig. 5). Structures were obtained for wild-type (wt) NO66·rpL8-G220C (complex 1), L299C/C300S-NO66·rpL8-G220C (complex 2), and S373C-NO66·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 that hydroxylated, i.e. His216, adopt near identical conformations (Cα rmsd, 0.29-0.36 Å) implying all 3 structures represent catalytically functional complexes (Extended Data Fig. 5). In the light of NO66·rpL8 structures, we identified a Mina53 residue (Tyr209) suitable for cross-linking: Y209C-Mina53 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 cross-linked structures comes from comparisons with the wt-$ycfD_{RM}·L16$ structure and kinetic studies demonstrating activities with most variants (Extended Data Fig. 6).

Mina53/NO66 bind their rpL27a/rpL8 substrates in a conserved manner (Cα rmsd, $rpL27a_{38-43}$, $rpL8_{215-220}$, 0.8 Å). Comparison of hROX and $ycfD_{RM}$ complexes reveals similarities in substrate binding particularly for the hydroxylated-residue and that to its *N*-terminus. In all ROX complexes, substrates bind with the same *N*-/*C*-directionality, as for FIH[17] and for one KDM, plant homeo domain finger 8/PHF8[21] (and likely other KDM2/7 subfamily members) but differing from that for most KDMs (JMJD2A/KDM4A[22], JMJD3/KDM6B[23], KDM6A[24]) (Fig. 2). The substrates bind in shallow channels on the ROX surfaces and form multiple interactions/hydrogen-bonds with residues from DSBH-βI, -βII, -βVIII, and the extended βIV-βV loop. While the *N*-terminal regions of rpL27a (aa 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 hROX than does L16 (83-85) with $ycfD_{RM}$ (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 $Arg42_{rpL27a}$ and $Asp333_{Mina53}$ 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 hROX, i.e. they bind in deep pockets and the positions of the hydroxylated β-methylenes nearly superimpose (Fig. 4). There are, however, clear differences in the way hROX and $ycfD_{RM}$ bind their target residue sidechains (Fig. 3). With hROX, the binding of $His39_{rpL27a}$/$His216_{rpL8}$ involve a series of hydrogen-bonds to backbone amides/sidechains of hROX residues: $Gln136_{Mina53}$/$Arg297_{NO66}$, $Asn165_{Mina53}$/$Asn326_{NO66}$, $Tyr167_{Mina53}$/$Tyr328_{NO66}$ and $Ser257_{Mina53}$/$Ser421_{NO66}$ (Fig. 3a, b). With $ycfD_{RM}$, the $Arg82$ 'slots' into a hydrophobic cleft defined by $Tyr137_{ycfDRM}$ and $Met120_{ycfDRM}$ sidechains and hydrogen-bonds to $Asp118_{ycfDRM}$ and $Ser208_{ycfDRM}$ (Fig. 3c). Mutagenesis studies on ROX 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-hydroxylases, including NO66 which has been assigned as both a hydroxylase[8] and a KDM[25]. Comparison of the ROX and KDM/FIH (Figs 2 and 4a) identifies distinctive structural features characteristic of JmjC-hydroxylases and -KDMs, in addition to the roles of the WH-domains (see below). This is important because it supports the assignment of hydroxylase (but not demethylase) activities for ROX and other human JmjC-hydroxylases e.g. FIH[17], JMJD6[4]. In our assays with isolated Mina53 and NO66 we have consistently not observed enzyme-catalyzed demethylation under conditions where JmjC-KDMs are active[8]. While we cannot rule out the possibility that some of the JmjC-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. In ROX and KDM4A·H3K9me$_2$ (PDB: 2OX0)[22] complex structures the different substrates bind with 'opposite' *N*- to *C*-directionalities with respect to the catalytic machinery. The histone K9me$_n$ side chain is positioned similarly to the ROX-hydroxylated residue sidechains; however, because KDM-catalyzed hydroxylations occur at $N^ε$-methyl lysine-residue termini, their target residues do not penetrate as far into the enzyme active site (Fig. 4a). The ROX also lack 2 flexible loops linking α4-βI (aa 164-175) and α9-α10 (aa 302-317) in KDM4A, which are conserved in KDMs[21-24] and which form important interactions with the Kme$_n$ sidechain, illustrating how the 'core' ROX-fold has been modified by evolution to accommodate the Kme$_n$ sidechain.

Like ROX, FIH catalyzes β-hydroxylation of an Asn-residue in its HIF-α transcription factor substrate[17] and of other residues including histidines in ankyrins[26]. Superimposition of hROX/FIH-substrate structures is interesting from catalytic and evolutionary perspectives. Although both FIH/hROX catalyze histidine 3*S*-hydroxylation, the positions of their substrate imidazoles is strikingly 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 $His39_{rpL27a}$/$His216_{rpL8}$ (hROX substrate) and $Asn803_{HIF1α}$ (FIH substrate) is ~50° (Fig. 4 and Extended Data Fig. 9), demonstrating that the reactive oxidizing intermediates (Fe(IV)=O)[15,27] react

from different coordination positions in different oxygenases. Studies with 2OG-dependent halogenases have led to the proposal that iron-bound reactive intermediates abstract a hydrogen from substrate and deliver a halogen/hydroxyl from different coordination positions to form products[28]. In contrast, our work implies flexibility in the coordination positions with respect to the hydrogen abstracted in different JmjC-hydroxylases from which the ferryl-oxo reacts. Together with other structural considerations, this observation has consequences for the evolution of the JmjC-enzymes.

The rpL8 (NO66 substrate) has an Asn at the −1 position relative to the hydroxylated His216 (ycfD/Mina53 substrates have hydrophobic residues at the analogous positions). The $Asn215_{rpL8}$ methylene is only slightly (0.5 Å) further from the metal than that of $His216_{rpL8}$, revealing exquisite sensitivity of oxygenase catalysis to geometric positioning. There is a striking correlation in the binding of $Asn215_{rpL8}$/$Asn803_{HIF1\alpha}$ to NO66/FIH, even though one residue is hydroxylated and one not; the primary amides of both $Asn215_{rpL8}$/$Asn803_{HIF-1\alpha}$ hydrogen-bond with primary amides, i.e. $Asn376_{NO66}$ and $Gln239_{FIH}$. Collectively these observations reveal 2OG-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 mode of ROX hydroxylations have likely evolved into the FIH/related JmjC Lys-hydroxylation and KDM type hydroxylation/ demethylation modes both by altering the coordination position from which the ferryl-oxo reacts 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/Mina53, reveal that the prokaryotic ycfDs evolved into NO66, which is a branchpoint leading to the eukaryotic JmjC-hydroxylases/demethylases. 2OG-oxygenases are amongst the most catalytically flexible of all enzyme-families. Recent work reveals that FIH manifests remarkable catalytic promiscuity, including the ability to oxidize Asn- and His-residues[29]. Our structural studies reveal that ROX react with substrates via a different but evolutionarily related binding mode to FIH. The catalytic capabilities of 2OG-oxygenases for protein oxidations thus likely extend beyond those presently identified.

METHODS

Recombinant protein production and enzyme assays

cDNA sequences encoding *N*-terminally truncated Mina53 (aa 26-465) and NO66 (aa 183-641) were PCR amplified from Mammalian Gene Collection (MGC) (accession no.: BC014928 and BC011350, respectively) and cloned into pNIC28-Bsa4 vector. Full length ycfD was cloned into pET-28a(+) vector (Novagen, Madison, WI, U.S.A.) as described[8]. *ycfD$_{RM}$* gene (NCBI GENE ID: 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_6-tagged proteins in *Escherichia coli* BL21(DE3) as described[8]. For crystallization experiments, selenomethionine (SeMet) derivatized enzymes, SeMet-Mina53 and SeMet-ycfD were produced in *E. coli* BL21(DE3)-R3-pRARE2 and BL21(DE3) strains, respectively. In general, cells were grown in Le Master media[31] (alternatively in SelenoMethionine Medium Base plus Nutrient Mix) supplemented with selenomethionine (40-50 mg·mL^{-1}) and kanamycin (30 μg·mL^{-1}) at 37 °C (while shaking at 200 rpm) until an OD_{600} (optical density at 600 nm) of 1.2 (SeMet-Mina53) or 0.6 (SeMet-ycfD) was reached. Protein expression was then induced with 0.2 mM (SeMet-Mina53) or 1.0 mM (SeMet-ycfD) isopropyl β-*D*-1-thiogalactopyranoside (IPTG) and allowed to continue for 18 h at 18 °C. All 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 ycfDs, 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 (ycfD) or SourceQ 16 (ycfD$_{RM}$) anion exchange chromatography. For Mina53 and NO66, the His_6-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 TCEP (tris(2-carboxyethyl)phosphine). 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[8].

Crystallization, data collection and processing

Crystals of Mina53, NO66, ycfD and ycfD$_{RM}$ 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 (ycfDs) before being cyro-cooled in liquid N_2.

As described in Supplementary Tables 2-4, data on native and SeMet-derivatized crystals were collected at 100K 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-BAKE[32] was used to identify 5 Se-positions in the SeMet-Mina53 dataset (*P*4$_3$32 space group); refinement of heavy atom parameters and phasing was carried out with SHARP[33] using the single ismorphous replacement with anomalous scattering (SIRAS) method with Mina53.NOG (native) as the native and SeMet-Mina53 as the derivative dataset (Supplementary Table 2). The electron density map after density modification with SOLOMON[34] 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 ~75%. Refinement was carried out with BUSTER[35] and after several cycles of manual rebuilding with COOT[36], the model converged to 19.7% R_{cryst} and 22.9% R_{free}. Atomic coordinates

and structure factors for this structure are deposited in the PDB database with the accession code 2XDV.

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 (SHELXCDE pipeline[37]/ CCP4 suite[38]) located 6 out of 8 possible Se-sites using a SeMet-Mina53 SAD data set. Refinement of substructure solution followed by density modification with SHELXE[37] resulted in good-quality initial phases to 2.8 Å resolution. Automated model building with BUCCANEER[39] resulted in a model where core regions including the JmjC and dimerization domains were built. Iterative refinement using CNS 1.3[40] and model building using COOT[36] continued until R_{free} was around 30%. Final rounds of manual fitting using COOT[36] and refinement using a combination of CNS 1.3[40] and PHENIX[41] continued until R_{cryst}/R_{free} no longer improved (Supplementary Table 2). This structure (deposited with PDB ID: 4BU2) was then used as a search model to solve the structure of Y209C Mina53 in complex with rpL27aG37C by molecular replacement with PHASER[42] ($P2_12_12_1$ space group, resolution, 2.05 Å). The quality of all Mina53 structures was validated using MOLPROBITY[43] with >95% of the residues in the favored region of the Ramachandran plot.

NO66 structures—An *N*-terminally truncated form of Mina53 (aa 30-260), comprising the JmjC domain, was used as a search model for MR using PHASER[42]. 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 RESOLVE[44] as implemented in PHENIX[41], which took advantage of the 2-fold non-crystallographic symmetry (in a $P2_12_12$ space group), led to a significant map improvement and allowed automated model building with BUCCANEER[39]. Refinement was carried out with REFMAC[45]; after several cycles of manual rebuilding with COOT[36], the model converged to 18.5% R_{cryst} and 23.1% R_{free}. Atomic coordinates and structure factors for this structure are deposited in the PDB (accession code 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/$P2_12_12$ structure (PDB ID: 4DIQ) as a search model. Iterative rounds of model building using COOT[36] and refinement using PHENIX[41] and/or CNS 1.3[40] were performed until the decreasing R_{cryst} and R_{free} no longer converged (Supplementary Table 3). All residues were in acceptable regions of Ramachandran plots as calculated by MOLPROBITY[43].

YcfD structures—SOLVE was used to locate 17 out of 22 possible Se-sites using the SeMet-ycfD dataset. 8 pairs of sites were related by non-crystallographic symmetry. The initial electron density map after solvent flattening density modification with RESOLVE[44] was of good quality and automated model building resulted in a model where core regions (60% of residues in the crystallized protein's sequence) of both molecules in the asymmetric unit were built. Refinement and fitting cycles were performed using PHENIX[41] and COOT[36] that converged to a final 19.5% R_{cryst} and R_{free} 25.0%. Phasing and refinement

statistics are summarized in Supplementary Table 4. Structures of ycfD$_{RM}$ in complex with IOX3[46] or substrate L16 were solved by MR using the ycfD structure as the search model. The structural refinement was carried out with PHENIX with iterative rebuilding of the models using COOT until R$_{cryst}$/R$_{free}$ converged to final values (Supplementary Table 4).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

REFERENCES

1. Klose RJ, Zhang Y. Regulation of histone methylation by demethylimination and demethylation. Nat. Rev. Mol. Cell. Biol. 2007; 8:307–318. [PubMed: 17342184]
2. Walport LJ, Hopkinson RJ, Schofield CJ. Mechanisms of human histone and nucleic acid demethylases. Curr. Opin. Chem. Biol. 2012; 16:525–534. [PubMed: 23063108]
3. Kaelin WG Jr. Ratcliffe PJ. Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. Mol. Cell. 2008; 30:393–402. [PubMed: 18498744]
4. Webby CJ, et al. Jmjd6 catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing. Science. 2009; 325:90–93. [PubMed: 19574390]
5. Fu Y, et al. The AlkB domain of mammalian ABH8 catalyzes hydroxylation of 5-methoxycarbonylmethyluridine at the wobble position of tRNA. Angew. Chem. Int. Ed. Engl. 2010; 49:8885–8888. [PubMed: 20583019]
6. Kato M, et al. Crystal structure of a novel JmjC-domain-containing protein, TYW5, involved in tRNA modification. Nucleic Acids Res. 2011; 39:1576–1585. [PubMed: 20972222]
7. van den Born E, et al. ALKBH8-mediated formation of a novel diastereomeric pair of wobble nucleosides in mammalian tRNA. Nat. Commun. 2011; 2:172. [PubMed: 21285950]
8. Ge W, et al. Oxygenase-catalyzed ribosome hydroxylation occurs in prokaryotes and humans. Nat. Chem. Biol. 2012; 8:960–962. [PubMed: 23103944]
9. Feng T, et al. Optimal translational termination requires C4 lysyl hydroxylation of eRF1. Mol. Cell. 2014; 53:645–654. [PubMed: 24486019]
10. Henri J, et al. Structural and functional insights into *Saccharomyces cerevisiae* Tpa1, a putative prolylhydroxylase influencing translation termination and transcription. J. Biol. Chem. 2010; 285:30767–30778. [PubMed: 20630870]
11. Loenarz C, et al. Hydroxylation of the eukaryotic ribosomal decoding center affects translational accuracy. Proc. Natl. Acad. Sci. U. S. A. 2014 doi:10.1073/pnas.1311750111.
12. Yosef N, et al. Dynamic regulatory network controlling T$_H$17 cell differentiation. Nature. 2013; 496:461–468. [PubMed: 23467089]
13. Teichmann M, Dumay-Odelot H, Fribourg S. Structural and functional aspects of winged-helix domains at the core of transcription initiation complexes. Transcription. 2012; 3:2–7. [PubMed: 22456313]
14. Hausinger RP. FeII/α-ketoglutarate-dependent hydroxylases and related enzymes. Crit. Rev. Biochem. Mol. Biol. 2004; 39:21–68. [PubMed: 15121720]

15. Clifton IJ, et al. Structural studies on 2-oxoglutarate oxygenases and related double-stranded beta-helix fold proteins. J. Inorg. Biochem. 2006; 100:644–669. [PubMed: 16513174]

16. Szilagyi A, Zavodszky P. Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. Structure. 2000; 8:493–504. [PubMed: 10801491]

17. Elkins JM, et al. Structure of factor-inhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF-1α. J. Biol. Chem. 2003; 278:1802–1806. [PubMed: 12446723]

18. Del Rizzo PA, Krishnan S, Trievel RC. Crystal structure and functional analysis of JMJD5 indicate an alternate specificity and function. Mol. Cell. Biol. 2012; 32:4044–4052. [PubMed: 22851697]

19. Yang CG, et al. Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA. Nature. 2008; 452:961–965. [PubMed: 18432238]

20. Woon EC, et al. Linking of 2-oxoglutarate and substrate binding sites enables potent and highly selective inhibition of JmjC histone demethylases. Angew. Chem. Int. Ed. Engl. 2012; 51:1631–1634. [PubMed: 22241642]

21. Horton JR, et al. Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases. Nat. Struct. Mol. Biol. 2010; 17:38–43. [PubMed: 20023638]

22. Ng SS, et al. Crystal structures of histone demethylase JMJD2A reveal basis for substrate specificity. Nature. 2007; 448:87–91. [PubMed: 17589501]

23. Kruidenier L, et al. A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. Nature. 2012; 488:404–408. [PubMed: 22842901]

24. Sengoku T, Yokoyama S. Structural basis for histone H3 Lys 27 demethylation by UTX/KDM6A. Genes Dev. 2011; 25:2266–2277. [PubMed: 22002947]

25. Sinha KM, Yasuda H, Coombes MM, Dent SY, de Crombrugghe B. Regulation of the osteoblast-specific transcription factor Osterix by NO66, a Jumonji family histone demethylase. EMBO J. 2010; 29:68–79. [PubMed: 19927124]

26. Yang M, et al. Factor-inhibiting hypoxia-inducible factor (FIH) catalyses the post-translational hydroxylation of histidinyl residues within ankyrin repeat domains. Febs J. 2011; 278:1086–1097. [PubMed: 21251231]

27. Hoffart LM, Barr EW, Guyer RB, Bollinger JM Jr. Krebs C. Direct spectroscopic detection of a C-H-cleaving high-spin Fe(IV) complex in a prolyl-4-hydroxylase. Proc. Natl. Acad. Sci. U. S. A. 2006; 103:14738–14743. [PubMed: 17003127]

28. Wong SD, et al. Elucidation of the Fe(IV)=O intermediate in the catalytic cycle of the halogenase SyrB2. Nature. 2013; 499:320–323. [PubMed: 23868262]

29. Yang M, et al. Substrate selectivity analyses of factor inhibiting hypoxia-inducible factor. Angew. Chem. Int. Ed. Engl. 2013; 52:1700–1704. [PubMed: 23296631]

30. Iyer LM, Abhiman S, de Souza RF, Aravind L. Origin and evolution of peptide-modifying dioxygenases and identification of the wybutosine hydroxylase/hydroperoxidase. Nucleic Acids Res. 2010; 38:5261–5279. [PubMed: 20423905]

31. Guerrero SA, Hecht HJ, Hofmann B, Biebl H, Singh M. Production of selenomethionine-labelled proteins using simplified culture conditions and generally applicable host/vector systems. Appl. Microbiol. Biotechnol. 2001; 56:718–723. [PubMed: 11601620]

32. Miller R, et al. On the application of the minimal principle to solve unknown structures. Science. 1993; 259:1430–1433. [PubMed: 8451639]

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 Biol. Crystallogr. 2003; 59:2023–2030. [PubMed: 14573958]

34. Abrahams JP, Leslie AG. Methods used in the structure determination of bovine mitochondrial F1 ATPase. Acta Crystallogr. D Biol. Crystallogr. 1996; 52:30–42. [PubMed: 15299723]

35. Smart OS, et al. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. Acta Crystallogr. D Biol. Crystallogr. 2012; 68:368–380. [PubMed: 22505257]

36. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 2010; 66:486–501. [PubMed: 20383002]

37. Sheldrick GM. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. Acta Crystallogr. D Biol. Crystallogr. 2010; 66:479–485. [PubMed: 20383001]

38. Winn MD, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 2011; 67:235–242. [PubMed: 21460441]

39. Cowtan K. The Buccaneer software for automated model building. 1 Tracing protein chains. Acta Crystallogr. D Biol. Crystallogr. 2006; 62:1002–1011. [PubMed: 16929101]

40. Brunger AT, et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. 1998; 54:905–921. [PubMed: 9757107]

41. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 2010; 66:213–221. [PubMed: 20124702]

42. McCoy AJ, et al. Phaser crystallographic software. J. Appl. Crystallogr. 2007; 40:658–674. [PubMed: 19461840]

43. Chen VB, et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 2010; 66:12–21. [PubMed: 20057044]

44. Terwilliger TC. Automated main-chain model building by template matching and iterative fragment extension. Acta Crystallogr. D Biol. Crystallogr. 2003; 59:38–44. [PubMed: 12499537]

45. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. 1997; 53:240–255. [PubMed: 15299926]

46. Chowdhury R, et al. Selective Small Molecule Probes for the Hypoxia Inducible Factor (HIF) Prolyl Hydroxylases. ACS Chem. Biol. 2013; 8:1488–1496. [PubMed: 23683440]

47. Bond CS. TopDraw: a sketchpad for protein structure topology cartoons. Bioinformatics. 2003; 19:311–312. [PubMed: 12538265]

48. Cockman ME, et al. Posttranslational hydroxylation of ankyrin repeats in IkappaB proteins by the hypoxia-inducible factor (HIF) asparaginyl hydroxylase, factor inhibiting HIF (FIH). Proc. Natl. Acad. Sci. U. S. A. 2006; 103:14767–14772. [PubMed: 17003112]

49. Tao Y, et al. Structural insights into histone demethylase NO66 in interaction with osteoblast specific transcription factor Osterix and gene repression. J. Biol. Chem. 2013 [PubMed: 23620590]

50. Lancaster DE, et al. Disruption of dimerization and substrate phosphorylation inhibit factor inhibiting hypoxia-inducible factor (FIH) activity. Biochem. J. 2004; 383:429–437. [PubMed: 15239670]

51. Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. Nucleic Acids Res. 2010; 38:W545–549. [PubMed: 20457744]

52. Regni CA, et al. How the MccB bacterial ancestor of ubiquitin E1 initiates biosynthesis of the microcin C7 antibiotic. EMBO J. 2009; 28:1953–1964. [PubMed: 19494832]

53. Schmeing TM, et al. The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. Science. 2009; 326:688–694. [PubMed: 19833920]

54. Wang S, Engohang-Ndong J, Smith I. Structure of the DNA-binding domain of the response regulator PhoP from *Mycobacterium tuberculosis*. Biochemistry. 2007; 46:14751–14761. [PubMed: 18052041]

55. Groisman EA. The pleiotropic two-component regulatory system PhoP-PhoQ. J. Bacteriol. 2001; 183:1835–1842. [PubMed: 11222580]

56. Aik W, McDonough MA, Thalhammer A, Chowdhury R, Schofield CJ. Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. Curr. Opin. Struct. Biol. 2012; 22:691–700. [PubMed: 23142576]

57. McDonough MA, Loenarz C, Chowdhury R, Clifton IJ, Schofield CJ. Structural studies on human 2-oxoglutarate dependent oxygenases. Curr. Opin. Struct. Biol. 2010; 20:659–672. [PubMed: 20888218]

58. Han MV, Zmasek CM. phyloXML: XML for evolutionary biology and comparative genomics. BMC bioinformatics. 2009; 10:356. [PubMed: 19860910]

59. Larkin MA, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23:2947–2948. [PubMed: 17846036]

60. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr. D Biol. Crystallogr. 2011; 67:271–281. [PubMed: 21460445]

61. Kabsch W. Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. D Biol. Crystallogr. 2010; 66:133–144. [PubMed: 20124693]

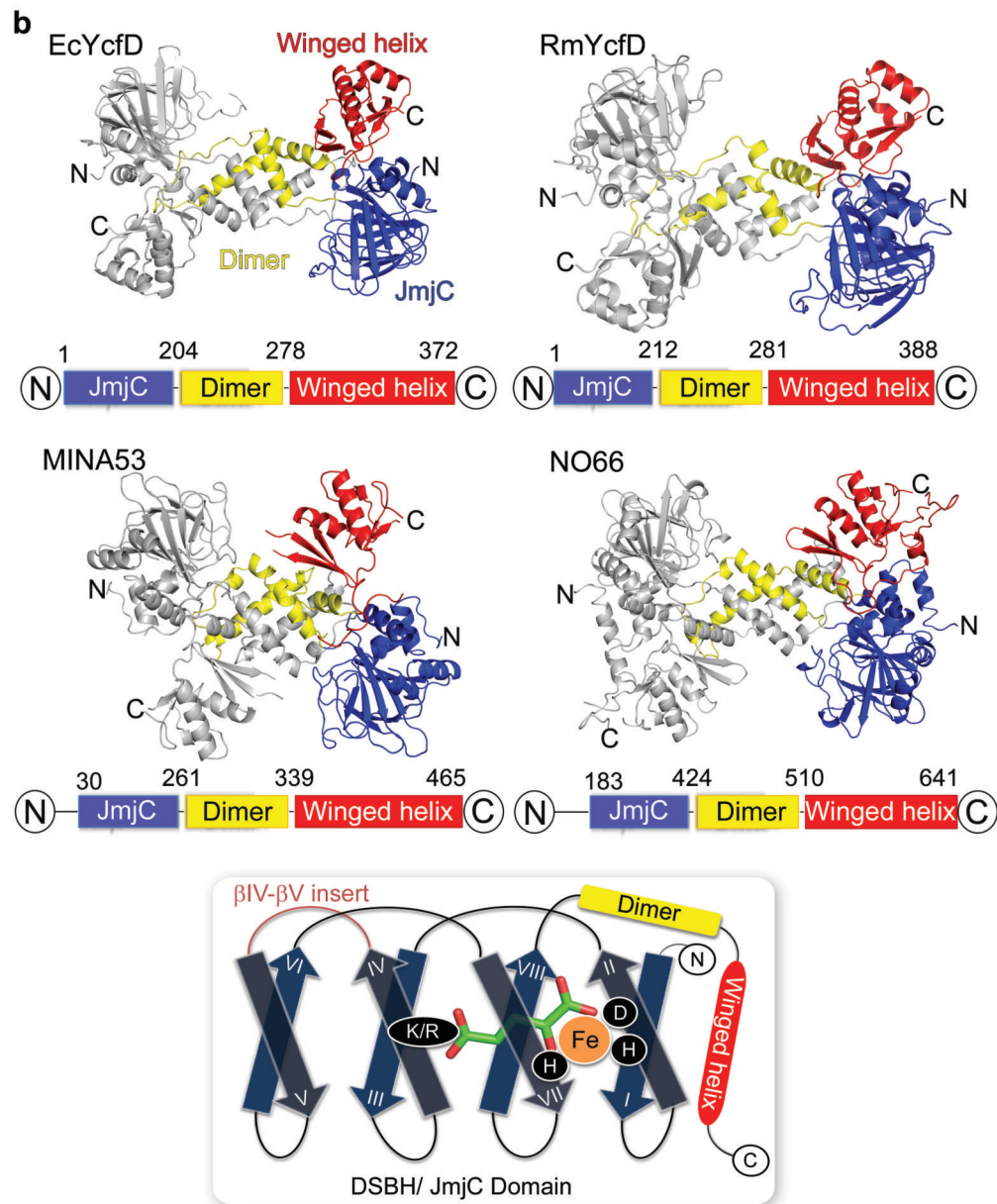62. Otwinowski, Z.; Minor, W.; C. W., C, Jr.. Methods in Enzymology. Vol. 276. Academic Press; 1997. p. 307-326.
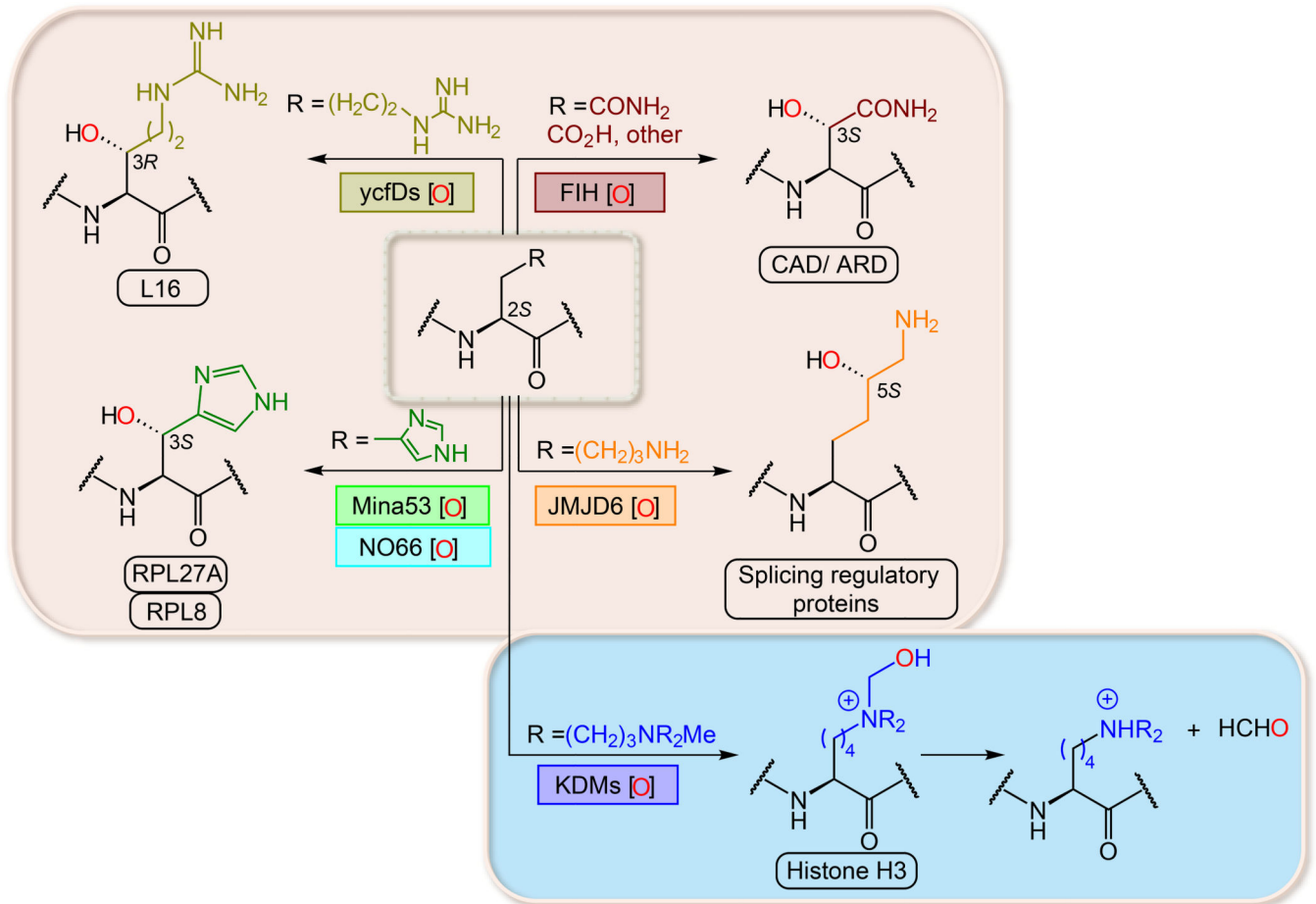
Figure 1. The overall folds of the ribosomal oxygenases

a, Reactions catalyzed by ROX and related oxygenases. CAD: *C*-terminal transactivation domain of HIF-α; ARD: Ankyrin repeat domain. **b**, Ribbons representations of ycfD, ycfD$_{RM}$, Mina53 and NO66 homodimers. The monomers contain a JmjC domain with the double-stranded-β-helix (DSBH) core present in all 2OG-oxygenases (blue) followed by dimerization (yellow) and *C*-terminal 'winged-helix' domains (red). Domain architecture and a schematic representation of the DSBH core β-strands (βI-VIII) that form major (grey, βI, VIII, III and VI) and minor sheets (blue, βII, VII, IV and V) is shown boxed. The insert between βIV and βV (purple) is involved in substrate binding. The 3 Fe-coordinating residues are on the βII and βVII strands (black circles). 2OG is in green sticks; the 2OG C5-carboxylate binding residue, Arg (ycfDs) or Lys (hROX) from βIV is a black circle.
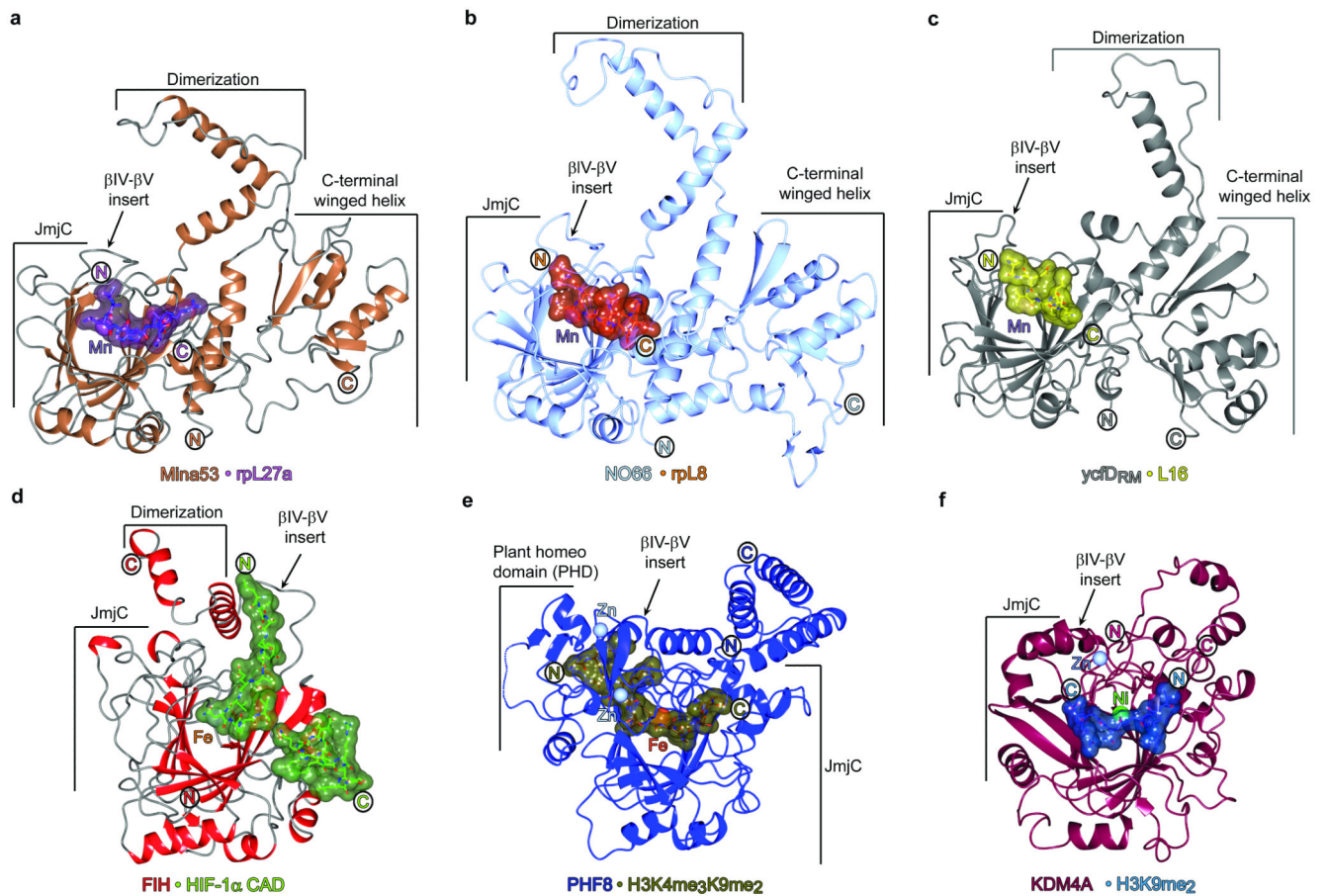
Figure 2. Comparison of the substrate structures for ROX/JmjC enzymes

Ribbons representations from ROX and related 2OG-oxygenase-substrate complexes: **a**, Mina53·Mn·2OG·rpL27a$_{(32-50)}$ ($P2_12_12_1$, 2.05 Å), **b**, NO66·Mn·NOG·rpL8$_{(205-224)}$ (C2, 2.35 Å), **c**, ycfD$_{RM}$·Mn·NOG·L16$_{(72-91)}$ ($P2_12_12_1$, 3.0 Å), **d**, FIH·Fe·NOG·HIF-1α$_{(786-826)}$ (PDB: 1H2K), **e**, PHF8·Fe·NOG·histone H3K4me3K9me2$_{(2-25)}$ (PDB: 3KV4), **f**, KDM4A·Ni·NOG·histone H3K9me2$_{(7-14)}$ (PDB: 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 color-coded spheres. Analyses of the structures reveal that the ROX overall folds (**a-c**), oligomerization states and active site architectures are evolutionarily conserved.
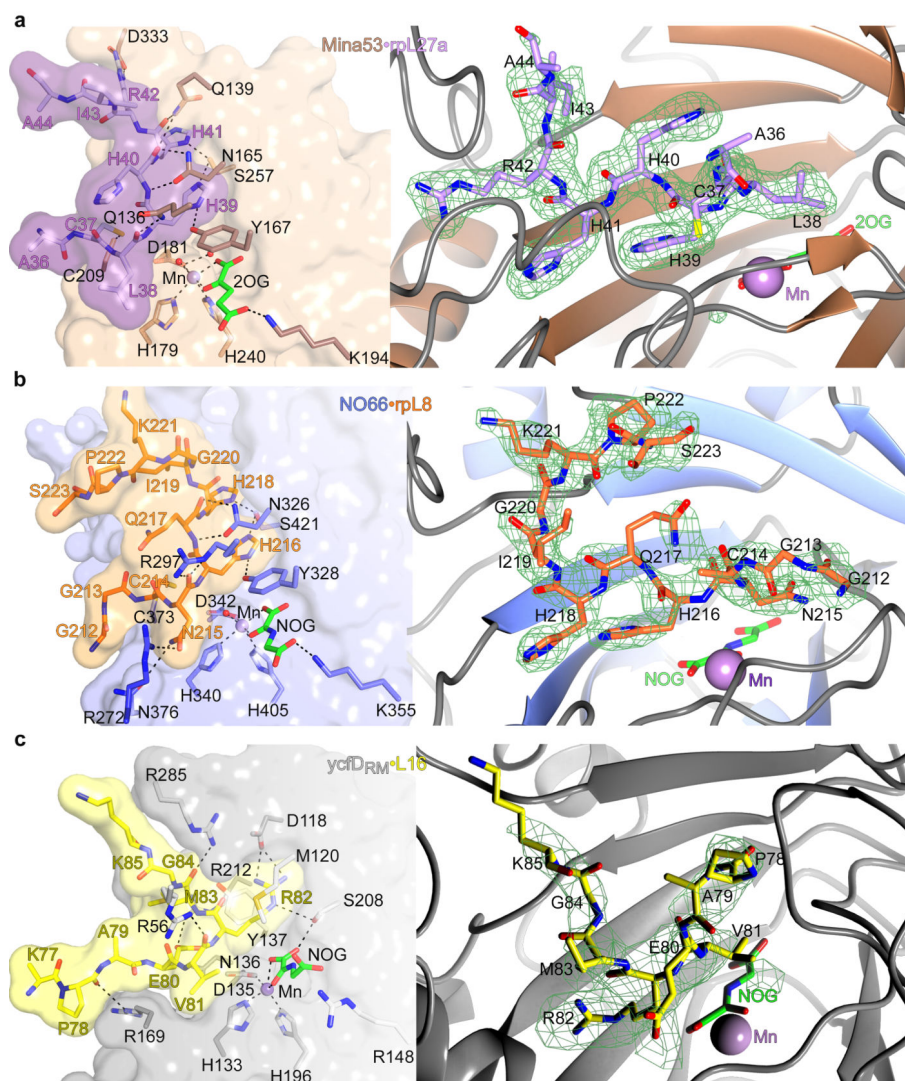
Figure 3. Features of ROX-substrate binding

Ribbons representations of Mina53 (**a**), NO66 (**b**) and ycfD$_{RM}$ (**c**) monomers showing difference electron density (*Fo-Fc* OMIT) for substrates contoured to 3σ (right panels). Left panels show active site surface representations, showing key hydrogen-bonds/polar interactions (dotted lines) with substrates. **a**, With Mina53, the His39$_{rpL27a}$ imidazole nitrogens form hydrogen-bonds with Tyr167/Ser257 (Nδ$_{His39}$-OH$_{Tyr167}$ 2.9 Å; Nε$_{His39}$-Oγ$_{Ser257}$ 3.1 Å). **b**, In NO66, His216$_{rpL8}$ is similarly bound in a deep pocket; the His216$_{rpL8}$ imidazole nitrogens form hydrogen-bonds with Tyr328/Ser421 (Nδ$_{His216}$-OH$_{Tyr328}$ 3.2 Å; Nε$_{His216}$-Oγ$_{Ser421}$ 2.7 Å) and hydrophobic interactions with Ile244, that project its pro-*S* hydrogen toward the metal (metal-β-CH$_2$, 4.4 Å). While Mina53 (**a**) uses 4 primary amides, Asn101, Gln136, Gln139 and Asn165 to interact with rpL27a backbone amides, NO66 (**b**) uses 2 arginines (272, 297) to hydrogen-bond with the Asn215$_{rpL8}$ sidechain and His216$_{rpL8}$ backbone. In the ycfD$_{RM}$·L16 complex (**c**), Arg82$_{L16}$ binds in a pocket defined by the Tyr137/Met120 sidechains, which form π-cation and hydrophobic interactions with Arg82$_{L16}$ sidechain. The Arg82 guanidino group makes electrostatic interactions with the

Asp118$_{\text{ycfDRM}}$ carboxylate (O-NH, 2.8-3.1 Å) and hydrogen-bonds to Ser208$_{\text{ycfDRM}}$ (N$\varepsilon_{\text{Arg82}}$-OH$_{\text{Ser208}}$ 3.5 Å; Nη_{Arg82}-CO$_{\text{Ser208}}$ 3.2 Å). Although Tyr167$_{\text{Mina53}}$/Tyr328$_{\text{NO66}}$ are not positionally related to Tyr137$_{\text{ycfDRM}}$, the role of the serine (Ser257$_{\text{Mina53}}$, Ser421$_{\text{NO66}}$, Ser208$_{\text{ycfDRM}}$, β-VIII) in binding the hydroxylated His/Arg is conserved in ROX. Substitutions of these residues cause significant loss of activity (see Extended Data Fig. 6).
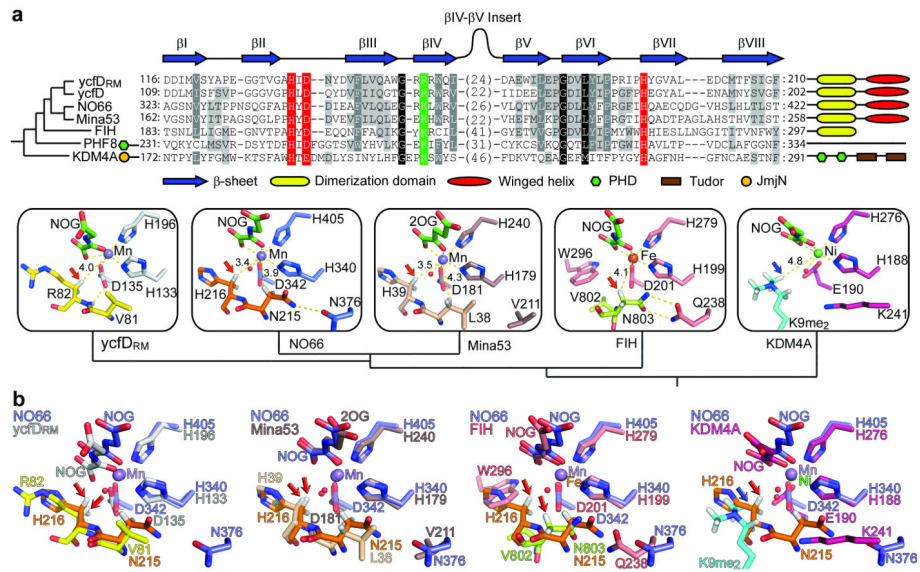
Figure 4. Proposed sequence of evolution of active metal chemistry of ROX and related JmjC 2OG-oxygenases

The figure compares views from active sites of representative JmjC-enzymes and suggests how the ROX fold evolved into JmjC-hydroxylases and -KDMs. Structurally informed cross-genomic bioinformatic analyses imply that the ROX are the earliest identified JmjC 2OG-oxygenases[30]; ycfD and NO66 both exist in prokaryotes but only NO66 is identified in eukaryotes. Coupled to the analyses of the active sites, these analyses imply NO66/close-relatives are the precursors of Mina53 and other JmjC-hydroxyalses and KDMs. **a**, Upper panel: structure based alignment of ROX, FIH, PHF8 and KDM4A with DSBH core labeled βI-VIII, iron-coordinating and the 2OG C5-carboxylate binding residues in red and green. Lower panels: analyses of active sites suggest conservation of metal-/2OG-binding in ROX, FIH and KDMs, note the 2OG C5-carboxylate binding residue (usually from βIV in JmjC-enzymes), changes from an Arg (in ycfDs) to a Lys (in hROX, JmjC-hydroxyalses/KDMs) (Extended Data Fig. 4). **b**, Overlays of the NO66/ycfD$_{RM}$, NO66/Mina53, NO66/FIH, NO66/KDM4A active site views. The hydroxylated β-methylenes nearly superimpose in ROX, such that the oxidized C-H bonds (red arrows, 3-pro-*R* in Arg82$_{L16}$ and 3-pro-*S* in His39$_{rpL27a}$ and His216$_{rpL8}$) project toward the metal. The spatial relationship of the hydroxylated C3/Nε-methyl carbon with respect to the metal (and associated reactive oxidizing species) is conserved in ROX and the demethylases, e.g. KDM4A, but not in FIH. Note the different hydroxylation positions, but the similar orientation of Asn803$_{CAD}$/FIH (hydroxylated) and Asn215$_{rpL8}$/NO66 (not hydroxylated).