RSC Chemical Biology

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *RSC Chem. Biol.*, 2025, **6**, 412

Bioinformatic and biochemical analysis uncovers novel activity in the 2-ER subfamily of OYEs†

Tamra C. Blue-Lahom, Stacey K. Jones and Katherine M. Davis *

Members of the old yellow enzyme (OYE) family utilize a flavin mononucleotide cofactor to catalyze the asymmetric reduction of activated alkenes. The 2-enoate reductase (2-ER) subfamily are of particular industrial relevance as they can reduce α/β alkenes near electron-withdrawing groups. While the broader OYE family is being extensively explored for biocatalytic applications, oxygen sensitivity and poor expression yields associated with the presence of an Fe/S cluster in 2-ERs have hampered their characterization. Herein, we explore the use of pseudo-anaerobic preparation as a route to more widespread study of these enzymes and apply bioinformatics approaches to identify a subset of 2-ERs containing unusual mutations in both a key catalytic residue and the Fe/S cluster-binding motif. Biochemical analysis of a representative member from *Burkholderia insecticola* (OYE*Bi*) reveals novel *N*-methyl-proline demethylation activity, which we hypothesize may play a role in osmotic stress regulation based on genomic neighborhood analysis.

Received 28th November 2024,
Accepted 18th January 2025

DOI: 10.1039/d4cb00289j

rsc.li/rsc-chembio

Introduction

Non-metalloenzyme members of the old yellow enzyme (OYE) family are single-domain flavin mononucleotide (FMN)-dependent oxidoreductases that utilize nicotinamide adenine dinucleotide (NAD(P)H) to drive the asymmetric reduction of activated alkenes, affording chiral products. Substrate conversion proceeds *via* a bi–bi–ping–pong mechanism wherein nicotinamide initially reduces the flavin cofactor, which subsequently reduces the primary substrate *via* hydride transfer to the β-carbon of the activated alkene. The resultant substrate anion is then protonated at the α-carbon by solvent or a proximal tyrosine (Tyr196 in OYE1) to yield a *trans*-hydrogenated product (Fig. 1).[1–3] As most hydrogenation reactions favor *syn*-addition, there is significant interest in the biocatalytic application of OYEs to sustainably produce industrially relevant compounds, such as chiral intermediates, polymer precursors, and diverse pharmaceutical products.[2]

Current synthetic routes to generate nylon, for example, often require harsh and environmentally hazardous conditions.[4,5] The synthesis of adipic acid, a nylon 6,6 monomer, releases the potent greenhouse gas nitrous oxide as a byproduct, making it a significant contributor (1–2%) to U.S. manufacturing-linked emissions.[5] Biocatalytic production of adipic acid is therefore highly desirable and has been demonstrated *via* the enzymatic reduction of *cis*,*cis*-muconic acid.[6,7] While non-metalloenzyme and engineered OYEs are unable to efficiently reduce α/β-unsaturated monocarboxylic acids and monoesters, limiting their utility in this area, 2-enoate reductase (2-ER) members of the family have been shown to catalyze the reduction of α/β alkenes adjacent to electron-withdrawing groups (EWGs), such as carboxylic acids.[2] Other reported non-canonical oxidative reactions among 2-ERs include the deamination of histamine[8] and demethylation of trimethylamine.[1,9] Despite this diverse and industrially significant reactivity, the subfamily remains underexplored.

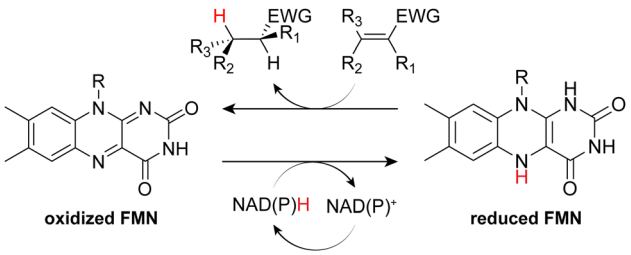

Fig. 1 Canonical OYE reaction.

2-ERs commonly possess a flavin adenine dinucleotide (FAD)-binding domain, in addition to the core $(\alpha,\beta)_8$ triosephosphate isomerase (TIM) barrel fold, which houses the catalytic FMN cofactor in OYEs.[10,11] Among characterized enzymes, a four-cysteine motif ($CX_2CX_{2-3}CX_{11}C$), located at the interface of these two domains, coordinates a [4Fe–4S] cluster believed to facilitate electron transfer between the two flavin cofactors (typically separated by a distance of ~15 Å). Sensitivity to

Department of Chemistry, Emory University, Atlanta, GA 30322, USA.
E-mail: katherine.davis@emory.edu

† Electronic supplementary information (ESI) available: SSN of the OYE family, domain and genomic neighborhood analysis of the 2-ER subfamily, species and sequence length distribution of 2-ERs, activity assay data, structural comparisons, EPR and oxygen sensitivity data, and plasmid maps. See DOI: https://doi.org/10.1039/d4cb00289j

molecular oxygen imparted by this metallocofactor complicates enzyme expression, purification, and handling. Fourth-wave biocatalytic approaches, however, such as bioinformatics, artificial intelligence and other computational tools, can help to more systematically approach identification and characterization of these enzymes.[12] The former has been employed to explore reactivity within the OYE family at large. A series of sequence similarity networks (SSNs) have been particularly useful for refining OYE classification and enabling the systematic investigation of sequence-function space for novel activity.[10,11,13] From these analyses, approximately one-quarter of the family are predicted to be 2-ERs. Unfortunately, previous studies intentionally excluded or limitedly explored 2-ER reactivity due to challenges associated with wet-lab manipulation.

Given the compelling biocatalytic and industrial relevance of OYE 2-ERs, we attempt to bridge this knowledge gap by directly investigating the functional diversity within the subfamily. Leveraging similar bioinformatic analyses, we identify a series of sequence features that may signal novel 2-ER catalytic function and select a subset of these for further study. Biochemical characterization of a representative enzyme with an altered Fe/S cluster-binding motif reveals unexpected dual reactivity, potentially suggesting a broader functional role in supporting cellular resilience for this and related enzymes. We additionally demonstrate the utility of a pseudo-anaerobic approach to yield active enzyme. These findings pave the way for expanded exploration of 2-ERs, both as versatile biocatalysts in sustainable industrial processes and as potential participants in cellular stress-response pathways.

Results and discussion

Bioinformatics reveals a non-canonical Fe/S-cluster binding motif

The number of sequences annotated as OYE-like enzymes in UniProt (PF00724) has grown dramatically in recent years, increasing by more than 60% since 2018. We, therefore, began our investigation by constructing an updated SSN of the family containing 115,314 homologous OYE sequences at an edge threshold of 10^{-87}. This approach grouped sequences with greater than 50% sequence identity (ID) into single nodes and recapitulated clustering behavior similar to that of earlier networks (Fig. S1, ESI†).[10] Although OYEs are widespread in both eukaryotes and prokaryotes, our analysis indicates that 2-ERs are primarily (~95%) bacterial in origin (Fig. S2A, ESI†). Notably, they have an average sequence length (~700 residues) approximately double that of more well-studied OYE subfamilies (Fig. S2B, ESI†). This observation is consistent with the bi-domain architecture observed in characterized enzymes, and is further supported by domain analysis, which indicates that 96% of 2-ER sequences contain a FAD-binding domain—the most common being Pfam PF07992 (Table S1, ESI†).[8,14,15]

To further investigate 2-ER diversity, associated sequence data were extracted to generate a higher-resolution 2-ER SSN (80% ID and *E*-value cutoff of 10^{-185}, see Methods). This analysis yielded twenty-seven clusters that each contain at least 100 sequences (Fig. 2(A)). Smaller clusters were excluded from further analysis due to limited statistical significance. Structurally characterized members of the 2-ER subfamily include 2,4-dienoyl CoA reductase (DCR) from cluster 1 of the higher-resolution SSN;[15] 6-hydroxypseudooxynicotine amine oxidase (HisD),[16] trimethylamine dehydrogenase (TMADH),[17] and histamine dehydrogenase (HADH)[8] from cluster 3; and 2-naphthoyl-CoA reductase (NCR) from an unnumbered sparse cluster.[18]

Deeper exploration reveals heterogeneity of key catalytic residues and cofactor binding motifs.[20] As is true for classical OYEs, the active site tyrosine, believed to serve as a proton donor for the enolate intermediate (Tyr196 in OYE1),[3] is strictly conserved within clusters 1 and 4. Clusters 2 and 3, by contrast, deviate from this norm. An active site histidine replaces tyrosine in cluster 3 and likely fulfills a similar proton-transfer role, while cluster 2 displays significant variability in this position (26% Ser, 23% Thr, 23% Gly, 15% Tyr and 13% His). The related FAD-dependent Type II UDP-*N*-acetylenolpyruvylglucosamine reductases (MurB) utilize serine or cysteine as analogous proton donors, while mutagenesis studies of chromate reductase (CrS) and other traditional OYEs suggest solvent may also be a suitable alternative.[21–23] Whether alkene reduction activity is retained among 2-ERs that lack the canonical catalytic Tyr, however, remains to be seen.

Heterogeneity is additionally observed within the four-cysteine Fe/S cluster-binding motif that may have implications for the catalytic role of the cluster.[3,14] These residues have been strictly conserved among OYE 2-ERs characterized to date, however, five clusters (8, 9, 17, 21, and 26) lack this motif altogether. Intriguingly, examination of their domain architectures indicates that they appear to have maintained an FAD-domain (Fig. S3, ESI†). Although these clusters comprise only 4% of the 2-ER subfamily, their sizes suggest that constituent enzymes are likely catalytically active. Further investigation is needed to evaluate their potential for catalyzing novel chemistries, including the possibility of oxygen-insensitive reactions that may be attractive for industrial applications.

Sequences in cluster 2, the second largest cluster in the 2-ER family, by contrast, overwhelmingly (~92%) possess a modified [4Fe–4S] cluster-binding motif depicting an alanine mutation at the second position (corresponding to Cys337 in DCR; Fig. 2(B)). Various other nonpolar amino acids occupy that position among the remaining 8% of sequences within the cluster. As there are no additional cysteines nearby that could indicate a shift in the motif, it seems likely that, if a [4Fe–4S] cluster is bound by these enzymes, it is either unusually coordinated to a non-thiolate residue at another position or retains an open coordination site. Unusual Fe/S-cluster coordination motifs are increasingly recognized as key features in enzyme reactivity;[24,25] however, no cluster 2 homologs have been biochemically or structurally characterized to date. Thus, we selected a representative for further investigation.

Expression of a representative enzyme from cluster 2

More specifically, we examined cluster 2 for sequences containing both the Cys-to-Ala mutation, as well as mutation of the

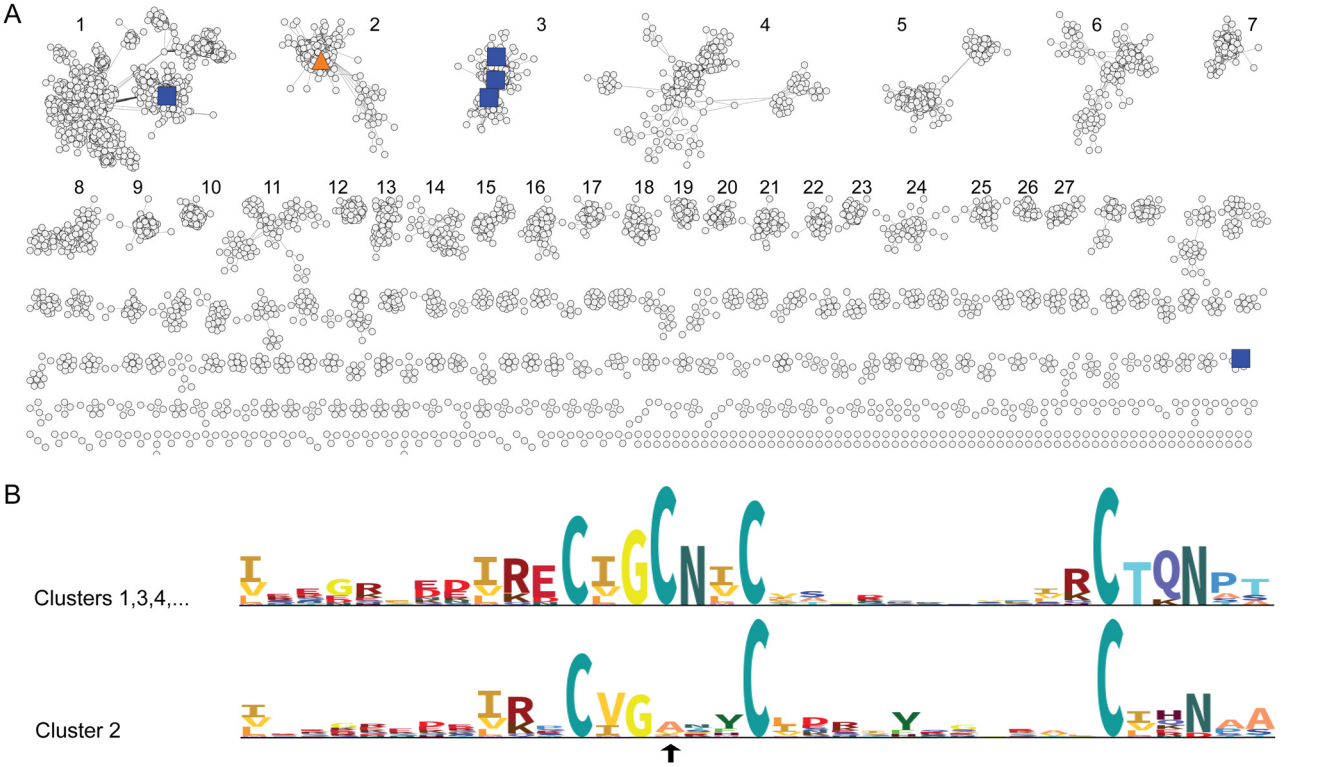

Fig. 2 Sequence analysis of the 2-ER subfamily of OYEs (A) SSN of 2-ERs. Structurally characterized sequences are highlighted with a blue square, and OYE*Bi* is marked with an orange triangle. (B) Sequence logo comparing the Fe/S-binding domains within differing 2-ER clusters. Note that the total height of each stack indicates the overall sequence conservation of that position, while the height of each letter corresponds to the relative frequency with which that amino acid occupies the given position.[19]

catalytic Tyr. To provide a higher probability of soluble expression, we focused on sequences at the lower end of the size distribution. OYE*Bi* (Transcript ID: BAO89787), encoded by the BRPE67_CCDS01850 gene in *Burkholderia insecticola*, is one such enzyme, which features a Tyr-to-His mutation (His175) analogous to that of sequences from cluster 3, along with the modified Fe/S-cluster binding motif containing alanine (Ala348) in place of a canonical Cys. Our interest in OYE*Bi* was further heightened by its production within a genus that includes strains of clinical and agricultural importance due to their pathogenicity.

Before attempting to express the enzyme, we first investigated structural implications of the aforementioned mutations *in silico*. The resultant high-confidence AlphaFold model suggests OYE*Bi* adopts a fold very similar to the canonical 2-ER, DCR (Fig. S4A, ESI†).[15,26] Overlay of the structures yields a root-mean-square deviation of the peptide backbone of 1.79 Å over 425 Cα atoms. Examination of the active site, likewise, reveals no largescale structural rearrangements. His175 overlays with the catalytic Tyr (Tyr166 in DCR), though its backbone is slightly shifted away (1.7 Å) from the substrate binding pocket (Fig. S4B, ESI†). While no other Tyr residue is close enough to serve as a proton donor, two other residues with proton-donating potential (His171 and Ser63) are positioned within ~3.5 Å. In the absence of a high-resolution substrate-bound structure, it is difficult to determine which of these three residues, if any, is catalytically relevant. However, their relative proximity may impart some functional redundancy, as has been observed in DCR.[15]

As expected, the unusual cysteine motif of OYE*Bi* lies at the domain interface and aligns well with the Fe/S cluster-binding motif of DCR (Fig. S4C, ESI†). Although amino acids besides cysteine have been reported to coordinate Fe/S clusters in other systems,[24] our AlphaFold model does not support such a binding mode in OYE*Bi*. Instead, Ala348 directly overlays with the second coordinating cysteine (Cys337) of DCR. This arrangement positions the open coordination site of the cluster closest to the FAD cofactor (Fig. S4C, ESI†), perhaps enhancing the directional transfer of electrons, as reduced electron density at this site may allow the unique Fe to act as an electron sink. Given no apparent structural impediments to cluster binding, we developed an expression and purification protocol under the assumption that OYE*Bi* can coordinate an Fe/S cluster.

2-ERs have historically been expressed and purified anaerobically in order to mitigate challenges associated with oxygen sensitivity of their [4Fe–4S] clusters.[14,15] However, these methods are difficult to implement on an industrial scale and do not address poor incorporation of the metallocofactor during standard heterologous expression. To further interrogate the impact of the alanine mutation on cluster binding and explore alternative approaches to generating 2-ERs more broadly, we surveyed three expression protocols: fully anaerobic, pseudo-anaerobic, and aerobic. Efficient incorporation of the oxygen-sensitive metallocofactor

was additionally facilitated *via* co-expression with cellular machinery associated with Fe/S-cluster assembly (see Methods).[27] This approach was inspired by heterologous expression methods routinely employed for the *in vitro* production of radical *S*-adenosyl-L-methionine (SAM) enzymes, which also incorporate a canonical [4Fe–4S] cluster.[20] Note that expression of Fe/S cluster assembly proteins was not induced during fully aerobic expression.

OYE*Bi* contains a [4Fe–4S] cluster

To verify the incorporation of an Fe/S cluster in OYE*Bi*, we first employed inductively coupled plasma mass spectrometry (ICP-MS) to estimate the number of Fe ions per enzyme. Results for both fully anaerobic expression conditions and pseudo-anaerobic expression followed by reconstitution are most consistent with incorporation of a [4Fe–4S] cluster (Table 1), in agreement with other FAD-containing 2-ER OYEs. In the absence of reconstitution, the number of Fe ions per molecule of OYE*Bi* expressed under pseudo-anaerobic conditions decreases significantly, perhaps indicative of oxidative damage.[28] Although fully aerobic expression unsurprisingly resulted in no detectable Fe bound to the enzyme, OYE*Bi* prepared this way was nonetheless stable in solution, suggesting that the Fe/S cluster does not play a crucial structural role.

To further validate our hypothesis that OYE*Bi* incorporates an intact cubane cluster, we additionally performed electron paramagnetic resonance (EPR) spectroscopy at 10 K. As isolated, 2-ERs contain oxidized flavins and a $[4Fe–4S]^{2+}$ cluster. OYE*Bi* is, therefore, EPR silent in the absence of reductant (Fig. 3(A)). Addition of 10 molar equivalents sodium dithionite, by contrast, produces a spectrum in which the isotropic feature (g = 2.005) associated with flavin semiquinones overlaps with the rhombic signal characteristic of $[4Fe–4S]^{+}$ (g-values of 1.9, 1.94, and 2.08).[29,30] Consistent with the relaxation of rate of paramagnetic iron species,[31] the rhombic signal is diminished at increasing temperatures (Fig. S5A, ESI†). Altogether, these analyses support the coordination of a [4Fe–4S] cluster in OYE*Bi*.

OYE*Bi* is oxygen sensitive and displays a preference for NADH

Before exploring the catalytic range of OYE*Bi*, assay conditions for optimal activity were determined with *R*-carvone as an exploratory substrate under fully anaerobic conditions. *R*-Carvone is efficiently converted to dihydrocarvone by most OYEs, making it a useful starting point for assessing OYE activity.[1,2,10] We began our investigation by determining the optimum buffer pH for reactivity. Toward that end, OYE*Bi* was reacted with substrate in a series of different conditions ranging from pH 5.8 to 10 (100 mM phosphate buffer for pH $\leq$ 8, and 50 mM CAPS for pH $\geq$ 8.5) with a 20 μM 1 : 1 NAD(P)H cocktail. Percent conversion was then assessed *via* gas-chromatography mass spectrometry (GC–MS) quantitation

Table 1 ICP-MS results displaying mol of Fe/mol of enzyme for each expression condition

	Anaerobic	Pseudo-anaerobic	Aerobic
(+) Reconstitution	—	4.82 ± 0.03	0.74 ± 0.00
(−) Reconstitution	4.18 ± 0.09	3.02 ± 0.20	0.41 ± 0.00

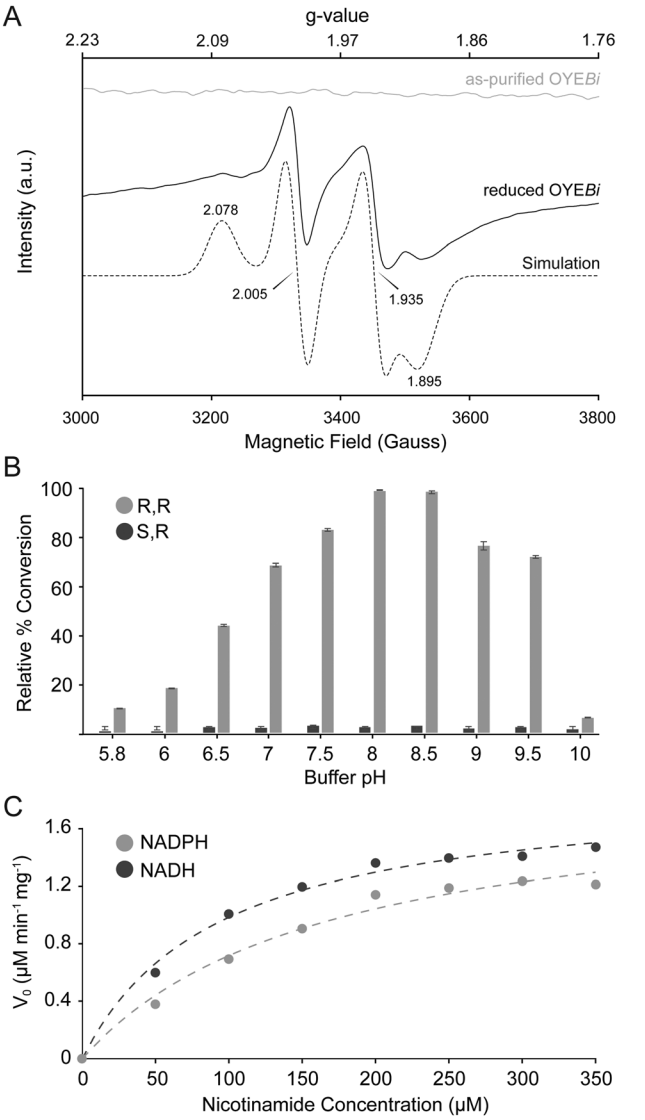

Fig. 3 Characterization of OYE*Bi* (A) EPR spectra. (B) pH-dependent conversion of *R*-carvone to *R,R*- and *S,R*-dihydrocarvone. (C) Steady-state kinetics of nicotinamide-dependence.

(see Methods). The most efficient activity was observed at pH 8 (Fig. 3(B)). The enzyme's preference for nicotinamide adenine dinucleotide (NADH) *versus* nicotinamide adenine dinucleotide phosphate (NADPH) was subsequently determined at this optimal pH. Spectrophotometric assays monitoring the consumption of NAD(P)H *via* its characteristic absorption at 340 nm yielded initial reduction rates of 0.37 μmol min^{-1} and 0.60 μmol min^{-1} for NADPH and NADH, respectively, demonstrating a preference for NADH as the proton source. Assuming Michaelis–Menten kinetics, K_m for NADPH is 167.25 ± 0.18 μM, while K_m of NADH is 92.84 μM ± 0.08 (Fig. 3(C)). No product formation was observed in the absence of enzyme or nicotinamide. Finally, we investigated the effect of oxygen on OYE*Bi* activity by exposing protein extract to air. Samples were taken at time intervals of 15 min for one hour at 4 °C, and percent conversion again assessed by GC–MS

(Fig. S5B, ESI†). Activity was reduced by ~25% following 1 h of exposure to air and was not restored by the reintroduction of anaerobic conditions. All subsequent experiments were performed in an anaerobic chamber with NADH at pH 8.0.

OYE*Bi* catalyzes unusual demethylations alongside canonical OYE and 2-ER activities

With optimal assay conditions in hand, we explored the substrate scope of OYE*Bi* using a diverse compound library of aliphatic and cyclic alkenes bearing ketones, aldehydes, monocarboxylic acid EWGs, and methyl-proline analogs. Each substrate was selected to interrogate a particular facet of OYE activity reported in the literature (Fig. 4(A)). Cinnamic acid (**1**) probed the ability of OYE*Bi* to catalyze 2-ER activity (*i.e.*, the conversion of alkenes adjacent to carboxylic acid EWGs). *R*-Carvone (**3**) and *S*-carvone (**4**) assessed whether the enzyme retains traditional OYE activity, while 2-methyl-2-cyclopentenone (**5**), 2-cyclopentenone (**6**), and the Wieland–Miescher ketone (**7**) determined how this traditional activity varies with size and steric hindrance. The remaining substrates were included to test for less common or novel OYE chemistries: 3-phenyl propionic acid (**2**) and levodione (**8**) for desaturase activity, along with *N*-methyl-L-proline (NMP; **9**) and stachydrine (**10**) for demethylase activity.

OYE*Bi* expressed under fully anaerobic conditions catalyzed the complete conversion of **1** to **2** with no detectable substrate remaining at the end of the reaction, confirming retention of 2-ER activity despite the altered Fe/S-cluster binding motif. Complete conversion is consistent with the activity of other 2-ERs tested *in vitro*, such as FldZ from *Clostridium sporogenes* DSM 795.[14] Pseudo-anaerobic expression conditions, by contrast, yielded 61.88 ± 2.12% conversion of **1** to **2**, and reconstitution could not fully restore activity (Fig. 4(B) and Table S2, ESI†). Surprisingly, OYE*Bi* expressed and purified under aerobic conditions was still capable of reducing **1** to **2**, albeit in substantially lower yields (6.52 ± 4.17% conversion). To our knowledge, OYE*Bi* is only the second 2-ER to display any such activity upon aerobic expression.[32]

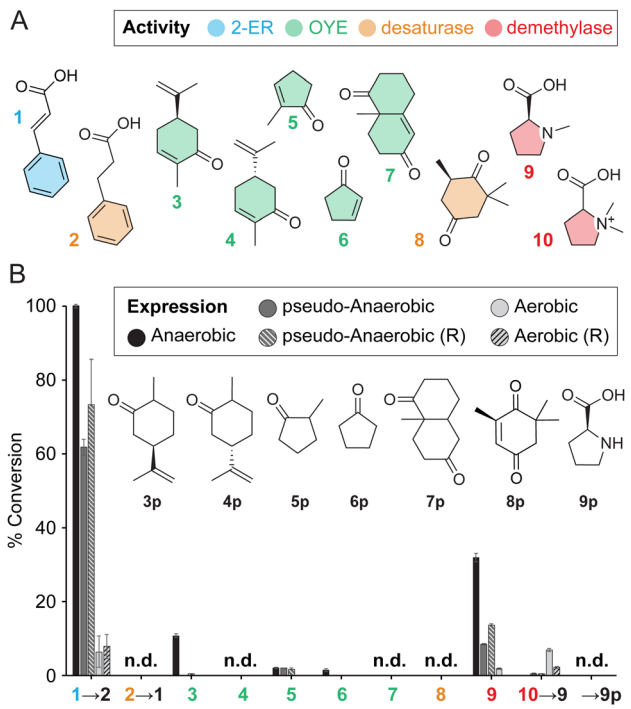

Fig. 4 Activity of OYE*Bi*. (A) Substrate panel. Compounds are colored by tested activity. (B) Percent conversion detected *via* MS (see Table S2, ESI† for exact values).

As expected from our optimization assays with **3**, OYE*Bi* is also able to catalyze traditional OYE reactivity. Although the presence of an Fe/S cluster is not required for such reactions by traditional OYEs, maximal reduction to dihydrocarvone (**3p**) was again consistently observed under fully anaerobic expression conditions (Fig. 4(B) and Table S2, ESI†). In contrast to the results with **1** described above, conversion of **3** by pseudo-anaerobic and aerobic preparations of OYE*Bi* were comparable and did not significantly improve upon reconstitution. We hypothesize that traditional OYE activity may be hindered by imperfect cluster incorporation under these conditions. However, further exploration is required to understand such behavior.

A comparison between **3** and **4** further reveals that OYE*Bi* exhibits enantiopreference for substrates in the *R*-conformation (Fig. 4(B) and Table S2, ESI†). While the enzyme did not appear to tolerate sterically hindered substrates—no conversion of **7** to **7p** was observed, it did show limited reactivity with the representative small substrates (**5** and **6**; Fig. 4(B)). Approximately 2% of the reduced product 2-methyl-2-cyclopentanone (**5p**) was detected when OYE*Bi* was expressed under anaerobic or pseudo-anaerobic conditions in which the [4Fe–4S] cluster is assumed to be predominantly intact (Table S2, ESI†). Cyclopentanone (**6p**), by contrast, was only observed following fully anaerobic expression. OYE*Bi* did not display desaturase activity with **2** or **8** to generate **1** and ketoisophorone (**8p**), respectively. As desaturase activity may require a cofactor capable of accepting electrons, we additionally tested these substrates with NAD^+. However, no conversion to NADH was observed, confirming the absence of desaturase activity irrespective of the nicotinamide redox state (Fig. S6, ESI†).

In addition to the reductive chemistry associated with traditional OYE and 2-ER reactivities, OYE*Bi* is also capable of performing oxidative demethylation. We observe up to 22.16 ± 2.24% conversion of **9** to L-proline (**9p**). Both singly and doubly demethylated products of **10** were also detected. Intriguingly, aerobic expression yielded the greatest conversion of **10** to **9**, perhaps suggesting that this reaction follows an oxygen-dependent mechanism. The second demethylation step to **9p**, by contrast, occurred at very low efficiencies difficult to detect across all expression conditions. While not the only 2-ER reported to perform oxidative demethylation chemistry,[8,33] to our knowledge, OYE*Bi* is the first confirmed to accept **9** and **10** as substrates.

OYE*Bi* likely utilizes both oxidative and reductive pathways

Based on the results of our substrate panel, we propose that OYE*Bi* operates *via* two distinct mechanistic pathways: a canonical reductive route associated with prototypical 2-ER activity and an oxidative pathway capable of demethylation. While the

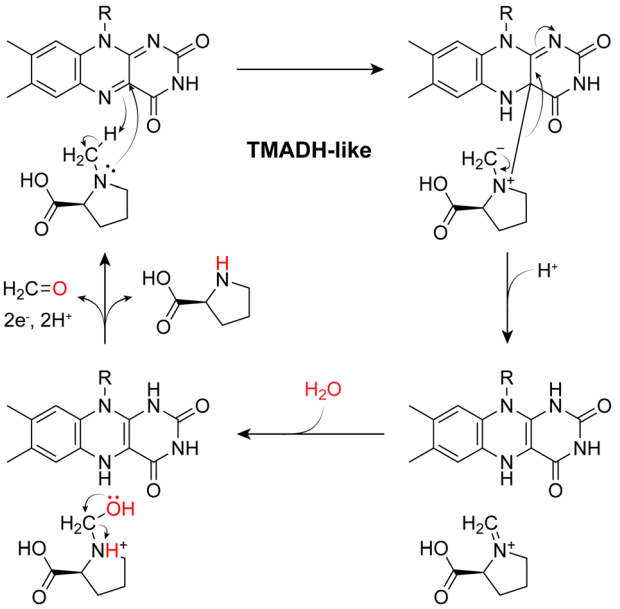

Fig. 5 Possible TMADH-like[39] mechanism for OYE*Bi*-catalyzed oxidative demethylation.

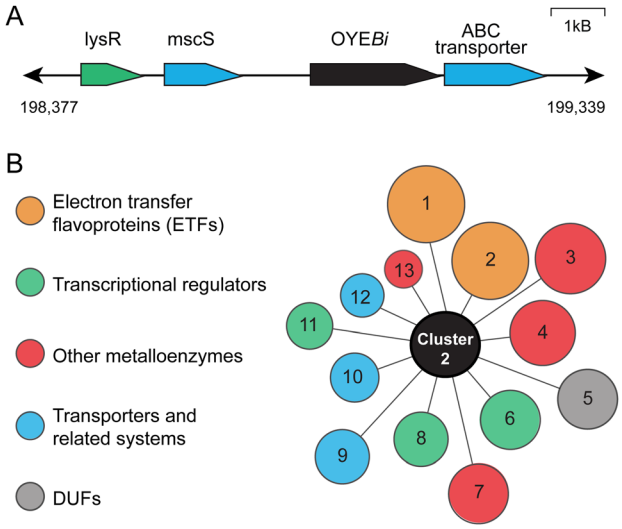

Fig. 6 Genomic context. (A) Genomic neighborhood of OYE*Bi*. (B) A 2-ER SSN cluster 2 hub-node GNN. Note that the diameter of the nodes corresponds to the CoF with cluster 2 enzymes.

reductive mechanism follows traditional OYE chemistry (Fig. 1), diminished activity upon oxygen exposure suggests that the [4Fe–4S] cluster plays a crucial, if indirect, role in this pathway. By contrast, we propose that the oxidative pathway relies on an oxidized FMN cofactor, in a manner analogous to the mechanism of trimethylamine dehydrogenase (TMADH; Fig. 5).[17,34,35] A distant homolog of OYE*Bi* (24.9% sequence identity), TMADH from *Methylotrophus methylophilus* is one of only two 2-ERs reported to perform oxidative demethylation chemistry.[8,33] More specifically, TMADH catalyzes the conversion of trimethylamine (TMA) to dimethylamine and formaldehyde.[36] Oxidation of TMA by FMN yields an imine intermediate that spontaneously hydrolyzes to form the demethylated product. It is difficult, however, to envision this mechanism accounting for the enhanced demethylation of **10** by aerobically purified enzyme. We speculate whether OYE*Bi* could instead utilize an alternate oxygen-dependent mechanistic strategy. There is evidence to suggest that molecular oxygen is capable of reoxidizing FMN, either directly or indirectly, for oxidative activity by members of the OYE family.[33,37] Such dependence on oxygen could help to explain the increased demethylation activity by aerobically prepared OYE*Bi*. As with the desaturation reaction, NAD^+ does not appear capable of serving this role, in further support of this hypothesis (Fig. S6, ESI†). Although Fe/S clusters are notoriously sensitive to molecular oxygen, the Rieske-type non-heme iron oxygenase stachydrine demethylase (Stc2) protects its [2Fe–2S] cluster *via* electron transfer from FMN, which also regenerates the flavin cofactor.[38] However, additional studies are needed to support any such mechanisms in OYE*Bi*.

Genomic neighborhood analysis suggests OYE*Bi* facilitates osmotic stress response

The gene encoding OYE*Bi* is located on the forward strand of chromosome 3 in *B. insecticola*. Upstream are genes for a LysR-type transcriptional regulator and a small-conductance mechanosensitive ion channel (MscS; UniProt accession code: A0A158J5F7), while downstream is a dipeptide ATP-binding cassette (ABC) transporter (Fig. 6(A)). Mechanosensitive ion channels are transmembrane proteins that sense mechanical changes in the lipid bilayer to help regulate the export of ions and osmoprotectants.[40] The movement of associated water-soluble substrates across the membrane is then achieved with ABC transporters, which hydrolyze ATP to power this process.[41–43] The colocalization of these genes is reminiscent of the operon associated with the well-documented stachydrine-dependent osmotic stress response in *Sinorhizobium meliloti*, a Gram-negative bacteria root symbiont of the alfalfa plant.[44]

Abundant in soil, the nonionic osmoprotectant stachydrine is taken up by ABC transporters during times of high external and low internal osmolarity to help *S. meliloti* retain water.[45] When stachydrine influx is reduced, a catabolic pathway is activated, leading to its demethylation to *N*-methlyproline and proline, which serve as a carbon and nitrogen sources for the cell.[46,47] This process is controlled by transcriptional factors similar to those regulating OYE*Bi*, including LysR,[44] while demethylation activity is facilitated by an analogous flavin-dependent multienzyme system comprised of Stc2 (the Fe/S-dependent enzyme described above), Stc3, Stc4, and StcD.[46] It, therefore, seems reasonable to suggest that OYE*Bi* plays a role in the osmotic stress response of *B. insecticola*.

Given the enzyme's somewhat unique activity profile, we set out to investigate whether this could be the native function of cluster 2 enzymes more broadly *via* genomic neighborhood network (GNN) analysis.[48] GNN analysis leverages the evolutionary tendency of bacteria to cluster functionally related genes into operons to predict enzyme function.[49,50] We, therefore, employed the Enzyme Function Initiative Genome Neighborhood Tool to examine the operons associated with all sequences within cluster 2 of the 2-ER SSN (Fig. 2(A)). A neighborhood of ±10 open reading

frames and a minimum 20% co-occurrence frequency (CoF) between cluster 2 sequences and neighboring genes were specified to generate the GNN.[51] This analysis identified 13 protein families that are commonly colocalized with OYE*Bi* and its homologs, with a CoF of at least 20% (Fig. 6(B) and Table S3, ESI†). A significant majority of these families are associated with cellular stress responses, including LysR (36% CoF) and other transcriptional regulators like TetR—also present in *S. meliloti*, ABC transporters or transporter components (21–34% CoF), and enzymes commonly associated with oxidative stress. While not conclusive, our GNN analysis suggests that 2-ERs from cluster 2 may be involved in responding to diverse cellular stressors.

Conclusions

The emergence of the fourth wave of biocatalysis, driven by advances in sequence analysis and bioinformatics tools, has enabled the discovery and functional characterization of previously unstudied enzyme subfamilies.[12,52] Herein, we demonstrate application of these approaches for the analysis of the industrially relevant 2-ER subfamily of OYEs, identifying significant heterogeneity in key catalytic residues and cofactor-binding motifs. Among a subset of these enzymes, we discovered a mutation in the highly conserved Fe/S cluster-binding motif and demonstrated that this change in motif does not impede cluster incorporation nor the proposed native functionality of a representative sequence from *B. inseticola* (OYE*Bi*). Inspired by studies of radical SAM enzymes, we demonstrate that pseudo-anaerobic expression achieves robust [4Fe–4S] cluster loading while reducing the complexity of experimental setup, providing a practical avenue for biocatalysis research where cell lysates are preferred for chemical transformations. We additionally identify a series of homologs lacking the Fe/S cluster-binding motif altogether, perhaps indicative of O_2-insensitive 2-ER activity that could further simplify biocatalytic applications. Although unreactive with *trans*-cinnamic acid and not currently classified in the same Pfam as the OYE family, the NADH-dependent 2-ene reductase HcrF from *Limosilactobacillus fermentum* was recently demonstrated to reduce hydroxycinnamic acids, highlighting the potential for metal-independent activity.[53] Biochemical assays with a tailored substrate library suggest that OYE*Bi* can catalyze both reductive chemistry characteristic of the OYE family, as well as oxidative demethylation reactions. To our knowledge, OYE*Bi* is the first to retain reduction activity with traditional OYE (>10% conversion), and 2-ER (>40% conversion) substrates, while also being able to perform an oxidative pathway (>30% conversion). The ability of OYE*Bi* to demethylate known osmoprotectants together with genomic analysis supports the hypothesis that these enzymes may play a role in cellular stress response systems.

Experimental procedures

Materials and methods

Unless otherwise noted, all substrates and cells were purchased from Sigma Aldrich. The pET-28a vector bearing the OYE*Bi* gene was obtained from Genscript, while the pDB1282 plasmid (*vide infra*) was kindly donated by Prof. Squire Booker. For anaerobic experiments, all solutions were degassed in an ultrasonic waterbath for 15 min before being introduced to the anaerobic chamber where they rested for a minimum of 4–5 h prior to use, unless otherwise specified.

SSN generation

The dataset of all sequences (115,314) associated with accession code PF00724 (corresponding to the OYE family) was obtained from the UniProt Protein Database in April 2023.[10,54] It was subsequently reduced by clustering to 50% identity using the Cluster Database at High Identity with Tolerance (CD-HIT) software.[55] At this point, unusually short (<100 residues) or long sequences (>1500 residues) were also removed. To ensure previously known and functionally characterized OYEs, as listed in Swiss-Prot, were represented, the following additional sequences were reintroduced after clustering: A0A1D8PPK, B5XRB0, C5H429, P41816, P42593, P54550, P77258, Q02899, Q03558, Q6Z965, Q8GYB8, Q8LAH7, Q9FEW9, Q9FUP0, Q9XG54, A0A0N9HP11, E9AGH7, P0DW92, P16099, P19410, P32370, P43084, Q6UEF0, Q84QK0, and W6Q9S9.[56] An all-*vs.*-all BLAST+ (Version 2.2.15) *E*-value matrix was generated with an edge cutoff of 1×10^{-87} and visualized in Cytoscope (Version 3.8.1).[57] This analysis was used to create the initial SSN of the entire OYE family with an alignment score threshold of 50%. The 2-ER SSN was subsequently generated by taking sequences (31,712) associated with cluster 2 of the OYE SSN and reclustering to 80% identity with an edge cutoff of 1×10^{-185} to yield 5744 representative nodes.

Transformation and expression of OYE*Bi*

The OYE*Bi* gene (European Nucleotide Archive accession code: BAO89787) from *Burkholderia* sp. RPE67, recently reclassified as *B. insecticola*, was codon-optimized for expression in *E. coli* and cloned into a pET-28a vector. The construct was designed with an N-terminal SUMO solubility tag linked to a strep II purification tag *via* a TEV cleavage site (Fig. S7A, ESI†). BL21(DE3) electrocompetent cells were co-transformed with the OEY*Bi* plasmid and pDB1282 for expression. pDB1282 imparts ampicillin resistance and encodes six genes from the *isc* operon associated with the assembly of Fe/S clusters in *Azotobacter vinelandii* under control of an arabinose-inducible promoter (Fig. S7B, ESI†). Associated cells were cultured in SOC outgrowth media (New England BioLabs) for 2 h at 37 °C before being transferred onto a Luria-Bertani (LB) agar plate containing ampicillin (100 μg mL^{-1}) and kanamycin (50 μg mL^{-1}) for incubation overnight at 37 °C. A single colony was subsequently used to inoculate a 10 mL culture tube containing 5 mL of LB broth. Upon reaching an optical density at 600 nm (OD_{600}) of 0.6, the cells were combined 1 : 1 with 50% glycerol and flash frozen in liquid nitrogen for storage at −80 °C.

All subsequent expression procedures utilized these cells. Detailed protocols are described below. In each case, OYE*Bi* expression was induced by the addition of 200 μM isopropyl β-D-1-thiogalactopyranoside (IPTG), and cultures were left to grow overnight (~16–18 h) while shaking at 180 rpm, unless

otherwise stated. Cells were harvested *via* centrifugation at 7500*g* for 15 min at 4 °C and flash frozen in liquid nitrogen for storage at −80 °C.

Anaerobic expression. A 50 mL culture tube containing 10 mL of LB media supplemented with ampicillin and kanamycin was inoculated with cells carrying the OYE*Bi* gene (*vide supra*) and incubated overnight at 37 °C while shaking. Aliquots (1 mL) of this culture were used to inoculate 250 mL flasks containing 100 mL of LB, which were again incubated and shaken at 37 °C and 180 rpm, this time to an OD_{600} of ∼1. After resting at 4 °C overnight, cultures were centrifuged at 4000*g* for 30 min. Resultant pellets were transferred to a vinyl anaerobic chamber (Coy), resuspended in LB media, and used to inoculate 2 L flasks each containing 1 L of deoxygenated Difco™ Terrific broth supplemented with the same antibiotics and excess FMN/FAD. Note that the broth was prepared by cooling freshly autoclaved media to room temperature (RT) in an oxygen-free environment. RT anaerobic cultures were stirred at 200 rpm until they reached an OD_{600} of ∼0.6. Expression of the *isc* operon encoded by pDB1282 was subsequently induced by addition of 0.2% w/v arabinose. Cells were simultaneously supplemented with $FeCl_3$ hexahydrate (∼150 μM) and L-cysteine hydrochloride (∼190 μM) and left to grow until the cultures reached an OD_{600} of ∼0.85 before inducing expression of OYE*Bi*.

Pseudo-anaerobic expression. A 1 L Erlenmeyer flask containing 250 mL of LB media supplemented with ampicillin and kanamycin was inoculated with cells carrying the OYE*Bi* gene and incubated overnight at 37 °C while shaking. Aliquots (6 mL) of this culture were used to inoculate 4 L flasks containing 1.5 L of LB supplemented with the same antibiotics and excess FMN/FAD. Cultures were again grown at 37 °C and shaken at 180 rpm until they reached an OD_{600} of ∼0.6, at which time the temperature was reduced to 18 °C and expression of the *isc* operon and OYE*Bi* induced as described above.

Aerobic expression. Aerobic expression followed the same procedure as the pseudo-anaerobic protocol, with the exception of *isc* operon induction. Instead, only expression of OYE*Bi* was initiated when cultures reached an OD_{600} of ∼0.6.

Purification of OYE*Bi*

Cell pellets were thawed and resuspended in chilled buffer A (2.5 mL g^{-1} pellet, 40 mM Tris–HCl, pH 8 and 10 mM NaCl) supplemented with DNase I (Millipore) and 1:1000 1× protease inhibitor cocktail (Thermo Scientific). After a 30–45 min incubation on ice, the mixture was sonicated for 45 min (1 min on/8 min off at 30% power) to lyse the cells. Centrifugation of the resultant lysate at 20 000*g* for 1 h at 4 °C yielded clarified lysate. Aerobic and pseudo-anaerobically generated crude extracts were subsequently incubated on ice for 1 h with excess FMN and FAD before being loaded on a 5 mL StrepTrap HP Strep-Tactin Sepharose column (Cytiva). The column was washed with two column volumes (CVs) of buffer A, followed by elution with 3 CVs of buffer B (40 mM Tris–HCl, pH 8.0, and 0.5 mM desthiobiotin). The presence of protein was monitored by absorbance at 280 nm. During aerobic purification, multiwavelength detection enabled simultaneous tracking of flavin absorbance at 450 nm. The eluate containing OYE*Bi* was concentrated and exchanged into buffer A for further purification by anion exchange chromatography on a 5 mL HiTrap Q FF column. The column was washed with two CVs of buffer A, followed by a linear gradient (15–18 CVs) containing 0% to 50% buffer C (40 mM Tris–HCl pH 8.0, and 2 M NaCl). Fractions contributing to the highest absorbing peak at 280 nm were collected and assessed for purity by SDS-PAGE before being concentrated to >5 mg mL^{-1} as quantified by a Bradford assay.

The much smaller culture volumes associated with anaerobic expression resulted in substantially lower yields that limited utility of column chromatography for purification. Instead, associated clarified lysate was passed through a series of membrane filters having molecular weight cutoffs of 30 kDa and 10 kDa, respectively. Note that purification of OYE*Bi* obtained through anaerobic and pseudo-anaerobic expression procedures was performed within an anaerobic chamber to encourage retention of the putative Fe/S cluster. Although centrifugation steps were performed outside of the chamber, tubes were tightly sealed with parafilm, and time spent outside the anaerobic environment was minimized. Aerobic chromatography was accomplished *via* the use of a Bio-Rad NGC Quest 10+ fast protein liquid chromatography (FPLC) system housed in a cold room maintained at 4 °C, while anaerobic chromatography was performed at RT with an ÄKTA go FPLC instrument housed within the anaerobic chamber.

Reconstitution of OYE*Bi*

All protein reconstitution steps were performed within the anaerobic chamber. OYE*Bi* was diluted to 100 μM in a degassed solution of 50 mM HEPES, pH 7.5, 300 mM KCl, 5 mM DTT, and 20% glycerol. For every additional 20 mL of solution, DTT was also added in three even increments separated by 20 min intervals to a final concentration of 5 mM, before incubating on ice for 1 h. $FeCl_3$ was subsequently added to the mixture in five even increments separated by 5 min intervals to a final concentration of 400 μM, before incubating for 30 min. Finally, Na_2S was added in five even increments separated by 20 min intervals to a final concentration of 400 μM, before incubating on ice overnight. The following day, protein precipitation was removed by centrifugation at 14 000*g* for 15 min. The reconstituted protein was immediately flash frozen.

Assessment of reactivity

All activity assays were conducted in an anaerobic chamber. Note that aerobically purified enzyme was not degassed to remove dissolved oxygen. Reaction vials were prepared in triplicate with 500 nM OYE*Bi* (from either purified enzyme or filter-purified clarified whole-cell lysate), 250 mM substrate, 10 μM NADH, 2 U glucose dehydrogenase, and 100 mM glucose in 100 mM phosphate buffer at pH 8. Note that substrates were dissolved in DMSO such that the final concentration in the reaction mixture was <0.5 v/v%. Glucose dehydrogenase and glucose were included to allow for nicotinamide regeneration. Reactions were left to proceed at RT for 24 h before being

quenched. Further specifics are presented by substrate below. Percent conversions were then analyzed by mass spectrometry.

NAD^+-dependent assays were performed in the absence of glucose dehydrogenase and glucose as previously described.[10] Reactions were prepared in triplicate with 250 μM NAD^+ or NADH, 250 μM substrate, and 250 nM enzyme. Excess NAD^+ and NADH were included to ensure assays achieved detectable product levels without the glucose dehydrogenase and glucose nicotinamide regeneration systems. Reactions were initiated with the enzyme and monitored *via* UV spectroscopy at 340 nm for consumption or production of NADH. Those displaying dramatic changes in nicotinamide were further analyzed by MS.

Substrates 3–8. Reactions were quenched with 0.5 mM cyclohexanone in ethyl acetate and analyzed *via* GC–MS (QP2010-SE from Shimadzu Scientifics Instruments, Kyoto Japan) with an after-column splitter, a flame ionization detector (FID) detector (detector temperature 200 °C, split ratio 1 : 1) and GC–MS detector with a flow rate column flow 3.69 mL min^{-1}. The GC–MS initial oven temperature was held at 55 °C, ramped up at 2.50 °C min^{-1} to 100 °C, again ramped up at 20 °C min^{-1} to 140 °C, then finally increased at 5.00 °C min^{-1} to 210 °C and held for 5 min. Retention times of the internal standard cyclohexanone, the remaining substrate post-reaction, and products produced during each reaction (if present) were compared to the retention times of substrate standards and confirmed by the mass peaks for each compound. Assay components were separated using a chiral CycloSil-B column (30 m × 0.32 mm, 0.25 μm film, Agilent, Santa Clara, CA).

Substrates 1, 2, 9, and 10. Reactions were quenched by freezing overnight at −20 °C to avoid degradation of associated products, and subsequently analyzed *via* liquid chromatography (LC)–MS. Components were separated on an Eclipse Plus C18 column (4.6 mm × 150 mm × 5 μm) equilibrated in 98% solvent A (water) and 2% solvent B (acetonitrile), with 10 μL injections at a flow rate of 1 mL min^{-1}. A multi-step gradient from 2–60% solvent B was applied over 10 minutes, followed by an increase to 95% over 3 minutes, with a 5-minute re-equilibration between injections. LC–MS was performed on an LTQ Orbitrap Velos (Thermo Fisher, Waltham, MA) with a Dionex Ultimate 3000 dual pump, diode array detector, and a Shimadzu SIL-20AC HT autosampler managed by a CBM20 controller. The instrument was handled by Xcalibur and Chomeleon Xpress software. Spectra were collected at a resolution of 60 000 with an automatic gain control setting of 50 000 and a maximum injection time of 100 ms. The system employed a heated electrospray ionization probe at 3.0 kV with a source temperature of 350 °C, capillary temperature of 320 °C, and the S-lens radio frequency (RF) level of 60%. Data for 1 and 2 were collected in positive mode, while 9 and 10 were recorded in negative mode.

Iron quantitation

All protein samples (1 μM−5 μM) in buffer B were thermally denatured at 95 °C and diluted 1 : 100 in oxygen-free water. The associated iron content was then determined *via* inductively coupled plasma (ICP)-MS (PerkinElmer NexION 2000) at the Mass Spectrometry Center in the Department of Chemistry at Emory University. Standards were generated by serial dilution of the Instrument Calibration Standard 2 from SPEX CertiPrep (catalog no. CL-CAL-2). External standards were 0, 1.000, 10.000, 100.000, and 1000.000 ppb Fe per mL, respectively.

EPR analysis

Oxidized $[4Fe–4S]^{2+}$ clusters are EPR silent as the iron ions are antiferromagnetically coupled to yield a total spin of zero. Samples were, therefore, prepared by incubating 250 μM OYE*Bi* with 2.5 mM sodium dithionite at 4 °C for 10 min to generate $[4Fe–4S]^{+}$ before transfer to EPR tubes for flash freezing in the anaerobic chamber. X-band (9.35 GHz) continuous wave spectra were recorded on a Bruker EMX-Plus spectrometer housed at the University of Georgia NMR facility. Data were collected using the following parameters: microwave frequency, 9.35 GHz; field modulation amplitude, 9.59 G at 100 kHz; microwave power, 1.0 mW. The sample temperature was controlled *via* Stinger (Cold-Edge) closed cycle He flow system. Spectra were simulated using EasySpin for MATLAB.[58]

Genomic neighborhood analysis of OYE*Bi*

Cluster 2 SSN generated as previously described, was uploaded into the Enzyme Function Initiative (EFI) server[51] and analyzed for 20% CoF. Resultant genomic neighborhood networks were visualized in Cytoscape.

Author contributions

Conceptualization: T. C. B.-L. & K. M. D.; investigation: T. C. B.-L. & S. K. J.; analysis: T. C. B.-L., S. K. J. & K. M. D. funding acquisition and supervision: K. M. D.; primary writing: T. C. B.-L. & K. M. D.; review & editing: T. C. B.-L., S. K. J. & K. M. D.

Data availability

The bioinformatics analysis herein was carried out using publicly available data from UniProt at **https://www.uniprot.org/** with accession number PF00724. Other data supporting this article have been included as part of the ESI.†

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was supported by Emory University start-up funds (K. M. D.), and the National Science Foundation Graduate Research Fellowship Program (grant no. 1937971 to S. K. J.). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. MS data were collected in the Mass Spectrometry Center at Emory University with support from Dr Frederick Strobel. EPR

data were collected with assistance from Jordan Dinning at the University of Georgia's NMR Facility, supported by the NSF MRI Program (grant no. 1827968). We additionally thank Dr Squire Booker for providing the pDB1282 plasmid, and Nicholas York for guidance in generating simulated EPR data.

Notes and references

1 H. S. Toogood, J. M. Gardiner and N. S. Scrutton, *ChemCatChem*, 2010, **2**, 892–914.
2 H. S. Toogood and N. S. Scrutton, *ACS Catal.*, 2018, **8**, 3532–3549.
3 R. M. Kohli and V. Massey, *J. Biol. Chem.*, 1998, **273**, 32763–32770.
4 A. Castellan, J. C. J. Bart and S. Cavallaro, *Catal. Today*, 1991, **9**, 237–254.
5 S. R. Nicholson, N. A. Rorrer, T. Uekert, G. Avery, A. C. Carpenter and G. T. Beckham, *ACS Sustainable Chem. Eng.*, 2023, **11**, 2198–2208.
6 K. A. Curran, J. M. Leavitt, A. S. Karim and H. S. Alper, *Metab. Eng.*, 2013, **15**, 55–66.
7 J. C. Joo, A. N. Khusnutdinova, R. Flick, T. Kim, U. T. Bornscheuer, A. F. Yakunin and R. Mahadevan, *Chem. Sci.*, 2017, **8**, 1406–1413.
8 T. Reed, G. H. Lushington, Y. Xia, H. Hirakawa, D. M. Travis, M. Mure, E. E. Scott and J. Limburg, *J. Biol. Chem.*, 2010, **285**, 25782–25791.
9 C. Loechel, A. Basran, J. Basran, N. S. Scrutton and E. A. H. Hall, *Analyst*, 2003, **128**, 166–172.
10 D. W. White, S. Iamurri, P. Keshavarz-Joud, T. Blue, J. Copp and S. Lutz, *BioRxiv*, 2023, preprint, DOI: **10.1101/2023.07.10.548207**.
11 Q. Shi, H. Wang, J. Liu, S. Li, J. Guo, H. Li, X. Jia, H. Huo, Z. Zheng, S. You and B. Qin, *Appl. Microbiol. Biotechnol.*, 2020, **104**, 8155–8170.
12 U. T. Bornscheuer, *Philos. Trans. R. Soc., A*, 2018, **376**, 20170063.
13 J. N. Copp, E. Akiva, P. C. Babbitt and N. Tokuriki, *Biochemistry*, 2018, **57**, 4651–4662.
14 P. M. Mordaka, S. J. Hall, N. Minton and G. Stephens, *Microbiology*, 2018, **164**, 122–132.
15 X. Tu, P. A. Hubbard, J.-J. P. Kim and H. Schulz, *Biochemistry*, 2008, **47**, 1167–1175.
16 G. Liu, W. Wang, F. He, P. Zhang, P. Xu and H. Tang, *Appl. Environ. Microbiol.*, 2020, **86**, e01559-20.
17 P. Trickey, J. Basran, L. Y. Lian, Z. Chen, J. D. Barton, M. J. Sutcliffe, N. S. Scrutton and F. S. Mathews, *Biochemistry*, 2000, **39**, 7678–7688.
18 M. Willistein, D. F. Bechtel, C. S. Müller, U. Demmer, L. Heimann, K. Kayastha, V. Schünemann, A. J. Pierik, G. M. Ullmann, U. Ermler and M. Boll, *Nat. Commun.*, 2019, **10**, 2074.
19 G. E. Crooks, G. Hon, J. M. Chandonia and S. E. Brenner, *Genome Res.*, 2004, **14**, 1188–1190.
20 N. D. Lanz, T. L. Grove, C. B. Gogonea, K. H. Lee, C. Krebs and S. J. Booker, *Methods Enzymol.*, 2012, **516**, 125–152.
21 M.-K. Kim, M. K. Cho, H.-E. Song, D. Kim, B.-H. Park, J. H. Lee, G. B. Kang, S. H. Kim, Y. J. Im, D.-S. Lee and S. H. Eom, *Proteins: Struct., Funct., Bioinf.*, 2007, **66**, 751–754.
22 T. E. Benson, C. T. Walsh and V. Massey, *Biochemistry*, 1997, **36**, 796–805.
23 D. J. Opperman, B. T. Sewell, D. Litthauer, M. N. Isupov, J. A. Littlechild and E. van Heerden, *Biochem. Biophys. Res. Commun.*, 2010, **393**, 426–431.
24 G. Caserta, L. Zuccarello, C. Barbosa, C. M. Silveira, E. Moe, S. Katz, P. Hildebrandt, I. Zebger and S. Todorovic, *Coord. Chem. Rev.*, 2022, **452**, 214287.
25 W. E. Broderick and J. B. Broderick, *J. Biol. Inorg. Chem.*, 2019, **24**, 769–776.
26 J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli and D. Hassabis, *Nature*, 2021, **596**, 583–589.
27 J. Frazzon and D. R. Dean, *Curr. Opin. Chem. Biol.*, 2003, **7**, 166–173.
28 F. W. Outten, *Nat. Chem. Biol.*, 2007, **3**, 206–207.
29 M. A. Swanson, R. J. Usselman, F. E. Frerman, G. R. Eaton and S. S. Eaton, *Biochemistry*, 2008, **47**, 8894–8901.
30 D. F. Becker, U. Leartsakulpanich, K. K. Surerus, J. G. Ferry and S. W. Ragsdale, *J. Biol. Chem.*, 1998, **273**, 26462–26469.
31 H. Rupp, K. K. Rao, D. O. Hall and R. Cammack, *Biochim. Biophys. Acta, Protein Struct.*, 1978, **537**, 255–269.
32 J. Sun, Y. Lin, X. Shen, R. Jain, X. Sun, Q. Yuan and Y. Yan, *Metab. Eng.*, 2016, **35**, 75–82.
33 W. Shi, J. Mersfelder and R. Hille, *J. Biol. Chem.*, 2005, **280**, 20239–20246.
34 M.-H. Jang, J. Basran, N. S. Scrutton and R. Hille, *J. Biol. Chem.*, 1999, **274**, 13147–13154.
35 J. Basran, M.-H. Jang, M. J. Sutcliffe, R. Hille and N. S. Scrutton, *J. Biol. Chem.*, 1999, **274**, 13155–13161.
36 D. J. Steenkamp and J. Mallinson, *Biochim. Biophys. Acta, Enzymol.*, 1976, **429**, 705–719.
37 Y. V. S. N. Murthy, Y. Meah and V. Massey, *J. Am. Chem. Soc.*, 1999, **121**, 5344–5345.
38 K. D. Daughtry, Y. Xiao, D. Stoner-Ma, E. Cho, A. M. Orville, P. Liu and K. N. Allen, *J. Am. Chem. Soc.*, 2012, **134**, 2823–2834.
39 J. Basran, M. J. Sutcliffe and N. S. Scrutton, *J. Biol. Chem.*, 2001, **276**, 42887–42892.
40 A. M. Bashir, T. Hoffmann, B. Kempf, X. Xie, S. H. J. Smits and E. Bremer, *Microbiology*, 2014, **160**(Pt 10), 2283–2294.
41 A. Ly, J. Henderson, A. Lu, D. E. Culham and J. M. Wood, *J. Bacteriol.*, 2004, **186**, 296–306.
42 F. J. Detmers, F. C. Lanfermeijer and B. Poolman, *Res. Microbiol.*, 2001, **152**, 245–258.

43 K. Tomii and M. Kanehisa, *Genome Res.*, 1998, **8**, 1048–1059.
44 N. Gurich and J. E. González, *J. Bacteriol.*, 2009, **191**, 4372–4382.
45 E. Bremer, *Comp. Biochem. Physiol., Part A: Mol. Integr. Physiol.*, 2000, **126**, 17.
46 G. Alloing, I. Travers, B. Sagot, D. L. Rudulier and L. Dupont, *J. Bacteriol.*, 2006, **188**, 6308–6317.
47 M. W. Burnet, A. Goldmann, B. Message, R. Drong, A. El Amrani, O. Loreau, J. Slightom and D. Tepfer, *Gene*, 2000, **244**, 151–161.
48 F. Cunningham, J. E. Allen, J. Allen, J. Alvarez-Jarreta, M. R. Amode, I. M. Armean, O. Austine-Orimoloye, A. G. Azov, I. Barnes, R. Bennett, A. Berry, J. Bhai, A. Bignell, K. Billis, S. Boddu, L. Brooks, M. Charkhchi, C. Cummins, L. Da Rin Fioretto, C. Davidson, K. Dodiya, S. Donaldson, B. El Houdaigui, T. El Naboulsi, R. Fatima, C. G. Giron, T. Genez, J. G. Martinez, C. Guijarro-Clarke, A. Gymer, M. Hardy, Z. Hollis, T. Hourlier, T. Hunt, T. Juettemann, V. Kaikala, M. Kay, I. Lavidas, T. Le, D. Lemos, J. C. Marugán, S. Mohanan, A. Mushtaq, M. Naven, D. N. Ogeh, A. Parker, A. Parton, M. Perry, I. Piližota, I. Prosovetskaia, M. P. Sakthivel, A. Imran, A. Salam, B. M. Schmitt, H. Schuilenburg, D. Sheppard, J. G. Pérez-Silva, W. Stark, E. Steed, K. Sutinen, R. Sukumaran, D. Sumathipala, M.-M. Suner, M. Szpak, A. Thormann, F. F. Tricomi, D. Urbina-Gómez, A. Veidenberg, T. A. Walsh, B. Walts, N. Willhoft, A. Winterbottom, E. Wass, M. Chakiachvili, B. Flint, A. Frankish, S. Giorgetti, L. Haggerty, S. E. Hunt, G. R. IIsley, J. E. Loveland, F. J. Martin, B. Moore, J. M. Mudge, M. Muffato, E. Perry, M. Ruffier, J. Tate, D. Thybert, S. J. Trevanion, S. Dyer, P. W. Harrison, K. L. Howe, A. D. Yates, D. R. Zerbino and P. Flicek, *Nucleic Acids Res.*, 2021, **50**, D988–D995.
49 J. G. Lawrence, *Cell*, 2002, **110**, 407–413.
50 S. Calhoun, M. Korczynska, D. J. Wichelecki, B. San Francisco, S. Zhao, D. A. Rodionov, M. W. Vetting, N. F. Al-Obaidi, H. Lin, M. J. O'Meara, D. A. Scott, J. H. Morris, D. Russel, S. C. Almo, A. L. Osterman, J. A. Gerlt, M. P. Jacobson, B. K. Shoichet and A. Sali, *eLife*, 2018, **7**, e31097.
51 R. Zallot, N. Oberg and J. A. Gerlt, *Biochemistry*, 2019, **58**, 4169–4182.
52 L. Poppe and B. G. Vértessy, *ChemBioChem*, 2018, **19**, 284–287.
53 G. Gaur and M. Gänzle, *Lett. Appl. Microbiol.*, 2024, **77**, ovae109.
54 Y. Wang, Q. Wang, H. Huang, W. Huang, Y. Chen, P. B. McGarvey, C. H. Wu, C. N. Arighi and UniProt Consortium, *PLoS Biol.*, 2021, **19**, e3001464.
55 L. Fu, B. Niu, Z. Zhu, S. Wu and W. Li, *Bioinformatics*, 2012, **28**, 3150–3152.
56 A. Bairoch and R. Apweiler, *Nucleic Acids Res.*, 2000, **28**, 45–48.
57 P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker, *Genome Res.*, 2003, **13**, 2498–2504.
58 S. Stoll and A. Schweiger, *J. Magn. Reson.*, 2006, **178**, 42–55.