Title

Co-occurrence of enzyme domains guides the discovery of an oxazolone synthetase

Journal

Nature Chemical Biology, 17(7)

ISSN

1552-4450

Authors

de Rond, Tristan
Asay, Julia E
Moore, Bradley S

Publication Date

2021-07-01

DOI

10.1038/s41589-021-00808-4

Peer reviewed

HHS Public Access

Author manuscript

Nat Chem Biol. Author manuscript; available in PMC 2022 January 01.

Published in final edited form as:

Nat Chem Biol. 2021 July ; 17(7): 794–799. doi:10.1038/s41589-021-00808-4.

Co-Occurrence of Enzyme Domains Guides the Discovery of an Oxazolone Synthetase

Tristan de Rond[1], **Julia E. Asay**[1], **Bradley S. Moore**[1,2]

[1]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093

[2]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA 92093

Abstract

Multidomain enzymes orchestrate two or more catalytic activities to carry out metabolic transformations with increased control and speed. Here, we report the design and development of a genome mining approach for the targeted discovery of biochemical transformations through the analysis of co-occurring enzyme domains (CO-ED) in a single protein. CO-ED was designed to identify unannotated multifunctional enzymes for functional characterization and discovery based on the premise that linked enzyme domains have evolved to function collaboratively. Guided by COED, we targeted an unannotated predicted ThiF-nitroreductase di-domain enzyme found in more than 50 proteobacteria. Through heterologous expression and biochemical reconstitution, we discovered a series of natural products containing the rare oxazolone heterocycle, and characterized their biosynthesis. Notably, we identified the di-domain enzyme as an oxazolone synthetase, validating CO-ED-guided genome mining as a methodology with potential broad utility for both the discovery of unusual enzymatic transformations and the functional annotation of multidomain enzymes.

Introduction

Our knowledge of nature's diversity of enzymatic transformations is crucial to advancing research in a multitude of disciplines. For instance, our ability to predict metabolic capacity from genome sequences enables new insights in human health and ecology[1], while the development of bioprocesses to produce chemicals in an environmentally benign fashion

Corresponding authors: Correspondence to Tristan de Rond tderond@ucsd.edu or Bradley S. Moore bsmoore@ucsd.edu.

AUTHOR CONTRIBUTIONS

TdR and BSM designed research; TdR and JEA performed research; TdR analyzed data; TdR and BSM wrote the paper

Code Availability

Jupyter notebooks containing Python code for CO-ED analysis and to generate the statistics shown in Supplementary Figs. 1 and 18 are publicly available at https://github.com/tderond/CO-ED.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests

SUPPLEMENTARY INFORMATION

The Supporting Information contains Supplementary Figures and a Supplementary Note containing information about structural characterization, chemical synthesis and spectral data.

Pre-generated CO-ED networks are available as Supplementary Data 1 and can be opened with Cytoscape

relies on the availability of a well-stocked biocatalytic toolbox[2]. Billions of years of evolution has resulted in immense natural genetic diversity, which we are rapidly starting to uncover using modern sequencing technologies. However, the functional assignment of the enzymes encoded by this sequence diversity is lagging[3,4]. Many enzymes catalyzing chemical transformations previously not known to occur in nature are still being discovered[5], and there is no end in sight of unannotated and mis-annotated enzymes in genomic databases[6]. The search for novel enzymes in this genomic wilderness can fulfill the dual purpose of functional gene annotation and biocatalyst discovery.

The computational searches for biosynthetic enzymes that underlie “genome mining” campaigns are most commonly based on gene homology to known core biosynthetic genes of well-established classes of medicinal natural products such as polyketides, non-ribosomal peptides and terpenoids[7]. This approach, while likely to identify gene clusters that produce bio-active products, is prone to the re-discovery of known enzymology. Therefore, we set out to develop a genome mining strategy that could guide enzyme discovery in a complementary manner.

One notable feature amongst specialized metabolism is an overabundance of enzymes harboring multiple catalytic domains (Supplementary Fig. 1). Multidomain enzymes are thought to arise through gene fusion of two or more single-domain enzymes, affording catalytic advantages such as coupled temporal and spatial regulation, a fixed active site stoichiometry, and channeling of reactive intermediates[8,9]. Over time, a multidomain enzyme may even evolve to orchestrate the reactivity of its constituent active sites to such an extent that it acquires a new function[10]. Classic examples of multidomain enzymes are the fatty acid synthase, polyketide synthase (PKS), and nonribosomal peptide synthetase (NRPS) assembly-line enzymes[11], some of which consist of dozens of domains. More recently, multidomain enzymes have been shown to catalyze epimerization[12] and *N*-nitrosation[13] reactions.

Here, we present a genome mining paradigm where we leverage the evolved co-occurrence of enzymatic domains (CO-ED) in proteins to inform enzyme discovery. Through CO-ED analysis, we discovered a series of new oxazolone-containing natural products, as well as the first reported oxazolone synthetase, a novel bifunctional cyclodehydratase–oxidoreductase.

Results

CO-ED–guided identification of a di-domain enzyme candidate

The CO-ED workflow takes a query set of proteins and generates a network representation of co-occurring enzymatic domains (Fig. 1). A node is drawn for each of a curated set of enzymatic Pfam[14] domains found in the query proteins, and an edge is drawn between two nodes for each two-domain combination of domains found together in at least one query protein. In parallel, a set of characterized enzymes, derived from the BRENDA[15], MIBiG[16], and Uniprot[17] databases, is subjected to the same analysis. If an edge in the network represents a domain pair found in a previously characterized multidomain enzyme, the edge is color coded according to the originating database. Remaining uncolored edges, therefore, represent enzyme domain combinations that are likely uncharacterized. A public web server

allowing for CO-ED analysis on user-supplied data sets is available at http://enzyme-analysis.org.

To validate the CO-ED workflow, we first applied it to the genome of the model organism *Escherichia coli* K12. The *E. coli* network (Fig. 2a,b, Supplementary Figs. 2 and 3, Supplementary Data 1) revealed 19 co-occurring pairs of domains, including some well-studied multifunctional enzymes including penicillin-binding proteins[18], the aldehyde dehydrogenase–alcohol dehydrogenase AdhE[19], the UDP-L-Ara4FN biosynthesis enzyme ArnA[20], several bifunctional amino acid biosynthesis enzymes, and the enterobactin NRPS EntF[21]. Every edge is colored, suggesting that the function of every multidomain enzyme in *E. coli* detected by CO-ED has likely already been established.

We next applied CO-ED analysis to the genome of the non-model marine gammaproteobacterium *Pseudoalteromonas rubra* DSM 6842, chosen because of our experience heterologously expressing enzymes from this and closely-related species[22–24]. The *P. rubra* CO-ED network (Fig. 2c, Supplementary Fig. 4, Supplementary Data 1) revealed numerous NRPSs, as can be inferred from the edge indicating that the "AMP-binding" (also known as adenylation) and "Condensation" Pfam domains co-occur in thirty-eight proteins, including two hybrid PKS/NRPS enzymes, as indicated by the two proteins containing both "ketoacyl-synt" and "Condensation" Pfam domains. The network also shows a putative bifunctional RibBA, known to be involved in riboflavin biosynthesis in other species[25], as well as some of the same bifunctional primary metabolic enzymes observed in *E. coli*. However, unlike the *E. coli* network, twelve domain pairs are unannotated, suggesting that *P. rubra* may harbor multidomain enzymes that catalyze undiscovered enzymatic transformations.

A *P. rubra* protein harboring a "ThiF"-"nitroreductase" domain pair (outlined in bottom right of Fig. 2c) caught our attention, because each of these domains is known to catalyze a diverse set of chemical transformations (Extended Data Fig. 1), yet their combination is unprecedented. Enzymes of the ThiF protein family catalyze various vital reactions that proceed through carboxylate adenylation, including the activation of ubiquitin and ubiquitin-like proteins by E1 enzymes[26], the cyclization of N^6-threonylcarbamoyladenosine in tRNA maturation[27], and the post-translational modification of ribosomally-synthesized peptide antibiotics[28]. ThiF family enzymes also play a role in the sulfur incorporation machinery for thiamin, the molybdenum cofactor and some natural products[29,30]. To date, all characterized substrates of ThiF family enzymes have been polypeptides or tRNA. The nitroreductase family consists of flavoenzymes that catalyze a remarkable variety of redox reactions[31], such as the reductive deiodination of thyroid hormones[32], the dehydrogenation of oxazolines and thiazolines[33], the fragmentation of flavin mononucleotide to form the vitamin B12 lower ligand 5,6-dimethylbenzimidazole[34], enol methylene reduction in cofactor F420 biosynthesis[35], diketopiperazine desaturation in albonoursin biosynthesis[36], and the reductive detoxification of various nitro functional groups[37].

Querying the Uniprot database with the di-domain ThiF-nitroreductase enzyme from *P. rubra,* which we named *oxzB*, revealed 56 proteins originating from alpha-, beta-, and gammaproteobacteria (Extended Data Fig. 1b, Supplementary Fig. 5) isolated from a variety

of (predominantly aquatic) environments (Extended Data Fig. 1c). In those cases where the genomic context of the *oxzB* homolog was known, it is typically accompanied by an *N*-acyltransferase homolog (*oxzA*). In *P. rubra* and many other species, *oxzA* is immediately upstream of *oxzB* on the same strand, forming an apparent two-gene *oxzAB* operon (Fig. 3a and Extended Data Fig. 1d).

Heterologous expression of *oxzAB* produces novel metabolites

To investigate the function of the *oxzAB* genes, we PCR amplified *oxzAB* from *P. rubra*, *Rheinheimera pacifica*, *Colwellia chukchiensis* (all gammaproteobacteria), *Skermanella aerolata* (alphaproteobacterium), and *Undibacterium pigrum* (betaproteobacterium), and heterologously expressed the gene pairs in *E. coli*. Colonies on solid media and cell pellets after growth in liquid media both turned visibly yellow (Supplementary Fig. 6). The color would briefly intensify when incubated above pH 9, suggesting the possibility of a phenolic, base-labile product (Supplementary Figs. 6 and 7). HPLC-UV-MS analysis of extracts of the pellets revealed several chromatographic peaks with absorbance in the 300–400 nm range which were not produced by *E. coli* natively (Fig. 3b). The product profiles of heterologously-expressed *oxzAB* genes originating from different species shared many constituents, and the inclusion of genes found up- or downstream from *oxzAB* did not substantially change the product profiles (Supplementary Fig. 8).

Through a combination of mass spectrometry and spectroscopic techniques, with notable insight provided by 1,1-ADEQUATE NMR, ^{1}H-^{15}N HMBC NMR, and the identity of degradation products formed in basic methanol, we were able to assign the products of *oxzAB* as a series of oxazolones (Fig. 3c, Supplementary Figs. 9 and 10, and Supplementary Note). We named these seemingly tyrosine- and phenylalanine-derived molecules tyrazolones and phenazolones, respectively. The major products were heptyl- nonyl-, undecyl-, and ω−6-undecenyltyrazolone (**2**, **4**, **5**, **6**, and **7**, respectively), which were found to occur in an equilibrium of (*E*) and (*Z*) isomers with the latter being predominant, and (*Z*)-heptylphenazolone (**9**). None of these products have previously been reported before in the literature, however an oxazolone that appears to be derived from tryptophan, almazolone (Fig. 3d), was isolated from a red alga[38] and its structure confirmed through synthesis. Almazolone similarly exists as an (*E*)/(*Z*) equilibrium, forms degradation products analogous to those formed by the tyrazolones and phenazolones, and has similar spectroscopic properties, lending credence to our structural assignments.

In vitro reconstitution of oxazolone biosynthesis by OxzAB

Considering the enzymatic domains contained within *oxzAB*, we hypothesized oxazolone biosynthesis to proceed as follows: OxzA forms an *N*-acyl amino acid using an acyl-CoA derived from the natural fatty acyl-CoA beta oxidation pool. In order to form the oxazolone, OxzB then catalyzes cyclization and oxidation of the *N*-acyl amino acid, in either order (Supplementary Fig. 11). The ThiF family enzyme TcdA is known to catalyze cyclodehydration of an amino acid moiety, and the nitroreductase family enzyme AlbA has been shown to oxidize amino acid-derived substrates (Extended Data Fig. 1a). It is thus conceivable that OxzB, which harbors both these domains, could act to form the oxazolone heterocycle.

To test our biosynthetic hypothesis, we heterologously expressed and purified *P. rubra* OxzA and OxzB in *E. coli* as N-terminal His6 and MBP fusion proteins, respectively (Supplementary Fig. 12), and assayed their activity by HPLC (Fig. 4). When incubated with L-tyrosine (**12**) and decanoyl-CoA (**13**), OxzA catalyzed the formation of *N*-decanoyl- L-tyrosine (**14**). OxzB was able to catalyze the formation of nonyltyrazolone both from synthetic *N*-decanoyltyrosine as well as when combined with OxzA, tyrosine and decanoyl-CoA. OxzB activity was only observed in the presence of ATP, as is typical for ThiF family enzymes. (*Z*)-nonyltyrazolone, the favored isomer of the (*E*)/(*Z*) equilibrium, was the major product upon extended incubation of the enzymatic reactions, however, shorter incubation times revealed a proportionally greater amount of (*E*)-tyrazolone (Supplementary Fig. 13), suggesting that the immediate product of OxzB is (*E*)-nonyltyrazolone.

Analysis of oxazolone production in the native hosts

To determine if the oxazolones are true natural products or merely artifacts of heterologous expression, we attempted to detect these molecules in the five bacteria whose *oxzAB* gene pairs we tested heterologously above. Under standard culturing conditions, oxazolones could be reliably detected only in *C. chukchiensis*. Encouraged by reports of chemical elicitation methods to activate silent gene clusters[39], we evaluated a series of antibiotics to induce oxazolone production in *P. rubra* or *C. chukchiensis* at sublethal concentrations (Extended Data Fig. 2 and Supplementary Fig. 14). Gratifyingly, we could elicit oxazolone production in *P. rubra* with erythromycin or chloramphenicol, both known inhibitors of protein synthesis[40], with oxazolone peak areas similar to those of uninduced *C. chukchiensis*. Oxazolone production in *C. chukchiensis* could be further induced by several antibiotics, only some of which inhibit protein synthesis. We were unable to detect tyrazolones or phenazolones in extracts of the other three species whose *oxzAB* gene pair we validated heterologously.

Discussion

Using a CO-ED–guided genome mining strategy, we discovered a series of novel oxazolone natural products as well as a unique oxazolone synthetase harboring ThiF and nitroreductase domains. This proof-of-principle experiment showcases the promise of CO-ED in selecting orphan multidomain enzymes for targeted functional discovery. Many computational workflows for the prioritization of unannotated genes for genome mining rely on identifying gene clusters based on sequence homology to known biosynthetic genes, and downstream analysis tools in turn often depend on these gene cluster annotations[7]. This approach risks overlooking genes that do not either themselves have homology to known biosynthetic genes, or are clustered with such genes. Alternatively, there exist workflows that allow the user to explore enzyme families related to a query sequence or domain, such as EnzymeMiner[41] and the tools developed by the Enzyme Function Initiative[42], but these might miss protein families that the user did not consider searching for. CO-ED is complementary to both the aforementioned approaches in that it considers unannotated enzyme domain co-occurrences in a relatively unbiased fashion.

The analysis of co-occurring protein domains has previously been applied in contexts other than genome mining, such as in evolutionary studies[43], to help identify antifungal drug targets[44], and in functional gene annotation (where it is sometimes called the “Rosetta Stone” method)[45,46]. Parallel to this, efforts have been made to connect protein domains to enzymatic reactions[42,47]. CO-ED represents, to our knowledge, the first time these concepts have been united to specifically guide the discovery of enzymes catalyzing novel transformations. The availability of a web tool at http://enzyme-analysis.org will allow any user to easily apply CO-ED to identify unstudied multi-domain enzymes present in their own datasets.

While CO-ED analysis has revealed that all multidomain enzymes in *E. coli* likely have known functions, we wished to explore whether this is the case for other model bacteria, particularly those that are promising sources of novel biocatalysts. To this end, we applied CO-ED to the genomes of *Streptomyces coelicolor* A3(2), *Salinispora tropica* CNB-440, and *Pseudomonas fluorescens* Pf-5 – model organisms whose remarkable biosynthetic capacities are well-studied. We found over a dozen unannotated domain pairs in each of these species (Supplementary Figs. 15–17). Investigating the functions of these multidomain enzymes may reveal yet more biochemical novelty even in these well-studied organisms.

CO-ED analysis of all proteins in the Uniprot[17] database (Supplementary Fig. 18 and Supplementary Data 1) shows that while the most abundantly distributed co-occurring domain pairs are represented by enzymes annotated in BRENDA, MIBiG, or Uniprot, there is a still a plethora of less-widespread multidomain enzymes whose functions are unassigned, with some notable examples shown in Supplementary Fig. 18b. Of the 252 enzymatic domain pairs found in at least 1000 proteins, 78.6% are annotated, whereas 82.8% of the 3381 domain pairs found in at least 10 proteins are unannotated (Supplementary Fig. 18c,d). These annotation rates may be an under-estimate because of a lag in the appearance of new annotated results in the databases, or an over-estimate if the same combination of enzymatic domains catalyzes different transformations in different proteins. Nevertheless, there are many more multidomain enzymes awaiting examination for new fundamental biochemical insight and as potential biocatalysts.

Oxazolones – also known as Erlenmeyer azlactones – are versatile synthetic intermediates[48], but no biocatalytic methods for their production have been reported. Biocatalytic approaches have the potential to exhibit better functional group tolerance than synthetic methods, which could shorten synthetic routes that rely on oxazolones. Besides OxzB, the only other known oxazolone-forming enzyme is MbnBC, which is involved in the biosynthesis of the chalkophore methanobactin[49]. MbnBC catalyzes a 4-electron oxidative rearrangement – a fundamentally different reaction from OxzB – and acts on a polypeptide substrate, which may pose challenges to developing of this enzyme as a biocatalyst. Considering that *P. rubra* OxzB is a soluble enzyme that acts on small molecule substrates, it and its homologs have the potential to become valuable biocatalysts.

Characterization of the product profiles of five *oxzAB* gene pairs led to the discovery of several novel oxazolone natural products. Dozens more *oxzAB* gene pairs are present in sequence databases, ready to be explored for further biosynthetic novelty. Given the

structural similarity of almazolone to the oxazolones described here, we expect a yet-undiscovered OxzB-like enzyme selective for a tryptophan derivative to be involved in almazolone biosynthesis. Some *oxzB* genes are clustered with predicted glycosyltransferases, PKSs and NRPSs, suggesting they may be part of biosynthetic gene clusters coding for more complex oxazolone-containing natural products (Supplementary Fig. 19).

The reaction catalyzed by OxzB can formally be understood to consist of a cyclization and an oxidation, presumably catalyzed by the ThiF-like and nitroreductase domains respectively. Ongoing mechanistic studies aim to determine the order in which these reactions occur, and whether any soluble intermediates are formed. OxzB is also the first characterized ThiF family enzyme that does not act on a macromolecular substrate but on a small molecule, potentially making it a useful model to gain insight into this important enzyme family.

We have shown that *P. rubra* and *C. chukchiensis* can produce oxazolone natural products, but it is unclear what the biological role of these molecules is. The oxazolone-containing methanobactins are chalkophores, but this function is mediated by thiolates that are not present in the tyrazolones or phenazolones. The observation that production of these molecules can be induced by antibiotics suggests that they may act as signals in response to stress. Like many signaling molecules, the oxazolones are somewhat labile in water; at pH 8 (e.g., seawater) they hydrolyze with a half-life on the order of hours. We have thus far observed no obvious phenotypes when bacteria (including the five species whose *oxzAB* genes we studied heterologously) are exposed to these molecules, both in the presence and absence of antibiotics. The genome neighborhoods of *oxzB* homologs occasionally reveal genes part of the PEP-CTERM suspected exopolysaccharide-associated protein sorting system[50], but no obvious patterns in transcriptional regulatory machinery or potential receptors for signaling molecules. As we continue to explore the distribution and repertoire of oxazolone biosynthetic genes, we hope to learn more about the biological functions of these newly discovered bacterial natural products.

Online Methods

Computational methods

A CO-ED analysis web tool is available at http://enzyme-analysis.org. For offline CO-ED analysis, code written in Python v3.6.10, Jupyter notebook v6.0.3, and pandas v1.0.3 is publicly available at https://github.com/tderond/CO-ED. Besides Python code, the Jupyter notebooks contain explanations of the CO-ED workflow's steps and example command-line instructions.

Briefly, the CO-ED workflow has the following stages. 1: Annotate domains (See "Domain annotation") in a set of "query" protein sequences, and a set of "known enzymes" (See "Selection of known enzymes"). 2: For every domain in the curated set of non-redundant enzymatic/catalytic domains (See "Curation of enzyme Domains for CO-ED" in SI), find its occurrences in each of the proteins in the query set. 3: Draw a node for each domain found in at least one protein in the query set, and label the number of proteins the domain was

found in. 4: For every **combination of two domains** in the curated set of non-redundant enzymatic/catalytic domains, find the occurrences of each combination in the query set and in the "known enzymes" sets. 5: Draw an edge for each combination of domains found in at least one protein in the query set, label it with the number of occurrences, and color the edge based on which set(s) of "known enzymes" contain at least one entry with this combination of domains. Networks were rendered in Cytoscape 3.5 or 3.8[51]

Query sequences—Protein sequences derived from the following genomes were downloaded from NCBI Genbank: *E. coli* K12 (assembly ASM584v2)[52], *P. rubra* DSM 6842 (assembly Prub_2.0)[53], *S. coelicolor* A2(3) (assembly ASM20383v1)[54], *S. tropica* CNB-440 (ASM1642v1)[55], *P. fluorescens* Pf-5 (ASM1226v1)[56].

Domain annotation—Pfam-A domain annotations were either retrieved from Uniprot, as was done for the "known" enzymes and for the all-of-uniprot analysis, or annotated using PfamScan 1.6 (which in turn made use of HMMER 3.3[57]), as was done for the single-genome queries, which gives the CO-ED workflow more flexibility, such as allowing for the analysis of genomes that are not in Uniprot, and the detection of domains that are not included in Pfam-A

Selection of "known" enzymes—Uniprot: All proteins in UniprotKB that **either** are manually annotated as having catalytic activity (i.e. match the query "annotation:(type:"catalytic activity" evidence:manual)") **or** are listed in Uniprot's pathway.txt file (https://www.uniprot.org/docs/pathway). This set is overlaps with, but is not identical to "Swissprot" (the subset of Uniprot marked "reviewed"): It contains some "unreviewed" entries, and does not include Swissprot entries whose catalytic activity was automatically assigned by homology. BRENDA: All entries in the BRENDA database that refer to Uniprot accessions. MIBiG: For each entry in MIBiG, an NCBI nucleotide region is defined. NCBI protein identifiers for all proteins in this region were obtained and mapped to Uniprot accessions. A small number of proteins that could not be mapped to Uniprot were not included in the final set.

Curation of enzyme Domains for CO-ED—The set of domains used to conduct CO-ED is intended to be enzymatic and non-redundant. With "enzymatic" we mean that only catalytic enzyme domains are included, and e.g. purely structural, regulatory, or docking domains are omitted. With "non-redundant" we mean that we wish to avoid "trivial" domain pairs where two crystallographic Pfam domains work together to catalyze a single reaction (such as "Terpene_synth" and "Terpene_synth_C" or "G6PD_N" and "G6PD_C"). This section describes how this set of domains was curated.

A set of Pfam domains was compiled by taking all domains annotated for entries in Uniprot (www.uniprot.org) that are also annotated in MIBiG (mibig.secondarymetabolites.org, all proteins), BRENDA (www.brenda-enzymes.org, all proteins), or in Uniprot's pathway.txt. Non-catalytic domains were removed. For pseudo-catalytic domains that together catalyze one reaction (often detected by performing CO-ED analysis on all proteins in Uniprot and finding domains that co-occur with another a high percentage of the time), the more abundant domain was included. Overlapping domains with similar catalytic functions are

often members of the same Pfam "Clan", causing only the best-matching domain to be annotated by PfamScan, but in some cases where both are annotated in a high proportion of proteins in Uniprot (e.g., because clan assignment has yet not been completed for those domains), only one of the proteins was included in our set. Lastly, many enzymatic domains acting on macromolecules and domains with unknown functions were annotated as such, and the analysis can be run with or without their consideration. Annotation categories are as follows: "m": nucleases, topoisomerases, transposases, helicases, polymerases, proteases, protein kinases and phosphatases, ATP-dependent transporters; "s": glycosyltransferases, glycosylhydrolases (cellulases, amylases, etc.); "e": enzymes in electron transport chains (oxidative phosphorylation, photosynthesis, etc.); "u": domains with unknown function; "y": all other enzymes, but only those transporters that couple transport to a reaction besides ATP hydrolysis; "n": determined to either not be catalytic or to comprise a catalytic domain together with a domain annotated in one of the above categories. For the analyses shown in this manuscript, annotation categories "y", "m", "s" and "e" were considered, totaling 1745 domains. We realize the domain curation process is somewhat subjective, and hence the CO-ED web tool allows for editing the set of domains considered in the submission interface, and the Jupyter notebook allows this by editing the pfamID_to_name_desc_longdesc.tsv file.

Bacteriological culture

Bacterial stocks were obtained from DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH): *P. rubra* DSM 6842 *C. chukchiensis* DSM 22576 and *R. pacifica* DSM 17616 cultured on Marine Agar or Broth 2216 at 30 °C, and *U. pigrum* DSM 19792 and *S. aerolata* DSM 18479 cultured on Reasoner's 2A (R2A) Agar or Broth at 30 °C.

Molecular cloning

Genomic DNA was isolated from overnight liquid cultures using the Zymo Research Quick-DNA Miniprep Kit. *oxz* genes were PCR amplified from this genomic DNA using NEB Q5 High-Fidelity polymerase. Cycling conditions were as follows: initial denaturation temperature of 98 °C for 30 seconds followed by 35 cycles of a denaturing temperature of 98 °C for 10 seconds, an annealing temperature of 68 °C for 30 seconds, and a 72 °C extension for 5 minutes. After cycling, a final extension at 72 °C for 2 minutes was performed. All amplicons were purified by gel extraction. *oxzAB* gene pairs were cloned using the Thermo Fischer ZERO Blunt TOPO kit, and Pr_*oxzA* and Pr_*oxzB* were cloned into pET28a without and with MBP respectively using the NEB Hi-Fi Assembly master mix. All plasmids were verified by Sanger sequencing.

All plasmid sequences used in this study are available at the following URL: https://benchling.com/tderond/f_/TUemteIN-de-rond-et-al-2020/. Primers used are also tabulated in Supplementary Table 1.

Metabolite analysis of heterologously expressed *oxzAB*

pTOPO-*oxzAB* plasmids were transformed into *E. coli* BLR(DE3) and grown up in 50 mL of terrific broth (TB), 10 mL/L glycerol and 50ug/mL kanamycin at 30 °C, induced with 0.1 mM IPTG at an OD of 0.8, and left to shake overnight at 30 °C. The cells were pelleted at

10,000 × g, and the pellets were frozen, lyophilized, extracted for two hours in ethyl acetate. The extracts were filtered through glass pipette cotton filters, evaporated under a stream of nitrogen gas, redissolved in 200 μL 35:65 acetonitrile:water, filtered through a 0.2 μm filter, and analyzed by HPLC-UV-MS as described in the SI.

Bulk heterologous production and isolation of oxazolones

E. coli BLR(DE3) containing pTOPO-Pr_*oxzAB* or pTOPO-Sa_*oxzAB* was grown in 3x 1.5L TB with 50 μg/mL kananmycin and 10 mL/L of glycerol for 3 days at 30 °C. The cultures were centrifuged at 10,000g for 20 minutes, and the pellets were frozen and lyophilized for two days. The dried biomass was crushed, extracted overnight in 100 mL ethyl acetate, filtered, and dried under reduced pressure. The residue was subjected to flash chromatography using hexanes:ethyl acetate on a CombiFlash EZPrep system using a 24g ReadySep gold silica gel column. UV active peaks were then further subjected to preparative reverse phase HPLC (Phenomenex Luna c18 column (100mm x 21.2 mm, 5 μm particle size) using isocratic conditions between 80:20 and 90:10 acetonitrile:water with 0.1% formic acid, and evaporated by rotary evaporation followed by lyophilization, yielding between 1 mg and 9 mg of purified oxazolone, compound characterization data for which can be found in the Supplementary Note.

Methanol adduct formation

2 mg **4** or **9** was dissolved in 1 mL of methanol, 10 mg of K_2CO_3 was added, upon which the reaction with **4** turned bright yellow (the reaction with **9** stayed colorless). The reactions were left to react for 5 minutes at room temperature, filtered, and purified by preparative reverse phase HPLC as described above, under isocratic conditions of 78:22 acetonitrile:water with 0.1% formic acid, compound characterization data for which can be found in the Supplementary Note

Protein purification and enzyme assays

Overnight cultures of *E. coli* BLR(DE3) harboring pET28a-His6-PrOxzA or pET28a-MBP-PrOxzB respectively were diluted into 1.5 L Terrific Broth with 50 μg/mL kanamycin and 10 mL/L of glycerol at 30 °C. When an OD of 0.8 was reached, the flasks were cooled to 18 °C and induced with 50 μM IPTG and left to shake at 18 °C overnight. The cells were pelleted at 4 °C for 15–20 minutes at 15,000 × g and resuspended in 40mL cold 50mM Tris, 200 mM NaCl and 10% glycerol at pH 7 (“lysis buffer”), and sonicated at 50% amplitude, 15 seconds on, 15 seconds off for 5 minutes. The proteins were purified using loose resin in 50 mL centrifuge tubes. For His6-PrOxzA, 2.5 mL Nickel-IDA resin was used, the resin washed twice with 40mL lysis buffer + 50 mM imidazole, eluted with 5 mL lysis buffer + 500 mM imidazole, and concentrated using a 15 kDa MWCO filter. For MBP-PrOxzB, 2.5 mL NEB amylose resin was used, the resin was washed twice with 40 mL lysis buffer, eluted in 5mL lysis buffer + 10 mM maltose, and concentrated using a 50 kDa MWCO filter. MBP-PrOxzB is visibly yellow, suggesting it binds a flavin cofactor, as expected from nitroreductase-family enzymes.

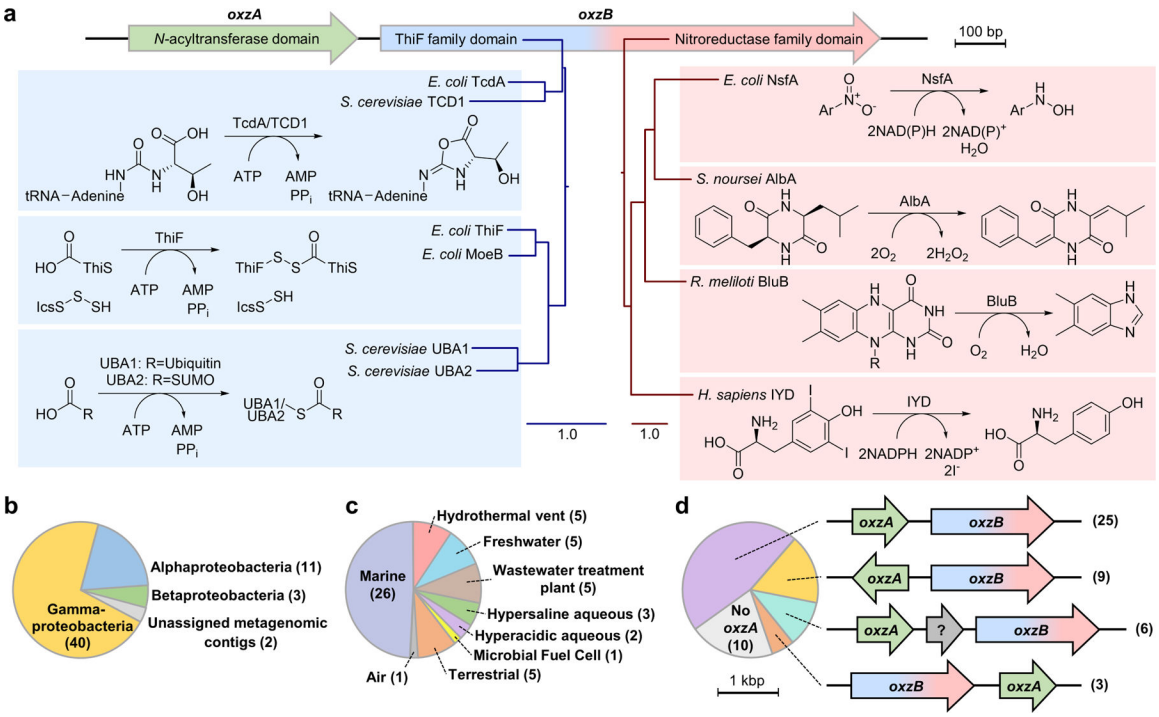
In vitro assays of His6-OxzA and MBP-OxzB were conducted in 200uL 50 mM potassium phosphate pH 7, 200 mM NaCl, 10% DMSO, with reactants at the following final

concentrations: 150 μM ATP, 150 μM decanoyl-CoA, 100 μM L-tyrosine, and 100 μM N-decanoyl-L-tyrosine (synthesis described in SI). After 1 hour, the reactions were quenched with 100 μL acetonitrile, filtered through 0.2 μM filters, and analyzed as described by LC-MS as described in the SI

Analysis of oxazolone production in *P. rubra* and *C. chukchiensis*

Cells were grown as a lawn on Marine Agar 2216 with or without a drop of antibiotic stock (see Supplementary Table 2 for amounts) in the middle of the plate. After 2 days of growth, biomass adjacent to the zone of inhibition was harvested by scraping the plate. An effort was made to harvest a roughly equivalent amount of biomass from each plate. The biomass was lyophilized, extracted with ethyl acetate for 2 hours, filtered through glass pipette cotton filters, evaporated under a stream of nitrogen gas, redissolved in 200 μL 35:65 acetonitrile:water, filtered through a 0.2 μm filter, and analyzed by HPLC-UV-MS as described in the SI.

Extended Data



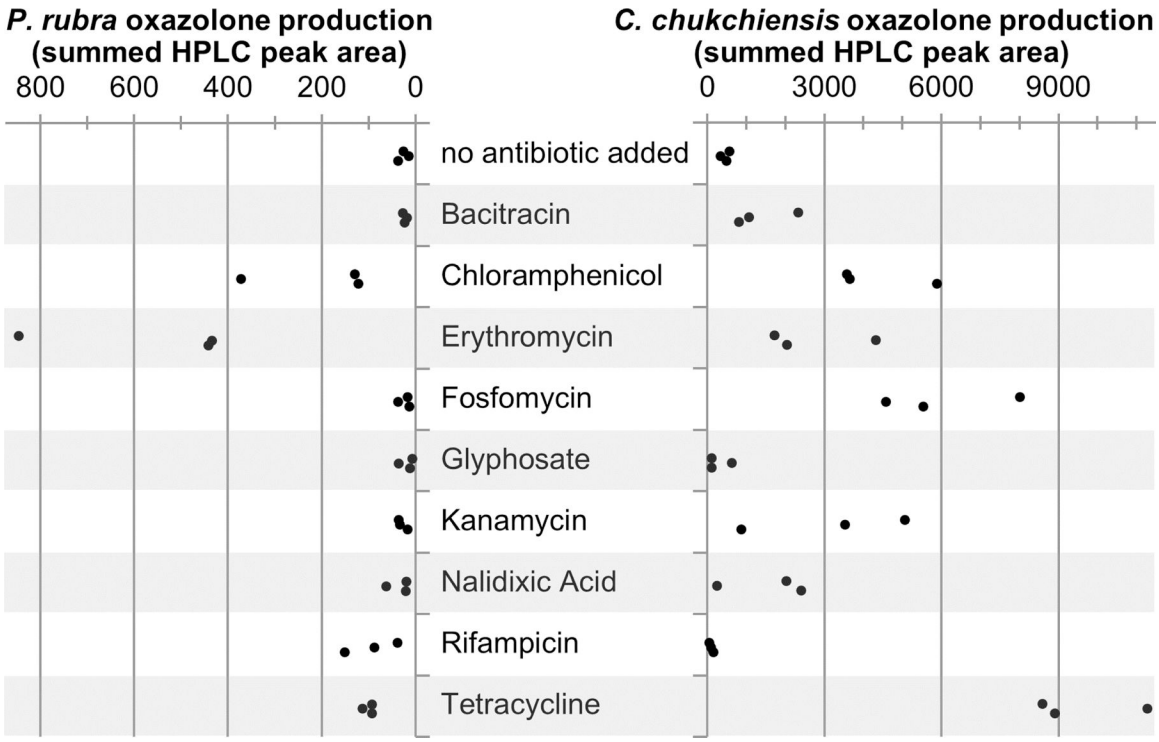
Extended Data Fig. 1. Survey of transformations known to be catalyzed by the ThiF and nitroreductase domains, and properties of organisms harboring a ThiF-nitroreductase di-domain enzyme identified by CO-ED.

a: OxzB exhibits homology to both the ThiF and nitroreductase enzyme families. Shown is a selection of characterized members of these enzyme families along with the transformations they catalyze, along with midpoint-rooted gene trees showing the relationships between each other and OxzB. ThiF family enzymes are known to catalyze ATP-dependent carboxylate activating reactions, while nitroreductase family enzymes catalyze a variety of redox reactions. Inferred phylogenies were generated from protein sequences using neighbor

joining on the MAFFT web server[58] and midpoint-rooted. Scale bars designate 1 substitution per site.

b,c: Phylogenetic distribution (**b**) and habitat (**c**) of the host organisms harboring genes encoding OxzB proteins represented in the Uniprot database. Organisms with unknown habitats are not included.

d: Genomic context of *oxzB* genes as determined using the Enzyme Function Initiative Genome Neighborhood Tool[42]. Most *oxzB* homologs are accompanied by *oxzA*, which codes for an *N*-acyltransferase. The arrow labeled "?" represents genes unrelated to *oxzA* or *oxzB*. Organisms with unknown *oxzB* genomic context (e.g., at the edge of a contig) are not included.



Extended Data Fig. 2. Induction of oxazolone production in *P. rubra* and *C. chukchiensis* by various antibiotics. Bacteria were grown as a lawn on Marine Agar 2216 with a drop of antibiotic.

A consistent amount of biomass adjacent to the zone of inhibition was harvested, extracted and analyzed by HPLC. Dots indicate summed peak areas between wavelengths of 300 nm and 400 nm. Three biological replicates were analyzed for each condition.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank our UCSD colleagues Jie Li, Vikram Shende, Timothy Fallon, and Brendan Duggan for helpful discussions. This work was supported by National Institutes of Health awards F32GM129960 to TdR and

R01GM085770 to BSM, as well as the American Society for Pharmacognosy Undergraduate Research Award and the UC San Diego “Eureka” Undergraduate Research Scholarship to JEA.

Data Availability

All data generated during this study are included in this article and its supplementary information files.

- The CO-ED networks used to generate Figure 2 and the several supplemental figures showing CO-ED networks, are available as Supplementary Data 1
- The data used to generate Extended Data Figs. 1 and 2 are available as Source Data
- The curated set of enzyme domains for CO-ED can be found as the node table of the all_of_uniprot network in Supplementary Data 1, as well as at https://github.com/tderond/CO-ED/blob/master/pfamID_to_name_desc_longdesc.tsv. This same set of domains is also the default setting for the web tool.
- NMR spectra of the newly reported structures are available in the Supplementary Note
- Tandem MS spectra of the newly reported structures are available in the Supplementary Note, and have also been deposited to the GNPS spectral library at the URLs shown in the Supplementary Note.

The following bioinformatic databases were employed in this study: BRENDA (www.brenda-enzymes.org), MIBiG (mibig.secondarymetabolites.org), Uniprot (www.uniprot.org) and Pfam-A (pfam.xfam.org)

References

1. Gu C, Kim GB, Kim WJ, Kim HU & Lee SY Current status and applications of genome-scale metabolic models. Genome Biol 20, 121 (2019). [PubMed: 31196170]
2. Bornscheuer UT The fourth wave of biocatalysis is approaching. Philos Transact A Math Phys Eng Sci 376, (2018).
3. Gerlt JA et al. The enzyme function initiative. Biochemistry 50, 9950–9962 (2011). [PubMed: 21999478]
4. Hanson AD, Pribat A, Waller JC & de Crécy-Lagard V Unknown proteins and orphan enzymes: the missing half of the engineering parts list – and how to find it. Biochem J 425, 1–11 (2009). [PubMed: 20001958]
5. Scott TA & Piel J The hidden enzymology of bacterial natural product biosynthesis. Nat. Rev. Chem 3, 404–425 (2019). [PubMed: 32232178]
6. Ellens KW et al. Confronting the catalytic dark matter encoded by sequenced genomes. Nucleic Acids Res 45, 11495–11514 (2017). [PubMed: 29059321]
7. Medema MH, de Rond T & Moore BS Mining genomes to illuminate the specialized chemistry of life. Nat Rev Genet (2021) In press.
8. Michael AJ Evolution of biosynthetic diversity. Biochem J 474, 2277–2299 (2017). [PubMed: 28655863]
9. Hagel JM & Facchini PJ Tying the knot: occurrence and possible significance of gene fusions in plant metabolism and beyond. J Exp Bot 68, 4029–4043 (2017). [PubMed: 28521055]
10. Bashton M & Chothia C The generation of new protein functions by the combination of domains. Structure 15, 85–99 (2007). [PubMed: 17223535]

11. Weissman KJ The structural biology of biosynthetic megaenzymes. Nat Chem Biol 11, 660–670 (2015). [PubMed: 26284673]
12. Winzer T et al. Plant science. Morphinan biosynthesis in opium poppy requires a P450-oxidoreductase fusion protein. Science 349, 309–312 (2015). [PubMed: 26113639]
13. Ng TL, Rohac R, Mitchell AJ, Boal AK & Balskus EP An N-nitrosating metalloenzyme constructs the pharmacophore of streptozotocin. Nature 566, 94–99 (2019). [PubMed: 30728519]
14. El-Gebali S et al. The Pfam protein families database in 2019. Nucleic Acids Res 47, D427–D432 (2019). [PubMed: 30357350]
15. Jeske L, Placzek S, Schomburg I, Chang A & Schomburg D BRENDA in 2019: a European ELIXIR core data resource. Nucleic Acids Res 47, D542–D549 (2019). [PubMed: 30395242]
16. Kautsar SA et al. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. Nucleic Acids Res 48, D454–D458 (2020). [PubMed: 31612915]
17. The UniProt Consortium. UniProt: the universal protein knowledgebase. Nucleic Acids Res 46, 2699 (2018). [PubMed: 29425356]
18. Ishino F, Mitsui K, Tamaki S & Matsuhashi M Dual enzyme activities of cell wall peptidoglycan synthesis, peptidoglycan transglycosylase and penicillin-sensitive transpeptidase, in purified preparations of Escherichia coli penicillin-binding protein 1A. Biochem Biophys Res Commun 97, 287–293 (1980). [PubMed: 7006606]
19. Goodlove PE, Cunningham PR, Parker J & Clark DP Cloning and sequence analysis of the fermentative alcohol-dehydrogenase-encoding gene of Escherichia coli. Gene 85, 209–214 (1989). [PubMed: 2695398]
20. Williams GJ, Breazeale SD, Raetz CRH & Naismith JH Structure and function of both domains of ArnA, a dual function decarboxylase and a formyltransferase, involved in 4-amino-4-deoxy-L-arabinose biosynthesis. J Biol Chem 280, 23000–23008 (2005). [PubMed: 15809294]
21. Rusnak F, Sakaitani M, Drueckhammer D, Reichert J & Walsh CT Biosynthesis of the Escherichia coli siderophore enterobactin: sequence of the entF gene, expression and purification of EntF, and analysis of covalent phosphopantetheine. Biochemistry 30, 2916–2927 (1991). [PubMed: 1826089]
22. De Rond T et al. Oxidative cyclization of prodigiosin by an alkylglycerol monooxygenase-like enzyme. Nat Chem Biol 13, 1155–1157 (2017). [PubMed: 28892091]
23. Agarwal V et al. Biosynthesis of polybrominated aromatic organic compounds by marine bacteria. Nat Chem Biol 10, 640–647 (2014). [PubMed: 24974229]
24. Ross AC, Gulland LES, Dorrestein PC & Moore BS Targeted capture and heterologous expression of the Pseudoalteromonas alterochromide gene cluster in Escherichia coli represents a promising natural product exploratory platform. ACS Synth Biol 4, 414–420 (2015). [PubMed: 25140825]
25. Herz S, Eberhardt S & Bacher A Biosynthesis of riboflavin in plants. The ribA gene of Arabidopsis thaliana specifies a bifunctional GTP cyclohydrolase II/3,4-dihydroxy-2-butanone 4-phosphate synthase. Phytochemistry 53, 723–731 (2000). [PubMed: 10783978]
26. Schulman BA & Harper JW Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. Nat Rev Mol Cell Biol 10, 319–331 (2009). [PubMed: 19352404]
27. Miyauchi K, Kimura S & Suzuki T A cyclic form of N6-threonylcarbamoyladenosine as a widely distributed tRNA hypermodification. Nat Chem Biol 9, 105–111 (2013). [PubMed: 23242255]
28. Regni CA et al. How the MccB bacterial ancestor of ubiquitin E1 initiates biosynthesis of the microcin C7 antibiotic. EMBO J 28, 1953–1964 (2009). [PubMed: 19494832]
29. Xi J, Ge Y, Kinsland C, McLafferty FW & Begley TP Biosynthesis of the thiazole moiety of thiamin in Escherichia coli: identification of an acyldisulfide-linked protein--protein conjugate that is functionally analogous to the ubiquitin/E1 complex. Proc Natl Acad Sci U S A 98, 8513–8518 (2001). [PubMed: 11438688]
30. Godert AM, Jin M, McLafferty FW & Begley TP Biosynthesis of the thioquinolobactin siderophore: an interesting variation on sulfur transfer. J Bacteriol 189, 2941–2944 (2007). [PubMed: 17209031]

31. Akiva E, Copp JN, Tokuriki N & Babbitt PC Evolutionary and molecular foundations of multiple contemporary functions of the nitroreductase superfamily. Proc Natl Acad Sci U S A 114, E9549–E9558 (2017). [PubMed: 29078300]
32. Mondal S, Raja K, Schweizer U & Mugesh G Chemistry and biology in the biosynthesis and action of thyroid hormones. Angew Chem Int Ed Engl 55, 7606–7630 (2016). [PubMed: 27226395]
33. Schneider TL, Shen B & Walsh CT Oxidase domains in epothilone and bleomycin biosynthesis: thiazoline to thiazole oxidation during chain elongation. Biochemistry 42, 9722–9730 (2003). [PubMed: 12911314]
34. Taga ME, Larsen NA, Howard-Jones AR, Walsh CT & Walker GC BluB cannibalizes flavin to form the lower ligand of vitamin B12. Nature 446, 449–453 (2007). [PubMed: 17377583]
35. Bashiri G et al. A revised biosynthetic pathway for the cofactor F420 in prokaryotes. Nat Commun 10, 1558 (2019). [PubMed: 30952857]
36. Gondry M et al. Cyclic dipeptide oxidase from Streptomyces noursei. Isolation, purification and partial characterization of a novel, amino acyl alpha,beta-dehydrogenase. Eur J Biochem 268, 1712–1721 (2001). [PubMed: 11248691]
37. Zenno S, Saigo K, Kanoh H & Inouye S Identification of the gene encoding the major NAD(P)H-flavin oxidoreductase of the bioluminescent bacterium Vibrio fischeri ATCC 7744. J Bacteriol 176, 3536–3543 (1994). [PubMed: 8206830]
38. Guella G, N'Diaye I, Fofana M & Mancini I Isolation, synthesis and photochemical properties of almazolone, a new indole alkaloid from a red alga of Senegal. Tetrahedron 62, 1165–1170 (2006).
39. Seyedsayamdost MR High-throughput platform for the discovery of elicitors of silent bacterial gene clusters. Proc Natl Acad Sci U S A 111, 7266–7271 (2014). [PubMed: 24808135]
40. Walsh C & Wencewicz T Antibiotics: challenges, mechanisms, opportunities (American Society of Microbiology, 2016). doi:10.1128/9781555819316
41. Hon J et al. EnzymeMiner: automated mining of soluble enzymes with diverse structures, catalytic properties and stabilities. Nucleic Acids Res 48, W104–W109 (2020). [PubMed: 32392342]
42. Gerlt JA Genomic enzymology: web tools for leveraging protein family sequence-function space and genome context to discover novel functions. Biochemistry 56, 4293–4308 (2017). [PubMed: 28826221]
43. Wuchty S & Almaas E Evolutionary cores of domain co-occurrence networks. BMC Evol Biol 5, 24 (2005). [PubMed: 15788102]
44. Barrera A, Alastruey-Izquierdo A, Martín MJ, Cuesta I & Vizcaíno JA Analysis of the protein domain and domain architecture content in fungi and its application in the search of new antifungal targets. PLoS Comput Biol 10, e1003733 (2014). [PubMed: 25033262]
45. Suhre K Inference of gene function based on gene fusion events: the rosetta-stone method. Methods Mol Biol 396, 31–41 (2007). [PubMed: 18025684]
46. Promponas VJ, Ouzounis CA & Iliopoulos I Experimental evidence validating the computational inference of functional associations from gene fusion events: a critical survey. Brief Bioinformatics 15, 443–454 (2014). [PubMed: 23220349]
47. Alborzi SZ, Devignes M-D & Ritchie DW ECDomainMiner: discovering hidden associations between enzyme commission numbers and Pfam domains. BMC Bioinformatics 18, 107 (2017). [PubMed: 28193156]
48. De Castro PP, Carpanez AG & Amarante GW Azlactone Reaction Developments. Chem. Eur. J 22, 10294–10318 (2016). [PubMed: 27245128]
49. Kenney GE et al. The biosynthesis of methanobactin. Science 359, 1411–1416 (2018). [PubMed: 29567715]
50. Haft DH, Paulsen IT, Ward N & Selengut JD Exopolysaccharide-associated protein sorting in environmental organisms: the PEP-CTERM/EpsH system. Application of a novel phylogenetic profiling heuristic. BMC Biol 4, 29 (2006). [PubMed: 16930487]
51. Shannon P et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13, 2498–2504 (2003). [PubMed: 14597658]
52. Blattner FR et al. The complete genome sequence of *Escherichia coli* K-12. Science 277, 1453–1462 (1997). [PubMed: 9278503]

53. Xie B-B et al. Genome sequence of the cycloprodigiosin-producing bacterial strain Pseudoalteromonas rubra ATCC 29570(T). J Bacteriol 194, 1637–1638 (2012). [PubMed: 22374963]

54. Bentley SD et al. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). Nature 417, 141–147 (2002). [PubMed: 12000953]

55. Udwary DW et al. Genome sequencing reveals complex secondary metabolome in the marine actinomycete Salinispora tropica. Proc Natl Acad Sci U S A 104, 10376–10381 (2007). [PubMed: 17563368]

56. Paulsen IT et al. Complete genome sequence of the plant commensal Pseudomonas fluorescens Pf-5. Nat Biotechnol 23, 873–878 (2005). [PubMed: 15980861]

57. Eddy SR Accelerated profile HMM searches. PLoS Comput Biol 7, e1002195 (2011). [PubMed: 22039361]

58. Katoh K, Rozewicki J & Yamada KD MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinformatics bbx108 (2017). doi:10.1093/bib/bbx108

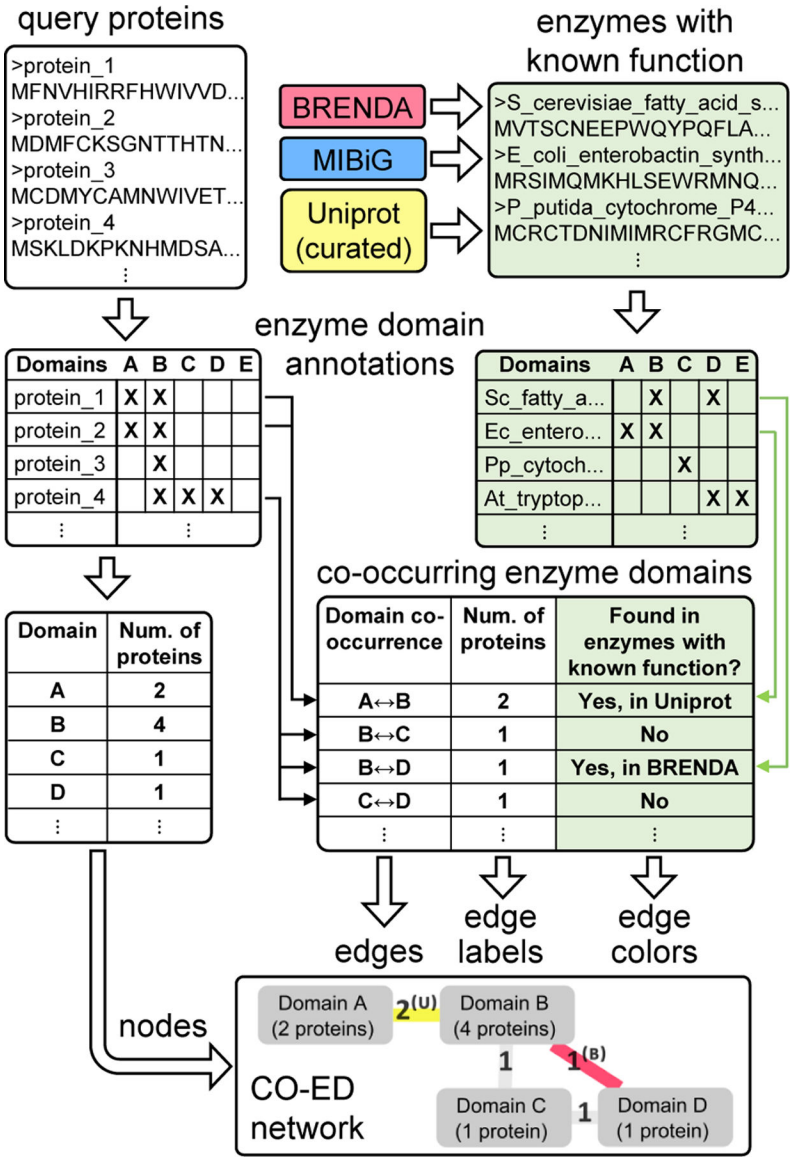


Figure 1.

Outline of the CO-ED workflow.

A CO-ED network consists of nodes representing a curated set of enzymatic domains present in a query set of proteins, and edges connecting two domains if they co-occur in the same protein. Edges are colored based on whether any known enzymes contain the pair of domains the edge represents. Known enzymes were compiled from BRENDA[15] MIBiG[16], and a curated set of entries from Uniprot[17], the edge color indicating the originating database(s) for the annotation.

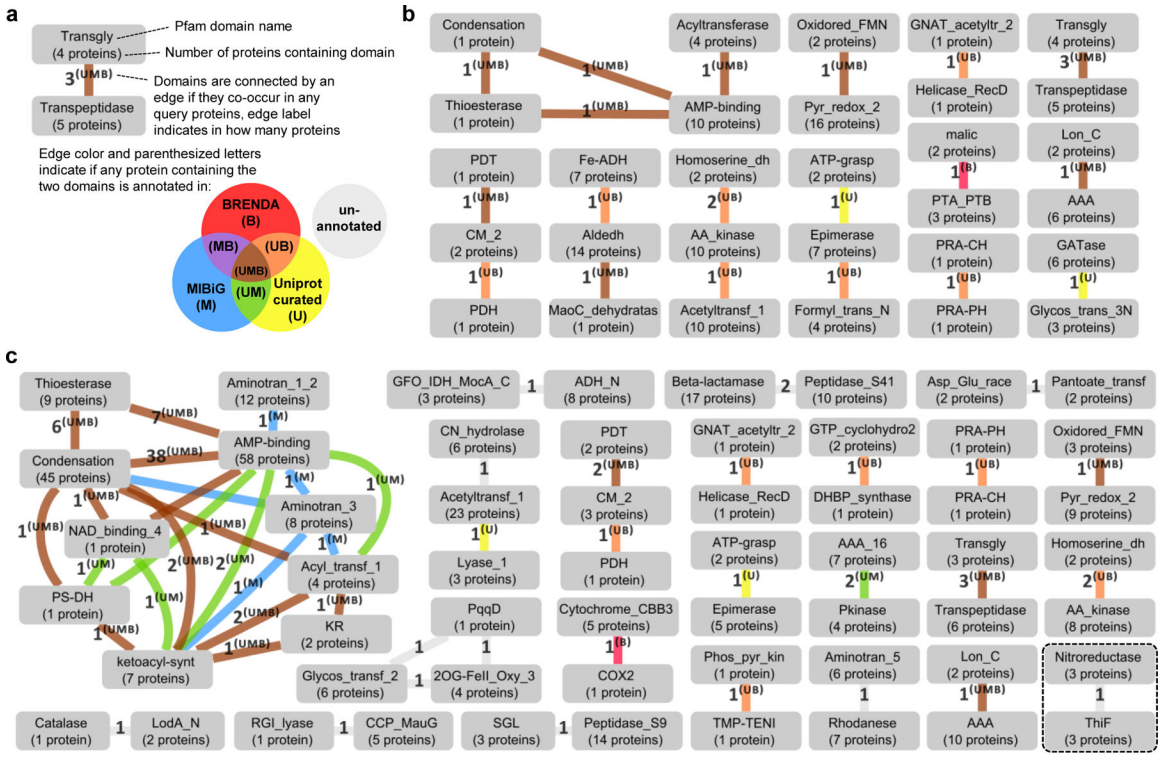


Figure 2.

CO-ED networks can help guide enzyme discovery.

a: Legend for the CO-ED networks shown in **b** and **c**.

b: CO-ED network for *E. coli* K12. All edges are colored, suggesting *E. coli* is of low priority for CO-ED–guided genome mining.

c: CO-ED network for *Pseudoalteromonas rubra* DSM 6842. Some overlapping edge labels are omitted for clarity. Many edges are unannotated (gray), suggesting there is much potential for exploring multidomain enzymes for novel biosynthetic capacity in this organism. The ThiF-nitroreductase domain pair investigated further in this study is outlined. Only nodes with edges are shown.

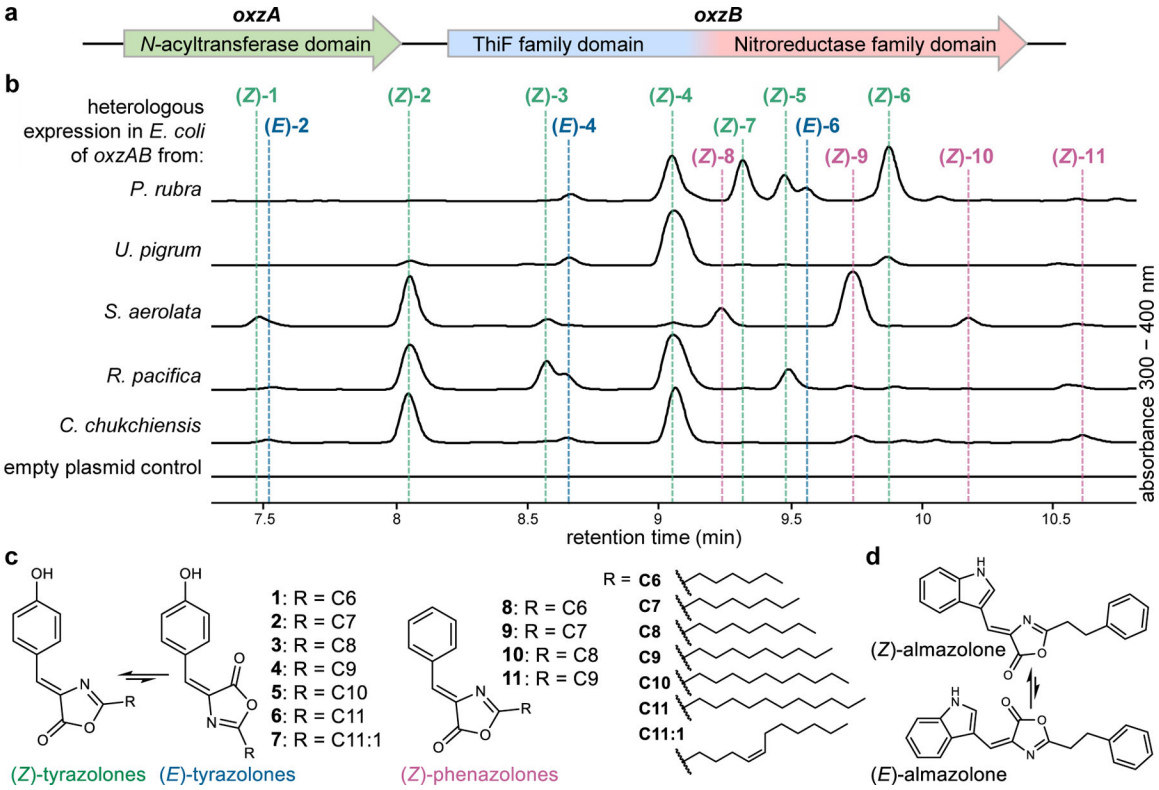


Figure 3.

Heterologous expression of *oxzAB* results in the production of a series of oxazolones.

a: Organization of *oxzAB* in the genomes of the five proteobacteria for which these genes were tested heterologously.

b: Reversed phase HPLC profiles of extracts of *E. coli* expressing *oxzAB* from five proteobacteria. An absorbance range of 300–400 nm was chosen to visualize all oxazolones reported in this manuscript (under our chromatography conditions (H_2O/acetonitrile with 0.1% formic acid), the λ_{max} of the (*E*)-tyrazolones is 360 nm, that of the (*Z*)-tyrazolones is 358 nm, and that of the (*Z*)-phenazolones is 330 nm).

c: Structures of the tyrazolones and phenazolones, which are found as a series with varying alkyl tails. Tyrazolones **4**, **5**, **6** and **7**, and phenazolone **9** were characterized by NMR, while all others were inferred by exact mass and MS/MS fragmentation patterns. The tyrazolones exist as (*E*)/(*Z*) isomers in equilibrium.

d: Almazolone, an oxazolone natural product previous isolated from an alga, also occurs as an (*E*)/(*Z*) equilibrium.

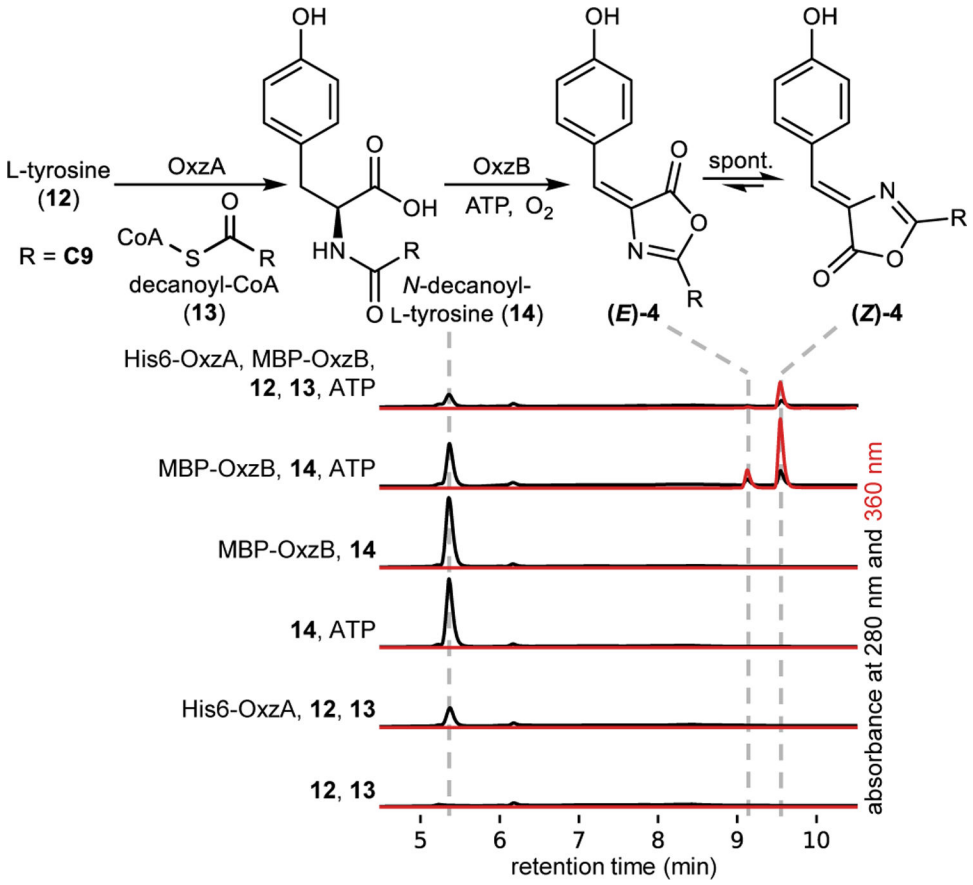


Figure 4.

In vitro characterization of oxazolone biosynthesis by *P. rubra* OxzA and OxzB.

Black HPLC chromatograms show absorbance at 280 nm (*N*-decanoyl-L-tyrosine); red chromatograms show absorbance at 360 nm (tyrazolones) and are scaled down 4.5-fold in relation to the 280 nm signal.