Cryptic indole hydroxylation by a non-canonical terpenoid cyclase parallels bacterial xenobiotic detoxification

Susann Kugel¹,*  Martin Baunach¹,*, Philipp Baer², Mie Ishida-Ito¹, Srividhya Sundaram¹, Zhongli Xu¹, Michael Groll² & Christian Hertweck¹,³

Terpenoid natural products comprise a wide range of molecular architectures that typically result from C–C bond formations catalysed by classical type I/II terpene cyclases. However, the molecular diversity of biologically active terpenoids is substantially increased by fully unrelated, non-canonical terpenoid cyclases. Their evolutionary origin has remained enigmatic. Here we report the in vitro reconstitution of an unusual flavin-dependent bacterial indoloterpenoid cyclase, XiaF, together with a designated flavoenzyme-reductase (XiaP) that mediates a key step in xiamycin biosynthesis. The crystal structure of XiaF with bound FADH₂ (at 2.4 Å resolution) and phylogenetic analyses reveal that XiaF is, surprisingly, most closely related to xenobiotic-degrading enzymes. Biotransformation assays show that XiaF is a designated indole hydroxylase that can be used for the production of indigo and indirubin. We unveil a cryptic hydroxylation step that sets the basis for terpenoid cyclization and suggest that the cyclase has evolved from xenobiotics detoxification enzymes.
It is a generally accepted model that the vast structural diversity of secondary metabolites has developed to address specific needs of the producing organism. Such natural products may, for example, serve as carrier of information or as chemical armour; they may be used to attract partners, to secure the niche, to protect against various types of stresses and to shape communities. Natural products are also a valuable source of molecular tools and therapeutics that have been pre-optimized for particular biological targets over millions of years. However, considering the impressive range of structures and functions, one of the remaining big questions is: how has this structural diversity evolved?

Research in the past decades has led to an impressive body of knowledge on biosynthetic pathways, their genetic basis and the biocatalysts involved. Typically, there are various levels of diversification in each biosynthetic pathway, such as the choice and assembly of basic building blocks, the processing of various intermediates and various ways to tailor the individual scaffolds. Typically, enzymes involved in secondary metabolite biosynthesis are rather specific. They are often fine-tuned to direct reactive intermediates into specific reaction channels, and small changes may lead to dramatic changes in the metabolic profile.

This principle is well exemplified in terpene biosynthesis, where designated cyclases orchestrate a series of cation–olefin cyclizations and Wagner–Meerwein rearrangements of intermediates to produce diverse carbocyclic systems. Beyond these classical type I/II terpene cyclases, there is a surprisingly broad range of non-canonical terpenoid cyclases.

Various atypical terpenoid cyclizations and diversifications are involved in the pathway leading to the xiamycin family of indolosequiterpenes (IST) (Fig. 1). These ‘plant-like’ terpenoids are actually produced by bacterial endophytes (Streptomyces spp.). Endowed with a range of antibiotic and antimicrobial activities, xiamycin and its congeners may aid in protecting the plant host from infections.

To elucidate the molecular basis of IST diversification, we investigated the xiamycin (xia) biosynthetic gene cluster in the endophyte. We found that xiamycin (3), the major IST component of the complex, is diversified by means of rare flavoenzyme-catalysed modifications to yield the cycloether dihydrocarbazole (5) and various N-C- and N,N-fused xiamycin dimers (such as 6a/6b), and sulfonyl-bridged xiamycin dimers (such as 7) (Fig. 1). At the heart of the biosynthetic pathway, however, we discovered an unusual cyclization cascade involving the non-canonical terpenoid cyclases XiaE and XiaF.

The first cyclization yields the decalin system of indosene (2), which is further cyclized into the spiro compound sespenene (4) or the carbazole ring system of xiamycin (Fig. 1).

This non-canonical terpenoid cyclization is mechanistically intriguing and has been the subject of several biosynthetic and synthetic studies. Through bioinformatics and mutational analysis, we found that XiaF is a flavin-dependent monoxygenase that catalyses the conversion of 2 to the carbazole scaffold of 3 (ref. 22). In an independent study, an orthologue from Streptomyces sp. SCSIO 02999 (XiaL, according to these authors’ nomenclature) was found to catalyse this reaction in vitro, albeit with unsatisfactory efficacy. Yet, the detailed enzymatic mechanism of this rare cyclization reaction and the evolutionary origin of this extraordinary monoxygenase/cyclase have remained elusive.

Here we report the in vitro reconstitution of XiaF and a designated flavoenzyme-reductase, which jointly mediate a cryptic indole hydroxylation step that has been proposed to be a key step in non-canonical terpenoid cyclization. By phylogenetic, functional and structural analyses, we also demonstrate the relatedness of this rare terpene cyclase with broad-specific xenobiotic-metabolizing enzymes.

**Results**

**Reconstitution of two-component terpenoid cyclization.** To gain insight into the biochemical function of the indosespene cyclase XiaF, we produced XiaF in *Escherichia coli*. Initially, the in vitro characterization of recombinant XiaF was hampered by its poor performance. Substantial biotransformation of 2 (100 μM) was only achieved after extended incubation times (12–24 h) with high concentrations of enzyme (25 μM) and cofactors FAD and NADH (1 and 2 mM). Similar observations were reported for enzyme assays using the XIAF orthologue Xial from *Streptomyces* sp. SCSIO 02999 (refs 21,26). However, even under these optimized non-physiological reaction conditions, the catalytic performance of XiaF was unreliable.

To overcome these shortcomings, we investigated the effect of added NAD(P)H:flavin oxidoreductase (flavin reductase) for cofactor recycling. In fact, terpene cyclization was efficiently reconstituted in vitro using Fre, the flavin reductase from *E. coli* (Supplementary Fig. 2). This finding strongly suggested that XiaF naturally acts in a two-component system. It is conceivable that XiaF variants are orphan oxidogenases that are supported by unrelated flavin reductases derived from the central metabolism. Alternatively, XiaF variants may have a designated reductase partner. Thus, we revisited the xia gene locus and related gene clusters of various other actinomycetes.

Genome mining revealed a cryptic gene cluster in *Amycolatopsis nigrescens* CSC17Ta-90 with high similarity to the xia gene locus. In contrast to the xia gene cluster in the endophytic *Streptomyces* sp. (HKI0576), a putative flavin reductase gene is located next to the xiaF orthologue (Fig. 2a). On the basis of this sequence, we screened the genome sequence of *Streptomyces* sp. (HKI0576) and discovered a homologous flavin reductase gene (named xiaP) located in the flanking region of the xia gene cluster (Fig. 2a; the xiaP orthologue from *A. nigrescens* CSC17Ta-90 was named xiaP*). The deduced gene products of both xiaP variants share homology with other flavin reductase partners of XIAF homologues (Supplementary Fig. 5).

To test whether XiaP functions as flavin reductase partner of XiaF, we cloned and heterologously expressed xiaP from *Streptomyces* sp. (HKI0576) in *E. coli*. Flavin reduction activity assays of XiaF with NAD(P)H and FAD, FMN or riboflavin confirmed its proposed function. They revealed a dependency of XiaP on NADH, whereas NADPH seems to be no suitable cofactor. In contrast, XiaP seems to have no marked preference for any flavin substrate tested (Fig. 3c; for a detailed kinetic analysis of XiaP, see the Supplementary Note 2). When substituting Fre by XiaP in the XIAF enzyme assays, we found that XiaP substantially increases the performance of XiaF (Fig. 3b,1). Since the corresponding genes are found in the same gene locus, XiaP likely acts as the natural flavin reductase for XiaF, although its function may be substituted by other flavin reductases.

Using this robust enzyme assay, we set out to analyse the course of the reaction. In the current biosynthetic model, an oxidative cyclization sequence is proposed in which hydroxylation of 2 is presumed to trigger nucleophilic attack of the exo-methylene group followed by dehydration to yield 3 (refs 20–22). Furthermore, the immediate product of the XIAF-catalysed reaction has been proposed to be a dihydrocarbazole, which most likely undergoes spontaneous oxidation to give the carbazole 3 (ref. 21). By liquid chromatography high-resolution mass spectrometry (LC-HRMS) monitoring of the XIAF assay, we found evidence for the transient intermediate (denoted with an asterisk, Supplementary Fig. 2b,1;
Phylogenetic relationship and genetic context of XiaF. To gain insight into the possible evolutionary origin of XiaF, we scrutinized related flavoenzymes, their catalytic function and biological role in silico. Through homology searches (BLAST), we compiled a diverse set of 28 XiaF homologues (Fig. 2b), most of which have been classified as group D flavin-dependent monooxygenases (FMOs). Group D FMOs are monooxygenases with an acetyl-CoA dehydrogenase (ACAD)-fold that receive reduced FAD/FMN from a NAD(P)H-dependent flavin reductase. Various group D FMOs have been shown to catalyze aromatic hydroxylation and N-hydroxylation reactions. Unexpectedly, apart from few secondary metabolite oxygenases, XiaF homologues identified in the homology search are involved in the catabolism of xenobiotics.

To investigate the phylogenetic position of XiaF, we used the neighbour-joining algorithm to construct a phylogenetic tree of all sequences from evolutionary distance data (Fig. 2b). XiaF forms a clade together with its orthologues from S. sp. SCSIO O2999 (XiaI) and A. nigriscens CSCP17Ta-90 (XiaF* #4; not yet characterized orthologues of XiaF were named XiaF and consecutively numbered from 1 to 6), as well as with five orthologues from other putative xia gene clusters from various actinomycetes.

Fig. 3b), showing the expected exact mass (m/z [M-H]− 364.1928 Da, calculated 364.1918 Da) (Supplementary Fig. 6). However, the proposed hydroxylated intermediate could not be detected in vitro.

Figure 1 | Model of bacterial IST biosynthesis. The biosynthetic scheme shows selected members of the xiamycin family of IST, their biogenetic relationship and their proposed biosynthetic pathway. IST biosynthesis comprises various levels of terpenoid diversification. An epoxidized linear IST-precursor has been proposed to be cyclized to yield preindosespene (1)21,22,26, which is subsequently oxygenated to the carboxylic acid indosespene (2)22,26. 2 has been proposed to be further cyclized into the spiro compound sespenine (4) or the carbazole ring system of xiamycin (3)22,23,26, of which the latter has been proposed to be further diversified to yield the cycloether oxiamycin (5)22,23, and various N,C- and N,N-fused xiamycin dimers (such as 6a/6b)18,22,23, and sulfonyl-bridged xiamycin dimers (such as 7)22. Genes encoding enzymes that are involved in the core biosynthesis of the IST scaffold of preindosespene (1) are coloured in blue (xiaEKLMM), genes for enzymes that tailor basic scaffolds to more complex ISTs are coloured in red (xiaFGHIJ) and genes with no proposed function are coloured in grey (xiaGHIJ). Atoms, bonds or functional groups that are added/are newly formed in biosynthetic reactions are highlighted in the same colour as the gene(s) encoding the gene product(s) that have been proposed to mediate these reactions18,19,22,26,46. The length of the scale bar represents a kilobase.
Figure 2 | Comparison of xia gene loci, and phylogeny of XiaF and its homologues. (a) Comparison of the xia gene cluster of Streptomyces sp. (HKI0576) with the xia gene cluster of Streptomyces sp. SCSIO 02999 and a putative xia gene locus of Amycolatopsis nigrescens CSC17Ta-90. Genes for the two-component system XiaF/XiaP and orthologues are framed in red. The 'detox gene cassette' xiaFGHIJ, which encodes homologues of xenobiotic-metabolizing enzymes that are mostly involved in IST tailoring reactions is framed in brown. The length of the scale bar represents a kilobase. (b) Phylogenetic tree of XiaF and related flavoenzymes. Enzymes that have been reported to catalyse indigo formation are coloured in blue; putative XiaF orthologues of other putative xia gene clusters that yet have not been characterized are coloured in grey. The evolutionary history was inferred using the neighbour-joining method. The oxygenase component HpaB of the 4-hydroxyphenylacetate 3-monooxygenase from Thermus thermophilus HB8 (ref. 90) was used as an outgroup. The length of the scale bar represents an evolutionary distance of 0.1 amino acid differences per site. (For further details, see the Supplementary Note 3 and Supplementary Figs 8–11.) (c) Structural comparison of XiaF (PDB ID: 5LVW), the 4-hydroxyphenylacetate 3-hydroxylase C2-HpaH (PDB ID: 2JBT)48, the secosteroid hydroxylase HsaA from Mycobacterium tuberculosis (PDB ID: 3AFE)32, the dibenzothiophene monooxygenase DszC from Rhodococcus erythropolis D-1 (PDB ID: 3X0Y)52 and the N-oxygenases KijD3 involved in D-kijanose biosynthesis (PDB ID: 4KCF)53. The structural alignment was calculated by the Dali server and shows the superposition of the different monomers. (For a more detailed structural comparison, see Supplementary Fig. 7.)
in the degradation of phenolic compounds. The deduced gene product of *xiaJ* is a cytochrome P450 monooxygenase (CYP450), which is closely related to hydroxylases implicated in bacterial detoxification processes. Except for *XiaG*, all enzymes encoded in this subcluster are proposed to be involved in IST tailoring reactions (Fig. 1).

**XiaF catalyses a cryptic indole hydroxylation.** To test XiaF's ability to detoxify typical xenobiotics, we added various aromatic compounds to the *in vitro* enzyme assay. We selected reported substrates of the XiaF homologues, 4-nitrotoluene, 4-hydroxyphenylacetate, 2-naphthoic acid, 3-hydroxybiphenyl, indole 3-acetic acid and indole. However, from this range of aromatic

---

**Figure 3 | In vitro reconstitution and characterization of XiaF and XiaP.** (a) Structures of substrates and products of XiaF enzyme assays. (b) HPLC-HRMS monitoring (PDA 300 nm) showing the results of (I) XiaF enzyme assays with indosespene (2) as the substrate in the presence and absence of XiaP (the asterisk denotes the proposed transient intermediate prexiamycin (8); for details, see Supplementary Fig. 6); (II) XiaF enzyme assays with indole (10) as the substrate in the presence and absence of XiaP. The profiles showing the reference compounds indosespene (2) and xiamycin (3) as well as indigo (11) and indirubin (12) are composed of different measurements of pure compounds (colour coded). (c) In vitro activity test of flavin reductase XiaP (0.1 μM). Conversion rates were followed by measuring the decrease in absorbance at 340 nm over time. Data represent mean values of three independent experiments, each conducted as duplicates. Error bars indicate s.d.
compounds, only indole (10) was transformed into an oxygenated product. LC-MS monitoring of the XiaF in vitro assay supplemented with indole gave diagnostic peaks for indigoid pigments indigo (11) and indirubin (12). Both compounds are known to result from oxidative dimerization of indoxyl (13), the immediate product of indole hydroxylation. Notably, these pigments were not formed in the negative control experiment using Fre and heat-denatured XiaF (Supplementary Fig. 2b,III). Likewise, no pigments could be detected in the chromatographic experiments, we found that the same indigoid pigments are also formed in vivo. On extended incubation of E. coli plate cultures that harbour the xiaF-expression plasmid (E. coli/pXU520), we noted a slight colouration of the colonies. Blue pigment production was substantially enhanced when E. coli/pXU520 was plated on agar supplemented with tryptophan (Fig. 4a). Since the colouration appeared without coexpression of fre or xiaP in E. coli, XiaF obviously recruits native E. coli enzymes for flavin reduction. The blue pigmentation was completely absent in control strains harbouring the empty vector only. For a full characterization of the blue pigments, cells were grown in liquid tryptophan medium. HPLC-HRMS analysis with authentic references unequivocally showed that the pigments are identical with indigo (11) and indirubin (12) (Fig. 4b,c).

**XiaF crystal structure and model of cryptic hydroxylation.** To further characterize the non-canonical terpenoid cyclase and the cryptic hydroxylation reaction, we carried out protein crystallization and modelling experiments. The structure of XiaF was solved at 2.4 Å resolution applying molecular replacement (R_free = 22.9%, search model: 4-hydroxyphenylacetate 3-hydroxylase, PDB ID: 2jbs, XiaF (PDB entry codes: 5LVV, 5LWV and 5MR6) assembles as a tetrahedral oligomer, which is in agreement to the molecular mass of 176 kDa as determined by size-exclusion chromatography. The intermolecular interactions between the individual subunits are solely formed between the C-terminal domains, whereas the N-terminal as well as the central β-sheet domains are directed towards the corners of the tetrahedron (Fig. 5a). The N-domain spans from Met1-Lys118, the β-sheet from Leu19-Ser213 and the C-domain from Gly214-Leu397 (Fig. 5b), which describes the typical architecture of a member of the ACAD superfamily (Supplementary Fig. 7).

Cocrystallizations of XiaF with flavin proved to be challenging; initially no FAD bound to crystals could be observed when the oxidized cofactor was used. A similar result has been reported for FMN and the XiaF homologue C2-HpaH (4-hydroxyphenylacetate 3-hydroxylase), which has been shown to bind reduced FMN considerably better than oxidized FMN. Eventually, on incubation of XiaF with its cognate reductase XiaP under anaerobic conditions, we succeeded in binding reduced FADH₂ to the active site (PDB entry code: 5MR6) displays a unique bent conformation of the isoalloxazin ring is embedded within a hydrophobic pocket mainly comprised by Trp87 (3.1 Å), Tyr144 (4.2 Å) and Trp186 (3.8 Å). The ring N5 nitrogen is coordinated by the hydroxyl group of Thr146 (3.1 Å), whereas N3 and O4 are both stabilized by the carbonyl group of Ser121 (both 3.0 Å). Interestingly, the XiaF structure in complex with FADH₂ and glyceraldehyde (PDB entry code: 5MR6) displays a unique bent conformation of the cofactor’s adenine moiety, thus blocking the substrate-binding channel. The phosphoribosyl linker is hereby exposed to the solvent (Fig. 5c and Supplementary Fig. 12). A similar conformation of FAD has been previously reported for an FAD synthase from yeast. This bent conformation is missing the distinct interactions between the FADH₂’s β-phosphate and the amino acids Tyr144, Arg266 and Ala350. Instead, three new intramolecular H-bonds between the ribose and adenine moieties of the cofactor are formed (3.0, 2.6 and 3.0 Å; Supplementary Fig. 14b). However, control assays with FMN, which lacks the adenine moiety, and the characterization of XiaF’s orthologue Xial (ref. 21) have shown that FAD and FMN can both be used by
Figure 5 | Structural analysis of XiaF and scheme of FADH₂-mediated indole hydroxylation. (a) Ribbon structure of the homotetramer (PDB entry code: 5LVU). (b) Close-up cartoon representation of a monomer with bound FADH₂ (PDB entry code: 5LVW). The N-terminal domain is coloured in white, the β-sheet domain in blue and the C-terminal domain in red. (c) Ribbon diagram of XiaF with the enlarged active site. The electron density represents a 2Fo-Fc map with the displayed FADH₂, glycerol (GOL) and three water molecules (W₁₋₃) omitted for phasing (PDB ID: 5MR6). The amino acid is labelled by the one-letter code and numbered according to the XiaF sequence. Black dashed lines indicate hydrogen bonds. (d) Docking of indole (carbon atoms coloured in green), using AutoDock Vina50, with XiaF in complex with the modelled reaction intermediate FAD–OOH. Distances in Å. (e) Catalytic cycle of XiaF consuming one molecule of oxygen and of reduced flavin provided by cognate reductase, yielding a C-3-hydroxylated indole derivative. The oxidized cofactor is transferred back to the reductase where it is recycled in a NAD(P)H/H⁺-dependent reaction. (f) Nucleophilic attack of the exo-methylene group on 3-hydroxyiminium cation 9 results in the pentacyclic intermediate 16. Deprotonation and loss of an equivalent of water yields prexiamycin (8), which aromatizes to yield 3. Spontaneous reactions are indicated by ‘s’.
the enzymes, a feature that is also shared by XiaF’s homologues C2-HpaH and HsaA from *Mycobacterium tuberculosis*[^24]. Therefore, it is unclear whether this unique bent conformation has a biological function. As described previously for the XiaF homologue C2-HpaH[^48], the highly conserved His371 (3.7 Å from C4) sits above the FADH2 ligand, perfectly positioned for specificity for hydroxylating indole. Position of the indole ring (2.9 Å), thus explaining XiaF’s high specificity for indoles (Fig. 5d). This modelling approach suggests that the hydroxylation of the indole ring and Phe123 could be a key interaction catalyzing the formation of xiamycin (Supplementary Fig. 15). This structural modelling highlights the role of Ser121 (3.2 Å) and Phe123 (3.6 Å) in positioning indole and coordinating residues are mainly located on the b-sheet domain and, while the substrate-binding pocket forming amino acids predominantly derive from the C domain. This structural separation of the cofactor and the ligand-binding elements might theoretically allow a structural plasticity of the substrate-binding pocket without disturbing the fine-tuned cofactor binding. In case of XiaF, the substrate-binding pocket is comprised by residues N91, L98, S121, F123, I237, M240, H371 and M373 (Supplementary Fig. 7a).

**Discussion**

In this study, we have gained insights into the structure, function and evolutionary relationship of a non-canonical terpenoid cyclase. Our findings illuminate the mechanism of the key step of this rare bacterial indoloterpenoid pathway, providing a bio catalyst for the production of indigoid pigments and shed more light on the evolution of metabolic diversity.

*In vitro* characterization, bioinformatics and structural investigations revealed that XiaF is a new member of the group D FMOs, monooxygenases that receive reduced FAD/FMN from a NAD(P)H-dependent flavin reductase and assemble in a tetrameric ACAD-fold composed of N- and C-terminal helical domains and a central *ß*-sheet domain[^35]. Genome mining and biochemical analyses revealed the designated flavin reductase partner of XiaF in a two-component system. The successful *in vitro* reconstitution of the two-component system has laid the foundation for studying bacterial indoloterpenoid biosynthesis in more detail. In the current model for xiamycin biosynthesis, it was proposed that a hydroxylated intermediate would set the basis for terpenoid cyclization[^22]. While bioinspired synthetic studies support this model[^54], experimental evidence for an enzymatic hydroxylation as trigger for the cyclization reaction has been missing. Using indole as substrate surrogate, we unveiled the ‘cryptic hydroxylation’ at C-3 of the indole ring, which is clearly indicated by the formation of indigoid pigments. Hydroxylation of indole followed by deprotonation of intermediate 14 (Fig. 5e) results in the formation of indoxyl, which is spontaneously oxidized to yield indigo (Fig. 4d). Alternatively, in an oxygen-rich environment, isatin (15) is formed from indoxyl as a side reaction. Condensation of indoxyl with isatin produces indirubin, which is a common by-product of indigo biosynthesis[^55]. These findings have important implications for the XiaF-catalysed reaction mechanism. Accordingly, in the terpenoid pathway C-3 hydroxylation of the indole ring of indosepsine by C4a-hydroperoxylavin as the electrophilic oxygenating species yields a reactive intermediate (most likely the 3-hydroxyiminium cation 9, although deprotonation to yield an hydroxyindolenine intermediate is also conceivable) that is prone to nucleophilic attack of the *exo*-methylene group, resulting in the pentacyclic intermediate 16 (Fig. 5f). Deprotonation and loss of an equivalent of water yields prexiamycin, which spontaneously aromatizes to yield 3.

Beyond these important mechanistic implications, our findings are also relevant from a biotechnological point of view. Bacterial indigo production is of considerable interest as a sustainable and environmentally friendly substitute for the synthetic production of indigo that depends on coal and oil and generates toxic byproducts[^56]. Also, the second indigoid pigment, indirubin, is a valuable natural product[^58]. Indirubin is not only an antimicrobial agent[^70], but also has therapeutic value against diseases and pathological conditions like cancer, Alzheimer’s disease and inflammation[^61]. In addition, indirubin is the main active ingredient of the plant mixture ‘Danggui Longhui Wan’, which has been used for millennia in traditional Chinese medicine to treat chronic diseases, likely because of its potent inhibition of cyclin-dependent kinases[^62,63]. Because of these important bioactivities, recently there has been a renaissance of oxygenase screening activities towards indirubin production[^37,47,60,64,65]. However, these enzymes generally seem to be rather promiscuous detoxification enzymes, not specialized indole oxygenases. In stark contrast, XiaF is to all appearance uniquely selective in hydroxylation of the indole ring, since aromatic model xenobiotics have not been accepted as substrates. There seems to be the general trend that broad substrate acceptance comes at the price of low reaction turnover numbers[^66,67]. Thus, a specialized indole oxygenase could be of particular interest for biotechnological applications, especially regarding large-scale indigoid pigment production. There is an unabated demand for indigo dye[^57], which is supplied by an annual world production of 80,000 tons of pigment[^68].
From an evolutionary point of view, the most intriguing discovery is that a terpenoid cyclase is functionally and structurally related to various enzymes from bacterial xenobiotic-degradation pathways. Xenobioc-metabolizing enzymes have existed for more than 2.5 billion years.66 Primordial cells have likely been exposed to many hostile chemical challenges and naturally occurring xenobiotics (for example, from volcano eruptions)70,71 and the requirement of protective mechanisms against cytotoxic and genotoxic chemicals are certainly older than secondary metabolism. Moreover, besides their role in detoxification, xenobioc-metabolizing systems have evolved in organisms like bacteria and fungi to tap new chemical resources by degradation of foreign aromatic hydrocarbons. Therefore, one may conceive a metabolic crosstalk of the myriad of ancient broad-specific detoxification/degradation enzymes that are often regarded as an indicator for general substrate tolerance towards aromatic hydrocarbons. On the other hand, this reaction could also be the result of natural selection. In recent years, evidence is mounting that indole hydroxylation catalysed by these enzymes is of physiological relevance, since indole is toxic to microorganisms, and indigo formation by indole oxygenases represents a detoxification mechanism. Moreover, the catalysis of indole formation provides the missing link that connects XIAF with most of these catabolic oxygenases at a functional level.

Indole hydroxylation and subsequent indigo formation is a side reaction of many xenobioc-metabolizing enzymes that is often regarded as an indicator for general substrate tolerance towards aromatic hydrocarbons. On the other hand, this reaction could also be the result of natural selection. In recent years, evidence is mounting that indole hydroxylation catalysed by these enzymes is of physiological relevance, since indole is toxic to microorganisms, and indigo formation by indole oxygenases represents a detoxification mechanism. Moreover, the catalysis of indole formation provides the missing link that connects XIAF with most of these catabolic oxygenases at a functional level.

On the basis of our structural data and docking experiments with indole, a potential binding mode of indole within the active site of XIAF could be revealed (Fig. 5d and Supplementary Fig. 13). A two-fold coordination of the substrate via an interaction of indole’s N1 and Ser121 as well as π-stacking between the aromatic substrate and Phe123 seems to be most important for molecular recognition of this chemical group. A modelling of the FAD–OOH reaction intermediate suggests a coordination sphere of indole, which places the hydroperoxide exactly above and in close distance to indole’s C3 position (Fig. 5d). This unique architecture may rationalize XIAF’s specificity for hydroxylation and subsequent indigo formation is a side reaction of many xenobioc-metabolizing enzymes that is often regarded as an indicator for general substrate tolerance towards aromatic hydrocarbons. On the other hand, this reaction could also be the result of natural selection. In recent years, evidence is mounting that indole hydroxylation catalysed by these enzymes is of physiological relevance, since indole is toxic to microorganisms, and indigo formation by indole oxygenases represents a detoxification mechanism. Moreover, the catalysis of indole formation provides the missing link that connects XIAF with most of these catabolic oxygenases at a functional level.

The remaining part of indolespene is bound to an open substrate channel (Supplementary Fig. 7). In case of the XIAF homologue 4-hydroxyphenylacetate 3-hydroxylase, the active site is limited by Arg263 (ref. 48). This amino acid is replaced in XIAF by Ile237, which opens up the binding channel for accepting the elongated indolespene substrate. Thus, the flexible active site architecture of these enzymes presumably builds the basis for the enzymes’ potential utilization of various substrates.

We suggest that the promiscuous indole-metabolizing reaction, shared by many catabolic XIAF homologues might be the evolutionary origin of the cyclization reaction of indolespene. The most parsimonious explanation would be that XIAF has evolved from a broad-specific xenobioc-metabolizing oxygenase that was able to accept the more complex indole compound 2, besides its ability to metabolize indole. In contrast to the formation of indoxyl, which results in the formation of indigoid pigments, hydroxylation of 2 at position C-3 triggers cyclization. In the course of evolution, XIAF likely specialized on this particular reaction. This is indicated by experiments to test the substrate flexibility of XIAF, for which typical xenobioc substrates of other XIAF homologues were not transformed.

Surprisingly, among the XIAF homologues are also the tentative polyketide-tailoring hydroxylases ActVA-ORF5 (ref. 78), and NcnH79 and the glycoside-modifying N-oxigenases KijD3 (ref. 53) and DnmZ80. Thus, it appears that various enzymes that catalyse secondary metabolite tailoring reactions could have evolved from detoxification pathways, yet this relationship has been overlooked.

Finally, our hypothesis that xiaF originates from xenobioc metabolism is substantiated by the fact that it clusters with genes (xiaGHII) that encode homologues of xenobioc-metabolizing enzymes. Therefore, it is tempting to speculate that this ‘detox gene cassette’ (Fig. 2a), which interrupts the core biosynthesis genes xiaE and xiaKLM originates from genes for broad-specific xenobioc-metabolizing enzymes that were recruited for the diversification of the IST core scaffold preceding by a scenario even more plausible. Moreover, these considerations along with the structural and phylogenetic analyses, the biochemical characterization and biotransformation experiments indicate that the specialized terpenoid cyclase presumably has evolved from the functionally related detoxification enzymes.

Methods

Organisms, cultivation conditions and plasmids. The E. coli strains Top 10 (Life Technologies Corporation), BL21 (DE3) and BL21-CodonPlus(DE3)-RP (Agilent Technologies GmbH) were used for cloning and heterologous expression experiments. Electrocompetence of cells was achieved by repeated washing with diluted glycerol. Transformation was carried out in 2 mm cuvettes at 2.5 kV in a Gene Pulser Xcell electroporator (Bio-Rad) or according to the manual for commercially available chemical competent Top 10 cells (Life Technologies Corporation). Cells were grown in Luria broth (LB medium, with appropriate antibiotics, whenever needed.

Plasmids used in this study were the pET-28a(+) based expression vector (Merck KGaA) pET28-fre carrying a NAD(P)H-dependent flavin reductase gene (fre) combined with a 5′-sequence coding for a hexahistidine-tag, the pET-28b(+) based expression vector pXU520 (for the production of recombinant XIAF with N-terminal histidine-tag), the pET-28c(+) based expression vector pET28c-xiaP (for the production of recombinant XIAF with N-terminal histidine-tag) and the pET-28b(+) based expression vector pET28b(+)-xiaF (for the production of recombinant XIAF with N-terminal histidine-tag used for crystallization).

DNA preparation and analysis. PCR or digested products were separated on agarose gels and purified from the gel using inuPREP Gel Extraction Kit (Analytic Jena). Digestion and ligation reactions were carried out by standard methods. Isolation of plasmid DNA was accomplished with standard alkaline lysis principle.

Construction of expression vectors for xiaF, xiaP and fre. For cloning the coding sequences of xiaF and xiaP downstream of the T7 promoter in pET28 vectors, the genes were PCR-amplified from cosmid 0B02 (from a genomic cosmid library of the endophyte) using primer sets XIAF_NdeI_fwd (5′-GGAGTTGGTG-CATATGCAGCA-3′), Ndel restriction site underlined) and XIAF_rv (5′-CGA-CATGCGATGAGAAA-3′) for xiaF and SK01 (5′-GCCATATGTC-ACGGGAGACCTTCG-3′) and SK02 (5′-CCGGACCCCATCACC- GCCG-3′, Ndel and BamHI restriction sites underlined) as primers. PCR was carried out with KAPAHiFi (PEQLAB) for xiaF and Phusion High-Fidelity DNA Polymerase (NEB) for xiaP. The amplicons were subcloned into the vector pCR-Blunt (Life Technologies) and the ligation mixture was introduced into electrocompetent E. coli Top 10 cells. Afterwards, plasmids were isolated from positive transformants and verified by sequencing. Excised DNA from correct plasmids was cloned by using the Ndel and BamHI restriction sites of pET-28b(+) or pET-28b(+) (for recombinant xiaF) and pET-28c(+) (for recombinant xiaP) to generate pXU520 and
pET26b (+) xiaF (for the production of recombinant XiaF with N-terminal histidine-tag) and pET28c xiaA (for the production of recombinant XIA with N-terminal hexahistidine tag). After introducing the mixture into LB, the Top 10 cells, plasmids were isolated from positive transformants and verified by sequencing. Correct plasmids were introduced into E. coli BL21 CodonPlus(DE3)-RP for heterologous expression.

Amplification and cloning of xia fre from BL21 (DE3) were carried out as described by Zeng et al.,31,32 with reference to a pET-28a (+)−based expression vector (pet28a-fre), which was introduced into electrocompetent E. coli BL21-CodonPlus(DE3)-RP cells for overexpression.

Production and purification of recombinant XIAF and XIAp. For heterologous protein production, E. coli BL21-CodonPlus(DE3)-RP carrying pXU520 or pET26b (+) xiaF for recombinant XIAF or pET28c xiaA for recombinant XIA were grown in LB medium supplemented with 34 mM 2.1 mm 3-chloroamphetamine and 50 μg ml−1 kanamycin at 30 °C and 180 r.p.m. to an OD_{600} of 0.4–0.6. Subsequently, cultures were induced with IPTG at a final concentration of 0.1 mM and incubated at 30 °C for 3 h. After collecting of cells, the pellet was resuspended in cold Fre lysis buffer (up to 50 ml for 1.6 l culture; 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 2 mM DTT) and lysed by means of Sonopuls HD 2,200 (BANDELIN). Separated soluble fraction was subjected onto a HiLoad 16/600 Superdex 200 prep grade column using a SEC buffer (100 mM NaH₂PO₄, 100 mM Na₂HPO₄, 2% glycerol, pH 7.0). The protein production, protein purification, and analysis were performed with XiaF assay buffer and substrate solution.

Activity assays of Fre and XIAp. The recombinant enzymes Fre (derived from E. coli) and XIAp (derived from Streptomyces sp. HK0576) were tested for flavin reductase activity. Reactions were set up in a standard assay buffer (20 μM Tris-HCl pH 8.0, 1 mM NADH) containing 50 μM NAD(P)H (Sigma-Aldrich), 50 μM FAD/FMN/riboflavin (Sigma-Aldrich) and 1 μM flavin reductase at room temperature, unless stated otherwise. Reaction rates were monitored by following the decrease of NAD(P)H absorbance at 340 nm caused by the oxidation to NAD(P). To test the catalytic activity of Fre, initial assays were performed with 25 μM Fre, 2 mM NADH, 1 mM FAD and 100 μM indoxyl in Fre assay buffer at 28 °C for 24 h, which were supplemented with 10 μM MnFRE during the course of the study. Control assays contained no flavin oxidoreductase. For analysis, the reactions were stopped with one volume of acetonitrile (indole, 4-nitrotoluene) or methanol (indole 3-acetic acid), dried under reduced pressure, dissolved in 0.5 ml of 0.5% acetic acid, and the reaction products were extracted with acetonitrile. The extract was filtered and analysed by means of HPLC-HRMS (Exactive Plus Orbitrap Mass Spectrometer, Thermo Scientific) operating with a 150 × 2.1 mm C18 column (Thermo Scientific) and a gradient from 5 to 15% solvent B in 15 min at 40 °C. The purity of His-tagged Fre (28.4 kDa) was checked on a Superdex 200 prep grade column (GE Healthcare) in manual handling. For Standardization of reaction, the reactions were performed with 25 μM Fre, 2 mM NADH, 1 mM FAD, 50 μM NiFRE and 100 μM indoxyl in Fre assay buffer at 28 °C for 24 h, which were supplemented with 10 μM MnFRE during the course of the study. Control assays contained no flavin oxidoreductase. For analysis, the reactions were stopped with one volume of acetonitrile (indole, 4-nitrotoluene) or methanol (indole 3-acetic acid), dried under reduced pressure, dissolved in 0.5 ml of 0.5% acetic acid, and the reaction products were extracted with acetonitrile. The extract was filtered and analysed by means of HPLC-HRMS (Exactive Plus Orbitrap Mass Spectrometer, Thermo Scientific) operating with a 150 × 2.1 mm C18 column (Thermo Scientific) and a gradient from 5 to 15% solvent B in 15 min at 40 °C. The purity of His-tagged Fre (28.4 kDa) was checked on a Superdex 200 prep grade column (GE Healthcare) in manual handling.

Indigoid pigment production in vivo. For pigment production on plate, an overnight culture of E. coli BL21-CodonPlus(DE3)-RP carrying the plasmid pXU520 (E.coli/pXU520) was streaked onto LB-agar plates supplemented with 0.1 mM IPTG, 10 μM tryptophan and 50 μg ml−1 kanamycin. The plates were incubated at room temperature for several days until blue pigmentation occurred. E. coli BL21-CodonPlus(DE3)-RP carrying pET28b (+) (E.coli/pET-28b (+)) was used as a negative control. After incubation at 30 °C for 3 days, a blue spot was grown in 100 ml tryptophan medium (2.1 g/l tryptophan, 5 g/l yeast extract, 10 g/l NaCl) supplemented with 50 μg ml−1 kanamycin at 30 °C and 200 r.p.m. After reaching an OD_{600} of 0.4, cells were induced with a final concentration of 0.1 mM IPTG and incubated for 48 h. Centrifugation of the production culture resulted in a dark blue pellet. Extraction of this pellet with dimethylformamide yielded a dark blue extract, which turned reddish after several days of incubation. In addition, for pigment characterization, blue pellet was dried and extracted with acetonitrile. The extract was filtered and analysed by means of HPLC-HRMS (Exactive Plus Orbitrap Mass Spectrometer, Thermo Scientific) operating with a 150 × 2.1 mm C18 column (Thermo Scientific) and a gradient from 5 to 98% solvent B in 15 min (solvent A, 0.3% (v/v) HCOOH in water; solvent B, 0.1% (v/v) HCOOH in acetonitrile) at a flow rate of 200 μl/min−1 and UV/HRMS detection. As alternative substrates, indole (conditions identical to indosespene assays) and substrates from homologous enzymes (4-nitrotoluene, 4-hydroxyphenylacetate, 2-naphthoic acid, indole 3-acetic acid, 3-hydroxyphenylacetate) were tested in XiaF biotransformation assays. In the latter case, reaction volumes were increased up to 300 μl. Maximum concentrations used were 50 μM Fre, 10 μM XiaF and 2 mM substrates. For analysis, the reactions were stopped with one volume of acetonitrile (indole, 4-nitrotoluene) or methanol (indole 3-acetic acid), dried under reduced pressure, dissolved in 0.5 ml of 0.5% acetic acid, and the reaction products were extracted with acetonitrile. The extract was filtered and analysed by means of HPLC-HRMS (Exactive Plus Orbitrap Mass Spectrometer, Thermo Scientific) operating with a 150 × 2.1 mm C18 column (Thermo Scientific) and a gradient from 5 to 15% solvent B in 15 min at 40 °C. The purity of His-tagged Fre (28.4 kDa) was checked on a Superdex 200 prep grade column (GE Healthcare) in manual handling. For Standardization of reaction, the reactions were performed with 25 μM Fre, 2 mM NADH, 1 mM FAD and 100 μM indoxyl in Fre assay buffer at 28 °C for 24 h, which were supplemented with 10 μM MnFRE during the course of the study. Control assays contained no flavin oxidoreductase. For analysis, the reactions were stopped with one volume of acetonitrile (indole, 4-nitrotoluene) or methanol (indole 3-acetic acid), dried under reduced pressure, dissolved in 0.5 ml of 0.5% acetic acid, and the reaction products were extracted with acetonitrile. The extract was filtered and analysed by means of HPLC-HRMS (Exactive Plus Orbitrap Mass Spectrometer, Thermo Scientific) operating with a 150 × 2.1 mm C18 column (Thermo Scientific) and a gradient from 5 to 15% solvent B in 15 min at 40 °C. The purity of His-tagged Fre (28.4 kDa) was checked on a Superdex 200 prep grade column (GE Healthcare) in manual handling.
UV/HRMS detection. Authentic references were used for comparison. e) E coli pET28b (+) was used as a control.

**Determination of kinetic parameters of XiaP.** Flavin reductase (19,220 Da, sequence see below) activity was determined aerobically in XiaF assay buffer (20 mM Tris-HCl pH 8.0, 10% glycerol) at 28 °C in the dark. The reaction mixtures (100 μl or 120 μl) contained constant NAD(P)H and varying flavin concentrations (FAD, FMN, riboflavin), or vice versa, and 0.1 μM XiaF. His-tagged XiaF was purified in three independent experiments by affinity chromatography (IMAC) followed by a gel filtration technique (active appearance as homodimer).

Concentrations of all XiaP enzyme solutions were determined using the Pierce BCA Protein Assay Kit (Thermo Scientific). Assays were initiated by addition of enzyme. The activity was measured by monitoring the NAD(P)H absorbance at 340 nm in 96-well plates on a Varioskan LUX plate reader (Thermo Scientific). Initial decrease in NAD(P)H concentration over time was used to calculate the conversion rate (μM min⁻¹ as v). Apparent kinetic parameters were determined from double reciprocal plotting of initial velocities (v/v) and substrate concentration ([S]u) fitted by linear regression calculated with GraphPad Prism5. Three independent experiments with two technical replicates were used for calculation of kinetic parameters. For the determination of apparent kinetic values of NADH, 42 mM FMN or riboflavin were added to 4-1, 12-5, 24, 41.8, 53, 83.3, 208, 216, 833, 1250, 1667, 2083 and 2500 μM of NADH (Vmax = 120 μM) and absorbance was monitored for maximal 24 min. The apparent kinetic values of FAD, FMN or riboflavin were determined with varying concentrations of flavins (5, 10, 20, 40, 80, and 150 μM) and 1.5 mM NADH. Absorbance measurement was started 5 min after enzyme addition for 45s. For each absorbance value, buffer absorbance was subtracted and final absorbance was converted into NAD(P)H concentration by a calibration curve. Since autocatalytic NAD(P)H oxidation in control assays (without XiaF) was minute in comparison to the XiaP enzyme assays (Supplementary Fig. 4a,d) the activity was not detected for the NAD(P)H concentration of the different kinds of flavins on the absorbance was considered to be equal and was neglected for the determination of the kinetic parameters of XiaP as well. Flavins that were putatively bound to XiaF after purification (noticeable by a very slight yellow colour of the solution) were neglected in the determination since control measurements with the buffer only showed NAD(P)H did not lead to a decrease in absorbance as well as none of the heat-denatured enzyme solutions revealed a detectable absorbance at 450 nm.

Amino-acid sequence of His-tagged XiaF used for determination of kinetic parameters:

**Phylogenetic analysis.** The evolutionary history was inferred using the neighbour-joining method36 and tested with the bootstrap test (1,000 replicates)34. The tree (Fig. 2b) is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the p-distances method and are representative of the number of amino-acid differences per site. Evolutionary analyses were conducted in MEGA6 (ref. 86). The analysis involved the following enzymes (NCBI/PDB accession number in brackets): XiaF from Streptomyces sp. HK0576 (CCHE63727)27, XiaF from Streptomyces sp. SCISO 02999 (AKF78075)21, XiaF #1 from Streptomyces sp. WP 031042 (ACFJ52313)38, XiaF #3 from Streptomyces fradiae (WP_043467134)28, XiaF #4 from Amycolatopsis nigerescens CSCI7ta-90 (WP_022673280)28, XiaF #5 from Streptomyces sp. AA0539 (Streptomyces sp. AA0539)29, XiaF #6 from Streptomyces sp. NRR-S15 (WP_031097791)28, XiaF #7 from Streptomyces flavigrises (WP_031220709)29, XiaF #8 from Micrococcus genitalis (WP_003290851)29, XiaF #9 from Streptomyces sp. JEF9 (WP_004423139)29, XiaF #10 from Streptomyces sp. E9 (WP_010330054)30, XiaF #11 from Cupriavidus necator (CAA35887)30, XiaF #12 from Streptomyces sp. PG22 (WP_011215578)31, XiaF #13 from Geobacillus stearothermophilus (AAAS85688)44, Idoa #1 from Pseudomonas algicola PA-10 (AF474963)45, Idoa #2 from Acinetobacter baumannii ATCC 19,606 (EEOX4702)25, lca #2 from Pseudomonas putida (ABY62757)26, Dsc #1 from Rhodococcus erythropolis IGTSB (AA9A9484)25, Dsc #2 from Rhodococcus erythropolis JF10 (AB467379)25, Dsc #3 from Rhodococcus erythropolis FLPP13 (ABC08118)25, IacA #1 from Rhodococcus erythropolis JF13 (AB467379)25, IacA #2 from Rhodococcus erythropolis JF11 (AAB97196)46, IacA #12 from Micrococcus samaricus (WPS1)46, IacA #13 from Micrococcus marismortui (WPS1)46, IacA #14 from Pseudomonas sp. X85 (WP_011228328) as an outgroup.90. The result of the phylogenetic tree was confirmed with the minimal evolution method31, which was used as an alternative method. For more detailed cladograms see Supplementary Figs 8 and 9.

**Crystallographic and structure determination.** Crystals of XiaF were grown at 20 °C within 2 days to their final size applying the hanging drop (1 μl × 1 μl) vapour diffusion method. For XiaF:FADH2 crystals, XiaF (127 μM), FAD (79 μM), NADH (500 μM) and FAD (150 μM) were anaerobically mixed inside a glove box and incubated for 25 min before setting up the crystallization trials. The XiaF:FADH2 crystals differed in 1.6 M (NH4)2SO4, 100 mM HEPES/NaOH, pH 7.5 and 100 mM NaCl. Subsequently, crystals were transferred in 4 μl drops containing the crystallization buffer plus 30% of glycerol. After 2 min, crystals were vitrified and stored in liquid nitrogen. For XiaF:FADH2 crystals, XiaF, FAD (193 μM), XIA (80 μM), NADH (2.75 mM), FAD (480 μM) and Na-dithionite (20 mM) were mixed and incubated for 60 min at 10 °C. Next, 1 μl of glycerol was added and incubated for 15 min. This mixture was used for setting up the crystallization trials. The XiaF:FADH2 crystals appeared in 1.6 M (NH4)2SO4, 100 mM HEPES/NaOH, pH 7.5, 100 mM NaCl and 5 mM Na-dithionite. Diffraction data sets were recorded using synchrotron radiation of λ = 1.0 Å at the beamline X06SA, Swiss Light Source (SLS), Paul Scherrer Institut, Villigen, Switzerland. Collected reflections were processed using the programme package XDS. The statistics are summarized in Supplementary Table 3.

Docking of the indole ligand into the active site of XiaF. For docking of the indole ligand, the XiaF structure in complex with FADH2 and glycerol was chosen (PDB ID: 5MR6). The position of the glycerol ligand served as search centre for the docking experiment. The docking grid was restricted to the active site, which is limited by FADH2’s adenine moiety. The indole ligand was built and energetically minimised with the Avogadro software.58 The docking grid was performed using AutoDock Vina50 as a plugin from the structure analysis software suite Chimera.59 To prepare XiaF for the docking experiment, the plugin Prep from the Chimera software suite was used. Hereby, solvent molecules were deleted and Gasteiger charges and hydrogens were added. For the latter ones, hydrogen bonds were considered as well, non-polar hydrogens were merged. Next, AutoDock Vina was run. Therefore, hydrogens were added to the indole ligand as well. The search was centred around the glycerol ligand coordinates (x = 49.1286, y = -47.4138, z = -0.4335),60 the docking grid had a size of 10.5886 × 9.5815 Å × 8.5837 Å. For the receptor (XiaF), all of the following hydrogen bond atoms were selected to be true: Merge charges and remove non-polar hydrogens, merge charges and remove lone pairs, ignore waters, ignore chains of non-standard residues, ignore all non-standard residues. For the ligand (indole), the following settings were selected to be true: Merge charges and remove non-polar hydrogens, merge charges and remove lone pairs, ignore waters, ignore chains of non-standard residues, ignore all non-standard residues. The orientation relative to the H2O molecule W1 (Fig. 5c), which presumably complies with the orientation of the hydroperoxide group during catalysis. The second best score, bloc the ligand, vertically flipped, increases the distance between the C3 position of indole and W1. The XiaF:FADH2:glycerol structure docked with indole was further analysed by modelling the reaction intermediate FAD-OOH using the graphic programme MAIN52 (Fig. 5d).

**Data availability.** The coordinates and structure factors of XiaF have been deposited in the Protein Data Bank [http://www.rcsb.org/pdb] with accession numbers 5LVU, 5LVV and 5MR6. Any additional data not shown in this published article or in its Supplementary Information are available from the corresponding authors on reasonable request.

**References**

1. Gershenzon, J. & Dudareva, N. The function of terpene natural products in the natural world. Nat. Chem. Biol. 3, 408–414 (2007).

2. Davies, J. E. Specialized microbial metabolites: functions and origins. J. Antibiot. 66, 11 (2013).

3. Welsch, M. E., Snyder, S. A. & Stockwell, B. R. Privileged scaffolds for library design and drug discovery. Curr. Opin. Chem. Biol. 14, 347–361 (2010).
4. Clardy, J. & Walsh, C. Lessons from natural molecules. *Nature* 432, 829–837 (2004).
5. Dewick, P. M. *Medicinal Natural Products: a Biosynthetic Approach* 550 (John Wiley & Sons, 2009).
6. Medema, M. H. et al. Minimum information about a biosynthetic gene cluster. *Nat. Chem. Biol.* 11, 625–631 (2015).
7. Hertweck, C. Decoding and reprogramming complex polyketide assembly lines: prospects for synthetic biology. *Trends Biochem. Sci.* 40, 189–199 (2015).
8. Floss, H. G. Combinatorial biosynthesis—potential and problems. *J. Biotechnol.* 124, 242–257 (2006).
9. Herrera, J. B. R., Wilson, W. K. & Matsuda, S. P. T. A tyrosine-to-threonine mutation converts cycloartenol synthase to an oxidosqualene cyclase that forms lanosterol as its major product. *J. Am. Chem. Soc.* 122, 6756–6766 (2000).
10. Morrone, D., Xu, M. M., Fulton, D. B., Determan, M. K. & Peters, R. J. Increasing complexity of a diterpene synthase reaction with a single residue switch. *J. Am. Chem. Soc.* 130, 5400–5401 (2008).
11. Roach, J. J. & Shenvi, R. A. Supramolecular catalysis: terpenes in tight spaces. *Nat. Chem.* 7, 187–189 (2015).
12. Dickschat, J. S. Bacterial terpene cyclases. *Nat. Prod. Rep.* 33, 87–110 (2016).
13. Baunach, M., Franke, J. & Hertweck, C. Terpenoid biosynthesis off the beaten track: unconventional cyclases and their impact on biomimetic synthesis. *Angew. Chem. Int. Ed.* 54, 2604–2626 (2015).
14. Li, S. et al. Hapalindole/ambiguine biosynthesis is mediated by a cype rearrangement, C – C bond-forming cascade. *J. Am. Chem. Soc.* 137, 15366–15369 (2015).
15. Kim, S. H., Ha, T. K., Oh, W. K., Shin, J. & Oh, D. C. Antiviral indolosesquiterpene from actinomycete Xiamycin C-E from a halophilic actinomycete. *J. Nat. Prod.* 79, 51–58 (2016).
16. Ding, L. et al. Xiamycin, a pentacyclic indolosesquiterpene with selective anti-HIV activity from a bacterial mangrove endophyte. *Biogeo. Med. Chem. Lett.* 20, 6685–6687 (2010).
17. Meng, Z. et al. Total synthesis and antiviral activity of indolosesquiterpenoids from the xiamycin and oridamycin families. *Nat. Commun.* 6, 6096 (2015).
18. Baunach, M., Ding, L., Bruhl, T., Bringmann, G. & Hertweck, C. Regiodivergent N-C and N-N aryl coupling reactions of indoletterpenes and cycoletterpenes mediated by a single bacterial flavoenzyme. *Angew. Chem. Int. Ed.* 52, 9040–9043 (2013).
19. Baunach, M., Ding, L., Willing, K. & Hertweck, C. Bacterial synthesis of unusual sulfoxonamide and sulfone antibiotics by flavoenzyme-mediated sulfur dioxide capture. *Angew. Chem. Int. Ed.* 54, 13279–13283 (2015).
20. Ding, L., Maier, A., Fiebig, H. H., Lin, W. H. & Hertweck, C. A family of multicyclc indolosesquiterpenes from a bacterial endophyte. *Org. Biomol. Chem.* 9, 4029–4031 (2011).
21. Li, H. et al. Identification and characterization of xiamycin A and oxiamycin gene cluster reveals an oxidative cyclization strategy tailoring indolosesquiterpene biosynthesis. *J. Am. Chem. Soc.* 134, 8996–9005 (2012).
22. Xu, Z., Baunach, M., Ding, L. & Hertweck, C. Bacterial synthesis of diverse indole terpene alkaloids by an unparalleled cyclization sequence. *Angew. Chem. Int. Ed.* 51, 10293–10297 (2012).
23. Zhang, Q. et al. N-N-coupled indolo-sesquiterpenate atropo-diastereomers from a marine-derived actinomycete. *Eur. J. Org. Chem.* 2012, 5256–5262 (2012).
24. Takeda, K., Kajiwara, H. & Imamura, N. Oridamycins A and B, anti-Saprolegnia parasitica indolosesquiterpenes isolated from Streptomyces sp. KS84. *J. Nat. Prod.* 73, 698–701 (2010).
25. Sun, Y. et al. Bioinspired total synthesis of seseptene. *Angew. Chem. Int. Ed.* 53, 9012–9016 (2014).
26. Li, H. et al. Elevulating the cyclization cascades in xiamycin biosynthesis by substrate synthesis and enzyme characterization. *Org. Lett.* 17, 306–309 (2015).
27. Nagayama, H. et al. Isolation of oxygenase genes for indigo-forming activity from actinomycete provides insight into the divergence of enzymes in the class of two-protein component aromatic hydroxylases. *Biochem. Biophys. Acta 1680*, 60–66 (2004).
28. Kim, I. C. & Oriel, P. J. Characterization of the Bacillus stearothermophilus BR219 phenol hydroxylase gene. *Appl. Environ. Microbiol.* 61, 1252–1256 (1995).
29. Arand, M. et al. Structure of Rhodococcus erythropolis limonene-1,2-epoxide hydroxylase reveals a novel active site. *EMBO J.* 22, 2583–2592 (2003).
30. Zhang, Q. et al. Carboxyl formation from methyl via triple hydroxylation by X3A in xiamycin A biosynthesis. *Org. Lett.* 14, 6142–6145 (2012).
31. Han, G. H., Gim, G. H., Kim, W., Seo, S. I. & Kim, S. W. Enhanced indirubin production in recombinant Escherichia coli harboring a flavin-containing monooxygenase gene by cyanine supplementation. *J. Biotechnol.* 164, 179–187 (2012).
32. Alifieri, A. et al. Structure of the monooxygenase component of a two-component flavoprotein monooxygenase. *Proc. Natl Acad. Sci. USA* 104, 1177–1182 (2007).
33. Lecoulot, N. et al. Crystal structure of yeast FAD synthetase (Fad1) in complex with FAD. *J. Mol. Biol.* 398, 641–646 (2010).
34. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31, 455–461 (2010).
35. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Trends Biochem. Sci.* 30, 58–60 (2005).
36. Maugard, T., Enaud, E., Choisy, P. & Legoy, M. D. Identification of an indigomine producing strain from isolate of *Isatis tinctoria* (woad). *Phytochemistry* 58, 897–904 (2001).
37. Berry, A., Dodge, T. C., Pepsin, M. & Weyer, W. Application of metabolic engineering to improve both production and use of biotech indigo. *Ind. Microbiol. Biotechnol.* 28, 127–133 (2002).
38. Chavan, R. B. in *Deninm: Manufacture, Finishing and Applications* (ed. Paul, R.) Ch. 3, 37–67 (Woodhead Publishing, 2015).
39. Han, X., Wang, W. & Xiao, M. Microbial biosynthesis and biotransformation of indigo and indigo-like pigments. *Chin. J. Biotechnol.* 24, 921–926 (2008).
59. O’Connor, K. E., Dobson, A. D. & Hartmans, S. Indigo formation by microorganisms expressing styrene monooxygenase activity. Appl. Environ. Microbiol. 63, 4287–4291 (1997).

60. MacNeil, I. A. et al. Expression and isolation of antimicrobial small molecules from soil DNA libraries. J. Mol. Microbiol. Biotechnol. 3, 301–308 (2001).

61. McClay, K., Boss, C., Keresztes, I. & Steffan, R. J. Mutations of toluene-4-monooxygenase that alter regiospecificity of indole oxidation and lead to production of novel indigoid pigments. Appl. Environ. Microbiol. 71, 5476–5483 (2005).

62. Hoessel, R. et al. Indirubin, the active constituent of a Chinese antileukaemia medicine, inhibits cyclin-dependent kinases. Nat. Cell Biol. 1, 60–67 (1999).

63. Wang, Y., Hoi, P. M., Chan, J. Y. & Lee, S. M. New perspective on the dual functions of indirubins in cancer therapy and neuroprotection. Anticancer Agents Med. Chem. 14, 1213–1219 (2014).

64. Zhang, X. Enzymes and operons mediating xenobiotic degradation in bacteria. Biotechnol. Lett. 15, 1419–1424 (1993).

65. McClay, K., Boss, C., Keresztes, I. & Steffan, R. J. Mutations of toluene-4-monooxygenase that alter regiospecificity of indole oxidation and lead to production of novel indigoid pigments. Appl. Environ. Microbiol. 71, 5476–5483 (2005).

66. Carbonell, P., Lecointre, G. & Faulon, J.-L. Origins of specificity and promiscuity in metabolic networks. J. Biol. Chem. 286, 43994–44004 (2011).

67. Jakoby, W. B. & Ziegler, D. M. The enzymes of detoxication. J. Biol. Chem. 265, 20715–20718 (1990).

68. Ghaly, A. E., Ananthashankar, R., Alhattab, M. & Ramakrishnan, V. V. Production, characterization and textile effluents: a critical review. J. Eng. Proc. Technol. 5, 1–18 (2014).

69. Nebert, D. W. & Dieter, M. Z. The evolution of drug metabolism. Pharmacol. Rev. 61, 124–135 (2000).

70. Coleman, J. O. D., BlakeKalff, M. M. A. & Davies, T. G. Detoxification of xenobiotics by plants: chemical modification and vacuolar compartmentation. Trends Plant Sci. 2, 144–151 (1997).

71. Menzie, C. A., Potocki, B. B. & Santodonato, J. Exposure to carcinogenic PAHs in the environment. Environ. Sci. Technol. 26, 1278–1284 (1992).

72. Mishra, V., Lal, R. & Srivinasan. Enzymes and operons mediating xenobiotic degradation in bacteria. Crit. Rev. Microbiol. 27, 133–166 (2001).

73. O’Connor, K. E. & Hartmans, S. Indigo formation by aromatic hydrocarbon-degrading bacteria. Biotechnol. Lett. 20, 219–223 (1998).

74. Fernández-Arrojo, L., Guazzaroni, M.-E., López-Cortés, N., Beloqui, A. & Ferrer, M. Metagenomic era for biocatalyst identification. Curr. Opin. Biotechnol. 21, 725–733 (2010).

75. Lee, J., Jayaraman, A. & Wood, T. K. Indole is an inter-species biofilm signal mediated by SdhA. BMC Microbiol. 7, 42 (2007).

76. Lee, J. H. & Lee, J. Indole as an intercellular signal in microbial communities. FEMS Microbiol. Rev. 34, 426–444 (2010).

77. Lin, G. H., Chen, H. P. & Shu, H. Y. Detoxification of indole by an indole-induced flavoprotein oxygenase from Acinetobacter baumannii. PLoS ONE 10, e0138798 (2015).

78. Taguchi, T. et al. Biosynthetic conclusions from the functional dissection of oxygenases for biosynthesis of actinorhodin and related Streptomyces antibiotics. Chem. Biol. 20, 510–520 (2013).

79. Brunker, P., Stener, O., Bailey, J. E. & Minas, W. Heterologous expression of the naphthocyclinone hydroxylase gene from Streptomyces averaeanus for production of novel hybrid polyketides. Anton.Leeuw. 79, 235–245 (2001).

80. Al-Mestarihi, A., Romo, A., Liu, H. W. & Bachmann, B. O. Nitrososynthase-triggered oxidative carbon-carbon bond cleavage in baumycin biosynthesis. J. Am. Chem. Soc. 135, 11457–11460 (2013).

81. Diaz, E. Bacterial degradation of aromatic pollutants: a paradigm of metabolic versatility. Int. Microbiol. 7, 173–180 (2004).

82. Springael, D. & Top, E. M. Horizontal gene transfer and microbial adaptation to xenobiotics: new types of mobile genetic elements and lessons from ecological studies. Trends Microbiol. 12, 53–58 (2004).

83. Zeng, J. & Zhan, J. A novel fungal flavin-dependent halogenase for natural product biosynthesis. ChemBioChem. 11, 2119–2123 (2010).

84. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791 (1985).

85. Wei, M. & Kumar, S. Molecular Evolution and Phylogenetics (Oxford University Press, 2000).

86. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729 (2013).

87. Lin, G. H. et al. Identification and characterization of an indigo-producing oxygenase involved in indole 3- acetate acid utilization by Acinetobacter baumannii. Anton. Leew. 101, 881–890 (2012).

88. Leevea, J. & Gerards, S. Discovery of a bacterial gene cluster for catabolism of the plant hormone indole 3-acetic acid. FEBS Microbiol. Ecol. 65, 238–250 (2008).

89. Drewlo, S., Bramer, C. O., Madkour, M., Mayer, F. & Steinbuchel, A. Cloning and expression of a Ralstonia eutropha HFB39 gene mediating indigo formation in Escherichia coli. Appl. Environ. Microbiol. 67, 1964–1969 (2001).

90. Kim, S. H. et al. Crystal structure of the oxygenase component (HpaB) of the 4-hydroxyphenylacetate 3-monooxygenase from Thermus thermophilus HB8. J. Biol. Chem. 282, 33107–33117 (2007).

91. Zeng, J. & Zhan, J. A novel fungal flavin-dependent halogenase for natural product biosynthesis. ChemBioChem. 11, 2119–2123 (2010).

92. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791 (1985).

93. Vagin, A. A. et al. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. Acta Crystallogr. D 69, 1342–1357 (2013).

94. Langer, G., Cohen, S. X., Laminz, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat. Protoc. 3, 1171–1179 (2008).

95. Hanwell, M. D. & Ferrer, M. Metagenomic era for biocatalyst identification. Curr. Opin. Biotechnol. 21, 725–733 (2010).

96. Pettersen, E. F. et al. UCSF chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612 (2004).

Acknowledgements
We thank A. F. F. and H. T. R. for HRMS measurements. M. P. C. for MALDI experiments and the staff of the beamline X06SA at the Paul Scherrer Institute, SLS, Villigen (Switzerland) for assistance during data collection. We are grateful for financial support by the SFB749 (M.G.) and the International Leibniz Research School for Microbial and Biomolecular Interactions (ILRS, to S.S.).

Author contributions
All authors involved in the design of the research; S.K., M.B., S.S., M.I.-F. and Z.X. performed biochemical experiments, M.B. and P.B. performed bioinformatic analyses; P.B. and M.G. performed protein crystallography and analysed data; P.B. performed docking simulations; all authors contributed to the early drafts of the manuscript; C.H. wrote the final version.

Additional information
Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Kugel, S. et al. Cryptic indole hydroxylation by a non-canonical terpenoid cyclase parallels bacterial xenobiotic detoxification. Nat. Commun. 8, 15804 doi: 10.1038/ncomms15804 (2017).

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