Published in final edited form as:

Nature. 2009 October 22; 461(7267): 1139–1143. doi:10.1038/nature08475.

Structural basis for biosynthetic programming of fungal aromatic polyketide cyclization

Jason M. Crawford^{1,*}, Tyler P. Korman^{2,*}, Jason W. Labonte¹, Anna L. Vagstad¹, Eric A. Hill¹, Oliver Kamari-Bidkorphe², Shiou-Chuan Tsai², and Craig A. Townsend¹

¹Department of Chemistry, Johns Hopkins University, Maryland 21218, USA.

²Department of Molecular Biology and Biochemistry, Department of Chemistry, University of California, Irvine, California 92697, USA.

Abstract

Polyketides are a class of natural products with diverse structures and biological activities. The structural variability of aromatic products of fungal nonreducing, multidomain iterative polyketide synthases (NR-PKS group of IPKs) results from regiospecific cyclizations of reactive poly- β -keto intermediates^{1–3}. How poly- β -keto species are synthesized and stabilized, how their chain lengths are determined, and, in particular, how specific cyclization patterns are controlled have been largely inaccessible and functionally unknown until recently⁴. A product template (PT) domain is responsible for controlling specific aldol cyclization and aromatization of these mature polyketide precursors, but the mechanistic basis is unknown. Here we present the 1.8 Å crystal structure and mutational studies of a dissected PT monodomain from PksA, the NR-PKS that initiates the biosynthesis of the potent hepatocarcinogen aflatoxin B₁ in *Aspergillus parasiticus*. Despite having minimal sequence similarity to known enzymes, the structure displays a distinct ‘double hot dog’ (DHD) fold. Co-crystal structures with palmitate or a bicyclic substrate mimic illustrate that PT can bind both linear and bicyclic polyketides. Docking and mutagenesis studies reveal residues important for substrate binding and catalysis, and identify a phosphopantetheine localization channel and a deep two-part interior binding pocket and reaction chamber. Sequence similarity and extensive conservation of active site residues in PT domains suggest that the mechanistic insights gleaned from these studies will prove general for this class of IPKs, and lay a foundation for defining the molecular rules controlling NR-PKS cyclization specificity.

Aflatoxin B₁ (Fig. 1a, **3**) biosynthesis is initiated by the NR-PKS PksA, which accepts a hexanoyl starter unit from a dedicated fungal fatty acid synthase (FAS) and extends it through seven iterative, malonyl-derived ketide extensions to norsolorinic acid anthrone ((**1**), noranthrone) (Fig. 1a)⁵. Application of the Udwy–Merski algorithm (UMA) afforded the unanticipated insight that PksA was composed of two unrecognized domains, in addition

Correspondence and requests for materials should be addressed to C.A.T. (ctownsend@jhu.edu) or S.-C.T. (sctai@uci.edu).

*These authors contributed equally to this work.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Author Contributions J.M.C. carried out biochemical experiments and provided recombinant proteins. A.L.V. carried out all mutational studies. T.P.K. assisted by O.K.-B. determined the PT X-ray crystal structures. E.A.H. prepared substrate and intermediate analogues for co-crystallization experiments. J.W.L. and T.P.K. conducted *in silico* docking studies. J.M.C., T.P.K. and J.W.L. analysed data and contributed to the writing of the paper, and J.W.L. refined the proposed PT mechanism. S.-C.T. and C.A.T. directed the research, provided funding and edited the final manuscript.

Author Information The atomic coordinates of palmitate-bound and HC8-bound PT have been deposited in the Protein Data Bank under accession code 3HRQ and 3HRR. Reprints and permissions information is available at www.nature.com/reprints.

to the known β -ketoacyl synthase (KS), malonyl-CoA:ACP transacylase (MAT), acyl-carrier protein (ACP) and thioesterase (TE) domains⁶. Deconstruction experiments on PksA established that the two domains with previously unknown function were a starter unit:ACP transacylase (SAT)⁷ and a product template (PT)⁴. The PT domain was necessary for binding the fixed-length linear intermediate **6** and catalysing the efficient, stepwise aldol cyclizations (**8** and **10**) and their dehydrations (**7** and **9**) to yield the bicyclic ACP-thioester intermediate **7**, which undergoes final Claisen/Dieckmann cyclization in the TE domain to release noranthrone (**1**) (Fig. 1a, c)⁴. When the TE domain was absent, naphthopyrone (**4**) was produced by self-condensation of the stalled bicyclic PT product **7** (Fig. 1b).

We crystallized PT in $P2_12_12_1$ and $P4_12_12$ space groups (Supplementary Table 1) and solved the PT crystal structure by multi-wavelength anomalous dispersion (MAD) using selenomethionine-derivatized PT followed by phase extension with the native data to 1.8 Å. The crystal structures from both crystal forms showed distinct electron density for palmitate (from *Escherichia coli* purification), which could be removed using lipid-binding resin followed by immediate incubation with the product analogue HC8 (**5**, Fig. 1b) with PT. The HC8-bound PT structure was solved by molecular replacement using the palmitate-bound structure to 1.9 Å. The PT domain (~39kDa) was isolated as a homodimer in solution and exists as a dimer in the crystal structure (Fig. 2a).

The PT primary sequence bears little similarity to any functionally characterized protein, except similarly placed domains in other fungal NR-PKSs. However, the PT crystal structure displays a DHD fold (Fig. 2a)⁸ that is a variant of those observed in the dehydrase (DH) domains of animal FASs and bacterial modular PKSs (Fig. 2b)^{9,10}. Two DHD monomers associate via a PT-specific sequence insertion (Supplementary Figs 1 and 2), leading to a molecule with dimensions 116×53×32 Å. On the basis of sequence similarity, the DHD fold and dimer interface are conserved in the PT domains from fungal NR-PKSs.

The palmitate- and HC8-bound PT structures identify a deep pocket that extends 30 Å from the surface to the bottom (Fig. 3a) with well defined electron density for palmitate (PLM, Fig. 3b) or HC8 (**5**, Fig. 3c). The two crystal structures have a root mean squared deviation of 0.38 Å, indicating that little conformational change occurs in the rigid PT binding pocket in the presence of either linear or bicyclic compounds (except for Asn 1568, see later). The internal pocket can be visually divided into three regions (Fig. 3a): (1) the phosphopantetheine (PPT)-binding region extends 14 Å from the protein surface into the pocket. While the downstream ACP domain is transiently docked to the PT surface, this region is proposed to bind the ~18 Å PPT arm that delivers the substrate **6** into the pocket. (2) The cyclization chamber (8×13.5 Å) can accommodate two aromatic rings (Fig. 3a, c) and contains the proposed catalytic dyad (Asp 1543/His 1345) positioned appropriately to initiate regiospecific cyclizations. The PTHC8-bound structure positioned HC8 near this chamber (detailed in Supplementary Information). The central location of the cyclization chamber suggests that the substrate binds in an extended conformation with a pre-cyclization 'kink' in the middle of the polyketide (Fig 3a and Fig 4). Notably, hydrophilic residues and a network of crystallographic water molecules constitute a distinctly 'wet' side to the reaction chamber. In contrast, the orthogonal pocket proposed to accommodate the aldol cyclization products is largely hydrophobic and 'dry'. (3) The hydrophobic hexyl-binding region (6×6 Å) lies at the bottom and is perfectly adapted to accept the substrate hexanoyl starter unit. The tail of the 16-carbon palmitate is positioned directly in this region, exemplifying its tight aliphatic binding capabilities. In PTs that do not bind a hexyl group, Gly 1491 at the bottom of the hexyl-binding region becomes a bulky residue and closes off the pocket (Supplementary Fig. 1). The above observations strongly support the view that the hexyl- and PPT-binding regions anchor the ends of the extended ACP-linked poly- β -keto substrate, and the cyclization chamber accommodates two regiospecific cyclizations.

The PT active site has three features distinct from other DHDs: (1) the catalytic histidine (His 1345) uses its N δ instead of N ϵ as the basic nitrogen (Fig. 4 and Supplementary Fig. 3). (2) The catalytic Asp 1543 polarizes His 1345. In DHs and hydratases, the active site His is polarized by a neighbouring backbone carbonyl oxygen, not by the Asp1112. (3) The oxyanion hole exists in an altered location compared to known DHDs, where the second half of the oxyanion hole typically is the backbone N–H of a glycyl residue before the ‘hot dog’ helix1112. This glycine is replaced by a proline (Pro 1355) in PksA—the only residue incapable of providing an amide hydrogen bond. These differences suggest that PksA PT has evolved substantially from a simple DH or hydratase.

The palmitate- and HC8-bound structures, docking simulations (detailed in Supplementary Information) and comparisons to related enzymes support the following proposed mechanism for PT-catalysed first- and second-ring cyclization and aromatization (Fig. 4 and Supplementary Fig. 3): the ACP-bound linear substrate **6** binds to the PT substrate pocket in an extended conformation, as exemplified in the palmitate-bound structure and simulations. The lack of a salt bridge between the carboxylate of palmitate and the catalytic His implies that His 1345 is present as histidine (not histidinium) and functions as the key catalytic base. Acting together with Asp 1543 (hydrogen-bonded to, and oriented by, Gln 1547), His 1345 deprotonates C4, leading to the enolate intermediate stabilized by the backbone N–H of N1568. Collapse of the enolate and subsequent C4 aldol addition to C9 carbonyl (C4–C9 cyclization event) leads to the first ring cyclization product **8**, with the network of water molecules bound to Ser 1356, Asp 1543, Asn 1568 and Thr 1546 stabilizing the oxyanion (altered PT oxyanion site). The extent of enolization in β -diketones is known to be highly dependent on the polarity of the surrounding medium¹³. 2,4-Pentanedione, for example, exists in its diketo form 5 or 6:1 in water, but is almost entirely enolized in cyclohexane (97–98%)¹⁴. The role of the ‘wet’ side of the reaction chamber becomes evident with its water network ensuring that the C9 carbonyl is both in its electrophilic keto tautomeric form and polarized through hydrogen bonding to activate it for intramolecular aldol cyclization, whereas the conserved Asn 1554 could anchor C11 or C13 by hydrogen bonding. First ring aromatization could occur with a loss of hydroxide (E1cb) similar to scytalone dehydrase¹⁵ and enoyl-CoA dehydrase¹⁶, but with a larger thermodynamic driving force. The first aromatized ring then translocates into the cyclization chamber, and the second ring (C2–C11) cyclization and aromatization occurs analogously to the first ring cyclization event. We take the view exemplified by recent experiments with *Mycobacterium tuberculosis* FabH that product release and substrate binding occur in an ‘open’ form of the protein, whereas catalysis takes place in the ‘closed’ form apparent in the crystal structures¹⁷.

The importance of the proposed active site residues was verified by independently mutating His 1345, Asp 1543, Gln 1547 and Asn 1554 to Ala, which resulted in no detectable activity (detailed in Supplementary Discussion and Supplementary Figs 3 and 4). The Thr 1546 backbone amide carbonyl and side-chain hydroxyl are each hydrogen bonded to one of the two ordered water molecules proposed to stabilize the substrate electrophilic centre of the aldol cyclization reactions (Fig. 4). Mutation of this residue to Ala disrupts one of these interactions and activity falls to approximately 40%. In the HC8-bound structure, the Asn 1568 side-chain amide is hydrogen bonded to one of these water molecules, whereas in the palmitate structure it moves to hydrogen bond to the fatty acid carboxylate—but the ordered water remains in place. Here mutation to Ala slightly raises production of the PT-released product **4** (approximately 120%), presumably by reducing steric constraints in the active site where substrate movement is occurring during the cyclization steps. Furthermore, Gly1491Leu represents a bulky mutation to block the hexyl-binding pocket, which also resulted in no apparent activity. Sequence alignments suggest that the catalytic mechanism

for PT domains in fungal PKSs is conserved (Supplementary Fig. 1). As expected, the largest variations lie within the cyclization chamber in keeping with their proposed function.

To evaluate the role of PT in the global domain organization of type I nonreducing IPKSs, the oligomerization states of independent mono-, di- and tridomains were determined by native PAGE (detailed in Supplementary Table 2), and the results support a structural organization of PksA with similarity to animal FAS⁹ (Fig. 2b). The above analyses coupled with previous phylogenetic results¹⁸ suggest that the dimeric PT domain evolved from ancient DH domains in reducing PKSs and would be positioned at the DH locus in animal FAS.

Two distinct cyclization patterns for aromatic polyketides from fungi and *Streptomyces* have been identified². The S-pattern, in which the first ring is cyclized between C7–C12 or C9–C14, is observed in type II PKSs from *Streptomyces*. Alternatively, the F-pattern denotes first ring cyclization between C4–C9⁴, C2–C7¹⁹, or C6–C11², which is observed exclusively in fungal aromatic IPKSs (Supplementary Fig. 5). Recently, the molecular basis of the S-folded cyclization pattern (C9–C14) was elucidated by solving the tetracenomycin (Tcm) ARO/CYC crystal structure (Fig. 3d) and conducting pocket residue mutagenesis²⁰. In the present work, we have shown that the PT domain is capable of driving first- and second-ring cyclization from a preassembled, fixed-length poly- β -keto chain to promote the F-folded cyclization pattern. The Tcm ARO/CYC structure exhibits a helix-grip fold (Supplementary Fig. 2)²⁰ that has topology similar to the DHD fold of PT. The crystal structures, mutagenesis experiments and earlier mass spectrometric results⁴ support the view that PT binds a fully extended ACP-bound poly- β -keto intermediate ‘kinked’ into the cyclization chamber, and catalyses formation of the F-folded cyclization pattern (Fig. 4; a rationale for the observed F-folded patterns is detailed in the Supplementary Information and illustrated in Supplementary Fig. 6). In contrast, the ARO/CYC probably bends the extended poly- β -keto intermediate into a hairpin and promotes closure to the S-folded pattern (Supplementary Fig. 6)²⁰. The high-resolution characterizations of a fungal PT and a bacterial ARO/CYC provide detailed views of these two engines of selective polyketide cyclization and aromatization and will guide the rational alteration of these cyclization domains to new synthetic tasks.

METHODS SUMMARY

All proteins used in the work were overproduced in *E. coli* BL21(DE3) or Rosetta2(DE3) (Novagen), and purified by Ni-affinity followed by FPLC chromatography. PT was crystallized at room temperature by sitting-drop vapour diffusion in 0.22M ammonium acetate (or 0.23M sodium citrate) and 20–22% PEG3350. Data were collected on beamline 9-1 at the Stanford Synchrotron Radiation Laboratory (SSRL) and beamline 8.2.2 at the Advanced Light Source (ALS) using HKL2000. The PT crystallographic phases were determined by MAD using selenomethionine-derivatized PT. Docking of polyketide substrates and ACP with PT was performed by GOLD21 and 3D-Dock22, respectively. CHARMm minimizations were performed with the Discovery Studio (DS) 2.1 suite. PksA global architecture was re-constructed using fragments of PksA detected by native gels and size-exclusion chromatography. Site-directed mutations were performed in the PT monodomain construct pEPT2.

METHODS

Protein production and purification

The PT monodomain was amplified by PCR using primers PT5.4 (5'-CGCATATGGAAATCAAGACCACCACGAC-3') and PT3.3n (5'-

GTAAGCGGCCGCCTAATCGGCCTTTTTCAGTGTAGTCT GC-3'). The amplification product was digested with NdeI/NotI and inserted into the corresponding sites in pET28a (Novagen) to generate pEPT3.3n, a construct encoding the PT domain with a thrombin cleavable N-terminal His₆-tag. The recombinant PksA PT-ACP didomain with a C-terminal His₆-tag fusion was produced and purified as previously described⁴. The PT monodomains was produced in *E. coli* BL21(DE3) (Novagen) by induction with 1mM IPTG at OD₆₀₀ = 0.6, and incubated at 20 °C overnight. Cell pellets were lysed by sonication in 50mM potassium phosphate (pH 7.6), 10mM imidazole, 300mM NaCl, and 10% glycerol. The His₆-tagged protein was purified by Qiagen Ni-NTA resin, cleaved with bovine thrombin (10 units per 1mg sample) for 48 h at 16 °C to remove the His₆-tag, dialysed against 50mM Tris-HCl pH 7.5, and further purified over a Q-Sepharose anion exchange column (10–500mM KCl gradient). The resulting pure protein was dialysed against 20mM Tris-HCl (pH 7.5), 5% glycerol, and 2mM dithiothreitol for crystallographic studies.

Se-Met-labelled protein production

pEPT3.3n was transformed into a methionine auxotrophic strain, B834(DE3) (Novagen). Overnight cultures were carried out in standard LB medium with 25 µg ml⁻¹ kanamycin. The cells were harvested, washed once and added to 200 ml of LeMaster's medium containing 25 µg ml⁻¹ kanamycin. The culture was grown at 37 °C to exchange the Met pool with Se-Met. Aliquots were used to inoculate 4–1 L LeMaster's cultures. The cultures were grown to an OD₆₀₀ = 0.6–0.8 at 37 °C and induced with 1mM IPTG. The induction temperature was reduced to 20 °C, and cultivated for an additional 16–18 h. Protein purifications were carried out as described for nonlabelled proteins.

Crystallization, structure determination and refinement

Crystals of the selenomethionine derivative of PksA PT-ACP construct were grown by vapour diffusion in sitting drops over 500 µl of the well solution at 25 °C. Drops were generated by mixing 2 µl purified protein at 10 mg ml⁻¹ with 2 µl well solution (0.23M sodium citrate, 22% PEG 3350). Analysis of PT-ACP crystals by SDS–PAGE showed that it degraded into the PT and ACP monodomains, and the PT domain selectively crystallized as a protein degradation product. Tetragonal crystals grew overnight. Orthorhombic crystals of the PT monodomain construct (different from the degradation product) grew in a similar condition (0.22 M ammonium acetate, 20% PEG3350) overnight at 25 °C. Crystals were soaked in crystallization solution plus 30% glycerol and flash frozen in liquid nitrogen before data collection. The HC8-bound PksA PT crystals were obtained by treating the monodomain PT with Amberlite XAD-2 resin for 2 h, followed by addition of HC8 (500 µM in acetone) in a 10:1 (HC8:PT) molar ratio. Orange crystals grew overnight in the same condition as the monodomain PT. Data were collected on beamline 9-1 at the Stanford Synchrotron Radiation Laboratory (SSRL), and subsequently integrated, indexed and scaled with HKL2000. Phases were determined by MAD using ShelX23 and Solve/Resolve24. All 20 Se sites except for the two N-terminal methionines were found. Electron density maps were generated using CNS and the model was built using COOT25. Refinement of the SeMet PT model was carried out with CNS26. The native palmitate-bound PT and HC8-bound PT crystal structures were solved by molecular replacement using the SeMet PT model as the template, using CNS (native PT) and CCP4 (HC8-PT)²⁷. Waters were added in COOT followed by refinement with CNS or CCP4, and edited by visual inspection. Initial coordinates and restraints for palmitate and HC8 were generated with PRODRG28 followed by manual fitting into well-defined positive $F_o - F_c$ electron density contoured at 4σ with COOT. Supplementary Table 1 lists a summary of phasing and refinement statistics. The refined structures were evaluated with PROCHECK²⁹ before coordinate submission.

Mutagenesis, *in vitro* PksA PT assay and HPLC product analysis

Three-part multidomain combination reactions (SAT-KS-MAT+PT*+*holo*-ACP) and subsequent product analysis by HPLC were carried out as previously described⁴. Site-directed mutations (*) in the PT monodomain construct (pEPT2) were performed using overlap-extension PCR³⁰.

Native PAGE analysis

The enzymes were dialysed against 10mM Tris-HCl pH 7.5, 10% glycerol and 2mM dithiothreitol to avoid negative salt effects on native PAGE. The proteins were compared to non-denatured protein from the Sigma-Aldrich molecular weight marker kit according to the manufacturer's instructions. Association states: SAT, monomer; SAT-KS-MAT, dimer; PT, dimer; ACP, monomer; TE, monomer; PT-ACP, dimer; ACP-TE, monomer; PT-ACP-TE, dimer/monomer (Supplementary Table 2).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank T. M. Harris for his gift of HC8. The work at Johns Hopkins was supported by the US National Institutes of Health grant ES001670 awarded to C.A.T. S.-C.T. is supported by the Pew Foundation. Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory (SSRL), a national user facility operated by Stanford University on behalf of the US Department of Energy, Office of Basic Energy Sciences. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under Contract No. DE-AC02-05CH11231. J.M.C. is currently a fellow supported by the Damon Runyon Cancer Research Foundation (DRG-2002-09, Harvard Medical School).

References

1. Staunton J, Weissman KJ. Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep* 2001;18:380–416. [PubMed: 11548049]
2. Thomas R. A biosynthetic classification of fungal and *Streptomyces* fused-ring aromatic polyketides. *ChemBioChem* 2001;2:612–627. [PubMed: 11828498]
3. Hertweck C. The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed* 2009;48:4688–4716.
4. Crawford JM, et al. Deconstruction of iterative multidomain polyketide synthase function. *Science* 2008;320:243–246. [PubMed: 18403714]
5. Townsend, CA.; Minto, RE. *Comprehensive Natural Products*. Elsevier; 1999. p. 443-471.
6. Udvary DW, Merski M, Townsend CA. Method for prediction of the locations of linker regions within large multifunctional proteins, and application to a type I polyketide synthase. *J. Mol. Biol* 2002;323:585–598. [PubMed: 12381311]
7. Crawford JM, Dancy BCR, Hill EA, Udvary DW, Townsend CA. Identification of a starter unit acyl-carrier protein transacylase domain in an iterative type I polyketide synthase. *Proc. Natl Acad. Sci. USA* 2006;103:16728–16733. [PubMed: 17071746]
8. Dillon SC, Bateman A. The Hotdog fold: wrapping up a superfamily of thioesterases and dehydratases. *BMC Bioinformatics* 2004;5:109. [PubMed: 15307895]
9. Maier T, Leibundgut M, Ban N. The crystal structure of a mammalian fatty acid synthase. *Science* 2008;321:1315–1322. [PubMed: 18772430]
10. Keatinge-Clay A. Crystal structure of the erythromycin polyketide synthase dehydratase. *J. Mol. Biol* 2008;384:941–953. [PubMed: 18952099]
11. Koski KM, Haapalainen AM, Hiltunen JK, Glumoff T. Crystal structure of 2-enoyl-CoA hydratase 2 from human peroxisomal multifunctional enzyme type 2. *J. Mol. Biol* 2005;345:1157–1169. [PubMed: 15644212]

12. Kimber MS, et al. The structure of (3*R*)-hydroxyacyl-acyl carrier protein dehydratase (FabZ) from *Pseudomonas aeruginosa*. *J. Biol. Chem* 2004;279:52593–52602. [PubMed: 15371447]
13. Hibbert F, Emsley J. Hydrogen bonding and chemical reactivity. *Adv. Phys. Org. Chem* 1991;26:255–379.
14. Mills SG, Beak P. Solvent effects on keto-enol equilibria: tests of quantitative models. *J. Org. Chem* 1985;50:1216–1224.
15. Jordan DB, Zheng YJ, Lockett BA, Basarab GS. Stereochemistry of the enolization of scytalone by scytalone dehydratase. *Biochemistry* 2000;39:2276–2282. [PubMed: 10694394]
16. Bahnson BJ, Anderson VE, Petsko GA. Structural mechanism of enoyl-CoA hydratase: three atoms from a single water are added in either an E1cb stepwise or concerted fashion. *Biochemistry* 2002;41:2621–2629. [PubMed: 11851409]
17. Sachdeva S, et al. Separate entrance and exit portals for ligand traffic in *Mycobacterium tuberculosis* FabH. *Chem. Biol* 2008;15:402–412. [PubMed: 18420147]
18. Kroken S, Glass NL, Taylor JW, Yoder OC, Turgeon BG. Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic *Ascomycetes*. *Proc. Natl Acad. Sci. USA* 2003;100:15670–15675. [PubMed: 14676319]
19. Ma SM, et al. Redirecting the cyclization steps of fungal polyketide synthase. *J. Am. Chem. Soc* 2008;130:38–39. [PubMed: 18072779]
20. Ames BD, et al. Crystal structure and functional analysis of tetracenomycin ARO/CYC: implications for cyclization specificity of aromatic polyketides. *Proc. Natl Acad. Sci. USA* 2008;105:5349–5354. [PubMed: 18388203]
21. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD. Improved protein-ligand docking using GOLD. *Proteins* 2003;52:609–623. [PubMed: 12910460]
22. Gabb HA, Jackson RM, Sternberg MJ. Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J. Mol. Biol* 1997;272:106–120. [PubMed: 9299341]
23. Sheldrick GM. A short history of SHELX. *Acta Crystallogr. A* 2008;64:112–122. [PubMed: 18156677]
24. Terwilliger TC. SOLVE and RESOLVE: automated structure solution, density modification and model building. *J. Synchrotron Radiat* 2004;11:49–52. [PubMed: 14646132]
25. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* 2004;60:2126–2132. [PubMed: 15572765]
26. Brünger AT, et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* 1998;54:905–921. [PubMed: 9757107]
27. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* 1994;50:760–763. [PubMed: 15299374]
28. Schüttelkopf AW, van Aalten DM. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D* 2004;60:1355–1363. [PubMed: 15272157]
29. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr* 1993;26:283–291.
30. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 1989;77:51–59. [PubMed: 2744487]

Figure 1. Biosynthesis of norsolorinic acid anthrone (1) by PksA

a, Native catalytic cycle. The SAT domain in PksA selects a hexanoyl starter unit. The MAT domain loads the free ACP with malonyl units. After seven successive condensation events with malonyl-ACP catalysed in KS, the linear ACP-bound polyketide **6** is cyclized (C4–C9 and C2–C11 cyclization events) and aromatized in the product template (PT) domain to give the bicyclic intermediate **7**. The TE catalyses C–C cyclization to release anthrone (**1**), which undergoes oxidation to the anthraquinone norsolorinic acid (**2**), to initiate the complex biosynthetic pathway to aflatoxin B₁ (**3**). Asterisks denote the C and O nucleophiles involved in on-path and off-path cyclizations, respectively. **b**, In the absence of TE, the PT product **7** undergoes spontaneous O–C cyclization to norpyrone (**4**). The PT monodomain structure was solved with intermediate/product mimic **5** (HC8). **c**, Conversions from **6** to **7** via intermediates **8**, **9** and **10**.

Figure 2. The crystal structure of the PT domain from PksA

PT contains a modified double hot dog (DHD) fold. **a**, Ribbon representation of the PT dimer with the interface between monomer A (green, with secondary structural elements labelled and the entrance to the pocket indicated with a red arrow) and monomer B (grey) stabilized by the elongated helices. Palmitate is shown in orange spheres. The α -helix and β -strand order is indicated on the left monomer unit. **b**, Proposed domain organization of PksA (right) and comparison to the structure of animal FAS (left). ACP, acyl-carrier protein; AT, acyl transacylase; DH, dehydrase; ER, enoyl reductase; KR, ketoreductase; KS, ketoacyl synthase; ψ MT, pseudomethyltransferase; SAT, starter unit:ACP transacylase; TE, thioesterase.

Figure 3. PT structures with bound palmitate or substrate analogue HC8

a, Surface representation of the PT monomer with the pocket shown (yellow outline). The hexyl-binding region, cyclization chamber and PPT-binding region are highlighted. **b**, The location and orientation of the hexyl-binding region is shown as a grey tube. The cyclization chamber location is indicated with a brown dashed line. The simulated annealing (SA) omit map calculated after omitting the bound palmitate is contoured at 1.3σ as blue mesh around the modelled palmitate ($F_o - F_c$). **c**, Same as **b**, except with HC8 bound. The SA omit map contoured at 1.5σ was calculated after omitting the bound HC8 and is shown as blue mesh. **d**, The bacterial Tcm ARO/CYC crystal structure reveals a reaction chamber (orange outline) shallower and broader than the fungal PksA PT that accommodates its C_{20} polyketide substrate in a 'hairpin' conformation leading to C9–C14 first ring cyclization.

Figure 4. Proposed mechanism of cyclizations catalysed by the PT domain

For clarity of presentation, liberties have been taken with the relative placement of active-site residues, but the principal interactions with the fully extended polyketide intermediate (left) are depicted, notably, to ordered crystallographic water molecules and a slight 'kink' towards the cyclization chamber. Hydrogens indicated by dotted circles form hydrogen bonds to acceptors on the catalytic Asp 1543. The His 1345/Asp 1543 catalytic dyad acts as the catalytic base, leading to enolate formation at C4 to initiate first ring cyclization (C4–C9). The second ring is formed in the same manner as the first to achieve C2–C11 closure (Supplementary Fig. 3).