Nitrogen deletion in the biosynthesis of HDAC-targeting anticancer drugs

Xinyun Jian[1,2,3,4,*], Jinlian Zhao[3], Edward M. Marschall[1,2], Douglas M. Roberts[3], Max J. Cryle[1,2,*], Lona M. Alkhalaf[3,4,*], and Gregory L. Challis[1,2,3,4,*]

[1] Biomedicine Discovery Institute, Monash University, Clayton, VIC 3800, Australia

[2] ARC Centre of Excellence for Innovations in Peptide and Protein Science, Monash University, Clayton, VIC 3800, Australia

[3] Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK

[4] Warwick Integrative Synthetic Biology Centre, University of Warwick, Coventry, CV4 7AL, UK

* Corresponding author

Abstract

Natural product biosynthetic pathways employ diverse enzymatic strategies for site-specific modifications of molecular scaffolds, tailoring their bioactivity. Here, we uncovered an unprecedented nitrogen deletion mechanism in the biosynthesis of the conserved pharmacophore of a natural depsipeptide histone deacetylase (HDAC) inhibitor family, exemplified by the clinical used anti-cancer drug romidepsin. Through combination of bioinformatics, genetic and biochemical characterization of their biosynthetic mechanism, we revealed novel catalytic functions within polyketide synthase and nonribosomal peptide synthetase assembly lines. Our findings demonstrate that a pair of multifunctional ketoreductase and dehydratase domains, coupled with a *trans*-acting phosphotransferase and a *trans*-acting flavin-dependent oxidoreductase, coordinate a series of transformations to remove the amino group introduced by the cysteine moiety in the pharmacophore of this HDAC inhibitor family. Deletion of this protruding amino group results in the formation of an aliphatic linker within the pharmacophore, mimicking the natural substrate of HDACs and conferring superior HDAC inhibitory activity of these compounds. Furthermore, we identified a novel condensation domain mediated a cryptic *S*-acylating protection mechanism in the biosynthesis of the conserved pharmacophore. Leveraging these insights, we established a highly efficient, one-pot enzymatic synthesis of the intact pharmacophore in its protected form, laying the groundwork for its scalable biocatalytic production. Our work expanded the mechanistic understanding of natural product site-specific editing enabled by novel functions of assembly line biosynthetic machineries. It also showcases how nature selectively sculpts molecular scaffolds to fine-tune bioactivity and diversify chemical space, offering new avenues for biocatalytic strategies in the development of this important family of HDAC inhibitors.

Main

Site-specific C–H bond functionalization, such as the addition of oxygen and nitrogen, is ubiquitous in the biosynthesis of natural products (NPs). Nature has evolved numerous enzymes capable of performing these reactions under mild condition, with extraordinary regio- and stereo-control. An example of oxygen insertion occurs in the biosynthesis of the potent β-lactamase inhibitor, Calvulanic acid, where a 2-oxoglutarate-dependent dioxygenase catalyses a site- and stereo-specific hydroxylation on the deoxyguanidinoproclavaminate intermediate (**Fig. 1a**).[1] One notable example of enzymatic nitrogen insertion is an unprecedented regioselective aromatic nitration catalyzed by the cytochrome P450 enzyme TxtE, in the biosynthesis of the antibiotic thaxtomin A.[2] Replicating such precise and efficient modifications through traditional organic synthesis remains challenging. [3] Consequently, over the past 30 years biocatalysts have played an increasingly important role in the production of many complex fine chemicals, particularly in the pharmaceutical industry.[4, 5]

Although site-specific removal of functional groups provides an effective strategy for molecular editing, this area remains relatively underexplored. In 2021, Levin and colleagues developed an innovative synthetic strategy to construct carbon-carbon bonds through skeletal nitrogen deletion. This approach expands the toolkit for molecular editing, enabling the modification of molecular frameworks by excising nitrogen atoms and forming new C–C linkages.[6] There are some examples of oxygen deletion in NP biosynthesis. For instance, polyketide synthases employ a multienzyme cascade to effectively remove oxygen atom in the polyketide chain through a serious stereo-controlled transformations, including ketoreduction, water elimination and enoyl reduction (**Fig. 1a**). Other examples of oxygen deletion include a two enzyme -cascade catalysed regioselective C3 deoxygenation of NDP-sugars.[7] However, removing of nitrogen atom in NPs has only been observed in the ammonium lyase-catalyzed reversible cleavage of C–N bonds in α-amino acids, which typically results in the release of ammonia and an unsaturated or cyclic derivative, rather than a direct amine removal product.[8]

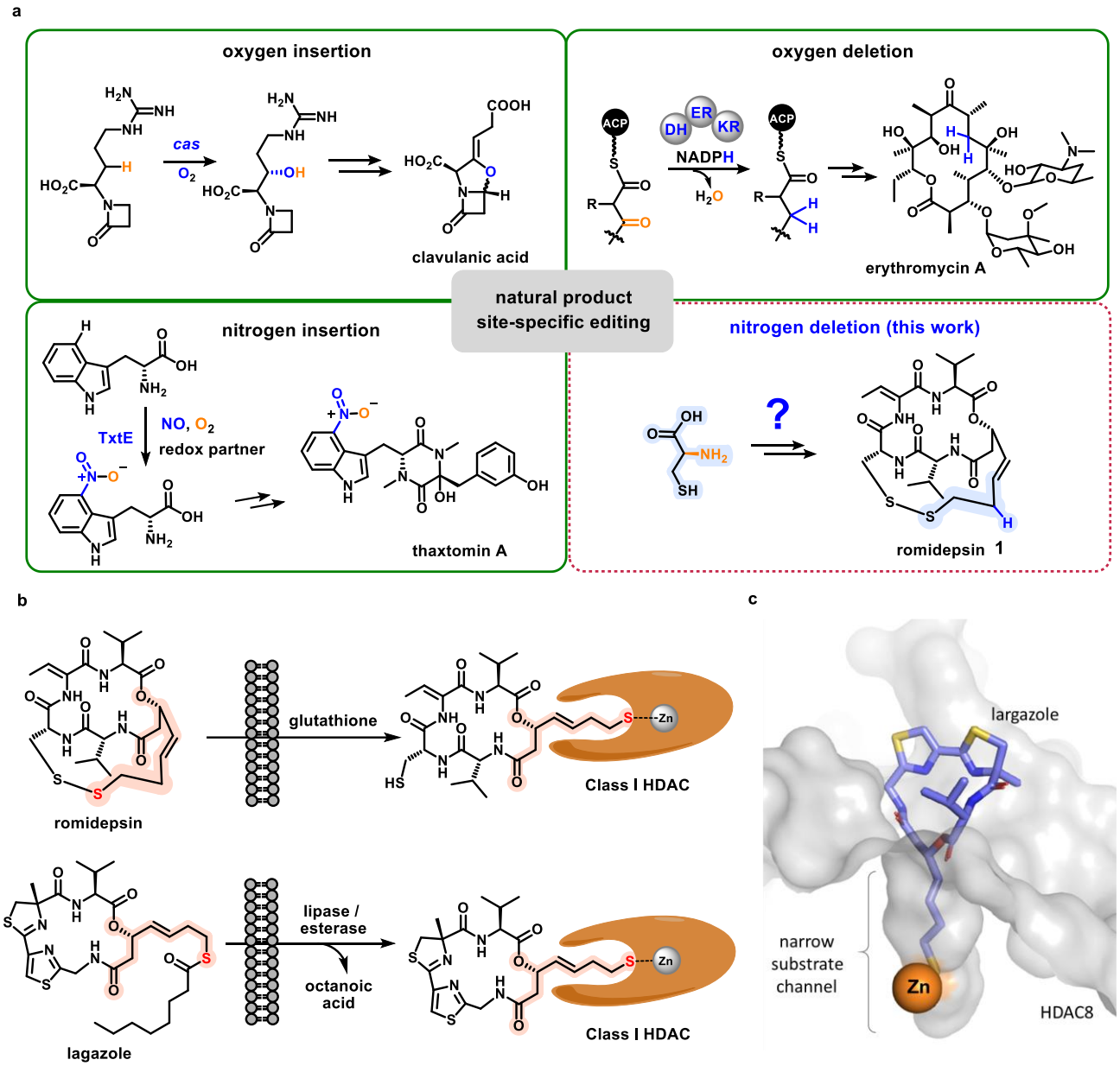


Figure 1 Site-specific editing in natural products biosynthesis and mode of action of natural cyclic depsipeptide HDACis. a, Examples of site-specific editing in natural product biosynthesis, oxygen insertion: clavaminate synthase, Cas catalysed hydroxylation on deoxyguanidinoproclavaminate in clavulanic acid biosynthesis; oxygen deletion: KR, DH and ER domains catalysed keto-reduction-dehydration-enoylreduction cascade reactions on β-keto thioester in erythromycin A biosynthesis; nitrogen insertion: cytochrome P450, TextE catalysed L-Trytophan nitration in thaxtomin A biosynthesis; nitrogen deletion: unknown mechanism involved in hydro-deamination in romidepsin biosynthesis (this work). **b**, Mode of action of natural cyclic depsipeptide HDACis. They undergo either intracellular reduction of disulfide (top) by glutathione or acyl hydrolysis (bottom) by lipase/esterase upon cell uptake to release a free thiol (in red) in the conserved pharmacophore (highlighted in orange), which chelates the Zn ion in the HDAC catalytic pocket. **c**, X-ray crystal structure of largazole-HDAC8 complex (3RQD) shows the pharmacophore extending into the narrow, hydrophobic substrate channel of HDAC8 while the peptide cap interacts with the top rim of the pocket.[18]

Histone deacetylases (HDACs) are epigenetic regulators that play a vital role in modulating chromatin conformation, gene expression and DNA repair by regulating the acetylation levels of histone proteins (**Extended Data Fig. 1A**). Aberrantly high expression and activity of HDACs have been found in many different cancer cells.[9] As a result, HDACs have emerged as promising targets in anti-cancer therapeutics.[10] There are 18 mammalian HDAC isomers that are divided into four classes based on structural homology, with three classes (class I, II and IV) are zinc-dependent, while class III is NADH-dependent.[10] Three families of zinc-dependent HDAC inhibiting anti-cancer drugs have been approved by FDA, including pan-inhibitory hydroxamic acids, and class I-selective benzamides and romidepsin **1** (**Extended Data Fig. 1 b &c**).[10]

Romidepsin, isolated from the soil bacterium, *Chromobacterium violaceum,*[11] is a member of a family of natural occurring cyclic depsipeptide HDACis, which are all potent inhibitors of class I HDACs, including burkholdacs, spiruchostatins, FR901375 and largazole (**Extended Data Fig. 1 b & c**).[12,13,14,15] They contain a conserved (*3S*,*4E*)-3-hydroxy-7-mercapto-4-heptenoic acid pharmacophore connected to a variable peptide cap. Variations in the amino acid composition of their peptidyl cap have been attributed to confer their different HDAC isomer selectivity. These HDACis are all produced as prodrugs by the bacteria, with the pharmacophore terminal thiol protected either as a disulfide or thioester. Intracellular disulfide reduction or acyl hydrolysis releases the pharmacophore thiol, which coordinates to the catalytically essential Zn^{2+} ion in the active site of HDACs. (**Fig. 1 b & c**).[16,17, 18] The narrow and hydrophobic cavity above Zn^{2+} ion in the substrate channel of Class I HDACs imposes significant structural constraints on the linker between the zinc-binding thiol and cap region.[19] The ultra-conserved aliphatic linker in the pharmacophore of this family of HDACis mimics the natural substrate, acetyl-lysine residue, explaining their exceptional class I HDAC inhibition activity.

PKSs and NRPSs are remarkable molecular factories which use simple building blocks (acyl-CoAs for PKSs, amino acids for NRPSs) to produce structural diverse and complex polyketide and peptide products with a wide range of pharmaceutical and agricultural applications. They organise repetitive catalytic domains into successive modules, with each module adding one monomer, acyl unit for PKS and amino acid for NRPS, to the growing chain that covalently tethered and shuttled along the assembly line.[20] Canonical transformations within a typical PKS module include an acyltransferase (AT) domain selectively loading an acyl-CoA extender unit onto a shuttler domain, acyl carrier protein (ACP) domain; followed by a ketosynthase (KS) domain catalysing chain elongation via decarboxylative condensation of the extender unit and the polyketide chain attached to the

upstream ACP domain; and optional domains including ketoreductase (KR) domain, dehydratase (DH) domain and enoylreductase (ER) domain further modifying the nascent β-keto group into β-hydroxyl, α, β-alkene and α, β-alkyl group, respectively, or methyltransferase (MT) domain installing a α-C-methyl group.[21] Accordingly, a typical NRPS module has an adenylation (A) domain selectively loading an amino acid onto peptidyl carrier protein (PCP) domain; followed by condensation (C) domain catalysing peptide bond formation between the loaded amino acid and the upstream peptidyl-PCP thioester, alternatively a cyclization (Cy) domain replacing a C domain to instal a thiazoline or an oxazoline ring via condensation and cyclodehydration. Optional modifications in NRPS include epimerization (E) domain altering the α-carbon stereochemistry and C- or N-MT domain incorporating a α-C-methyl or N-methyl group.[22]

The biosynthetic gene clusters (BGCs) of bicyclic depsipeptide HDACis (**1**, **2a**/**b**, **3a**/**b** and **4**) have been identified, encoding highly-related hybrid assembly lines consisting of polyketide synthases (PKSs) and non-ribosomal peptide synthetases (NRPSs), regulators and a conserved set of tailoring enzymes, including, a flavin-dependent oxidase (FDO, DepF and homologs), a phosphotransferase (PT, DepG and homologs), a flavin-dependent oxidoreductase (OR, DepH and homologs) and a hydrolase/deacylase (HYD, DepI and homologs) (**Fig. 2a**). [12,23, 24,25]

It has been hypothesized that the biosynthesis of the bicyclic depsipeptide HDACis starts with assembly of the conserved pharmacophore by the ultra-conserved PKS and NRPS subunits (DepA-C and homologs), using one unit of L-cysteine and two units of malonyl-coenzyme A (**Fig. 2b**).[23] An unprecedented deletion of the cysteine α-amino group on the pharmacophore was proposed to occur during the biosynthesis. This represents the first example of hydro-deamination in the natural product biosynthesis. However, the specific underlying mechanism remains elusive, despite the significant potential of these molecules. Given the likely conserved biosynthetic pathway for the pharmacophore across all four members of this compound family, it is proposed that the conserved set of tailoring enzymes play a key role in its biosynthesis (**Fig. 2b**).

Following the conserved pharmacophore assembly, their biosynthetic pathways diverge, incorporating different building blocks to form the distinctive peptide chain by the downstream NRPSs or hybrid NRPS-PKS subunits. (**Extended Data Fig.2**). Our previous work has characterized a short linear motif – β-hairpin docking domain (SLiM-βHD domain) pair at the interface between the pharmacophore and peptide cap assembling subunits, mediating the subunits interaction at the branching point of their biosynthesis. Once the linear chain is full assembled and released from the

assembly line their biosynthetic pathways converge again to install the prodrug mechanism, forming the disulfide bridge between the pharmacophore thiol and peptide cysteine thiol.[26]

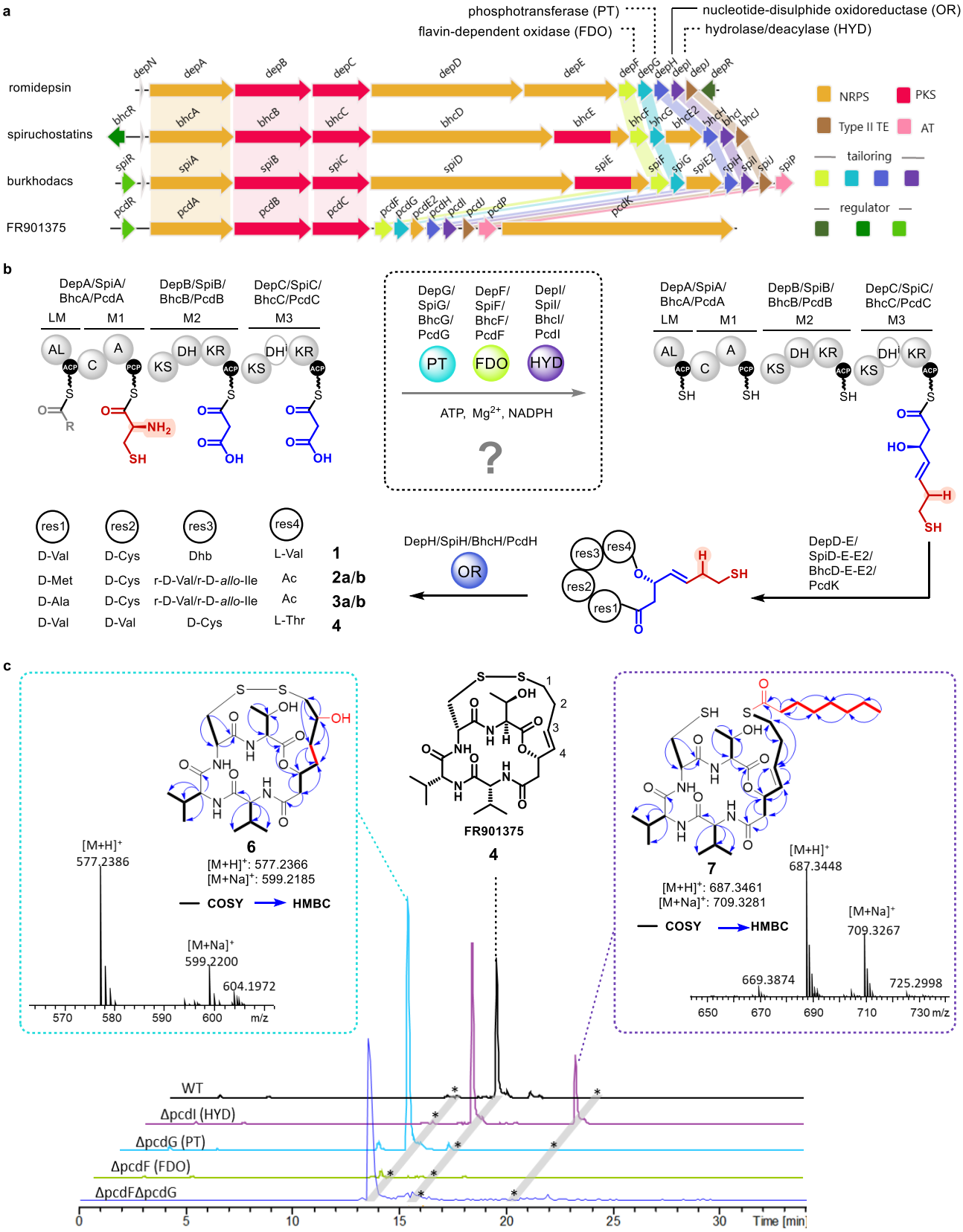


Figure 2 Romidepsin-like HDACis biosynthetic gene clusters and genetic characterization of their conserved tailoring gene. a, Comparison of biosynthetic gene clusters of romidepsin, spiruchostatins, burkholdacs, and FR901375 shows highly conserved PKS-NRPS encoding genes (A-B-C) and a set of tailoring enzymes (F-G-H-I-J) with function of DepH has been characterized while the rest of three with predicted functions shown. **b**, Proposed biosynthetic assembly lines of **1**, **2a/b**, **3a/b** and **4**. Predicted building blocks for pharmacophore are shown bound to the carrier protein in the conserved first four modules. The three tailoring enzymes, PT, FDO and HYD are proposed to be involved in pharmacophore biosynthesis, including a key step of hydrodeamination of cysteine α-NH_2. Different peptides are incorporated by respective downstream NRPS-PKS assembly lines followed by disulfide formation by OR. **c,** UHPLC-ESI-QTOF-HRMS and NMR analysis of shunt metabolites from in-frame gene deletion mutants of FR901375 (**4**) producer, *Pseudomonas chlororaphis* DSM21509. Extracted ion chromatograms at *m/z* of 559.2260 ± 0.002 (**4**), 577.2366± 0.002 (**6**) and 687.3448 ± 0.002 (**7**) show ΔpcdG, ΔpcdF and ΔpcdGΔpcdF abolish parental product **4** production while ΔpcdG and ΔpcdGΔpcdF accumulate shunt metabolite **6**. ΔpcdI produces **7** other then **4.**

Facilitated by advances in bioinformatics and computational approaches over the last two decades, the PKS and NRPS paradigms are largely expanded through the discoveries of non-canonical catalytic domains, and novel modifications by canonical catalytic domains and standalone enzymes *trans*-acting on the assembly lines.[27] Yet Nature has evolved myriads of ingenious ways to incorporate structure diversity, the complexity and versatility of these biosynthetic systems are still vastly underexplored. This work presents the first example of several such unprecedented mechanisms employed by bacteria to delete a nitrogen in the conserved pharmacophore of romidepsin-like HDAC inhibitors, addressing the long-standing question of how their pharmacophores are biosynthesised. Deleting the protruding amino group leads to the aliphatic linker in the pharmacophore of this family of HDACis, which closely mimics the natural substrates of HDACs. Characterization of this novel nitrogen deletion mechanism expanded our understanding of site-specific modifications in natural product biosynthesis, highlighting how bacteria selectively sculpt molecular scaffolds to fine-tune bioactivity and diversify chemical space.

Genetic characterization of tailoring enzymes

We first investigated the role of the conserved tailoring genes through a genetic characterization approach. The oxidoreductase (OR) DepH has previously been shown to catalyse formation of the disulfide of romidepsin so we focused on the genes encoding the other three tailoring enzymes.[26] An in-frame deletion was introduced in *pcdF* (FDO), *pcdG* (PT) and *pcdI* (HYD) in the FR901375 (**4)** producer, *Pseudomonas chlororaphis* DSM21509. UHPLC-ESI-QTOF-HRMS analysis of metabolite extracts from P.*chlororaphis* ΔpcdF and P.*chlororaphis* ΔpcdG showed production of **4** was abolished.

Interestingly, P.*chlororaphis* ΔpcdG accumulated a shunt metabolite **6.** Isolation of **6**, and characterisation by NMR spectroscopy and mass spectrometry, showed it was an analogue of **4** containing a hydroxyl group at C2 and an alkene at C2-C3(**Fig. 2c**). The same metabolite profile was observed in the double deletion mutant P.*chlororaphis* ΔpcdFΔpcdG (**Fig. 2c**). This confirmed PcdF and PcdG are essential for FR901375 biosynthesis and PcdF acts downstream of PcdG in the pathway, likely involved in removing a phosphate group installed by PcdG.

UHPLC-ESI-QTOF-HRMS analysis of extracts from *P.chlororaphis*ΔpcdI showed the presence of a different shunt metabolite **7**, in which the disulfide is reduced, and the resultant pharmacophore thiol is acylated as an octanoyl thioester (**Fig. 2c**). This suggests that PcdI hydrolyzes a cryptic octanoyl moiety from the pharmacophore thiol. Given the high reactivity of the free thiol, we hypothesized that this unit serves as a protective cap, installed at the onset of biosynthesis to prevent disulfide formation on premature intermediates. It is then removed just before the PcdH catalysing disulfide formation between the pharmacophore thiol and backbone cysteine thiol as the final step of biosynthesis. If this is the case, this reveals an interesting link to the structure and prodrug mechanism of largazole (**5**), in which the pharmacophore thiol is similarly capped by an octanoyl unit. We proposed that the biosynthesis of largazole shares an evolutionary synergy with that of the other four HDACi family members, wherein the pharmacophore thiol is capped immediately upon installation and remains acylated as the protection mechanism due to the absence of a free cystine residue in the peptide backbone to form a disulfide bond.

We made in-frame deletion in the homologs of *pcdF*, *pcdG* and *pcdI* in the other bicyclic depsipeptide HDACis producers and observed similar metabolite profiles (**Extended Data Fig. 4**), confirming the common role of these tailoring enzymes in biosynthesis of the conserved pharmacophore.

Bioinformatic analysis of assembly lines

The discovery of an acylated intermediate prompted us to revisit the annotation of module 1-3 of the FR901375 PKS-NRPS. This revealed the acyl-CoA ligase (AL) domain, the C domain in module 1 were previously misannotated as inactive, even though containing all the conserved amino acids necessary for catalysis (**Extended Data Fig. 3a**). Furthermore, the presence of an ACP domain in-between the AL domain and C domain of module 1 was overlooked entirely. This domain organisation has also been found conserved in the other three HDACi assembly lines. Phylogenetic analysis of the C domains from the first module of all four HDACi assembly lines showed they form a distinct clade

closely related to Cy domains, which catalyse condensation and cyclodehydration between a Cys or Thr/Ser and an acyl unit to form a thiazoline and oxazoline, respectively (**Extended Data Fig. 3b**). Sequence alignment of the module 1 C domains and canonical Cy domains showed variations in the conserved DX_5D motif and previously reported catalytically important residues (**Extended Data Fig. 3c**).[28] Together with the observation in the genetic characterization of *pcdI*, suggesting an acyl thioester is a likely intermediate in HDACi biosynthesis, these findings led to a hypothesis that the AL domain and the newly discovered ACP domain constitute a loading module, responsible for incorporating an octanoyl unit as the starter unit, which module 1 C domain then utilizes to catalyse S-acylation of the L-Cys loaded on the cognate PCP domain.

The DH domain in the module 2 was also found to be misannotated to be inactive (**Extended Data Fig. 3a**). Our sequence analysis identified the intact catalytic His-Asp dyad at the active site, suggesting they are functional. Additionally, phylogenetic analysis of the DH domains from module 2 of all four HDACi assembly lines showed they form a single clade with closest homology to a dual functional dehydratase-enoyl-isomerase (DH^{ei}) domain $GphDH_1$ and diene synthase (DH^{d}) domains (**Extended Data Fig. 3d**).[29, 30] Previous studies have identified a characteristic substitution of proline in the HX_8P motif in enoyl-isomerizing DH domains (DH^{i} and DH^{ei}).[30] This substitution was also found in the HX_8P motif of the module 2 DH domains from all the HDACis suggesting they are likely performing enoyl isomerising in addition to their dehydration function in the pharmacophore biosynthesis.

***In vitro* reconstitution of the L-cysteine S-acylation in module 1**

To validate our hypothesis, we next attempted to reconstitute the pharmacophore biosynthesis *in vitro* using purified His_6-tagged recombinant proteins overproduced in *E. coli* (**Extended Data Fig. 5**). Due to solubility issues with constructs cloned from the FR901375 BGC, constructs used in this work were cloned from the romidepsin assembly line. This included the excised loading module of DepA, the excised module 1 of DepA, DepA, DepB (module 2), DepC (module 3), the DepG (PT) and DepF (FDO). We first employed intact protein mass spectrometry to analyze the reconstitution assays, a powerful tool for capturing transformations of intermediates, particularly unexpected ones, attached to the assembly line.[31,32,33]

The loading module and module 1 were first converted to their respective *holo* forms (RXN1 and 3) via promiscuous PPtase Sfp and coenzyme A in the presence of $MgCl_2$. Following incubation of the

holo-loading module with octanoic acid and ATP (RXN2), UHPLC-ESI-QTOF-ESI-MS analysis showed a new peak with a mass consistent with addition of an octanoyl unit onto the loading module (**Fig. 3A**). Similarly, L-Cys could be loaded onto *holo*-module 1 (RXN4) (**Fig. 3b**). Incubation of L-Cys-*holo*-module 1 with the octanoyl-*holo*-loading module resulted in transfer of the octanoyl unit from the loading module to module 1. (RXN5) (**Fig. 3b**). No further cyclisation was observed as occurs in typical Cy domains. This result is consistent with the module 1 C domain catalysing a novel *S*-acylation instead of a cyclodehydration as in canonical Cy domains. No further transformation of the S-acylated-cysteinyl thioester was observed by intact protein MS analysis when DepG, ATP and DepF were added to the reaction, suggesting DepG and DepF do not act in module 1 and deamination may occur further downstream on the assembly line.

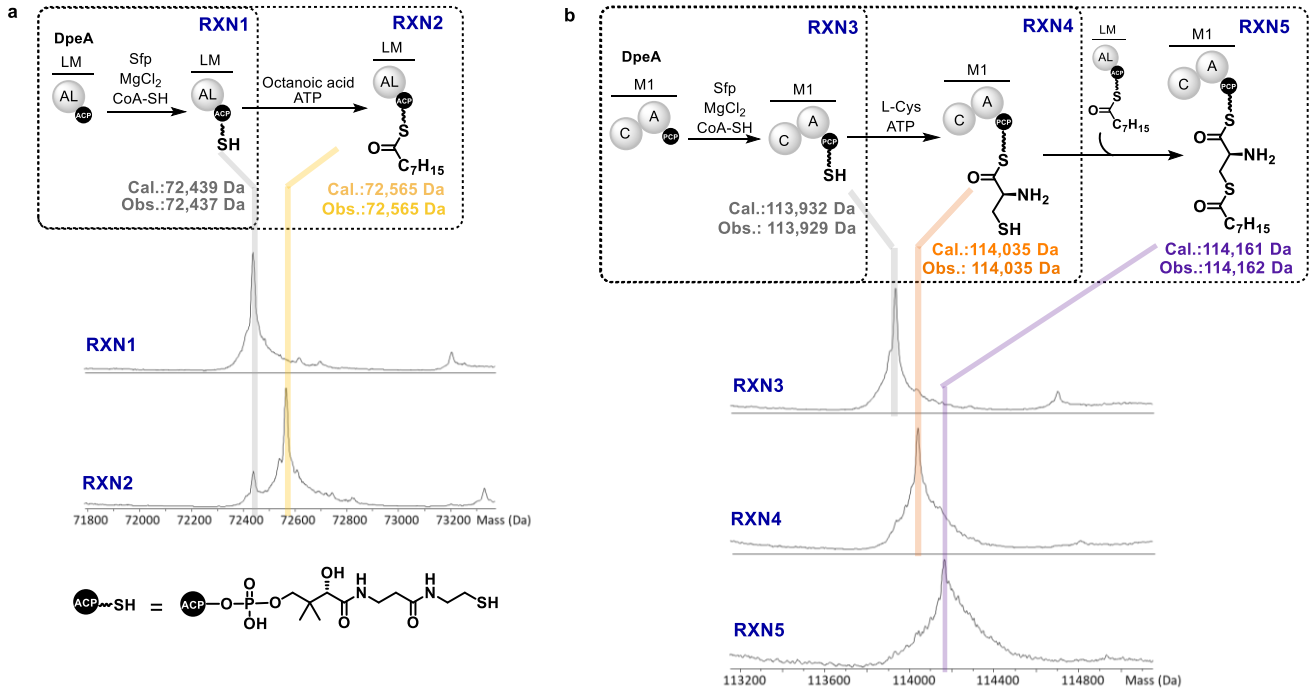


Figure 3 *In vitro* reconstitution and intact protein mass spectrometry analysis of biosynthetic steps in loading module, module 1 and module 2. a, Schematic representation of *in vitro* reconstitution of *holo* modification (RXN1) and following octanoyl acyl unit loading in loading module (RXN2). Deconvoluted mass spectra from UHPLC-ESI-QTOF-HRMS analysis of assays showing species with observed MW of 72,437 Da in RNX 1 (consistent with *holo*-loading module, calculated for 72,439 Da) and 72,565 Da in RNX2 (consistent with octanoyl-*holo*-loading module, calculated for 72,565 Da). **b**, Schematic representation of *in vitro* reconstitution of *holo* modification of module 1(RXN3), L-Cys incorporation (RXN4) and S-acylation (RXN5) in module 1. Deconvoluted mass spectra from UHPLC-ESI-QTOF-HRMS analysis of assays showing observed species with MW of 113,929 Da in RXN3 (consistent with *holo*-module 1, calculated for 113,932 Da), 114,035 Da in RXN4 (consistent with L-Cys-*holo*-module 1, calculated for 114,035 Da) and 114, 162 Da in RXN5 (consistent with S-Octanoyl-Cys-*holo*-module 1, calculated for 114,161 Da).

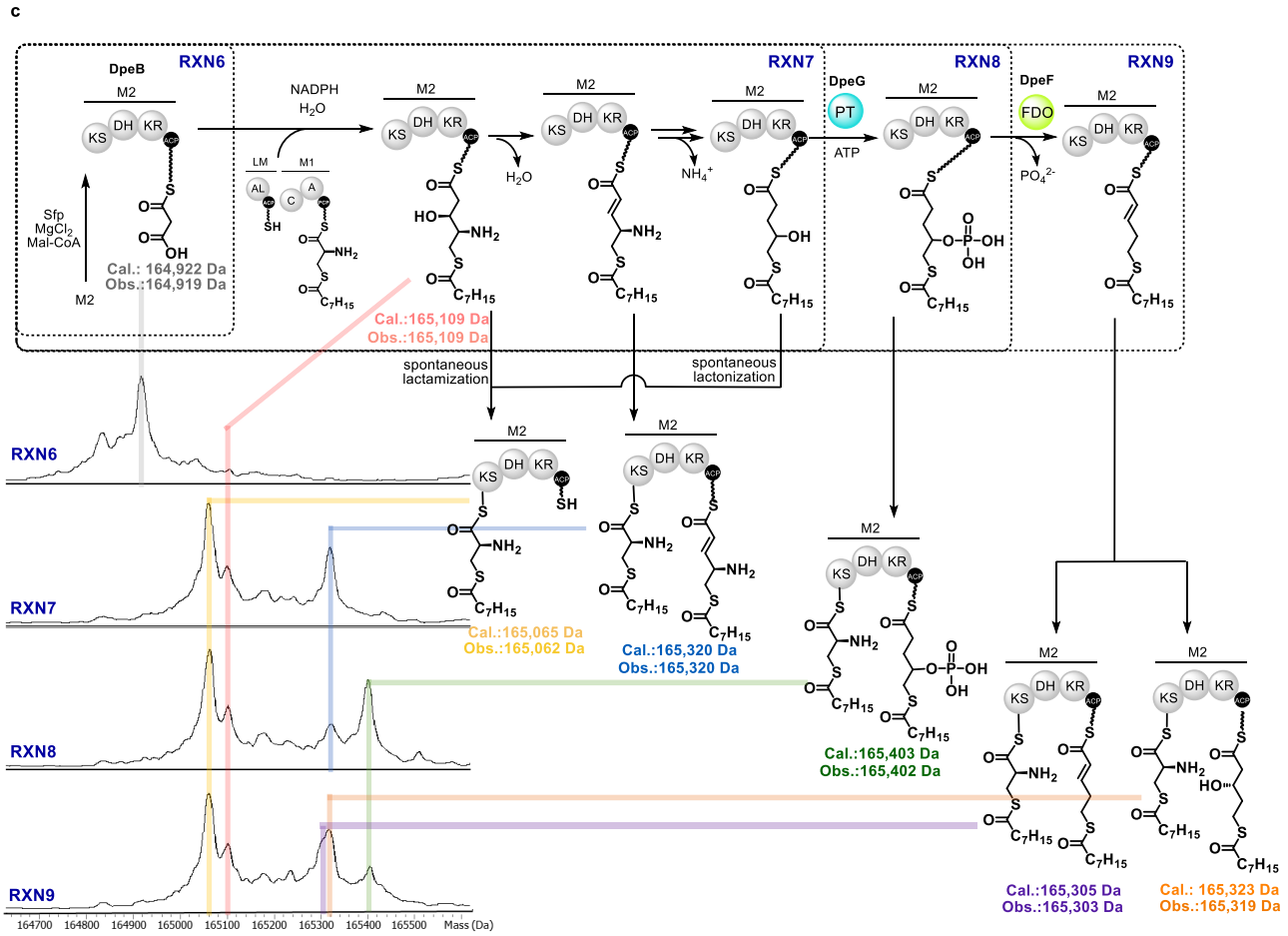


Figure 3 (continued) ***In vitro*** **reconstitution and intact protein mass spectrometry analysis of biosynthetic steps in loading module, module 1 and module 2. c**, Schematic representation of *in vitro* reconstitution assay of malonyl unit loading (RXN6), canonical chain elongation and modifications (RXN7), further modification by DepG (RXN8), and DepF (RXN9) in module 2. Deconvoluted mass spectra from UHPLC-ESI-QTOF-HRMS analysis of assays showing species with MW of 165,062 Da, 165,109 Da and 165,320 Da in RXN7 (consistent with module 2 with simple transferred of upstream product, calculated for 165,065 Da, or module 2 with intermediate underwent canonical chain elongation and ketoreduction attached, calculated for 165,109 Da, and module 2 with both upstream product and intermediate underwent chain elongation and canonical modification attached, calculated for 165,320 Da ,respectively) in RXN7. RXN8 revealed new species with MW of 165,402 (consistent with module 2 with both upstream product and chain elongated, α, β C-C bound saturated and γ-O-phosphorylated product attached, calculated for 165,403 Da). RXN9 revealed new species with MW of 165,303 Da and 165,319 corresponding to α, β-alkene and related rehydrated intermediate attached on module 2, calculated for 165,305 Da and 165,323 Da, respectively.

In vitro reconstitution of nitrogen deletion from S-acylated L-cysteine

We then attempted to reconstitute the canonical chain elongation and modification steps proposed to be catalysed by module 2. We first loaded the ACP domain of module 2 with a malonyl extender

unit using Sfp (RXN6), then incubated it with *S*-octanoyl-L-Cys-*holo*-module 1 and NADPH (RXN7) (**Fig. 3c**). UHPLC-ESI-QTOF-ESI-MS analysis revealed three distinct species. One of these species, with molecular weight (MW) 165,320 Da, is consistent with the ACP domain of module 2 being attached to the expected product of one round of chain extension with the product of module 1, followed by ketoreduction and dehydration, and the KS domain of module 2 attached to the product of module 1 (as if primed for chain extension). We can also identify a species with MW of 165,109 Da, consistent with an intermediate in which chain extension and ketoreduction has occurred but not dehydration. Interestingly, we observe a significant amount of a species with MW of 165,062 Da, consistent with simple transfer of the product of module 1 onto the KS domain of module 2. This is presumably due to hydrolysis of the ACP-tethered intermediate, either by water or by intramolecular cyclisation. Together, these species support that module 2 is capable of catalysing elongation, ketoreduction and dehydration as in a canonical PKS module.

We then included the phosphotransferase DepG in the reaction. Adding DepG and ATP into the reaction (RXN8) led to the consumption of the canonical product of module 2 and appearance of a species with MW: 165,402 Da, in which the α-β C-C bond has been reduced and the amine is replaced by a phosphorylated alcohol (**Fig. 3c**). Given the structure of shunt metabolite **6** produced by *P.chlororaphis*ΔpcdG we propose that following canonical chain elongation, ketoreduction and dehydration in module 2, the dehydratase domain in module 2 catalyses isomerisation of the alkene to allow imine formation (**Fig. 4**). Hydrolysis of the imine followed by reduction by the ketoreductase domain in module 2 would give the γ-hydroxylated intermediate which is observed to be incorporated in **6**. This intermediate would also be readily cyclised and removed from the ACP domain in module 2 via spontaneous lactonization, which explains why it is not observed in RXN7. To confirm this hypothesis, we used UHPLC-ESI-QTOF-ESI-MS analysis of the reaction extract to identify the small molecule products of RXN7 and identified a species with *m/z* corresponding to the lactonized product **14** (**Extended Data Fig. 6**). Furthermore, when we conducted RXN7 in the presence of 50% H_2O^{18} an appearance of species in *m/z* of **15** was observed, consistent with this hydroxyl group being incorporated via hydrolysis of the imine. To confirm the ability of the ketoreductase domain to catalyse reduction at the γ-carbon we synthesised a CoA-thioester substrate mimic **16**, in which the octanoyl thioester is replaced by a thioether and used Sfp to load this substrate onto the ACP domain of module 2 (RXN10) (**Extended Data Fig. 7**). In the presence of NADPH (RXN11), cysteamine offloading of products attached on the ACP domain followed by analysis

by UHPLC-ESI-QTOF-MS identified two products **19** and **20** in which the ketone has been reduced to an alcohol (resulting from cysteamine cleavage from the ACP domain or hydrolysis via lactonization, respectively). These products were not observed in the absence of NADPH or in a control reaction in which the KR domain was inactivated via a point mutation (Y328F).

Following production of the γ-hydroxylated intermediate by module 2 we propose this is directly phosphorylated by DepG (**Fig. 4**). Indeed, phosphorylated intermediate **21** could be observed following cysteamine cleavage of our octanyl thioether substrate mimic when DepG was included in this reaction (RXN 12). DepG belongs to a family of aminoglycoside-3'-phosphotransferase that primarily catalyse the addition of phosphate from ATP to the 3'-hydroxyl group of a 4,6-disubstituted aminoglycoside, such as kanamycin.[34] To our knowledge, DepG is the first example of a member of this family acting in *trans* on a PKS-bound intermediate.

Finally, we conducted a reaction containing all the necessary enzymatic machinery to make the phosphorylated intermediate as well as DepF (RXN9). This led to consumption of the phosphorylated species and the emergence of a species with MW of 165,305 Da consistent with phosphate elimination resulting in the formation of an alkene (**Fig. 3c**). As a further rehydrated species (observed MW: 165,319 Da) was found, presumably via the reaction of the dehydratase domain, which are known to act in reverse, we proposed the alkene is likely between Cα and Cβ rather than Cβ and CΥ. To confirm the location of the double bond we conducted UHPLC-ESI-QTOF-HRMS analysis of the cysteamine offloaded products using our octanyl thioether analogue in the presence of DepF (RXN13) (**Extended Data Fig. 7**). This showed the presence not only of **23**, resulting from cysteamine cleavage of the rehydrated product from the ACP domain, but also **22**, resulting from Michael addition of cysteamine to the unsaturated thioester, which could only occur if the double bond is between the α and β carbons. These compounds were not observed in an inactivated mutant of DepF in which the active site Asp is replaced by an alanine (D367A). Whether the intermediate containing the double bond between Cα and Cβ is the direct product of the DepF catalysed elimination of phosphate, or the result of a subsequent rearrangement requires further investigation.

Together, these results demonstrate the key steps in HDACi pharmacophore biosynthesis, the hydro-deamination, is completed in module 2 via an unprecedented mechanism involving novel modifications catalyzed by the multifunctional module 2 KR and DH domains, along with the non-canonical activity of tailoring enzymes DepG (PT) and DepF (FDO) that *trans*-act in module 2 (**Fig. 4**).

All that remains to complete the biosynthesis of the HDACi pharmacophore is chain extension and ketoreduction by module 3 through canonical PKS enzymology.

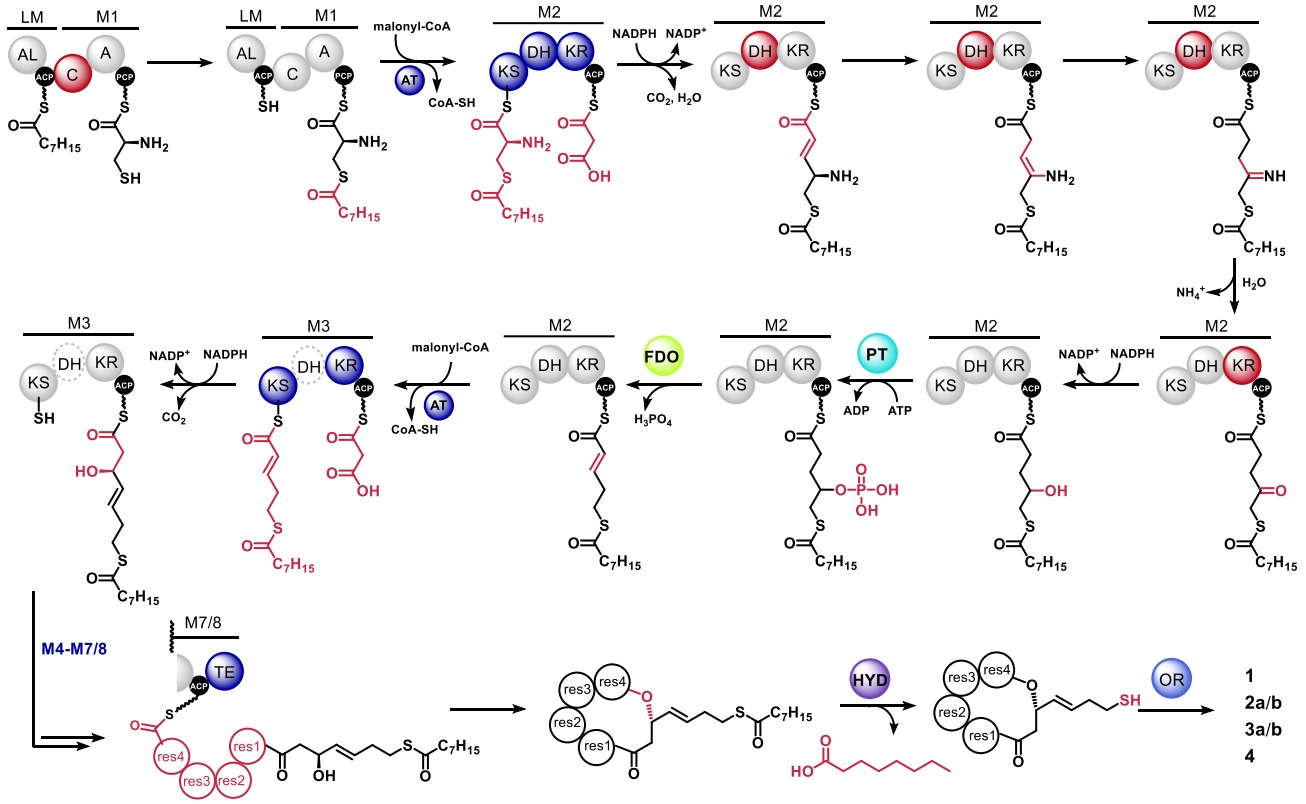


Figure 4 Proposed biosynthetic pathway of bicyclic depsipeptide HDACis. Domains or enzymes responsible for the transformation (highlighted in red in structures) in each step are highlighted, where dark blue denotes canonical function of domains and red denotes novel function of domains discovered in this work. Detailed steps proposed for conserved pharmacophore biosynthesis is shown.

One-pot biosynthesis of pharmacophore

Chemical synthesis of the pharmacophore has proven challenging and cost-inefficient, limiting the structural diversification of this important family of HDAC inhibitors.[35] We thus aimed to produce the intact pharmacophore in its natural protected form in a one-pot enzymatic reaction. To promote multiple turnovers and mimic the natural assembly process, we employed an AT domain (FabD1), reported to be responsible for fatty acid biosynthesis in the romidepsin producer to continuously supply malonyl extender units to module 2 and module 3,[36] along with the type II TE domain (PcdJ) from FR901375 biosynthesis, to cleave the mature product from the assembly line, preventing product stalling. In total, the one-pot reaction consisted of eight enzymes, including a PPtase (Sfp),

AT domain (FabD1), the NRPS-PKSs (DepA, DepB, DepC), the phosphotransferase (DepG), the flavin dependent monooxygenase (DepF) and a type IITE (PcdJ); four substrates, including octanoic acid, CoA, L-Cys and malonyl-CoA; as well as three cofactors, $MgCl_2$, ATP and NADPH (RXN14) (**Fig. 5**). The reaction was allowed to proceed for 3h at room temperature. UHPLC-ESI-QTOF-HRMS analysis of the reaction mixture revealed a prominent peak at 21.7 min in the base peak chromatogram, with *m/z* corresponding to $[M+H]^-$ for the *S*-protected pharmacophore **26**. This result confirms that module 3 successfully completes the biosynthesis of the pharmacophore in its protected form, via a final round of polyketide chain elongation and β-keto reduction.

While the reaction will require optimisation, as some shunt products and intermediates were detected, **26** is clearly the most abundant peak, demonstrating an exceptionally efficient synthesis of the pharmacophore via a 15-step sequential transformations under mild conditions. This highlights the system's immense potential for scalable and sustainable production of the pharmacophore in its protected form. Coupled with solid phase peptide synthesis of variable peptide caps, it can enable efficient structural diversification of this important family of HDACis.

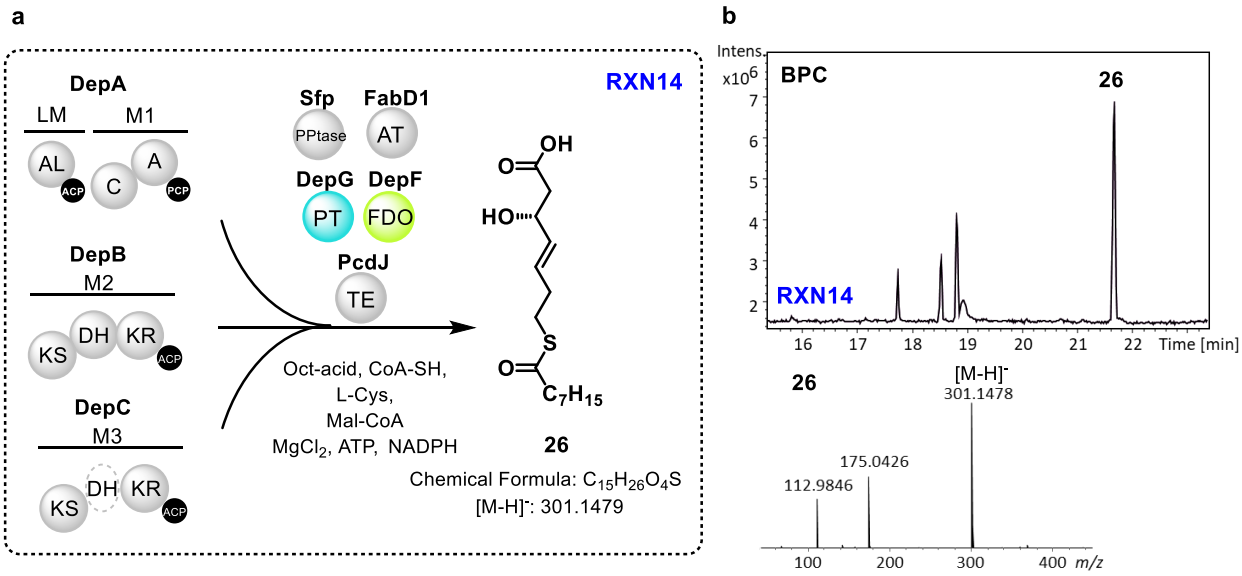


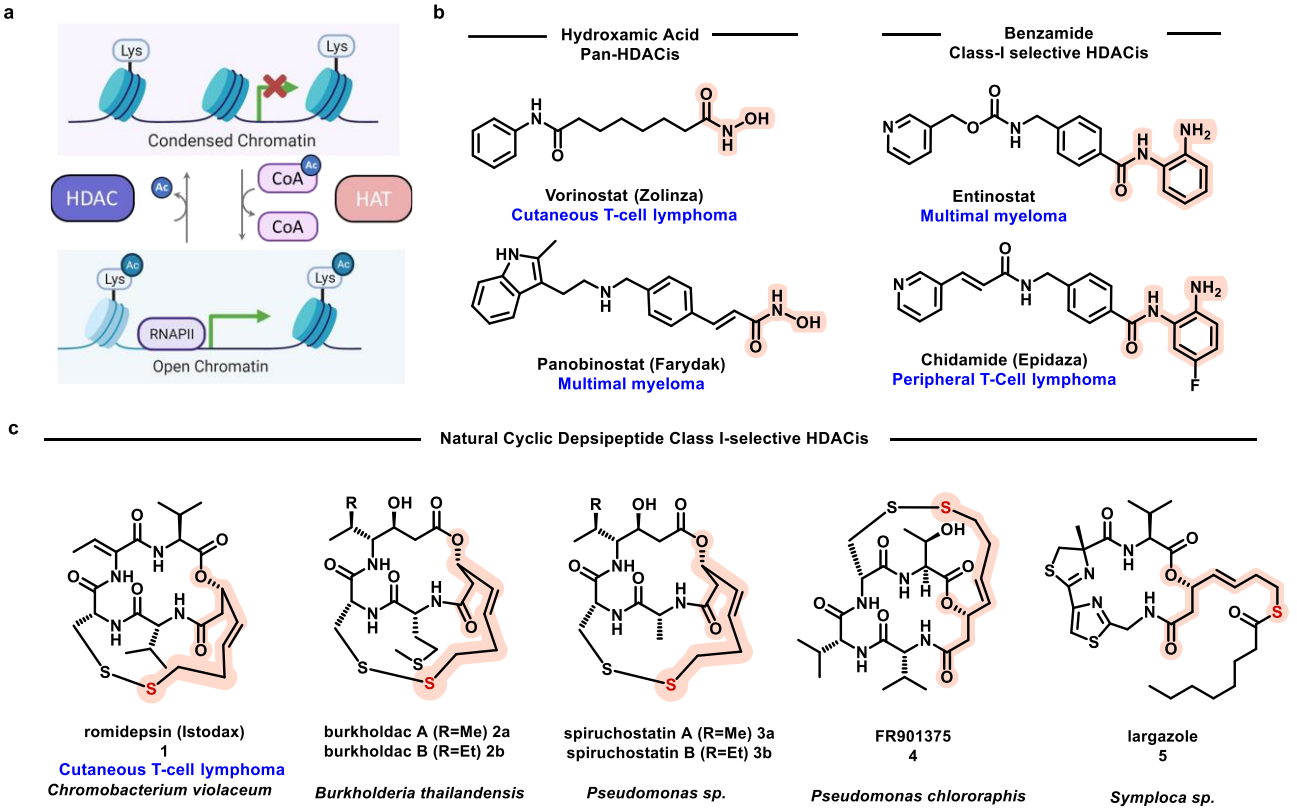
Figure 5 One-pot enzymatic synthesis of intact pharmacophore biosynthesis and UHPLC-ESI-QTOF-HRMS analysis. a, schematic representation of one-pot reconstitution of *S*-protected pharmacophore biosynthesis (RXN14); **b**, Base peak chromatogram from HUPLC-ESI-QTOF-HRMS analysis of one-pot reconstitution reaction showing a prominent peak at 21.7 min with observed m/z: 301.1478, consistent with the $[M-H]^-$ ion of **26**, calculated for 301.1479.

Conclusions

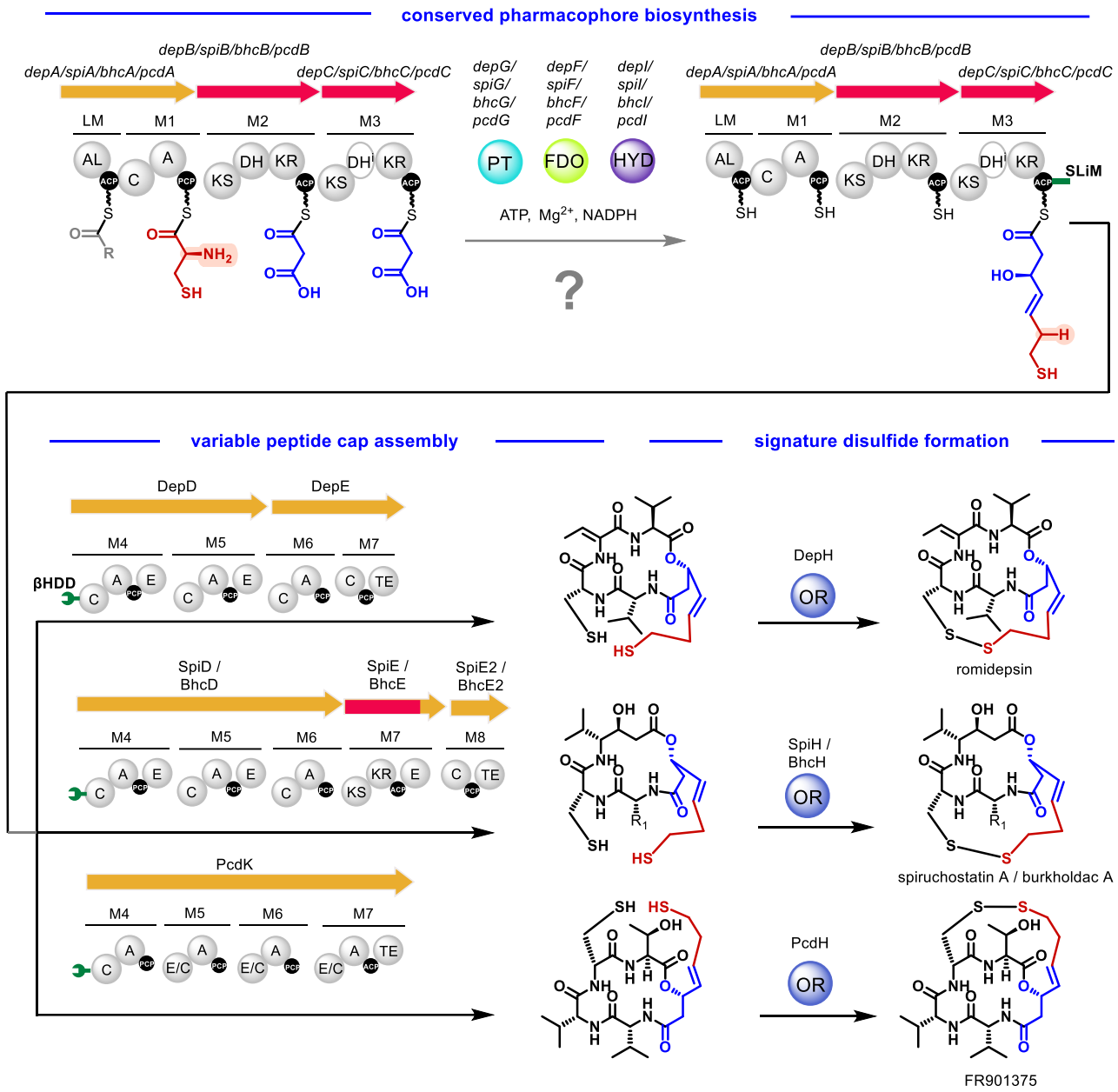
In this study, we elucidated the mechanism underlying the biosynthesis of the conserved pharmacophore of HDAC-inhibitory depsipeptides, revealing the first example of nitrogen deletion in the natural product biosynthesis. Our study uncovered a cryptic biosynthetic protection mechanism in the bicyclic depsipeptide members of this HDACi family. The assembly line employs a Cy-domain-like C domain to catalyze a novel S-acylation reaction at the onset of assembly, masking the pharmacophore thiol with an octanoyl unit, which is later removed by an esterase at a late stage of biosynthesis. This cryptic mechanism is reminiscent of the prodrug mechanism in another subgroup of this HDACi family, largazole, where the pharmacophore thiol is similarly capped by an octanoyl unit and remains acylated form. This finding suggests that the biosynthesis of different subgroup members of this HDACi family likely shares an evolutionary synergy. Furthermore, we identified an unprecedented nitrogen deletion mechanism, which involves the coordinated action of novel catalytic functions of PKS and its recruited tailoring enzymes: (i) a β- and ϒ-keto reducing dual-functional KR domain, (ii) a multifunctional DH domain capable of dehydration, alkene isomerization, and hydrolysis, (iii) a *trans*-acting and γ-alcohol phosphorylating phosphotransferase, and (iv) a *trans*-acting and dephosphorylating flavin-dependent oxidoreductase. This work provides significant insights into the complexity and versatility of PKS and NRPS biosynthetic machineries. These enzymatic activities demonstrate an intricate biochemical logic that allows for precise structural modifications, thereby expanding the paradigm of natural product site-specific editing.

Leveraging the insights into the biosynthetic mechanism, we established a highly efficient one-pot enzymatic synthesis of the intact pharmacophore of this family HDAC inhibitors in its protected form. This biocatalytic strategy provides a sustainable and efficient alternative for pharmacophore generation compared to conventional synthetic approaches. While our study demonstrates the feasibility of enzymatic synthesis, there remains room for further optimization to enhance scalability. Coupling the optimised enzymatic synthesis approach with solid-phase peptide synthesis to incorporate diverse peptide caps will enable the efficient generation of structural analogues of this important family of HDAC inhibitors.

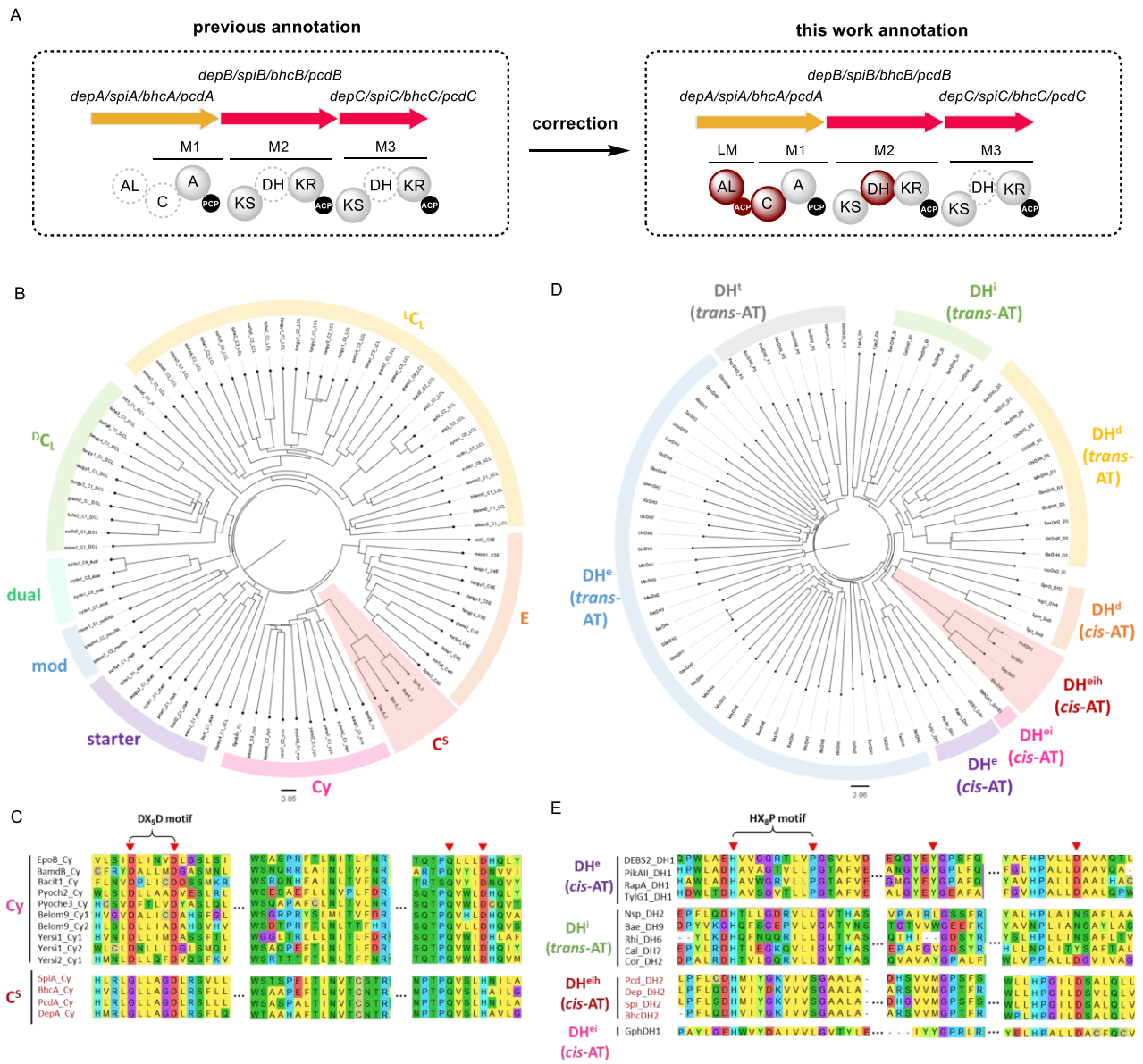
Extended Data Figures



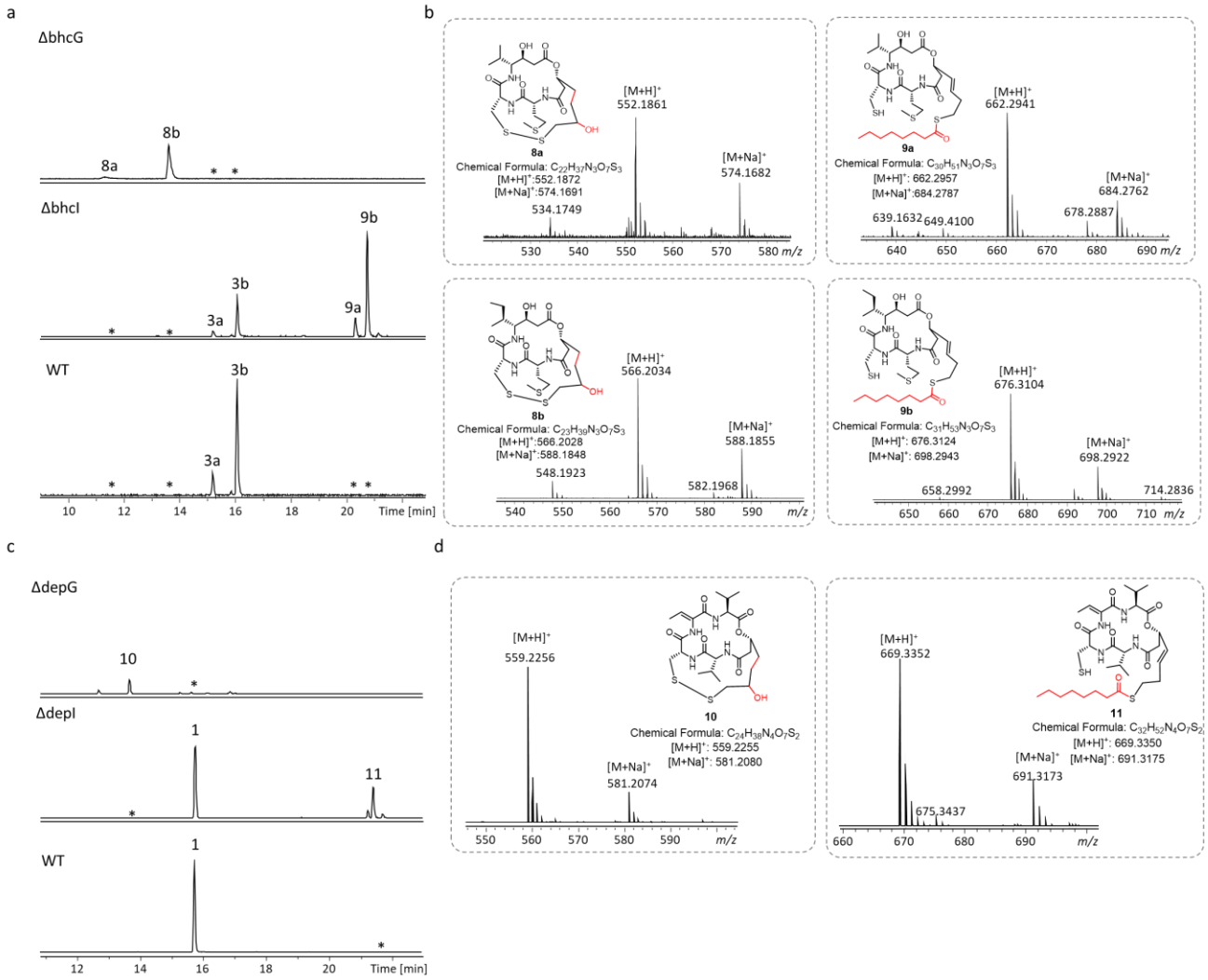
Extended Data Figure 1 HDACs and HDACis in anticancer therapeutics. a, Mechanisms of HDAC and HAT epigenetic regulations. **b**, Prototypical of synthetic HDACis with zinc binding group highlighted; **c**, Natural cyclic depsipeptide HDACis with zinc binding group highlighted.



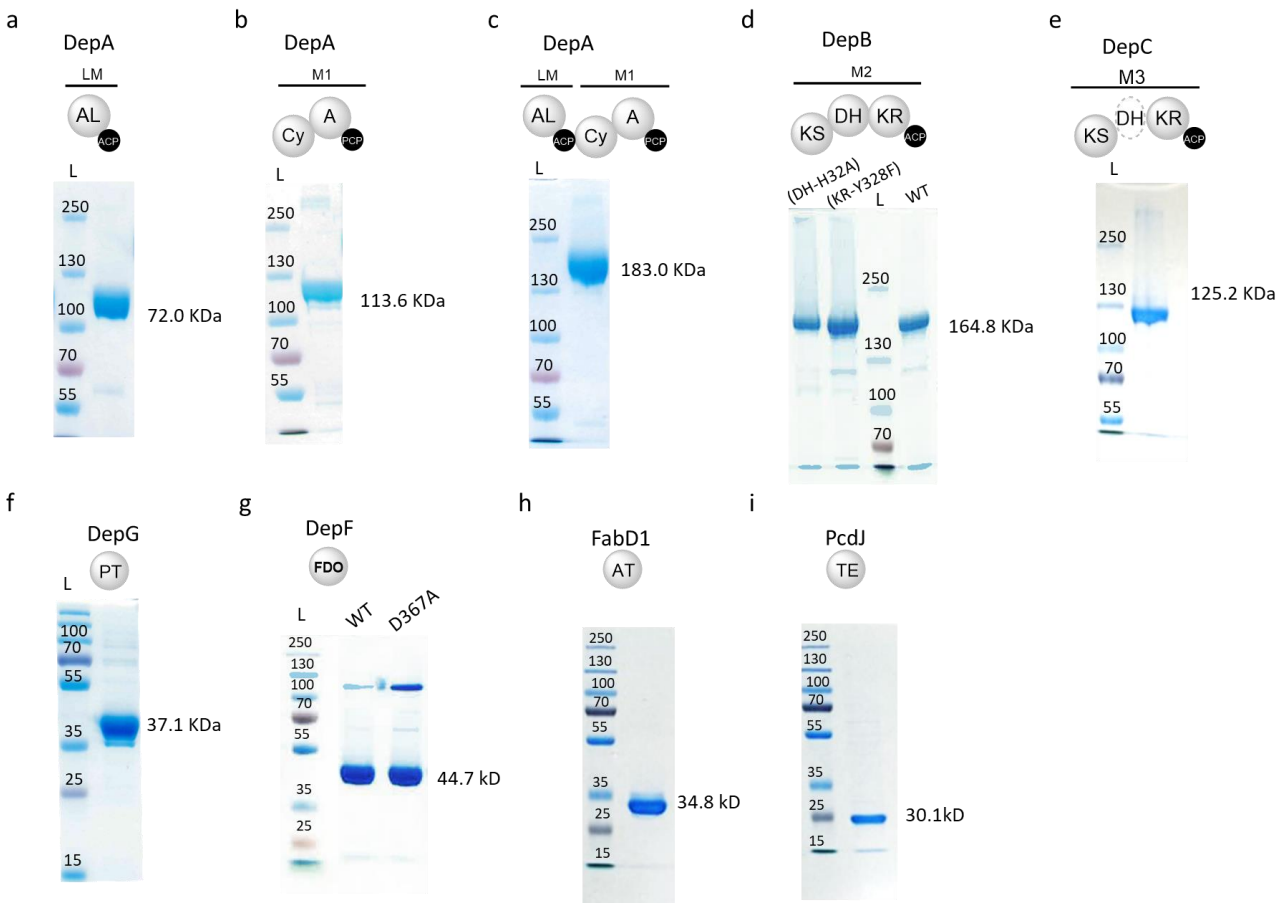
Extended Data Figure 2 Complete biosynthetic assembly lines of bicyclic depsipeptide HDAC inhibitors.



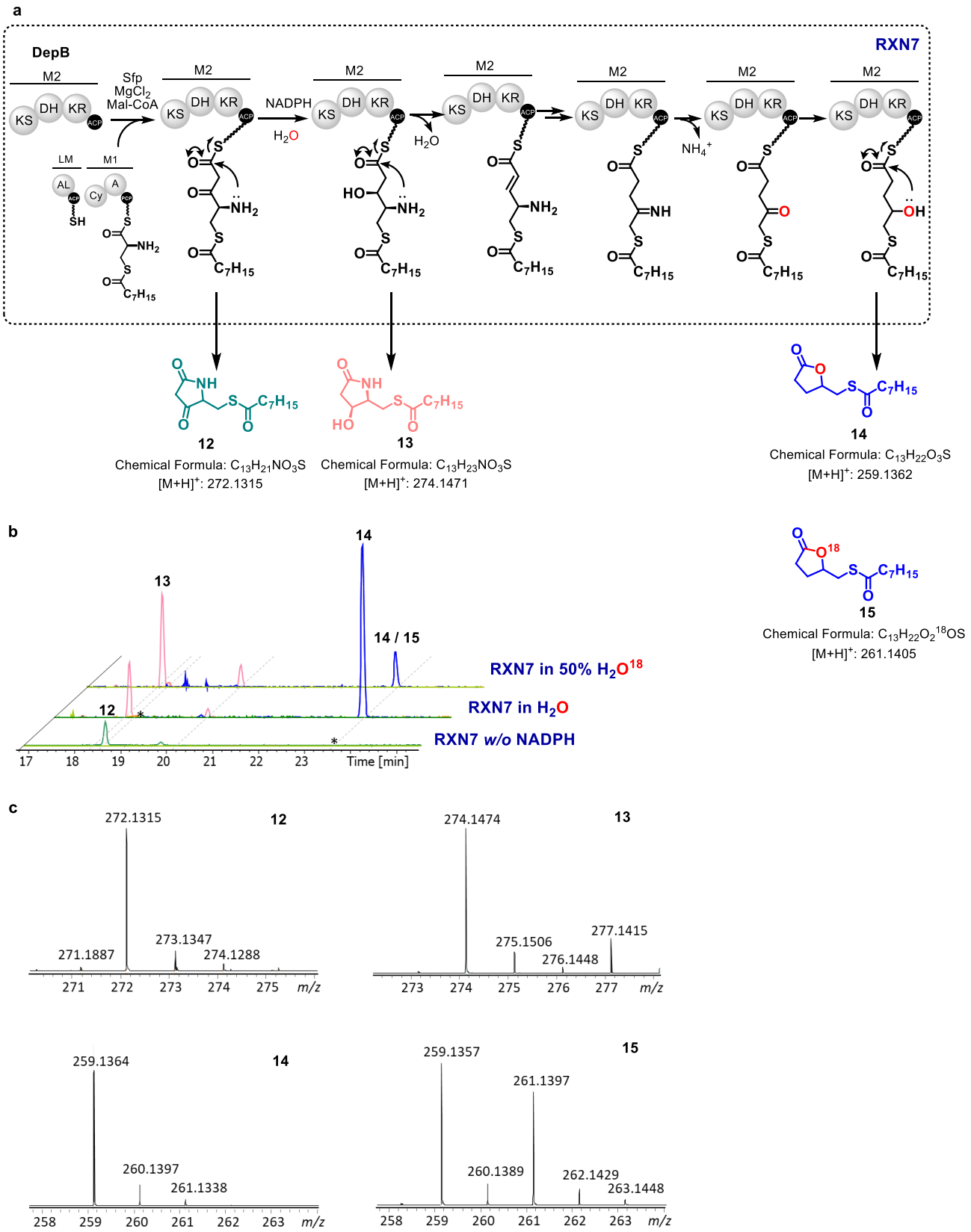
Extended Data Figure 3 Bioinformatic analysis of module 1-3 of bicyclic depsipeptide HDACi assembly lines. a, Annotation revision of loading module and module 1-3, corrections of domain annotations are highlighted in dark red. **b**, Phylogenetic analysis of module 1 C domain, $^{L}C_{L}$: L-aminoacyl-PCP condensing C domains, $^{D}C_{L}$: D-aminoacyl-PCP and a L-aminoacyl-PCP condensing C domains, E: epimerisation domains, Cy: heterocyclization domains, starter: starter C domains, mod: dehydrating C domains, dual: epimerising and condensing dual functional C domains, C^{S}: thioester synthesising C domains (module 1 C domains in this work). **c**, sequence alignment of reported Cy domains and module 1 Cy domains. **d**, Phylogenetic analysis of module 2 DH domains. DH^{e}: enoyl synthesising DH domains, DH^{d}: 2, 4-dienoyl synthesising DH domains, DH^{T}: tetrahydropyran/tetrahydrofuran synthesising DH domains, DH^{i}: enoyl isomerising DH domains, DH^{ei}: enoyl synthesising and isomerising dual functional DH domains, DH^{eih}: enoyl synthesising, isomerising and amino hydrolysing trifunctional DH domains (module 2 DH domains in this work), *trans*-AT: DH domain from *trans*-AT PKSs, *cis*-AT: DH domains from *cis*-AT PKSs. **e**, sequence alignment of reported DH-like domains and module 2 DH domains.



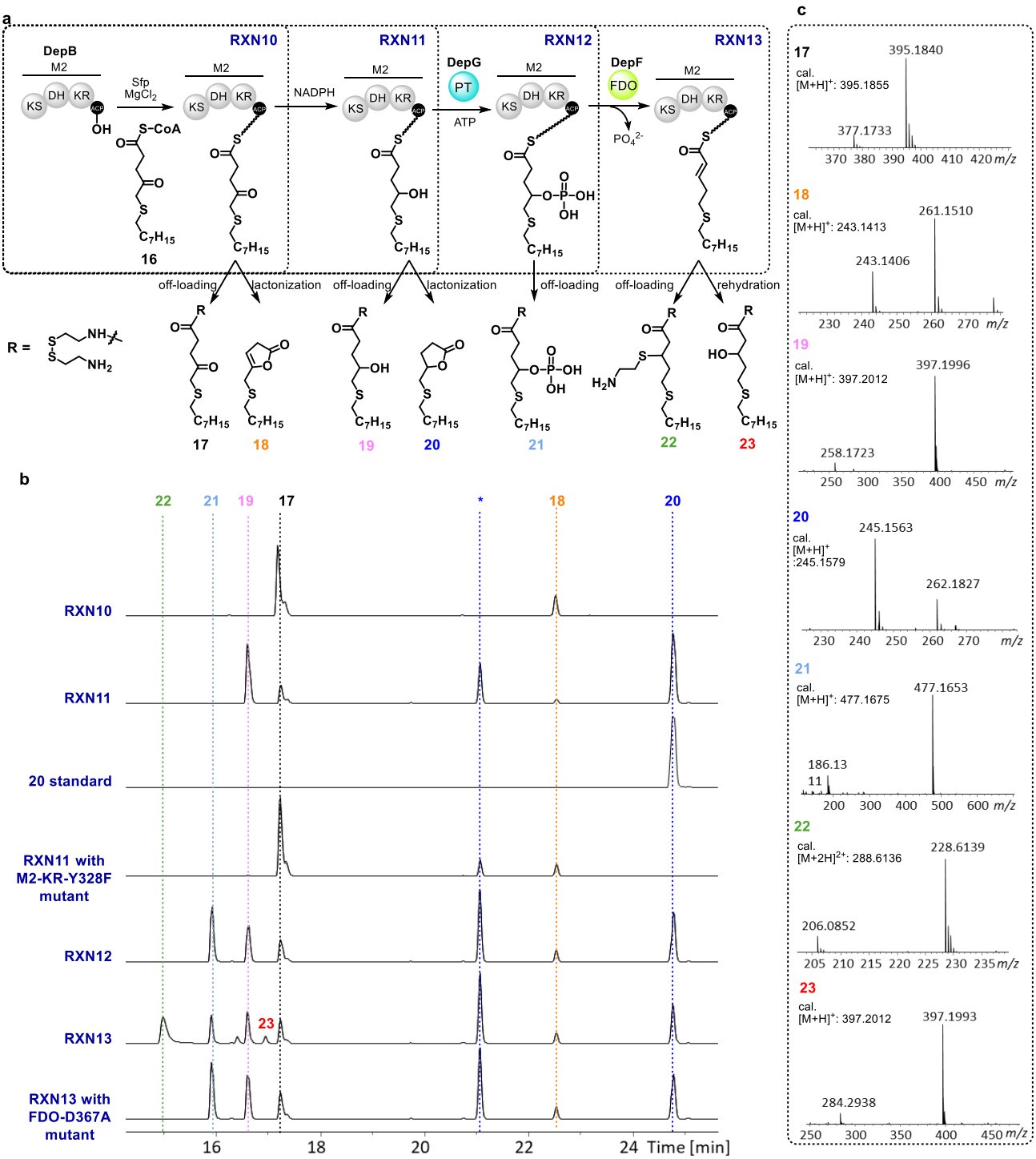
Extended Data Figure 4 UHPLC-ESI-QTOF-HRMS analysis of shunt metabolites produced by in frame gene deletion mutants. a, EIC of $[M+H]^+$ ions of shunt metabolite **8a**/**b**, **9a**/**b** and burkholdac A/B **3a**/**b** from *bhcG* deletion mutant (top), *bhcI* deletion mutant (middle) and WT producer *B. thailandensis* DSM13276 (bottom). **b**, Observed mass spectra of $[M+H]^+$ ions of **8a/b** and **9a**/**b**. **c**, EIC of $[M+H]^+$ ions of shunt metabolite **10, 11** and romidepsin **1** from *depG* deletion mutant (top), *depI* deletion mutant (middle) and WT producer *C.chromobacterium* (bottom); **d**, Observed mass spectra of $[M+H]^+$ ions of **10** and **11.**



Extended Data Figure 5 SDS-PAGE analysis of purified recombinant proteins used in this work.



Extended Data Figure 6 UHPLC-ESI-Q-TOF-MS analysis of spontaneous lactamized or lactonized products in the module 2 chain elongation assay (RXN7).



Extended Data Figure 7. UHPLC-ESI-QTOF-HRMS of *in vitro* assays dissecting module 2 late-stage steps with synthetic probe OTKP-CoA thioester 16. a, Schematic representation of assays and expected products. **b**, EICs of $[M+H]^+$ ions pf off-loaded cysteamine adducts and lactonized products **17**-**23**. **c**, HRMS spectra of observed $[M+H]^+$ ion and corresponding calculated *m/z*.

References

1. Paradkar, A. Clavulanic acid production by Streptomyces clavuligerus: biogenesis, regulation and strain improvement. *J. Antibiot. (Tokyo).* **66**, 411–420 (2013).

2. Barry, S. M. *et al.* Cytochrome P450-catalysed L-tryptophan nitration in thaxtomin phytotoxin biosynthesis. *Nat. Chem. Biol.* **8**, 814 (2012).

3. Li, F., Zhang, X. & Renata, H. Enzymatic C–H functionalizations for natural product synthesis. *Curr. Opin. Chem. Biol.* **49**, 25–32 (2019).

4. Buller, R. *et al.* From nature to industry: Harnessing enzymes for biocatalysis. *Science (80-.).* **382**, eadh8615 (2023).

5. Pyser, J. B., Chakrabarty, S., Romero, E. O. & Narayan, A. R. H. State-of-the-Art Biocatalysis. *ACS Cent. Sci.* **7**, 1105–1116 (2021).

6. Kennedy, S. H., Dherange, B. D., Berger, K. J. & Levin, M. D. Skeletal editing through direct nitrogen deletion of secondary amines. *Nat. 2021 5937858* **593**, 223–227 (2021).

7. Chang, C. W. T. *et al.* Characterization of a unique coenzyme B6 radical in the ascarylose biosynthetic pathway. *J. Am. Chem. Soc.* **122**, 4239–4240 (2000).

8. Parmeggiani, F., Weise, N. J., Ahmed, S. T. & Turner, N. J. Synthetic and therapeutic applications of ammonia-lyases and aminomutases. *Chem. Rev.* **118**, 73–118 (2018).

9. Glozak, M. A. & Seto, E. Histone deacetylases and cancer. *Oncogene 2007 2637* **26**, 5420–5432 (2007).

10. Falkenberg, K. J. & Johnstone, R. W. Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat. Rev. Drug Discov.* **13**, 673–691 (2014).

11. Shigematsu, N. *et al.* FR901228, a novel antitumor bicyclic depsipeptide produced by *chromobacterium violaceum* no. 968. II. Structure determination. *J. Antibiot. (Tokyo).* **47**, 311–314 (1994).

12. Biggins, J. B., Gleber, C. D. & Brady, S. F. Acyldepsipeptide HDAC inhibitor production induced in *Burkholderia thailandensis. Org. Lett.* **13**, 1536–1539 (2011).

13. Masuoka, Y. *et al.* Spiruchostatins A and B, novel gene expression-enhancing substances produced by *Pseudomonas* sp. *Tetrahedron Lett.* **42**, 41–44 (2001).

14. Chen, Y., Gambs, C., Abe, Y., Wentworth, P. & Janda, K. D. Total synthesis of the depsipeptide FR-901375. *J. Org. Chem.* **68**, 8902–8905 (2003).

15. Taori, K., Paul, V. J. & Luesch, H. Structure and activity of largazole, a potent antiproliferative agent from the Floridian marine cyanobacterium *Symploca* sp. *J. Am. Chem. Soc.* **130**, 1806–1807 (2008).

16. Furumai, R. *et al.* FK228 (depsipeptide) as a natural prodrug that inhibits class I histone deacetylases. *Cancer Res.* **62**, 4916–4921 (2002).

17. Ying, Y., Taori, K., Kim, H., Hong, J. & Luesch, H. Total synthesis and molecular target of largazole, a histone deacetylase inhibitor. *J. Am. Chem. Soc.* **130**, 8455–8459 (2008).

18. Cole, K. E., Dowling, D. P., Boone, M. A., Phillips, A. J. & Christianson, D. W. Structural basis of the antiproliferative activity of largazole, a depsipeptide inhibitor of the histone deacetylases. *J. Am. Chem. Soc.* **133**, 12474–12477 (2011).

19. Micelli, C. & Rastelli, G. Histone deacetylases: Structural determinants of inhibitor selectivity. *Drug Discov. Today* **20**, 718–735 (2015).

20. Weissman, K. J. The structural biology of biosynthetic megaenzymes. *Nat. Chem. Biol.* **11**, 660–670 (2015).

21. Grininger, M. Enzymology of assembly line synthesis by modular polyketide synthases. *Nat. Chem. Biol.* (2023) doi:10.1038/s41589-023-01277-7.

22. Hur, G. H., Vickery, C. R. & Burkart, M. D. Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. *Nat. Prod. Rep.* **29**, 1074–1098 (2012).

23. Cheng, Y. Q., Yang, M. & Matter, A. M. Characterization of a gene cluster responsible for the biosynthesis of anticancer agent FK228 in *Chromobacterium violaceum* No. 968. *Appl. Environ. Microbiol.* **73**, 3460–3469 (2007).

24. Potharla, V. Y., Wang, C. & Cheng, Y. Q. Identification and characterization of the spiruchostatin biosynthetic gene cluster enable yield improvement by overexpressing a transcriptional activator. *J. Ind. Microbiol. Biotechnol.* **41**, 1457–1465 (2014).

25. Wang, C. *et al.* Thailandepsins: Bacterial products with potent histone deacetylase inhibitory activities and broad-spectrum antiproliferative activities. *J. Nat. Prod.* **74**, 2031–2038 (2011).

26. Wang, C., Wesener, S. R., Zhang, H. & Cheng, Y. Q. An FAD-Dependent pyridine nucleotide-disulfide oxidoreductase Is Involved in disulfide bond formation in FK228 anticancer depsipeptide. *Chem. Biol.* **16**, 585–593 (2009).

27. Keatinge-Clay, A. T. The uncommon enzymology of *cis*-acyltransferase assembly lines. *Chem. Rev.* **117**, 5334–5366 (2017).

28. Dowling, D. P. *et al.* Structural elements of an NRPS cyclization domain and its intermodule docking domain. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12432–12437 (2016).

29. Hobson, C. *et al.* Diene incorporation by a dehydratase domain variant in modular polyketide synthases. *Nat. Chem. Biol. 2022 1812* **18**, 1410–1416 (2022).

30. Dodge, G. J., Ronnow, D., Taylor, R. E. & Smith, J. L. Molecular basis for olefin rearrangement in the gephyronic acid polyketide synthase. *ACS Chem. Biol.* **13**, 2699–2707 (2018).

31. Jenner, M. *et al.* An unusual *Burkholderia gladioli* double chain-initiating nonribosomal peptide synthetase assembles 'fungal' icosalide antibiotics. *Chem. Sci.* **10**, 5489–5494 (2019).

32. Jian, X. *et al.* Antibiotic skeletal diversification via differential enoylreductase recruitment and module iteration in *trans*-acyltransferase polyketide synthases. *J. Am. Chem. Soc.* **146**, 6114–6124 (2024).

33. Jenner, M. *et al.* Elucidating the molecular programming of a nonlinear non-ribosomal peptide synthetase responsible for fungal siderophore biosynthesis. *Nat. Commun. 2023 141* **14**, 1–12 (2023).

34. Wright, G. D. Aminoglycoside phosphotransferases: proteins, structure, and mechanism. *Front. Biosci.* **4**, d9 (1999).

35. Maolanon, A. R., Kristensen, H. M. E., Leman, L. J., Ghadiri, M. R. & Olsen, C. A. Natural and Synthetic Macrocyclic Inhibitors of the Histone Deacetylase Enzymes. *ChemBioChem* **18**, 5–49 (2017).

36. Wesener, S. R., Potharla, V. Y. & Cheng, Y. Q. Reconstitution of the FK228 biosynthetic pathway reveals cross talk between modular polyketide synthases and fatty acid synthase. *Appl. Environ. Microbiol.* **77**, 1501–1507 (2011).