A widespread family of bacterial gene clusters produces ClpP inhibitors

Gerry Wright (wrightge@mcmaster.ca)
McMaster University https://orcid.org/0000-0002-9129-7131
Elizabeth Culp
McMaster University
David Sychantha
Christian Hobson
McMaster University
Andrew Pawlowski
McMaster University
Gerd Prehna
University of Manitoba

Biological Sciences - Article

Keywords: intracellular proteolytic complexes, ClpP inhibitors, drug discovery, Clipibicyclene

Posted Date: May 27th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-464484/v1

Version of Record: A version of this preprint was published at Nature Microbiology on March 4th, 2022.
See the published version at https://doi.org/10.1038/s41564-022-01073-4.

Abstract

Intracellular proteolytic complexes play an essential role in modeling the proteome in both bacteria and eukaryotes. ClpP is the protease subunit of one such highly conserved proteolytic complex that, despite its potential, remains unexploited as a drug target. Here we describe a target-directed genome mining strategy to identify ClpP targeting compounds from the bacterial order Actinomycetales. By searching for biosynthetic gene clusters that contain duplicated copies of ClpP as putative antibiotic resistance genes, we identify a family of ClpP-associated clusters that are widespread across phyla, including environmental and pathogenic bacteria. While numerous bacterial pyrrolizidine alkaloids produced by these gene clusters are known, their connection to ClpP has never been made. We show that these previously characterized molecules do not affect ClpP function but are shunt metabolites derived from the genuine product of these gene clusters, a reactive covalent ClpP inhibitor. Focusing on one such cryptic gene cluster from *Streptomyces cattleya* DSM 46488, we use heterologous expression to purify the relevant ClpP inhibitor, which we name clipibicyclene. We show *in vitro* and *in vivo* that clipibicyclene is a potent covalent inhibitor of ClpP and that cluster-associated ClpPs provide resistance. ClpP inhibition results in antibacterial activity against actinobacteria, including *Mycobacterium smegmatis*, and inhibition of virulence factor production by *Staphylococcus aureus.* Finally, we solve the crystal structure of clipibicyclene-modified *Escherichia coli* ClpP. Clipibicyclene's discovery deconvolutes the actual function of a family of natural products widespread in nature. It provides a novel scaffold for therapeutic ClpP inhibitor development, making these findings significant from the perspective of their discovery and their clinical potential.

Main

The Caseinolytic protease (ClpP) complex is a central regulator of protein homeostasis in both bacteria[1,2] and eukaryotic mitochondria[3]. As a highly conserved complex, it has been extensively characterized for its role in the degradation of mistranslated proteins and for specifically targeted proteins for general regulation of the proteome[1–3]. To achieve these controlled functions, ClpP's activity is tightly regulated by the formation of a barrel-shaped tetradecameric complex of two stacked heptameric rings that shelter fourteen internal serine protease catalytic sites[4]. Access requires passage through axial openings of the barrel too small for native proteins to pass. To gain access, substrates are selected and unfolded by an Hsp100 family AAA+ superfamily ATPase, including ClpA, ClpX, and ClpC, that form heptameric rings and dock on the axial openings[1,2]. By interacting with different adaptor proteins and AAA+ ATPases, specific substrates can be selected.

Most organisms encode a single ClpP isoform that forms a homo-tetradecameric complex that is dispensable for growth. Actinobacteria, including *Mycobacterium* spp. and *Streptomyces* spp., encode at least two ClpP isoforms essential for growth. Functional proteolytic complexes formed in these organisms are composed of two stacked homo-heptameric rings that form a hetero-tetradecamer (e.g., ClpP1 heptamer with ClpP2 heptamer).

Targeting of ClpP is attractive in the discovery of both anticancer and antibacterial therapeutic candidates. For example, several synthetic β-lactones and phenyl esters have been developed that inhibit bacterial and mitochondrial ClpP[5–9]. In bacterial pathogens where it is dispensable (e.g., *Staphylococcus aureus*), ClpP inhibition can be used as an anti-virulence strategy[6,10], while if essential (e.g., *Mycobacteria tuberculosis*), it can act as a traditional antibiotic[5]. In the context of anticancer drugs, human mitochondrial ClpP is dispensable but appears to be more important to the function of cancer cells than normal cells[11]. For example, ClpP inhibition can selectively kill acute myeloid leukemia cells versus normal hematopoietic cells[12].

While ClpP inhibition is the aim of many drug discovery campaigns, activation of the complex is also attractive. By mimicking AAA+ ATPase binding and inducing widening of the ClpP tetradecamer's axial pore, ClpP activators stimulate nonspecific intracellular proteolytic activity and cause cell death[13,14]. Imipiridones such as ONC201[15] activate mitochondrial ClpP and are currently being studied in numerous anticancer clinical trials[16]. ONC201 also has antibacterial activity via the same mechanism[17]. Most ClpP campaigns focus on synthetic compounds, and only a single natural product, acyldepsipeptide antibiotic (ADEP)[18], is known to specifically target this important protease complex. ADEP appears rare in nature as the producer organism, *Streptomyces hawaiiensis* NRRL15010, remains the only organism in GeneBank containing the ADEP biosynthetic gene cluster (BGC)[19]. Given ClpP's promise as a therapeutic target, we wondered if other natural products that target ClpP were yet to be discovered. Such compounds could identify new chemical scaffolds with application in antibacterial and anticancer drug discovery.

Target-directed genome mining for ClpP directed natural products

To avoid self-intoxication, BGCs contain resistance genes against their encoded antibiotic. When this resistance gene takes the form of an antibiotic-insensitive copy of the targeted protein, searching for BGCs containing that target gene can specifically lead to the identification of new antibiotics. So-called "target-directed genome mining" was applied in the discovery of thiolactomycin inhibitors of fatty acid synthases FabB /F[20,21]. The ADEP BGC encodes a resistant copy of ClpP, demonstrating that this protection mechanism from self-intoxication is also relevant to ClpP[19]. We, therefore, applied target-directed genome mining to identify novel natural products targeting ClpP.

To identify ClpP-associated clusters, we used the collection of Actinobacterial ClpP proteins from pFAM as queries to extract 5928 *clpP* genes from RefSeq (Figure 1a). We examined genomic regions 50 kb upstream and downstream of each *clpP* homolog and identified BGCs in 1096 of these regions using AntiSMASH[22]. These BGCs were then filtered for those that contained ClpP within cluster boundaries, as identified through tblastn of *S. coelicolor* ClpP1, leaving 145 candidate hits. We next sought to differentiate cluster-associated from housekeeping *clpPs* coincidentally in proximity to a BGC, an exercise complicated because actinomycetes encode up to six *clpP* homologs per genome. Using the characteristic genetic context of each *clpP* homolog to filter out housekeeping copies (see methods), we landed on a final prioritized list of ten BGCs associated with ClpPs (Figure 1a, Supplementary Table 1).

While our target-directed genome mining approach is similar to the idea employed by the Antibiotic Resistant Target Seeker (ARTS)[21], the latter focuses on a single genome/group of genomes, and using a reference 'core' genome, identifies any duplicated essential gene within these genomes' BGCs. Conversely, our approach focuses on a single target, ClpP, and searches all genomes in RefSeq. Furthermore, given the multiple and a variable number of *clpP* copies per genome, specific knowledge of genetic context was vital to differentiating housekeeping from cluster-associated paralogs.

Interestingly, six of our ten final prioritized BGCs contained a bimodular non-ribosomal peptide synthetase (NRPS) predicted to activate proline and serine. We identified 306 BGCs in unique species' genomes from RefSeq with this bimodular NRPS (Supplementary Table 2). These BGCs were missed in our initial analysis due to cluster fragmentation, upload date, or lacking ClpP association. As a comparison, BGCs for streptothricin-family compounds, which are estimated to be produced by 1 in 10 antibiotic-producing *Streptomyces* spp.[23], are present in RefSeq only 50 times. In addition to its abundance, this bimodular NRPS is widespread across genera of the order Actinomycetales (*Streptomyces, Saccharothrix, Nocardia,* etc.) as well as across phyla in important human pathogens, including *Pseudomonas aeruginosa, Acinetobacter baumannii,* and *Enterobacter cloacae* (Supplementary Table 2). While not all of the identified BGCs were associated with ClpPs, many were located close to other serine hydrolases such as S8-family peptidases and β-lactamases. These results suggest that this bimodular NRPS BGC family is unusually common, widespread in environmental and clinical bacteria, and may have a conserved function of inhibiting Ser hydrolases.

To visualize the relationship between ClpP association and BGCs in this family, we selected 30 exemplar BGCs and built a multi-locus phylogeny using the tool CORASON (CORe Analysis of Syntenic Orthologs to prioritize Natural product biosynthetic gene clusters)[24]. This analysis effectively revealed three distinct groups among our collection of BGCs (Figure 1b). In the first group, the bimodular NRPS is adjacent to a Baeyer-Villiger flavin-containing monooxygenase (FMO) in a pseudomonal strain, with members of this group encoding the known phospholipase inhibitor SB-253514 (also called brabantamide, Supplementary figure 1a)[25,26]. In the second group, the bimodular NRPS is adjacent to a condensation domain containing protein, with no product having been characterized. The third group includes BGCs associated with an FMO and ClpP, with members known for producing a family of bacterial pyrrolizidine alkaloids (Supplementary figure 1a). The biosynthesis of some of these compounds has been characterized, such as for the bohemamines produced by *Streptomyces* sp. CB02009[27] and azabicyclenes (also called azetidomonamide B) produced by *P. aeruginosa* PAO1[28,29]. However, to the best of our knowledge, the obligate association of these BGC's with ClpP has not been previously explored. We, therefore, set out to further examine this widespread family of clusters and investigate their association with ClpP.

In surveying the unknown members of this family of ClpP-associated BGCs, we became interested in a cluster from *Streptomyces cattleya* DSM 46488 that, in addition to the bimodular NRPS, also contains a type I PKS (Figure 1c, Supplementary Table 4). We named the cluster *cac* for ClpP Associated Cluster. The *cac* BGC was transcriptionally silent under several culturing conditions tested. Consequently, we used

heterologous expression by directly capturing the 44 kb cluster on a shuttle vector (pCGW-cac) and moving it to the superhost *Streptomyces coelicolor* M1154. To transcriptionally activate the cluster, we first attempted to overexpress the cluster-situated XRE family transcriptional activator, Cac15. Still, we did not observe production of any new metabolites from this engineered strain, consistent with incomplete transcriptional activation (Figure 1c, Supplementary Figure 1b). Therefore, we refactored *cac* by inserting non-native promoters in front of four putative transcriptional units (Supplementary Figure 1c; pCGW-cac-LHK). This approach successfully activated the full BGC and supported the production of numerous metabolites (Supplementary Figure 1b, Figure 1d).

We purified and solved the structure of major peaks by NMR and/or LC-MS/MS (Supplementary Figures 2-6, Supplementary Tables 5-8, Supplementary Discussion). The primary product identified was a novel azabicyclene with a hydroxylated decatriene acyl tail (Figure 1e, compound **5**). We also identified congeners lacking the hydroxyl group (compound **6**), harboring a ketone at the same position (compound **4**), and primary amide variants of the respective acyl tails (compounds **1-3**). We named these molecules azabicyclenes B (**4**), C (**5**), and D (**6**), following the nomenclature of the compound azabicyclene (which we will hereto refer to as azabicyclene A) produced by *P. aeruginosa* PAO1 (Supplementary Figure 1a)[28,29]. The biosynthesis of azabicyclenes B-D is predicted to be similar to bohemamines and azabicyclene A[27,29] (Supplementary Discussion, Supplementary Figure 7).

An elusive product of the *cac* BGC results in ClpP inhibition

Having identified the azabicyclenes as metabolites produced by *cac,* we next tested the hypothesis that they disrupt the function of ClpP. However, azabicyclene C and D did not robustly inhibit the growth of eukaryotic cells or bacterial species where ClpP is essential (Supplementary Figure 8a). Azabicyclene D has weak growth inhibitory activity (MIC = 64-128 µg/mL) but is equally active against species where ClpP is dispensable (*S. aureus, Bacillus subtilis)* and *clpP* knockout strains (*B. subtilis* Δ*clpP*), indicating that its activity is independent of ClpP. Azabicyclenes C and D also do not impact ClpP activity *in vitro* (Supplementary Figure 8b). Similarly, bohemamines purified from *Streptomyces sp.* NBRC110035 did not affect ClpP (Supplementary Figure 8c), consistent with no previously reported biological activity[30].

Since azabicyclenes and previously identified pyrrolizidine alkaloids do not perturb ClpP, we hypothesized that another undetected product of the BGC might. To comprehensively assay metabolites produced by *cac,* we, therefore, probed ClpP function directly in the heterologous host, *S. coelicolor* M1154 pCGW-cac-LHK. In *S. coelicolor,* as in many *Streptomyces* spp., there are five copies of *clpP* in three operonic units: *clpP1clpP2, clpP3clpP4,* and *clpP5.* Each operon encodes subunits capable of forming a functional hetero-tetradecameric complex except for ClpP5, whose catalytic triad is irregularly spaced, and proteolytic activity remains unknown[31]. ClpP1P2 complexes are susceptible to ADEP activation while ClpP3P4 is resistant (Figure 2a)[32]. Due to this differential susceptibility, ADEP can be used to probe the inhibition of ClpP1P2 in *S. coelicolor*, if ClpP1P2 is functional, nonspecific degradation induced by ADEP will result in cell death, while if ClpP1P2 is inhibited, ADEP will be unable to trigger nonspecific degradation and ClpP3P4 will support the growth of cells. *S. coelicolor* M1154 pCGW-cac-LHK was more

resistant to ADEP than the empty vector control, *S. coelicolor* M1154 pCGW, indicating that ClpP1P2 was inhibited in the heterologous producer (Figure 2b). Furthermore, deletion of the cluster-associated ClpPs homologous to *clpP1* and *clpP2*, *cac16,* and *cac17* respectively, provided further ADEP resistance. This finding supports that Cac16/17 are susceptible to ADEP but resistant to inhibition by *cac* BGC products (Figure 2a).

Having shown that harboring the *cac* BGC results in ClpP1P2 disruption in the heterologous producer, we next sought to develop an assay to easily determine whether this was through the production of a secreted metabolite. To this end, we made use of the transcriptional regulation of ClpP homologs in *S. coelicolor.* Under normal conditions, ClpP1P2 is the primary housekeeping complex, and in conjunction with ClpX targets substrates for degradation, including ClgR, its own transcriptional activator, and PopR, the activator of *clpP3clpP4* (Figure 2c)[31,33,34]. Since PopR is usually degraded, the *clpP3clpP4* operon is not expressed and is only activated upon PopR accumulation due to ClpP1P2 disruption[31]. Validating this reasoning, we observed activation of *clpP3* by reverse-transcriptase PCR (RT-PCR) in the heterologous producer (Supplementary Figure 9a). Therefore, to easily probe for ClpP1P2 disruption in *S. coelicolor* not harboring *cac,* we created the construct pIJGUS-pClpP3, where the *clpP3* promoter drives *gusA* expression, a β-glucuronidase that hydrolyzes the chromogenic substrate X-gluc resulting in the production of a blue pigment[35]. Co-streaking this indicator strain next to *S. coelicolor* M1154 pCGW-cac-LHK resulted in P_{clpP3} activation, as shown by formation of the blue pigment (Figure 2e). Moreover, deletion of a biosynthetic gene, *cac8,* but not *cac16-17,* abolished P_{clpP3} induction (Figure 2e and Supplementary Figure 9b,c). Consistent with our previous results, azabicyclenes C-D did not stimulate P_{clpP3} (Supplementary Figure 9d). These results show that a secreted metabolite produced by *cac* inhibits ClpP1P2. Using this same assay in the context of another ClpP-associated BGC known to produce bohemamines, which have no ClpP activity (Supplementary figure 9d), co-streaking the producer organism *S.* sp. NBRC110035 revealed P_{clpP3} activation (Supplementary Figure 9e). Therefore, the production of ClpP inhibitors is a widespread trait of this family of BGCs.

Next, we tested whether our elusive ClpP1P2 inhibitor could kill actinobacterial species where ClpP is essential. However, in Kirby-Bauer assays using agar plugs inoculated with *S. coelicolor* M1154 pCGW-cac-LHK, minimal zones of inhibition were produced against wildtype *S. coelicolor* (Figure 2e). We hypothesized that these small zones resulted from intrinsic resistance provided by ClpP3P4, as has been observed for other β-lactone ClpP inhibitors[5]. Consistent with this hypothesis, *S. coelicolor* M1154 pCGW-cac-LHK Δ*cac16-17* is viable even though ClpP1P2 is inhibited in this strain. Moreover, *S. coelicolor* Δ*clpP3P4* exconjugants with pCGW-cac-LHK Δ*cac16-17* could not be isolated. Indeed, in agar plug assays with *S. coelicolor* M1154 pCGW-cac-LHK, *S. coelicolor* Δ*clpP3P4* was more susceptible than wildtype to growth inhibition (Figure 2e). In a collection of six other *Streptomyces* spp., those that naturally lacked *clpP3clpP4* were overall more susceptible to growth inhibition than those that encode these homologs (Supplementary Figure 9f), suggesting that this is a general resistance mechanism in the environment. Notably, the growth of *Mycobacterium smegmatis* was also inhibited by the ClpP inhibitor (Figure 2e).

Covalent modification of ClpP guides purification of Clipibicyclene

Given our strong evidence for the production of a ClpP inhibitor by *cac* family BGCs, we wondered why we and others had been previously unable to detect this bioactivity or the responsible metabolite. The ClpP1P2 inhibitory activity of spent media from *S. coelicolor* M1154 pCGW-cac-LHK decreased throughout a five-day fermentation and was lost upon incubation overnight at room temperature or *n*-butanol extract generation (Supplementary Figure 9g). In contrast, azabicyclenes accumulate in fermentation broth over time (results not shown), suggesting that they are not responsible for the activity, and are in fact, shunt metabolites. The decreasing activity we observe from spent media and butanolic extract over time indicates that the active metabolite is unstable. This led us to hypothesize that the compound may possess a reactive warhead that inhibits ClpP by covalent modification of the active site Ser. The active metabolite's instability compared to inert azabicyclenes explains how it evaded detection and may explain why previous research has focused on stable pyrrolizidine alkaloids like bohemamine rather than the active metabolites produced by these BGCs.

To test for covalent modification of ClpP, we first reconstituted the *in vitro* activity of recombinant *S. cattleya* ClpP1 and ClpP2 hetero-tetradecamers (hereafter referred to as $ClpP1P2_{scatt}$). Peptidase activity required both subunits to be present and was stimulated by ADEP (Figure 3a), similar to *M. smegmatis* ClpP1P2[36]. ClpP proteins are produced as pro-peptides and require N-terminal processing to form active peptidases. ClpP1 and ClpP2 were co-incubated to predict processing sites and then analyzed by intact protein liquid chromatography-mass spectrometry (LC-MS). We observed no evidence of $ClpP1_{scatt}$ processing while $ClpP2_{scatt}$ was cleaved only in the presence of $ClpP1_{scatt}$, indicating cross-processing (Supplementary Figure 10a). Processing did not require the presence of cofactors such as ADEP, corresponding closely to what is observed in *M. smegmatis*[36]. We further characterized the role of agonist peptides, substrate selection, and asymmetric interaction with AAA+ ATPases (Supplementary discussion, Supplementary Figure 10b,c). This represents the first time ClpP from an Actinomycetaceae species has been reconstituted *in vitro.*

Next, we incubated recombinant $ClpP1P2_{scatt}$ in spent media from *S. coelicolor* M1154 pCGW-cac-LHK containing the active metabolite. Based on our time course results (Supplementary figure 9g), spent media was harvested after just 24 hours, diverging from our standard protocol for fermentation involving growth for 5-14 days. After recovery of ClpP from spent media, peptidase activity was assayed. The activity of both $ClpP1P2_{scatt}$ and *Escherichia coli* ClpP ($ClpP_{ec}$) was abolished (Figure 3b), and analysis by intact protein LC-MS revealed a mass increase of $ClpP1_{scatt}$ and $ClpP_{ec}$ of ~328 Da (Figure 3b,c). Peptide mapping of $ClpP1_{scatt}$ identified the same 328 Da modification on the active site serine, and fragments observed from MS/MS analysis were consistent with the acyl tail of azabicyclenes (Figure 3d, Supplementary Figure 6). These results are congruent with our hypothesis that the active metabolite derived from the *cac* BGC inactivates ClpP by covalent modification of the active site Ser.

We then analyzed spent media containing the active compound by LC-MS and detected a peak with an *m/z* of 329.14 Da, which corresponds to the mass shift observed from the modification of ClpP. This peak was associated with a parent ion with an *m/z* = 347.15, suggesting *m/z* = 329.14 Da is a dehydration product (Figure 3e). Further, this species was removed from spent media after incubation with recombinant enzyme, confirming that it was the active metabolite (Supplementary Figure 11a). We subsequently isolated the active compound, which we named clipibicyclene, by organic extraction of the supernatant after a 24 hr fermentation. The pure compound was relatively stable once removed from the fermentation broth and could covalently modify and inhibit ClpP1P2$_{scatt}$ and ClpP$_{ec}$ *in vitro* (Supplementary Figure 11b,c).

Complete structure elucidation of clipibicyclene revealed a bicyclocarbamate warhead attached to an acyl tail matching azabicyclene C (Figure 3f, Supplementary Table 9, Supplementary Figure 12, Supplementary Discussion). A similar bicyclocarbamate was previously identified from the ClpP-associated BGC in *P. aeruginosa* PAO1 (Supplementary Figure 1a) but was not recognized as a ClpP inhibitor[28]. To account for the observed mass shift from intact protein LC-MS, we propose a mechanism whereby clipibicyclene's reactive carbamate moiety undergoes nucleophilic attack by the active site Ser of ClpP. This results in elimination of the hydroxyl group attached to the azetidine ring (Figure 3g). Future studies will focus on the biosynthesis of the bicyclocarbamate warhead and the installation of the hydroxyl group on azabicyclene C's acyl tail (See supplementary discussion). Notably, the discovery that clipibicyclene is a ClpP inhibitor represents the identification of the biologically relevant metabolite produced by *cac,* and more broadly, this BGC family.

Since we observed growth inhibition of *Streptomyces* and *Mycobacterial* strains using agar plugs, we tested whether purified clipibicyclene could also kill bacteria in liquid culture. However, we were unable to observe growth inhibition at concentrations up to 512 µg/mL by microbroth dilution. Clipibicyclene was also non-toxic to mammalian cells (HEK293) up to 250 µg/mL. This lack of activity may result from the instability of clipibicyclene in media over the >24 hr incubations required for these assays (Supplementary Figure 11c). We also tested clipibicyclene's ability to inhibit *S. aureus* production of hemolysin, which is attenuated in ClpP mutants[6]. We observed that 64-128 µg/mL of purified clipibicyclene reduced hemolysis induced by *S. aureus* MRSA strains (Supplementary Figure 11d). We hypothesize that modifying clipibicyclene's scaffold may afford more stable derivatives for further clinical development.

Cluster-associated ClpPs are resistant to clipibicyclene inhibition

A fundamental postulate of ClpP-directed genome mining is that cluster-associated ClpPs, in this case, Cac16 and Cac17, provide resistance to the produced antibiotic. Consistently, expression of *cac16-17* in *S. coelicolor* provided resistance to growth inhibition by clipibicyclene (Figure 4a). It also protected cells from P$_{clpP3}$ activation using our *S. coelicolor* pIJGUS-pClpP3 indicator strain (Figure 2d), demonstrating that it performs an analogous function to ClpP1P2 in the ClgR regulatory network (Figure 2b).

To investigate clipibicyclene resistance *in vitro,* we first reconstituted the activity of recombinant Cac16 and Cac17 and characterized their activity. Both Cac16 and Cac17 were N-terminal processed in *trans* (Supplementary Figure 10a), and analogous to ClpP1P2$_{scatt}$, peptidase activity required both subunits to be present and was stimulated by ADEP (Figure 3a). Substrate selectivity and asymmetric interaction with ADEP were also assessed (see Supplementary Discussion and Supplementary Figure 10b,c). Measuring peptidase activity, Cac16/17 complexes were significantly more resistant to clipibicyclene inhibition than either ClpP1P2$_{scatt}$ or ClpP$_{ec}$ (Figure 4b). We also assessed the susceptibility of different ClpP isoforms to covalent modification by intact protein LC-MS using 100 μM clipibicyclene (Figure 4c). ClpP1$_{scatt}$ was modified to a greater extent than ClpP2$_{scatt}$, and ClpP2$_{scatt}$ was not modified at all when co-incubated with ClpP1$_{scatt}$. Cac16 was almost wholly resistant to modification by 100 μM clipibicyclene, while Cac17 was partially modified. Together with cell-based assays, these results show that Cac16/17 provide resistance to clipibicyclene through reduced susceptibility to covalent inactivation. Clipibicyclene's specificity for different ClpP subunits, including ClpP$_{ec}$, ClpP1, ClpP2, Cac16, and Cac17, may result from differences in active site shape or reactivity and warrants further investigation.

We next wondered whether both Cac16 and Cac17 are required for resistance to clipibicyclene. To address this, we tested growth inhibition of *S. coelicolor* Δ*clpP3P4,* expressing various combinations of Cac16, Cac17, and catalytically dead mutants (Cac16$_{S108}$, Cac17$_{S129A,S130A}$). Cac16 alone was necessary and sufficient to provide resistance to clipibicyclene and required that the enzyme be catalytically active (Figure 4a). In contrast, Cac17 was insufficient to provide resistance. We sought to explain this discrepancy by first investigating the minimal ClpP subunits required to support the growth of *Streptomyces.* To this end, we placed the *clpP1clpP2* operon under control of a tightly repressed cumate inducible promoter[37] at its native locus (P$_{A26\text{-}cmtO}$:*clpP2clpP2;* Figure 4d). In a Δ*clpP3P4* background, this allows controlled expression of these essential genes. Complementation was then assessed in the absence of inducer by defined ClpP subunits expressed in *trans.* Since hetero-complexes are required for tetradecamer formation, we tested which subunits were essential for growth by generating inactive variants of each isoform's partner. Both ClpP1$_{S99A}$/ClpP2 and ClpP1/ClpP2$_{S132A}$ supported growth (Figure 4e), showing that only a single catalytically active ClpP isoform is required *S. coelicolor* growth. Co-expression of Cac16/17 supported growth, as did Cac16/Cac17$_{S129A,S130A}$, but Cac16$_{S108}$/Cac17 did not (Figure 4e). Therefore, Cac17 lacks the required catalytic activity or biological function for it to support growth. Combined with *in vitro* data showing partial modification of Cac17 by clipibicyclene, these observations explain Cac17's inability to provide resistance.

Finally, we wondered whether functional tetradecamers were formed strictly between ClpP1/ClpP2 and Cac16/Cac17 or if 'mixed-complexes' of ClpP1/Cac17 and Cac16/ClpP2 were also relevant for resistance and growth. Supporting this hypothesis, mixed complexes were active in *in vitro* peptidase assays (Figure 4f). While complexes containing Cac16 or Cac17 were less active *in vitro* than ClpP1P2$_{scatt}$, this may be an artifact of sub-optimal reaction conditions. *In vivo,* using our inducible system for testing ClpP function in a clean background, Cac16/ClpP2 and ClpP1/Cac17 could support growth (Figure 4e), while only the former provided resistance, consistent with our previous results (Figure 4a). Therefore, ClpP

subunits can mix and match in the cell, and these mixed complexes are physiologically relevant. This provides the first evidence of the complexity of ClpP function in an organism containing more than two isoforms.

Crystal structure of $ClpP_{ec}$ in complex with clipibicyclene

We next determined an x-ray crystal structure of $ClpP_{ec}$ in complex with clipibicyclene to uncover the molecular basis for the observed irreversible inhibition. $ClpP_{ec}$ was modified using pure clipibicyclene until intact MS analysis determined it to be fully modified. Using this material, a $ClpP_{ec}$:clipibicyclene complex was solved to 2.95 Å resolution using the apo form of $ClpP_{ec}$ as a search model[38] (PDB 1TYF, (Supplementary Table 10). The $ClpP_{ec}$ inhibitor complex crystallized with two complete tetradecamers in the asymmetric unit with all catalytic serine residues (S97) modified (Supplementary Figure 13a). The individual tetradecamer complexes were formed by two heptameric rings that produced the typical barrel-like quaternary structure of ClpP[38]. However, the N-terminal residues (7-17) that form each heptameric ring's axial pores were disordered and could not be fully modeled (Supplementary Figure 13b). The N-terminal region of ClpP is known to exhibit high conformational flexibility, and thus it is unlikely that disorder in these residues was caused by clipibicyclene binding[39]. Moreover, each $ClpP_{ec}$ monomer (18-193) adopted the typical α/β fold and did not significantly differ from the apo form as judged by superposition of the tetradecamer (RMSD = 0.38 Å over 16,148 atoms; Supplementary Figure 13c).

Positive peaks in the electron density map extended from the catalytic Ser97 residues in all 28 chains, indicating covalent modification and full occupancy of the clipibicyclene adduct (Supplementary Figure 14). The structures of the modified Ser97 residues were built and restrained under our proposed reaction mechanism and fit well within the electron density map (Figure 3g). The adducts were defined by a carbamoyl linkage between the O□ of Ser97 and the nitrogen of the azetine ring (Figure 5a). As expected, the seven-membered ring of clipibicyclene is no longer intact. Instead, it is replaced by an imide moiety connecting the azetine portion of the adduct to the linear aliphatic chain, which showed a high degree of flexibility (Figure 5a). Many of the aliphatic chains possessed elevated B-factors and weak electron density, which precluded them from being built. However, three complete adduct complexes could be placed (chains A, O, and W), revealing different aliphatic tail conformations (Supplementary Figure 14). We use the adduct from chain A as the representative conformer to describe the $ClpP_{ec}$:clipibicyclene complex.

In the active site context, the clipibicyclene adduct points clockwise relative to the interior circumference of the heptameric ring of $ClpP_{ec}$ (Figure 5b). The azetine ring is positioned between sheets β4 and β9 in the hydrophobic S1 subsite and forms van der Waals contacts with I70, M98, P124, and L125 (Figure 5c,d). The carbonyl oxygen of the carbamate linkage forms a critical hydrogen bond with the backbone amide of Gly68, which usually serves as an H-bond donor in the oxyanion hole (Figure 5d). The linear portion of the adduct extends toward the S’ subsites over the carbonyl carbon across the β4 strand and partially protrudes into the solvent. To our knowledge, this is the first example of an inhibitor that

occupies this region of a ClpP active site (Figure 5c, d). Only one potential H-bond is formed between the side chain of Q134 and the imide moiety, and a van der Waals contact is likely made with P66 and the aliphatic tail. This limited interaction network along the tail probably explains the disorder associated with the linear aliphatic portion of the adduct (Figure 5c,d). Except for a few outliers, these interactions were observed in the majority of chains. Based on this model, irreversible inhibition can be partly explained by both the inherent stability of the carbamoyl moiety and/or steric exclusion of water from the catalytic base (His122), preventing any base-catalyzed hydrolysis of the adduct.

Discussion

By applying target-directed genome mining, here we describe the association of a widespread family of BGCs encoding bimodular NRPSs with the highly conserved and often essential protease ClpP. Despite being a relatively well-known BGC family for their production of pyrrolizidine alkaloids, the function and biologically relevant molecules produced by this group of BGCs has previously evaded discovery. Using a combination of heterologous expression, *in vivo* assays focused on *Streptomyces* biology and *in vitro* enzyme reconstitution to guide inhibitor purification, we identified clipibicyclene, a representative ClpP inhibitor, as the active metabolite produced by *cac*. The production of ClpP inhibitors is likely a governing property of this family of BGCs. The presence of peptidases in other bimodular NRPS BGCs suggests that they may produce other Ser hydrolase inhibitors such as the brabantamides[25,26]. Ser hydrolase inhibitors hold potential for various clinical indications, and as a ClpP inhibitor, clipibicyclene may hold promise as a scaffold for antibiotics against the important human pathogen, *Mycobacterium tuberculosis,* or virulence inhibitors for *S. aureus*. In addition to drug development, the abundance and widespread nature of this family of BGCs, from *Streptomyces* to *Pseudomonas,* is unusual compared to other BGCs families and suggests that the metabolites produced are important to bacterial physiology. In addition to informing eco-evolutionary hypotheses about the ubiquity of these BGCs, our identification of clipibicyclene as a ClpP inhibitor is essential to deciphering their function in human pathogens, including *P. aeruginosa, K. pneumoniae,* and *E. cloacae.* Further investigation is warranted, as production of metabolites from the ClpP-associated *aze* BGC in *P. aeruginosa* PAO1 causes reduced growth and virulence in a *Galleria mellonella* model[28].

Numerous antibiotics act on and remodel the transcriptional and translational landscape of the cell, and by doing so, play essential roles during infection, adaptation, and survival[40–42]. Post-translational regulation is likely no less critical in coordinating cellular activities, and so it should be expected that specialized metabolites have also evolved to modulate this level of regulation. Fully elucidating the function, biosynthesis, and mechanism of clipibicyclene and related compounds will be the first step towards understanding post-translational control of the proteome through ClpP.

Methods

Strains and Culture Conditions

All primers, gBlocks, and strains used in this study are listed in Supplementary Tables 11-13, respectively. *Streptomyces* strains were routinely grown at 30°C, 250rpm, in TSBY (tryptone-soya-broth + 5 g/L yeast extract; BD Biosciences) for starter cultures and genomic DNA preparation, SM+$MgCl_2$ (2% each D-mannitol, soya flour, agar, 10mM $MgCl_2$) for sporulation and conjugation, and Bennett's media (1% potato starch, 0.2% casamino acids, 0.18% yeast extract, 0.02% KCl, 0.02% $MgSO_4 \cdot 7H_2O$, 0.024% $NaNO_3$, 0.0004% $FeSO_4 \cdot 7H_2O$) for fermentation. Cloning was performed in *E. coli* Top10 using standard conditions and routinely grown in LB media (Bioshop). Antibiotics were supplemented as necessary (50 µg/mL kanamycin, 100 µg/mL ampicillin, 50 µg/mL apramycin, 35 µg/mL chloramphenicol, 25 µg/mL nalidixic acid, 100 µg/mL or 50 µg/mL hygromycin B for *E. coli* and *Streptomyces* spp., respectively).

All pIJ10257, pSET152, and pCRISPomyces-2 derived constructs were introduced into *S. coelicolor* M1154 via biparental conjugation using *E. coli* ET12567/pUZ8002 *neo::bla* using standard protocols[43]. All pCGW derived constructs were introduced via triparental conjugation using *E. coli* ET12567/pCGW-XXX and *E. coli* ET12567/pR9406 by standard protocols[43].

ADEP was purified from *S. hawaiiensis* NRRL15010 as previously described with minor modifications[19]. Bohemamines were isolated from *S.* sp. NBRC110035 by fermentation in SMP media (2.5% soluble starch, 0.146% glutamine, 0.1% K_2HPO_4, 0.1% NaCl, 0.05% $MgSO_4 \cdot 7H_2O$, 4 × 10−6% $ZnCl_2$, 2 × 10−5% $FeCl_3 \cdot 6H_2O$, 1 × 10−6% each $CuCl_2 \cdot 2H_2O$, $MnCl_2 \cdot 4H_2O$, $Na_2B_4O_7 \cdot 10H_2O$, $(NH_4)6Mo_7O_{24} \cdot 4H_2O$) for 7 days and purified using a procedure similar to that for azabicyclenes described below.

MICs were assessed using standard microbroth dilution protocols. *Streptomyces* spp. were assayed in TSB, *M. smegmatis* in 7H9 media, and all other strains in Mueller-Hinton Broth (BD Biosciences).

In silico detection and analysis of ClpP associated clusters

First, we collected all ClpP homologs from Actinobacteria in PFAM (PF00574) to assemble a complete and diverse set of relevant ClpP-like proteins. This collection of ClpP protein sequences were used to query prokaryotic RefSeq proteins using blastp, and the top five hits for each query were taken. Proteins were then mapped to RefSeq genomes with the identical protein report function. Using a custom python script, 50 kb upstream and downstream of each *clpP* hit was collected, and BGCs were identified using AntiSMASH[22]. Further filtering was performed in Geneious V8.9.1. Clusters containing ClpP homologs were identified by tblastn of *S. coelicolor* ClpP1 (NP_626855.1). Distinctive genetic context for *clpP* paralogs is as follows: *clpP1clpP2* is associated with the AAA+ ATPase *clpX, clpP3clpP4* is associated with the transcriptional regulator *popR,* and *clpP5* is associated with a distinctive gene encoding a protein of unknown function. Tblastn was used to identify these paralogs from protein sequences: ClpX (NP_626853.1), PopR (NP_631335.1, WP_067791847.1, WP_028564420.1), ClpP5-associated hypothetical protein (NP_625527.1). Eight BGCs where *clpP* was located on the contig edge were also discounted.

To count how many Ser-Pro bimodular NRPS or streptothricin-type BGCs were present in RefSeq, key proteins (WP_014143997.1 and WP_037694042.1, respectively) were taken as BlastP queries against the NCBI non-redundant protein sequence database (accessed January 7, 2021), allowing up to 5000 hits and excluding *P. aeruginosa* due to its overwhelming abundance. Amino acid % identity cut-offs were set manually by determining which cut-off separated hits belonging to the BGC family from unrelated clusters. To this end, hits were first sorted by genus. Identical protein sequences in RefSeq were taken to obtain genomic context, and ~10 kb on either side of the query protein was analyzed by AntiSMASH to determine if the hit belonged to the BGC family. Hits above the % identity cut-off were taken, and identical sequences were removed to give a final count of BGCs in unique species.

30 BGCs containing the biomodular NRPS for CORASON analysis were collected using an arbitrary amino acid identity cut-off of 54% vs. Cac9 (WP_014143997.1; Supplementary Table 2). Five *Pseudomonas* spp. BGCs with amino acid identity <54% were also included for comparison. CORASON was run on contigs containing these 30 BGC using Cac9 (NRPS) as the query gene, *cac* from *S. cattleya* as the reference BGC, and a bit score cut-off of 1000.

Cloning cac using TAR

pCGW is a capture construct derived from pCAP03[20], and modified to use the 'oriV-ori2-repE-sopABC' single copy origin of replication from pBAC-lacZ as previously described[44]. pCGW was maintained in *E. coli* EPI300, which controls the expression of *trfA* required for high copy amplification from *oriV* with an arabinose inducible promoter. When necessary for miniprep and plasmid mapping, high copy number was induced using 1 mM arabinose. The vector backbone contains all the elements required for propagation and selection in *E. coli* (pUC *ori* and Kan^R), yeast (ARSH4/CEN6 and TRP1 auxotrophic marker), and *Streptomyces* (φC31 integration, *oriT,* and Kan^R).

The boundaries of *cac* were defined by comparison with homologous clusters in *Streptomyces pini* PL19 NRRL B-24728, *Streptomyces barkulensis* RC1831, and *Kitasatospora phosalacinea* NRRL B-16228 (Supplementary Figure 15). 50 bp homology arms flanking *cac* (*orf-3* to *cac18*) were concatenated with an MssI site in between, and 18 bp overlaps with the pCGW backbone were added to either end. The gBlock (Integrated DNA Technologies; IDT) was cloned into pCGW linearized with NdeI/XhoI by Gibson assembly (Supplementary Table 12). Before TAR cloning, the capture vector was digested with MssI to release the hooks.

HMW genomic DNA was prepared from *S. cattleya* DSM 46488 and digested with XhoI and NdeI to cut surrounding, but not within, the BGC. *Saccharomyces cerevisiae* VL648N sphereoplasting and transformation were carried out as previously described[45]. Twenty-two colonies were restreaked on SD-Trp plates and screened by colony PCR for the desired insertion using diagnostic primers located inside *cac,* giving one positive colony. DNA was extracted from the positive yeast colony by zymolyase treatment, followed by phenol/chloroform extraction and ethanol precipitation. This DNA was transformed into *E. coli* EPI300 cells by electroporation and plasmids were selected with kanamycin. *E.*

coli EPI300 was induced for high copy number using 1 mM arabinose, and the resulting construct, pCGW-*cac,* was confirmed by restriction mapping.

Refactoring pCGW-cac

pCGW-*cac* was refactored using a combination of yeast recombination and λ-red recombineering in *E. coli.* All primers, gblocks, strains, and plasmids are listed in Supplementary Tables 11-13. First, Leu2 and His3 auxotrophic selection cassettes were constructed to contain orthogonal promoters and terminators using pSASS series plasmids, a gift from Dr. Mike Tyers (University of Montreal). The *leu2* open reading frame was PCR amplified from pRS316 and inserted into pSASS5, and *his3* was PCR amplified from pYAC10 (Mike Tyers) and inserted into pSASS4, using Gibson assembly. Selection cassettes containing the marker along with yeast promoter and terminator sequences were subsequently PCR amplified from pSASS5-*leu2* and pSASS4-*his3.* For replacement of the *cac14* promoter, XNR_1700p was PCR amplified from *Streptomyces albus* gDNA[46], and stitched to the *leu2* selection cassette with 500 bp homology arms to *cac* on either side through overlap extension PCR. The resulting refactoring cassette was co-transformed with pCGW-cac into *Saccharomyces cerevisiae* SASy31/SASy35 by standard lithium acetate/single-stranded carrier DNA/PEG mediated transformation[47]. Recombinants were selected on SD-trp-leu plates, screened by colony PCR, and the resulting construct, pCGW-*cac*-L was recovered and confirmed in *E. coli* EPI300, similar to initial TAR cloning. Replacement of *cac8* and *cac9* promoters were achieved similarly using the His3 selection marker sandwiched by synthetic *Streptomyces* promoters, A26 and A35, incorporated into primer sequences[48]. This time, *S. cerevisiae* SASy31 transformants were selected using SD-trp-leu-his plates, and the resulting construct, pCGW-cac-LH, was moved to *E. coli* EPI300.

After refactoring *cac8, cac9,* and *cac14,* RT-PCR revealed that *cac5* and the putative operon, including *cac4, cac3, cac2,* and *cac1,* was still poorly transcribed (data not shown). Since few selection markers were left for use in *S. cerevisiae,* we chose to use *E. coli* λ-Red recombineering. The promoter *kasO*p* and T7 terminator were designed adjacent to the apramycin resistance gene *aac*(3)IV flanked by PmeI/HpaI restriction sites to remove *aac*(3)IV after recombination. Thirty-nine bp homology arms were added to each end to direct recombination. The recombineering cassette was synthesized as a gBlock by IDT, and recombineering was carried out using standard protocols[49]. pCGW-cac-LH was transformed into *E. coli* BW25113/pKD46, selected using kanamycin and ampicillin, and grown at 30°C to maintain pKD46. The strain was grown in LB overnight at 30°C, then sub-cultured in SOB without $MgSO_4$ with the addition of 10 mM arabinose to induce expression of *red* genes from pKD46. After reaching OD_{600} = 0.6, the cells were recovered by centrifugation, washed twice with ice-cold 10% glycerol, and electroporated with 1 µg of linear refactoring cassette. Successfully recombinants were selected with apramycin and grown at 37°C to promote loss of pKD46. The resulting construct, pCGW-cac-LHK-apra, was extracted from *E. coli* BW25113, transformed into *E. coli* EPI300, and verified by restriction mapping. Finally, *aac*(3)IV was removed by digestion with HpaI and PmeI followed by intramolecular blunt end ligation with T4 DNA ligase, generating pCGW-cac-LHK.

Recombineering pCGW-cac-Δcac16-17 and pCGW-cac-Δcac8

In-frame deletions of *cac16-cac17* and *cac8* that leave downstream genes intact were created using λ-Red recombineering as described above and primers listed in Supplementary Table 11. A refactoring cassette with *aac*(3)IV, HpaI/PmeI restriction sites, and 40 bp homology arms was PCR amplified from the apra cassette used to recombine *cac5* then transformed into *E. coli* BW25113/pKD46/pCGW-cac-LHK. The *aac*(3)IV cassette was again removed by HpaI and PmeI digestion followed by blunt-end ligation, generating pCGW-cac-LHK-Δ*cac16-17* and pCGW-cac-LHK-Δ*cac8.*

RT-PCR measurement of cac and clpP regulon expression

Strains for analysis were inoculated 1:100 from a saturated TSBY seed culture into 60 mL Bennett's media in 250 mL baffled flasks, and pellets were taken at 24 hr. Cells were lysed by bead beating mycelium with 4 mm glass beads in 5 mL TRIzol reagent (Invitrogen), and RNA was extracted using the manufacturer's recommendations. RNA from the resulting aqueous phase was extracted a second time using acid phenol/chloroform, then combined with a half volume of anhydrous ethanol, and finally purified using PureLink RNA Mini Kit (Invitrogen). Maxima H Minus First Strand cDNA synthesis kit with dsDNase (Thermo Scientific) was used for cDNA synthesis, and PowerUp SYBR Green master mix (Applied Biosystems) was used for RT-PCR quantification on a BioRad CFX96 real-time system. Primers targeting genes of interest (Supplementary Table 11) were designed, and 80-100% efficiency was verified before quantification. Analysis was performed on three or four independent fermentations and quantified in technical duplicate. Technical duplicates for each biological replicate were averaged, then fold change expression for each replicate was calculated by normalizing to *hrdB* expression using the ΔCt method. Statistical analysis of *clpP3* gene expression was performed using GraphPad Prism V6. Multiple comparisons to pCGW were made using a two-sided Kruskal-Willis analysis with Dunn's test for multiple comparisons ($n = 3$).

Crude extract LCMS analysis

To prepare crude extracts, strains were fermented in 50 mL media for 7 days, unless otherwise indicated, extracted with 15 mL *n*-butanol, dried under vacuum, and resuspended in 100 µL DMSO. Extracts were analyzed on an Agilent 1290 UPLC with G6550A Q-TOF using a ZORBAX StableBond C18 column (Agilent, 4.6 x 150 mm, 3.5 µm, 80 Å, 0.4 mL/min, Buffer A water + 0.1% formic acid, Buffer B acetonitrile + 0.1% formic acid) and the following gradient: 0-1 min 95% B, 1-45 min 25-75% B, 45-50 min 75-95% B, 50-57 min 95% B, 57-59 min 95-5% B.

Purification and structural elucidation of azabicyclenes

S. coelicolor pCGW-cac-LHK seed culture was grown in TSB with 50 µg/mL kanamycin for 3 days, then spent media was removed, and 50 mL worth of mycelium was inoculated into each of 25x600 mL Bennett's media in 3 L flasks. Fermentations were grown for 7 days at 30°C, 250 rpm. 13 L of spent media was harvested and extracted with 390 g HP-20 resin (Diaion). The resin was washed with 10% methanol

(MeOH), then eluted with 100% MeOH and concentrated under vacuum. Dry material was extracted with ethyl acetate:MeOH (1:1) and dried onto 5 g silica gel (Sigma) under vacuum. Normal phase vacuum liquid chromatography was performed using the following stepwise gradient: 1. hexanes, 2. ethyl acetate, 3. ethyl acetate:MeOH (3:1), 4. ethyl acetate:MeOH (1:1). Fractions 2-4 were combined, dried, and using DMSO, applied to an 86 g reverse-phase CombiFlash ISCO (RediSep Rf C18, Teledyne) eluted with a linear gradient system (5-45% water/acetonitrile, 0.1% formic acid). Fractions containing azabicyclenes were pooled and further purified by semipreparative HPLC using a ZORBAX StableBond C18 column (Agilent, 9.4 x 150 mm, 5 µm, 3 mL/min, Buffer A water + 0.1% formic acid, Buffer B acetonitrile + 0.1% formic acid). Compound **2** was purified by the following gradient: 0-3 min 25.5% B, 3-13.5 min 25.5-26.3% B, 13.5-14 min 26.3-95% B, 14-18 min 95% B, 18-19 min 95-25.5% B. Azabicyclene C and Compound 1 were purified by the following gradient: 0-1 min 30% B, 1-3 min 30-38% B, 3-8 min 38-43% B, 8-21 min 43-44.1% B, 21-22 min 41.1-95% B, 22-24 min 95% B, 24-25 min 95-30% B. Azabicyclene D (11.7 mg) was purified by the following gradient: 0-1 min 13% B, 1-5 min 13-20% B, 5-25 min 20-23.7% B, 25-26min 23.7-95% B, 26-30 min 95% B, 30-31 min 95-13% B. Care was taken to protect compounds from light.

For structural elucidation, compounds **1**, **2**, azabicyclene C (**5**), and azabicyclene D (**6**) were subjected to 1D and 2D NMR and MS/MS. See Supplementary Discussion, Supplementary Figures 2-5, and Supplementary Tables 5-8. NMR spectra were recorded using a Bruker AVIII 700 MHz instrument equipped with a cryoprobe. Chemical shifts are reported in parts per million (ppm) referenced to d^4-MeOH (δ_H: 3.31 ppm and δ_C: 49.0 ppm). Coupling constants (J) are rounded to the nearest 0.5 Hertz (Hz). Multiplicities are given as multiplet (m), singlet (s), doublet (d), triplet (t), quartet (q), quintet (quin.), sextet (sext.), septet (sept.), octet (oct.) and nonet (non.). ^{1}H and ^{13}C assignments were established based on COSY, DEPT, HSQC, and HMBC correlations. HR-ESI-MS data were acquired using an Agilent 1290 UPLC separation module and qTOF G6550A mass detector in positive-ion mode. The structure of azabicyclene B (**4**) was predicted from LC-MS/MS fragmentation and precedence given by the structure of compound 2 (Supplementary Figure 6).

Purification and structural elucidation of clipibicyclene

S. coelicolor pCGW-cac-LHK culture was grown in the same way as for azabicyclene purification, except for that the fermentation was harvested after 24 hr. After the growth of a 2 L culture, mycelium was removed, and the supernatant was extracted twice by a total volume of 4 L dichloromethane. The organic phases were combined and concentrated gently *in vacuo* to obtain a yellow oil. The crude extract was resuspended in 200 µL DMSO and purified on a Waters XSelect CSH preparative C-18 HPLC column (10 × 100 mm, 5 µm, 130 Å, 3 mL/min, Buffer A water + 0.1% formic acid, Buffer B acetonitrile + 0.1% formic acid)) using the following method: 0-2 min 25% B, 2-19 min 25-100% B, 19-20 min 100% B, 20-21 min 100-25% B. This afforded 2.0 mg of clipibicyclene, which was subject to LCMS and NMR structural elucidation as described for azabicyclenes (see Supplementary discussion; Supplementary Figure 12, Supplementary Table 9). Chemical shifts are reported in parts per million (ppm) referenced to d^6-DMSO (δ_H: 2.50 ppm and δ_C: 39.5 ppm).

Kirby-Bauer assays

Indicator *Streptomyces* strains were grown for 2-3 days in TSBY media with antibiotic selection as necessary until saturated. Cultures were diluted to OD_{600} = 0.1 and streaked on Bennett's agar using a sterile cotton swab. For assays involving ADEP, 50 µg compound in DMSO was placed on a sterile cellulose disk. For assays involving agar plugs, *S. coelicolor* pCGW-*cac*-LHK or *S. coelicolor* pCGW was inoculated on Bennett's agar from a saturated TSBY starter culture and grown for 24 hr before removal of an agar plug that was placed on a plate inoculated with the indicator strain.

GusA indicator strain construction and assays

A cassette consisting of the T7 terminator placed upstream of P_{clpP3}:*gusA* was constructed in the pIJ10257 backbone. Using the primers listed in Supplementary Table 11, parts PCR were amplified as follows: *clpP3* promoter region lacking its RBS from *S. coelicolor* gDNA*, gusA* including its RBS from pGUS[35], and T7 terminator from pDR3K[50]. PCR products were cloned into pIJ10257 digested with EcoRV using Gibson assembly to generate the construct pIJGUS-pClpP3. The construct was moved into *S. coelicolor* M1154 in addition to pSET152 (empty vector) or pSET152-cac16-17, cloned between XbaI and EcoRI restriction sites using primers listed in Supplementary Table 11.

To perform assays for P_{clpP3} activation, saturated starter cultures of *S. coelicolor* M1154 pIJGUS-pClpP3 pSET152 and *S. coelicolor* pCGW-XXX were swabbed next to each other on 0.25x Bennett's agar containing 200 µg/mL X-gluc (Alfa Aesar). SMP agar instead of Bennett's was used for *S.* sp. NBRC110035. Plates were imaged after 3-4 days (*S. coelicolor)* or 7 days (*S.* sp. NBRC110035).

For time-course analysis of clipibicyclene production, 250 mL flasks containing 50 mL Bennett's media were inoculated with *S. coelicolor* M1154 pCGW-*cac*-LHK on sequential days. Cultures that had been growing for 1-5 days were harvested simultaneously, and mycelium was removed by centrifugation. Cultures were split and either stored overnight at -80°C, at room temperature, or extracted with butanol, dried under vacuum, and resuspended at 1/150 the original volume in DMSO. 2.5 µL of each sample was spotted on an agar plate to test for P_{clpP3} activation as above.

Hemolysis assay

S. aureus hemolysis assays were performed with strains USA-300 JE2, NE912 (JE2 *clpP*::NΣ)[51], and the clinical isolate CMRSA-3, which was isolated from Mount Sinai Hospital, Toronto, ON. Overnight cultures of each strain were subcultured in MHB medium at 37 ºC until OD_{600} of 0.6 was reached. The bacteria were then pelleted by centrifugation and resuspended in an equal volume of fresh MHB medium. The cell suspension was dispensed in a 96-well round bottom plate and mixed with 2-fold serial dilutions of inhibitor or vehicle. The final concentration of DMSO did not exceed 1% (v/v). The bacteria were grown for a further 5 hr and then centrifuged (5000 x g, 10 min) to remove the cells. The supernatant was added to a 5% (w/v) suspension of sheep erythrocytes in phosphate buffered saline and incubated at 37 ºC for 1 hr. To determine the relative extent of hemolysis, any intact erythrocytes were pelleted by centrifugation

(5000 x g, 10 min) and the released haemoglobin was measured in the supernatant by monitoring absorbance at a wavelength of 545 nm. The experiments were performed on two independent occasions.

Expression and purification of ClpPs

Untagged *E. coli* ClpP was expressed from *E. coli* BL21(DE3) pET9a-EcClpP, a kind gift from Dr. Walid Houry (University of Toronto), and purified as previously described with minor modifications[52]. *S. cattleya* ClpP1 (SCATT_17350), ClpP2 (SCATT_17340), Cac16 (SCATT_32700) and Cac17 (SCATT_32710) were codon-optimized for expression in *E. coli* and synthesized as gBlocks by IDT (Supplementary Table 12). Synthesized DNA was cloned into pET-28 digested with NcoI and NotI using Gibson assembly to be expressed with a C terminal 6xHis tag. Constructs were transformed into *E. coli* BL21(DE3) 1146D (Δ*clpP::*cm^r). For protein expression, overnight cultures were inoculated into 1 L LB with appropriate antibiotics in a 3 L flask, grown at 37°C until cultures reached OD_{600} = 0.6-0.8, then induced with 0.5 mM IPTG. Cultures were grown for 18 hr at 17°C before harvesting cell pellets by centrifugation.

Cell pellets were resuspended in 20 mL lysis buffer (20 mM Tris (pH 8), 300 mM KCl, 10 mM imidazole, 10% glycerol), treated with 10 mg/mL lysozyme and 5 µg/mL DNase, and lysed on ice by sonication. Clarified lysates were loaded onto equilibrated 2 mL Ni-NTA agarose (Qiagen), washed (20 mM Tris, pH 8, 300 mM KCl, 25 mM imidazole, 10% glycerol), and eluted (20 mM Tris, pH 8, 300 mM KCl, 250 mM imidazole, 10% glycerol). $ClpP1_{scatt}$ and $ClpP2_{scatt}$ were dialyzed into buffer A for further purification by anion exchange. Anion exchange was performed on a HiTrap Q HP 5mL column (GE Healthcare) (Buffer A: 20 mM Tris-HCl, pH 8, 10% glycerol, 1 mM DTT; Buffer B: 20 mM Tris-HCl, pH 8, 10% glycerol, 1 mM DTT, 1 M KCl). Cac16 and Cac17 were further purified by gel filtration on a HiLoad 16/60 Superdex 200 column (GE healthcare) (Buffer C: 50 mM Tris-HCl, pH 8, 10% glycerol, 200 mM KCl). Protein purity was >95% as assessed by SDS-PAGE analysis, and proteins were buffer exchanged and concentrated into Buffer C using Amicon Ultra centrifugal filters with 50 K cut-off.

In vitro ClpP assays

Reactions involving $ClpP_{ec}$ were carried out under the following conditions: 2 µM ClpP (monomer), 200 µM N-Succinyl-Leu-Tyr-7-amido-4-methylcoumarin (SLY-AMC; Cayman chemicals), ClpP reaction buffer (50 mM HEPES, pH 8, 100 mM KCl, 10% glycerol). Reactions involving *S. cattleya* ClpPs were carried out under the following conditions: 5 µM each ClpP subunit (monomer, unless otherwise indicated), 50 µM ADEP, 200 µM substrate in ClpP reaction buffer. Unless otherwise indicated, N-Succinyl-Leu-Leu-Val-Tyr-AMC (S-LLVY-AMC; R&D Systems) or (Suc-LLVY)2-Rhodamine110 (AAT Bioquest) were used as substrates. Quenching of AMC fluorescence was observed above 50 µM azabicyclene/bohemamines, so (Suc-LLVY)2-Rhodamine110 was used for all assays containing these compounds. Where indicated, agonist peptides (Sigma) were tested at 500 µM. Where indicated, clipibicyclene, azabicyclene, or bohemamines were preincubated with enzyme for 10 min before substrate addition. 100 µL reactions were initiated with the addition of substrate, carried out at room temperature, and tracked by fluorescence

excitation/emission 360 nm/460 nm (AMC) or 485 nm/525 nm (rhodamine 110). Rates were calculated using the slope over the first 15 min of the reaction.

Thermal shift assays for ADEP binding

Melt curves were performed using the following conditions: 5 µM ClpP subunit (monomer), 5x SYPRO orange dye (5000x stock, ThermoFisher), 100 µM ADEP (or DMSO), in ClpP reaction buffer. 50 µL per well were assayed on a BioRad CFX96 real-time system, ramping from 35°C to 95°C in increments of 0.5°C for 10 seconds, reading SYPRO fluorescence after every increment. Melting temperatures were calculated using CFX Maestro software.

Intact protein LCMS

Intact protein LCMS was performed on an Agilent 6546 LC/Q-TOF or Agilent 1290 UPLC and G6550A Q-TOF in positive ion mode, with a ZORBAX StableBond 300 C3 column (Agilent, 3.0 x 150 mm, 3.5 µm) using the following method. Gas Temp: 200°C, Gas Flow: 14 l/min, Fragmentor: 380, Buffer A: Water + 0.1% formic acid, Buffer B: Acetonitrile + 0.1% formic acid, flow rate: 0.4 mL/min, gradient: 0-1 min 95% B, 1-4 min 5-38% B, 4-20 min 38-55% B, 20-24 min 55-95% B, 24-27 min 95% B, 27-28 min 95-5% B.

In general, 1 µL of ~30-40 µM protein was injected for analysis. Spectra were deconvoluted, and figures were generated using UniDec software[53]. To assess cross-processing, subunits were mixed in a 1:1 molar ratio and incubated overnight at room temperature. Processing sites were predicted using ExPASy FindPept.

To label ClpPs in spent media, the fermentation broth was first filtered through an Amicon Ultra centrifugal filters with 5 K cut-off to remove secreted proteins present in the media and neutralized by adding 50 mM Tris-HCl pH 7.5. Next, 5 µg/mL ClpP, either total ($ClpP_{ec}$) or each subunit ($ClpP1P2_{scatt}$), was added and incubated for 1.5 hr at room temperature. ClpP protein was recovered, and buffer exchanged using an Amicon Ultra centrifugal filters with 30 K cut-off before LCMS analysis or peptidase assays. To label ClpP with pure clipibicyclene, 20 µM of each ClpP subunit was first incubated in ClpP reaction buffer for 3 hr at room temperature to allow processing to occur before the addition of 100 µM clipibicyclene and LCMS analysis.

Peptide mapping LC-MS/MS

To identify the site of ClpP1 modification, modified enzyme was prepared using a modified version of the method used prior to intact protein LC-MS. ClpP1 (50 µg) was incubated with spent media filtrate of *S. coelicolor* pCGW or *S. coelicolor* pCGW-*cac*-LHK cultures (10 mL) for 1.5 h at 25º C and concentrated to 1 mg/mL. Five times the sample volume of cold acetone (-20 ºC) was added, and the mixture was and incubated for 10 min at -20 ºC to completely precipitate the protein. The protein floc was collected by centrifugation (12,000 x g, 2 min), and the pellet was resuspended in cold acetone (1 mL). This was repeated three times to remove residual media components. The supernatant was removed from the final

pellet, and the protein was dissolved in 50 mM ammonium bicarbonate buffer pH 8 (20 µL). To proteolytically digest ClpP1 in-solution for peptide mapping, sequencing-grade trypsin:ClpP1 ratio of 1:25 (w/w) was used, and the reaction was incubated at 37 ºC for 16 h. The resultant peptides were analyzed by LC-MS/MS with an Agilent G6550A Q-TOF in positive ion mode. Peptides were separated with a C18 column (Agilent XDB C18 100 mm x 2.1 mm; 2.7 µm) equilibrated with 5% acetonitrile in 0.1 % formic acid. A linear gradient to 95% acetonitrile in 0.1 formic acid was applied over 8 min at a flow rate of 0.5 mL/min. The mass spectrometer was operated with the following parameters: Gas Temp: 350°C, Gas Flow: 14 l/min, Capillary voltage: 4.0 kV, Nebulizer pressure: 40 psi, Fragmentor: 150. Data analysis was performed using Mmass (www.mmass.org).

CRISPR-editing and complementation of S. coelicolor ClpPs

CRISPR-editing constructs containing sgRNA targeting sequences were cloned using the pCRISPomyces-2 backbone using primers listed in Supplementary Table 11 via golden-gate assembly[54]. *S. coelicolor clpP3clpP4* homology arms were designed to introduce a scarless in-frame deletion and cloned using Gibson assembly. For inducible control of *clpP1clpP2*, a cassette consisting of the cumate-responsive repressor *cymR,* the strong synthetic promoter A26, and the *cymR* operator sequence *cmtO* was designed and synthesized as a gBlock (IDT; Supplementary Table 12)[37,48]. This cassette was PCR amplified using primers with suitable homology arms for Gibson assembly. Construction of the final pCRISPomyces-2 construct was performed in two Gibson assembly steps: first homology arms for *clpP1clpP2* were introduced flanking a HindIII site, and second, the promoter cassette was introduced into the HindIII site.

CRISPR-editing was performed as previously described[55]. Briefly, pCRISPomyces constructs were conjugated to *S. coelicolor* M1154; exconjugants were restreaked and screened using PCR for successful editing, pCRISPomyces-2 was cured by growth at 37°C, and exconjugants were verified again by PCR. Conjugation and maintenance of the *clpP1clpP2* edited strain were performed on media containing 100 µM cumate.

Complementation plasmids pIJ-ncac16-17, pIJ-nclpP1P2, pIJ-nclpP3P4, and pIJ-clpP1-cac17 were cloned by PCR amplifying inserts from *S. coelicolor* (*clpP1clpP2* and *clpP3clpP4*) or *S. cattleya* (*cac16*-17) gDNA using primers listed in Supplementary Table 11 before Gibson assembly into pIJ10257 digested with KpnI and HindIII. For pIJ-cac16-clpP2, the plasmid backbone including *cac16* was PCR amplified from pIJ-ncac16-17, and *clpP2* was PCR amplified from gDNA for Gibson assembly. All remaining constructs, including pIJ-ncac16, pIJ-ncac17, and catalytically inactive variants (pIJ-ncac16d17, pIJ-cac16-17d, pIJ-clpP1dP2, pIJ-clpP1P2d) were cloned using site-directed mutagenesis by PCR amplifying the full constructs followed by DpnI digestion and ligation. We observed that native promoters provided superior expression than the commonly used constitutive *Streptomyces* promoter ermEp*, so they were retained in these constructs. Since Cac17 contains two adjacent serine residues in its active site, both were converted to alanine to ensure the enzyme was catalytically inactive.

Crystallization and structure determination of $ClpP_{ec}$:clipibicyclene

Preparation of $ClpP_{ec}$:clipibicyclene for crystallization involved incubating purified enzyme (2 mg/ml) with 0.5 mM clipibicyclene (5-fold molar excess) in ClpP reaction buffer for 1 hour at room temperature. Intact protein ESI-MS verified complete covalent modification. Before crystallization, the covalent complex was buffer exchanged into 10 mM HEPES pH 7.5, 100 mM NaCl using a BioGel P6-DG column, and was concentrated to 12 mg/ml. Crystals of $ClpP_{ec}$:clipibicyclene were grown at a 2:1 precipitant to protein ratio using 0.1 M sodium acetate pH 5.6, 30% (*v/v*) MPD, and were flash cooled in a N_2 cryo stream.

Data collection was conducted at the Canadian Light Source CMCF-BM (08IB1), Saskatoon, SK, Canada. The x-ray data was processed using autoPROC[56], XDS[57], and CCP4[58]. The structure of $ClpP_{ec}$:clipibicyclene determined by molecular replacement with Phenix[59] using apo $ClpP_{ec}$ (PDB ID: ITYF) as the search model. Model building and refinement were carried out using Coot[60] and Phenix[59] with TLS groups determined automatically using the TLSMD webserver[61]. Ramachandran statistics were calculated in Phenix using MolProbity, which gave 97.4 % total favored assignments and 0.21% outliers. To ensure that the stereochemistry of the clipibicyclene adduct remained well-restrained during refinement, complete amino acid restraints were generated for a modified serine residue using the GradeWebServer (http://grade.globalphasing.org). Data statistics are listed in supplementary table 10. The coordinates and structure factors have been deposited (PDB code 7MK5) in the Protein Data bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NY. Molecular graphics and analysis were performed using Pymol.

Declarations

Acknowledgments

We thank Professor Mike Tyers for the gift of pSASS vectors and *Saccharomyces cerevisiae* SASy31 and SAS735, Professor Walid Houry for the gift of *E. coli* BL21(DE3) 1146D and pET9a-EcClpP, Professor Heike Brötz-Oesterhelt for *B. subtilis* Δ*clpP,* and Professor Vladimir Larionov for *S. cerevisiae* VL648N. This research was funded by a Canadian Institutes of Health Research grant (FRN-148463), the Ontario Research Fund, and a Canada Research Chair to G.D.W. E.J.C. was supported by a CIHR Vanier Canada Graduate Scholarship. D.S. was supported by a CIHR Postdoctoral Fellowship. Protein structural studies were performed using beamline CMCF-BM at the Canadian Light Source at the University of Saskatchewan, which is supported by the Canada Foundation for Innovation, the Natural Sciences and Engineering Research Council, the National Research Council (NRC), the Canadian Institutes of Health Research, the Government of Saskatchewan, and the University of Saskatchewan.

Author Contributions

E.J.C. and G.D.W. conceived the study and designed experiments. A.C.P. and E.J.C. performed target-directed genome mining. G.P. and D.S. performed X-ray data collection, processing, and model building. D.S. carried out structural analysis and performed peptide mapping. C.H. performed structural elucidation

of azabicyclenes and clipibicyclene and purified clipibicyclene. E.J.C. performed all other experiments. E.J.C. and G.D.W. prepared the manuscript.

Competing Interests

The authors declare no competing interests.

Data availability statement

All data generated or analyzed during this study are included in this published article (and its supplementary information files). The *S. cattleya* DSM 46488 whole genome sequence is available in GenBank under accession number NC_017586. The structure of $ClpP_{ec}$ in complex with clipibicyclene is available in PDB with accession number 7MK5.

References

1. Culp, E. & Wright, G. D. Bacterial proteases, untapped antimicrobial drug targets. *J. Antibiot.* **70**, 366–377 (2017).
2. Sauer, R. T. & Baker, T. A. AAA+ Proteases: ATP-fueled machines of protein destruction. *Annu. Rev. Biochem.* **80**, 587–612 (2011).
3. Glynn, S. E. Multifunctional mitochondrial AAA proteases. *Frontiers in Molecular Biosciences* **4**, (2017).
4. Alexopoulos, J. A., Guarné, A. & Ortega, J. ClpP : A structurally dynamic protease regulated by AAA + proteins. *J. Struct. Biol.* **179**, 202–210 (2012).
5. Compton, C. L., Schmitz, K. R., Sauer, R. T. & Sello, J. K. Antibacterial activity of and resistance to small molecule inhibitors of the clpp peptidase. *ACS Chem. Biol.* **8**, 2669–2677 (2013).
6. Ju, Y. *et al.* Discovery of Novel peptidomimetic boronate ClpP inhibitors with noncanonical enzyme mechanism as potent virulence blockers in vitro and in vivo. *J. Med. Chem.* **63**, 3104–3119 (2020).
7. Hackl, M. W. *et al.* Phenyl esters are potent inhibitors of Caseinolytic Protease P and reveal a stereogenic switch for deoligomerization. *J. Am. Chem. Soc.* **137**, 8475–83 (2015).
8. Böttcher, T. & Sieber, S. A. Structurally refined β-lactones as potent inhibitors of devastating bacterial virulence factors. *ChemBioChem* **10**, 663–666 (2009).
9. Lakemeyer, M. *et al.* Tailored peptide phenyl esters block ClpXP proteolysis by an unusual breakdown into a heptamer–hexamer assembly. *Angew. Chemie Int. Ed.* **58**, 7127–7132 (2019).
10. Böttcher, T. & Sieber, S. A. β-Lactones as specific inhibitors of ClpP attenuate the production of extracellular virulence factors of *Staphylococcus aureus*. *J. Am. Chem. Soc.* **130**, 14400–14401 (2008).
11. Seo, J. H. *et al.* The mitochondrial unfoldase-peptidase complex ClpXP controls bioenergetics stress and metastasis. *PLoS Biol.* **14**, (2016).

12. Cole, A. *et al.* Inhibition of the mitochondrial protease ClpP as a therapeutic strategy for human acute myeloid leukemia. *Cancer Cell* **27**, 864–876 (2015).
13. Li, D. H. S. *et al.* Acyldepsipeptide antibiotics induce the formation of a structured axial channel in ClpP: A model for the ClpX/ClpA-bound state of ClpP. *Chem. Biol.* **17**, 959–969 (2010).
14. Lee, B. *et al.* Structures of ClpP in complex with acyldepsipeptide antibiotics reveal its activation mechanism. *Nat. Struct. Mol. Biol.* **17**, 471–478 (2010).
15. Graves, P. R. *et al.* Mitochondrial protease ClpP is a target for the anticancer compounds ONC201 and related analogues. *ACS Chem. Biol.* **14**, 1020–1029 (2019).
16. National Cancer Institute. Clinical Trials Using Akt/ERK Inhibitor ONC201. Available at: https://www.cancer.gov/about-cancer/treatment/clinical-trials/intervention/akterk-inhibitor-onc201. (Accessed: 20th April 2021)
17. Jacques, S. *et al.* Imipridone anticancer compounds ectopically activate the ClpP protease and represent a new scaffold for antibiotic development. *Genetics* **214**, 1103–1120 (2020).
18. Michel, K. H. & Kastner, R. E. A54556 antibiotics and process for production thereof. US 4492650 A (1985).
19. Thomy, D. *et al.* The ADEP biosynthetic gene cluster in *Streptomyces hawaiiensis* NRRL 15010 reveals an accessory ClpP gene as a novel antibiotic resistance factor. *Appl. Environ. Microbiol.* **85**, e01292-19 (2019).
20. Tang, X. *et al.* Identification of thiotetronic acid antibiotic biosynthetic pathways by target-directed genome mining. *ACS Chem. Biol.* **10**, 2841–2849 (2015).
21. Alanjary, M. *et al.* The Antibiotic Resistant Target Seeker (ARTS), an exploration engine for antibiotic cluster prioritization and novel drug target discovery. *Nucleic Acids Res.* **45**, W42–W48 (2017).
22. Blin, K. *et al.* antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **45**, W36–W41 (2017).
23. Baltz, R. H. Marcel Faber Roundtable: Is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration? *J. Ind. Microbiol. Biotechnol.* **33**, 507–513 (2006).
24. Navarro-Muñoz, J. C. *et al.* A computational framework to explore large-scale biosynthetic diversity. *Nat. Chem. Biol.* **16**, 60–68 (2020).
25. Schmidt, Y. *et al.* Biosynthetic origin of the antibiotic cyclocarbamate brabantamide A (SB-253514) in plant-associated *Pseudomonas*. *ChemBioChem* **15**, 259–266 (2014).
26. Johnston, C. W., Zvanych, R., Khyzha, N. & Magarvey, N. A. Nonribosomal assembly of natural lipocyclocarbamate lipoprotein-associated phospholipase inhibitors. *ChemBioChem* **14**, 431–435 (2013).
27. Liu, L. *et al.* Activation and characterization of bohemamine biosynthetic gene cluster from *Streptomyces* sp. CB02009. *Org. Lett.* **22**, 4614–4619 (2020).
28. Hong, Z. *et al.* Azetidine-containing alkaloids produced by a quorum-sensing regulated nonribosomal peptide synthetase pathway in Pseudomonas aeruginosa. *Angew. Chemie* **131**, 3210–3214 (2019).

29. Patteson, J. B., Lescallette, A. R. & Li, B. Discovery and biosynthesis of azabicyclene, a conserved nonribosomal peptide in *Pseudomonas aeruginosa*. *Org. Lett.* **21**, 4955–4959 (2019).
30. Bugni, T. S., Woolery, M., Kauffman, C. A., Jensen, P. R. & Fenical, W. Bohemamines from a marine-derived Streptomyces sp. *J. Nat. Prod.* **69**, 1626–1628 (2006).
31. Viala, J. & Mazodier, P. ClpP-dependent degradation of PopR allows tightly regulated expression of the *clpP3 clpP4* operon in *Streptomyces lividans*. *Mol Microbiol* **44**, 633–643 (2002).
32. Gominet, M., Seghezzi, N. & Mazodier, P. Acyl depsipeptide (ADEP) resistance in Streptomyces. *Microbiology* **157**, 2226–2234 (2011).
33. Bellier, A., Gominet, M. & Mazodier, P. Post-translational control of the Streptomyces lividans ClgR regulon by ClpP. *Microbiology* **152**, 1021–1027 (2006).
34. Bellier, A. & Mazodier, P. ClgR, a novel regulator of *clp* and *lon* expression in *Streptomyces*. *Society* **186**, 3238–3248 (2004).
35. Myronovskyi, M., Welle, E., Fedorenko, V. & Luzhetskyy, A. β-Glucuronidase as a sensitive and versatile reporter in Actinomycetes. *Appl. Environ. Microbiol.* **77**, 5370–5383 (2011).
36. Nagpal, J. *et al.* Molecular and structural insights into an asymmetric proteolytic complex (ClpP1P2) from *Mycobacterium smegmatis*. *Sci. Rep.* **9**, 1–12 (2019).
37. Ji, C. H., Kim, H. & Kang, H. S. Synthetic inducible regulatory systems optimized for the modulation of secondary metabolite production in *Streptomyces*. *ACS Synth. Biol.* **8**, 577–586 (2019).
38. Wang, J., Hartling, J. A. & Flanagan, J. M. The structure of ClpP at 2.3 Å resolution suggests a model for ATP- dependent proteolysis. *Cell* **91**, 447–456 (1997).
39. Bewley, M. C., Graziano, V., Griffin, K. & Flanagan, J. M. The asymmetry in the mature amino-terminus of ClpP facilitates a local symmetry match in ClpAP and ClpXP complexes. *J. Struct. Biol.* **153**, 113–128 (2006).
40. Davies, J., Spiegelman, G. B. & Yim, G. The world of subinhibitory antibiotic concentrations. *Curr. Opin. Microbiol.* **9**, 445–453 (2006).
41. Yim, G., Spiegelman, G. B. & Davies, J. E. Separate mechanisms are involved in rifampicin upmodulated and downmodulated gene expression in *Salmonella typhimurium*. *Res. Microbiol.* **164**, 416–424 (2013).
42. Vázquez-Laslop, N. & Mankin, A. S. How macrolide antibiotics work. *Trends in Biochemical Sciences* **43**, 668–684 (2018).
43. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. *Practical Streptomyces Genetics*. (John Innes Foundation, 2000).
44. Xu, M. *et al.* GPAHex-A synthetic biology platform for Type IV–V glycopeptide antibiotic production and discovery. *Nat. Commun.* **11**, 1–12 (2020).
45. Yamanaka, K. *et al.* Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proc. Natl. Acad. Sci.* **111**, 1957–1962 (2014).

46. Luo, Y., Zhang, L., Barton, K. W. & Zhao, H. Systematic identification of a panel of strong constitutive promoters from *Streptomyces albus*. (2015). doi:10.1021/acssynbio.5b00016
47. Gietz, R. D. & Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2**, 31–34 (2007).
48. Ji, C.-H., Kim, J.-P. & Kang, H.-S. Library of synthetic *Streptomyces* regulatory sequences for use in promoter engineering of natural product biosynthetic gene clusters. *ACS Synth. Biol.* **7**, 1946–1955 (2018).
49. Gust, B., Kieser, T. & Chater, K. PCR targeting system in *Streptomyces coelicolor* A3(2). (2002).
50. Wang, W. *et al.* An engineered strong promoter for streptomycetes. *Appl. Environ. Microbiol.* **79**, 4484–4492 (2013).
51. Fey, P. D. *et al.* A genetic resource for rapid and comprehensive phenotype screening of nonessential Staphylococcus aureus genes. *MBio* **4**, (2013).
52. Ahsan, B. Understanding the activation of bacterial protease ClpP by acyldepsipeptide antibiotics. (McMaster University, 2014).
53. Marty, M. T. *et al.* Bayesian deconvolution of mass and ion mobility spectra: From binary interactions to polydisperse ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).
54. Cobb, R. E., Wang, Y. & Zhao, H. High-efficiency multiplex genome editing of *Streptomyces* species using an engineered CRISPR/Cas system. *ACS Synth. Biol.* **4**, 723–728 (2015).
55. Culp, E. J. *et al.* Hidden antibiotics in actinomycetes can be identified by inactivation of gene clusters for common antibiotics. *Nat. Biotechnol.* **37**, 1149–1154 (2019).
56. Vonrhein, C. *et al.* Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 293–302 (2011).
57. Kabsch, W. *XDS*. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132 (2010).
58. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography* **67**, 235–242 (2011).
59. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 352–367 (2012).
60. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
61. Painter, J. & Merritt, E. A. *TLSMD* web server for the generation of multi-group TLS models. *J. Appl. Crystallogr.* **39**, 109–111 (2006).

Figures

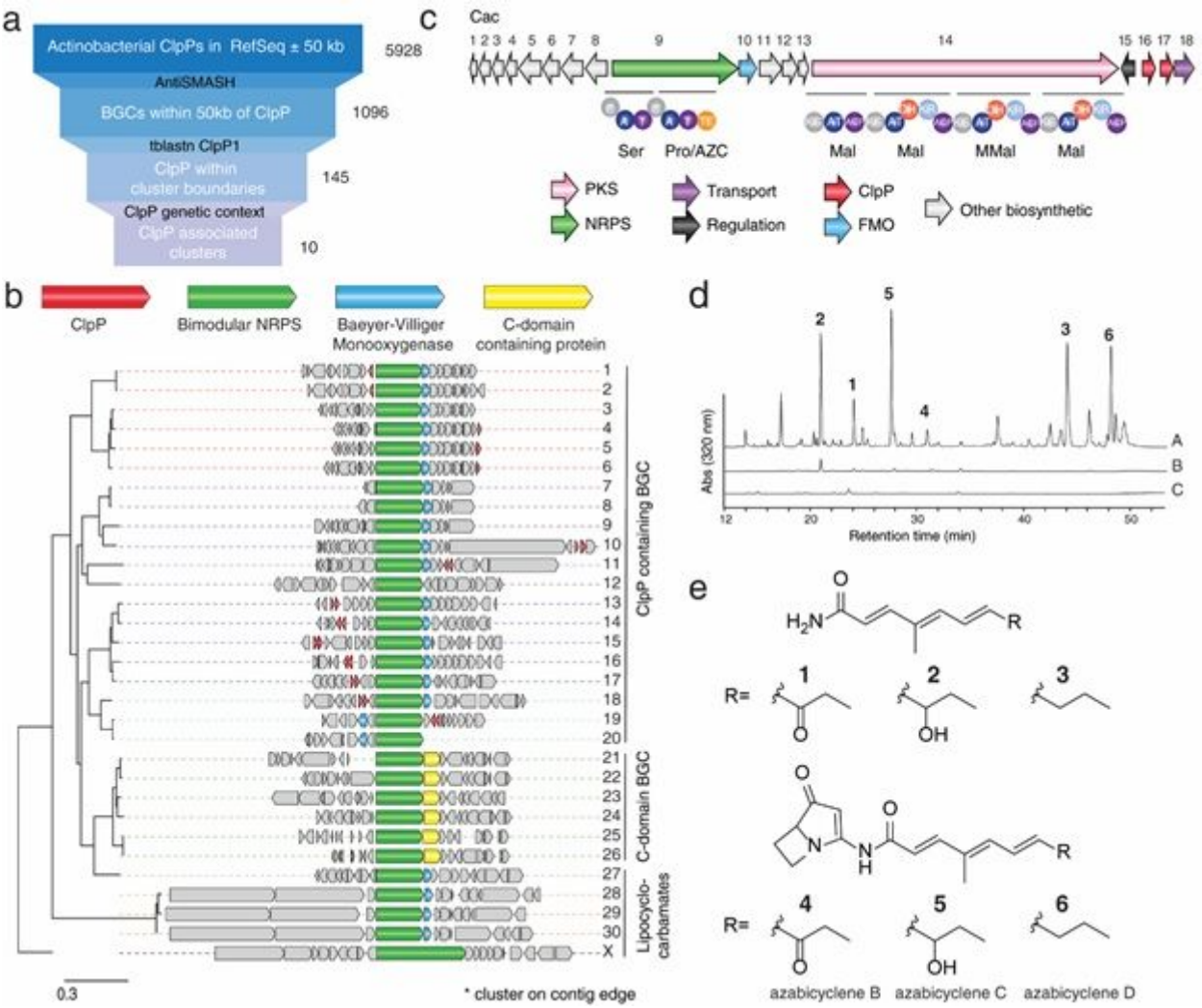

Figure 1

Genome mining identifies a family of ClpP-associated BGCs (a) Summary of target-directed genome mining approach used to identify ClpP associated clusters from genomes available in GenBank. Numbers to the right of the workflow indicate the number of hits at each step. (b) Cladogram and cluster alignment built using CORASON24 on 30 select clusters containing a bimodular NRPS. Clusters are numbered as in Supplementary Table 3, and an unrelated BGC identified in Amycolatopsis sp. CA-126428 marked X is used to root the tree. (c) Schematic of cac identified in S. cattleya DSM 46488 and putative gene functions. (d) HPLC chromatogram of n-butanolic extracts from S. coelicolor M1154 strains carrying either pCGW-cac-LHK (A), pCGW-cac pIJ10257-cac15 (B) or pCGW (C). Fermentations were performed in triplicate, and a representative chromatogram is shown. Numbered peaks were identified as azabicyclenes family compounds or hydrolyzed acyl tails, as shown in (e).

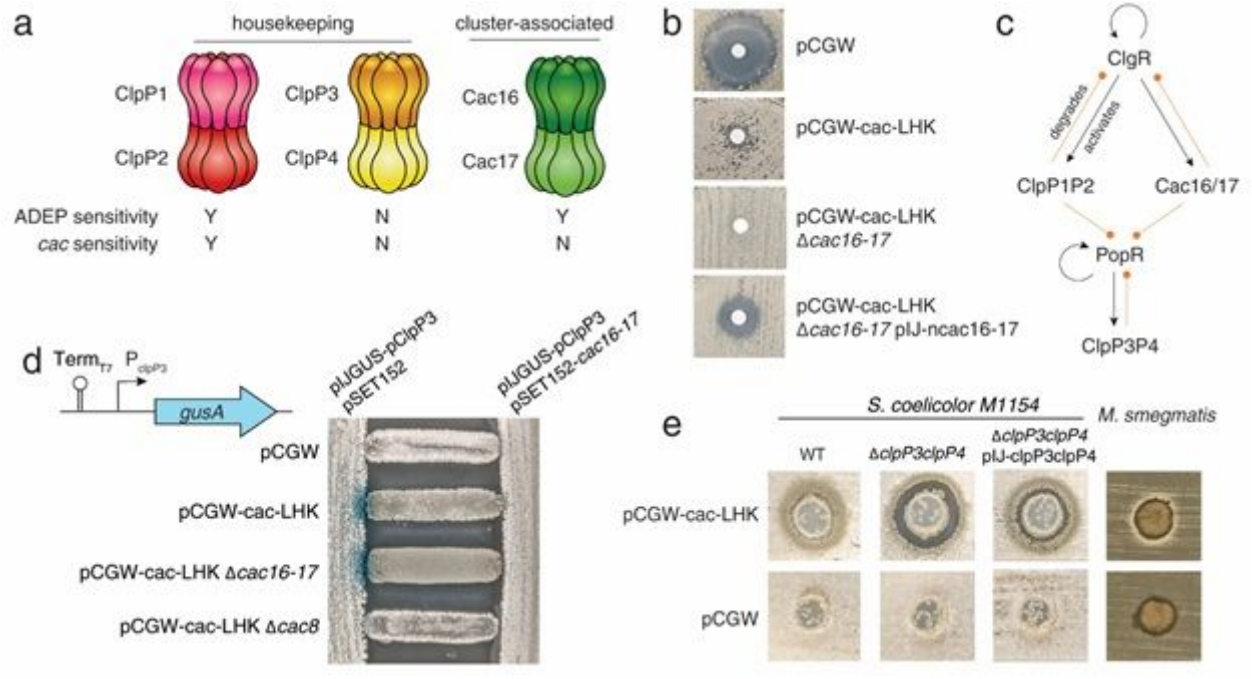

Figure 2

Biological activity of an elusive ClpP inhibitor (a) Summary of ClpP isoforms present in S. coelicolor M1154 pCGW-cac-LHK and their susceptibility to ADEP activation or cac metabolite inhibition. (b) Kirby-Bauer assays with ADEP show the sensitivity of S. coelicolor M1154 strains harboring the constructs indicated. Complementation of cac16-17 is provided by the plasmid pIJ-ncac16-17. Each disk contains 50 µg ADEP. (c) Schematic of ClpP regulation by the ClgR regulon in S. coelicolor. ClgR is a transcriptional activator that regulates the clpP1clpP2 operon as well as its own expression33,34. Subsequently, ClpP1/P2/X degrades ClgR and the transcriptional activator for the clpP3clpP4 operon, PopR. A similar negative feedback loop occurs between ClpP3P4 and PopR31. The promoter region of cac16-17 contains a ClgR binding motif34 and is also likely to be regulated by ClgR. (d) Co-streaking of S. coelicolor M1154 strains harboring constructs for heterologous production (streaked horizontally) next to indicator strains (streaked vertically) reveals PclpP3 activation. The schematic shows the design of the plasmid pIJGUS-pClpP3. Indicator strains also contain empty vector (pSET152) or that for expressing cac16-17 (pSET152-cac16-17). Strains are streaked on media containing 200 µg/mL X-gluc. (e) Kirby-Bauer assays for growth inhibition make use of agar plugs inoculated with S. coelicolor strains as indicated on the left. Strains that are being tested for susceptibility are labeled along the top. Complementation of clpP3clpP4 expression is provided by the plasmid pIJ-clpP3clpP4. Experiments in panels (b), (d), and (e) were performed on at least two independent occasions with similar results.

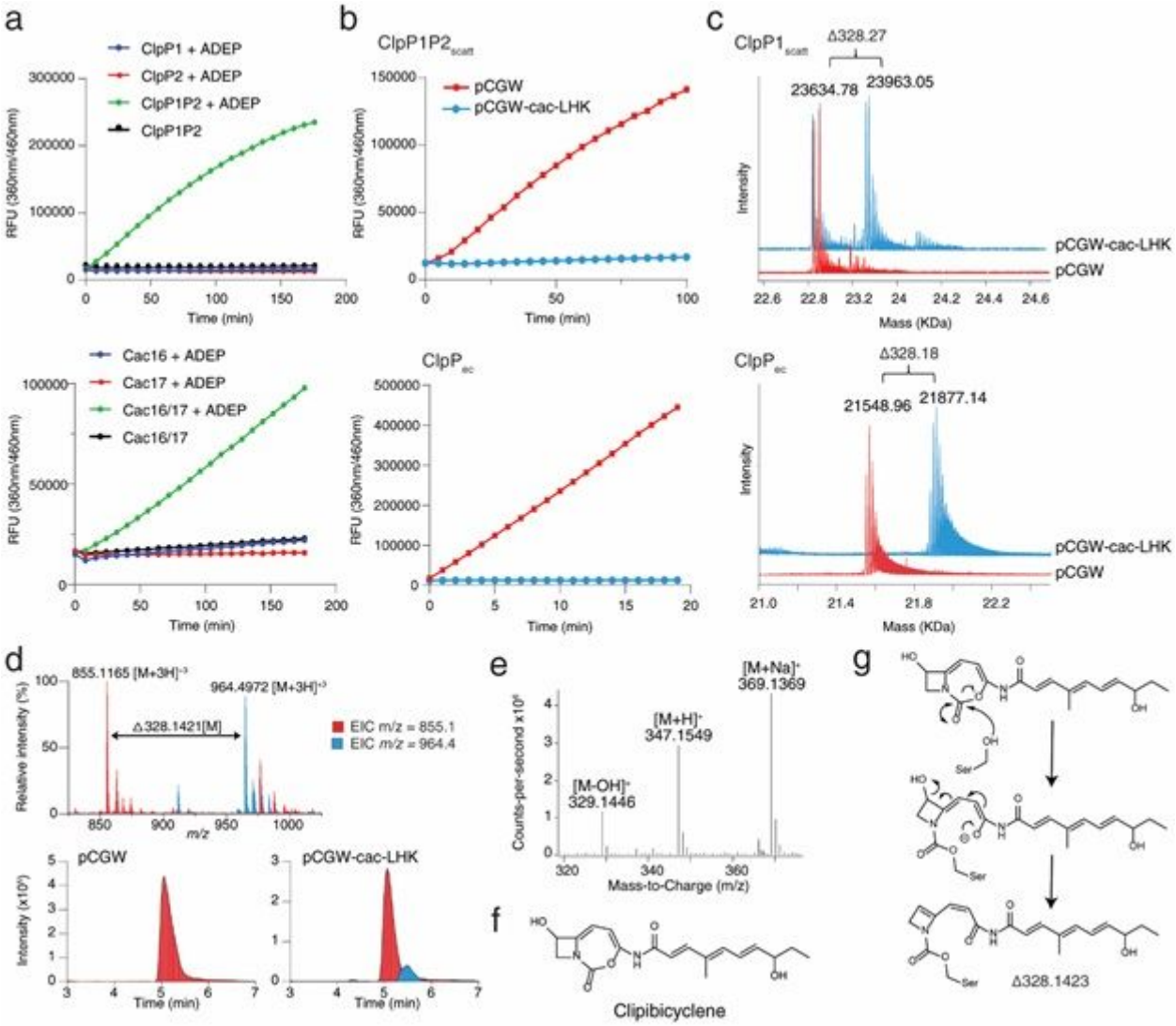

Figure 3

In vitro activity and inhibition of ClpP by clipibicyclene (a) ClpP1P2scatt (top) and Cac16/17 (bottom) form active heterocomplexes in vitro when combined with ADEP. Assays are performed using the fluorogenic peptide substrate S-LLVY-AMC and 2 μM of each subunit. (b-d) After incubation of ClpP1P2scatt (top) or ClpPec (bottom) in spent media from S. coelicolor M1154 strains, as indicated, enzymes were analyzed for peptidase activity (b), by intact protein LC-MS (c), or trypsin digested for peptide analysis (d). (d) The top MS spectrum shows the ClpP1 peptide containing the active site Ser99 in its unmodified (red) and clipibicyclene-modified (blue) form, corresponding to a mass increase of 328.14 Da. Bottom panels show extracted ion chromatograms of the corresponding peptides from ClpP1 incubated in spent media, as indicated. (e) LC-MS spectra from S. coelicolor M1154 pCGW-cac-LHK shows a metabolite with [M+H]+ = 347.1549 and corresponding dehydrated or sodiated adducts that we identify as (f) clipibicyclene (calculated [M+H]+ = 347.1607, [M-OH]+ = 329.1501, [M+Na]+ = 369.1426). (g) Proposed reaction of clipibicyclene with ClpP's active site Ser resulting in a mass shift of the intact enzyme of 328.1423 Da. Experiments in panels (a)-(d) were performed on at least two independent

occasions with similar results. For panels (a) and (b), the mean of technical triplicate reactions with error bars representing standard deviation (not visible at this scale) is shown.

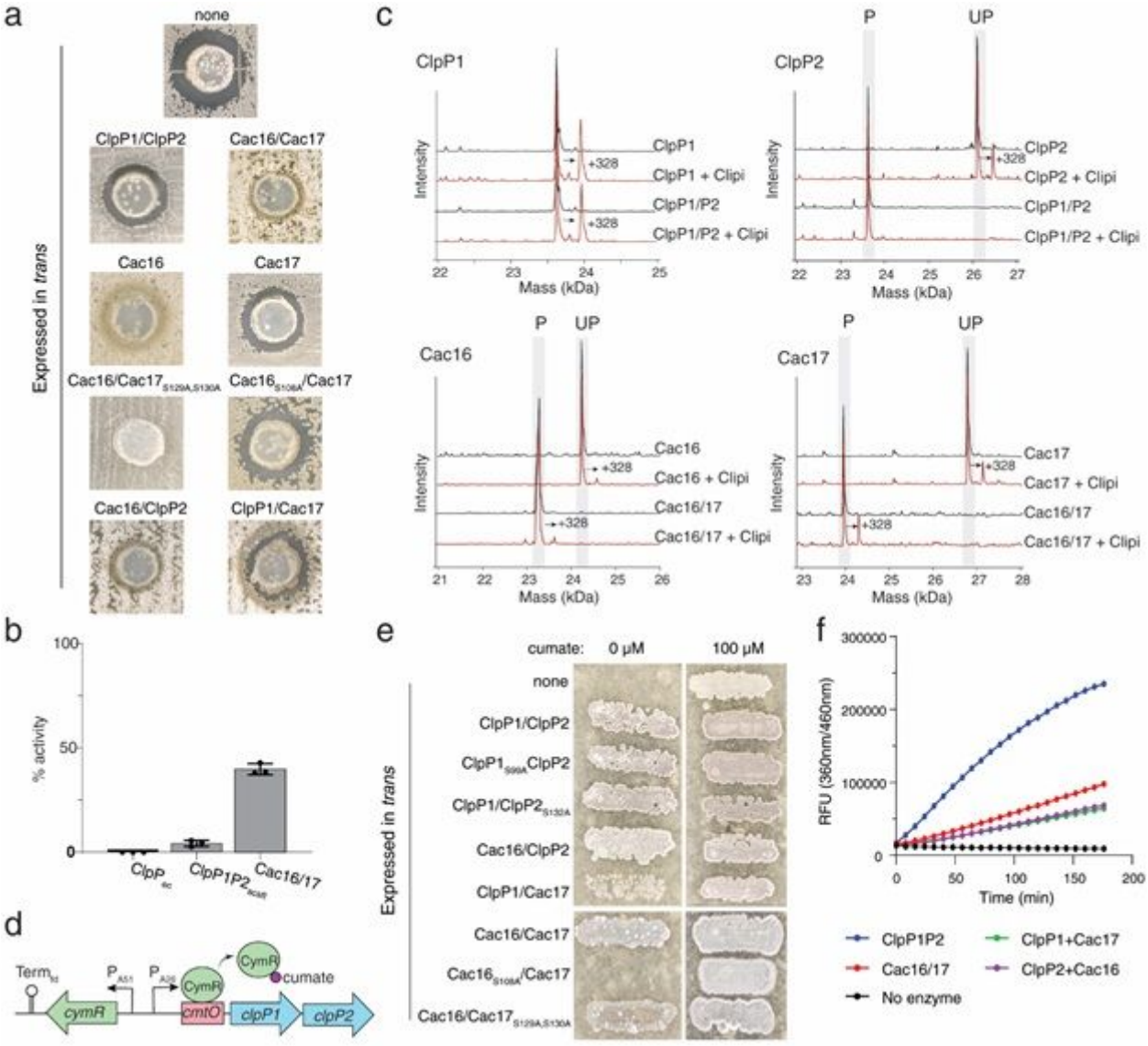

Figure 4

Cac16/17 provide resistance to clipibicyclene (a) Kirby-Bauer assays using agar plugs inoculated with S. coelicolor M1154 pCGW-cac-LHK showing growth inhibition of S. coelicolor ΔclpP3clpP4 strains. Alternate ClpPs are expressed in trans, as indicated, from pIJ10257 constructs. (b) In vitro peptidase assays show percent peptidase activity compared to DMSO control of ClpPs treated with 50 µM clipibicyclene. (c) Intact protein LC-MS of individual ClpP subunits shows susceptibility to modification by 100 µM clipibicyclene (clipi). “P” indicates the unmodified, processed form of the subunit, and “UP” indicates the unmodified, unprocessed form of the subunit. Arrows indicate a mass shift of ~328 Da by clipibicyclene modification. (d) Schematic of inducible regulation of the clpP1clpP2 operon at its native locus, PA26-cmtO:clpP2clpP2. The operator sequence cmtO is bound by the repressor CymR, providing strict repression in the absence of cumate37. PA26 and PA51 are strong, constitutive synthetic promoters48. (e) Growth of S. coelicolor ΔclpP3clpP4 PA26-cmtO:clpP2clpP2 in the presence/absence of

cumate. The ability of various ClpPs to support growth is tested by expression in trans, as indicated, from pIJ10257 constructs. Strains were streaked on SM+MgCl2 agar. (f) Mixed complexes of Cac16/ClpP2 and ClpP1/Cac17 form active heterocomplexes in vitro. Assays were performed with ADEP using the fluorogenic peptide substrate S-LLVY-AMC and 2 μM of each subunit. For panels (b) and (f), the mean of technical triplicate reactions with error bars representing standard deviation (not visible in panel f at this scale) is shown. All experiments were performed on at least two independent occasions with similar results.

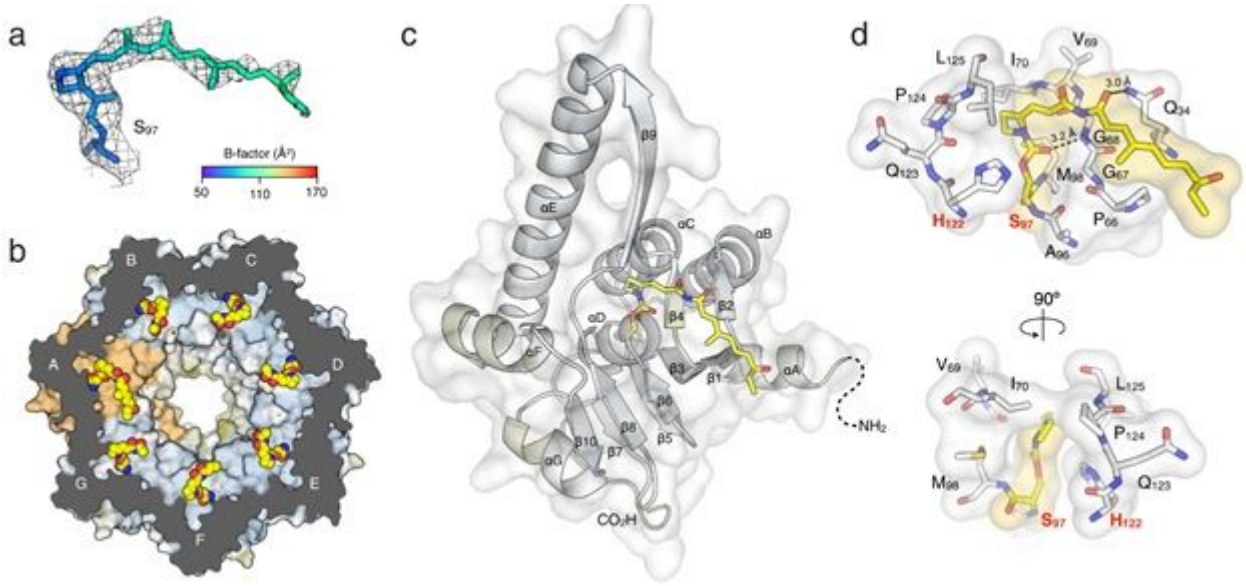

Figure 5

Structure of the ClpPec:clipibicyclene complex (a) Stick model of S97 covalently bound to clipibicyclene colored by individual atomic B-factors. The color bar depicts the B-factor scale. The 2mFo-DFc map is shown in black mesh and contoured at 0.9 σ. (b) Interior cavity of the ClpPec tetradecamer is shown as a surface cutaway with the clipibicyclene adducts depicted in sphere representation. The chains are labeled alphabetically (A-G), and the representative chain used in panels (c) and (d) is highlighted in orange. (c) Structure of the ClpPec:clipibicyclene monomer shown in surface/cartoon representations with the adduct depicted by yellow sticks. (d) Detailed interactions with clipibicyclene and the residues of active site cleft in surface/stick representation. Amino acids and clipibicyclene are colored white and yellow, respectively. H-bonds are shown as black dashed lines.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryinfosubmission1.docx