



Published in final edited form as:

J Am Chem Soc. 2012 August 15; 134(32): 13192–13195. doi:10.1021/ja3052156.

Malleilactone, a polyketide synthase-derived virulence factor encoded by the cryptic secondary metabolome of *Burkholderia pseudomallei* group pathogens

John B. Biggins, Melinda A. Ternei, and Sean F. Brady*

Laboratory of Genetically Encoded Small Molecules, The Rockefeller University and Howard Hughes Medical Institute, 1230 York Ave, New York, NY 10068

Abstract

Sequenced bacterial genomes are routinely found to contain gene clusters that are predicted to encode metabolites not seen in fermentation based studies. Pseudomallei group *Burkholderia* are emerging pathogens whose genomes are particularly rich in cryptic natural product biosynthetic gene clusters. We systemically probed the influence of the cryptic secondary metabolome on the virulence of these bacteria and found that the disruption of the MAL gene cluster, which is natively silent in laboratory fermentation experiments and conserved across this group of pathogens, attenuates virulence in animal models. Using a promoter exchange strategy to activate the MAL cluster we identified malleilactone, a polyketide synthase-derived cytotoxic siderophore encoded by this gene cluster. Small molecules targeting malleilactone biosynthesis either alone, or in conjunction with antibiotics, could prove useful as next-generation therapeutics for combating melioidosis and glanders.

Pseudomallei group pathogens including *Burkholderia pseudomallei* (*BP*), *Burkholderia mallei* (*BM*) and *Burkholderia thailandensis* (*BT*) are a closely related collection of Gram negative bacteria¹. *BP* is the causative agent of melioidosis in humans, *BM* is the causative agent of glanders in horses, and while *BT* is not generally considered to be a human pathogen, it is infectious in a number of model laboratory organisms. Melioidosis is endemic in parts of Southeast Asia and Northern Australia and is the third most frequent cause of mortality from infectious disease after HIV and tuberculosis in Northeast Thailand.² Current therapies for this neglected disease are inadequate as mortality rates from melioidosis have been reported to approach 50% even with antibiotic treatment.² Additionally, these pathogens have been classified as potential bio-terrorism threats by the CDC, and in spite of the considerable attention paid to this group, their virulence determinants are still not well understood. Small molecules, including signaling molecules, siderophores, and toxins, are known to play important roles in both the establishment and propagation of bacterial infections.³ Biosynthetic gene clusters to which no small molecule has yet been assigned are frequently encountered in bacterial genome sequencing projects.⁴ These cryptic gene clusters represent the pool of biosynthetic pathways from which additional small molecule virulence factors might be characterized. In an attempt to better understand the virulence factors used by pseudomallei group pathogens, we have investigated the influence of the cryptic secondary metabolism on *Burkholderia* pathogenicity by systematically disrupting individual cryptic biosynthetic pathways in *BT* and assessing these mutants for changes in virulence. Disruption of the MAL gene cluster, a

*Corresponding Author sbrady@rockefeller.edu.

Supporting Information. All experimental protocols, supplementary data and NMR spectra. This material is available free of charge via the Internet at <http://pubs.acs.org>.

polyketide synthase based-cluster that is conserved across this group of pathogens, was found to attenuate virulence. Here we describe the structural and functional characterization of malleilactone, a MAL gene cluster encoded cytotoxic siderophore and virulence factor.

The pangenome of the pseudomallei group of pathogens⁵ is rich in polyketide synthase (PKS) and nonribosomal peptide synthetase (NRPS) gene clusters, two common biosynthetic systems that are easily identified bioinformatically in sequenced genomes. Many clusters are shared among these three pathogens and the majority of molecules encoded by these clusters remain cryptic in structure and/or biological function (Fig. 1a and Tables S1 and S2). *BM* and *BP* are classified as BSL3 level pathogens, and due to their potential use in biological warfare, there are significant regulatory issues governing their genetic manipulation. *BT* is strikingly similar genetically to *BP*,^{5b,c} possessing the majority of gene clusters found within *BP* (Table S1); however there are no similar restrictions on its genetic manipulation.

To address the potential role of secondary metabolites in Burkholderia pathogenesis, we created a collection of *BT* mutants wherein individual cryptic PKS/NRPS clusters were disrupted either by suppressing transcription through promoter exchange or by deleting one of the large modular PKS/NRPS genes within the gene cluster itself (Table S1). These strains were then assessed for virulence using a *Caenorhabditis elegans* nematode co-culture model, which has been successfully used to assess pathogenicity and identify virulence factors in bacterial pathogens.⁶ In this co-culture virulence model, both *BT* and *BP* elicit a rapid toxicity to *C. elegans*, which is thought to arise from specific *in vivo* host-bacteria interactions.⁷ In co-culture assays using wild-type *BT* worms display either no movement or lack any significant loco-motor behavior within 24 hours (Fig. 1b and c and Video S1). With one exception, *C. elegans* grown on cryptic gene cluster disruption strains were similarly dead or dying within 24 hours. The disruption, either by promoter exchange or gene deletion, of a cryptic gene cluster that we have termed the MAL cluster results in worms that uniformly continue to move and forage beyond 24 hours (Fig. 1b and c). Worms feeding upon these MAL knockout strains are larger than worms fed wild type *BT*, retain their active feeding and foraging behavior, and grow into adulthood as evidenced by the accumulation of eggs and hatched larvae on assay plates (Fig. 1b and c and Video S1).

To assess the generality of the importance of the MAL gene cluster to pseudomallei group virulence, we also tested MAL disruption mutants in an amoeba co-culture model.⁸ The response of the social amoeba *Dictyostelium discoideum* to pathogenic bacteria is hypothesized to closely mimic the response of the mammalian immune system to these organisms, as *D. discoideum* similarly feeds upon bacteria through phagocytosingestion.^{8,9} In this model, *D. discoideum* cells foraging on pathogenic bacteria often fail to develop past their unicellular state into multicellular aggregates and fruiting bodies. In many instances, this inhibition has been observed to be dependent on the same set of virulence factors that are required for mammalian infections.^{8,9} When co-cultured with wild-type *BT* amoebae die as unicellular organisms, preceding any observable social coordination. In identical experiments using MAL gene cluster knockout strains amoebae aggregated as expected and developed completely through the formation of terminal fruiting bodies (Fig. 1d). As observed in the *C. elegans* model, the MAL gene cluster appears to be a key component of *BT*'s virulence toward *D. discoideum*.

The MAL gene cluster spans ~35 kb and is predicted to contain 13 open reading frames, two of which (*malA* and *malF*) encode large modular PKSs (Fig. 1e). It is one of only three PKS/NRPS gene clusters that is conserved across the pseudomallei group, making it a conspicuous candidate for encoding a communal virulence factor. MAL gene clusters from all three pathogens have the same gene organization and content (Fig. 1e), with the

exception of a hypothetical gene (*malB_b*) found in *BM* and *BP* but not in *BT*. Bioinformatics comparisons to previously characterized gene clusters revealed no close relatives that could provide insight into the structure or function of the MAL-encoded metabolite.

Culture broth extracts from MAL gene cluster knockout strains of *BT* did not show any obvious differences when compared to wild type control extracts (Fig. 2a) indicating that in the laboratory setting this gene cluster is silent. Many secondary metabolite gene clusters are thought to have remained cryptic even after extensive investigation of the organisms in which they reside, because these clusters are only expressed in response to specific environmental cues and therefore silent under standard laboratory fermentation conditions. In previous studies with *BT*, we found that silent gene clusters could be activated via transcriptional activation strategies.¹⁰ The most universal of such approaches is likely to be promoter exchange, where a native silent promoter governing gene cluster expression is replaced with a model inducible promoter. The MAL gene cluster is predicted to be composed of a set of thirteen genes that are unidirectional and potentially organized into a single large operon (Fig. 1e). We reasoned that inducing the transcription of this operon could lead to the production of detectable levels of the product encoded by the MAL gene cluster. The promoter region in front of the first ORF in the MAL operon (*malA*) was therefore replaced by homologous recombination with the rhamnose inducible promoter, P_{RhaB},¹¹ to yield *BT*:P_{RhaB}-MAL (Figs. 2b and Fig. S1). Ethyl acetate extracts from cultures of *BT*:P_{RhaB}-MAL induced with rhamnose contained one major metabolite that was not seen in extracts from similarly treated wild type cultures (**1**, Fig. 2c). This peak is not seen in extracts from wild-type *BT* cultures, nor is it seen in extracts from rhamnose induced cultures of *BT*:P_{RhaB}-MAL(Δ *malF*), a *BT*:P_{RhaB}-MAL strain where the PKS2 gene (*malF*) is disrupted, thus directly linking MAL operon induction to the production of **1**.

Compound **1**, which we have named malleilactone, was purified from ethyl acetate extracts of large-scale (6L) cultures of rhamnose-induced *BT*:P_{RhaB}-MAL using a modified Kupchan scheme followed by three rounds of silica gel chromatography. The structure of **1** was determined through a combination of HRMS (HRMS-TOF *m/z*: [M+H]⁺ calcd for C₁₈H₂₇O₄, 307.1910; found 307.1903), NMR and UV data (Fig. 3a). All 18 of the carbons predicted by HRMS to be present in **1** are seen in the ¹³C spectra. Empirical chemical shift data and ¹H-¹³C HMQC experiments indicated the presence of 2 carbonyl/enol (194, 189 ppm), 5 olefin (140, 139, 137, 123, 100 ppm), 1 ester (174 ppm), 7 methylene and 3 methyl carbons (Fig. 3a). Three spin systems could be resolved in the ¹H-¹H COSY spectrum, two of which are joined by HMBC correlations into an unbroken saturated 7-carbon chain that is connected to the methyl substituted C-8/C-9 olefin through HMBC correlations. The *trans* geometry of this double bond is supported by the chemical shift of the C-8a tertiary methyl (13.9 ppm).¹² HMBC correlations from the C-9 methine proton and the C-8a methyl protons to C-7 connect this substructure to the predicted C-7 enol. At the other end of the molecule, the final 2-carbon ¹H-¹H COSY spin system can be connected by HMBC correlations to the C-3 carbonyl and C-4, and HMBC correlations from H-5 to C-3 and C-4 define the position of the C-4/C-5 olefin. These two substructures are linked through C-6 based on the large collection of HMBC correlations involving H-5 that is shown in Fig. 3a. A final three-bond correlation from H-5 to the C-6a carbonyl yields the complete 18 carbon skeleton of **1**. Based on the unsaturation index and empirical chemical shift data, the final unsaturation is satisfied by closing the γ -butyrolactone ring at C-4 to give **1**. The presence of the conjugated unsaturated system seen in the final structure of **1** is supported by the UV_{max} observed at 373 nm.¹³

A detailed examination of **1** suggested that it likely arise from the condensation of two separate polyketide chains, one running from the C-16 methyl through the C-6a carbonyl and a second running from the C-1 methyl through the C-5 olefin (Fig. 3b). The formation

of two such polyketide precursors and their subsequent condensation into **1** can be easily rationalized based on the collection of domains found in the MAL PKSs, MalA and MalF (Fig. 3b, c). In our biosynthetic proposal, one polyketide precursor is produced on MalA from propionic acid and a hydroxylmalonyl extender unit. The use of the rare oxidized extender unit, which is necessary for the subsequent lactone formation, is supported by the presence of an FkbH homolog (MalH) in the MAL cluster (Fig S5). FkbH proteins are predicted to load a glycolytic pathway intermediate onto a thiolation domain (T), which then serves as a substrate for the biosynthesis of oxidized extender units (Fig S5).¹⁴ The second of the two required polyketide precursors is predicted to be synthesized on MalF from a caprylic acid starter unit, followed by two rounds of elongation: one involving incorporation, reduction and dehydration of a methylmalonyl extender unit and the second involving incorporation of a non-reduced malonyl extender. While MalA contains an acyltransferase (AT) domain, MalF is predicted to be a *trans* AT system that use the AT activity of MalL. The formation of **1** from these two PKS precursors requires that they be linked through an ester bond and an additional C-C bond to form the central five membered lactone. There are now a handful of NRPS condensation (C) domains that are predicted to form esters instead of amides, and MalF contains the conserved “HHXXXDD” active-site motif shared among this set of domains.¹⁵ We therefore propose that the diketide from MalA is transferred to the final T-domain in MalF and linked via a MalF C-domain catalyzed ester bond. MalA contains two thioesterase (TE) domains, both of which are predicted to be type II, or proof reading, TE domains and therefore not responsible for releasing the polyketide. Release of the polyketide from MalF is predicted to instead occur through the action of the terminal reductase (R) domain, which have increasingly been shown to catalyze the terminal release and cyclization of polyketides.¹⁶ Reductive cleavage of the condensed intermediate found on the T₄ domain of MalF, followed by intramolecular cyclization, provides an intermediate that at upon dehydration would yield **1**. The best of our knowledge **1** has not been previously reported as a natural or synthetic compound.

In all three pseudomallei group pathogens, the MAL cluster resides directly adjacent to a predicted LuxR-type transcription factor, *malR*(*btaR4*¹⁷). LuxR homologs function as receptors for *N*-acylhomoserine lactones (AHLs) in quorum sensing circuits and are often responsible for regulating virulence gene expression.^{3a} As seen in experiments with the MAL PKS deletion strain *BT*(Δ *malF*), *C. elegans*, and *D. discoideum* co-cultured with *malR*-deletion strains show extended lifespans (Fig 1b and Supporting data), suggesting that MAL expression is mediated through MalR. A consensus LuxR-like binding element (*lux* box) is present in the promoter region immediately upstream of *malA* (Fig. S5). In rare cases, LuxRs are ligand independent;¹⁸ however, in cases where a native ligand has been defined, they are controlled by AHLs. Thus, malleilactone production, like the production of many virulence factors, appears to be governed by AHL-dependent quorum sensing.

Malleilactone was assayed for toxicity against a panel of cell types. It shows low micromolar cytotoxicity to human cell lines (e.g. IC₅₀ 19 μ M) and in disk diffusion assays it inhibits the growth of Gram-positive bacteria at as low as 2.5 μ g per disk (Table S4). When incorporated in growth media (>100 μ g-mL⁻¹), malleilactone had no obvious deleterious effects upon *C. elegans*/*E. coli* co-cultures, nor did worms co-cultured with MAL mutants show increased distress upon malleilactone reintroduction, suggesting that its bioactivity might extend beyond localized cytotoxicity. The central core of malleilactone closely resembles a tetronic (or tetramic) acid substructure, and while reported bio-activities for tetronic acids vary widely, they are predicted to share the common phenomenon of chelating cations, in particular iron.¹⁹ Using a chromogenic iron chelation assay, we observed iron-binding activity with malleilactone (Fig. S6). Malleilactone may serve therefore both as an iron-chelator to assist with iron acquisition as well as a toxin when secreted at the site of infection. It is also possible that *in vivo* malleilactone serves an additional undetermined role

beyond either of these functions, as so far the MAL cluster has proven recalcitrant to disruption in *BP* and *BM*.²⁰ Given that malleilactone contains a γ -butyrolactone core, a structural motif found in a variety of bacterial signaling molecules including AHLs, it is also plausible that malleilactone is a signaling molecule capable of activating downstream virulence factors.

As iron is tightly sequestered in eukaryotes, bacterial pathogens have in many cases evolved ironchelation systems in order to survive in these low iron environments. Two additional PKS/NRPS gene clusters predicted to encode siderophores (malleobactin/ornibactin and pyochelin) are shared between *BP* and *BT* genomes (*BM* lacks pyochelin). It has been theorized that the production of multiple siderophores with different structures and polarities serves to benefit pathogens by providing alternate routes for securing iron and possibly circumventing host defense mechanisms that neutralize conserved microbial siderophores.²¹ The therapeutic importance of this class of secondary metabolites as potential anti-infective targets is highlighted by the use of the siderophore (mycobactin) biosynthesis inhibitor, *p*-aminosalicylate, as an anti-tuberculosis drug.²² Small molecules targeting malleilactone biosynthesis could similarly prove useful as next-generation therapeutics for combating both melioidosis and glanders.

The systematic gene cluster knockout/promoter exchange strategy we used to identify malleilactone is easily generalizable, and should permit the functional characterization of small molecule virulence factors encoded by silent gene clusters found in the genomes of a diverse collection of additional bacterial pathogens.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

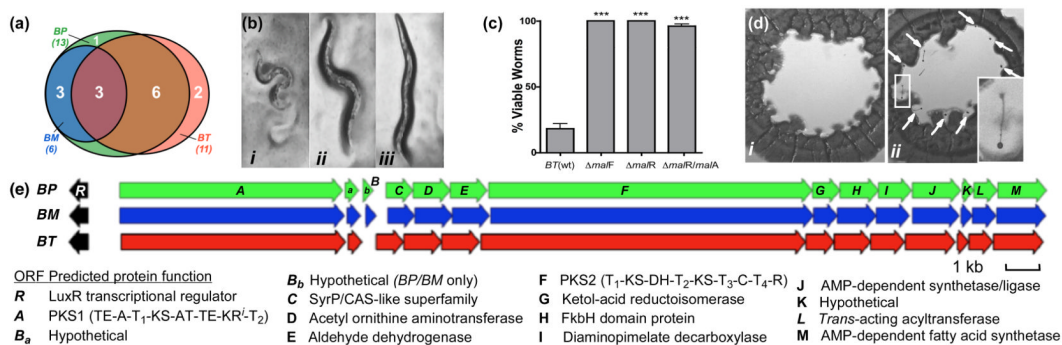
Acknowledgments

This work was supported by the Northeast Biodefense Center (U54-A1057158) and NIH (GM077516). SFB is an HHMI early career scientist. Vector pSCrhaB2 was a gift from Katie Nurse and Dr. Miguel Valvano (University of Western Ontario, Canada). *Dictyostelium* was obtained from Laura Macro and Prof. Sanford Simon (The Rockefeller University).

References

1. Galyov EE, Brett PJ, DeShazer D. *Annu Rev Microbiol.* 2010; 64:495–517. [PubMed: 20528691]
2. Limmathurotsakul D, Wongratanacheewin S, Teerawattansook N, Wongsuvan G, Chaisuksant S, Chetchotisakd P, Chaowagul W, Day NP, Peacock SJ. *J Am J Trop Med Hyg.* 2010; 82:1113–7.
3. (a) Bassler BL, Losick R. *Cell.* 2006; 125:237–46. [PubMed: 16630813] (b) Gomez JE, Clatworthy A, Hung DT. *Crit Rev Biochem Mol Biol.* 2011; 46:41–66. [PubMed: 21250782]
4. Zerikly M, Challis GL. *ChemBiochem.* 2009; 10:625–33. [PubMed: 19165837]
5. (a) Nierman WC, et al. *Proc Natl Acad Sci U S A.* 2004; 101:14246–51. [PubMed: 15377793] (b) Kim HS, Schell MA, Yu Y, Ulrich RL, Sarria SH, Nierman WC, DeShazer D. *BMC Genomics.* 2005; 6:174. [PubMed: 16336651] (c) Yu Y, et al. *BMC Microbiol.* 2006; 6:46. [PubMed: 16725056]
6. Sifri CD, Begun J, Ausubel FM. *Trends Microbiol.* 2005; 13:119–27. [PubMed: 15737730]
7. (a) O'Quinn AL, Wiegand EM, Jeddeloh JA. *Cell Microbiol.* 2001; 3:381–93. [PubMed: 11422081] (b) Lee SH, Ooi SK, Mahadi NM, Tan MW, Nathan S. *PLoS One.* 2011; 6:e16707. [PubMed: 21408228]
8. Hasselbring BM, Patel MK, Schell MA. *Infect Immun.* 2011; 79:2079–88. [PubMed: 21402765]
9. (a) Steinert M, Heuner K. *Cell Microbiol.* 2005; 7:307–14. [PubMed: 15679834] (b) Boz-zaro S, Eichinger L. *Curr Drug Targets.* 2011; 12:942–54. [PubMed: 21366522]

10. (a) Biggins JB, Liu X, Feng Z, Brady SF. *J Am Chem Soc.* 2011; 133:1638–41. [PubMed: 21247113] (b) Biggins JB, Gleber CD, Brady SF. *Org Lett.* 2011; 13:1536–9. [PubMed: 21348454]
11. Cardona ST, Valvano MA. *Plasmid.* 2005; 54:219–28. [PubMed: 15925406]
12. Crews, P.; Rodriguez, J.; Jaspars, M. *Organic Structure Analysis.* Oxford University Press; New York: 1998.
13. Pretch, E.; Bühlmann, P.; Affolter, C. *Structure Determination of Organic Compounds: Tables of Spectral Data.* Springer; Berlin: 2000.
14. (a) Emmert EA, Klimowicz AK, Thomas MG, Handelsman J. *Appl Environ Microbiol.* 2004; 70:104–13. [PubMed: 14711631] (b) Chan YA, Boyne MT 2nd, Podevels AM, Klimowicz AK, Handelsman J, Kelleher NL, Thomas MG. *Proc Natl Acad Sci U S A.* 2006; 103:14349–54. [PubMed: 16983083]
15. Lin S, Van Lanen SG, Shen B. *Proc Natl Acad Sci U S A.* 2009; 106:4183–8. [PubMed: 19246381]
16. Du L, Lou L. *Nat Prod Rep.* 2010; 27:255–78. [PubMed: 20111804]
17. Ulrich RL, Hines HB, Parthasarathy N, Jeddelloh JA. *J Bacteriol.* 2004; 186:4350–60. [PubMed: 15205437]
18. Subramoni S, Venturi V. *Microbiology.* 2009; 155:1377–85. [PubMed: 19383698]
19. Schobert R, Schlenk A. *Bioorg Med Chem.* 2008; 16:4203–21. [PubMed: 18334299]
20. Attempts to disrupt the MAL cluster in BP and BM by conventional mutagenesis protocols were unsuccessful, suggesting a more prominent role of malleilactone in the physiology of these highly pathogenic species. This observation is supported by the inability to generate malR deletions in either strain, previously detailed in Ulrich RL, Deshazer D, Brueggemann EE, Hines HB, Oyston PC, Jeddelloh JA. *J Med Microbiol.* 2004; 53:1053–64. [PubMed: 15496380] Ulrich RL, Deshazer D, Hines HB, Jeddelloh JA. *Infect Immun.* 2004; 72:6589–96. [PubMed: 15501791]
21. Miethke M, Marahiel MA. *Microbiol Mol Biol Rev.* 2007; 71:413–51. [PubMed: 17804665]
22. Nagachar N, Ratledge C. *FEMS Microbiol Lett.* 2010; 311:193–9. and references within. [PubMed: 20735479]

**Figure 1.**

(a) Venn diagram showing the relationship of number of NRPS/PKS gene clusters shared among pseudomallei group pathogens *B. pseudomallei* K96243, *B. mallei* ATCC 23344 and *B. thailandensis* E264 (*BP*, *BM* and *BT* respectively). (b) *C. elegans* after 24 h co-culture with: *i*. wild type *BT* (dead), *ii*. *BT*($\Delta malF$), a PKS2 gene deletion mutant and *iii*. *BT*($\Delta malR$), a transcription factor deletion mutant (40x magnification). (c) *C. elegans* survival after 24 h exposure to wild type *BT*, *BT*($\Delta malF$), *BT*($\Delta malR$) and *BT*($\Delta malR/malA$), a *malR*-PKS1 double mutant. *** $P < 0.001$, two-tailed *t*-test. (d) *D. discoideum* co-culture at 120 h with: *i*. wild type *BT* and *ii*. *BT*($\Delta malF$). Arrows highlight *D. discoideum* aggregation and differentiation into fruiting bodies with MAL cluster disruption (white box - close-up of a fruiting body). (e) MAL clusters from *BP*, *BM* and *BT* with predicted functions for each MAL protein.

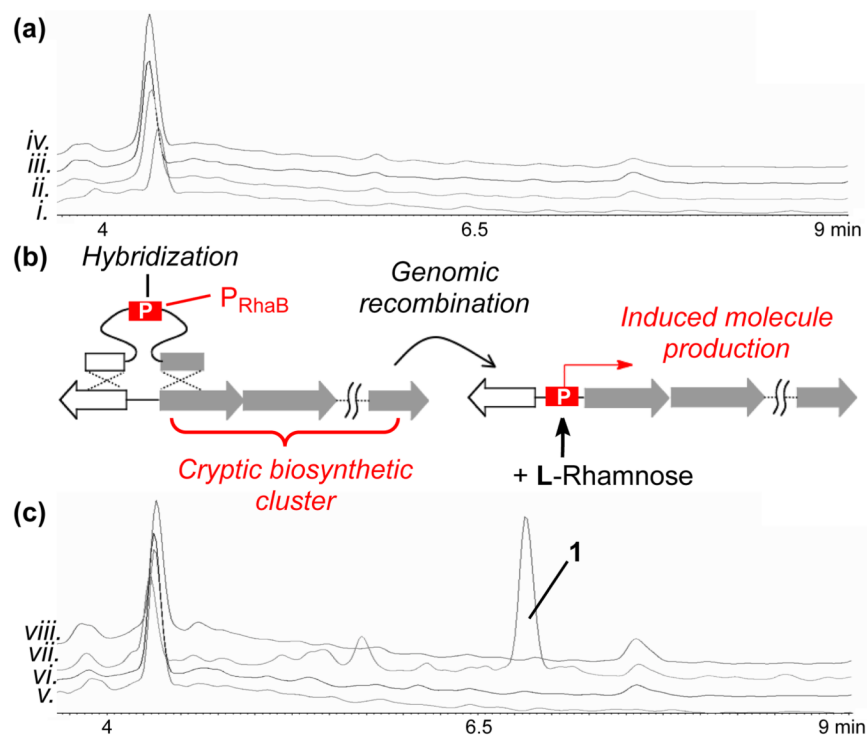


Figure 2.

(a) HPLC traces (total diode array: 210-450 nm) of *BT* culture broth ethyl acetate extracts. *i.* wild type *BT*, *ii.* *BT*($\Delta malF$), *iii.* *BT*($\Delta malR$), *iv.* *BT*($\Delta malR/\Delta malA$). (b) Promoter exchange strategy used to induce cryptic gene cluster expression. In brief, *BT* cells are transformed with a cassette designed to hybridize and recombine with the native promoter region resulting in the insertion of the rhamnose-inducible promoter P_{RhaB} directly upstream of the cryptic biosynthetic gene cluster. Small molecule production is then induced with the addition of 0.2% L-rhamnose to the culture media. (c) HPLC traces of culture broth extracts, showing that the induction of the MAL cluster leads to the production of the novel metabolite (malleilactone, **1**). Cultures were grown with or without 0.2% L-rhamnose: *v.* wild type *BT* with rhamnose, *vi.* *BT*: P_{RhaB} -MAL, no rhamnose, *vii.* *BT*: P_{RhaB} -MAL with rhamnose, *viii.* *BT*: P_{RhaB} -MAL($\Delta malF$) with rhamnose.

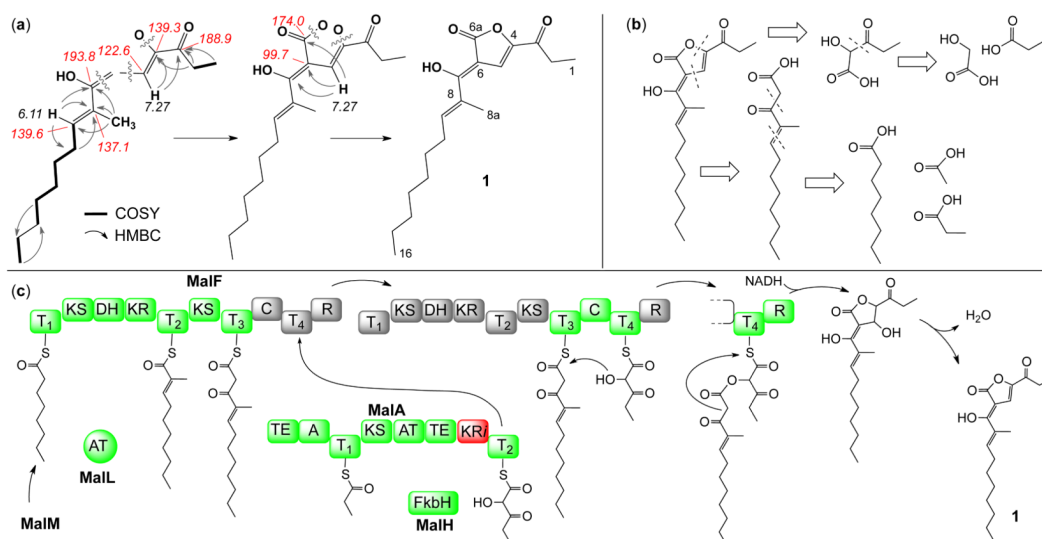


Figure 3.

(a) Key NMR arguments used to define the structure of malleilactone (1). (b) Retro-biosynthetic analysis of 1. (c) Biosynthetic proposal for 1. AT, acyltransferase; KS, ketosynthase; DH, dehydratase; KRⁱ, inactive ketoreductase; C, condensation domain; R, reductase domain; T, thiolation domain.