LETTER

doi:10.1038/nature14137

Metabolic coupling of two small-molecule thiols programs the biosynthesis of lincomycin A

Qunfei Zhao[1]*, Min Wang[1]*, Dongxiao Xu[1], Qinglin Zhang[2] & Wen Liu[1,2]

Low-molecular-mass thiols in organisms are well known for their redox-relevant role in protection against various endogenous and exogenous stresses[1–3]. In eukaryotes and Gram-negative bacteria, the primary thiol is glutathione (GSH), a cysteinyl-containing tripeptide. In contrast, mycothiol (MSH), a cysteinyl pseudo-disaccharide, is dominant in Gram-positive actinobacteria, including antibiotic-producing actinomycetes and pathogenic mycobacteria. MSH is equivalent to GSH, either as a cofactor or as a substrate, in numerous biochemical processes[4], most of which have not been characterized, largely due to the dearth of information concerning MSH-dependent proteins. Actinomycetes are able to produce another thiol, ergothioneine (EGT), a histidine betaine derivative that is widely assimilated by plants and animals for variable physiological activities[5]. The involvement of EGT in enzymatic reactions, however, lacks any precedent. Here we report that the unprecedented coupling of two bacterial thiols, MSH and EGT, has a constructive role in the biosynthesis of lincomycin A, a sulfur-containing lincosamide (C8 sugar) antibiotic that has been widely used for half a century to treat Gram-positive bacterial infections[6–9]. EGT acts as a carrier to template the molecular assembly, and MSH is the sulfur donor for lincomycin maturation after thiol exchange. These thiols function through two unusual *S*-glycosylations that program lincosamide transfer, activation and modification, providing the first paradigm for EGT-associated biochemical processes and for the poorly understood MSH-dependent biotransformations, a newly described model that is potentially common in the incorporation of sulfur, an element essential for life and ubiquitous in living systems.

Mycothiol (MSH, Fig. 1a) mediated detoxification typically relies on the conjugation of MSH to an electrophilic toxin. An amidase, Mca, then hydrolyses the resulting MSH *S*-conjugate to produce a pseudo-disaccharide unit, 1-*O*-glucosamine-D-*myo*-inositol (GlcN-Ins), and a mercapturic acid derivative, which is an excretive *N*-acetyl-cysteinyl product common in GSH-mediated metabolism[10–12] (Fig. 1b). In actinomycetes, *mca* orthologues have been found in several biosynthetic gene clusters of antibiotics[4,11], including that of lincomycin A, suggesting that MSH *S*-conjugation is associated with the production of these antibiotics. Lincomycin A consists of an *N*-methylated 4-propyl-L-proline (PPL) moiety and lincosamide, an unusual eight-carbon sugar decorated with a methylmercapto group at C-1 (Fig. 1d). Cell protection against the activity of lincomycin A depends largely on methylation of the bacterial ribosome, whereby the molecule mimics the 3′ end of (de)acetyl-tRNA and blocks protein synthesis at the initial stage of the elongation cycle[13,14]. This fact, in combination with the methylmercapto group found in the structure, leads to the proposal that MSH *S*-conjugation has a constructive role in lincomycin biosynthesis by supplying sulfur rather than a protective role in antibiotic detoxification. To validate this hypothesis, we inactivated *lmbE*, a *mca* orthologue that is located in the lincomycin biosynthetic gene cluster (*lmb*) in *Streptomyces lincolnensis*[15,16] (Extended Data Fig. 2a).

As anticipated, the *ΔlmbE* mutant strain accumulated a MSH-associated lincomycin analogue, **1** (Fig. 2). In this molecule, the C8-sugar unit is appended to MSH via an α-*S*-linkage (Supplementary Text), identical to that of lincomycin A in configuration. The *ΔlmbE* mutant strain still produced lincomycin A, albeit in a lower yield, indicating that an additional *mca* orthologue is present outside the *lmb* cluster and partially

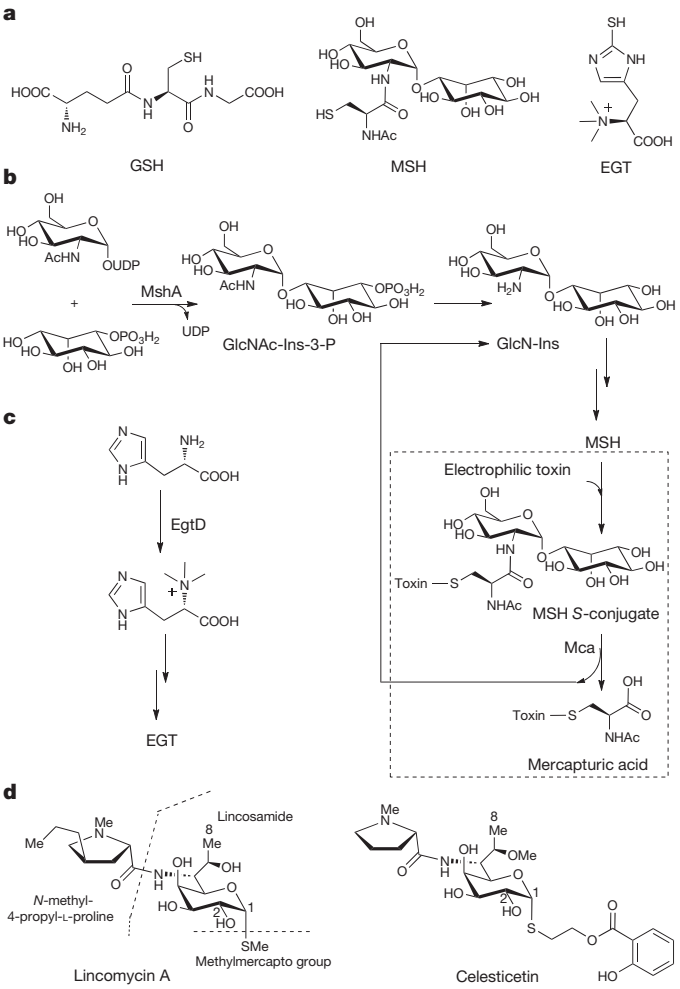


Figure 1 | Representative low-molecular-mass thiols, relevant metabolic pathways, and associated lincosamide natural products. **a**, Structures of the thiols GSH, MSH and EGT. **b**, Biosynthetic pathway of MSH and its typical associated detoxification process (shown in the dashed rectangle). GlcN-Ins is recyclable as an intermediate or product. **c**, Biosynthetic pathway of EGT. EgtD is an *S*-adenosyl methionine-dependent protein that catalyses the first reaction. **d**, Structures of the lincosamide antibiotics lincomycin A and celesticetin.

[1]State Key Laboratory of Bioorganic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 345 Lingling Road, Shanghai 200032, China. [2]Huzhou Center of Bio-Synthetic Innovation, 1366 Hongfeng Road, Huzhou 313000, China.
*These authors contributed equally to this work.

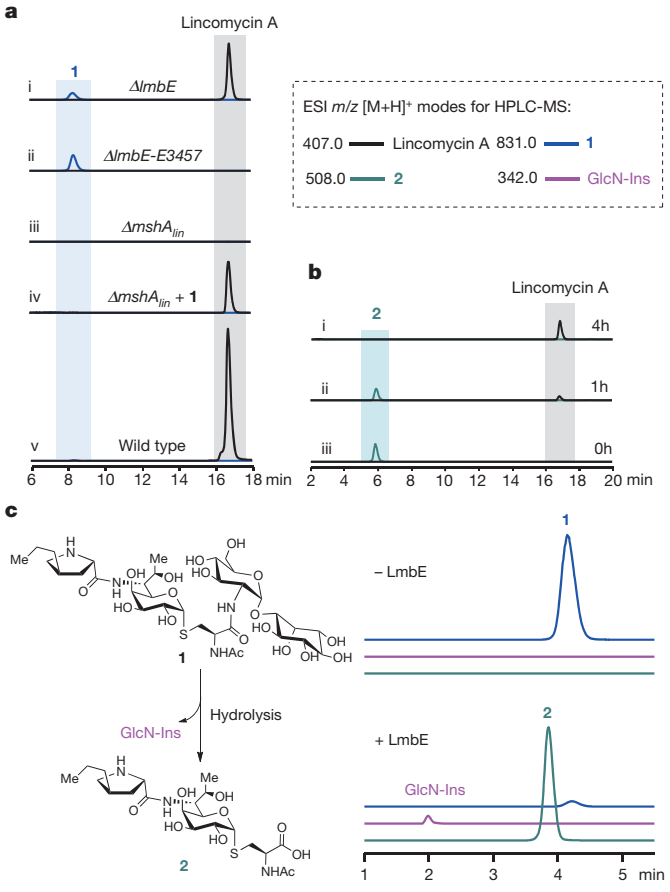


Figure 2 | Characterization of LmbE as a pathway-specific Mca protein to process 1, the MSH *S*-conjugated intermediate, in the biosynthesis of lincomycin A. For high-performance liquid chromatography-mass spectrometry (HPLC-MS) analysis, the electrospray ionization (ESI) *m/z* $[M + H]^+$ modes are indicated in the dashed rectangle. **a**, *In vivo* product profiles of *S. lincolnensis* strains, including the mutants (i, for *ΔlmbE*; ii, for *ΔlmbE-E3457*; iii, for *ΔmshA$_{lin}$*; and iv, for *ΔmshA$_{lin}$* supplemented with ~0.20 mM intermediate **1**) and the wild-type control (v). **b**, Time-dependent biotransformation of mercapturic acid **2** (~0.20 mM) into lincomycin A using the cell homogenate of the *ΔmshA$_{lin}$* mutant strain. **c**, *In vitro* hydrolysis of intermediate **1** to generate **2** and GlcN-Ins (left) in the absence (top right) and presence (lower right) of LmbE.

compensates for the loss of *lmbE*. Sequencing the genome of *S. lincolnensis* revealed three *lmbE* homologues: *lmbE80*, *lmbE447* and *lmbE3457*. The individual inactivation of these genes was performed in the *ΔlmbE* mutant strain, but only the *ΔlmbE-E3457* double mutant strain completely lost lincomycin producibility and had a concomitant increase in **1** (Fig. 2a and Extended Data Fig. 3). Moreover, we expressed and purified LmbE from *Escherichia coli*. This recombinant protein rapidly converted **1** into two products (Fig. 2c): GlcN-Ins and **2**, a mercapturic acid derivative (Supplementary Text). Thus, the involvement of LmbE as a specific Mca protein in lincomycin biosynthesis was established.

MSH biosynthesis begins with the glycosyltransferase (GTase) MshA, which catalyses the formation of 1-*O*-(2-*N*-acetyl)-glucosamine-D-*myo*-inositol-3-phosphate (GlcNAc-Ins-3-P) to afford the essential pseudodisaccharide unit[12] (Fig. 1b). To validate the necessity of MSH for lincomycin biosynthesis, we identified a *mshA* orthologue (Extended Data Fig. 2b), *mshA$_{lin}$*, from the *S. lincolnensis* genome and inactivated it in the wild-type strain. The *ΔmshA$_{lin}$* mutant strain failed to produce MSH along with lincomycin A (Fig. 2a and Extended Data Fig. 4). MSH *S*-conjugate **1** was then added to this mutant, leading to the restoration of lincomycin production (Fig. 2a). Unambiguously, **1** is a key intermediate rather than a detoxified antibiotic in the biosynthetic pathway. Notably, adding ~0.20 mM **1** to the cells yielded ~1.08 mM lincomycin A after a 5 day cultivation period, representing a ~4-fold increase in product concentration compared with the precursor. This increase could result from the regeneration of intermediate **1** by restoring the MSH cycle, because GlcN-Ins, the product originating from **1** through LmbE-catalysed hydrolysis, is the intermediate in MSH biosynthesis[4,11] (Fig. 1b). The cell homogenate of the *ΔmshA$_{lin}$* mutant strain was capable of transforming ~0.20 mM **2** into a nearly equal amount of lincomycin A (Fig. 2b), further confirming the essentiality of the LmbE-catalysed reaction in the lincomycin biosynthetic pathway, which hydrolyses **1** to generate recyclable GlcN-Ins and intermediate **2**. Processing **2**, including C–S bond cleavage of the cysteinyl group (mechanistically similar to that in the GSH-associated biosynthesis of gliotoxin[17]) and subsequent *S*-methylation of the exposed sulfhydryl group, may eventually afford the sulfur appendage of lincomycin A. Notably, the generation of excretive thiomethyl products has previously been found as an alternative to the mercapturate pathway for xenobiotic detoxification in GSH-mediated metabolism[18].

We next focused on how and when MSH is incorporated to generate *S*-conjugate **1**. Sequence analysis of the *lmb* cluster revealed two closely linked, functionally unassigned genes (Extended Data Fig. 2a), *lmbT* and *lmbV*. *lmbT* encodes a protein belonging to the GTase superfamily. Recently, lincosamide formation has proven to involve the generation of GDP-D-α-D-octose and its associated modifications[19–21]. The phosphonucleotidyl group on the resulting product, GDP-D-α-D-lincosamide (**3**) (Fig. 3), facilitates the attack by nucleophiles such as MSH, probably requiring a GTase activity; however, direct transfer of lincosamide from GDP onto MSH appears unlikely, because GTase-catalysed glycosylation is often envisioned proceeding through S_N2 nucleophilic displacement, which would not explain the same α-linkage that is predicted in **3** and found in intermediate **1**. *lmbV* encodes a protein classified into the DinB-2 superfamily (Extended Data Fig. 5), which now includes over ten thousand members with activities presumably related to various low-molecular-mass thiols[22]. These proteins share a conserved DinB-2-like domain combined with variable functional domain(s), indicating that thiols potentially act in different biochemical processes. Notably, a clade phylogenetically relevant to LmbV contains the MSH-maleylpyruvate isomerase (Extended Data Fig. 5), one of the few MSH-dependent proteins that have been biochemically characterized for the isomerization of maleylpyruvate to fumarylpyruate[23,24]. Accordingly, LmbV was proposed to catalyse a MSH-dependent reaction, although the role of MSH remained unknown.

We next established the relevance of LmbT and LmbV to lincomycin biosynthesis, as the inactivation of *lmbT* or *lmbV* completely abolished lincomycin production, which was partially restored by complementing each of the genes in the corresponding mutant strain (Fig. 3a). Surprisingly, a lincomycin analogue, EGT *S*-conjugate **4**, was found in the *ΔlmbV* mutant (Fig. 3 and Supplementary Text). Re-examination of the *ΔmshA$_{lin}$* mutant strain showed a similar product profile in which **4** was the major product (Fig. 3a); evidently, EGT *S*-conjugation is independent of MSH. To correlate this thiol with lincomycin production, we surveyed the *S. lincolnensis* genome and identified the EGT biosynthetic genes[25] (Extended Data Fig. 2c), in which *egtD$_{lin}$*, responsible for triple *N*-methylation of histidine in EGT biosynthesis (Fig. 1c), was chosen for inactivation. The *ΔegtD$_{lin}$* mutant strain produced lincomycin A, the yield of which, however, was significantly reduced at the indicated time points (Fig. 3b). Analysis of the mutant cells revealed a trace amount of EGT (with a yield ~16.8% of that in the wild-type cells, Extended Data Fig. 4), which probably resulted from the exogenous assimilation of the culture medium that included plant-derived, EGT-containing components[4,5]. Over time, the production of lincomycin A from the mutant became much lower than that from the wild-type strain (Fig. 3b), presumably because of the limiting amount of EGT. Altogether, EGT is necessary for lincomycin biosynthesis, whereby its *S*-conjugate **4** is an intermediate instead of a shunt product in the pathway.

Notably, in EGT *S*-conjugate **4**, lincosamide is attached to the thiol via a β-*S*-linkage (Fig. 3 and Supplementary Text). If **1** is the product of

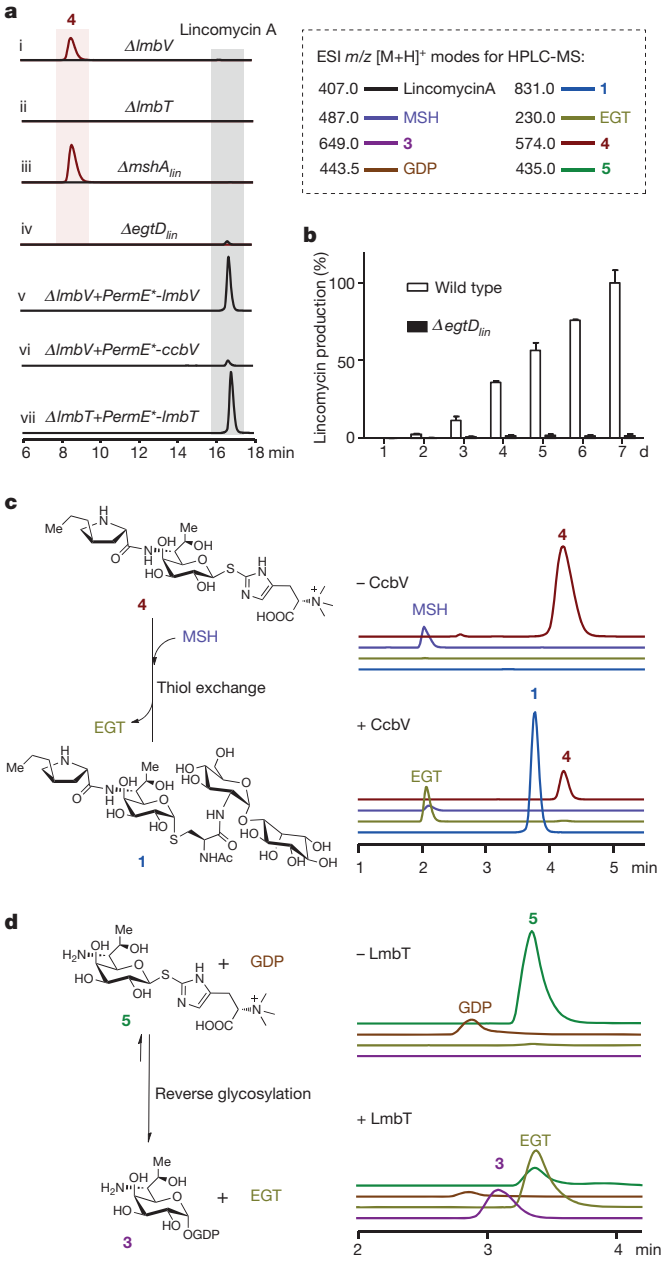


Figure 3 | Characterization of MSH- and EGT-associated biotransformations. For HPLC-MS analysis, the ESI *m/z* $[M + H]^+$ modes are indicated in the dashed rectangle. **a**, *In vivo* product profiles of *S. lincolnensis* strains, including the mutants *ΔlmbV* (i), *ΔlmbT* (ii), *ΔmshA*$_{lin}$ (iii) and *ΔegtD*$_{lin}$ (iv); the *ΔlmbV* mutant harbouring homologous *lmbV* (v) or heterologous *ccbV* (vi); and the *ΔlmbT* mutant harbouring homologous *lmbT* (vii). *PermE** is the constitutive promoter for expressing the erythromycin-resistance gene in *Saccharopolyspora erythraea*. **b**, Time-course analysis of lincomycin production in the wild-type (white) and *ΔegtD*$_{lin}$ mutant (black) strains. Mean ± s.e.m. ($n = 3$) are shown. **c**, *In vitro* thiol exchange for converting EGT *S*-conjugate **4** into MSH *S*-conjugate **1** (left) in the absence (top right) and presence of CcbV (lower right). **d**, *In vitro* reverse glycosylation for converting EGT *S*-conjugate **5** into GDP-mediated C8 sugar **3** (left) in the absence (top right) and presence of LmbT (lower right).

LmbV, MSH is probably the agent that performs the nucleophilic attack at C-1 of the lincosamide unit, which is activated by EGT rather than (d)NDP, and the resulting inversion of configuration is consistent with the glycosylation reaction that proceeds via an S_N2 displacement to generate the α-*S*-linkage in MSH *S*-conjugate **1**. Therefore, we chemically synthesized MSH (Supplementary Methods), making use of the scaled-up precursor GlcN-Ins, for an *in vitro* assay of LmbV activity.

Despite our numerous attempts, LmbV was highly refractory to various methods of expression. We then selected CcbV, a homologue of LmbV (57% identity) in celesticetin biosynthesis[26], as the catalyst. Celesticetin is a naturally occurring analogue of lincomycin A that possesses the same α-*S*-linkage (Fig. 1d). Heterologous expression of the gene *ccbV* in the *ΔlmbV* mutant strain partially restored the production of lincomycin A (Fig. 3a), confirming that CcbV functionally substitutes for LmbV. The *in vitro* conversion of **4** occurred in the presence of CcbV, resulting in the production of **1** along with the release of EGT (Fig. 3c and Extended Data Fig. 6). This reaction is irreversible, as MSH *S*-conjugate **1** was not converted back to EGT *S*-conjugated **4** in the presence of EGT. Therefore, LmbV and its counterpart CcbV form a new type of GTase that catalyses a unique *S*-glycosylation for thiol exchange between EGT and MSH.

We thus considered that the other GTase, LmbT, could transfer lincosamine from GDP onto EGT via S_N2 nucleophilic displacement. LmbT was expressed and purified, and to determine its activity, the reverse glycosylation reaction was conducted according to a method established previously using saturated GDP as the nucleophilic agent[27]. Because LmbT did not convert EGT *S*-conjugate **4** to the corresponding GDP-mediated sugar product (Extended Data Fig. 7a), we re-examined the *ΔlmbT* mutant strain and found that the PPL moiety had accumulated extensively (Extended Data Fig. 8a). This finding suggested that the *S*-glycosylation of EGT precedes PPL incorporation. Consistently, inactivation of *lmbC*, which encodes an adenylation protein characterized in PPL activation[28], resulted in the production of PPL and an EGT *S*-conjugated lincosamine, **5** (Extended Data Fig. 8a and Supplementary Text). Compared with **4**, **5** lacks the PPL moiety. A similar result was found in the *ΔlmbN* and *ΔlmbD* mutant *S. lincolnensis* strains (Extended Data Fig. 8a), supporting the corresponding proteins, LmbC, LmbN and LmbD, being responsible for incorporating PPL to transform **5** into **4**. LmbN is a bi-functional protein, and its 1,2-isomerization activity has recently been shown in lincosamine formation[20]. Careful analysis of the protein sequence revealed a peptidyl carrier protein (PCP) domain present at its amino terminus. LmbC is thus considered to activate PPL with ATP and transfer it onto this PCP domain, followed by LmbD-catalysed condensation with EGT *S*-conjugate **5** to afford **4** (Extended Data Fig. 8b). Consequently, LmbC, LmbN (or LmbN-PCP, the N-terminal PCP domain) and CcbD (the homologue in celesticelin biosynthesis[26] that is functionally identical to LmbD) were expressed and purified, and *in vitro* assays showed that these proteins indeed coordinate PPL attachment to generate **4** (Extended Data Fig. 8). Clearly, functionalization of the lincosamine unit by PPL occurs in an EGT *S*-conjugated form, and the PPL-lacking compound **5** probably serves as the product of LmbT-catalysed *S*-glycosylation.

As anticipated, in the presence of LmbT and the co-substrate GDP, **5** was efficiently transformed to GDP-D-α-D-lincosamide **3**, accompanied by EGT release (Fig. 3d and Extended Data Fig. 9). We then validated the forward activity of LmbT to produce **5** and GDP using the substrates **3** and EGT (Extended Data Fig. 7b). The LmbT-catalysed reaction exhibited an equilibrium constant K_{eq} of 1.94 (Extended Data Fig. 9d), indicating that the reverse and forward conversions are comparable; however, the activities of downstream enzymes may drive the pathway forward to produce lincomycin A. Consequently, LmbT represents a new enzyme that employs EGT as a sugar acceptor and catalyses *S*-glycosylation with the naturally rare C8 sugar to generate EGT *S*-conjugate **5**.

The *ΔlmbC*, *ΔlmbN* or *ΔlmbD* mutant strain produced a minor product, **6** (Extended Data Fig. 8a), which is a MSH *S*-conjugated lincosamide lacking the PPL moiety (Supplementary Text), suggesting that LmbV tolerated **5** as a substrate for thiol exchange. CcbV, the homologue of LmbV, converted **5** into **6** *in vitro* (Extended Data Fig. 10a); however, the rate of this reaction was much lower than that for the transformation of **4** into **1**. The Mca protein LmbE was not active on **6** and failed to produce mercapturic acid for further processing (Extended Data Fig. 10b). Clearly, lincomycin biosynthesis involves the EGT-mediated

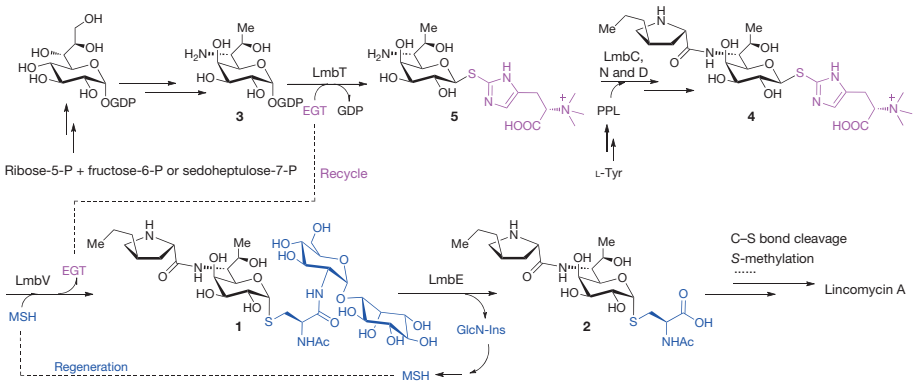


Figure 4 | The EGT (pink) and MSH (blue) programmed biosynthetic pathway of lincomycin A. The recycling of EGT and the regeneration of MSH are shown as dashed lines.

assembly of lincosamide with the building blocks, PPL and MSH, and the MSH-associated post-modifications in a highly ordered process.

We uncovered a constructive role of two different bacterial thiols in *S. lincolnensis* and clarified the biosynthetic pathway of lincomycin A, in which EGT and MSH-associated metabolism creates an extraordinary biochemical strategy for the formation of active molecule (Fig. 4 and Extended Data Fig. 1). Lincomycin biosynthesis involves a convergent pathway to synthesize the L-Tyr-derived PPL moiety and the GDP-activated C8 sugar lincosamide. The thiol EGT serves as a carrier, via the first *S*-glycosylation (reversible), to channel the lincosamine unit and mediates its condensation with PPL. The thiol MSH goes through the second *S*-glycosylation (irreversible) to associate with the lincomycin intermediate resulting from EGT and then acts as the sulfur donor for affording the methylmercapto group. Both thiols are recyclable or reproducible, thus maintaining the biosynthetic route to lincomycin A.

The characterization of the low-molecular-mass thiol-programmed biosynthetic pathway largely expands our knowledge regarding the intrinsic, versatile functions of thiols, which are apparently not limited to a protective role against oxidative stress and the neutralization of electrophilic toxins. The lincosamide antibiotic celesticetin probably shares this biosynthetic strategy based on the thiols EGT and MSH, despite their differences in processing the MSH appendage[26]. The involvement of EGT in C8-sugar transfer and activation represents the first biochemical evidence of this thiol in enzymatic reactions and generates interest in nucleotide-independent sugar modifications and associated glycosylations, which have been less appreciated to date[29]. Sulfur is one of the most abundant elements in living organisms and contributes to a large number of biologically active natural products; however, incorporation of this atom has not been well established. Complementing the recent advance of co-opting the sulfur-delivery system of primary metabolism for thiosugar formation[30], we demonstrated that MSH serves as a different source for sulfur incorporation (Fig. 4). This could be a general paradigm because the biosynthetic pathways of several sulfur-containing natural products involve homologues of Mca, the protein responsible for processing the MSH *S*-conjugate[4,11]. The findings reported here represent a key step towards elucidating the biochemical mechanisms of numerous MSH and EGT-dependent but poorly understood proteins and exploring new features of thiols with regard to their currently unknown associated biochemical processes.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 18 July; accepted 3 December 2014.
Published online 14 January 2015.

1. Fahey, R. C. Glutathione analogs in prokaryotes. *Biochim. Biophys. Acta* **1830,** 3182–3198 (2013).
2. Deponte, M. Glutathione catalysis and the reaction mechanisms of glutathione-dependent enzymes. *Biochim. Biophys. Acta* **1830,** 3217–3266 (2013).
3. Newton, G. L. *et al.* Bacillithiol is an antioxidant thiol produced in Bacilli. *Nature Chem. Biol.* **5,** 625–627 (2009).
4. Rawat, M. & Av-Gay, Y. Mycothiol-dependent proteins in actinomycetes. *FEMS Microbiol. Rev.* **31,** 278–292 (2007).
5. Cheah, I. K. & Halliwell, B. Ergothioneine; antioxidant potential, physiological function and role in disease. *Biochim. Biophys. Acta* **1822,** 784–793 (2012).
6. Mason, D. J., Dietz, A. & DeBoer, C. Lincomycin, a new antibiotic. I. Discovery and biological properties. *Antimicrob. Agents Chemother.* 554–559 (1962).
7. Herr, R. R. & Bergy, M. E. Lincomycin, a new antibiotic. II. Isolation and characterization. *Antimicrob. Agents Chemother.* 560–554 (1962).
8. Lewis, C., Clapp, H. W. & Grady, J. E. *In vitro* and *in vivo* evaluation of lincomycin, a new antibiotic. *Antimicrob. Agents Chemother.* 570–582 (1963).
9. Mason, D. J. & Lewis, C. Biological activity of the lincomycin related antibiotics. *Antimicrob. Agents Chemother.* 7–12 (1964).
10. Newton, G. L., Buchmeier, N. & Fahey, R. C. Biosynthesis and functions of mycothiol, the unique protective thiol of *Actinobacteria*. *Microbiol. Mol. Biol. Rev.* **72,** 471–494 (2008).
11. Jothivasan, V. K. & Hamilton, C. J. Mycothiol: synthesis, biosynthesis and biological functions of the major low molecular weight thiol in actinomycetes. *Nature Prod. Rep.* **25,** 1091–1117 (2008).
12. Fan, F., Vetting, M. W., Frantom, P. A. & Blanchard, J. S. Structures and mechanisms of the mycothiol biosynthetic enzymes. *Curr. Opin. Chem. Biol.* **13,** 451–459 (2009).
13. Reusser, F. Effect of lincomycin and clindamycin on peptide chain initiation. *Antimicrob. Agents Chemother.* **7,** 32–37 (1975).
14. Wilson, D. N. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. *Nature Rev. Microbiol.* **12,** 35–48 (2014).
15. Peschke, U., Schmidt, H., Zhang, H.-Z. & Piepersberg, W. Molecular characterization of the lincomycin-production gene cluster of *Streptomyces lincolnensis* 78–11. *Mol. Microbiol.* **16,** 1137–1156 (1995).
16. Koběrská, M. *et al.* Sequence analysis and heterologous expression of the lincomycin biosynthetic cluster of the type strain *Streptomyces lincolnensis* ATCC 25466. *Folia Microbiol.* **53,** 395–401 (2008).
17. Scharf, D. H. *et al.* Epidithiol formation by an unprecedented twin carbon-sulfur lyase in the gliotoxin. *Angew. Chem. Int. Ed.* **124,** 10211–10215 (2012).
18. Cooper, A. J. L. *et al.* Cysteine *S*-conjugate β-lyases: important roles in the metabolism of naturally occurring sulfur and selenium-containing compounds, exnobiotics, and anticancer agents. *Amino Acids* **41,** 7–27 (2011).
19. Brahme, N. M. *et al.* Biosynthesis of the lincomycins. 2. Studies using stable isotopes on the biosynthesis of methylthiolincosaminide moiety of lincomycin A. *J. Am. Chem. Soc.* **106,** 7878–7883 (1984).
20. Sasaki, E., Lin, C.-I., Lin, K.-Y. & Liu, H.-w. Construction of the octose 8-phosphate intermediate in lincomycin A biosynthesis: characterization of the reactions catalyzed by LmbR and LmbN. *J. Am. Chem. Soc.* **134,** 17432–17435 (2012).
21. Lin, C.-I., Sasaki, E., Zhong, A. & Liu, H.-w. *In vitro* characterization of LmbK and LmbO: identification of GDP-D-*erythro*-α-D-*gluco*-octose as a key intermediate in lincomycin A biosynthesis. *J. Am. Chem. Soc.* **136,** 906–909 (2014).
22. Newton, G. I., Leung, S. S., Wakabayashi, J. I., Rawat, M. & Fahey, R. C. The DinB superfamily includes novel mycothiol, bacillithiol, and glutathione *S*-transferases. *Biochemistry* **50,** 10751–10760 (2011).
23. Feng, J. *et al.* The gene *ncgl2918* encodes a novel maleylpyruvate isomerase that needs mycothiol as cofactor and links mycothiol biosynthesis and gentisate assimilation in *Corynebacterium glutamicum*. *J. Biol. Chem.* **281,** 10778–10785 (2006).
24. Wang, R. *et al.* Crystal structures and site-directed mutagenesis of a mycothiol-dependent enzyme reveal a novel folding and molecular basis for mycothiol-mediated maleylpyruvate isomerization. *J. Biol. Chem.* **282,** 16288–16294 (2007).
25. Seebeck, F. P. *In vitro* reconstitution of mycobacterial ergothioneine biosynthesis. *J. Am. Chem. Soc.* **132,** 6632–6633 (2010).
26. Čermák, L. *et al.* Hybridization analysis and mapping of the celesticetin gene cluster revealed genes shared with lincomycin biosynthesis. *Folia Microbiol.* **52,** 457–462 (2007).
27. Zhang, C. *et al.* Exploiting the reversibility of natural product glycosyltransferase-catalyzed reactions. *Science* **313,** 1291–1294 (2006).

28. Kadlčík, S. *et al.* Adaptation of an L-proline adenylation domain to use 4-propyl-L-proline in the evolution of lincosamide biosynthesis. *PLoS ONE* **8,** e84902 (2013).
29. Gantt, R. W., Peltier-Pain, P. & Thorson, J. S. Enzymatic methods for glyco (diversification/randomization) of drugs and small molecules. *Natural Prod. Rep.* **28,** 1811–1853 (2011).
30. Sasaki, E. *et al.* Co-opting sulphur-carrier proteins from primary metabolic pathways for 2-thiosugar biosynthesis. *Nature* **510,** 427–431 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported in part by grants from NSFC (81302674, 31430005, 91213303, 21472231 and 91413101), STCSM (14JC1407700 and 13XD1404500) and MST (2012AA02A706) of China.

Author Contributions Q.Zhao, M.W. and Q.Zhang performed *in vivo* investigations. M.W. and D.X. prepared and characterized the chemical compounds. Q.Zhao and W.L. conducted sequence analysis. Q.Zhao and M.W. performed *in vitro* enzymatic investigations. Q.Zhao, M.W. and W.L. analysed the data and wrote the manuscript. W.L. directed the research.

Author Information The sequences of the genes *lmbE80*, *lmbE447*, *lmbE3457*, $mshA_{lin}$ and $egtD_{lin}$ are deposited in GenBank with the NCBI accession numbers KJ958528, KJ958529, KJ958530, KJ958531 and KJ958532, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to W.L. (wliu@mail.sioc.ac.cn).

METHODS

General materials and methods. Biochemicals and media were purchased from Sinopharm Chemical Reagent Co., Ltd (China), Oxoid Ltd (UK) or Sigma-Aldrich Corporation (USA) unless otherwise stated. Restriction endonucleases were purchased from Thermo Fisher Scientific Co. Ltd (USA). Chemical reagents were purchased from standard commercial sources.

DNA isolation and manipulation in *Escherichia coli* or *Streptomyces* strains were carried out according to standard methods[31,32]. PCR amplifications were carried out on an Applied Biosystems Veriti Thermal Cycler using either Taq DNA polymerase (Vazyme Biotech Co. Ltd, China) for routine genotype verification or Phanta Max Super-Fidelity DNA Polymerase (Vazyme Biotech Co. Ltd, China) for high fidelity amplification. Primer synthesis was performed at Shanghai Sangon Biotech Co. Ltd (China), and DNA sequencing was performed at Shanghai Majorbio Biotech Co. Ltd or Shenzhen BGI in China.

High performance liquid chromatography (HPLC) analysis was carried out on the Agilent 1200 HPLC system (Agilent Technologies Inc., USA). HPLC Electrospray ionization MS (HPLC-ESI-MS) analysis was performed on the Thermo Fisher LTQ Fleet ESI-MS spectrometer (Thermo Fisher Scientific Inc., USA), and the data were analysed using Thermo Xcalibur software. ESI-high resolution MS (ESI-HR-MS) analysis was carried out on the 6230B Accurate-Mass TOF LC/MS System or 6530 Accurate-Mass Q-TOF LC/MS System (Agilent Technologies Inc., USA) and the data were analysed using Agilent MassHunter Qualitative Analysis software. NMR data were recorded on the Bruker DRX400 and Bruker AV500 spectrometers (Bruker Co. Ltd, Germany), or on the Agilent 500 MHz PremiumCompact+ NMR spectrometer (Agilent Technologies Inc., USA).

No statistical methods were used to predetermine sample size.

Gene inactivation and complementation. The inactivation of each gene in *S. lincolnensis* was performed by in-frame deletion to exclude polar effects on downstream gene expression. For complementation *in trans*, the target gene was under the control of *PermE**, the constitutive promoter for expressing the erythromycin-resistance gene in *Saccharopolyspora erythraea*. The genomic DNA of the *S. lincolnensis* wild-type strain served as the template for PCR amplification unless otherwise stated.

For *lmbE* deletion, the 2.52-kb fragment obtained using primers lmbE-L-for and lmbE-L-rev was digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1001. The 2.34-kb fragment obtained using primers lmbE-R-for and lmbE-R-rev was digested by XbaI and HindIII and cloned into the same site of pLL1001 to yield the recombinant plasmid pLL1002, in which a 549-bp in-frame coding region of *lmbE* was deleted. To transfer pLL1002 into the lincomycin-producing strain *S. lincolnensis*, conjugation between *E. coli* ET12567-Streptomyces was carried out following the standard procedure[1]. The colonies that were apramycin resistant at 37 °C were identified as integrating mutants, in which a single-crossover homologous recombination event took place. These mutants were cultured for several rounds in the absence of apramycin, and the resulting apramycin-sensitive isolates were subjected to PCR amplification to examine the genotype, as judged by the formation of the desired 1.1-kb product when using primers lmbE-gt-for and lmbE-gt-rev. Further sequencing of this PCR product confirmed the genotype of LL1001, in which *lmbE* was in-frame deleted.

For *lmbE-E80* double deletion, the 1.84-kb fragment obtained using primers lmbE80-L-for and lmbE80-L-rev was digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1003. The 2.06-kb fragment obtained using primers lmbE80-R-for and lmbE80-R-rev was digested by XbaI and HindIII and cloned into the same site of pLL1003, yielding the recombinant plasmid pLL1004. Following the procedure described above, pLL1004 was introduced into LL1001 for double-crossover recombination, yielding the recombinant strain LL1002, in which the 447-bp internal fragment of *lmbE80* was deleted in frame. Primers lmbE80-gt-for and lmbE80-gt-rev were used for genotype validation by PCR amplification. Further sequencing of this PCR product confirmed the genotype of LL1002, in which *lmbE80* was also in-frame deleted.

For *lmbE-E447* double deletion, the 1.81-kb fragment obtained using primers lmbE447-L-for and lmbE447-L-rev was digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1005. The 1.82-kb fragment obtained using primers lmbE447-R-for and lmbE447-R-rev was digested by XbaI and HindIII and cloned into the same site of pLL1005, yielding the recombinant plasmid pLL1006. Following the procedure described above, pLL1006 was introduced into LL1001 for double-crossover recombination, yielding the recombinant strain LL1003, in which the 459-bp internal fragment of *lmbE447* was deleted in frame. Primers lmbE447-gt-for and lmbE447-gt-rev were used for genotype validation by PCR amplification. Further sequencing of this PCR product confirmed the genotype of LL1003, in which *lmbE447* was also in-frame deleted.

For *lmbE-E3457* double deletion, the 1.79-kb fragment obtained using primers lmbE3457-L-for and lmbE3457-L-rev was digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1007. The 1.89-kb fragment obtained using primers lmbE3457-R-for and lmbE3457-R-rev was digested by XbaI and HindIII and cloned into the same site of pLL1007, yielding the recombinant plasmid pLL1008. Following the procedure described above, pLL1008 was introduced into LL1001 for double-crossover recombination, yielding the recombinant strain LL1004, in which the 402-bp internal fragment of *lmbE3457* was deleted in frame. Primers lmbE3457-gt-for and lmbE3457-gt-rev were used for genotype validation by PCR amplification. Further sequencing of this PCR product confirmed the genotype of LL1004, in which *lmbE3457* was also in-frame deleted.

For $mshA_{lin}$ deletion, the 2.25-kb fragment obtained using primers mshA-L-for and mshA-L-rev was digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1009. The 2.18-kb fragment obtained using primers mshA-R-for and mshA-R-rev was digested by XbaI and HindIII and cloned into the same site of pLL1009 to yield the recombinant plasmid pLL1010, in which a 633-bp in-frame coding region of $mshA_{lin}$ was deleted. Following the procedure described above, pLL1010 was introduced into the *S. lincolnensis* wild-type strain for double-crossover recombination. The resulting strain LL1005 was then subjected to PCR amplification to examine the genotype, as judged by the formation of the desired 1.09-kb product when using primers mshA-gt-for and mshA-gt-rev. Further sequencing of this PCR product confirmed the genotype of LL1005, in which $mshA_{lin}$ was in-frame deleted.

For *lmbT* deletion, the 2.56-kb fragment obtained using primers lmbT-L-for and lmbT-L-rev was digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1011. The 2.56-kb fragment obtained using primers lmbT-R-for and lmbT-R-rev was digested by XbaI and HindIII and cloned into the same site of pLL1011 to yield the recombinant plasmid pLL1012, in which a 462-bp in-frame coding region of *lmbT* was deleted. Following the procedure described above, pLL1012 was introduced into the *S. lincolnensis* wild-type strain for double-crossover recombination. The resulting strain LL1006 was then subjected to PCR amplification to examine the genotype, as judged by the formation of the desired 1.04-kb product when using primers lmbT-gt-for and lmbT-gt-rev. Further sequencing of this PCR product confirmed the genotype of LL1006, in which *lmbT* was in-frame deleted.

For *lmbV* deletion, the 2.29-kb fragment obtained using primers lmbV-L-for and lmbV-L-rev was digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1013. The 2.30-kb fragment obtained using primers lmbV-R-for and lmbV-R-rev was digested by XbaI and HindIII and cloned into the same site of pLL1013 to yield the recombinant plasmid pLL1014, in which a 399-bp in-frame coding region of *lmbV* was deleted. Following the procedure described above, pLL1014 was introduced into the *S. lincolnensis* wild-type strain for double-crossover recombination. The resulting strain LL1007 was then subjected to PCR amplification to examine the genotype, as judged by the formation of the desired 1.23-kb product when using primers lmbV-gt-for and lmbV-gt-rev. Further sequencing of this PCR product confirmed the genotype of LL1007, in which *lmbV* was in-frame deleted.

For $egtD_{lin}$ deletion, the 1.98-kb fragment obtained using primers egtD-L-for and edtD-L-rev was digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1019. The 1.75-kb fragment obtained using primers egtD-R-for and egtD-R-rev was digested by XbaI and HindIII and cloned into the same site of pLL1019 to yield the recombinant plasmid pLL1020, in which a 465-bp in-frame coding region of $egtD_{lin}$ was deleted. Following the procedure described above, pLL1020 was introduced into the *S. lincolnensis* wild-type strain for double-crossover recombination. The resulting strain LL1010 was then subjected to PCR amplification to examine the genotype, as judged by the formation of the desired 2.05-kb product when using primers egtD-gt-for and egtD-gt-rev. Further sequencing of this PCR product confirmed the genotype of LL1010, in which $egtD_{lin}$ was in-frame deleted.

For *lmbC* deletion, the 1.69-kb fragment obtained by using primers lmbC-L-for and lmbC-L-rev was digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1023. The 1.82-kb fragment obtained by using primers lmbC-R-for and lmbC-R-rev was digested by XbaI and HindIII and cloned into the same site of pLL1023 to yield the recombinant plasmid pLL1024, in which a 708-bp in-frame coding region of *lmbC* was deleted. Following the procedure described above, pLL1024 was introduced into the *S. lincolnensis* wild-type strain for double-crossover recombination. The resulting strain LL1012 was then subjected to PCR amplification to examine the genotype, as judged by the formation of the desired 0.82-kb product when using primers lmbC-gt-for and lmbC-gt-rev. Further sequencing of this PCR product confirmed the genotype of LL1012, in which *lmbC* was in-frame deleted.

For *lmbD* deletion, the 2.54-kb fragment obtained using primers lmbD-L-for and lmbD-L-rev was digested by EcoRI and XbaI, and cloned into the same sites of pKC1139 to yield plasmid pLL1031. The 2.51-kb fragment obtained using primers lmbD-R-for and lmbD-R-rev was digested by XbaI and HindIII and cloned into the same site of pLL1031 to yield the recombinant plasmid pLL1032, in which a

459-bp in-frame coding region of *lmbD* was deleted. Following the procedure described above, pLL1032 was introduced into the *S. lincolnensis* wild-type strain for double-crossover recombination. The resulting strain LL1016 was then subjected to PCR amplification to examine the genotype, as judged by the formation of the desired 0.60-kb product when using primers lmbD-gt-for and lmbD-gt-rev. Further sequencing of this PCR product confirmed the genotype of LL1016, in which *lmbD* was in-frame deleted.

For site-specific mutation of *lmbN*, the 2.01-kb fragment obtained using primers lmbN-L-for and lmbN-L-rev was digested by EcoRI and NehI, and cloned into the same sites of pKC1139 to yield plasmid pLL1033. The 1.94-kb fragment obtained using primers lmbN-R-for and lmbN-R-rev was digested by NheI and HindIII and cloned into the same site of pLL1033 to yield the recombinant plasmid pLL1034, which encoded a S37A mutation that replaced the conserved motif XXXDSL with XXXDAL and had a mutation that replaced CTC (encoding L38) with CTA (resulting no coding change) to introduce the NheI site. Following the procedure described above, pLL1034 was introduced into the *S. lincolnensis* wild-type strain for double-crossover recombination. The resulting strain LL1017 was then subjected to PCR amplification to give a 0.86-kb product (using primers lmbN-gt-for and lmbN-gt-rev). NheI digestion was then carried out on this PCR product to determine the genotype, as judged by the release of the desired 0.11-kb and 0.75-kb fragments. Further sequencing of this PCR product confirmed the genotype of LL1017, in which *lmbN* was site-specifically mutated.

For complementation of *lmbT* in LL1006, the 1.5-kb *lmbT*-containing fragment was amplified by PCR using primers lmbT-C-for and lmbT-C-rev, and then cloned into pMD19-T to yield pLL1015. After digestion with BamHI and XbaI, this 1.5-kb DNA fragment was recovered from pLL1015 and ligated to a 0.45-kb EcoRI/BamHII fragment from pWHM79, and the resulting product was ligated into the EcoRI/XbaI site of pKC1139, yielding the recombinant plasmid pLL1016, in which *lmbT* was under the control of the constitutive promoter *PermE**. pLL1016 was then introduced into LL1006 (*ΔlmbT* mutant) by conjugation, generating the corresponding recombinant strain LL1008 that expressed *lmbT in trans*.

For complementation of *lmbV* in LL1007, a 1.03-kb *lmbV*-containing fragment was amplified by PCR using primers lmbV-C-for and lmbV-C-rev, and then cloned into pMD19-T to yield pLL1017. After digestion with BamHI and XbaI, this 1.03-kb DNA fragment was recovered from pLL1017 and ligated to a 0.45-kb EcoRI/BamHII fragment from pWHM79, and the resulting product was then ligated into the EcoRI/XbaI site of pKC1139, yielding the recombinant plasmid pLL1018, in which *lmbV* was under the control of the constitutive promoter *PermE**. pLL1018 was then introduced into LL1007 (*ΔlmbV* mutant) by conjugation, generating the corresponding recombinant strain LL1009 that expressed *lmbV trans*.

For heterologous complementation of *ccbV* in LL1007, the 1.15-kb *ccbV*-containing fragment was amplified from the genomic DNA of the celesticetin-producing strain *S. caelestis* NRRL2418 by PCR using primers ccbV-C-for and ccbV-C-rev, and then cloned into pMD19-T to yield pLL1021. After digestion with BglII and XbaI, this 1.03-kb DNA fragment was recovered from pLL1021 and subsequently ligated to a 0.45-kb EcoRI/BamHII fragment from pWHM79; the resulting product was ligated into the EcoRI/XbaI site of pKC1139, yielding the recombinant plasmid pLL1022, in which *ccbV* was under the control of the constitutive promoter *PermE**. pLL1022 was then introduced into LL1007 (*ΔlmbV* mutant) by conjugation, generating the corresponding recombinant strain LL1011 that heterologously expressed *ccbV in trans*.

Production and analysis of lincomycin A, intermediate or shunt product. The *S. lincolnensis* wild-type strain or its derivative was spread on agar plates, which was composed of 19 g of starch, 5 g of soybean meal, 0.5 g of K_2HPO_4, 0.5 g of $MgSO_4 \cdot 7H_2O$, 1.0 g of KNO_3, 0.5 g of NaCl, 0.01 g of $FeSO_4 \cdot 7H_2O$ and 20.0 g of agar per litre (pH 7.0~7.5), and then incubated at 28 °C for sporulation and growth. Approximately 1 cm^2 of the sporulated agar of *S. lincolnensis* was cut, chopped, and inoculated into 25 ml of the seed medium, which was composed of 20 g of starch, 10 g of glucose, 10 g soybean of meal, 30 g of corn steep liquor, 1.5 g of $(NH_4)_2SO_4$ and 5 g of $CaCO_3$ per litre (pH 7.0). After incubation at 28 °C and 220 rpm for 36 h, 5 ml of the seed culture broth was transferred into 50 ml of the fermentation medium, which was composed of 100 g of glucose, 25 g of soybean meal, 2 g of corn steep liquor, 8 g of $(NH_4)_2SO_4$, 0.2 g of KH_2PO_4, 8 g of $NaNO_3$, 5 g of NaCl and 8 g of $CaCO_3$ per litre (pH 7.0). Further incubation was carried out at 28 °C and 220 rpm for 7 days.

For product examination, 500 µl of each fermentation broth was mixed with an equal volume of methanol, and after centrifugation to remove the residue, the supernatant was subjected to HPLC analysis on an Agilent Zorbax column (SB-C18, 5 µm, 4.6 × 250 mm, Agilent Technologies Inc., USA) by isocratic elution of 40% 5 mM ammonium acetate (NH_4Ac) and 60% MeOH for 20 min at a flow rate of 0.6 ml min^{-1}. HPLC-ESI-MS analysis of the supernatant diluted five times with 50% MeOH was carried out on the same column by gradient elution of solvent A (10 mM NH_4Ac) and solvent B (CH_3CN) at a flow rate of 1 ml min^{-1} over a 22-min period as follows: $t = 0$ min, 10% B; $t = 9$ min, 10% B; $t = 17$ min, 60% B; $t = 18$ min, 60% B; $t = 20$ min, 10% B; and $t = 22$ min, 10% B (mAU at 210 nM). The data were analysed using Thermo Xcalibur software. The concentration of lincomycin A was estimated by HPLC using the commercially available product as the standard.

For biotransformation of **2** (the mercapturic acid derivative) into loncomycin A by the cell homogenate of the $\Delta mshA_{lin}$ mutant strain, LL1005, the mycelia from 50 ml of the 6-day fermentation culture broth was collected, washed twice, and then re-suspended in 10 ml of 50 mM Tris-HCl buffer (pH 7.5). Sonication for 20 min on ice, followed by centrifugation at 4 °C to remove the cellular debris, resulted in the supernatant that was used for the conversion of **2**. Each biotransformation was conducted in 100 µl of the mixture, which contained 1 µl of **2** (giving a final concentration of 0.20 mM) and 99 µl of the supernatant of LL1005, and incubated at 30 °C for 0, 1 or 4 h. After quenching the reaction by addition of an equal volume of methanol, 20 µl of each biotransformation mixture was centrifuged and subjected to HPLC-ESI-MS analysis of **2** consumption and of lincomycin A production.

For feeding of the MSH *S*-conjugate **1** to LL1005 ($\Delta mshA_{lin}$), **1** was added into the culture broth of LL1005 on the 4th day of fermentation (giving a final concentration of approximately 0.20 mM). In a further 5-day incubation period, the production of lincomycin A was monitored daily by HPLC and HPLC-ESI-MS as described above.

Production and analysis of MSH and EGT in *S. lincolnensis*. Approximately 1 cm^2 of the sporulated agar of *S. lincolnensis* or its derivative was cut, chopped, and inoculated into 25 ml of YEME medium and then incubated at 28 °C and 220 rpm for 36 h. 5 ml of the resulting culture broth was added into 100 ml of the same medium for scale-up and then further incubated at 28 °C and 220 rpm for 3 days. The mycelia were then harvested by centrifugation at 4 °C and 5,000 rpm for 15 min.

For thiol extraction, derivatization and detection, the procedure described by Fahey and Newton[33] was used with slight modifications. Approximately 200 mg (wet weight) of the freshly harvested mycelia was weighed in a 5 ml microcentrifuge tube and then re-suspended in 2 ml of a mixture of 50% CH_3CN and 50% 2 mM monobromobimane (mBBr) dissolved in 20 mM Tris-HCl (pH 8.0) buffer for sonication. The mixture was incubated in the dark at 60 °C for 15 min, acidified with 5 µl of methanesulfonic acid (5N) and then centrifuged at 12,000 rpm for 10 min to remove the debris before storing at −80 °C. HPLC-ESI-HR-MS analysis of the resulting thiol-mBBr derivatives was carried out on a Agilent Zorbax column (Extend-C18, 1.8 µm, 2.1 × 50 mm, Agilent Technologies Inc., USA) by gradient elution of solvent A (H_2O) and solvent B (CH_3CN) at a flow rate of 0.2 ml min^{-1} over a 15 min period as follows: $t = 0$ min, 5% B; $t = 5$ min, 5% B; $t = 10$ min, 50% B; $t = 12$ min, 50% B; $t = 13$ min, 5% B; and $t = 15$ min, 5% B (mAU at 370 nm).

Protein expression and purification. The recombinant proteins LmbE, CcbV, LmbT, LmbC, LmbN, LmbN-PCP and CcbD, all of which were in an N-terminal 6× His-tagged form, were overproduced in *E. coli* BL21(DE3) and purified by Ni-affinity followed by desalting.

The genes of recombinant proteins LmbE, CcbV, LmbT, LmbC, LmbN, LmbN-PCP and CcbD were PCR amplified from *S. lincolnensis* or *S. caelestis* genomic DNA using primers with engineered NedI and HindIII restriction sites. The PCR-amplified gene fragments were purified, digested with NdeI and HindIII and ligated into a pET28a(+) vector (Novagen) that had been digested with the same enzymes. The resultant plasmids were used to transform *E. coli* BL21(DE3) for protein over-expression. The *E. coli* transformant cultures were incubated in Luria-Bertani (LB) medium containing 50 µg ml^{-1} kanamycin at 37 °C and 250 rpm until the cell density reached 0.6–0.8 at $OD_{600\,nm}$. To induce protein expression, IPTG (0.1 mM) was added to the cultures, which were further incubated at 16 °C for 40–48 h. The cells were harvested by centrifugation and stored at −80 °C before lysis. The thawed cells were re-suspended in lysis buffer containing 50 mM Tris-HCl (pH 7.6), 300 mM NaCl, 10 mM imidazole and 10% (v/v) glycerol. After disruption by a low-temperature ultra-high-pressure cell disrupter (FB-110X, Shanghai Litu Mechanical Equipment Engineering Co., Ltd, China or JN-02HC, JNBIO, China), the soluble fraction was collected, subjected to purification of each target protein by using a HisTrap FF column (GE Healthcare, USA) and then desalted using a PD-10 Desalting Column (GE Healthcare, USA) according to the manufacturers' protocols. The resulting proteins were concentrated and stored at −80 °C for *in vitro* assays. The purity of the proteins was examined by 12% SDS–PAGE analysis, and the concentration was determined by Bradford assay using bovine serum albumin (BSA) as the standard. For LmbN-PCP, HPLC-HR-ESI-Q-TOF MS (Agilent Technologies Inc., USA) analysis indicated that the recombinant protein purified from *E. coli* BL21(DE3) was fully phosphopantetheinylated into a holo form (m/z $[M + H]^+$ calculated 11,842.94, found 11,843.42).

Determination of the amidase activity of LmbE *in vitro*. The assays were carried out at 30 °C for 5 h in a 50 µl reaction mixture containing 50 mM Tris-HCl (pH 7.5) and 2 mM substrate (MSH *S*-conjugate) in the presence of 2 µM (for **1**) or 20 µM (for **6**) LmbE. Reactions in the absence of the enzyme were used as negative controls.

An equal volume of methanol was added into each mixture to terminate the reaction. After removal of the denatured protein by centrifugation, the reaction mixtures were subjected to HPLC-ESI-MS analysis on a Phenomenex Luna C18(2) column (5 μm, 4.6 × 250 mm, USA) by isocratic elution of 90% 10 mM NH_4Ac and 10% CH_3CN for 15 min period at a flow rate of 1 ml min^{-1}.

Characterization of the CcbV-catalysed reaction *in vitro*. The assays were carried out at 30 °C for 2 h in a 50 μl reaction mixture. For substrate **4**, the mixture contained 50 mM Tris-HCl (pH 8.0), 1 mM MSH, 2 mM TCEP, 1 mM EGT *S*-conjugate (**4**) and 40 μM CcbV. For substrate **5**, the mixture contained 50 mM Tris-HCl (pH 8.0), 2 mM MSH, 2 mM TCEP, 2 mM EGT *S*-conjugate (**5**) and 100 μM CcbV. Reactions in the absence of the enzyme were used as negative controls. The termination of each reaction and analysis of the resulting MSH *S*-conjugate and EGT were carried out according to the methods described above for LmbE-catalysed conversion. HPLC analysis of EGT production was performed on a COSMOSIL HILIC Packed Column (5 μm, 4.6 × 250 mm, Nacalai Tesque Inc., Japan) by isocratic elution of 30% 10 mM NH_4Ac and 70% CH_3CN for 20 min at a flow rate of 1 ml min^{-1} (mAU at 260 nM). The commercially available EGT (Enzo Life Sciences Inc., USA) was used as the standard.

To evaluate the pH dependence, each reaction was performed in triplicate at 30 °C for 1 h in a 25 μl reaction mixture containing 1 mM MSH, 2 mM TCEP, 1 mM **4** and 20 μM CcbV in 50 mM PIPES (pH 6.0–7.0) or 50 mM Tris-HCl (pH 7.5–9.0) buffer.

For the kinetic analysis, a time course was carried out to determine the initial rate conditions in a 25 μl of reaction mixture containing 50 mM Tris-HCl buffer (pH 8.0), 1 mM MSH, 1 mM TCEP, 1 mM **4** and 20 μM CcbV. The reactions were initiated by the addition of CcbV, incubated at 30 °C, and then terminated by 25 μl of methanol at 3, 5, 7, 10, 16, 25, 45 and 60 min. The samples were subjected to the same work-ups and HPLC analysis of EGT production as described above. The production of EGT, linear with respect to time during 0–20 min, was fitted into a linear equation to obtain the initial velocity. To determine the kinetic parameters for substrate **4**, the reactions were carried out at 30 °C for 20 min, each in 25 μl of mixture containing 50 mM Tris-HCl buffer (pH 8.0), 2 mM MSH, 2 mM TCEP, and 20 μM CcbV, and varying the concentration of substrate **4** (0.02, 0.05, 0.10, 0.20, 0.50, 1.00 and 2.00 mM). To determine the kinetic parameters for the substrate MSH, the reactions were carried out at 30 °C for 20 min, each in 25 μl of mixture containing 50 mM Tris-HCl buffer (pH 8.0), 5 mM **4**, 2 mM TCEP, and 20 μM CcbV, and varying the concentration of substrate MSH (0.02, 0.05, 0.10, 0.20, 0.50 and 1.00 mM). All assays were performed in triplicate, and each conversion was analysed by HPLC for EGT production as described above. The resulting initial velocities were then fitted to the Michaelis–Menten equation using GraphPad Prism5 software (GraphPad Software, Inc., USA) to extract the K_m and k_{cat} parameters.

Characterization of the LmbT-catalysed reaction *in vitro*. The reverse glycosylation assays of LmbT were carried out at 30 °C for 2 h in a 50 μl reaction mixture. For substrate **5**, the mixture contained 50 mM Tris-HCl (pH 7.5), 2 mM $MgCl_2$, 1 mM **5**, 2 mM GDP and 4 μM LmbT. For substrate **4**, the mixture contained 50 mM Tris-HCl (pH 7.5), 2 mM $MgCl_2$, 2 mM **4**, 4 mM GDP and 20 μM LmbT. The forward glycosylation assays were carried out at 30 °C for 2 h in a 50 μl reaction mixture containing 50 mM Tris-HCl (pH 7.5), 2 mM $MgCl_2$, 1 mM **3**, 1 mM EGT and 4 μM LmbT. Reactions in the absence of the enzyme were used as negative controls. The termination of each reaction was carried out according to the methods described above for LmbE-catalysed conversion. The reaction mixtures containing **5** were subjected to HPLC-ESI-MS analysis on a Phenomenex Luna C18(2) column by isocratic elution of 95% 10 mM NH_4Ac and 5% CH_3CN for 15 min at a flow rate of 1 ml min^{-1}.

To evaluate the pH dependence, each reaction was performed in triplicate at 30 °C for 1 h in 25 μl of reaction mixture containing 2 mM $MgCl_2$, 4 mM GDP, 2 mM **5** and 4 μM LmbT in 50 mM PIPES (pH 6.0–7.0) or 50 mM Tris-HCl (pH 7.5–9.0) buffer.

For the kinetic analysis, a time course was carried out to determine the initial rate conditions in 25 μl of reaction mixture containing 50 mM Tris-HCl (pH 7.5), 2 mM $MgCl_2$, 1 mM **5**, 2 mM GDP and 4 μM LmbT. The reactions were initiated by the addition of LmbT, incubated at 30 °C, and then terminated by 25 μl of methanol at 3, 5, 7, 10, 16, 25, 45 and 60 min. The samples were subjected to the same work-ups and HPLC analysis of EGT production as described above. The production of EGT, linear with respect to time during 0–20 min, was fitted into a linear equation to obtain the initial velocity. To determine the kinetic parameters for substrate **5**, the reactions were carried out at 30 °C for 20 min, each in 25 μl of mixture containing 50 mM Tris-HCl buffer (pH 7.5), 2 mM $MgCl_2$, 2 mM GDP, and 4 μM LmbT, and varying the concentration of substrate **5** (0.02, 0.05, 0.10, 0.20, 0.50, 1.00, 2.00 and 4.00 mM). To determine the kinetic parameters for substrate GDP, the reactions were carried out at 30 °C for 20 min, each in 25 μλ of mixture containing 50 mM Tris-HCl buffer (pH 7.5), 2 mM $MgCl_2$, 4 mM **5**, and 4 μM LmbT, and varying the concentration of substrate GDP (0.02, 0.05, 0.10, 0.20, 0.50, 1.00 and 2.00 mM). All assays were performed in triplicate, and each conversion was analysed by HPLC for EGT production as described above. The resulting initial velocities were then fitted to the Michaelis–Menten equation using GraphPad Prism5 software (GraphPad Software, Inc., USA) to extract the K_m and k_{cat} parameters.

To determine the equilibrium constant (K_{eq}) of the LmbT-catalysed reaction, the experiment was performed according to the method described previously[34]. K_{eq} was measured by performing a series of saturated reactions, in which the concentration ratio of [GDP]/[**3**] varied from 1/3 to 5 on the premise that the ratio of [**5**]/[EGT] was fixed at 1. The total concentrations of [**3**] + [GDP] and [**5**] + [EGT] were both kept at 4 mM. The reaction was performed in a 25 μl mixture containing 50 mM Tris-HCl (pH 7.5), 2 mM $MgCl_2$ and 4 μM LmbT at 30 °C for 3 h to reach equilibrium. The change of [EGT] was monitored by HPLC as described previously and plotted against the ratio of [GDP]/[**3**]. The equilibrium constant was subsequently determined from the equation $K_{eq} = ([GDP]/[\mathbf{3}]) \times ([\mathbf{5}]/[EGT])$.

Characterization of PPL incorporation *in vitro*. To convert holo-LmbN-PCP into PPL-acylated LmbN-PCP, the reaction was carried out at 30 °C for 3 h in a 50 μl reaction mixture containing 50 mM Tris-HCl (pH 7.5), 2 μM LmbC, 2 mM PPL, 5 mM ATP, 10 mM $MgCl_2$, 2 mM TCEP, and 1 mM CoA in the presence of 100 μM LmbN-PCP. Reactions in the absence of LmbC or ATP were used as negative controls. For product examination, the reaction mixtures were subjected to HPLC-HR-ESI-Q-TOF MS (Agilent Technologies Inc., USA) analysis on an Agilent Zorbax column by gradient elution of solvent A (H_2O containing 0.1% formic acid) and solvent B (CH_3CN containing 0.1% formic acid) at a flow rate of 0.2 ml min^{-1} over a 45-min period as follows: $t = 0$ min, 10% B; and $t = 45$ min, 95% B (mAU at 210 nm).

For PPL incorporation, the assays were carried out at 30 °C for 3 h in a 50 μl reaction mixture containing 50 mM Tris-HCl (pH 7.5), 2 μM LmbC, 2 mM PPL, 5 mM ATP, 10 mM $MgCl_2$, 2 mM TCEP, 1 mM CoA, and 2 mM substrate **5** (MSH *S*-conjugate) in the presence of 50 μM LmbN or LmbN-PCP and 10 μM CcbD. Reactions in the absence of each enzyme were used as negative controls. The termination of each reaction and HPLC-ESI-MS analysis was carried out according to the methods described above for LmbE-catalysed conversion.

Compound isolation and purification. For compounds **1** and **4** (from the *ΔlmbE* and *ΔlmbV* mutant strain, respectively), 100 g of Amberlite XAD-2 resin (Rohm and Haas Co., USA) was incubated with 1 l of each fermentation culture broth overnight to absorb the target compound. After filtration, the resin was backwashed with 2 l of water, and then eluted with 3 l of MeOH. The eluent was evaporated *in vacuo* to a crude extract. The resultant residue was re-dissolved in 5 ml of water, and then loaded onto a Sephadex LH20 column (3.5 × 200 cm, GE Healthcare, USA) by eluting 500 ml of MeOH at a flow rate of 0.5 ml min^{-1}. According to ESI-MS analysis, the fractions containing the target compound were combined, evaporated *in vacuo* and then loaded onto an Agilent Zorbax column (SB-C18, 5 μm, 9.4 × 250 mm, Agilent Technologies Inc., USA) by isocratic elution of 40% 5 mM NH_4Ac and 60% MeOH for 12 min at a flow rate of 2 ml min^{-1} (mAU 210 nm). After a similar work-up for fractionation and concentration, further purification was carried out on a COSMOSIL HILIC Packed Column (5 μm, 10 × 250 mm, Nacalai Tesque Inc., Japan) by isocratic elution of 21% 10 mM NH_4Ac and 79% CH_3CN for 70 min at a flow rate of 2 ml min^{-1} (mAU 210 nm).

For PPL (from the *ΔlmbT* mutant strain), after a similar procedure with Amberlite XAD-2 resin, further HPLC semi-preparation purification was carried out twice on an Agilent Zorbax column by isocratic elution of 20% 5 mM NH_4Ac and 80% MeOH for 30 min at a flow rate of 2 ml min^{-1} (mAU at 210 nm).

For compounds **5** and **6** (from the *ΔlmbC* mutant strain), 100 g of Amberlite XAD-16 resin (Rohm and Haas Co., USA) was incubated with 1 l of fermentation culture broth overnight to remove most of the impurity. After filtration and concentration, the resultant residue was then loaded onto a Sephadex LH20 column by eluting 500 ml of H_2O:MeOH (1:1) at a flow rate of 0.5 ml min^{-1}. According to ESI-MS analysis, the fractions containing compound **5** or **6** were combined, evaporated *in vacuo* and then loaded onto a COSMOSIL HILIC Packed Column by isocratic elution of 30% 10 mM NH_4Ac and 70% CH_3CN for 30 min at a flow rate of 2 ml min^{-1} (mAU at 210 nm). For compound **6**, further purification was carried out on a Sephadex G10 column (1.5 × 120 cm, GE Healthcare, USA) by eluting H_2O at a flow rate of 0.2 ml min^{-1}.

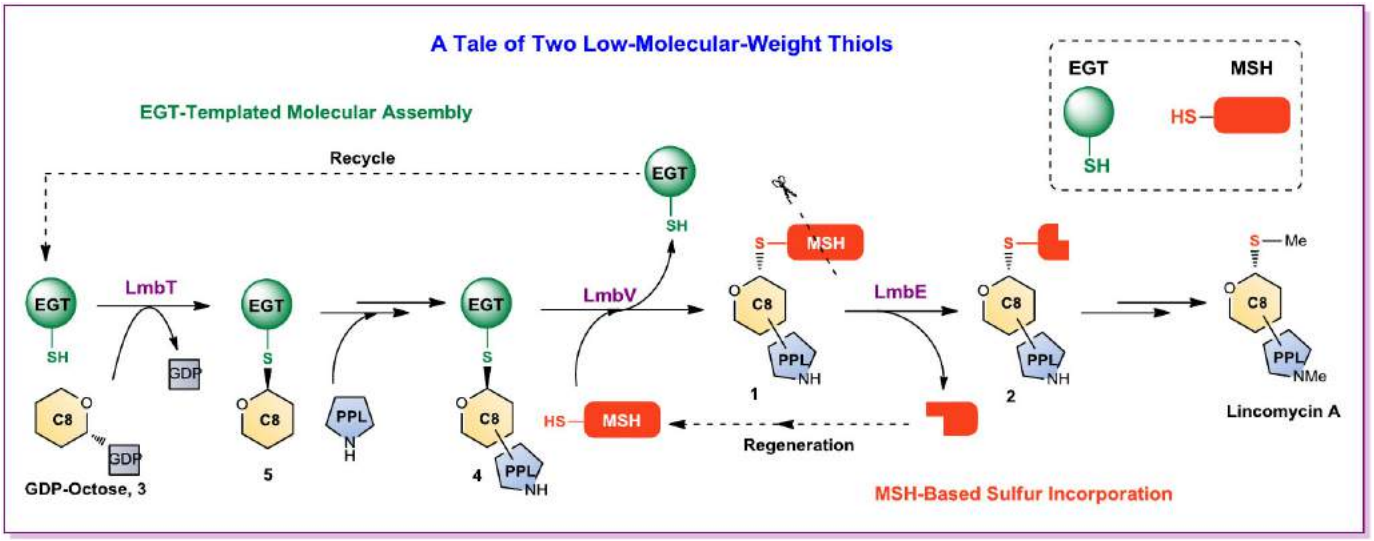
For compounds **2** and GlcN-Ins (from the LmbE-catalysed reaction), the assay was scaled up and carried out at 30 °C for 2 h in 20 ml of mixture containing 200 μM LmbE, 1 mM pure compound **1** and 50 mM Tris-HCl buffer (pH 7.5). After filtration with an ultra-filtration membrane (Amicon YM-30, Millipore) to remove protein, the solution was loaded onto an Agilent Zorbax column by gradient elution of solvent A (H_2O) and solvent B (CH_3CN) at a flow rate of 2 ml min^{-1} over a 18-min period as follows: $t = 0$ min, 5% B; $t = 5$ min, 5% B; $t = 10$ min, 55% B; $t = 16$ min, 55% B; and $t = 18$ min, 5% B (mAU at 210 nm). The fraction containing GlcN-Ins was concentrated and then loaded onto a COSMOSIL HILIC Packed Column by

isocratic elution of 25% 10 mM NH_4Ac and 75% CH_3CN for 40 min at a flow rate of 2 ml min^{-1}. GlcN-Ins was examined using a refractive index (RI) detector. For compound **2**, further purification was carried out on the same column by the same isocratic elution for 60 min (mAU at 210 nm).

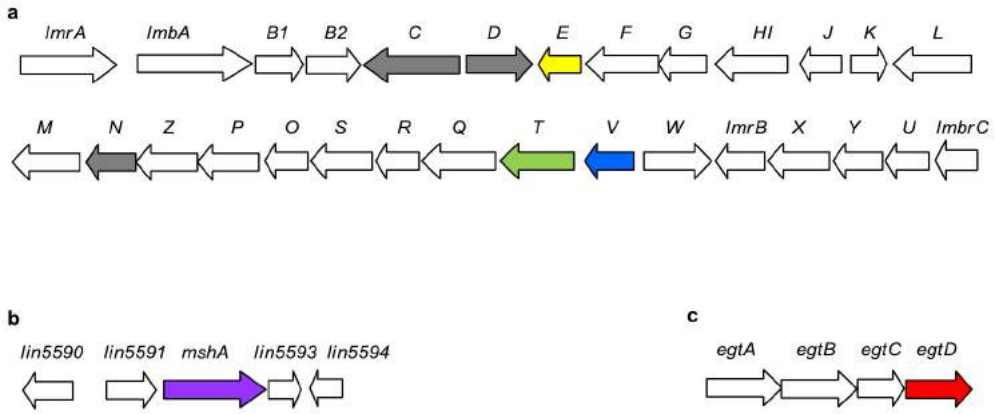
For compound **3** (from the LmbT-catalysed reverse glycosylation reaction), the assay was scaled up and carried out at 30 °C for 5 h in 10 ml of mixture containing 4 μM LmbT, 2 mM pure compound **5**, 2 mM GDP, 1 mM $MgCl_2$ and 50 mM Tris-HCl buffer (pH 7.5). After filtration with an ultra-filtration membrane to remove protein, the solution was concentrated and then subjected to a Sephadex G10 column by eluting H_2O at a flow rate of 0.1 ml min^{-1}.

Chemical synthesis of MSH. The synthesis was carried out according to the procedures described previously[35,36] with slight modifications. For details, see Supplementary Methods.

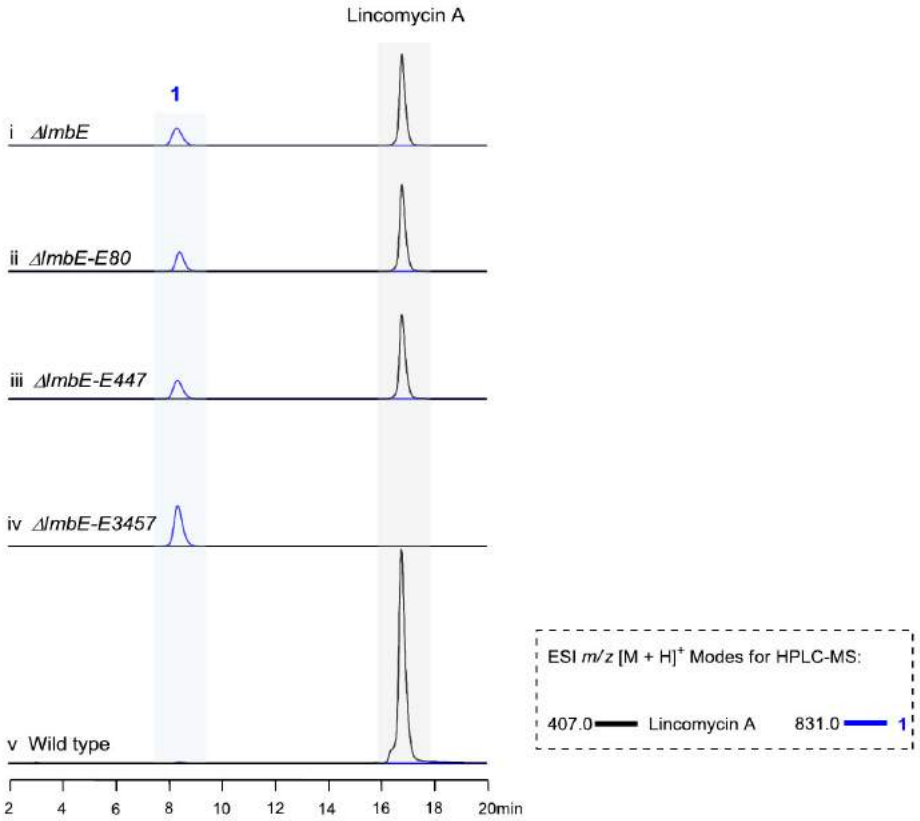
31. Kieser, T. *et al. Practical Streptomyces Genetics* (John Innes Foundation, 2000).
32. Green, M. R. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* 4th edn (Cold Spring Harbor Laboratory Press, 2012).
33. Fahey, R. C. & Newton, G. L. Determination of low molecular weight thiols using monobromobimane fluorescent labeling and high-performance liquid chromatography. *Methods Enzymol.* **143,** 85–96 (1987).
34. Zhang, C. *et al.* Exploiting the reversibility of natural product glycosyltransferase-catalyzed reactions. *Science* **313,** 1291–1294 (2006).
35. Lee, S. & Rossaza, J. P. First total synthesis of mycothiol and mycothiol disulfide. *Org. Lett.* **6,** 365–368 (2004).
36. Ajayi, K., Thakur, V. V., Lapo, R. C. & Knapp, S. Intramolecular α-glucosaminidation: synthesis of mycothiol. *Org. Lett.* **12,** 2630–2633 (2010).



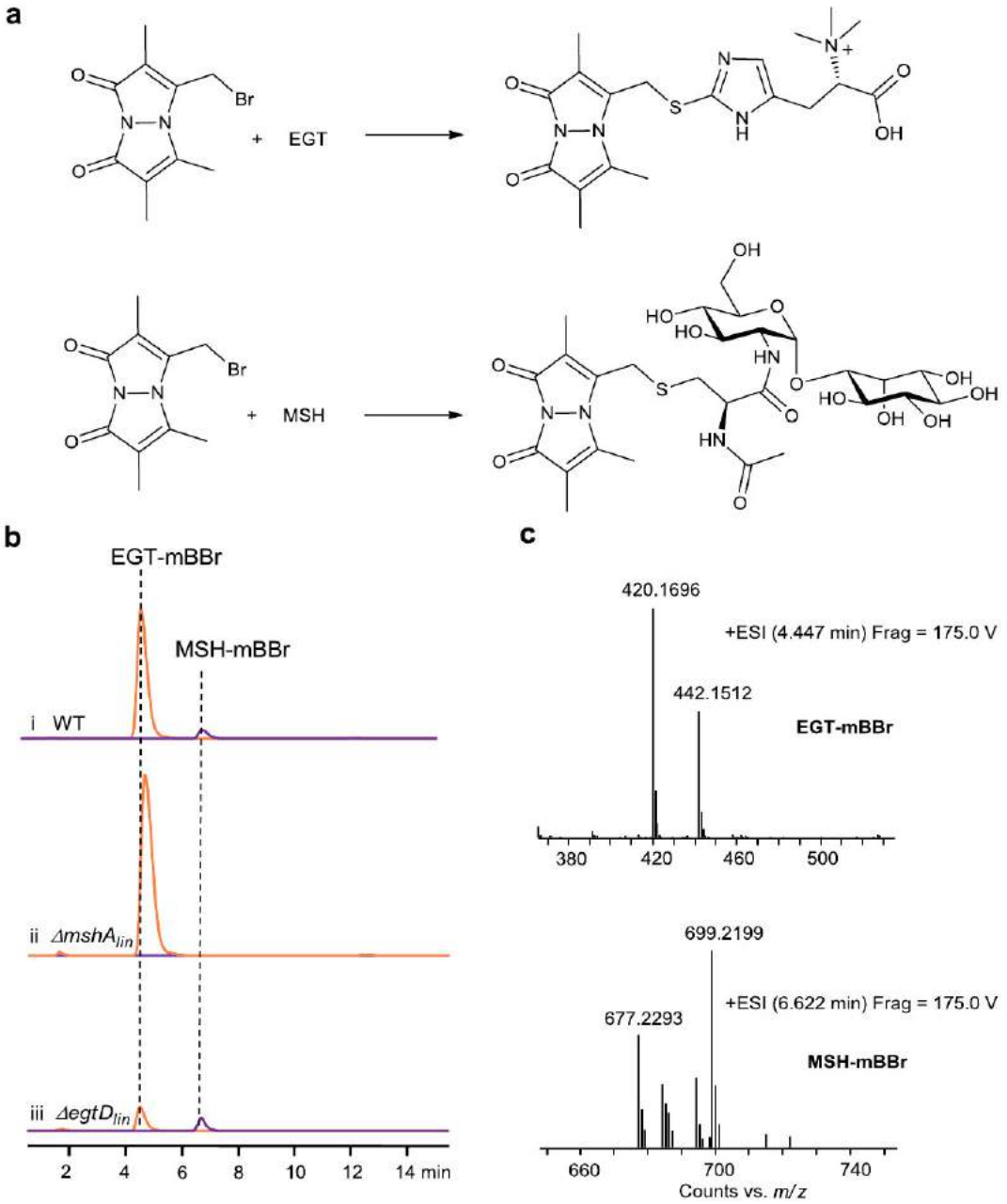
Extended Data Figure 1 | Constructive role of MSH (orange) and EGT (green) in lincomycin biosynthesis (shown as cartoon models).



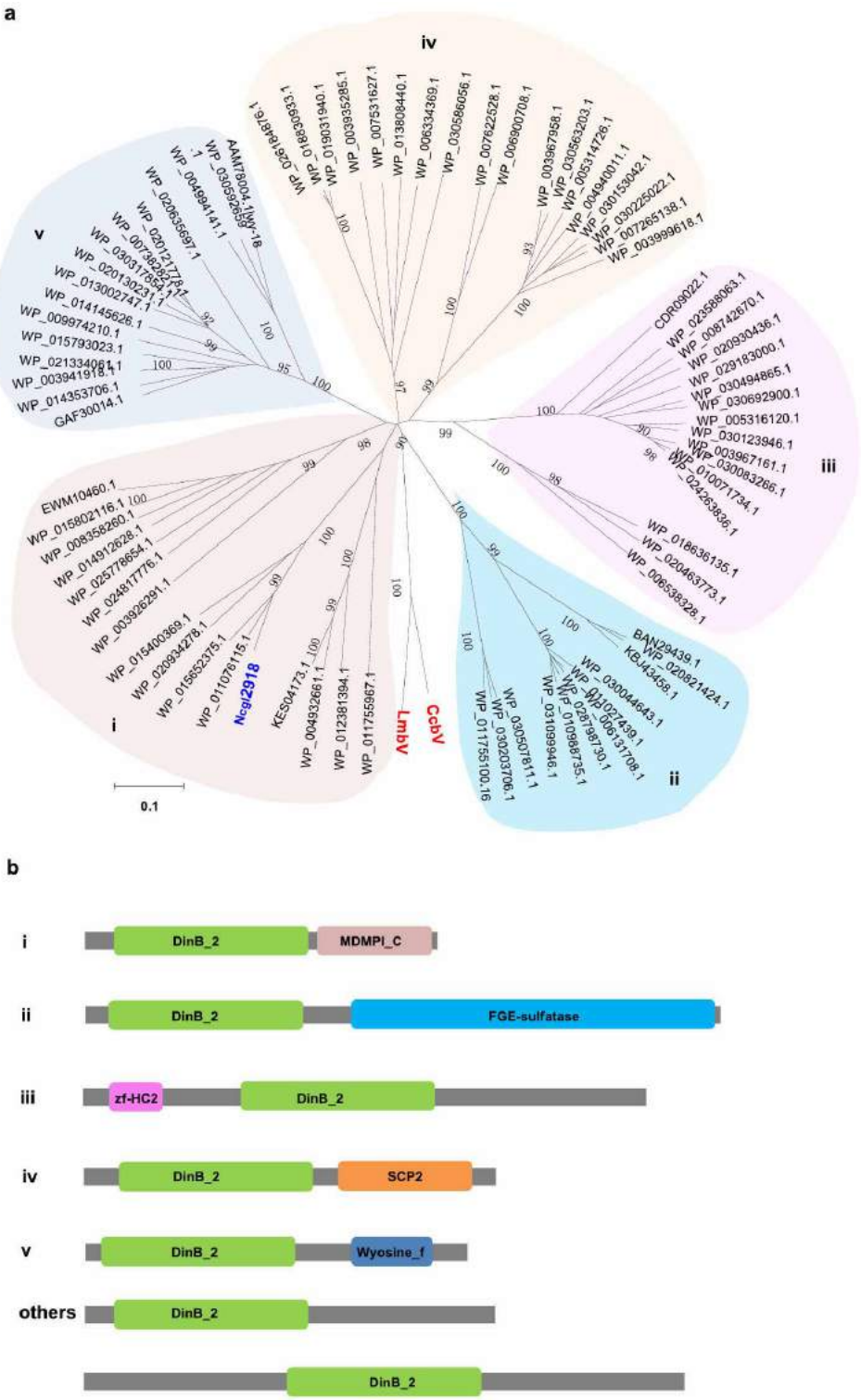
Extended Data Figure 2 | Genes and/or clusters relevant to lincomycin biosynthesis in *S. lincolnensis*. **a**, Biosynthetic gene cluster of lincomycin A. The *mca* homologue *lmbE* and two GTase-encoding genes, *lmbV* and *lmbT*, are shown in yellow, blue and green, respectively. The genes responsible for PPL incorporation (*lmbC*, *lmbD* and *lmbN*) are shown in grey. **b**, The location of $mshA_{lin}$ (shown in purple), which is not clustered with the other genes responsible for MSH biosynthesis, in the genome of *S. lincolnensis*. The flanking genes *lin5590*, *lin5591*, *lin5593* and *lin5594* share sequence homology with genes encoding the methylase (WP_026151264.1) from *S. prunicolor*, the DUF899-like protein (WP_004004056.1) from *S. viridochromogenes*, the YbjN-like protein (WP_020130332.1) from *Streptomyces* sp. 303MFCol5.2 and the hypothetical protein (WP_004000160.1) from *S. viridochromogenes*, respectively. **c**, The biosynthetic gene cluster of EGT. The gene $egtD_{lin}$ is shown in red.



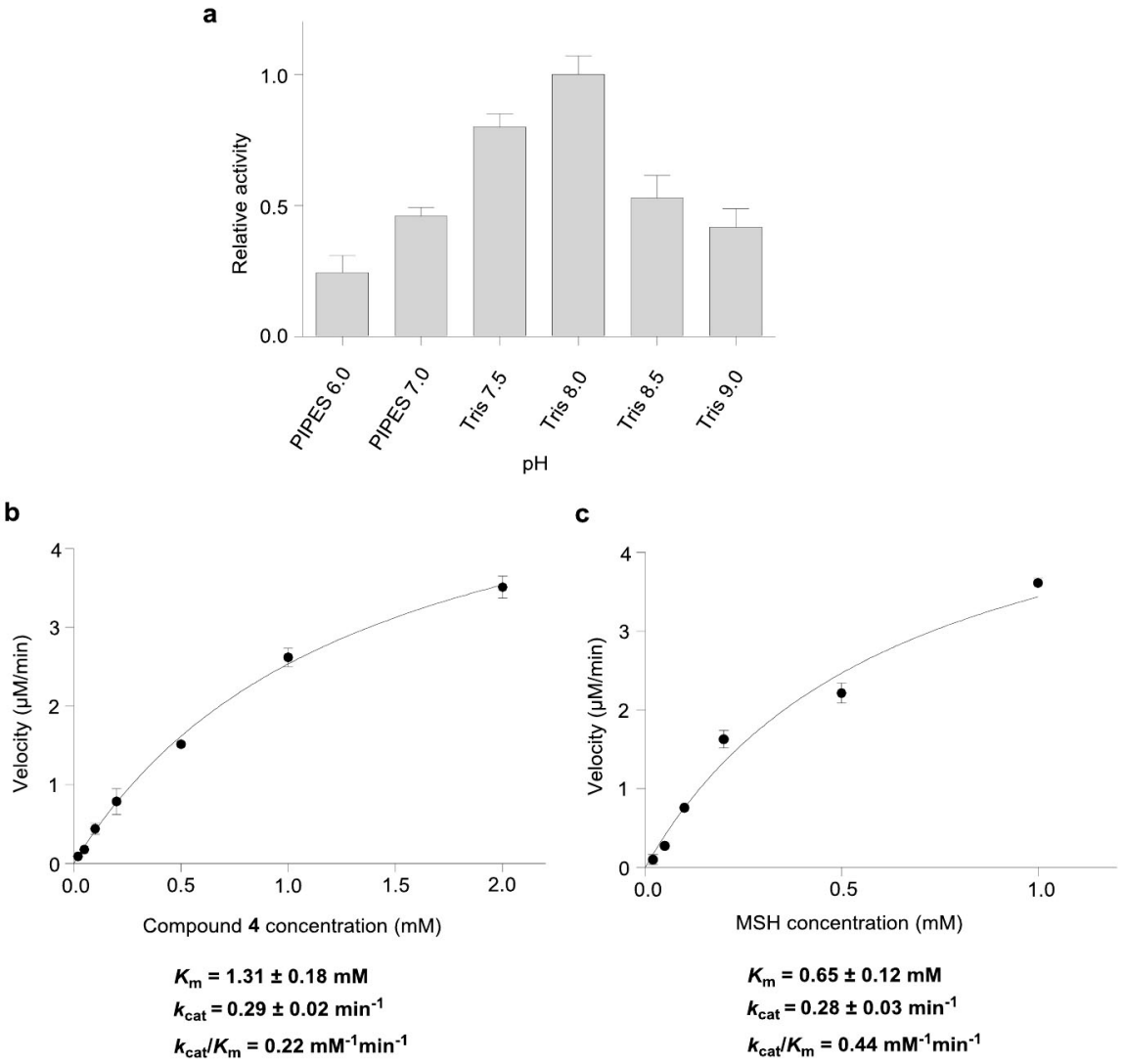
Extended Data Figure 3 | The production of MSH *S*-conjugate 1 and lincomycin A in various *S. lincolnensis* strains. The strains include the mutants (i, for *ΔlmbE*; ii, for *ΔlmbE-E80*; iii, for *ΔlmbE-E447*; and iv, for *ΔlmbE-E3457*) and the wild-type control (v). For HPLC-MS analysis, the ESI *m/z* $[M + H]^+$ modes are indicated in the dashed rectangle.



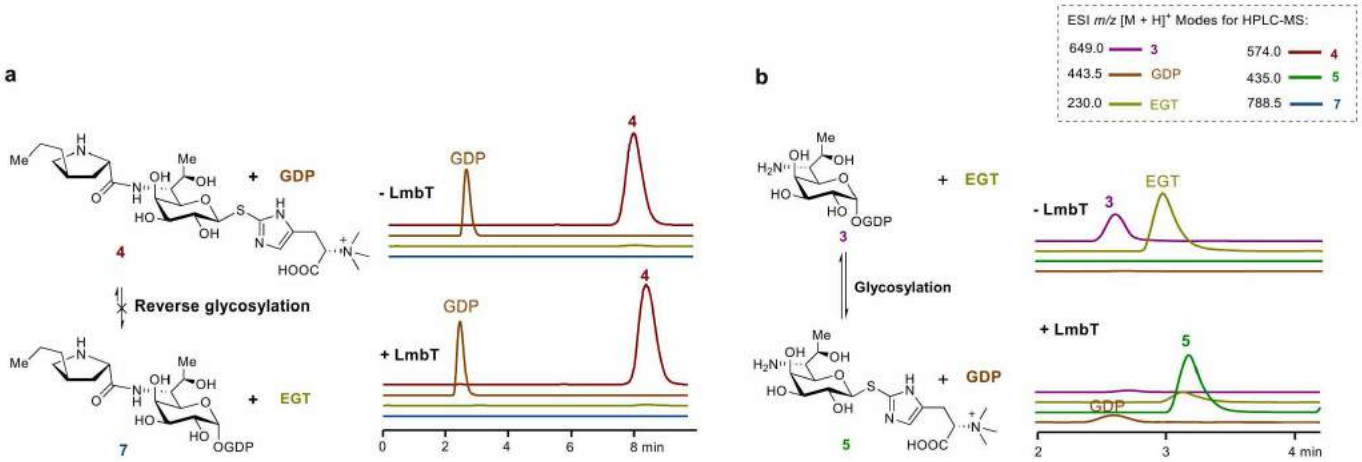
Extended Data Figure 4 | Analysis of MSH and EGT production in the *ΔmshA$_{lin}$* and *ΔegtD$_{lin}$* mutant *S. lincolnensis* strains. **a**, Derivatization of MSH and EGT with mBBr to generate the corresponding *S*-conjugates. **b**, HPLC-HR-MS analysis of MSH-mBBr (calculated for $C_{27}H_{41}N_4O_{14}S^+$ $[M + H]^+$ 677.2340) and EGT-mBBr (calculated for $C_{19}H_{26}N_5O_4S^+$ $[M + H]^+$ 420.1706) in the wild-type control (i), LL1005 (ii, *ΔmshA$_{lin}$* mutant) and LL1010 (iii, *ΔegtD$_{lin}$* mutant). **c**, The HR-ESI-MS spectra of EGT-mBBr (top) and MSH-mBBr (below).



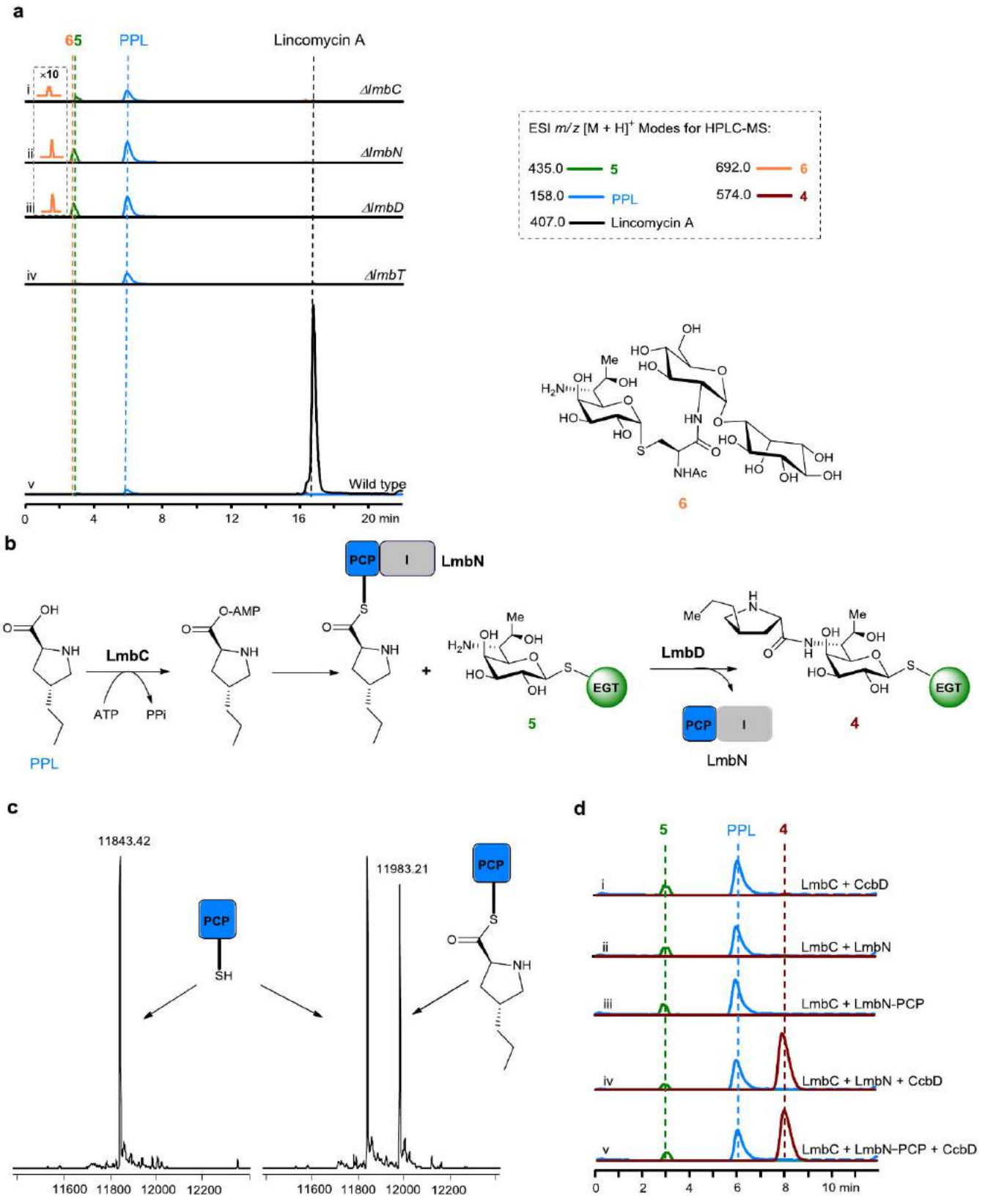
Extended Data Figure 5 | Phylogenetic analysis in the DinB-2 superfamily. **a**, LmbV (from *S. lincolnensis*) and CcbV (from *S. caelestis*), shown in red, with selected DinB-2-like proteins (from various actinomycetes, in which the thiol MSH is dominant) in the phylogenetic tree. The evolutionary distances were computed using the p-distance method. The support for grouping clades i, ii, iii, iv and v (shaded in different colours) is indicated by bootstrap values. The known MSH-maleylpyruvate isomerase Ncgl2918 is shown in blue. **b**, Typical domain organization of the DinB-2-like proteins. The conserved DinB-2 domain is shown in green. Clade i features the C-terminal MDMPI-C domain. Proteins containing this domain include the MSH-maleylpyruvate isomerases, such as Ncgl2918. Clade ii features the C-terminal FGE-sulfatase domain, which is found in eukaryotic proteins required for post-translational modification to produce sulfatases, which are essential for the degradation and remodelling of sulfate esters. Clade iii features the N-terminal zf-HC2 domain, which contains a putative zinc-finger binding motif and is found in some anti-sigma factor proteins. Clade iv features the C-terminal SCP2 domain involved in binding sterols. Clade v features the *C*-terminal wyosine_f domain. Some proteins containing this domain appear to be important in wyosine base formation in a subset of phenylalanine-specific tRNAs. 'Others' indicate a number of DinB-2-like proteins that possess an unknown domain(s) either at the C terminus or at both the C and N termini of the proteins.



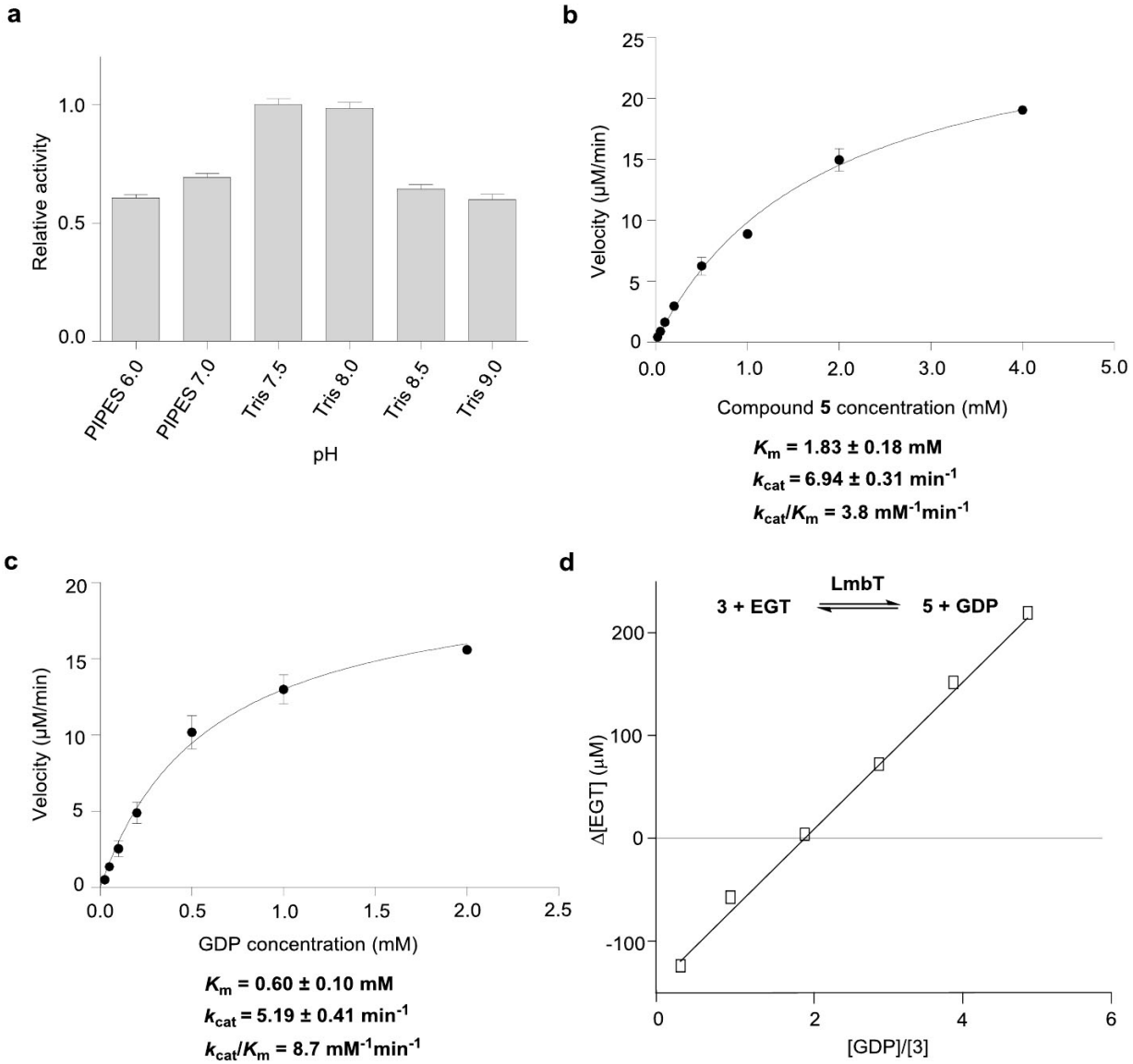
Extended Data Figure 6 | Kinetic analysis of CcbV-catalysed thiol exchange. **a**, pH dependence. The activity of CcbV in 50 mM PIPES (pH 6.0-7.0) or 50 mM Tris-HCl (pH 7.5–9.0) buffer was measured. **b**, **c**, Determination of the steady-state kinetic parameters for substrate **4** and for MSH, respectively. The error bars are standard error of mean ($n = 3$).



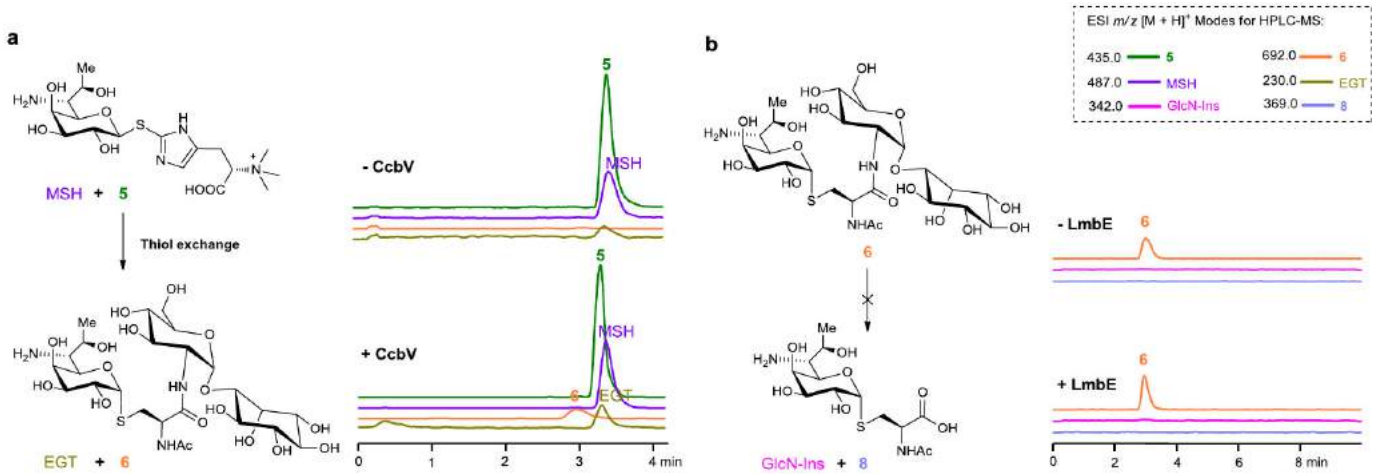
Extended Data Figure 7 | Characterization of LmbT-catalysed reverse and forward glycosylation. For HPLC-MS analysis, the ESI *m/z* $[M + H]^+$ modes are indicated in dashed rectangle. **a**, Examination of the acylated C8-sugar transfer in the presence of LmbT, which showed that LmbT was unable to utilize **4** as a substrate for reverse glycosylation to generate the predicted GDP-D-α-D-sugar **7** (left) in the absence (top right) and in the presence (lower right) of LmbT. **b**, Characterization of LmbT-catalysed forward glycosylation. LmbT used **3** as a substrate for glycosylation to generate the GDP-D-α-D-sugar **5** (left) in the absence (top right) and in the presence (lower right) of LmbT.



Extended Data Figure 8 | Characterization of PPL incorporation in lincomycin A biosynthesis. For HPLC-MS analysis, the ESI *m/z* $[M + H]^+$ modes are indicated in the dashed rectangle. **a**, *In vivo* product profiles of *S. lincolnensis* strains, including the mutants (i, for *ΔlmbC*; ii, for *ΔlmbN*; iii, for *ΔlmbD*; and iv, for *ΔlmbT*) and the wild-type control (v). **b**, Process of the incorporation of PPL (with EGT *S*-conjugate **5**) into intermediate **4**. PCP (blue), peptidyl carrier protein; and I (grey), isomerase. **c**, HPLC-MS examination of LmbC-catalysed conversion of holo-LmbN-PCP (*m/z* $[M + H]^+$ calculated 11,842.94, found 11,843.42) into PPL-acylated LmbN-PCP (*m/z* $[M + H]^+$ calculated 11,982.01, found 11,983.21) in the absence (left) and in the presence (right) of ATP. **d**, *In vitro* analysis of the condensation between PPL and **5** to generate **4**. The catalyst systems included LmbC + CcbD (i), LmbC + LmbN (ii), LmbC + LmbN-PCP (iii), LmbC + LmbN + CcbD (iv), and LmbC + LmbN-PCP + CcbD (v).



Extended Data Figure 9 | Kinetic analysis of LmbT-catalysed reversible glycosylation. **a**, pH dependence. The activity of LmbT in 50 mM PIPES (pH 6.0–7.0) or 50 mM Tris-HCl (pH 7.5–9.0) buffer was measured. **b**, **c**, Determination of the steady-state kinetic parameters for substrate **5** and for GDP, respectively, in LmbT-catalysed reverse glycosylation. The error bars are standard error of mean ($n = 3$). **d**, Determination of the equilibrium constant (K_{eq}) of LmbT-catalysed glycosylation. $K_{eq} = ([GDP]/[\mathbf{3}]) \times ([\mathbf{5}]/[EGT]) = 1.94 \times 1 = 1.94$.



Extended Data Figure 10 | Validation of the highly ordered process involving EGT-mediated assembly and MSH-associated post-modifications in lincomycin biosynthesis. For HPLC-MS analysis, the ESI *m/z* $[M + H]^+$ modes are indicated in the dashed rectangle. **a**, Examination of the thiol exchange using **5** as a substrate in the absence (top right) and in the presence (bottom right) of CcbV, showing that CcbV was able to convert **5** to MSH *S*-conjugate **6** (left). **b**, Determination of the hydrolysis reaction in the absence (top right) and in the presence (bottom right) of LmbE, showing that this enzyme was unable to utilize **6** as a substrate and convert it to the predicted mercapturic acid **8** (left).