







Please cite the Published Version

Xu, G , Torri, D , Cuesta-Hoyos, S, Panda, D, Yates, LRL , Zallot, R , Bian, K, Jia, D, Iorgu, AI, Levy, C, Shepherd, SA  and Micklefield, J  (2024) Cryptic enzymatic assembly of peptides armed with -lactone warheads. *Nature Chemical Biology*, 20 (10). pp. 1371-1379. ISSN 1552-4450

DOI: <https://doi.org/10.1038/s41589-024-01657-7>

Publisher: Springer

Version: Published Version

Downloaded from: <https://e-space.mmu.ac.uk/636700/>

Usage rights:  [Creative Commons: Attribution 4.0](https://creativecommons.org/licenses/by/4.0/)

Additional Information: The version of record of this article, first published in *Nature Chemical Biology*, is available online at Publisher's website: <http://dx.doi.org/10.1038/s41589-024-01657-7>

Data Access Statement: The coordinates of the CysF X-ray crystal structures were deposited to PDB 8RA0. The AlphaFold structure of BhCysFE can be accessed at <https://www.alphafold.ebi.ac.uk/entry/A0A562R406> (AF-A0A562R406-F1-model_v4). Structures used for modeling and docking studies can be accessed from PDB 5BSM, 5BSR, 5WM3, 5IE3, 4FUT, 4GXR and 4GXQ. All proteins characterized in this study can be accessed from UniProt using the accession codes presented in Supplementary Tables 1–3 and their synthetic gene sequences are provided in Supplementary Data 1. The remaining data are available in the main text or the Supplementary Information. Correspondence and requests for materials should be addressed to J.M.

Enquiries:

If you have questions about this document, contact openresearch@mmu.ac.uk. Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)

Cryptic enzymatic assembly of peptides armed with β -lactone warheads

Received: 13 December 2023

Accepted: 29 May 2024

Published online: 1 July 2024

Check for updates

Guangcai Xu , Daniele Torri , Sebastian Cuesta-Hoyos, Deepanjan Panda, Luke R. L. Yates , Rémi Zallot, Kehan Bian, Dongxu Jia, Andreea I. Iorgu, Colin Levy, Sarah A. Shepherd & Jason Micklefield

Nature has evolved biosynthetic pathways to molecules possessing reactive warheads that inspired the development of many therapeutic agents, including penicillin antibiotics. Peptides armed with electrophilic warheads have proven to be particularly effective covalent inhibitors, providing essential antimicrobial, antiviral and anticancer agents. Here we provide a full characterization of the pathways that nature deploys to assemble peptides with β -lactone warheads, which are potent proteasome inhibitors with promising anticancer activity. Warhead assembly involves a three-step cryptic methylation sequence, which is likely required to reduce unfavorable electrostatic interactions during the sterically demanding β -lactonization. Amide-bond synthetase and adenosine triphosphate (ATP)-grasp enzymes couple amino acids to the β -lactone warhead, generating the bioactive peptide products. After reconstituting the entire pathway to β -lactone peptides *in vitro*, we go on to deliver a diverse range of analogs through enzymatic cascade reactions. Our approach is more efficient and cleaner than the synthetic methods currently used to produce clinically important warhead-containing peptides.

Covalent inhibitors with reactive functionality (warheads) that bond irreversibly with a biological target have been widely used as therapeutic agents and are becoming increasingly important in modern drug discovery^{1–6}. The design of this class of inhibitors is often inspired by natural products^{7–10}. For example, the discovery of penicillin, over 90 years ago, led to many semisynthetic antibiotics with reactive β -lactam rings, which remain the most widely prescribed class of antibiotics today (Fig. 1a)^{11,12}. Peptide natural products with epoxyketone warheads inspired the discovery of carfilzomib, a proteasome inhibitor that was approved for the treatment of cancer¹⁰ and provided payloads for antibody–drug conjugates¹³. Peptides armed with aldehydes, vinyl sulfones, nitriles and other electrophilic warheads have also provided many other antimicrobial, anticancer and antiviral agents^{1–10,14}, including nirmatrelvir, which was recently approved to treat COVID-19 (ref. 6). Most of these important therapeutic agents are currently produced by chemical synthesis. The chemical synthesis of peptides, particularly

those containing reactive functionality, can be challenging. Although peptide synthesis works well on a small scale, multistep processes using an excess of deleterious reagents, atom-inefficient protecting groups and large volumes of solvents are problematic to implement at the scale required for peptide therapeutics^{15–17}. By virtue of their benign operating conditions, selectivity and evolvability¹⁸, enzymes could provide more efficient and sustainable routes, making these medically valuable agents more widely available. Exploring the biosynthesis of warhead-containing peptide natural products could, therefore, provide important new enzymes that can be used to produce peptide therapeutics.

Several natural products have been discovered with β -lactone warheads¹⁹, including antiobesity drug lipstatin²⁰ and the proteasome inhibitor salinosporamide, which is in clinical trials for the treatment of cancer²¹. β -Lactone-containing peptides cystargolide (cys) and belactosin (bel) have also been isolated from *Kitasatospora cystarginea* and *Streptomyces* sp. UCK14, respectively^{22,23}. The promising antitumor

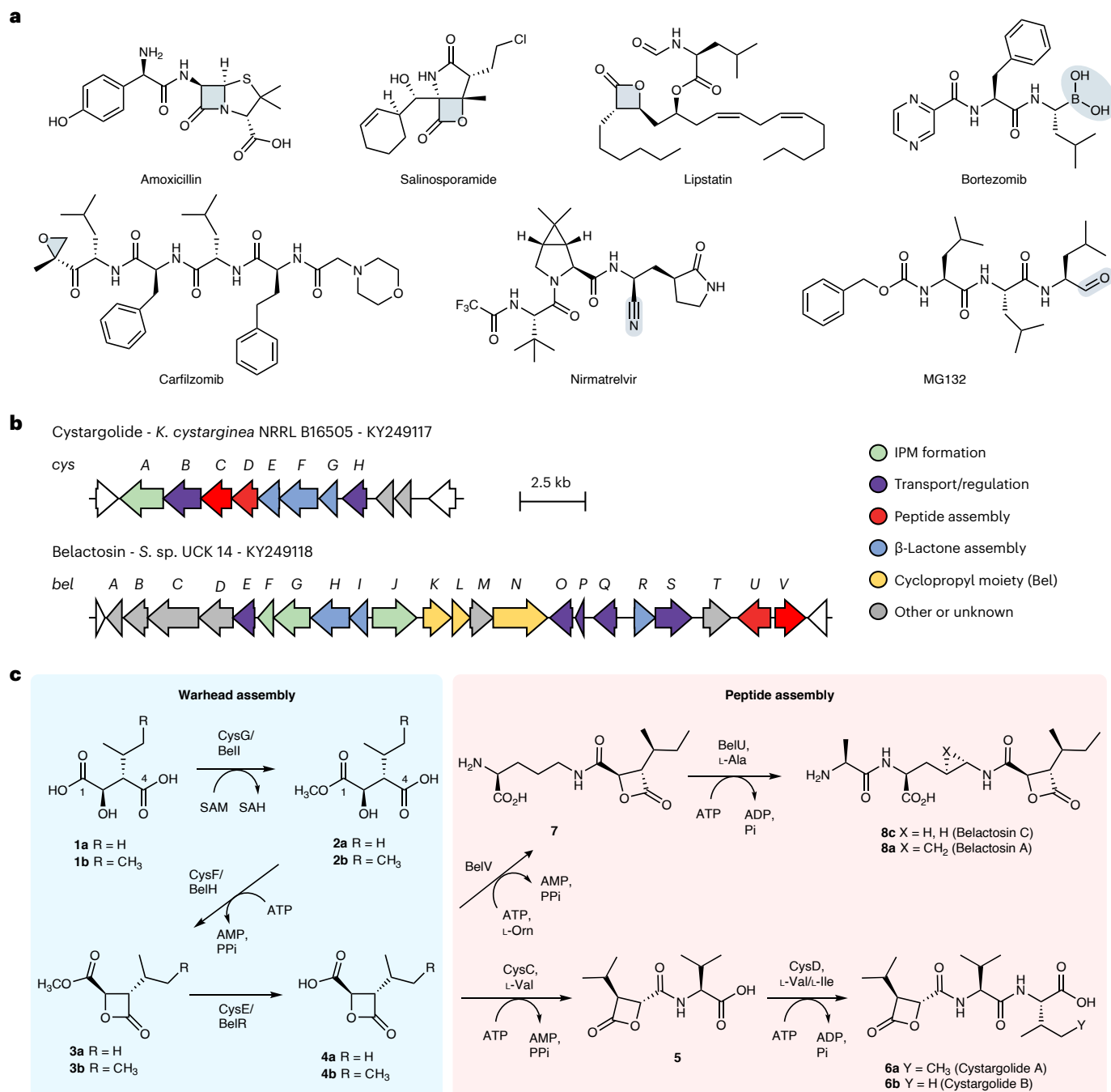


Fig. 1 | Warhead-containing therapeutic agents and biosynthesis of β -lactone-containing peptides. **a**, Examples of natural products and synthetic derivatives bearing electrophilic warheads. **b**, The *cys* and *bel* BGCs, including revisions based on this work. IPM, isopropyl malic acid. **c**, The biosynthetic

pathways to *cys* and *bel* based on the in vitro characterization of the complete set of ten pathway enzymes CysGFEC and BellHRVU, as described in this study. AMP, adenosine monophosphate; ADP, adenosine diphosphate; SAH, S-adenosylhomocysteine.

activity of *cys* and *bel* stimulated extensive medicinal chemistry efforts leading to the synthesis of improved derivatives^{24–31}, including *cys* analogs that were more potent proteasome inhibitors than carfilzomib when tested in vitro^{29,30}. The *cys* and *bel* biosynthetic gene clusters (BGCs) have been sequenced and the putative functions of pathway enzymes were proposed on the basis of bioinformatics analysis (Fig. 1b)³². The biosynthetic origin of the malic acid and cyclopropyl ornithine precursors of *cys* and *bel* have also been explored^{32–36}. However, the enzymes required for assembly of the β -lactone-containing peptides have not been characterized.

In this paper, we provide complete in vitro characterization of the entire set of ten enzymes required for *cys* and *bel* biosynthesis,

revealing cryptic pathways to reactive β -lactone intermediates and unanticipated timing of peptide-bond-forming steps (Fig. 1c). Our experiments show that the functions of all *cys* and *bel* biosynthetic enzymes differ from those predicted in silico, illustrating the unpredictable and unusual nature of the biosynthetic pathways. We also demonstrate how the *cys* and *bel* enzymes can be combined in cascade reactions to produce many peptides with different warheads.

Results

CysF catalyzes β -lactone formation

The presence of orthologs of the well-known primary metabolic enzyme isopropyl malic acid synthetase (IPMS) within the *cys* and *bel* BGCs

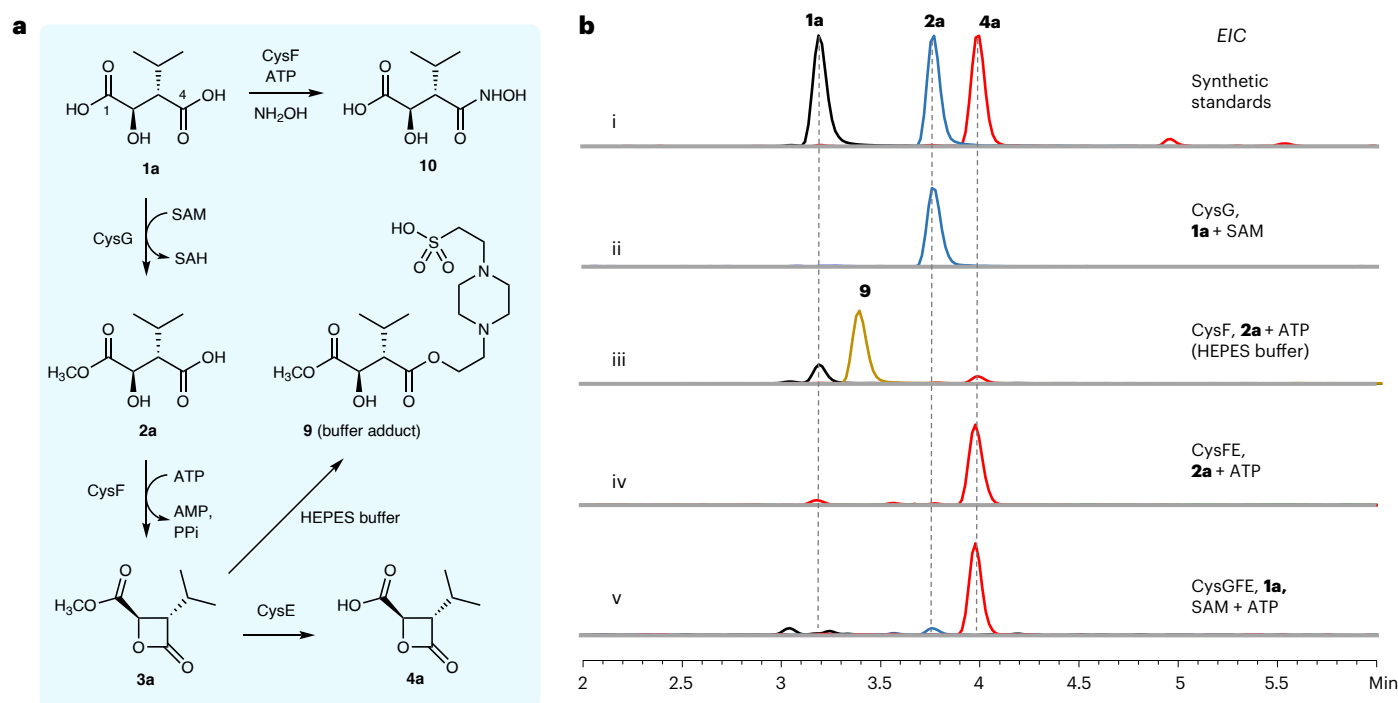


Fig. 2 | In vitro characterization of the β -lactone assembly enzymes CysGFE. a, Confirmed pathway for β -lactone assembly based on in vitro characterization of the CysGFE enzymes. **b**, LC-MS analysis for the in vitro reactions of CysGFE. EIC, extracted ion chromatogram.

led to the proposal that malic acid derivatives (**1a** and **1b**) are likely β -lactone precursors³². It was suggested that the putative methyltransferases CysG and BelI methylate the C4-carboxylates of **1a** and **1b**, respectively, to generate C4-methyl ester intermediates that are cyclized to β -lactones **4a** and **4b** by putative hydrolases CysE and BelR³². Alternatively, it was suggested that CysC and BelV, which show similarity to adenosine triphosphate (ATP)-dependent acyl-CoA synthetases, may activate **1a** and **1b**, forming C4-CoA thioesters facilitating cyclization to **4a** and **4b**³². To explore the cys and bel pathways, the proposed (2*R*,3*S*)-malic acid precursors (**1a** and **1b**; Fig. 1c) were synthesized as described previously (Supplementary Fig. 1)³⁷. Ten enzymes potentially involved in the assembly of bel and cys from the malic acid precursors (CysCDEFG and BelHIRUV) were overproduced in *Escherichia coli* and purified (Supplementary Fig. 2 and Supplementary Tables 1 and 2). First, we incubated **1a** with the putative acyl-CoA synthetase CysC and ATP, with and without CoASH but this failed to produce β -lactone **4a** as proposed³² (Supplementary Fig. 3). Next, we incubated **1a** with the putative methyltransferase CysG, *S*-adenosyl methionine (SAM) and hydrolase CysE, which also failed to produce the β -lactone **4a**. The CysG reaction of **1a** and SAM did result in the formation of a methyl ester with the same mass as the proposed C4-methyl ester intermediate. However, detailed nuclear magnetic resonance (NMR) analysis and comparisons with an authentic synthetic standard showed that CysG exclusively produced the C1-methyl ester **2a** from **1a** and SAM (Fig. 2a and Extended Data Figs. 1 and 2).

A SAM-dependent methyltransferase has been shown to activate a carboxylic acid precursor for subsequent amide-bond formation³⁸. We are unaware, however, of any other related examples of SAM-dependent carboxyl activation. The cyclization of malic acid C4-methyl ester to a strained β -lactone (**4a**), as suggested³², is also thermodynamically unfavorable. We, therefore, reasoned that CysG, rather than activating the C4-carboxyl³², may function to mask the C1-carboxylic acid as a methyl ester (**2a**) with one of the other putative ATP-dependent enzymes in the cys BGC (CysC, CysD or CysF) catalyzing β -lactonization of **2a** to **3a** instead (Fig. 2a). Incubation of **2a** with CysC or CysD and ATP failed to give any products. The reaction

of CysF, a putative adenylate (AMP)-forming enzyme, with **2a** and ATP did result in a new product but liquid chromatography–mass spectrometry (LC–MS) and NMR analysis showed that this was a C4-adduct formed from a reaction with the HEPES buffer component (**9**) (Fig. 2a,b and Supplementary Fig. 4). This suggested that lactone ester **3a** may have been produced but was attacked by the HEPES nucleophile because of the reactive nature of the β -lactone ring. To address this, we changed the buffer to MOPS or potassium phosphate and **2a** was again consumed in the CysF reaction with ATP. No β -lactone ester **3a** or adducts were evident by LC–MS, which may have been because of instability or poor ionization of **3a** under the analysis conditions. Gas chromatography (GC)–MS analysis did, however, reveal a product with the same retention time, mass and fragmentation pattern as a synthetic standard of **3a** (Extended Data Fig. 3 and Supplementary Fig. 1). A preparative-scale reaction of **2a** and ATP catalyzed by CysF was carried out, extracted with diethyl ether and rapidly subjected to ¹H NMR analysis, confirming the structure of the hydrolytically labile β -lactone **3a** (Extended Data Fig. 4). These results showed that CysF is responsible for β -lactonization, rather than functioning as an amide or peptide synthetase as initially proposed³².

β -Lactone assembly involves a cryptic three-step sequence

Given that the methyl ester of **3a** must be hydrolyzed or transformed directly to an amide or peptide product whilst retaining the labile β -lactone ring structure, we reasoned that **3a** may be the substrate of putative hydrolase CysE. Although overproduction of CysE proved problematic, a small amount of the soluble enzyme was obtained. Incubating **2a** in a tandem reaction with CysF and CysE, including ATP, resulted in the β -lactone carboxylic acid **4a** (Fig. 2a,b). With each step for β -lactone formation confirmed, the cascade reaction of malic acid **1a** with the three enzymes, CysGFE, was carried out on a larger scale, with SAM and ATP, to give **4a** in 94% isolated yield. The NMR spectra of **4a** were identical to a sample of **4a** that was produced by asymmetric synthesis using Evan's oxazolidinone (Extended Data Fig. 5 and Supplementary Fig. 1)³⁷. Assays with the set of enzyme orthologs from the bel pathway (BelIHR) also resulted in the complete conversion

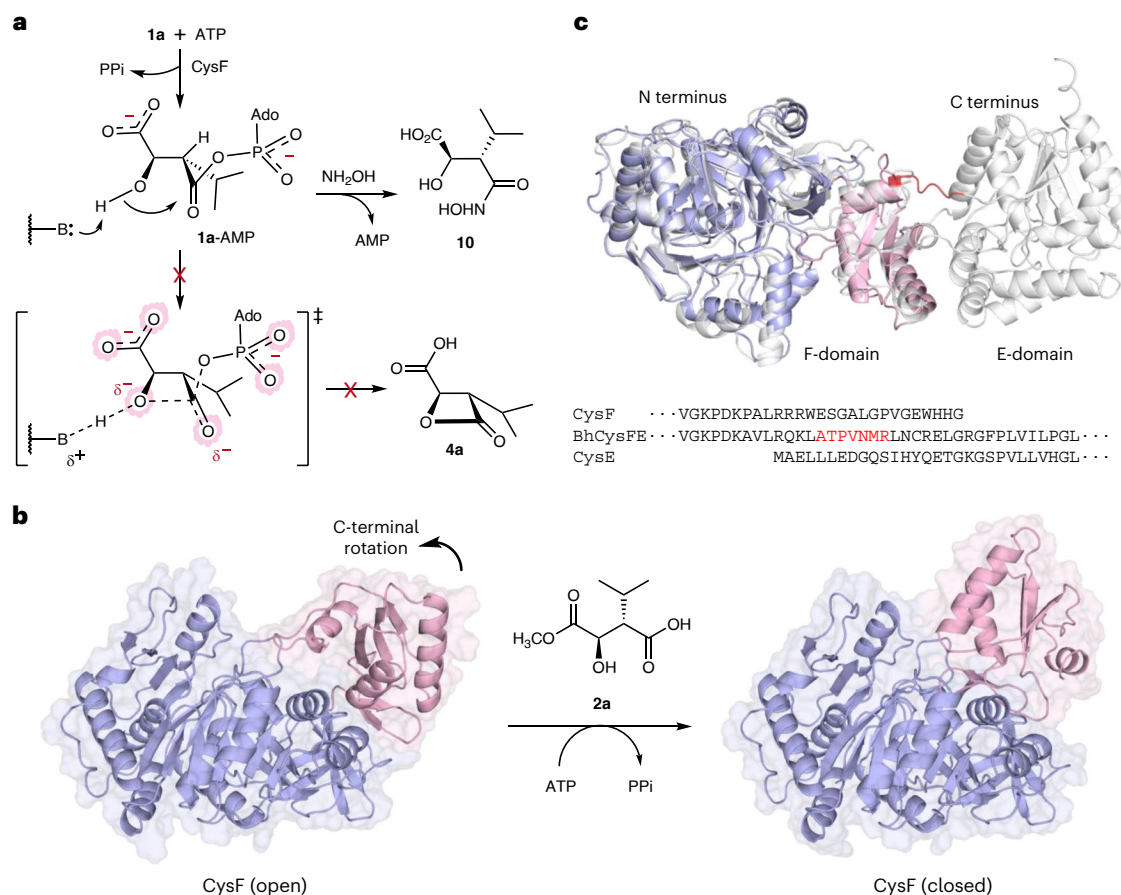


Fig. 3 | Structural and mechanistic considerations for CysF β -lactone formation. **a**, Adenylation of **1a** by CysF results in **1a-AMP**, which can be intercepted by hydroxylamine to give **10**, but does not result in β -lactone formation (**4a**), possibly because of electrostatic repulsion in the strained four-membered ring transition state. **b**, X-ray crystal structure of CysF (1.9 Å) in the open adenylation state (PDB 8RA0). CysF has a large N-terminal region (residues 1–378), shown in blue, and a small C-terminal region (residues 379–491), shown in

pink. A model of CysF in the closed state is also shown, which is based on similar conformational changes in structurally related adenylate-forming enzymes. **c**, AlphaFold model of the bifunctional fusion enzyme BhCysFE in gray (Extended Data Fig. 8), with the CysF X-ray crystal structure overlaid in blue and pink. The protein sequence alignment indicates a short linker (highlighted in red) in the fused enzyme. Protein structures were visualized with PyMOL (version 2.5.7).

of 3-*sec*-butyl malic acid **1b** to give **4b** (Fig. 1c), which was identical to the synthetic **4b** (Supplementary Fig. 1 and Extended Data Fig. 6). The cascade reaction of **1b** with BelHR, SAM and ATP was also performed on a preparative scale to give **4b** in 87% isolated yield. The fact that three enzymatic steps consuming two high-energy cofactors (SAM and ATP) are required to transform malic acids (**1a** and **1b**) to the β -lactone (**4a** and **4b**) was unexpected. In principle, the ATP-dependent activation of **1a** or **1b** by CysF or BelH may enable direct cyclization to **4a** or **4b** in a single step, rather than three (Figs. 1c and 2a). To explore why nature evolved this cryptic methylation pathway (**1**→**2**→**3**→**4**), we incubated **1a** with CysF and ATP alone. Although the CysF reaction with **1a** did not afford **4a**, the inclusion of hydroxylamine in this assay did lead to the hydroxamic acid derivative **10** in ~50% yield (Figs. 2a and 3a and Supplementary Fig. 5). This indicates that CysF can efficiently activate the C4-carboxyl of **1a** but the resulting acyl-adenylate (**1a-AMP**) is not subsequently cyclized to form the lactone **4a**. Cyclization of **1a-AMP** and the C1-methyl ester **2a-AMP** would involve similar steric constraints (Fig. 3a). However, with **1a-AMP**, the negatively charged C1-carboxylate results in additional unfavorable electrostatic interactions in the sterically hindered and strained transition state, which presumably impedes the direct cyclization to β -lactone **4a** (Fig. 3a). We, therefore, suggest that nature may have evolved this unusual cryptic methylation sequence, at the expense of SAM and ATP, to reduce electrostatic repulsion and stabilize the strained transition state for the sterically demanding β -lactonization reaction to proceed.

We determined a 1.9-Å-resolution X-ray crystal structure of CysF, which revealed a large N-terminal domain with a smaller C-terminal domain (Fig. 3b), similar to structures of the related acyl-CoA and amide-bond synthetases in the open adenylation state^{39–43}. As with the related adenylate-forming enzymes^{39–43}, it is likely that the C-terminal domain undergoes a rotation to a closed state, which may facilitate β -lactonization of the intermediate **2a-AMP** (Fig. 3a,b). In silico docking (ICM Molsoft) was performed with CysF to investigate the potential binding mode of the intermediate **2a-AMP** (Extended Data Fig. 7). The docked conformation of **2a-AMP** was consistent with the acyl-AMP-binding or ATP-binding modes observed in related enzymes^{39–43} and showed that the C2-OH of **2a-AMP** was in close proximity to H192, which may act as a general base facilitating β -lactone formation. Bioinformatics analysis of available bacterial genome sequences also revealed several BGCs similar to the *cys* BGC, except where CysF and CysE are produced as a single fusion protein with a linker sequence between the two (Fig. 3c, Extended Data Fig. 8 and Supplementary Fig. 6). One of the fusions (BhCysFE) was overproduced in *E. coli* (Supplementary Fig. 2 and Supplementary Table 3), purified and then incubated with **2a** and ATP, leading to a quantitative production of **4a** (Extended Data Fig. 8). The β -lactone ring of methyl ester **3a** is more labile than corresponding acid **4a**, which may be because of the additional electrostatic repulsion provided by the charged carboxylate group of **4a** hindering attack by water or other nucleophiles (Supplementary Fig. 7). CysFE may, therefore, have evolved as a bifunctional

fusion protein to facilitate channeling of the more unstable β -lactone **3a** between the two catalytic active sites (Fig. 3c and Extended Data Fig. 8). Various channeling mechanisms are proposed to occur when different enzyme active sites are in close proximity and these mechanisms serve to sequester unstable intermediates, limiting diffusion and subsequent degradation or toxic effects^{44–46}. Further studies would be required to explore whether channeling operates between the two active sites of CysFE.

Characterization of the cys and bel peptide assembly enzymes

Having established the pathway to the β -lactone warheads, we next sought to elucidate the remaining steps in the biosynthesis of cys and bel. Previously, it was proposed that a putative ATP-grasp enzyme CysD ligates L-valine to generate the dipeptide L-valine-L-valine, which is then coupled with the lactone acid **4a**, by CysF³². We incubated CysD with L-valine and ATP, but this failed to produce L-valine-L-valine. Similarly, CysF did not catalyze ligation of **4a** with L-valine or the dipeptide (L-valine-L-valine), which was unsurprising after we showed that CysF catalyzes β -lactonization. Next, we tested the other putative ATP-dependent enzymes in the cys BGC (CysC and CysD) to establish whether they catalyzed the ligation of β -lactone **4a** and L-valine. Whilst CysD failed to generate any products, the reaction of CysC with **4a**, L-valine and ATP gave **5** in excellent yields (Fig. 4a,b). CysC exhibits sequence similarity to other adenylate-forming enzymes, including amide-bond synthetases such as CfaL, which also *N*-acylates hydrophobic amino acids⁴³. By a process of elimination, we confirmed that the remaining ligase, CysD, catalyzes coupling of **5** with L-valine to complete the biosynthesis of cys (**6**) (Fig. 4a,b). Although other ATP-grasp ligases, such as CysD, are known to couple two amino acids to generate dipeptides⁴⁷, a pathway to cys through L-valine-L-valine³² is unlikely to have evolved in nature, given that this dipeptide would be particularly susceptible to proteolysis. The route that nature has selected (Fig. 4a) proceeds through a more stable intermediate (**5**) and, thus, conserves ATP. In the predicted bel pathway, it was suggested that the ATP-grasp enzyme BelU couples L-alanine with L-ornithine to generate a dipeptide (L-alanine-L-ornithine) that is then acylated with β -lactone **4b** by BelH to give **8c** (ref. 32). We tested both enzymes but neither possessed the proposed activity. Instead, we found that BelV, which is most similar to CysC, ligates **4b** with the side chain amino group L-ornithine to produce **7**. BelU then activates L-alanine for ligation with **7** to complete the biosynthesis of bel (Fig. 4c,d). We also found that potassium ions are required for the activity of BelU, which has been observed for other ATP-grasp enzymes⁴⁸.

Enzymatic cascades to cys and bel analogs

After elucidating the pathways to cys and bel, we set out to explore the substrate scope of the pathway enzymes and develop cascade reactions to cys and bel derivatives, including non-natural precursors similar to those used in the synthesis of promising analogs^{26,29,30}. We first reconstituted both pathways *in vitro*. Using the entire set of enzymes, CysGFEDC or BelIHRVU, total *in vitro* biosynthesis of cys and bel was achieved in one-pot cascade reactions from commercially available malic acid derivatives **1a** or **1b** and amino acid substrates (L-valine or L-ornithine and L-alanine) with SAM and ATP (Fig. 5a and Supplementary Fig. 8). The enzymatic synthesis of cys (**6b**) in 64% yield, in a single unoptimized reaction from commercially available starting material, compares favorably with the chemical syntheses of cys and bel, which typically require 9–11 steps^{26–30,37}, with an average overall yield of 10% for cys³⁷. Next, we tested the amino acid scope of the peptide assembly enzymes CysC and CysD. Whilst both enzymes had preference for hydrophobic amino acids, as expected, they exhibited considerable flexibility, accepting a wide range of non-natural amino acids commonly found in peptide therapeutics (Fig. 5b,c). CysC also accepted 11 of the proteinogenic amino acids as substrates (Extended Data Fig. 9). *O*-Benzyl-L-serine is also a CysD substrate, producing the

analog **34** in low yield. An ester derivative of **34** was previously shown to be a promising lead compound and is a notably more potent proteasome inhibitor *in vitro* than the parent cys or the widely used anticancer agent carfilzomib²⁹. Given that promising cys analogs, such as **34**, require laborious syntheses²⁹, we sought to exploit the promiscuity of CysC and CysD to develop cleaner, more efficient enzymatic cascade reactions to other analogs. First, we demonstrated that the five-enzyme, one-pot CysGFEDC cascade could transform simple malic acid **1a** to a range of cys analogs (**11–17**) in one step, with L-valine replaced by alternative nonproteinogenic amino acids (Fig. 5a). We then developed a chemoenzymatic approach to assemble cys analogs with control over the amino acid sequence, including both native and non-natural warheads (Fig. 5d). In these examples, a synthetic β -lactone warhead (**4a–4c**) is incubated with CysC along with the first amino acid (1.0 eq.) and ATP. After 12 h, the ligase CysD and a second, different amino acid (1.25 eq.) are added to the reaction. By simply changing the order of addition of amino acids, without isolating intermediates, we can control the final peptide sequence, delivering cys A (**6a**) and various analogs (**39–43**) on a preparative scale in sufficient yields without any process optimization. Although the cys analogs prepared in this work possess C-terminal carboxylic acids, they could be easily transformed to C-terminal amides or esters, which have been shown to be more potent proteasome inhibitors.

Similar cascade reactions were also used to produce bel analogs with L-lysine or L-2,4-diaminobutyrate replacing L-ornithine and various alternative amino acids in place of L-alanine (Extended Data Fig. 10). Unlike CysCD, the BelVU ligases had substantially different (orthogonal) amino acid selectivity. Consequently, BelVU cascade reactions could be performed with both enzymes and different combinations of amino acids added concomitantly. The substantial number of β -lactone peptides that we generated enzymatically (over 40 peptides; Fig. 5 and Extended Data Figs. 9 and 10) is not exhaustive or optimized and we envisage that many more analogs could be accessible using just the native cys and bel enzymes. It should also be possible, by exploiting established engineering approaches¹⁸, to evolve the cys and bel enzymes to increase their substrate scope, thereby creating more analogs, or to enhance enzyme activity and selectivity for the synthesis of specific target (lead) compounds for future drug discovery.

Discussion

We deciphered the cryptic enzymatic logic that underpins the biosynthesis of therapeutically relevant β -lactone-containing peptide natural products, revealing several surprising features, including carboxyl methylation for transient protection facilitating β -lactone assembly. Other pathways to β -lactone-containing natural products have been investigated^{19,21,49,50} but none of these rely on cryptic carboxyl methylation. Aside from a few volatile methyl esters known to be produced by SAM-dependent carboxyl methyltransferases in plants⁵¹, methyl esters are scarce in nature because of their hydrolytic instability, particularly in cells where hydrolase enzymes are abundant. There are rare examples where methyl esters have been shown to exist transiently in the biosynthetic pathways to capuramycin and biotin^{38,52}. However, we are unaware of any examples where nature has used cryptic methylation to facilitate energetically demanding enzymatic reactions, such as β -lactonization. Our observation that CysFE exists as a bifunctional fusion is also noteworthy, suggesting that these enzymes collaborate closely to transfer reactive β -lactone intermediates between active sites. In addition to well-characterized examples of physical tunneling to facilitate the transfer of intermediates between active sites, other proximity channeling mechanisms have been proposed^{44–46}. Although channeling is thought to be prevalent in primary metabolism^{44–46}, there are only a few examples where it is suggested to occur in secondary metabolism⁵³. It will be interesting to further explore the structure and properties of CysFE to establish whether the channeling phenomenon occurs in this case.

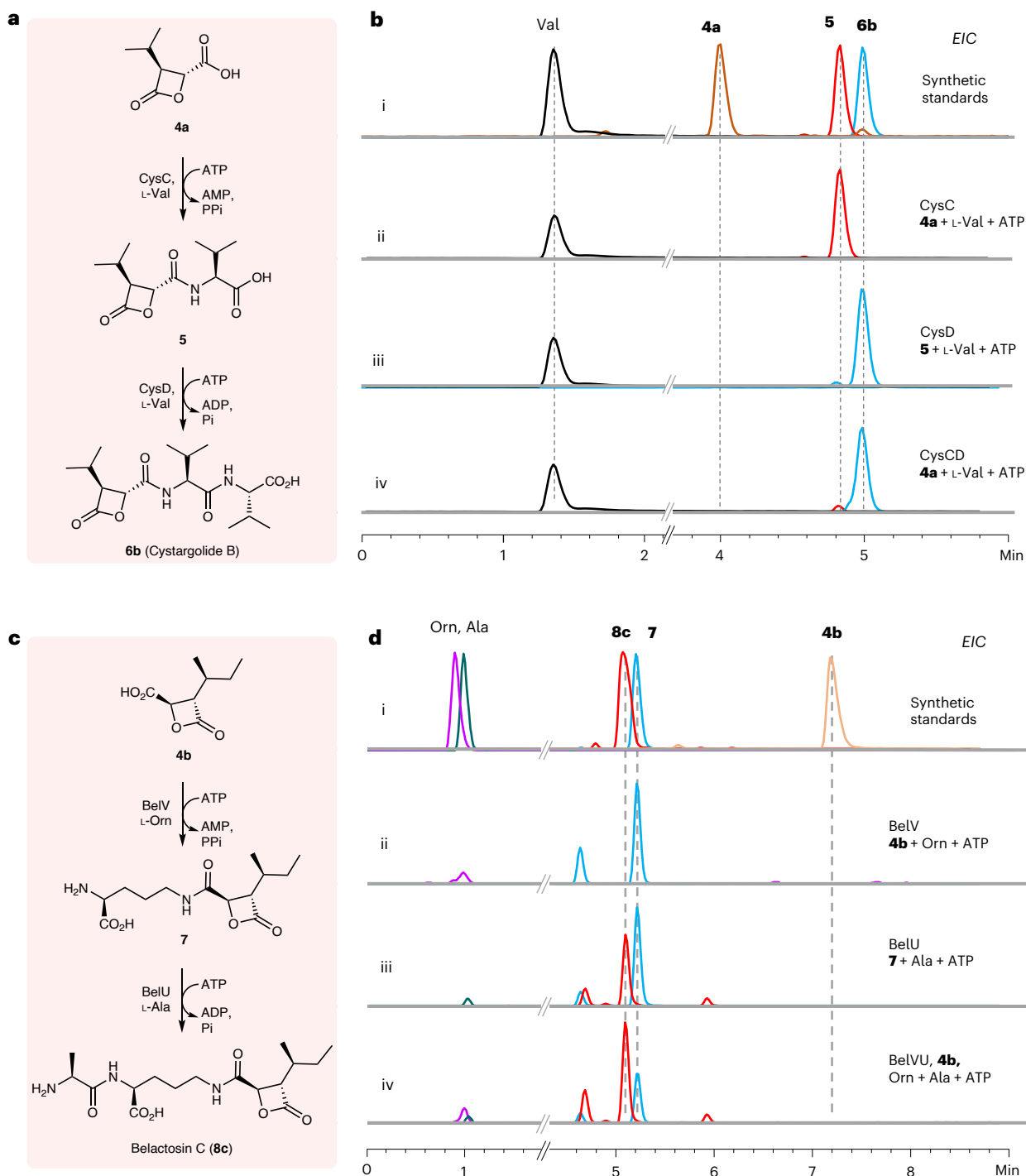


Fig. 4 | In vitro characterization of the peptide assembly enzymes CysCD and BelVU. a, The confirmed pathway for cys peptide assembly. **b**, LC-MS analysis for the reaction catalyzed by the CysCD enzymes. **c**, The confirmed pathway for bel peptide assembly. **d**, LC-MS analysis for the reaction catalyzed by the BelVU enzymes.

The pathways to cys and bel are also unusual as the reactive warhead is synthesized first, followed by peptide assembly catalyzed by standalone ligases. Other peptide natural products armed with warheads (for example, epoxyketones and aldehydes) use nonribosomal peptide synthetase (NRPS) enzymes to first assemble the peptide on carrier proteins before addition of the unstable warhead^{54–57}. The unusual ligase assembly pathways described here offer notable advantages over NRPS systems for the production of analogs and future engineering. Using native pathway enzymes, we showed how a large and diverse range of cys and bel analogs can be produced by in vitro biosynthesis through scalable enzymatic

cascade reactions. NRPS enzymes, on the other hand, are large and complex multimodular enzyme systems that are more difficult to use for synthesis in vitro. Misfolding can cause problems with the expression of active soluble NRPS and only analytical-scale quantities of NRPS products are typically produced in vitro^{58,59}. Although major advances in NRPS engineering have been made in recent years, the diversity of new products and yields obtained remain low^{58–60}. The use of standalone ligases, including ligases from different pathways and improved engineered variants^{18,43}, could offer greater flexibility and scope for producing a wider range of bioactive peptides for future drug development. Our approach for delivering analogs in

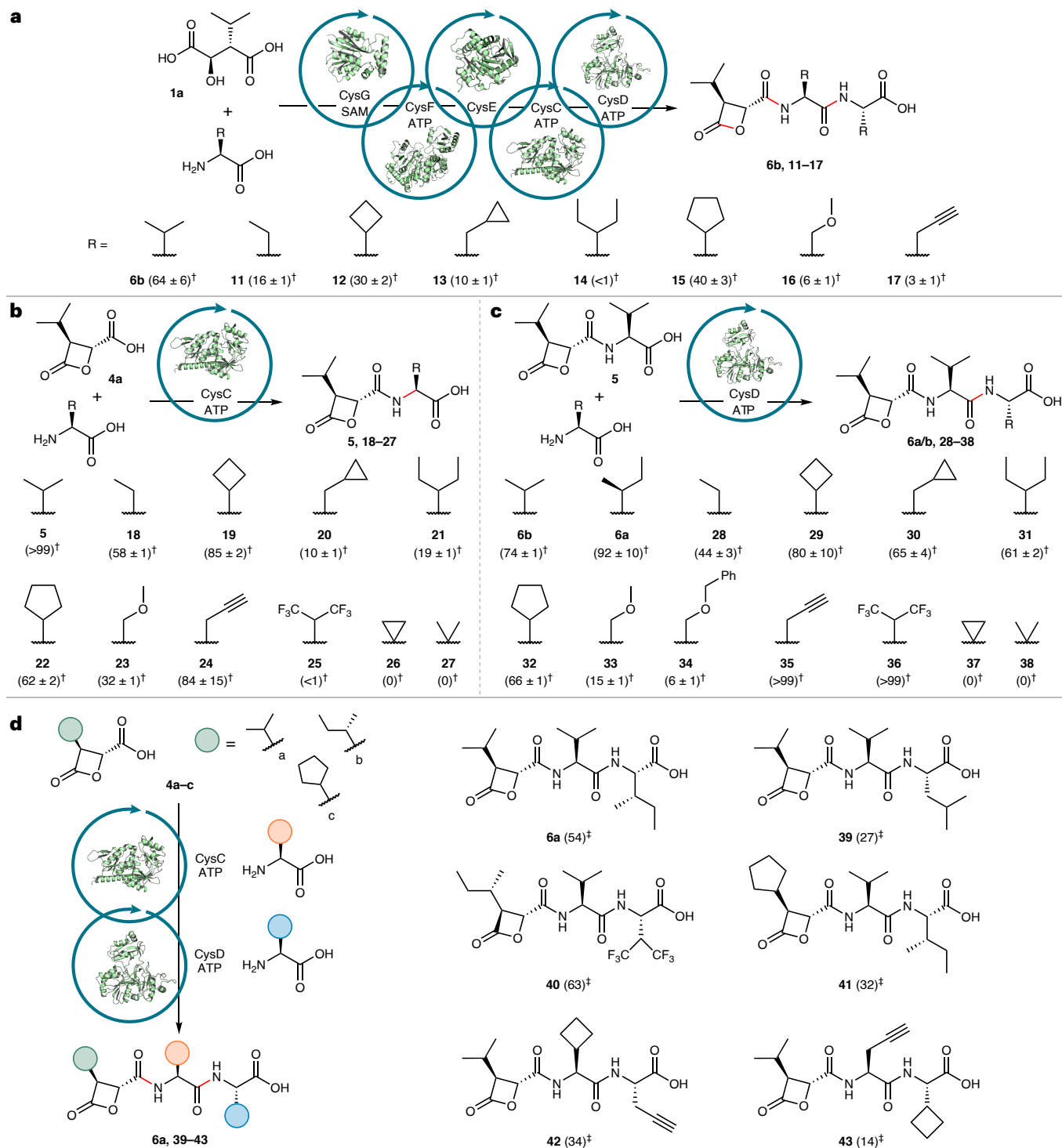


Fig. 5 | Synthetic scope of CysGFECd and enzymatic cascades to cys derivatives. a, Total in vitro biosynthesis of cys B (**6b**) and analogs using the entire set of enzymes in the cys pathway. Assays were carried out with 20 μ M of each enzyme (CysG, BhCysFE, CysC and CysD), SAM (3 mM), ATP (9 mM), **1a** (2 mM) and amino acids (6 mM), with incubation for 6 h at 25 °C. **b, c**, Assays to determine the amino acid scope of CysC and CysD were carried out with 20 μ M CysC or CysD, ATP (3 mM), **1a** or **5** (2 mM) and amino acids (3 mM), with incubation for 5 h at 25 °C. **d**, Preparative-scale CysCD enzymatic cascade

reactions were carried out at 25 °C with **4a–4c** (0.08 mmol, 4.0 mM), the first amino acid (4.0 mM), ATP (4.4 mM) and CysC (20 μ M) with incubation for 12 h, before the second amino acid (5 mM), ATP (4.4 mM) and CysD (20 μ M) were added for a further 12 h before quenching. [†]Percentage conversions were determined in triplicate by LC–MS using calibration curves with synthetic standards. [‡]Isolated yields were obtained after product isolation and purification (see Methods and the Supplementary Information for full experimental details and Extended Data Fig. 10 for similar bel cascade reactions).

one-pot scalable cascade reactions from available malic acid or β -lactone precursors is also much cleaner and more efficient than the multistep chemical syntheses currently used to manufacture peptide drugs of this type.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions

and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41589-024-01657-7>.

References

- Boike, L., Henning, N. J. & Nomura, D. K. Advances in covalent drug discovery. *Nat. Rev. Drug Discov.* **21**, 881–898 (2022).
- Sutanto, F., Konstantinidou, M. & Dömling, A. Covalent inhibitors: a rational approach to drug discovery. *RSC Med. Chem.* **11**, 876–884 (2020).
- Singh, J. The ascension of targeted covalent inhibitors. *J. Med. Chem.* **65**, 5886–5901 (2022).
- Gehring, M. & Laufer, S. A. Emerging and re-emerging warheads for targeted covalent inhibitors: applications in medicinal chemistry and chemical biology. *J. Med. Chem.* **62**, 5673–5724 (2019).
- Li, H. et al. Structure- and function-based design of *Plasmodium*-selective proteasome inhibitors. *Nature* **530**, 233–236 (2016).
- Owen, D. R. et al. An oral SARS-CoV-2 Mpro inhibitor clinical candidate for the treatment of COVID-19. *Science* **374**, 1586–1593 (2021).
- Gersch, M., Kreuzer, J. & Sieber, S. A. Electrophilic natural products and their biological targets. *Nat. Prod. Rep.* **29**, 659–682 (2012).
- Kaysler, L. Built to bind: biosynthetic strategies for the formation of small-molecule protease inhibitors. *Nat. Prod. Rep.* **36**, 1654–1686 (2019).
- Hubbell, G. E. & Tepe, J. J. Natural product scaffolds as inspiration for the design and synthesis of 20S human proteasome inhibitors. *RSC Chem. Biol.* **1**, 305–332 (2020).
- Kim, K. B. & Crews, C. M. From epoxomicin to carfilzomib: chemistry, biology, and medical outcomes. *Nat. Prod. Rep.* **30**, 600–604 (2013).
- Mora-Ochomogo, M. & Lohans, C. T. β -Lactam antibiotic targets and resistance mechanisms: from covalent inhibitors to substrates. *RSC Med. Chem.* **12**, 1623–1639 (2021).
- Hamed, R. B. et al. The enzymes of β -lactam biosynthesis. *Nat. Prod. Rep.* **30**, 21–107 (2013).
- Almaliti, J. et al. Exploration of the carmaphycins as payloads in antibody drug conjugate anticancer agents. *Eur. J. Med. Chem.* **161**, 416–432 (2019).
- Kisselev, A. F., Van Der Linden, W. A. & Overkleeft, H. S. Proteasome inhibitors: an expanding army attacking a unique target. *Chem. Biol.* **19**, 99–115 (2012).
- Isidro-Llobet, A. et al. Sustainability challenges in peptide synthesis and purification: from R&D to production. *J. Org. Chem.* **84**, 4615–4628 (2019).
- Ferrazzano, L. et al. Sustainability in peptide chemistry: current synthesis and purification technologies and future challenges. *Green Chem.* **24**, 975–1020 (2022).
- Pattabiraman, V. R. & Bode, J. W. Rethinking amide bond synthesis. *Nature* **480**, 471–479 (2011).
- Buller, R. et al. From nature to industry: harnessing enzymes for biocatalysis. *Science* **382**, eadh8615 (2023).
- Robinson, S. L., Christenson, J. K. & Wackett, L. P. Biosynthesis and chemical diversity of β -lactone natural products. *Nat. Prod. Rep.* **36**, 458–475 (2019).
- Kumar, P. & Dubey, K. K. Current trends and future prospects of lipstatin: a lipase inhibitor and pro-drug for obesity. *RSC Adv.* **5**, 86954–86966 (2015).
- Bauman, K. D. et al. Enzymatic assembly of the salinosporamide γ -lactam- β -lactone anticancer warhead. *Nat. Chem. Biol.* **18**, 538–546 (2022).
- Gill, K. A. et al. Cystargolides, 20S proteasome inhibitors isolated from *Kitasatospora cystarginea*. *J. Nat. Prod.* **78**, 822–826 (2015).
- Asai, A., Hasegawa, A., Ochiai, K., Yamashita, Y. & Mizukami, T. Belactosin A, a novel antitumor antibiotic acting on cyclin/CDK mediated cell cycle regulation, produced by *Streptomyces* sp. *J. Antibiot.* **53**, 81–83 (2000).
- Asai, A. et al. A new structural class of proteasome inhibitors identified by microbial screening using yeast-based assay. *Biochem. Pharmacol.* **67**, 227–234 (2004).
- Groll, M., Larionov, O. V., Huber, R. & de Meijere, A. Inhibitor-binding mode of homobelactosin C to proteasomes: new insights into class I MHC ligand generation. *Proc. Natl Acad. Sci. USA* **103**, 4576–4579 (2006).
- Korotkov, V. S. et al. Synthesis and biological activity of optimized belactosin C congeners. *Org. Biomol. Chem.* **9**, 7791–7798 (2011).
- De Meijere, A. et al. Synthesis and biological activity of simplified belactosin C analogues. *Org. Biomol. Chem.* **10**, 6363–6374 (2012).
- Kawamura, S., Unno, Y., Akira, A., Mitsuhiro, A. & Satoshi, S. Development of a new class of proteasome inhibitors with an epoxyketone warhead: rational hybridization of non-peptidic belactosin derivatives and peptide epoxyketones. *Bioorg. Med. Chem.* **22**, 3091–3095 (2014).
- Niroula, D. et al. Design, synthesis, and evaluation of cystargolide-based β -lactones as potent proteasome inhibitors. *Eur. J. Med. Chem.* **157**, 962–977 (2018).
- Viera, C. R. et al. Cystargolide-based amide and ester Pz analogues as proteasome inhibitors and anti-cancer agents. *R. Soc. Open Sci.* **9**, 220358 (2022).
- Illigmann, A. et al. Structure of *Staphylococcus aureus* ClpP bound to the covalent active-site inhibitor cystargolide A. *Angew. Chem. Int. Ed. Engl.* **63**, e202314028 (2024).
- Wolf, F. et al. Biosynthesis of the β -lactone proteasome inhibitors belactosin and cystargolide. *Angew. Chem. Int. Ed. Engl.* **56**, 6665–6668 (2017).
- Beller, P. et al. Characterization of the cystargolide biosynthetic gene cluster and functional analysis of the methyltransferase CysG. *J. Biol. Chem.* **300**, 105507 (2024).
- Shimo, S. et al. Stereodivergent nitrocyclopropane formation during biosynthesis of belactosins and hormaomycins. *J. Am. Chem. Soc.* **143**, 18413–18418 (2021).
- Li, X., Shimaya, R., Dairi, T., Chang, W. & Ogasawara, Y. Identification of cyclopropane formation in the biosyntheses of hormaomycins and belactosins: sequential nitration and cyclopropanation by metalloenzymes. *Angew. Chem. Int. Ed. Engl.* **61**, e202113189 (2022).
- Engelbrecht, A. et al. Discovery of a cryptic nitro intermediate in the biosynthesis of the 3-(*trans*-2'-aminocyclopropyl)alanine moiety of belactosin A. *Org. Lett.* **24**, 736–740 (2022).
- Tello-Aburto, R., Hallada, L. P., Niroula, D. & Rogelj, S. Total synthesis and absolute stereochemistry of the proteasome inhibitors cystargolides A and B. *Org. Biomol. Chem.* **13**, 10127–10130 (2015).
- Funabashi, M. et al. An ATP-independent strategy for amide bond formation in antibiotic biosynthesis. *Nat. Chem. Biol.* **6**, 581–586 (2010).
- Hisanaga, Y. et al. Structural basis of the substrate-specific two-step catalysis of long chain fatty acyl-CoA synthetase dimer. *J. Biol. Chem.* **279**, 31717–31726 (2004).
- Scagolione, A. et al. Structure of the adenylation domain Thr1 involved in the biosynthesis of 4-chlorothreonine in *Streptomyces* sp. OH-5093—protein flexibility and molecular bases of substrate specificity. *FEBS J.* **284**, 2981–2999 (2017).
- Shelton, C. et al. Rational inhibitor design for *Pseudomonas aeruginosa* salicylate adenylation enzyme PchD. *J. Biol. Inorg. Chem.* **27**, 541–551 (2022).

42. Petchey, M. et al. The broad aryl acid specificity of the amide bond synthetase McbA suggests potential for the biocatalytic synthesis of amides. *Angew. Chem. Int. Ed. Engl.* **57**, 11584–11588 (2018).
43. Winn, M. et al. Discovery, characterization and engineering of ligases for amide synthesis. *Nature* **593**, 391–398 (2021).
44. Pareek et al. Metabolic channeling: predictions, deductions, and evidence. *Mol. Cell* **81**, 3775–3785 (2021).
45. Castellana, M. et al. Enzyme clustering accelerates processing of intermediates through metabolic channeling. *Nat. Biotechnol.* **32**, 1011–1018 (2014).
46. Wheeldon, I. et al. Substrate channelling as an approach to cascade reactions. *Nat. Chem.* **8**, 299–309 (2016).
47. Ogasawara, Y. & Dairi, T. Biosynthesis of oligopeptides using ATP-grasp enzymes. *Chem. Eur. J.* **23**, 10714–10724 (2017).
48. Pederick, J. L., Thompson, A. P., Bell, S. G. & Bruning, J. B. D-Alanine-D-alanine ligase as a model for the activation of ATP-grasp enzymes by monovalent cations. *J. Biol. Chem.* **295**, 7894–7904 (2020).
49. Kreitler, D. F. et al. The structural basis of N-acyl- α -amino- β -lactone formation catalyzed by a nonribosomal peptide synthetase. *Nat. Commun.* **10**, 1–13 (2019).
50. Christenson, J. K. et al. β -Lactone synthetase found in the olefin biosynthesis pathway. *Biochemistry* **56**, 348–351 (2017).
51. Ward, L. C., McCue, H. V. & Carnell, A. J. Carboxyl methyltransferases: natural functions and potential applications in industrial biotechnology. *ChemCatChem* **13**, 121–128 (2021).
52. Lin, S., Hanson, R. E. & Cronan, J. E. Biotin synthesis begins by hijacking the fatty acid synthetic pathway. *Nat. Chem. Biol.* **6**, 682–688 (2010).
53. Faylo, J. L., Ronnebaum, T. A. & Christianson, D. W. Assembly-line catalysis in bifunctional terpene synthases. *Acc. Chem. Res.* **54**, 3780–3791 (2021).
54. Zabala, D. et al. A flavin-dependent decarboxylase–dehydrogenase–monooxygenase assembles the warhead of α,β -epoxyketone proteasome inhibitors. *J. Am. Chem. Soc.* **138**, 4342–4345 (2016).
55. Chen, Y., McClure, R. A., Zheng, Y., Thomson, R. J. & Kelleher, N. L. Proteomics guided discovery of flavopeptins: anti-proliferative aldehydes synthesized by a reductase domain-containing non-ribosomal peptide synthetase. *J. Am. Chem. Soc.* **135**, 10449–10456 (2013).
56. Mulleney, M. W., McClure, R. A., Robey, M. T., Kelleher, N. L. & Thomson, R. J. Natural products from thioester reductase containing biosynthetic pathways. *Nat. Prod. Rep.* **9**, 847–878 (2018).
57. Zhao, Q. et al. Characterization of the azinomycin B biosynthetic gene cluster revealing a different iterative type I polyketide synthase for naphthoate biosynthesis. *Chem. Biol.* **15**, 693–705 (2008).
58. Bozhüyük, K. A. J. et al. De novo design and engineering of non-ribosomal peptide synthetases. *Nat. Chem.* **10**, 275–281 (2017).
59. Bozhüyük, K. A. J. et al. Modification and de novo design of non-ribosomal peptide synthetases using specific assembly points within condensation domains. *Nat. Chem.* **11**, 653–661 (2019).
60. Thong, W. L. et al. Gene editing enables rapid engineering of complex antibiotic assembly lines. *Nat. Commun.* **12**, 1–10 (2021).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

Methods

General information

Chemicals were purchased from Sigma-Aldrich, Fluorochem, Acros Organics, Fisher Scientific UK, Bachem, Alfa Aesar or Apollo Scientific and used without further purification unless otherwise stated. Molecular biology enzymes were purchased from New England Biolabs unless otherwise stated. Codon-optimized genes were synthesized by Twist Bioscience. Proteins were analyzed by SDS-PAGE on precast gels (Invitrogen Novex WedgeWell 8–16% Tris-glycine gel). Low-resolution LC-MS was performed on an Agilent 1260 LC system fitted with an Agilent 6130 Quadrupole MS. High-resolution LC-MS and tandem LC-MS/MS were recorded on an Agilent 6560 quadrupole time-of-flight (Q-TOF) + Agilent 1290 Infinity LC system. GC-MS was performed on an Agilent GC 7890B coupled with 5975 Series mass selective detector. NMR spectra were recorded on Bruker Avance III spectrometers (400 MHz, 500 MHz or 800 MHz) with TopSpin and IconNMR. Deuterated solvents used included CDCl₃, CD₃OD, DMSO-d₆, D₂O and D₂O with 0.1% DCl. NMR data were processed using MestReNova version 11 software.

Cloning

Synthetic genes for CysF, BelH, CysD and BelU were initially cloned into the pET28-a(+) plasmid (using NdeI and XhoI sites) with an N-terminal histidine tag sequence. However, N-terminal histidine-tagged CysF produced insoluble protein. N-terminal histidine-tagged CysD and BelU produced soluble expression but with no activity. CysF, CysD and BelU were, therefore, cloned into the pET21-a(+) plasmid (using NdeI and XhoI sites) to insert a C-terminal histidine tag (Supplementary Tables 1 and 2). Synthetic genes for CysG, CysE, CysC, Bell, BelR, BelV and BhCysFE were inserted into the pET28-a(+) plasmid (using NcoI and XhoI sites) carrying C-terminal histidine tag sequences (Supplementary Tables 1 and 2).

Protein expression

Luria-Bertani medium (10 ml) with 50 µg ml⁻¹ kanamycin (pET28-a(+)) or 100 µg ml⁻¹ ampicillin (pET21-a(+)) was inoculated with *E. coli* BL21(DE3) containing a plasmid and incubated for 18 h at 37 °C with shaking (180 r.p.m.). The resulting seed culture (8 ml) was then used to inoculate 800 ml of autoinduction 2YT broth (with trace elements from Formedium) containing 50 µg ml⁻¹ kanamycin (pET28a(+)) or 100 µg ml⁻¹ ampicillin (pET21-a(+)). After incubation for 4 h at 37 °C with shaking (180 r.p.m.), the temperature was reduced to 20 °C and incubation was continued for a further 18 h (180 r.p.m.). The cultures were then harvested by centrifugation (3,000g, 4 °C, 10 min) and cells were resuspended in 80 ml of PBS, transferred into two 50-ml falcon tubes, pelleted by centrifugation (3,000g, 4 °C, 15 min) and then frozen until further usage. CysE (with C-terminal histidine tag) produced mainly insoluble protein. A small amount of soluble CysE was, however, obtained (-5 mg of protein from 1 L of cell culture) after the standard protein purification method described below. For the three fused bifunctional CysFE enzymes tested, only BhCysFE was produced in soluble form. MsCysFE afforded a very low level of expression, whilst MpCysFE showed good expression but formed inclusion bodies.

Protein purification

Frozen cell pellets were thawed and resuspended in 40 ml of lysis buffer (50 mM NaH₂PO₄, 300 mM NaCl, 10 mM imidazole and 10% glycerol; pH 7.8) and lysed by sonication (10 min, 50% pulse, 60% amplitude). The lysate was then cleared by centrifugation (23,224g, 40 min, 4 °C), 1 ml of pre-equilibrated Ni-NTA resin was added and the tube was incubated at 4 °C for 30 min. The lysate with resin was loaded onto a gravity-flow column (Bio-Rad) and flowthrough was collected. The resin was washed with 10 ml of lysis buffer, followed by two washes with 10 ml of wash buffer (50 mM NaH₂PO₄, 300 mM NaCl, 40 mM imidazole and 10% glycerol; pH 7.8), and eluted with 10 ml of elution buffer (50 mM NaH₂PO₄, 300 mM NaCl, 100 mM imidazole and 10% glycerol; pH 7.8). The eluent

from the column was monitored by Bio-Rad protein assay and elution continued until no further protein could be seen eluting. Samples of wash and elution fractions were analyzed by SDS-PAGE. The wash and elution fractions that contained the protein of interest (with good purity) were combined and concentrated down to 2.5 ml using a centrifugal concentrator (Sartorius, Vivaspin 20; mol. wt. cutoff, 30 kDa or 10 kDa depending on the size of the protein of interest). The 2.5-ml solution was applied to an equilibrated PD-10 desalting column (GE Healthcare, performed according to the manufacturer's instructions). The column was eluted with 3.5 ml of protein storage buffer (20 mM NaH₂PO₄, 300 mM NaCl and 10% glycerol; pH 7.8). If necessary, the resulting eluent was further concentrated to approximately 10 mg ml⁻¹ or more. Protein was aliquoted into single-use tubes, flash-frozen using liquid nitrogen and stored at -80 °C.

Further purification for protein crystallization

For the protein crystallization, the plasmid encoding the CysF gene was used to transform NiCo21(DE3) competent *E. coli* cells. Following transformation, glycerol stocks were prepared from the overnight cultures of a single colony (sequencing confirmed). The enzyme was expressed and first purified using Ni-NTA column according to the general procedure described above. The wash and elution protein fractions with good purity (judged by SDS-PAGE) were combined and concentrated using a Vivaspin centrifugal concentrator (mol. wt. cutoff, 30 kDa) up to 20 mg ml⁻¹. For further purification, approximately 0.6 ml of this concentrated protein solution was injected into a 2-ml loop attached to the ÄKTA pure protein purification system with GE Healthcare UNICORN software (version 7.3). The protein was passed down a Superdex 200 Increase 10/300 GL gel filtration column that was pre-equilibrated with gel filtration buffer (50 mM Tris-HCl, 300 mM NaCl, 10 mM MgCl₂ and 10% glycerol; pH 7.5). CysF was eluted with gel filtration buffer and buffer-exchanged to crystallization buffer (50 mM Tris-HCl, 100 mM NaCl, 10 mM MgCl₂ and 5% glycerol; pH 7.5) using a PD-10 desalting column (pre-equilibrated with crystallization buffer). The resulting fraction was concentrated to 8–10 mg ml⁻¹ using an Amicon Ultra 0.5-ml concentrator (mol. wt. cutoff, 30 kDa).

Crystallogensis

Single crystals of CysF were prepared by mixing 200 nl of 8 mg ml⁻¹ protein in crystallization buffer with equal volumes of precipitant. All trials were conducted by sitting-drop vapor diffusion and incubated at 4 °C. Crystals of the apo protein were formed in 0.1 M sodium citrate pH 5.5, 20% w/v polyethylene glycol (PEG) 3000 (JCSG + Eco A2, Molecular Dimensions). Individual crystals were cryoprotected in mother liquor supplemented with 25% PEG 200 before flash-cooling in liquid nitrogen. Data were collected from single crystals at the Diamond Light Source (MX31850-i04) and subsequently scaled and reduced with Xia2. Preliminary phasing was performed by molecular replacement in Phaser using a search model generated in AlphaFold^{61,62}. Iterative cycles of rebuilding and refinement were performed in Coot and Phenix.refine, respectively. Structure validation with MolProbity and PDBREDO was integrated into the iterative rebuild and refinement process. The resolution cut of the data was determined by paired refinement as implemented in PDBREDO. Complete data collection and refinement statistics can be found in Supplementary Table 4. Coordinates and structure factors were deposited in the Protein Data Bank (PDB) under accession code [8RA0](https://doi.org/10.1038/s41589-024-01657-7).

Bioinformatic analysis of CysF and discovery of fused CysFE

Bioinformatic analysis was conducted on the cys and bel pathway enzymes using the Enzyme Function Initiative tools⁶³. The analyses used databases UniProt 2022-04, InterPro 91 and corresponding information from the European Nucleotide Archive. The identification of the fused CysFE configuration began with a basic local alignment search tool query using UniProt [AOAIW6R555](https://doi.org/10.1038/s41589-024-01657-7) for CysF to generate a

sequence similarity network (SSN) with default settings. An alignment score threshold of 10^6 was chosen to identify a subnetwork containing both CysF and BelH as they are isofunctional in their enzymatic capabilities. The obtained SSN was further analyzed with the Genome Neighborhood Tool, leading to the obtention of genome neighborhood diagrams (Supplementary Fig. 6). Manual inspection of these diagrams revealed three genomes exhibiting a CysF and CysE fusion, indicating a bifunctional enzyme. The structure of BhCysFE was modeled using AlphaFold (AF-AOA562R406-F1-model_v4, <https://www.alphafold.ebi.ac.uk/entry/AOA562R406>)^{61,62}.

Modeling of CysF closed conformation

The crystal structure of CysF was used to search for homologs through Foldseek and Dali. Identified homologs were observed that displayed structures in both open and closed conformations. Crystal structures of 4-coumarate CoA ligase in both open and closed conformations (PDB 5B5M, open; PDB 5BSR, closed) were used as templates upon which to model a closed state of CysF. The 4-coumarate CoA ligase structure revealed that domain closure is achieved through a hinge-like rigid-body reorganization of the two domains. A closed model of CysF was, therefore, modeled by superposition of the N-terminal domain of CysF with the corresponding N-terminal domain of the template (root-mean-square deviation (r.m.s.d.) of 1.86 Å after secondary-structure alignment). Subsequently, the C-terminal domain of CysF was superimposed with the C-terminal domain of the template in the closed conformation (r.m.s.d. of 3.03 Å after secondary-structure alignment) followed by an energy minimization (Yasara).

CysF docking studies

A model of **2a**-AMP was constructed and docked into the crystal structure of CysF. The pocket for docking was identified using ICM Pocket Finder as implemented in ICM-Pro⁶⁴. The top hit from the docking as ranked by the RTCNN (radial convolution neural network with topological convolutions) score (−29.78) is presented in Extended Data Fig. 7. The docked pose was compared to structures of similar enzymes (PDB 5WM3, 5IE3, 4FUT, 4GXR and 4GXQ). A multiple-sequence alignment of these enzymes is presented in Extended Data Fig. 7.

Chromatography

LC–MS analysis was performed with a Kinetex XB-C18 Core Shell column (100 mm × 4.6 mm, 5 μm, Phenomenex), using a flow rate of 1.0 ml min^{−1}, column oven temperature of 40 °C, mobile phase A of water with 0.1 (v/v) formic acid and mobile phase B of acetonitrile with 0.1% (v/v) formic acid. The mobile phase program for method 1 was as follows: 0–6 min, 5–95% B; 6–7 min, 95% B; 7–7.5 min, 95–5% B; 7.5–11 min, 5% B. The mobile phase program for method 2 was as follows: 0–15 min, 5–95% B; 15–16 min, 95% B; 16–16.5 min, 95–5% B; 16.5–20 min, 5% B.

GC–MS analysis was performed with a VF-5ht column (30 m × 0.25 mm, 0.1 μm, J&W), using a flow rate of 1 ml min^{−1}, inlet temperature of 240 °C, split ratio of 100:1, inlet pressure of 7.6 psi and helium as the carrier gas. The oven temperature program was an initial temperature of 50 °C, held for 2 min, followed by a temperature ramp at 30 °C min^{−1} to 350 °C, held for 3 min.

Enzyme assays for testing the proposed pathway³²

CysC was initially assayed for the lactonization of **1a** to **4a** (Fig. 2a) in a reaction mixture (100 μl) containing 2 mM **1a**, 5 mM ATP, with or without 5 mM CoASH, 10 mM MgCl₂ and 20 μM CysC in 100 mM potassium phosphate buffer (pH 7.8) incubated at 25 °C for 12 h. The reaction was quenched by addition of 100 μl of acetonitrile and the precipitated protein was removed by centrifugation. The supernatant was subjected to LC–MS analysis (method 1).

CysGE and BelIR were tested for methylation and lactonization activity (**1a** to **4a**, Fig. 2a) in reactions (100 μl) containing 2 mM **1a**, 5 mM SAM, 10 mM MgCl₂, 20 μM CysG or Bell and 20 μM CysE or BelR

in 100 mM potassium phosphate buffer (pH 7.8) incubated at 25 °C for 12 h. The reactions were quenched and analyzed by LC–MS analysis (as above).

The proposed amide-bond-forming activities of CysF and BelH (Fig. 4a,d) were tested in reactions (100 μl) containing 2 mM **4a** (CysF) or **4b** (BelH), 5 mM L-valine-L-valine (CysF) or L-alanine-L-ornithine (**12**, BelH), 5 mM ATP, 10 mM MgCl₂ and 20 μM CysF or BelH in 100 mM potassium phosphate buffer (pH 7.8, with 300 mM NaCl and 10% v/v glycerol) incubated at 25 °C for 12 h. The reactions were quenched and subjected to LC–MS analysis (as above).

The proposed amide-bond-forming activities of CysD and BelU (Fig. 4a,d) were also tested in reactions (100 μl) with 2 mM L-valine (for CysD) or 2 mM L-ornithine and 2 mM L-alanine (for BelU), 5 mM ATP, 10 mM MgCl₂ and 20 μM CysD or BelU in 100 mM potassium phosphate buffer (pH 7.8) incubated at 25 °C for 12 h. The reactions were quenched and analyzed (as above).

Enzyme assays for the new pathways (Figs. 2 and 4)

The CysG and Bell assay mixtures (100 μl) containing 2 mM **1a** or **1b**, 5 mM SAM, 10 mM MgCl₂, and 20 μM CysG or Bell in 100 mM potassium phosphate buffer (pH 7.8) were incubated at 25 °C for 12 h. The reactions were quenched by addition of 100 μl acetonitrile and the precipitated protein was removed by centrifugation. The supernatant was subjected to LC–MS analysis (method 1).

The CysGF and BelIH cascade reactions (100 μl) containing 2 mM **1a** or **1b**, 5 mM SAM, 5 mM ATP, 10 mM MgCl₂, 20 μM CysG or Bell and 20 μM CysF or BelH in 100 mM HEPES or potassium phosphate buffer (pH 7.8) were incubated at 25 °C for 12 h. The reaction was quenched by addition of 100 μl of acetonitrile and the precipitated protein was removed by centrifugation. The supernatant was subjected to LC–MS analysis (method 1). For GC–MS analysis, the reaction was scaled up to 200 μl and the reaction time was 5 h. The reaction was quenched by extracting with ethyl acetate (200 μl). The organic layer was dried over MgSO₄ and analyzed by GC–MS.

The CysFE and BelHR cascade reactions (100 μl) containing 2 mM **2a** or **2b**, 5 mM ATP, 10 mM MgCl₂, 20 μM CysF or BelH and 20 μM CysE or BelR in 100 mM potassium phosphate buffer (pH 7.8) were incubated at 25 °C for 12 h. The reaction was quenched by addition of 100 μl of acetonitrile and the precipitated protein was removed by centrifugation. The supernatant was subjected to LC–MS analysis (method 1).

The CysGE and BelIHR cascade reactions (100 μl) containing 2 mM **1a** or **1b**, 5 mM SAM, 5 mM ATP, 10 mM MgCl₂, 20 μM CysG or Bell, 20 μM CysF or BelH and 20 μM CysE or BelR in 100 mM potassium phosphate buffer (pH 7.8) were incubated at 25 °C for 12 h. The reaction was quenched by addition of 100 μl of acetonitrile and the precipitated protein was removed by centrifugation. The supernatant was subjected to LC–MS analysis (method 1).

The CysC and BelV assays (100 μl) containing 2 mM **4a** or **4b**, 4 mM L-valine or L-ornithine, 5 mM ATP, 10 mM MgCl₂ and 20 μM CysC or BelV in 100 mM potassium phosphate buffer (pH 7.8) were incubated at 25 °C for 2 h. The reaction was quenched and analyzed (see directly above).

The CysD and BelU assays (100 μl) containing 2 mM **5** or **7**, 4 mM L-valine or L-alanine, 5 mM ATP, 10 mM MgCl₂ and 20 μM CysD or BelU in 100 mM potassium phosphate buffer (pH 7.8) were incubated at 25 °C for 2 h. The reaction was quenched (as above) and subjected to LC–MS analysis (method 1 for CysD and method 2 for BelU reactions). We noted that potassium ions served as an activator for CysD catalysis and were essential for the activity of BelU.

The CysCD and BelVU cascade reactions (100 μl) containing 2 mM **4a** or **4b**, 6 mM L-valine (CysCD) or 4 mM L-ornithine with 4 mM L-alanine (BelVU), 10 mM ATP, 10 mM MgCl₂, 20 μM CysC or BelV and 20 μM CysD or BelU in 100 mM potassium phosphate buffer (pH 7.8) at 25 °C for 12 h. The reaction was quenched (as above) and subjected to LC–MS analysis (method 1 for CysCD and method 2 for BelVU reactions).

Substrate scope analysis (Fig. 5)

The CysGFEC enzyme cascade reactions (100 μ l) containing 2 mM **1a**, 3 mM SAM, 9 mM ATP, 6 mM L-valine (or other amino acids), 10 mM MgCl₂, 20 μ M CysG, 20 μ M BhCysFE, 20 μ M CysC and 20 μ M CysD in 100 mM potassium phosphate buffer (pH 7.8) were incubated at 25 °C for 6 h. The reaction was quenched by addition of 100 μ l of acetonitrile and the precipitated protein was removed by centrifugation. The supernatant was subjected to LC–MS analysis (method 1). Reaction yields (average of three replicates) were determined from calibration curves using synthetic product standards.

The BellHRVU enzyme cascade reactions (100 μ l) containing 2 mM **1b**, 3 mM SAM, 9 mM ATP, 5 mM L-ornithine (or other amino acids), 5 mM L-alanine (or other amino acids), 10 mM MgCl₂, 15 μ M Bell, 15 μ M BelH, 15 μ M BelR, 15 μ M BelV and 15 μ M BelU in 100 mM potassium phosphate buffer (pH 7.8) were incubated at 25 °C for 6 h. The reactions were the quenched and analyzed (as above).

The CysC and BelV reaction mixtures (100 μ l) containing 2 mM **4a** or **4b**, 3 mM amino acid, 3 mM ATP, 10 mM MgCl₂ and 20 μ M CysC or BelV in 100 mM potassium phosphate buffer (pH 7.8) were incubated at 25 °C for 5 h. The reaction was quenched by addition of 100 μ l of acetonitrile and the precipitated protein was removed by centrifugation. The supernatant was diluted ten times with acetonitrile–H₂O (50:50) and subjected to LC–MS analysis (method 1). Product yields for CysC reactions (average of three replicates) were determined from calibration curves using synthetic product standards. The yield of **25** was determined by ¹⁹F NMR.

The CysD and BelU reaction mixtures (100 μ l) containing 2 mM **5** or **7**, 3 mM amino acid, 3 mM ATP, 10 mM MgCl₂ and 20 μ M CysD or BelU in 100 mM potassium phosphate buffer (pH 7.8) were incubated at 25 °C for 5 h. The reaction was quenched as described above. The supernatant was diluted ten times with acetonitrile–H₂O (50:50) and subjected to LC–MS analysis (method 1). Product yields for CysD reactions (average of three replicates) were determined using calibration curves. The yield of **36** was determined by ¹⁹F NMR.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The coordinates of the CysF X-ray crystal structures were deposited to PDB **8RA0**. The AlphaFold structure of BhCysFE can be accessed at <https://www.alphafold.ebi.ac.uk/entry/AOA562R406> (AF-AOA562R406-F1-model_v4). Structures used for modeling and docking studies can be accessed from PDB **5BSM**, **5BSR**, **5WM3**, **5IE3**, **4FUT**, **4GXR** and **4GXQ**. All proteins characterized in this study can be accessed from UniProt using the accession codes presented in Supplementary Tables 1–3 and their synthetic gene sequences are provided in Supplementary Data 1. The remaining data are available in the main text or the Supplementary Information. Correspondence and requests for materials should be addressed to J.M.

References

61. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

62. Varadi, M. et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).
63. Zallot, R., Oberg, N. & Gerlt, J. A. The EFI web resource for genomic enzymology tools: leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. *Biochemistry* **58**, 4169–4182 (2019).
64. An, J., Totrov, M. & Abagyan, R. Pocketome via comprehensive identification and classification of ligand binding envelopes. *Mol. Cell. Proteom.* **4**, 752–761 (2005).
65. Blake, L. I. & Cann, M. J. Carbon dioxide and the carbamate post-translational modification. *Front. Mol. Biosci.* **9**, 825706 (2022).

Acknowledgements

We thank the Diamond Light Source for beamtime access (proposal MX31850 Beamline i04) and the Michael Barber Center for Collaborative MS for access to MS instrumentation. This work was funded by the Engineering and Physical Sciences Research Council grant code EP/V048929/1 (J.M.), the Biotechnology and Biological Sciences Research Council grant codes BB/V016083/1 (J.M.) and BB/V008552/1 (J.M.) and the European Commission Marie Skłodowska-Curie Actions fellowship code EP/Y023714/1 (G.X. and J.M.). We also acknowledge the Chinese Scholarship Council for supporting the secondment of D.J. to Manchester from Zhejiang University of Technology and CoEBio3 for a PhD studentship (D.T.).

Author contributions

G.X. and J.M. conceptualized the study. G.X., D.T., S.C.H., D.P., L.Y., R.Z., K.B. and D.J. performed the experimental investigation. A.I. and C.L. performed crystallography. G.X., S.S. and J.M. acquired funding. J.M. and S.S. provided project administration. J.M. provided supervision. G.X. and J.M. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

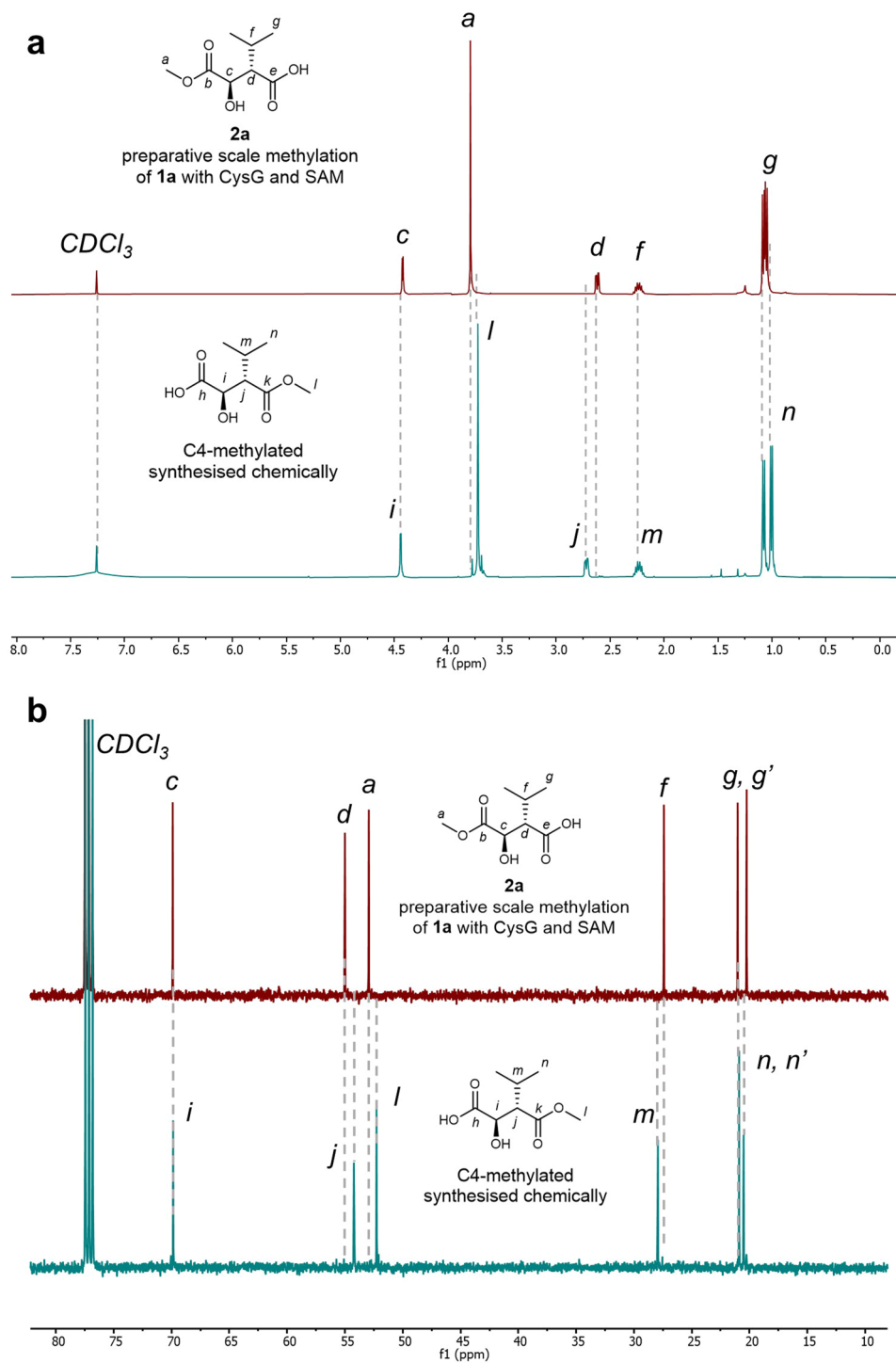
Extended data is available for this paper at <https://doi.org/10.1038/s41589-024-01657-7>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41589-024-01657-7>.

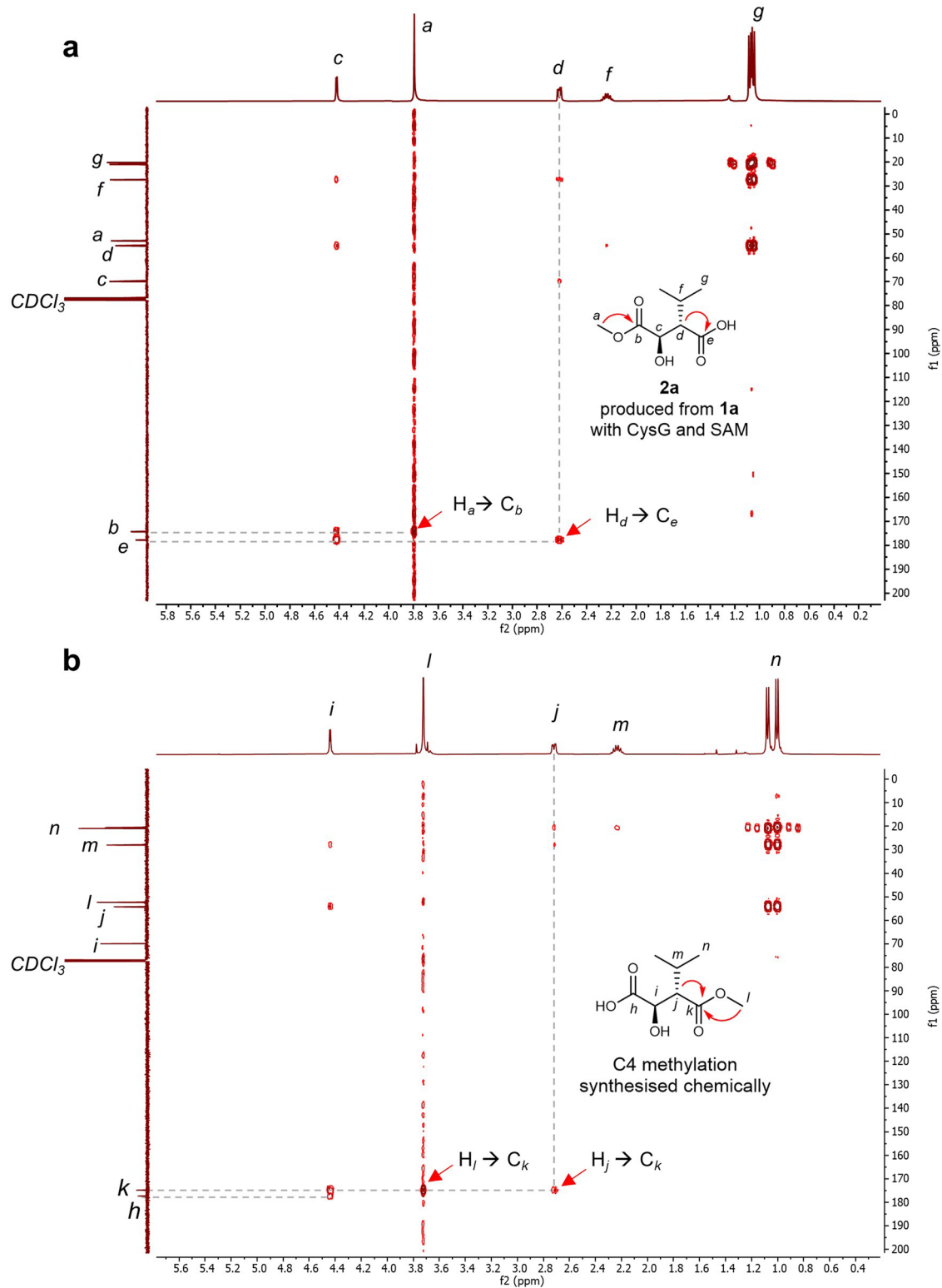
Correspondence and requests for materials should be addressed to Jason Micklefield.

Peer review information *Nature Chemical Biology* thanks Ariane Bertonha, Daniela Trivella and the other, anonymous reviewer(s) for their contribution to the peer review of this work.

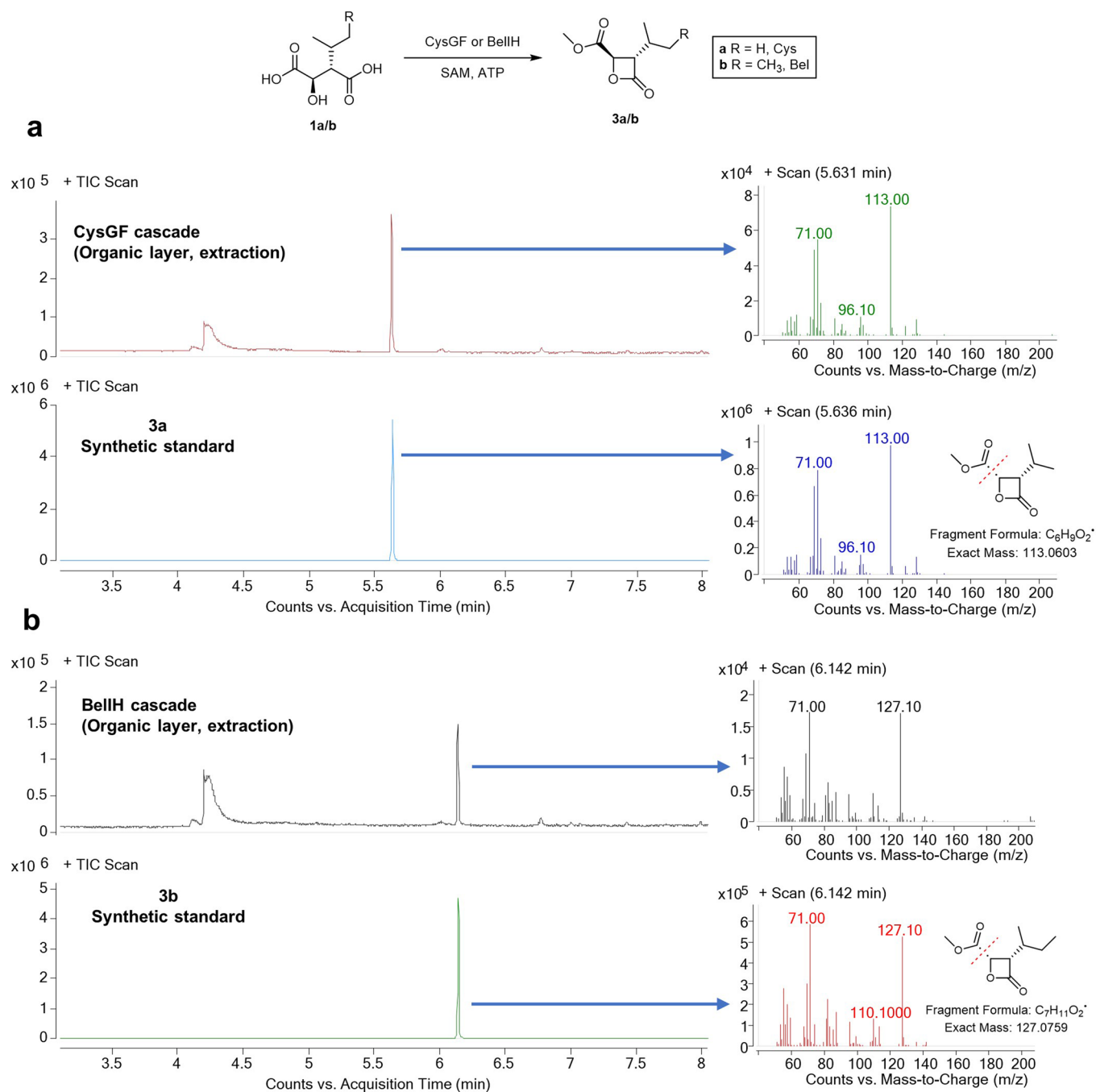
Reprints and permissions information is available at www.nature.com/reprints.



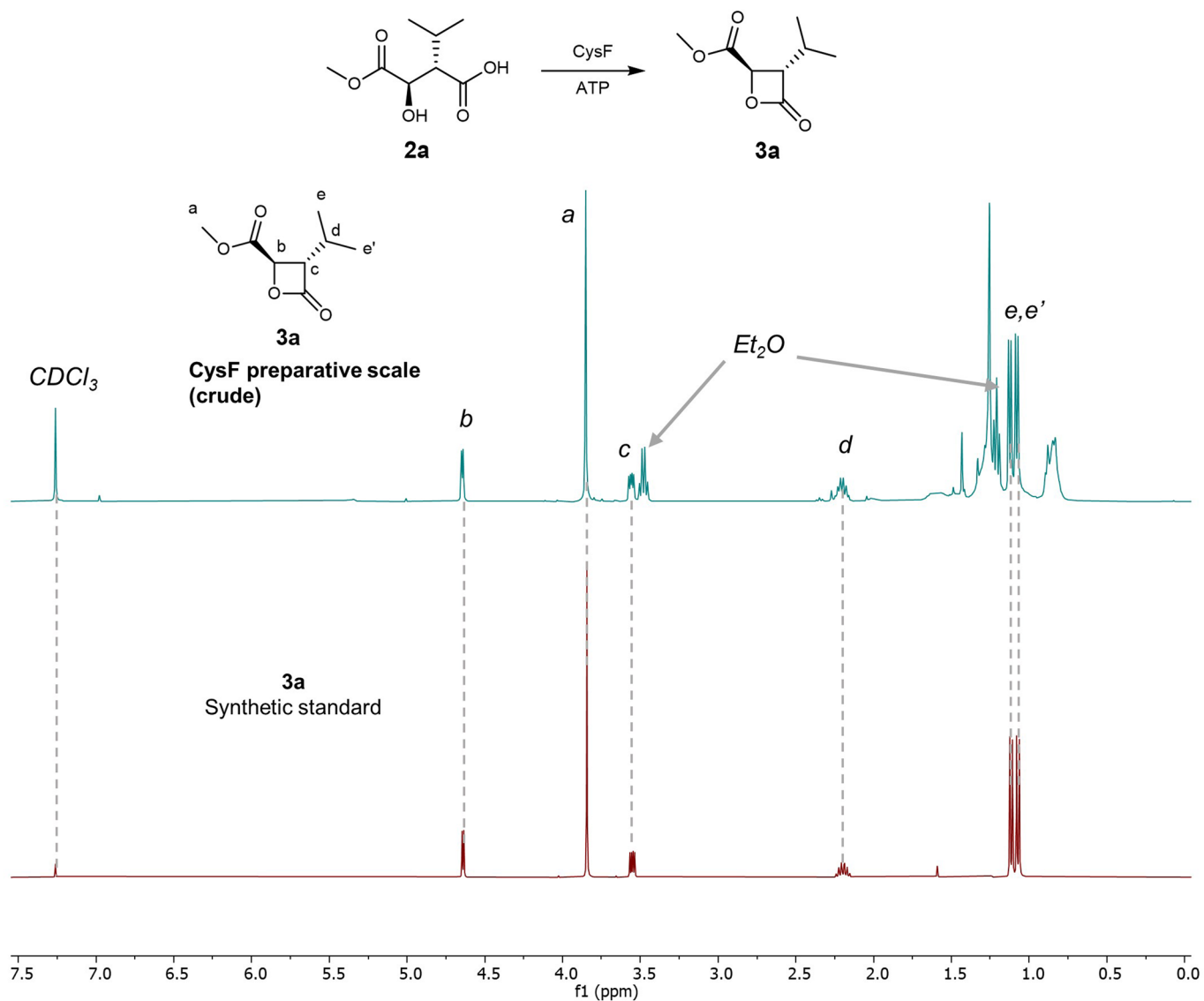
Extended Data Fig. 1 | Comparison of the NMR spectra between enzymatic (CysG) product **2a and the synthetic C4-methyl ester. a, 1H NMR spectra. b, ^{13}C NMR spectra.**



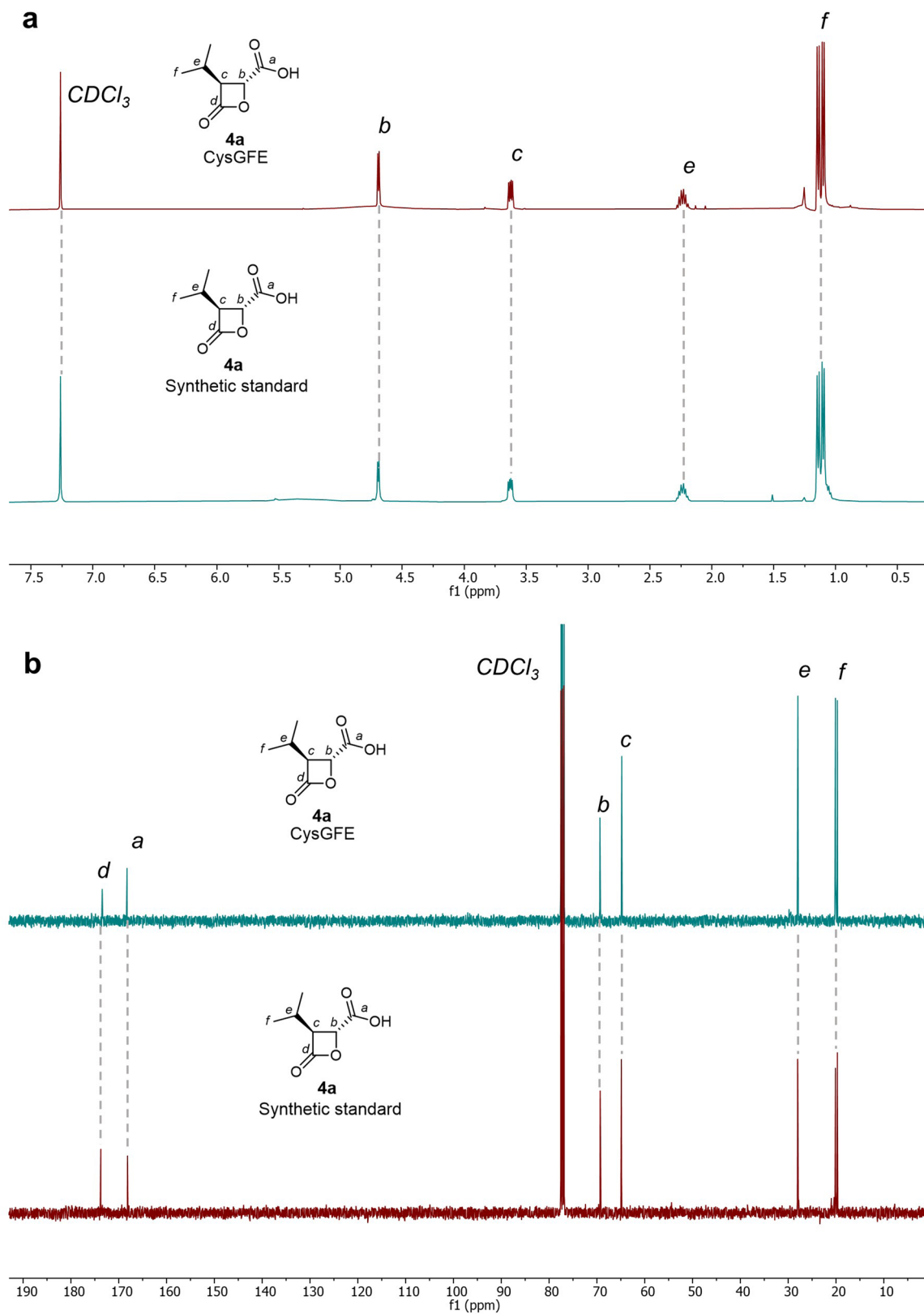
Extended Data Fig. 2 | Comparison of the HMBC spectra of **2a (enzymatic) and C4 methyl ester. a**, HMBC spectra of **2a** produced from **1a** with CysG and SAM. **b**, HMBC spectra of synthetic C4-methyl ester standard.



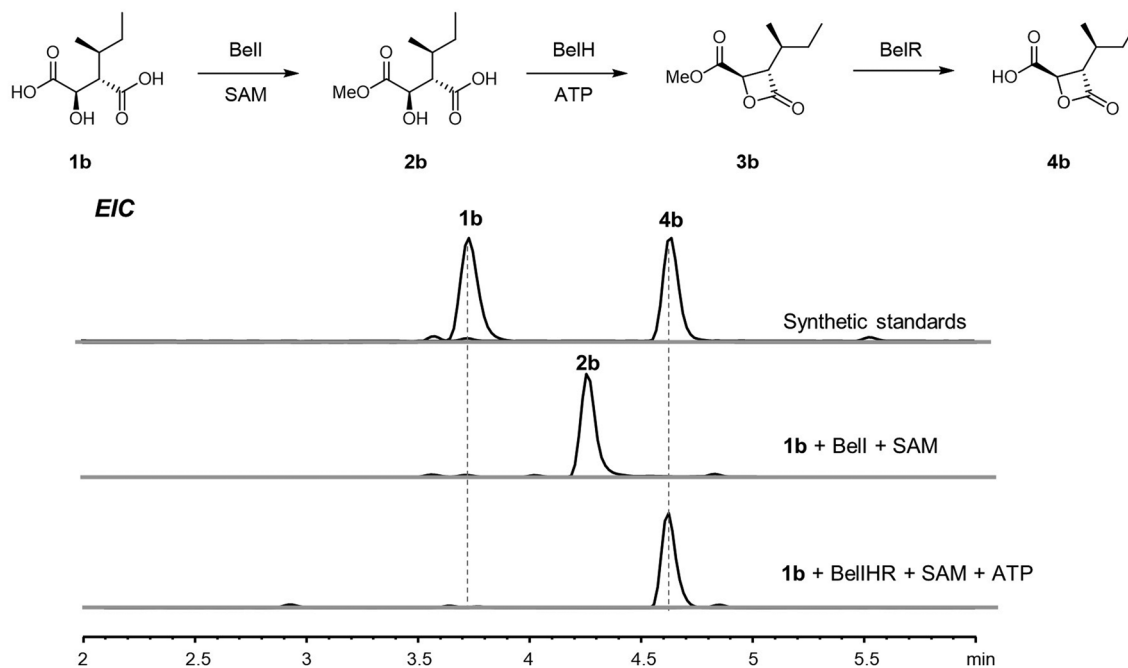
Extended Data Fig. 3 | GC-MS analysis of the CysGF and BellH cascade confirmed the formation of 3a/b. **a**, Comparison of the GC-MS spectra between CysGF cascade and 3a synthetic standard. **b**, Comparison of the GC-MS spectra between BellH cascade and 3b synthetic standard.



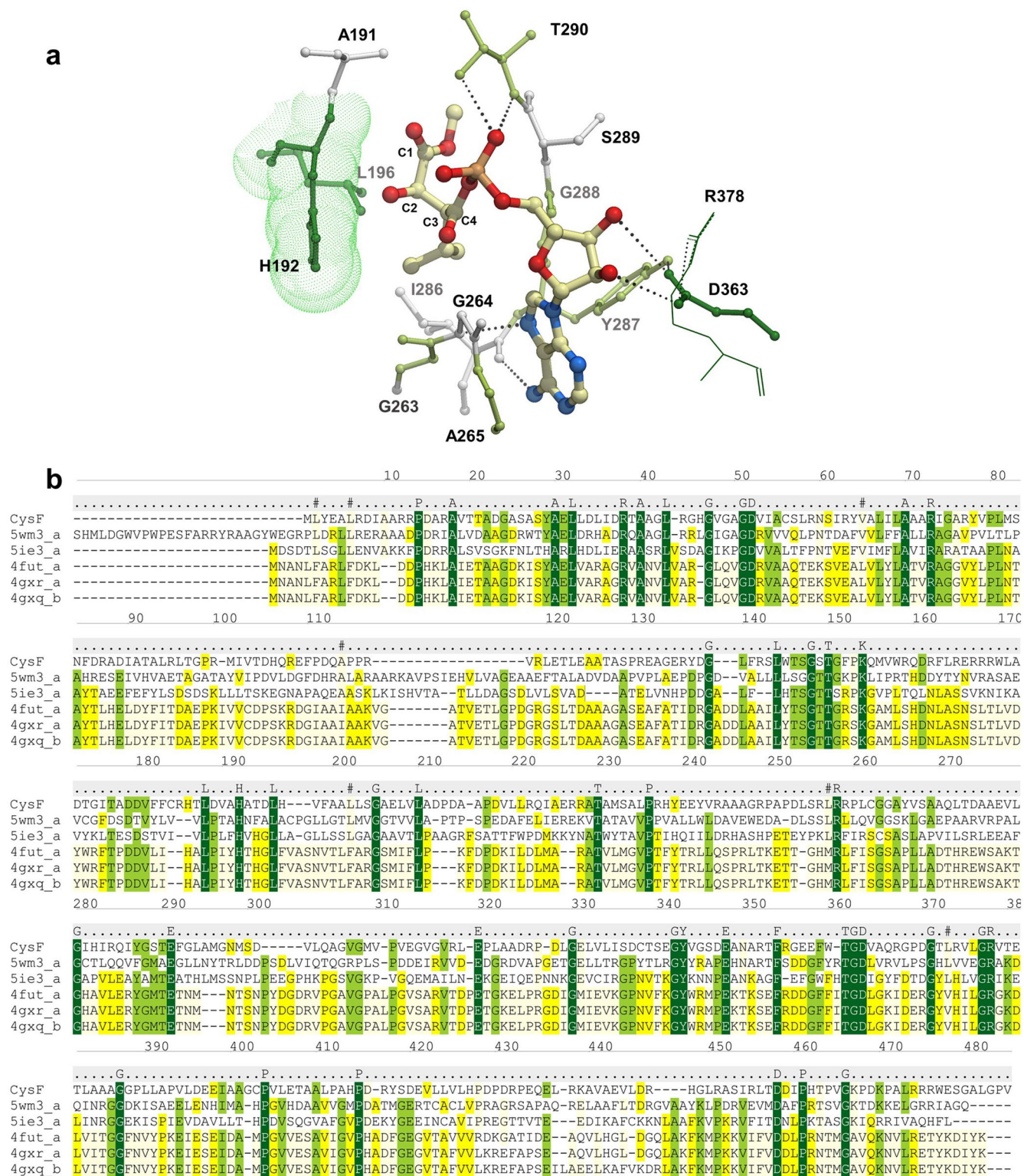
Extended Data Fig. 4 | Comparison of the ¹H NMR of **3a prepared enzymatically and chemically. **3a** was synthesised enzymatically with CysF, **2a**, and ATP, and chemically by methylation of **4a** using TMS-diazomethane (column purified). See supplementary methods for detail experiment procedures.**



Extended Data Fig. 5 | Comparison of the NMR spectra of 4a prepared enzymatically and chemically. 4a was synthesised enzymatically with CysGFE, 1a, SAM, and ATP, and chemically using previously reported procedure³⁷. **a**, ¹H-NMR comparison. **b**, ¹³C-NMR comparison.

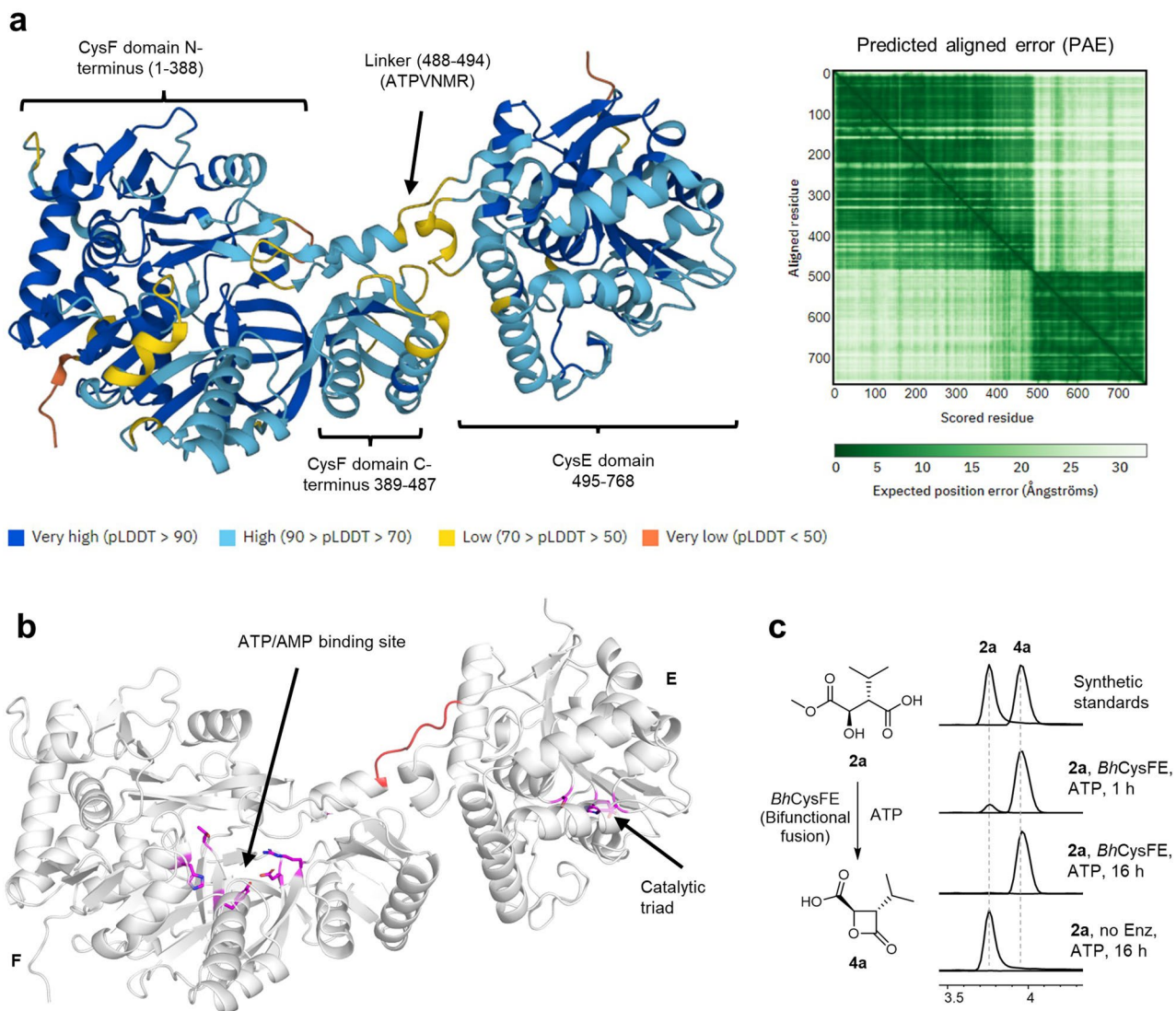


Extended Data Fig. 6 | Bel pathway β -lactone formation LC-MS assay. β -lactone warhead assembly in the bel pathway follows the same methylation-lactonisation-hydrolysis sequence as the cys pathway. See Method section (enzyme assay) for experimental details.



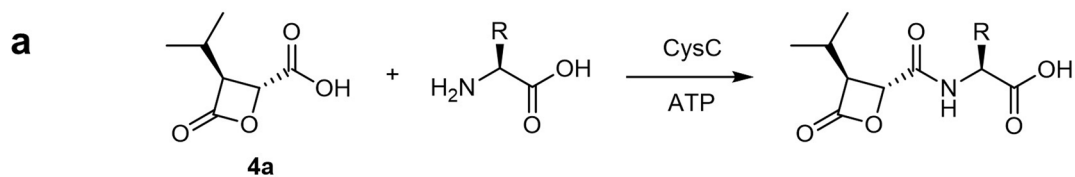
Extended Data Fig. 7 | Docking and sequence alignment of CysF. a, *in silico* docking of 2a-AMP in the crystal structural of CysF. **b**, sequence alignment of CysF with other adenylate-forming enzymes. Pairwise levels of sequence identities between CysF and the related adenylate forming enzymes (5IE3, 4GXR, SWM3, 4GXQ & 4FUT)³⁹⁻⁴³ are between 18–30% and the key residues involved in the binding of AMP are highly conserved (T290, D363, R378). The docked conformation of 2a-AMP suggests H192 may perform the role of a general base during β -lactone formation. The close proximity of the C1-methyl ester to the phosphoryl group, of the docked 2a-AMP, further underscores the potential

for electrostatic repulsion if the C1-carboxylate group was not methylated. Polyethylene glycol (PEG4) is observed to occupy a region of the putative active site in the crystal structure of CysF, which is likely to be an artifact of the crystallisation process. Additional electron density is present in the crystal structure which is consistent with carboxylation resulting in an N-terminal carbamate stabilised by interactions to R131. Carboxylation is suggested to be a common, spontaneous and reversible post translation modification⁶⁵. Given the CysF N-terminus is not located near the active site, it is unlikely that this modification has any effect on the catalytic mechanism of the enzyme.



Extended Data Fig. 8 | AlphaFold structural model and activity assay of bifunctional enzyme *BhCysFE*. **a**, Structural prediction by AlphaFold, AF-A0A562R406-F1-model_v4 (<https://www.alphafold.ebi.ac.uk/entry/A0A562R406>), revealed two domains similar to CysF (aa. 1-487) and CysE (aa. 495-768), connected by a short linker (sequence ATPVNMR) colour coded showing regions with lower (orange) to higher accuracy (blue). **b**, The position

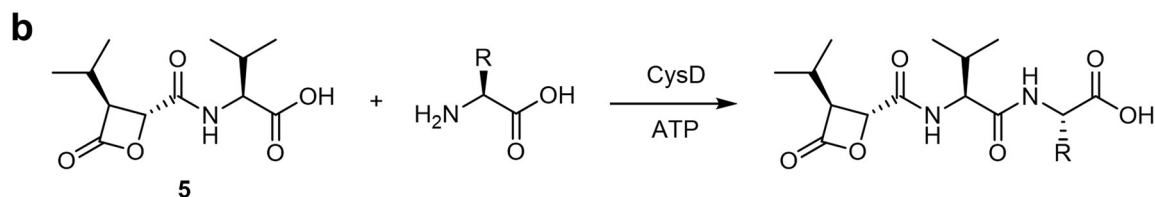
of the predicted active site residues of the F domain are highlighted in magenta based on sequence alignment and docking as described in Extended Data Fig. 7. The predicted catalytic triad present in the E domain Glu607, His739 and Ser582, which is highly conserved in hydrolase enzymes is also highlighted in magenta. **c**, Assay of the bifunctional fusion *BhCysFE* with **2a** results in direct formation of **4a**.



Accepted amino acids (Product confirmed by LCMS) L-amino acids: Val (**5a**), Ile (**44**), Leu (**45**), Met (**46**), Ala (**47**), Ser (**48**), Thr (**49**), Asn (**50**), Arg (**51**), His (**52**), Lys (**53**)

Not accepted
(No product detected)

L-amino acids: Phe, Tyr, Trp, Cys, Gly, Pro, Gln, Asp, Glu
D-amino acids: Val



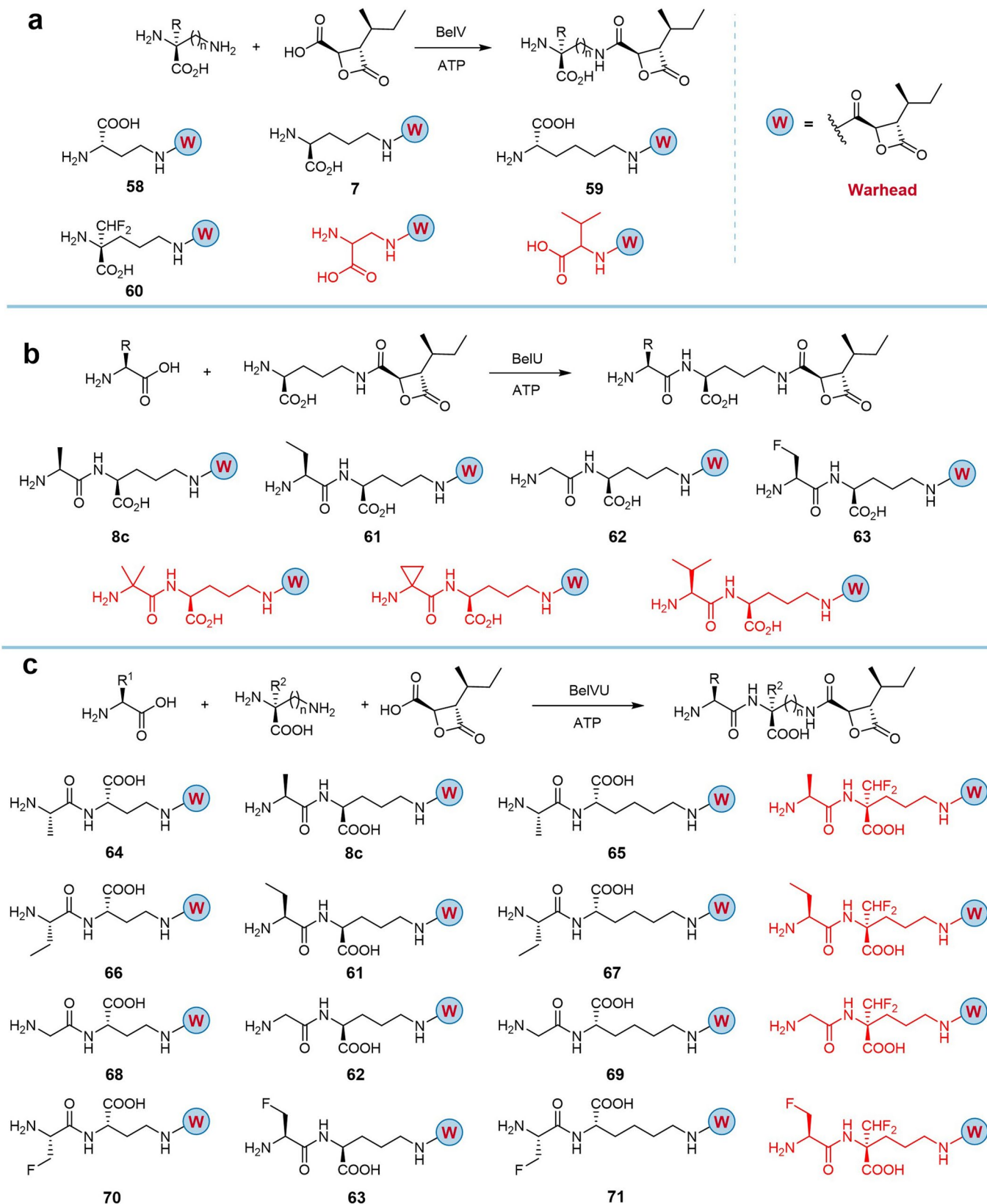
Accepted (Product confirmed by LCMS) L-amino acids: Val (**6b**), Ile (**6a**), Leu (**39**), Met (**54**), Ala (**55**), Gly (**56**), Thr (**57**)

Not accepted
(No product detected)

L-amino acids: Phe, Tyr, Trp, Cys, Pro, Ser, Asn, Gln, Arg, His, Lys, Asp, Glu
D-amino acids: Val

Extended Data Fig. 9 | Proteinogenic amino acids scope of ligase CysC and CysD. **a**, Ligation of natural donor carboxylic acid **4a** with acceptor proteinogenic amino acids catalysed by CysC. **b**, Ligation of natural donor carboxylic acid **5** with acceptor proteinogenic amino acids catalysed by CysD.

Reactions were carried out on analytical scale (see method section enzyme assays). Tandem mass (MS/MS) data of products (in bracket) are shown in supplementary methods (LC-MS/MS data).



Extended Data Fig. 10 | Amino acid scope tested for the bel pathway ligases.

a, Amino acid substrates tested for BelV reaction. **b**, Amino acid substrates tested for BelU reaction. **c**, Substrate scope tested for BelVU cascade reaction. The reaction was analysed by LC-MS. Black colour compounds have anticipated

product mass detected, while the red colour indicated no product mass was detected. For assay conditions see Methods section (enzyme assays). Tandem mass (MS/MS) data of products are shown in supplementary methods: LC-MS/MS data.

Corresponding author(s): Jason MicklefieldLast updated by author(s): May 21, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a | Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection Agilent ChemStation (mass spectrometry + HPLC) C.01.07, TopSpin v3.6.3 & IconNMR (NMR) v5.0.11, GE healthcare UNICORN v7.3

Data analysis MestReNova v11 (NMR), Agilent ChemStation (HPLC) C.01.07, Agilent MassHunter (mass spectrometry) B.07, ChemDraw Professional v21 (chemical structures and exact mass calculations), PyMOL v2.5.7 (protein structure), GE Healthcare UNICORN v7.3, COOT v0.9.8.92, Foldseek, Dali, Yasara, Phenix.refine, Alphafold2

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Data Availability: The coordinates of the CysF X-ray crystal structures have been deposited in the protein data bank (PDB) with ID code 8RA0. The AlphaFold structure of BhCysFE can be accessed on <https://www.alphafold.ebi.ac.uk/entry/A0A562R406> (AF-A0A562R406-F1-model_v4). Structures used for modelling and

docking studies can be accessed on rcsb.org using ID 5BSM, 5BSR, 5WM3, 5IE3, 4FUT, 4GXR, and 4GXQ. All proteins characterised in this study can be accessed on uniprot.org using the accession code presented in Supplementary Table 1-3, and their synthetic gene sequence are provided as source data file. All the remaining data are available in the main text or the supplementary information. Supplementary methods, supplementary figures (1–8), tables (1–4) and references (1-13) are provided in the supplementary information. Correspondence and requests for materials should be addressed to JM.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

| | |
|--|-----|
| Reporting on sex and gender | N/A |
| Reporting on race, ethnicity, or other socially relevant groupings | N/A |
| Population characteristics | N/A |
| Recruitment | N/A |
| Ethics oversight | N/A |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|-----------------|---|
| Sample size | No sample size calculation was used. Quantitative assays were performed in triplicate and mean and standard deviation values calculated. The sample size chosen is based on our past experience with the in vitro experiments and three times were sufficient as the results are highly reproducible. |
| Data exclusions | no data was excluded |
| Replication | Experiments were carried out in at least triplicate and standard deviation provided. Each measurement is from a separate experiment but carried out at the same time. All attempts at replication were successful. |
| Randomization | No randomization was performed during this study as it was not applicable for our experiments. Experiments were performed with the same procedure, controls and data analysis method. |
| Blinding | No blinding was involved in this study as it does not involve animal or human subjects or group allocation. |

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

- n/a | Involved in the study
- Antibodies
 - Eukaryotic cell lines
 - Palaeontology and archaeology
 - Animals and other organisms
 - Clinical data
 - Dual use research of concern
 - Plants

Methods

- n/a | Involved in the study
- ChIP-seq
 - Flow cytometry
 - MRI-based neuroimaging

Plants

Seed stocks

N/A

Novel plant genotypes

N/A

Authentication

N/A