Structures of two distinct conformations of holo-non-ribosomal peptide synthetases

Eric J. Drake1,2, Bradley R. Miller1,2, Ce Shi3, Jeffrey T. Tarasch4, Jesse A. Sundlov1,2, C. Leigh Allen1, Georgios Skiniotis4, Courtney C. Aldrich3 & Andrew M. Gulick1,2

Many important natural products are produced by multidomain non-ribosomal peptide synthetases (NRPSs)1–4. During synthesis, intermediates are covalently bound to integrated carrier domains and transported to neighbouring catalytic domains in an assembly line fashion5. Understanding the structural basis for catalysis with non-ribosomal peptide synthetases will facilitate bioengineering to create novel products. Here we describe the structures of two different holo-non-ribosomal peptide synthetase modules, each revealing a distinct step in the catalytic cycle. One structure depicts the carrier domain cofactor bound to the peptide bond-forming condensation domain, whereas a second structure captures the installation of the amino acid onto the cofactor within the adenylation domain. These structures demonstrate that a conformational change within the adenylation domain guides transfer of intermediates between domains. Furthermore, one structure shows that the condensation and adenylation domains simultaneously adopt their catalytic conformations, increasing the overall efficiency in a revised structural cycle. These structures and the single-particle electron microscopy analysis demonstrate a highly dynamic domain architecture and provide the foundation for understanding the structural mechanisms that could enable engineering of novel non-ribosomal peptide synthetases.

A non-ribosomal peptide synthetase (NRPS) module incorporates a single residue into a peptide natural product. Each module contains a peptidyl carrier protein (PCP) that is post-translationally modified with a phosphopantetheine cofactor6, an adenylation domain that loads the amino-acid substrate onto the PCP cofactor, and a condensation domain that catalyses peptidyl bond formation. NRPSs then use a carboxy (C)-terminal thioesterase or reductase domain to catalyse product release. Structures of individual domains can provide insight into the NRPS structural mechanism. Interestingly, the adenylation domains have been shown to adopt two catalytic conformations7. First, the adenylation domain mediates delivery of the peptide intermediate to the adenylation domain. These results provide views of two distinct steps in the NRPS catalytic cycle and demonstrate how the domain rotation within the adenylation domain mediates the delivery of the PCP between the two catalytic domains.

The structures of AB3403 were determined at 2.7 and 2.9 Å resolution (Extended Data Table 1). No prior structure exists of an NRPS condensation domain bound to a ligand; the holo-AB3403 protein shows the pantetheine cofactor residing in the active site (Fig. 2 and Extended Data Fig. 3a). The two lobes of the condensation domain adopt the closed orientation seen recently in the CDA synthetase condensation domain17. Contacts are made between the pantetheine and the helix running from Glu20 to Leu30, in particular Tyr26 and Ile27, which forms one wall of the tunnel through which the pantetheine approaches the active site (Fig. 2b). Additionally, Tyr37 forms a hydrogen bond with the amide of the cysteine moiety of the pantetheine cofactor. As the main chain carbonyl of Tyr37 hydrogen bonds to the main chain amide of the catalytic His145, this is a critical interaction to close the two lobes and bring the active histidine into proper position.

Holo-AB3403 therefore illustrates the conformation that is adopted to properly deliver the pantetheine of the PCP to the condensation domain. The PCP is rotated ~30° relative to the orientation of the PCP domain of SrfA-C (Extended Data Fig. 4). The AB3403 PCP interface with the condensation domain is composed of residues from helix α2, the helix that follows the pantetheinylation site at Ser1006, and the loops that precede and follow this helix. In particular, residues Phe999 to Tyr1032 face the pantetheine cofactor. As the main chain carbonyl of Tyr37 hydrogen bonds within the active site to the pantetheine cofactor, and the single-particle electron microscopy analysis demonstrate a highly dynamic domain architecture and provide the foundation for understanding the structural mechanisms that could enable engineering of novel non-ribosomal peptide synthetases.

1Hauptman-Woodward Medical Research Institute, 700 Ellicott Street, Buffalo, New York 14203, USA. © 2016 Macmillan Publishers Limited. All rights reserved
2Department of Structural Biology, University at Buffalo, Buffalo, New York 14203, USA.
3Center for Drug Design and Department of Medicinal Chemistry, University of Minnesota, Minneapolis, Minnesota 55455, USA.
4Life Sciences Institute and Department of Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109, USA.
*These authors contributed equally to this work.
condensation domain, which is positioned on an insertion compared with SrfA-C, interacts with the phosphate from the cofactor.

The AB3403 adenylation domain (Fig. 2c) is precisely positioned in the adenylate-forming conformation, unlike the adenylation domain of SrfA-C, which is in an open conformation that may be used for substrate binding or release\(^5\). The lysine of the conserved catalytic A10 motif\(^7,18\) interacts with a phosphate oxygen from AMP and a carboxylate oxygen with glycine and superimposes with the homologous lysine in the gramicidin synthetase domain. In SrfA-C, the homologous lysine is \(\sim 12 \text{ Å} \) away.

The thioesterase domain of AB3403 is structurally similar to the homologous domains of both SrfA-C and EntF (Extended Data Fig. 5), the latter of which has been characterized by NMR and crystallography in complex with the upstream PCP domain\(^9,20\). Despite the similarities in domain structure, the thioesterase domain of AB3403 is in a markedly different location compared with SrfA-C (Fig. 3a). Interestingly, in this new position the thioesterase domain cradles the back face of the PCP domain. The thioesterase domains of SrfA-C or AB3403 do not make substantial contacts with the other catalytic domains.

We next examined the delivery of the holo-PCP to the adenylation domain in a different NRPS protein. We have previously used targeted mechanism-based inhibitors, harbouring a vinylsulfonamide moiety that traps the thioester-forming reaction\(^11\) to characterize functional adenylation-PCP di-domain interactions\(^8,10\). These inhibitors mimic the native aminoacyl adenylate, but contain a Michael acceptor positioned to react with the pantetheine thiol. EntF crystallized only in the presence of the serine adenosine vinylsulfonamide (Ser-AVS) inhibitor (Fig. 2d and Extended Data Fig. 6) that limits conformational flexibility.
the amino (N)-terminal subdomains of the adenylation domain. Alignment of the structures on the basis of
adenylation C-terminal subdomain to the thioester-forming conformation. If EntF is only 780 Å², resulting from the rotation of the ade-
domains suggested they constitute a catalytic platform, upon which the intact protein was present in the
condensation domain omitted for clarity.

The four-stage catalytic cycle of an NRPS module. The pantetheine cofactor is represented by the wavy
dot, and the aminoacyl adenylate intermediate is represented by AA-AMP. The thioester between the amino acid and the
cofactor is shown as S-AA. Finally, the peptide bound to the upstream
carrier protein (purple) is abbreviated Pep. Following the condensation
reaction, the peptide is extended by one amino acid (Pep + 1) and
presented to the thioesterase domain. The revised NRPS structural cycle is
highlighted in yellow showing that only three structural states are required.

Figure 4 | Dynamics of the NRPS cycle. The four-stage catalytic cycle of
an NRPS module. The pantetheine cofactor is represented by the wavy
dot, and the aminoacyl adenylate intermediate is represented by AA-AMP. The thioester between the amino acid and the
cofactor is shown as S-AA. Finally, the peptide bound to the upstream
carrier protein (purple) is abbreviated Pep. Following the condensation
reaction, the peptide is extended by one amino acid (Pep + 1) and
presented to the thioesterase domain. The revised NRPS structural cycle is
highlighted in yellow showing that only three structural states are required.

The basic NRPS catalytic cycle requires that the PCP visits three
adjacent catalytic domains in a coordinated manner. The two cata-
lytic conformations of the adenylation domain require that the full
cycle has four catalytic structural states (Fig. 4). Specifically, (I) the
adenylation domain catalyses amino-acid adenylation, (II) the PCP is
triggered to promote crystallization. Crystals of the EntF protein diffract to 2.8 Å
(Extended Data Table 2). No electron density was observed for the
thioesterase domain although the intact protein was present in the
crystal lattice (Extended Data Fig. 7).

The condensation domain of EntF is similar to the closed AB3403
conformation (Fig. 2a). The adenylation domain adopts the catalytic
thioester-forming conformation of prior adenylation-PCP proteins,
demonstrating that the conformation is compatible with a full NRPS
module. The active site of the EntF adenylation domain identifies
conserved residues (Fig. 2d) that have been shown to play important
catalytic roles in other members of this enzyme superfamily. Arg863
interacts with the cofactor phosphate, while Gly864 and Gln865 form
one wall of the pantetheine tunnel. Interactions with the nucleotide
occurs between Asp840 and the ribose hydroxyls, and between Tyr746
and Tyr852 and the adenine ring. The inhibitor serine binds in the
binding pocket formed by Asp648, Ser722, and Asp754 (Fig. 2d).

The lack of density for the thioesterase domain in EntF suggested
multiple conformations in the crystal lattice. This is not surprising
given the limited interactions in SrfA-C and AB3403 between the
thioesterase domains and the other domains. To assess thioesterase
conformational mobility, we examined EntF by negative-stain electron
microscopy followed by classification and averaging of single-particle
projections (Extended Data Fig. 8). The class averages revealed pri-
marily a tri-lobed density with two neighbouring globular densities of
similar size attributed to the condensation and adenylation domains
and a smaller lobe attributed to the thioesterase domain (Fig. 3b). The
positioning of the thioesterase domain assumes a surprisingly wide
range of distances and angles relative to the other domains.

The large interface of the SrfA-C condensation and adenylation
domains suggested they constitute a catalytic platform, upon which the
other domains move. We therefore compared the interfaces of the
three NRPS modules (Fig. 3c). The interface in AB3403 is 1,023 Å²,
comparable in size to the 1,097 Å² interface of SrfA-C. In contrast, the
interface in EntF is only 780 Å², resulting from the rotation of the ade-
nylation C-terminal subdomain to the thioester-forming conformation.

Additionally, the conformation of the interface is not conserved
between all three proteins. Alignment of the structures on the basis of
the amino (N)-terminal subdomains of the adenylation domain shows
that the condensation domain of both AB3403 and EntF differ slightly
from each other and more significantly from SrfA-C. In AB3403 and
EntF, the condensation domains are rotated by ~25° relative to the
adenylation domains. Furthermore, the EntF condensation domain is
shifted closer towards the adenylation domain. Structural comparisons
suggest that this alternative conformation in EntF may not be compat-
ible with the adenylation-forming conformation. The three different
condensation–adenylation domain conformations, the adenylation-forming
incompatibility seen in EntF, and the multiple extended and compact
conformations seen in the electron microscopy data suggest that the
condensation–adenylation domain platform may be more dynamic
than previously proposed.

The new structures confirm the hypothesis that the adenylation
domain conformational change is a structural mechanism to guide
the PCP between active sites in the context of complete NRPS mod-
ules. The rotation of the adenylation domain C-terminal subdomain
from the adenylation-forming conformation in AB3403 to the thio-
ester-forming conformation of EntF delivers the PCP into the adenylation
domain for loading. The recent structure of loaded holo-PCP has shown
the interaction of the substrate with the PCP core which may help to
promote release of the substrate from the adenylation domain. This
interaction also alters the surface electrostatic potential of regions
that interact with the neighbouring catalytic domains, including α2 and
α3, and may influence the PCP delivery to neighbouring catalytic
domains. Finally, this transfer is further assisted by the linker region
that joins the adenylation C-terminal subdomain with the PCP domain,
which includes important contacts that are preserved in the adenylate-
and thioester-forming conformations, as well as the open conforma-
tion of SrfA-C.

Conformational dynamics in NRPS modules. a, Alternative
locations of the thioesterase domain SrfA-C and AB3403. b, Representative
electron microscopy class averages of EntF. The smaller thioesterase (TE)
domain is observed in various positions relative to the condensation
(C)-adenylation (A) di-domain. Overall EntF adopts a variety of extended
(top) to compact (bottom) conformations. c, The interface between the
condensation C-terminal subdomain and the adenylation domain is shown for
SrfA-C, AB3403, and EntF. The adenylation surface is shown in white,
highlighting in red the regions that interact with the condensation domain.
The right panel shows this interface, rotated by 90° around the y axis, with
the condensation domain omitted for clarity.
delivered to the adenylation domain for thioester-formation to load the PCP, (III) the PCP is delivered to the condensation domain to receive the upstream peptide, and finally (IV) the peptide is delivered to a downstream condensation, thioesterase, or reductase domain for release.

Our results show that states I and III are identical and only three distinct conformations are required to accommodate the four catalytic states of the NRPS cycle (Fig. 4, yellow). The protein first adopts an adenylate-forming conformation, seen in AB3403, state III, to catalyse amino-acid adenylation. Through the domain rotation of the adenylation C-terminal subdomain, the PCP is delivered to the adenylation domain to load the pantetheine cofactor, as seen in the crystal structure of EntF state II. Return of the PCP to the condensation domain delivers the loaded PCP for receipt of the upstream peptide, state III. Critically, as seen in AB3403, the adenylation domain can activate a second amino acid to prime the system for another cycle. The ability to simultaneously catalyse peptide bond formation and amino-acid adenylation at two active sites significantly increases the overall catalytic efficiency and throughput of the NRPS module. Finally, although no structure exists of a full NRPS module with the PCP directed into the thioesterase or other downstream domain in state IV, the structure of AB3403 also offers a new view of the thioesterase domain and suggests the peptide-loaded PCP could be delivered to the downstream thioesterase domain through a simple rotation. The modular architecture of NRPSs as well as their capacity to catalyse unusual chemistry23,24 offer the potential for generating novel products through engineering enzyme activity and the combination of heterologous domains. These efforts have been limited by deficiencies in our understanding of the functional interactions between domains and within active sites. The new views of two essential catalytic states in the NRPS cycle, an appreciation of the greater dynamics of NRPS systems, and the structures of holo-NRPS proteins with relevant ligands will provide the necessary insights to guide these engineering efforts. In addition, these studies complement the recent visualization of modular polyketide synthases by cryo-electron microscopy25 to set the stage for investigations of the structural foundation of even larger, multi-modal biosynthetic proteins.

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. Acknowledgements We thank R. Sanishvili for assistance with data collection. This work was funded in part by National Institutes of Health GM-068440 (to A.M.G.) and GM-115601 (to G.S.), and Award W81XWH-11-2-0218 from the Telemedicine and Advanced Technology Research Center of the US Army Medical Research and Materiel Command (A.M.G.). Data were collected at the GM/CA beamline of the Advanced Photon Source, which is funded by the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006) under Department of Energy contract number DE-AC02-06CH11357 to A.P.S. A Stafford Fellowship (to B.R.M.) and support from the Hauptman-Woodward Institute is acknowledged. Authors Contributions C.L.A. characterized activity of and initially crystallized AB3403. J.A.S. initially crystallized EntF; E.J.D. and B.R.M. optimized crystal, A.M. designed and synthesized the mechanism-based inhibitor. J.T.T. and A.M.G. designed and synthesized the mechanism-based inhibitor. J.T.T. and A.M.G. designed and synthesized the mechanism-based inhibitor.

Supplementary Information is available in the online version of the paper.
Diffraction data were indexed, merged, and scaled using iMOSFLM29 in space group P4_2_2. Structure determination was performed with PHENIX30 using a combination of experimental single-wavelength anomalous diffraction (SAD) phasing and phased molecular replacement. A partial molecular replacement solution was positioned through Phaser with a sculpted (PHENIX sculptor) model derived from PheA (PDB accession number 1AMU)31 and CytC1 (PDB 3VNR). Using this partial molecular replacement model, the selenium sites were identified with the SAD data from SeMet-labelled crystals. An initial model was co-constructed with PHENIX_AutoBuild that contained ~65% of the protein molecule spread across multiple symmetry related molecules. This model was combined into a single protein chain, built and refined iteratively against native data using ARP-WARP32, COOT33, and PHENIX refine.

The final refinements were performed with translation-libration-screwrotation (TLS) parameterization34 with groups consisting of residues 1-191, 191-1445, 446-480, 481-862, 863-959, 960-973, 974-1044, and 1054-1218, roughly defining the NRPS domain (or subdomain) boundaries. The protein is complete from residues Asn2 to Pro1319 with two small disordered loops in the adenylation domain at Asn500–Asp501 and Gly627–Gly630. The latter loop is part of the conserved serine/threonine- and glycine-rich P-loop that is involved in binding the triphosphate of the nucleotide35. Additionally, the condensation domain contains electron density for a diacylglycerol lipid molecule that co-purified with the protein and potentially derived from the bacterial membrane during cell disruption. Diffraction and refinement statistics are presented in Extended Data Table 1. Experimental electron densities of the ligands of both structures are presented in stereo format in Extended Data Fig. 3.

Purification of EntF. The enterobactin biosynthetic cluster of E. coli has been used as a model system in many studies. The full-length EntF, containing the condensation, adenylation, PCP, thioesterase domain architecture, loads serine onto the PCP domain. The condensation domain recognizes the external carrier molecule Ent that has been loaded with 2,3-dihydroxybenzoate (DHB) by the activity of the freestanding adenylation domain EntE. The DHB-serine amide is then transferred to the thioesterase domain while two additional cycles of synthesis complete the enterobactin trilactone.

The EntF protein used in this study (GenBank P11454) was described previously24,25. The entf gene was PCR amplified from E. coli JM109 and cloned into a PET15b-TEV expression vector27 and confirmed by DNA sequencing. The vector provides a His6-tag, linker, and tobacco etch virus (TEV) protease recognition site that, upon treatment with TEV protease, yields a final recombinant product with glycine and histidine preceding the initial methionine residue.

The AB3403 PET15b-TEV construct was transformed into E. coli (BL21-DE3) cells. Transformed cells were grown in LB media to an absorbance at 600 nm (A600 nm) of 0.6 at 37 °C. Protein expression was induced by addition of 0.5 mM isopropyl-β-D-thiogalactoside (IPTG) and cells were incubated overnight at 16 °C. Cells were harvested by centrifugation, flash-frozen in liquid nitrogen, and stored at −80 °C. Selenomethionine-labelled protein was generated in M9 minimal media using a metabolic inhibition protocol25. All purification steps were identical to the native protein.

For purification, cells were resuspended in a buffer containing 50 mM HEPES (pH 7.5), 250 mM NaCl, 10 mM imidazole, 0.2 mM TCEP. Cells were lysed by mechanical disruption (Bomem Sonicator) and the resulting lysate was clarified by centrifugation at 235,000g for 45 min. The cell lysate was passed over a His-trap (GE Healthcare) immobilized metal ion affinity column and washed with lysis buffer containing 50 mM imidazole. Bound proteins were eluted with the same buffer containing 300 mM imidazole. The protein was incubated with TEV protease and dialysed against a TEV cleavage buffer (50 mM HEPES (pH 8.0), 250 mM NaCl, 0.2 mM TEV, and 0.5 mM EDTA) for 16 h at 4 °C. This purified protein was then phosphopantetheinylated by incubation with His6-tagged non-specific phosphopantetheinyl transferase Sfp (10 nM), 12.5 mM MgCl2, and 1 mM CoA for 60 min at 20 °C. The clarified protein was then passed over the His-trap column a second time to remove uncleaved protein, the TEV protease, Sfp, and other contaminating proteins. The holo-AB3403 protein in the column flow-through was pooled, dialysed against a size exclusion buffer containing 50 mM HEPES (pH 7.5), 150 mM NaCl, 0.2 mM TCEP, and further purified by gel filtration (Superdex200). Protein concentration was assessed after dialysis against a crystallization buffer (25 mM HEPES (pH 7.5), 50 mM NaCl, 0.2 mM TCEP) using an extinction coefficient at 280 nm of 157,570 M−1 cm−1.

Crystallization conditions for holo-AB3403 were identified from a sparse matrix screen at 20 °C. Final crystals for native and SeMet-labelled holo-AB3403 (95% purity) were grown at 14 °C by hanging-drop vapour diffusion for 45 min. A high-resolution region of 0.75–0.95 M potassium citrate, 0.01–0.025 M glycine, and 0.05–0.25 M tri-sodium citrate (BTP) (pH 8.0). Highest-quality native crystals were obtained using a protein concentration of 5.5 mg ml−1 with a protein to cocktail ratio of 1:5.1. SeMet protein was crystallized in the same manner with a protein concentration of 7.5 mg ml−1 and 1:1 protein to cocktail ratio. To obtain crystals in the presence of ligands, the protein was pre-incubated for 45 min at 4 °C with 2 mM MgCl2 and 1.5-fold molar excess of ATP and glycine.

Structure determination of AB3403. Crystals of holo-AB3403 were cryoprotected by stages using either ethylene glycol or potassium citrate for native and SeMet protein. Holo-AB3403 was cryoprotected to 30% ethylene glycol (BTP) (pH 8.0). Highest-quality native crystals were obtained using a protein concentration of 1.5 mg ml−1 with a protein to cocktail ratio of 1.5:1. SeMet protein was crystallized in the same manner with a protein concentration of 7.5 mg ml−1 and 1:1 protein to cocktail ratio. To obtain crystals in the presence of ligands, the protein was pre-incubated for 45 min at 4 °C with 2 mM MgCl2 and 1.5-fold molar excess of ATP and glycine.

Crystallization conditions for the Ser-AVS inhibited EntF were first identified using the Hauptman-Woodward high-throughput screen36. Large diffraction-quality native crystals were grown using hanging drop vapour diffusion at 20 °C. A crystallization cocktail, consisting of 100 mM BTP pH 7.5, 125–150 mM MgCl2, and 22–28% PEG 4000, was diluted 1:1 with the final dialysis buffer. The hanging drops then combined protein at 30 mg ml−1 and the undiluted crystallization cocktail at 1:2. The ‘batch mimick’ limited the differences between the drop and reservoir and has been successful with other protein samples in our laboratory36.
glycerol to a final concentration of ~20%. Diffraction data were collected on APS beamline 23-IDB using the rastering option to find the optimal spots on both the native and the SeMet crystals. Diffraction data were indexed, merged, and scaled using iMOSFLM in space group P4_2_2_1. Structure determination for the SeMet inflection data was performed in PHENIX using a PhaserEP MR-SAD with a partial molecular replacement solution that was obtained using a sculpted model generated from the Pseudomonas aeruginosa bidomain adenylation-PCP protein PA1221 (PDB 4DG9). Automated model building with BUCCANEER led to 65% of the structure being refined. This partial model from the SeMet data was used as a molecular replacement model for the native data, and the remaining portion of the protein was built by hand (excluding the thiosterase domain, which was unresolved and constituted about 19%). This model was built and refined iteratively using COOT and PHENIX. TLS refinement was used in final stages with groups consisting of residues 5186, 187-429, 430-444, 445-857, 858-964, 965-971, and 972-1045.

The final model showed density for the condensation, adenylation, and PCP domains of EntF; no density was observed for the thiosterase domain. Diffraction and refinement statistics are presented in Extended Data Table 2.

In general, the overall quality of the density was weaker for the N-terminal subdomain of the condensation domain, residues 1–186, probably reflecting the higher mobility of this region of the protein. The average B-factors for different regions of the protein (Extended Data Table 2) support this conclusion.

Negative-electron microscopy analysis of EntF. EntF, purified as described above, was prepared for electron microscopy using the conventional negative staining protocol, and imaged at room temperature with a Tecnai T12 electron microscope operated at 120 kV using low-dose procedures. Images were recorded at a magnification of ×71,138 and a defocus value of ~1.5 μm on a Gatan US4000 CCD camera. All images were binned (2 pixels × 2 pixels) to obtain a pixel size of 4.16 Å on the specimen level. Particles were manually excised using e2boxer (part of the EMAN 2 software suite). Two-dimensional reference-free alignment and classification of particle projections was performed using ISAC40. A total of 17,431 projections of EntF were subjected to ISAC, producing 133 classes consistent over two-waing imaging and accounting for 53,443 particle projections (Extended Data Fig. 8B).

Synthesis of serine adenosine vinylsulfonamide. Ser-avs was synthesized using the protocol summarized in (Extended Data Fig. 6). All reactions were performed under an inert atmosphere of dry Ar in oven-dried (150°C) glassware. 1H and 13C NMR spectra were recorded on a Varian 600 MHz spectrometer. Proton chemical shifts are reported in parts per million from an internal standard of residual chloroform (7.26 ppm). Proton chemical data are reported as follows: chemical shift, coupling constant. High-resolution mass spectra were obtained on an Agilent TOF mass spectrometer with ESI or APCI interface. Thin-layer chromatography (TLC) analyses were performed in toluene-ethyl acetate-hexanes as a solvent system. The purity of products was determined by 1H NMR (400 MHz, CD3OD) and 13C NMR (100 MHz, CD3OD). Standard reaction conditions contained 50 mM HEPES (pH 7.5), 100 mM MgCl2, 1 mM 2-oxoglutarate, 1 mM vinylsulfonamide, and 5 mM substrate. Reactions (100 μl) were preincubated for 10 min at 37°C, then quenched with 0.5 mM 1,2- diol, 0.1 mM unlabeled PPI, and 0.35 mM perchloric acid. The charcoal was pelleted by centrifugation, washed twice with 1 ml H2O, and resuspended in 0.5 ml H2O for scintillation counting.

To determine the apparent kinetic constants for ATP and glycine for the holo-AB3403 adenylation domain, the NADH oxidase activity monitored at 340 nm (ref. 44) and full-length AB3403. Hydroxylamine was used as a surrogate for the pantheine in the second partial reaction to displace AMP for use in the coupled consumption assay. Standard reactions contained 50 mM HEPES (pH 7.5), 15 mM MgCl2, 2 mM ATP, 3 mM phosphoenolpyruvate, 0.2 mM NADH, 5 U myokinase, 5 μM pyruvate kinase, 6.5 U lactate dehydrogenase, and 150 μM buffered hydroxylamine. Apparent kinetic constants were determined by varying concentrations of ATP or glycine with the one or the other in excess. Reactions were initiated by the addition of 0.001 mM enzyme. Calculations were done using PRISM software.

26. Adams, M. D. et al. Comparative genome sequence analysis of multidrug-resistant Acinetobacter baumannii ATCC 19606. J. Bacteriol. 190, 8053–8064 (2008).
27. Doublier, S. Preparation of selenomethionyl proteins for phase determination. Methods Enzymol. 276, 523–530 (1997).
28. Battey, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with cerim. J. Struct. Biol. 166, 213–221 (2010).
29. Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nature Protocols 3, 1171–1179 (2008).
30. Emley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta Crystallogr. D 66, 212–221 (2010).
31. Uznanski, A., Atkinson, P. V. & Adams, P. D. TLS from fundamentals to practice. Crystallogr. Rev. 19, 230–270 (2013).
32. Roche, E. D. & Walsh, C. T. Dissection of the EntF condensation domain boundary and active site residues in nonribosomal peptide synthesis. J. Biol. Chem. 242, 1334–1344 (2005).
33. Jare, J. R. et al. A deliberate approach to screening for initial crystallization conditions of biological macromolecules. J. Struct. Biol. 142, 170–179 (2003).
34. Sundlov, J. A. & Gulick, A. M. Structure determination of the functional interaction domain of a chimeric nonribosomal peptide synthetase from a challenging crystal with noncrystallographic translational symmetry. Acta Crystallogr. D 69, 1482–1492 (2013).
35. Lowe, H. S. & Cowtan, K. B Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr. D 62, 1002–1011 (2006).
36. Ohi, M., Li, Y., Cheng, Y. & Walz, T. Negative staining and image classification—powerful tools in modern electron microscopy. BioL Proced. Online 6, 23–34 (2004).
39. Tang, G. et al. EMAN2: an extensible image processing suite for electron microscopy. J. Struct. Biol. 157, 38–46 (2007).
40. Yang, Z., Fang, J., Chittuluru, J., Asturias, F. J. & Penczek, P. A. Iterative stable alignment and clustering of 2D transmission electron microscope images. Structure 20, 237–247 (2012).
41. Ikeuchi, H., Meyer, M. E., Ding, Y., Hiratake, J. & Richards, N. G. A critical electrostatic interaction mediates inhibitor recognition by human asparagine synthetase. Bioorg. Med. Chem. 17, 6641–6650 (2009).
42. Rusnak, F., Faraci, W. S. & Walsh, C. T. Subcloning, expression, and purification of the enterobactin biosynthetic enzyme 2,3-dihydroxybenzoate-AMP ligase: demonstration of enzyme-bound (2,3-dihydroxybenzoyl)adenylate product. Biochemistry 28, 6827–6835 (1989).
43. Horswill, A. R. & Escalante-Semerena, J. C. Characterization of the propionyl-CoA synthetase (PrpE) enzyme of Salmonella enterica: residue Lys592 is required for propionyl-AMP synthesis. Biochemistry 41, 2379–2387 (2002).
44. Reger, A. S., Carney, J. M. & Quink, A. M. Biochemical and crystallographic analysis of substrate binding and conformational changes in acetyl-CoA synthetase. Biochemistry 46, 6536–6546 (2007).
45. Wilson, D. J. & Aldrich, C. C. A continuous kinetic assay for adenylation enzyme activity and inhibition. Anal. Biochem. 404, 96–63 (2010).
46. Reuter, D. C., McIntosh, J. E., Guinn, A. C. & Madera, A. M. Synthesis of vinyl sulfonamides using the Horner reaction. Synthesis 2003, 2321–2324 (2003).
Extended Data Figure 1 | Structure-based alignment of EntF, AB3403, and SrfA-C. Condensation, adenylation, PCP, and thioesterase domains are represented with bars in grey, pink, green–cyan, and blue. Conserved motifs and catalytically important residues are highlighted with the same colours, including the HHxxxD motif of the condensation domains, the aspartic acid hinge that separates the N- and C-terminal subdomains of the adenylation domain, the GGHS motif that is the site of pantetheinylation in the PCP, and the catalytic nucleophile of the thioesterase domain. The SrfA-C, AB3403, and EntF proteins share approximately 26% sequence identity. The adenylation and PCP domains are more well-conserved, sharing ~35% identity, whereas the condensation (21%) and thioesterase (25%) domains are less well conserved. Domain boundaries are described in the table below.

| Domain     | EntF | AB3403 | SrfA-C |
|------------|------|--------|--------|
| Condensation | 1-429 | 1-452 | 1-437 |
| Adenylation  | 446-964 | 467-964 | 452-962 |
| PCP         | 972-1052 | 972-1048 | 969-1044 |
| Thioesterase | 1055-1293 | 1051-1317 | 1047-1275 |
| Hinge       | Asp857 | Asp866 | Asp861 |
| PPant       | Ser1006 | Ser1006 | Ser1003 |
Extended Data Figure 2 | Substrate specificity of full-length AB3403.
Amino-acid specificity of AB3403 was recorded for all 20 proteinogenic amino acids, as well as 4-chlorobenzoate (4CB) and 4-hydroxybenzoate (4HB). Average values and standard deviations are shown for three replicates with each substrate; results were recorded as micromoles of radiolabelled ATP incorporated per minute per milligram of enzyme. Apparent kinetic constants are also shown for ATP and glycine calculated from duplicate measurements for four to six substrate concentrations.

| Substrate | $k_{cat}$ (min$^{-1}$) | $K_M$ (μM) | $k_{cat}/K_M$ (μM$^{-1}$ min$^{-1}$) |
|-----------|-----------------|------------|---------------------------------|
| Glycine   | 3.6             | 1117 ± 16  | 3.2 x 10$^2$                    |
| ATP       | 2.5             | 375 ± 11   | 6.6 x 10$^2$                    |
Extended Data Figure 3 | Stereo representations of electron density figures shown in Fig. 2. To better visualize the active sites and electron density quality, stereo figures are included in the extended data. In all panels, density is shown with coefficients of the form \((F_o - F_c)\) calculated before inclusion of ligands and contoured at 3σ. 

a, Stereo representation of electron density of AB3403 condensation domain shows the phosphopantethine on Ser1006 approaching His145 within the condensation domain pocket. Inhibitor carbon atoms in green, carbons of residues within 5 Å of inhibitor in grey, nitrogen in blue, oxygen in red, sulphur in yellow, and water in light blue.

b, Electron density of the nucleotide binding pocket of AB3403 bound to glycine and AMP. Stereo representation of electron density shows the AMP, glycine, and Mg\(^{2+}\) present in the active site of the adenylation domain. Ligand carbon atoms are in green, carbons of residues within 5 Å of inhibitor in grey, nitrogen in blue, oxygen in red, phosphorus in orange, and the Mg\(^{2+}\) cofactor in purple.

c, Stereo representation of the electron density shows the phosphopantethine on Ser1006 covalently attached to the Ser-AVS inhibitor in the active site of the adenylation domain. Inhibitor carbon atoms in green, carbons of residues within 4 Å of inhibitor in grey, nitrogen in blue, oxygen in red, phosphorus in orange, sulphur in yellow, and water in light blue.
Extended Data Figure 4 | Comparison of AB3403 and SrfA-C PCP-condensation domain interaction. Stereo representation illustrating different orientations of the PCP domains of SrfA-C and AB3403 relative to the condensation domains with which they interact. AB3403 is shown with a white condensation domain and a green-cyan PCP. SrfA-C is shown with a yellow condensation domain and a pale blue PCP. The pantetheine of AB3403 is shown bound to Ser1006. The position of Ser1003, mutated to an alanine residue in SrfA-C, is also highlighted.
Extended Data Figure 5 | Comparison of AB3403 thioesterase domain to the functional PCP–thioesterase interaction. Stereo representation of the thioesterase (blue) domain of AB3403 interacts with the back face of the PCP domain in AB3403. The functional interaction between the EntF thioesterase domain and its holo-PCP, trapped crystallographically, illustrates that the same face of the thioesterase domain interacts functionally (PDB 3TEJ). A 28-residue insertion of AB3403 is coloured yellow.
Extended Data Figure 6 | Synthesis of Ser-AVS. The Ser-AVS probe was synthesized following similar protocols described elsewhere\textsuperscript{41,46}. Garner’s aldehyde 1 was coupled with 2 using LiHMDS to exclusively furnish the (E)-vinylsulfonamide 3. Mitsunobu coupling of 3 with bis-Boc adenosine 4 afforded 5, which was globally deprotected using 80% aqueous trifluoroacetic acid to yield Ser-AVS.
Extended Data Figure 7 | Electrophoretic mobility of EntF. a, Native gel electrophoresis. Lane 1: EntF. Lane 2: EntF incubated with fourfold molar excess of Ser-AVS inhibitor. Lane 3: EntF Crystals. Lane 4: novex NativeMark labelled in kilodaltons. b, Denaturing gel electrophoresis using loading buffer with SDS and β-mercaptoethanol. Gel lane 1: EntF Lane 2: EntF incubated four times with Ser-AVS inhibitor. Lane 3: Life Technologies Mark12 labelled in kilodaltons. The native gel shows the inhibited EntF in a more compact conformation compared with EntF without the inhibitor.
Extended Data Figure 8 | Negative-stain electron microscopy analysis of EntF. a, Raw electron microscopy image of negative-stained EntF. b, Class averages of EntF particles.
Extended Data Table 1 | Diffraction data statistics and refinement statistics for AB3403

|                          | SeMet_AB3403 | AB3403 | Liganded AB3403 |
|--------------------------|--------------|--------|-----------------|
| **PDB Code**             | 4ZXH         | 4ZHI   |
| **Beamline**             | APS 23-ID-B  | APS 23-ID-B |
| **Wavelength (Å)**       | 0.9793       | 0.9796 | 1.0332          |
| **Space group**          | P4₃2·2       | P4₃2·2 | P4₃2·2          |
| **Unit cell a, b, c (Å)**| 116.19 116.19 348.61 | 116.19 116.19 348.61 | 116.10 116.10 342.02 |
| **Molecules per ASU**    | 1            | 1      | 1               |
| **Resolution range (Å)** | 29.75–3.35   | 49.80–2.70 | 45.03–2.90     |
|                          | (3.52–3.35)  | (2.79 - 2.70) | (3.00 – 2.90) |
| **Total reflections**    | 137397 (16096) | 416923 (21743) | 257325 (25582) |
| **Unique reflections**   | 34599(4299)  | 66559 (6495) | 52900 (5187)   |
| **Multiplicity**         | 4.0 (3.7)    | 6.3 (3.4) | 4.9 (4.9)      |
| **Completeness (%)**     | 98.9 (94.6)  | 99.96 (100.00) | 99.99 (100.00) |
| **Mean I/σ(I)**          | 11.9 (3.5)   | 9.91 (2.49) | 8.47 (2.19)    |
| **R<sub>merge</sub>**    | 0.090 (0.359) | 0.125 (0.511) | 0.130 (0.641) |
| **R<sub>meas</sub>**     | 0.116        | 0.143  | 0.162           |
| **CC<sub>1/2</sub>**    | 0.993 (0.798) | 0.991 (0.685) | 0.991 (0.635) |
| **CC<sup>+</sup>**       | 0.998 (0.937) | 0.998 (0.902) | 0.998 (0.881) |

**Structure Refinement**

|                          |              |          |                |
|--------------------------|--------------|----------|----------------|
| **R<sub>factor</sub>**   | 0.179 (0.248) | 0.174 (0.307) |
| **R<sub>free</sub>**     | 0.234 (0.322) | 0.225 (0.369) |
| **No. atoms**            | 10301        | 10198    |
| **RMSD bond distances (Å)** | 0.009      | 0.009   |
| **RMSD bond angles**     | 1.18         | 1.18     |
| **Wilson B-factor (Å<sup>2</sup>)** | 41.15  | 50.83   |
| **Average B-Factor (Å<sup>2</sup>)** |            |          |
| | Protein | 46.00 | 53.60 |
| | Ligand   | 54.40 | 55.30 |

**Ramachandran analysis**

|                          |              |          |                |
|--------------------------|--------------|----------|----------------|
| **Favored (%)**          | 97.0         | 96.0     |
| **Allowed (%)**          | 2.3          | 3.4      |
| **Outliers (%)**         | 0.7          | 0.6      |
| **Molprobity ClashScore** | 4.51      | 6.75     |

Values in parentheses are for the highest resolution shell.
Extended Data Table 2 | Diffraction data statistics and refinement statistics for EntF

|                          | SeMet EntF | EntF  |
|--------------------------|------------|-------|
| **PDB Code**             | 4ZXJ       |       |
| **Beamline**             | APS 23-ID-B| APS 23-ID-B |
| **Wavelength**           | 0.9796     | 1.0332|
| **Space group**          | P4_2_2     | P4_2_2|
| **Unit cell a, b, c (Å)**| 127.55 127.55 186.72 | 127.71 127.71 186.94 |
| **Molecules per ASU**    | 1          | 1     |
| **Resolution range (Å)** | 60 – 2.9 (3.0 – 2.9) | 81.31 – 2.8 (2.9 – 2.8) |
| **Total reflections**    | 152578 (15129) | 175399 (17288) |
| **Unique reflections**   | 34693 (3380) | 38753 (3800) |
| **Multiplicity**         | 4.4 (4.5)  | 4.5 (4.5) |
| **Completeness (%)**     | 99.66 (99.41) | 99.96 (99.92) |
| **Mean I/σ(I)**          | 9.86 (2.49) | 9.85 (2.11) |
| **R-merge**              | 0.1153 (0.6165) | 0.0979 (0.6484) |
| **R-meas**               | 0.1312 | 0.1109 |
| **CC1/2**                | 0.995 (0.598) | 0.997 (0.629) |
| **CC* (Å²)**             | 0.999 (0.865) | 0.999 (0.879) |

**Structure Refinement**

|                          |       |
|--------------------------|-------|
| **R_factor**             | 0.183 (0.290) |
| **R_free**               | 0.230 (0.324) |
| **No. protein/ligand atoms** | 7898/49 |
| **RMSD bond distances (Å)** | 0.008 |
| **RMSD bond angles**     | 1.23 |
| **Wilson B-factor (Å²)** | 62.88 |
| **Average B-Factor (Å²)** |       |
| Protein                  | 74.6 |
| Ligand                   | 49.8 |
| Condensation N-terminal Subdomain | 90.5 |
| Condensation C-terminal Subdomain | 84.8 |
| Adenylation N-terminal Subdomain | 54.9 |
| Adenylation C-terminal Subdomain | 54.8 |
| PCP Domain               | 78.8 |
| **Ramachandran analysis**|       |
| Favored (%)              | 95.0 |
| Allowed (%)              | 4.13 |
| Outliers (%)             | 0.87 |
| **Molprobity ClashScore**| 6.03 |

Values in parentheses are for the highest resolution shell.