Published in final edited form as:

Nature. 2016 January 14; 529(7585): 235–238. doi:10.1038/nature16163.

Structures of Two Distinct Conformations of *holo*-Nonribosomal Peptide Synthetases

Eric J. Drake[‡,1,2], **Bradley R. Miller**[‡,1,2], **Ce Shi**[3], **Jeffrey T. Tarrasch**[4], **Jesse A. Sundlov**[1,2], **C. Leigh Allen**[1], **Georgios Skiniotis**[4], **Courtney C. Aldrich**[3], and **Andrew M. Gulick**[1,2,*]

[1]Hauptman-Woodward Medical Research Institute, 700 Ellicott St. Buffalo, NY 14203

[2]Department of Structural Biology, University at Buffalo. Buffalo, NY 14203

[3]Center for Drug Design and Department of Medicinal Chemistry, University of Minnesota, Minneapolis, MN 55455

[4]Life Sciences Institute and Department of Biological Chemistry, University of Michigan, Ann Arbor MI 48109

Abstract

Many important natural products are produced by multidomain nonribosomal peptide synthetases (NRPSs)[1–4]. During synthesis, intermediates are covalently bound to integrated carrier domains and transported to neighboring catalytic domains in an assembly line fashion[5]. Understanding the structural basis for catalysis with NRPSs will facilitate bioengineering to create novel products. Here we describe the structures of two different *holo*-NRPSs 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 single-particle electron microscopy analysis demonstrate a highly dynamic domain architecture and provide the foundation for understanding the structural mechanisms that could enable engineering novel NRPSs.

An 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

*Corresponding Author: gulick@hwi.buffalo.edu.
‡These authors contributed equally to this manuscript.

Author 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, solved and refined the models of AB3403 and EntF, respectively. C.S. and C.C.A. designed and synthesized the mechanism-based inhibitor. J.T.T. and G.S. performed and analyzed the single particle electron microscopy. A.M.G., E.J.D., B.R.M., G.S., J.T.T., C.C.A., and C.S. analyzed the results and wrote the manuscript. All authors have seen and approve of the manuscript.

PDB Deposition. The coordinates and structure factors have been deposited with the PDB: **4ZXH**, *holo*-AB3403; **4ZXI**, *holo*-AB3403 bound to AMP and glycine; **4ZXJ**, *holo*-EntF bound to Ser-AVS.

phosphopantetheine cofactor[6], an adenylation domain that loads the amino acid substrate onto the PCP cofactor, and a condensation domain that catalyzes peptide bond formation. NRPSs then use a C-terminal thioesterase or reductase domain to catalyze product release. Structures of individual domains[1] provide insight into the NRPS structural mechanism. Interestingly, the adenylation domains have been shown to adopt two catalytic conformations[7]. First the adenylate-forming conformation activates the amino acid substrate using ATP to form an aminoacyl adenylate and pyrophosphate. A C-terminal subdomain then rotates by ~140° to form the thioester-forming conformation that is used to install the amino acid onto the PCP[7]. These two functional states have been observed in structures of the phenylalanine activating adenylation domain of gramicidin synthetase[8] and the complexes between adenylation and carrier proteins obtained with mechanism-based inhibitors[9,10]. Once loaded, both the pantetheine and loaded substrate have been shown to interact transiently with the core of the carrier protein[11,12]. The structure of SrfA-C, the terminal module from surfactin biosynthesis, contains a condensation-adenylation-PCP-thioesterase architecture and is to date the only structure of an intact NRPS module[13]. The condensation and adenylation domains share an extensive interface and were proposed to form the core of the module[13]. Lacking the pantetheine modification, this *apo*-structure shows the PCP domain directed towards the condensation domain. The other active sites are 40–60 Å from the pantetheinylation site, indicating that extensive domain rearrangements are required to complete the NRPS catalytic cycle. Movement of the PCP domain, potentially coupled to the adenylation C-terminal subdomain rotation[7], is necessary for delivery of the peptide intermediates to the different catalytic domains.

We determined structures of two NRPSs with the same architecture as SrfA-C (Extended Data Fig. 1), but with *holo*-proteins that show functional interactions between the PCP and catalytic domains (Fig. 1). First we present two structures of AB3403 from the human pathogen *Acinetobacter baumannii* (protein annotation ABBFA_003403 in strain AB307-0294) that belongs to an uncharacterized biosynthetic pathway implicated in motility[14], and biofilm[15] and pellicle[16] formation. We describe the structures of *holo*-AB3403 obtained without ligands and also upon crystallization in the presence of Mg•ATP and glycine, which among the proteinogenic amino acids serves as the best substrate (Extended Data Fig. 2). Second, we present the structure of EntF from *Escherichia coli*, showing the PCP cofactor covalently trapped with a mechanism-based inhibitor to model thioester formation within 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 2.9Å 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 domain[17]. Contacts are made between the pantetheine and the helix running from Glu20-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 cysteamine 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 through Tyr1032 face the condensation domain. Leu1007 and Val1010 form a hydrophobic interaction with Leu22 and Ile80 of the condensation domain. The side chain of Lys1011 forms a hydrogen bond with the main chain carbonyl of Gln78. Finally, Val1026, Ala1027, and Ala1030 on the PCP helix α3 form a hydrophobic interaction with Tyr26 and Leu30. Arg344 of the condensation domain, which is positioned on an insertion compared to SrfA-C, interacts with the phosphate from the cofactor.

The AB3403 adenylation domain (Fig. 2c) is precisely positioned in the 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 ~12 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[19,20]. Despite the similarities in domain structure, the thioesterase domain of AB3403 is in a markedly different location compared to 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, harboring a vinylsulfonamide moiety that traps the thioester-forming reaction[21] to characterize functional adenylation-PCP didomain interactions[9,10]. 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 to promote crystallization. Crystals of the EntF protein diffract to 2.8Å (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[9,10], 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[7]. Arg863 interacts with the cofactor phosphate, while Gly864 and Gln865 form one wall of the pantetheine tunnel. Interactions with the nucleotide occur 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 primarily a tri-lobed density with two neighboring 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[13], 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 1023 $Å^2$, comparable in size to the 1097 $Å^2$ interface of SrfA-C. In contrast, the interface in EntF is only 780 $Å^2$, resulting from the rotation of the adenylation 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 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 alternate conformation in EntF may not be compatible with the adenylate-forming conformation. The three different condensation-adenylation domain conformations, the adenylate-forming incompatibility seen in EntF, and the multiple extended and compact conformations seen in the EM data, suggest that the condensation-adenylation domain platform may be more dynamic then previously proposed[13].

The new structures confirm the hypothesis[7] that the adenylation domain conformational change is a structural mechanism to guide the PCP between active sites in the context of complete NRPS modules. The rotation of the adenylation domain C-terminal subdomain from the adenylate-forming conformation in AB3403 to the thioester-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[11]. This interaction also alters the surface electrostatic potential of regions that interact with the neighboring catalytic domains, including α2 and α3, and may influence the PCP delivery to neighboring 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[22], as well as the open conformation of SrfA-C.

The basic NRPS catalytic cycle requires that the PCP visits three adjacent catalytic domains in a coordinated manner. The two catalytic conformations of the adenylation domain[7] require that the full cycle has four catalytic structural states (Fig. 4). Specifically, (I) the adenylation domain catalyzes amino acid adenylation, (II) the PCP is 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 catalyze 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 catalyze 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 catalyze unusual chemistry[23,24] offers 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 also 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 microscopy[25] to set the stage for investigations of the structural foundation of even larger, multi-modular biosynthetic proteins.

Methods

Expression, purification, and crystallization of AB3403

The human pathogen *Acinetobacter baumannii* contains an uncharacterized NRPS cluster that has been implicated in motility and biofilm formation; the product of this operon is unknown. This operon contains eight genes. In strain AB307-0294[26], from which the NRPS

gene was cloned, this operon consists of genes ABBFA_003399 through ABBFA_003406. In the more commonly used ATCC17978 strain, the same genes are encoded by A1S_0119 through A1S_0112. The ABBFA_003403 (designated AB3403 herein) protein sequence is available at Genbank accession # ACJ56070.1.

The gene encoding AB3403 was PCR-amplified from AB307-0294 genomic DNA[26] (courtesy of Dr. Thomas A. Russo, University at Buffalo). The amplified fragment was cloned into the pET15b-TEV expression vector[27] and confirmed by DNA sequencing. The vector provides a His_5-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 OD_{600} of 0.6 at 37 °C. Protein expression was induced by addition of 0.5 mM 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 labeled protein was generated in M9 minimal media utilizing a metabolic inhibition protocol[28]. 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 (Branson Sonifier) and the resulting lysate was clarified by centrifugation at 235,000 g for 45 min. The cell lysate was passed over a His-trap (GE-Healthcare) immobilized metal ion affinity (IMAC) 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 dialyzed against a TEV cleavage buffer (50 mM HEPES (pH 8.0), 250 mM NaCl, 0.2 mM TCEP, and 0.5 mM EDTA) for 16 h at 4 °C. This partially purified protein was then phosphopantetheinylated by incubation with His_6-tagged nonspecific phosphopantetheinyl transferase Sfp (10 nM), 12.5 mM $MgCl_2$, and 1 mM CoA for 60 minutes at 20 °C. The clarified protein was then passed over the Histrap 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, dialyzed 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 initially identified from a sparse matrix screen at 20 °C. Final crystals for native and SeMet labeled *holo*-AB3403 were grown at 14 °C by hanging-drop vapor diffusion against 0.75–0.95 M potassium citrate, 0.01–0.025 M glycine, and 0.05 M BisTrisPropane (BTP) (pH 8.0). Highest quality native crystals were obtained using a protein concentration of 5.5 mg/mL 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 and 1:1 protein to cocktail ratio. To obtain crystals in the presence of ligands, the protein was preincubated for 45 min at 4 °C with 2 mM $MgCl_2$, and 1.5-fold molar excess of ATP and glycine.

Structure determination of AB3403

Crystals of *holo*-AB3403 were cryoprotected by stages utilizing either ethylene glycol or potassium citrate for native and SeMet protein, respectively. The native protein crystals were cryo-protected with cocktails containing 1.0 M potassium citrate, 0.3 M glycine, 0.05 M BTP (pH 8.0), and increasing [8, 16, and 24 %] v/v ethylene glycol. The SeMet-labeled protein was cryo-protected with cocktails containing 0.3 M glycine, 0.05M BTP (pH 8.0) and increasing [1.0, 1.2, 1.4, and 1.6 M] potassium citrate. Crystals derived from protein co-crystallized with ligands included the same concentration of $MgCl_2$, ATP, and glycine in the cryo-protectant cocktails.

Diffraction data were collected on APS beamline 23-IDB. The native data (2.7 Å) was collected using a multi-crystal, multi-dataset strategy using two crystals. A complete low resolution scan was taken for one crystal followed by a higher resolution scan of the best diffracting crystal. A high resolution region of the second crystal was combined with the two scans from the first crystal. The optimal regions were identified with the JBLU-ICE software at the GM/CA beamline. A single peak wavelength dataset (3.35 Å) was collected for SeMet labeled protein. The liganded protein crystal was collected with a single crystal.

Diffraction data were indexed, merged, and scaled using iMOSFLM[29] in space group $P4_x2_12$. Structure determination was carried out with PHENIX[30] using a combination of experimental 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 **1AMU**)[8] and CytC1 (PDB **3VNR**). Using this partial molecular replacement model, the selenium sites were identified with the SAD data from SeMet labeled crystals. An initial model was produced 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-WARP[31], COOT[32] and PHENIX refine.

The final refinements were performed with TLS parameterization[33] with groups consisting of residues 1:191, 191:445, 446:480, 481:862, 863:959, 960:973, 974:1044, and 1054:1318, roughly defining the NRPS domain (or subdomain) boundaries. The protein is complete from residue Asn2 through 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 nucleotide[7].

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 density 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 then recognizes the external carrier protein EntB 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 previously[22,34]. The *entf* gene was PCR amplified from *E. coli* JM109 and cloned into a pET15-TEV vector with a N-terminal 5x His-tag and a TEV protease cleavage site[22]. The *entf* vector was transformed into *E. coli* (BL21-DE3) cells for protein expression. Cells were grown in LB media to an OD_{600} of 0.6 at 37 °C prior to protein induction with 1mM IPTG. Cells were grown overnight at 16 °C, and collected by centrifugation. The cell pellets were flash frozen in liquid nitrogen. Selenomethionine labeled EntF was expressed in M9 minimal media as described[28].

For purification of both native and SeMet labeled protein, cells were resuspended in lysis buffer containing 50 mM Tris-HCl pH 7.5, 400 mM NaCl, 0.2 mM TCEP, 10% glycerol, and 10 mM imidazole. Cells were lysed via sonication and centrifuged at 235,000 g for 45 minutes. Initial purification was achieved with a His-trap IMAC column. Protein was eluted using lysis buffer with 300 mM imidazole. EntF was incubated with TEV protease overnight at 4°C in a cleavage buffer containing 50 mM Tris pH 7.5, 400 mM NaCl, 0.2 mM TCEP, 10% glycerol, and 0.5 mM EDTA. Although expressed in *E. coli*, phosphopantetheinylation was assured by the addition of 10 nM Sfp, 1 mM CoA, and 12.5 mM $MgCl_2$. The reaction was incubated at room temperature for 1-2 hours. The *holo*-EntF was run over an IMAC column once more to remove uncleaved protein along with Sfp. A final polishing step was performed with a Superdex 200 16/600 in a final dialysis buffer containing 50 mM EPPS pH 8.0, 150 mM NaCl, 0.2 mM TCEP, 1 mM $MgCl_2$, and 10% glycerol. Prior to crystallization, the Ser-AVS inhibitor was added at a concentration 4x that of EntF and allowed to incubate for 2–4 hours at room temperature.

For electron microscopy, native EntF was purified the same as above with the exception that a minimal dialysis buffer was used, which contained 50 mM EPPS pH 8.0, 100 mM NaCl, and 0.2 mM TCEP. No inhibitor was added.

Crystal conditions for the Ser-AVS inhibited EntF were first identified using the Hauptman-Woodward high throughput screen[35]. Large diffraction quality native and SeMet crystals were grown using hanging drop vapor diffusion at 20 °C. A crystallization cocktail, consisting of 100 mM BTP pH 7.5, 125–150 mM $MgCl_2$, and 22–28% PEG 4000, was diluted 1:1 with the final dialysis buffer. The hanging drops then combined protein at 30 mg/mL and the undiluted crystallization cocktail at a ratio of 1:2. This “batch mimic” limited the differences between the drop and reservoir and has been successful with other protein samples in our lab[36].

Structure determination of EntF

Native EntF crystals were cryoprotected by that addition of 2,3-butanediol directly to the crystallization drop to a final concentration of ~10%. SeMet crystals were cryoprotected

similarly except with glycerol to a final concentration of ~20%. Diffraction data was collected on APS beamline 23-IDB using the rastering option to find the optimal spots on both the native the SeMet crystals. Diffraction data were indexed, merged, and scaled using iMOSFLM[29] in space group $P4_x2_12$. Structure determination for the SeMet inflection data was carried out in PHENIX[30] using a PhaserEP MR-SAD with a partial molecular replacement solution that was obtained using a sculpted model (generated with PHENIX sculptor) derived from the *P. aeruginosa* bidomain Adenylation-PCP protein PA1221 (PDB: **4DG9**)[9]. Automated model building with BUCCANEER was used to build ~65% of the structure[37]. 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 thioesterase domain, which was unresolved and comprises about 19%). This model was built and refined iteratively using COOT[32] and PHENIX refine. TLS refinement[33] was utilized in final stages with groups consisting of residues 5:186, 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 thioesterase 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, likely 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 stain EM analysis of EntF

EntF, purified as described above, was prepared for electron microscopy using the conventional negative staining protocol[38], 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 71,138x and a defocus value of ~1.5μm on a Gatan US4000 CCD camera. All images were binned (2 × 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][39]. 2D reference-free alignment and classification of particle projections was performed using ISAC[40]. 17,431 projections of EntF were subjected to ISAC producing 133 classes consistent over two-way matching and accounting for 5,344 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. ^{1}H and ^{13}C NMR spectra were recorded on a Varian 600 MHz spectrometer. Proton chemical shifts are reported in ppm from an internal standard of residual chloroform (7.26 ppm) or methanol (3.31 ppm), and carbon chemical shifts are reported using an internal standard of residual chloroform (77.3 ppm) or methanol (49.1 ppm). Proton chemical data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, p = pentet, m = multiplet, br = broad), integration, coupling constant.

High resolution mass spectra were obtained on an Agilent TOF II TOF/MS instrument equipped with either an ESI or APCI interface. TLC analyses were performed on TLC silica gel 60F254 from EMD Chemical Inc., and were visualized with UV light or 10% PMA solution. Purifications were performed by flash chromatography on silica gel (Dynamic Adsorbents, 60A).

Materials—Chemicals, reagents and solvents were purchased from Sigma Aldrich Company, Chem-Impex or Acros Organic Fischer Company, and were used as received. An anhydrous solvent dispensing system (J. C. Meyer) using 2 packed columns of neutral alumina was used for drying THF, Et_2O, while 2 packed columns of molecular sieves were used to dry DMF and the solvents were dispensed under argon. Compound **1** was purchased from Chem-Impex and used as received. Compound **2**[41] and **4**[10] were synthesized according to the reported procedures.

***tert*-Butyl(*R*,*E*)-4-(2-(*N*-(*tert*-butoxycarbonyl)sulfamoyl)vinyl)-2,2-dimethyloxazolidine-3-carboxylate (3)—**To a solution of *tert*-butyl [(diphenylphosphoryl)methyl]sulfonylcarbamate **2** (395 mg, 1.0 mmol, 2.0 equiv) in 1:3 DMF–THF (4 mL) at −78 °C, was added a 1 M solution of LiHMDS in THF (2.0 mL, 4.0 equiv) dropwise over 15 min and the solution was stirred at −78 °C for an additional 15 min. Next, Garner's aldehyde **1** (115 mg, 0.5 mmol, 1.0 equiv) in THF (1 mL) was added to the reaction over 15 min. The solution was gradually warmed to 25 °C and stirred for 15 h. The solvent was removed *in vacuo* and the mixture was taken up in H_2O (30 mL). The pH was adjusted to 3–4 with 1 N aqueous HCl, then was extracted with EtOAc (3 × 20 mL). The combined organic layers were washed with H_2O (30 mL), saturated aqueous NaCl (30 mL), dried ($MgSO_4$) and concentrated. Purification by flash chromatography (10% EtOAc–hexane to 50% EtOAc–hexanes) afforded the title compound **3** as colorless oil (150 mg, 74%): R_f = 0.50 (50:50 EtOAc–hexanes); $[\alpha]^{23}_D$ +0.9 (*c* 0.02, CH_2Cl_2); ^{1}H NMR (600 MHz, CD_3OD) ^{1}H NMR (600 MHz, CD_3OD) δ 1.45 (s, 3H), 1.48 (m, 9H), 1.51 (s, 9H), 1.60 (s, 3H), 3.83–3.85 (m, 1H), 4.15 (dd, *J* = 12.0, 6.0 Hz, 1H), 4.56–4.58 (m, 1H), 6.64 (d, *J* = 18 Hz, 1H), 6.77–6.81 (m, 1H); ^{13}C NMR (150 MHz, CD_3OD) δ 28.41, 28.47, 28.80, 28.81, 58.7, 68.3, 84.19, 84.22, 95.8, 130.6, 145.7, 152.2, 152.7; HRMS (ESI−) calcd for $C_{17}H_{29}N_2O_7S$ $[M - H]^-$ 405.1701, found 405.1721 (error 4.9 ppm).

Ser-AVS—To a solution of N^6, N^6-bis(*tert*-butoxycarbonyl)-2′,3′-*O*-isopropylideneadenosine **4** (73 mg, 0.14 mmol, 1.1 equiv), vinylsulfonamide **3** (52 mg, 0.13 mmol, 1.0 equiv) and PPh_3 (56 mg, 0.21 mmol, 1.7 equiv) in THF (1 mL) at 0 °C, was added a solution of DIAD (42 μL, 0.21 mmol, 1.7 equiv) in THF (1 mL) over 1 h using a syringe pump. The solution was gradually warmed up to 23 °C and stirred overnight. The mixture was filtered over a short pad of silica gel, which was washed with 20% EtOAc–hexanes (100 mL). The filtrate was concentrated to afford crude **5** (R_f = 0.45, 50:50 EtOAc–hexanes), which was used in the next step without further purification. To a solution of crude **5** from the previous step was added 80% aqueous TFA (1 mL) at 0 °C. The solution was stirred for 6 h at 0 °C then concentrated. Recrystallization from 1:20 MeOH–Et_2O (5 mL) afforded the title compound (32 mg, 47%) as colorless film: $[\alpha]^{23}_D$ −10.3 (*c* 0.600, MeOH); ^{1}H NMR (600 MHz, CD_3OD) δ 3.30–3.39 (m, 2H), 3.67–3.70 (m, 1H), 3.83 (dd, *J*

= 11.6, 4.1 Hz, 1H), 4.05–4.08 (m, 1H), 4.22–4.25 (m, 1H), 4.34–4.35 (m, 1H), 4.77–4.81 (m, 1H), 5.94 (d, J = 6.2 Hz, 1H), 6.70 (dd, J = 15.4, 6.5 Hz, 1H), 6.77 (d, J = 15.4 Hz, 1H), 8.27 (s, 1H), 8.29 (s, 1H); ^{13}C NMR (150 MHz, CD_3OD) δ 45.8, 54.1, 62.3, 72.9, 74.8, 85.8, 91.7, 121.3, 134.8, 137.0, 143.2, 149.9, 151.3, 156.1; HRMS (ESI+) calcd for $C_{14}H_{22}N_7O_6S$ $[M + H]^+$ 416.1347, found 416.1339 (error 1.9 ppm).

Kinetic analysis of AB3403

Substrate preference for the adenylation domain of *holo*-AB3403 was established by the pyrophosphate exchange assay[42] allowing radiolabeled PP_i to be incorporated into ATP in the reverse reaction. 1.0 μM holo-AB3403 was added to 2 mM ATP, 0.2 mM NaPPi, 50 mM HEPES (pH 7.5), 100 mM NaCl, 10 mM $MgCl_2$, 0.15 μCi ^{32}P-PPi, and 5 mM substrate. Reactions (100 μl) were incubated for 10 min at 37°C, then quenched with 0.5 ml 1.2% charcoal, 0.1 M unlabeled PPi, and 0.35 M perchloric acid. The charcoal was pelleted by centrifugation, washed twice with 1 ml H_2O, and resuspended in 0.5 ml H_2O for scintillation counting.

To determine the apparent kinetic constants for ATP and glycine for the *holo*-AB3403 adenylation domain, the $NADH^+$ consumption assay monitored at OD_{340}[43,44] with full length AB3403. Hydroxylamine was used as a surrogate for the pantetheine in the second partial reaction to displace AMP for use in the coupled consumption assay[45]. Standard reactions contained 50 mM HEPES (pH 7.5), 15 mM MgCl2, 2 mM ATP, 3 mM phosphophenyl pyruvate, 0.2 mM $NADH^+$, 5 U myokinase, 5 U pyruvate Kinase, 6.5 U lactate dehydrogenase and 150 mM 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 0.0025 mM enzyme. Calculations were done using PRISM software.

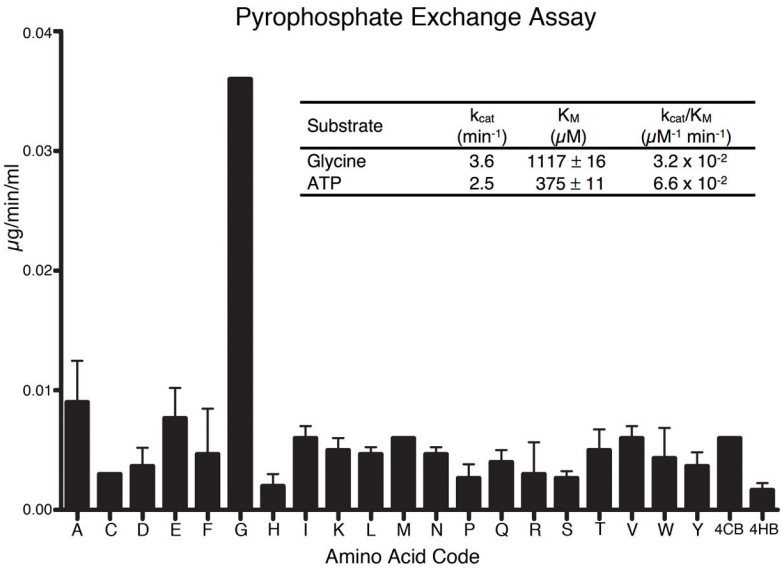
Extended Data

	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 1. Structure-based alignment of EntF, AB3403, and SrfA-C

Condensation, Adenylation, PCP, and Thioesterase domains are represented with bars in grey, pink, greencyan, and blue. Conserved motifs and catalytically important residues are highlighted with the same colors, 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

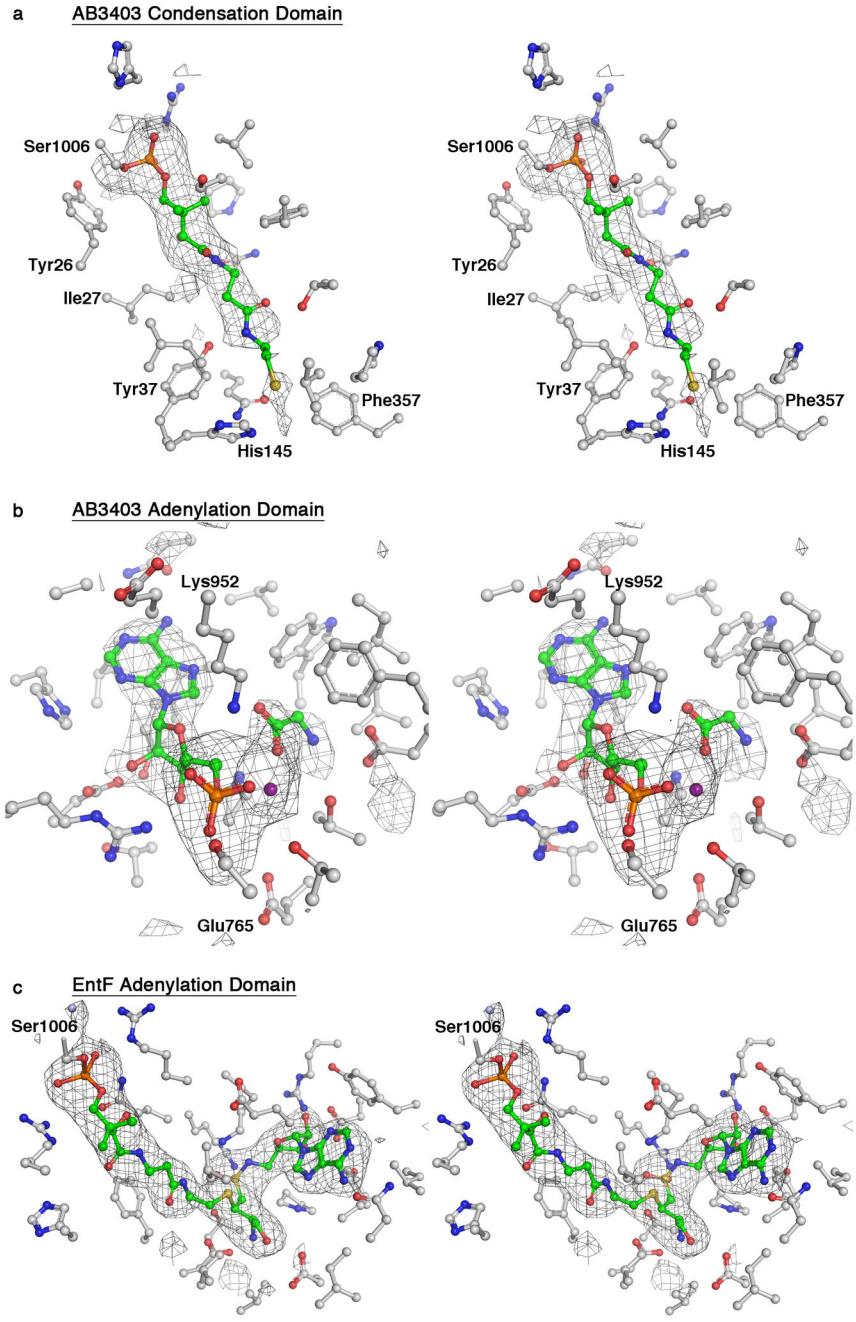
approximately 26% sequence identity. The adenylation and PCP domains are more well-conserved, sharing ~35% identity, while the condensation (21%) and thioesterase (25%) domains are less well conserved. Domain boundaries are described in the table below.



Substrate	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M (μM^{-1} min^{-1})
Glycine	3.6	1117 ± 16	3.2×10^{-2}
ATP	2.5	375 ± 11	6.6×10^{-2}

Extended Data Figure 2. Substrate Specificity of full-length AB3403

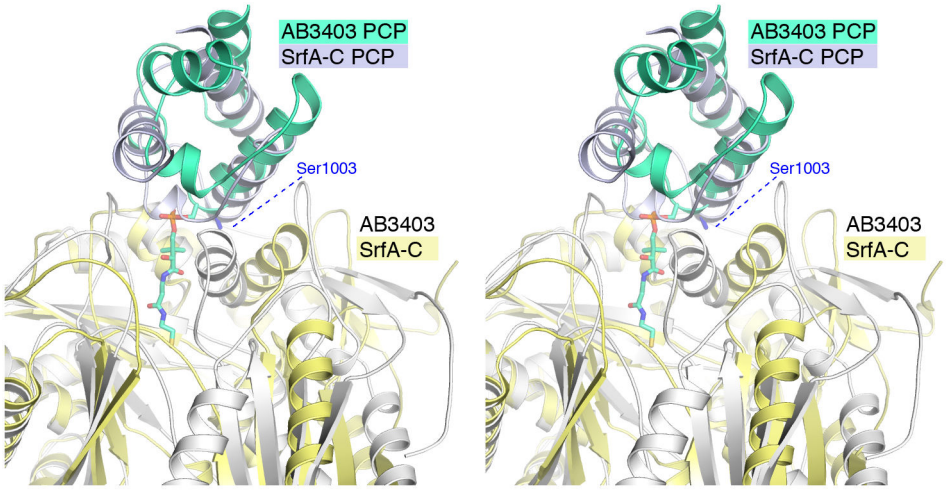
Amino acid specificity of AB3403 was recorded for all 20 proteinogenic amino acids, as well as 4-chlorobenzoate (4CB) and 4-hydroxybenzoate (4HB). Average values and standard deviations are shown for three replicates with each substrate; results were recorded as μmol radiolabeled ATP incorporated per min per mg of enzyme. Apparent kinetic constants are also shown for ATP and glycine.



Extended Data Figure 3. Stereorepresentations of Electron density figures shown in Fig. 2

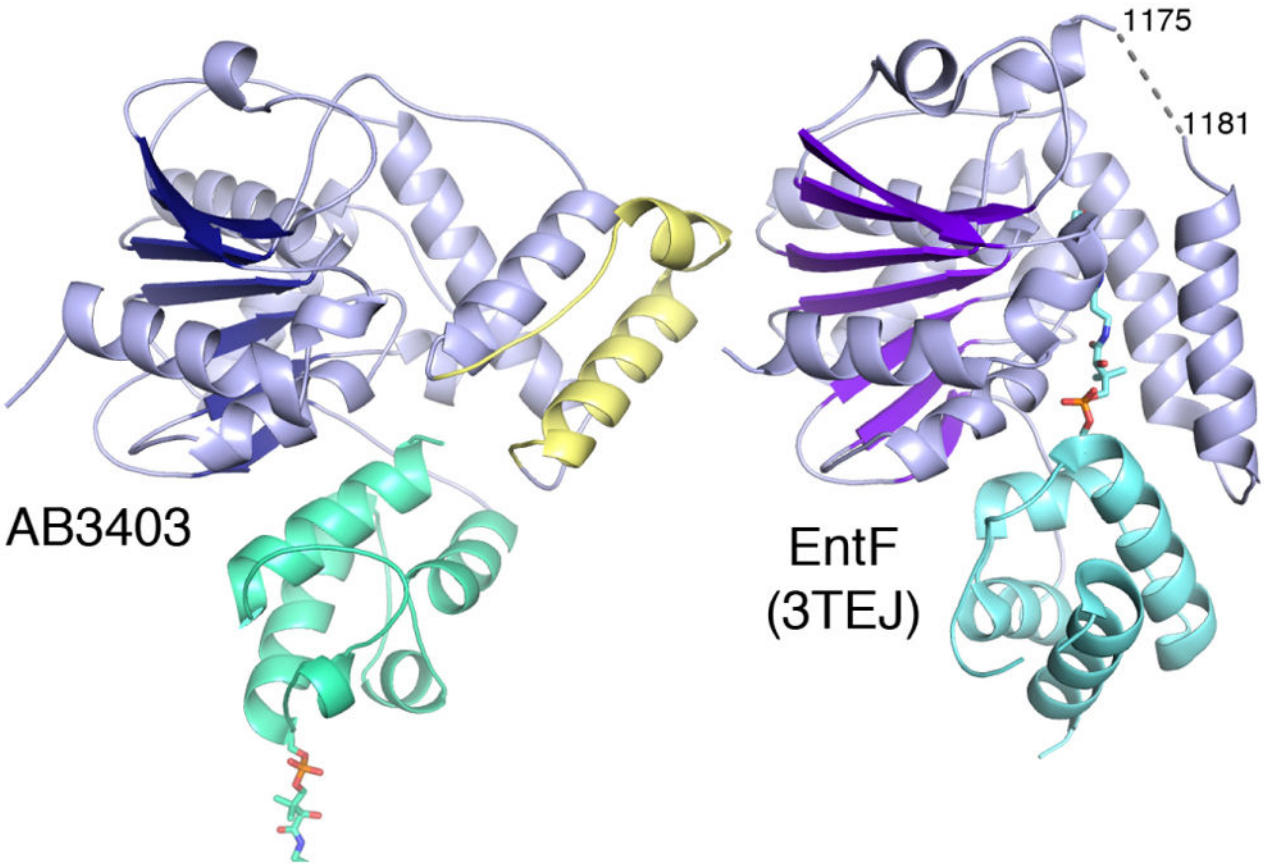
To better visualize the active sites and electron density quality, stereofigures are included in the extended data. In all panels, density is shown with coefficients of the form (F_O-F_C) calculated prior to inclusion of ligands and contoured at 3σ. (a) Stereorepresentation of electron density of AB3403 condensation domain shows the phosphopantethine on Ser1006 approaching His145 within the Condensation Domain pocket. Inhibitor carbon atoms in green, carbons of residues within 5Å of inhibitor in gray, nitrogen in blue, oxygen in red, sulfur in yellow, and water in light blue. (b) Electron density of the nucleotide binding

pocket of AB3403 bound to glycine and AMP. Stereorepresentation of electron density shows the AMP, glycine, and Mg^{+} present in the active site of the adenylation domain. Ligand carbon atoms are in green, carbons of residues within 5Å of inhibitor in gray, nitrogen in blue, oxygen in red, phosphorous in orange, and the Mg^{+} cofactor in purple. (c) Stereorepresentation of the electron density shows the phosphopantethine on Ser1006 covalently attached to the Ser-vinylsulfonamide inhibitor in the active site of the adenylation domain. Inhibitor carbon atoms in green, carbons of residues within 4Å of inhibitor in gray, nitrogen in blue, oxygen in red, phosphorous in orange, sulfur in yellow, and water in light blue.



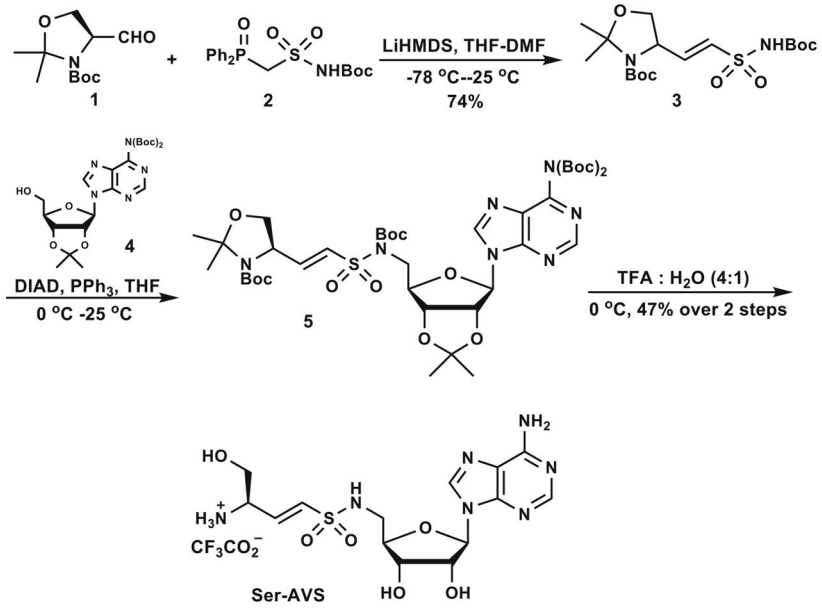
Extended Data Figure 4. Comparison of AB3403 and SrfA-C PCP-Condensation domain interaction

Stereorepresentation 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.



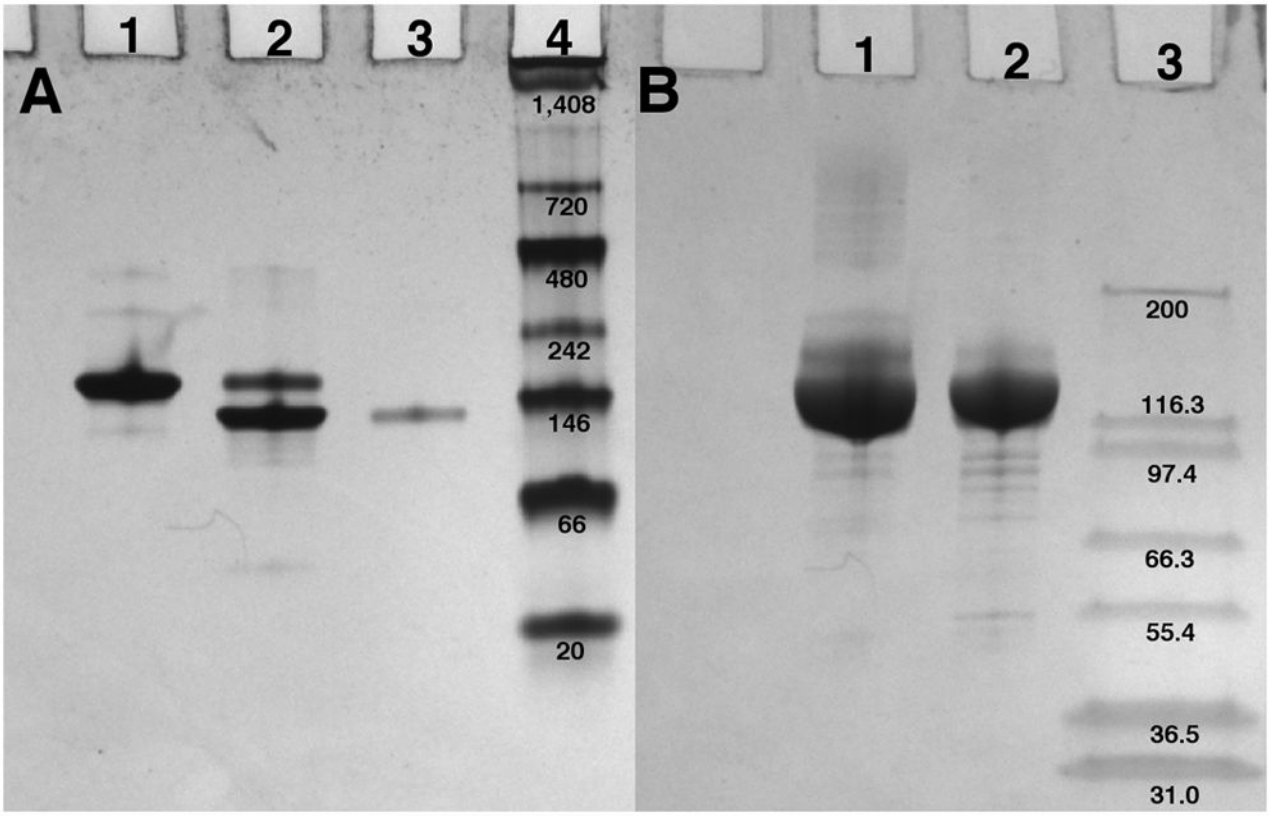
Extended Data Figure 5. Comparison of AB3403 thioesterase domain to the functional PCP-TE interaction

Stereorepresentation 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 colored yellow.



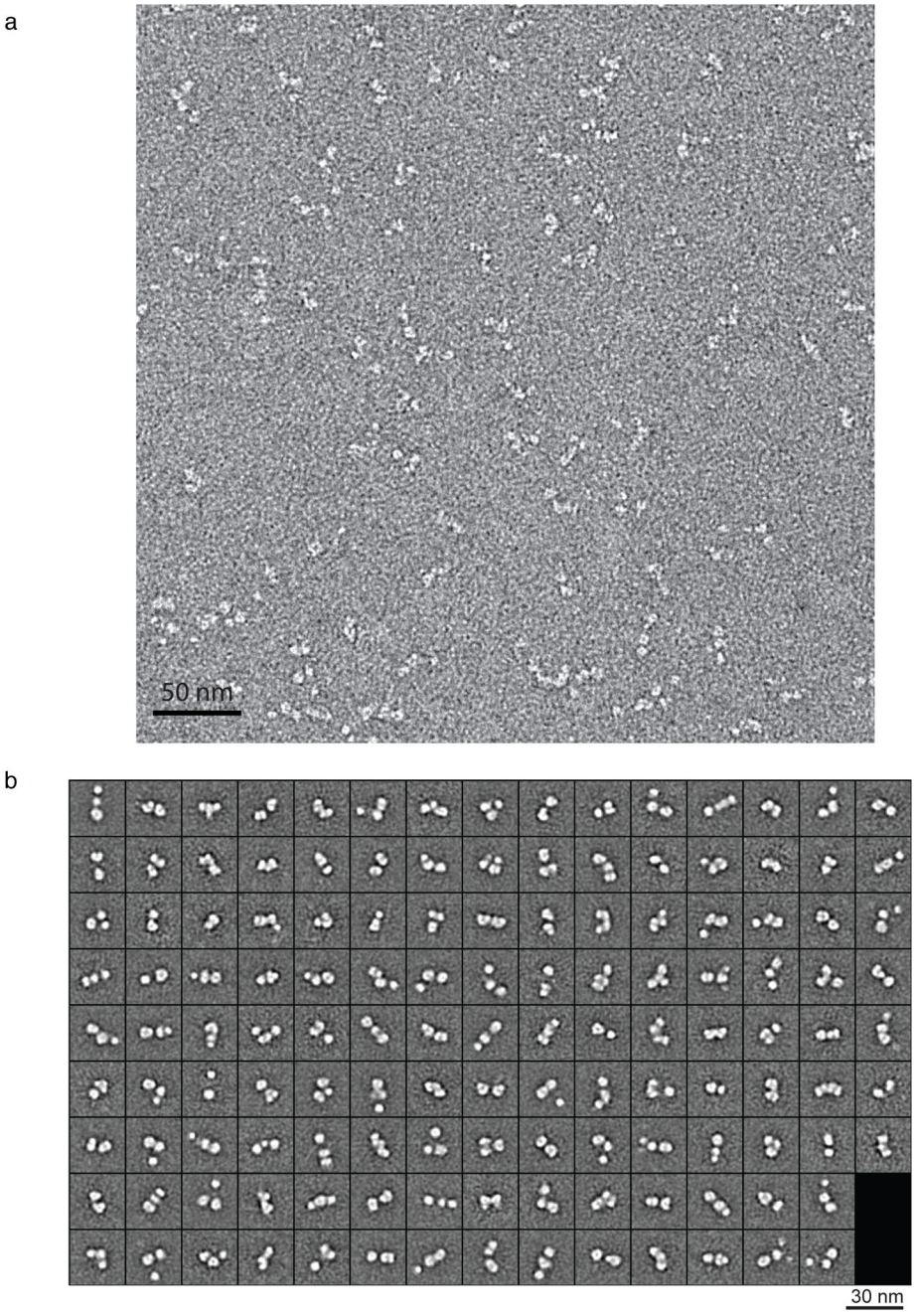
Extended Data Figure 6. Synthesis of Ser-AVS

The Ser-AVS probe was synthesized following similar protocols described elsewhere[41,46]. Garner's aldehyde **1** was coupled with **2** employing 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 TFA to yield **Ser-AVS.**



Extended Data Figure 7. Electrophoretic mobility of EntF

(A) Native gel electrophoresis. Lane 1: EntF. Lane2: EntF incubated with 4-fold molar excess of Ser-AVS inhibitor. Lane 3: EntF Crystals. Lane 4: novex NativeMark labeled in kDa. (B) Denaturing gel electrophoresis using loading buffer with SDS and β-mercaptoethanol. Gel Lane 1: EntF, Lane 2: EntF incubated 4x Ser-AVS inhibitor, and Lane 3: Life Technologies Mark12 labeled in kDa. The native gel shows the inhibited EntF in a more compact conformation compared to EntF without the inhibitor.



Extended Data Figure 8. Negative stain EM analysis of EntF

(a) Raw EM 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**

	SeMet_AB3403	AB3403	Liganded AB3403
Beamline	APS 23-ID-B	APS 23-ID-B	APS 23-ID-B
Wavelength (Å)	0.9793	0.9796	1.0332
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$
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 (3.52-3.35)	49.80-2.70 (2.79 - 2.70)	45.03-2.90 (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/sigma(I)	11.9 (3.5)	9.91 (2.49)	8.47 (2.19)
R_{merge}	0.090 (0.359)	0.125 (0.511)	0.130 (0.641)
R_{meas}	0.116	0.143	0.162
CC1/2	0.993 (0.798)	0.991 (0.685)	0.991 (0.635)
CC*	0.998 (0.937)	0.998 (0.902)	0.998 (0.881)
Structure Refinement			
R_{factor}		0.179 (0.248)	0.174 (0.307)
R_{free}		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^2)		41.15	50.83
Average B-Factor (Å^2)			
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	$P4_12_12$	$P4_12_12$
Unit cell a, b, c (Å)	127.55 127.55 186.72	127.71 127.71 186.94
Molecules per ASU	1	1

	SeMet EntF	EntF
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/sigma(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*	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 (Å^2)		62.88
Average B-Factor (Å^2)		
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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Ruslan Sanishvili for assistance with data collection. This work was funded in part by National Institutes of Health GM-068440 (AMG) and GM-115601 (GS), and Award #W81XWH-11-2-0218 from the Telemedicine and Advanced Technology Research Center of the US Army Medical Research and Materiel Command (AMG). Data were collected at the GM/CA beamline of the Advanced Photon Source, which is funded by the NCI (ACB-12002) and the NIGMS (AGM-12006) under DOE Contract No DE-AC02-06CH11357 to APS. A Stafford Fellowship (BRM) and support from the Hauptman-Woodward Institute is gratefully acknowledged.

References

1. Koglin A, Walsh CT. Structural insights into nonribosomal peptide enzymatic assembly lines. Nat Prod Rep. 2009; 26:987–1000. [PubMed: 19636447]
2. Fischbach MA, Walsh CT. Antibiotics for emerging pathogens. Science. 2009; 325:1089–93. [PubMed: 19713519]
3. Walsh CT. The chemical versatility of natural-product assembly lines. Acc Chem Res. 2008; 41:4–10. [PubMed: 17506516]
4. Walsh CT, Fischbach MA. Natural products version 2.0: connecting genes to molecules. J Am Chem Soc. 2010; 132:2469–93. [PubMed: 20121095]
5. Strieker M, Tanovic A, Marahiel MA. Nonribosomal peptide synthetases: structures and dynamics. Curr Opin Struct Biol. 2010; 20:234–40. [PubMed: 20153164]
6. Mercer AC, Burkart MD. The ubiquitous carrier protein--a window to metabolite biosynthesis. Nat Prod Rep. 2007; 24:750–73. [PubMed: 17653358]
7. Gulick AM. Conformational dynamics in the acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. ACS Chem Biol. 2009; 4:811–827. [PubMed: 19610673]
8. Conti E, Stachelhaus T, Marahiel MA, Brick P. Structural basis for the activation of phenylalanine in the non- ribosomal biosynthesis of gramicidin S. EMBO J. 1997; 16:4174–83. [PubMed: 9250661]
9. Mitchell CA, Shi C, Aldrich CC, Gulick AM. Structure of PA1221, a nonribosomal peptide synthetase containing adenylation and peptidyl carrier protein domains. Biochemistry. 2012; 51:3252–63. [PubMed: 22452656]
10. Sundlov JA, Shi C, Wilson DJ, Aldrich CC, Gulick AM. Structural and functional investigation of the intermolecular interaction between NRPS adenylation and carrier protein domains. Chem Biol. 2012; 19:188–98. [PubMed: 22365602]
11. Goodrich AC, Harden BJ, Frueh DP. Solution Structure of a Nonribosomal Peptide Synthetase Carrier Protein Loaded with Its Substrate Reveals Transient, Well-Defined Contacts. J Am Chem Soc. 2015; 137:12100–9. [PubMed: 26334259]
12. Jaremko MJ, Lee DJ, Opella SJ, Burkart MD. Structure and Substrate Sequestration in the Pyoluteorin Type II Peptidyl Carrier Protein PltL. J Am Chem Soc. 2015; 137:11546–9. [PubMed: 26340431]
13. Tanovic A, Samel SA, Essen LO, Marahiel MA. Crystal Structure of the Termination Module of a Nonribosomal Peptide Synthetase. Science. 2008; 321:659–663. [PubMed: 18583577]
14. Clemmer KM, Bonomo RA, Rather PN. Genetic analysis of surface motility in Acinetobacter baumannii. Microbiology. 2011; 157:2534–44. [PubMed: 21700662]
15. Rumbo-Feal S, et al. Whole Transcriptome Analysis of Acinetobacter baumannii Assessed by RNA-Sequencing Reveals Different mRNA Expression Profiles in Biofilm Compared to Planktonic Cells. PLoS One. 2013; 8:e72968. [PubMed: 24023660]
16. Giles SK, Stroeher UH, Eijkelkamp BA, Brown MH. Identification of genes essential for pellicle formation in Acinetobacter baumannii. BMC Microbiol. 2015; 15:116. [PubMed: 26047954]
17. Bloudoff K, Rodionov D, Schmeing TM. Crystal Structures of the First Condensation Domain of CDA Synthetase Suggest Conformational Changes during the Synthetic Cycle of Nonribosomal Peptide Synthetases. J Mol Biol. 2013; 425:3137–50. [PubMed: 23756159]
18. Marahiel MA, Stachelhaus T, Mootz HD. Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis. Chem Rev. 1997; 97:2651–2674. [PubMed: 11851476]
19. Frueh DP, et al. Dynamic thiolation-thioesterase structure of a non-ribosomal peptide synthetase. Nature. 2008; 454:903–6. [PubMed: 18704088]
20. Liu Y, Zheng T, Bruner SD. Structural basis for phosphopantetheinyl carrier domain interactions in the terminal module of nonribosomal peptide synthetases. Chem Biol. 2011; 18:1482–8. [PubMed: 22118682]
21. Qiao C, Wilson DJ, Bennett EM, Aldrich CC. A mechanism-based aryl carrier protein/thiolation domain affinity probe. J Am Chem Soc. 2007; 129:6350–1. [PubMed: 17469819]

22. Miller BR, Sundlov JA, Drake EJ, Makin TA, Gulick AM. Analysis of the linker region joining the adenylation and carrier protein domains of the modular nonribosomal peptide synthetases. Proteins. 2014; 82:2691–702. [PubMed: 24975514]
23. Gaudelli NM, Long DH, Townsend CA. beta-Lactam formation by a non-ribosomal peptide synthetase during antibiotic biosynthesis. Nature. 2015; 520:383–387. [PubMed: 25624104]
24. Maruyama C, et al. A stand-alone adenylation domain forms amide bonds in streptothricin biosynthesis. Nat Chem Biol. 2012; 8:791–7. [PubMed: 22820420]
25. Dutta S, et al. Structure of a modular polyketide synthase. Nature. 2014; 510:512–7. [PubMed: 24965652]
26. Adams MD, et al. Comparative genome sequence analysis of multidrug-resistant Acinetobacter baumannii. J Bacteriol. 2008; 190:8053–64. [PubMed: 18931120]
27. Kapust RB, et al. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. Protein Eng. 2001; 14:993–1000. [PubMed: 11809930]
28. Doubliè S. Preparation of selenomethionyl proteins for phase determination. Methods Enzymol. 1997; 276:523–530. [PubMed: 9048379]
29. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr D Biol Crystallogr. 2011; 67:271–81. [PubMed: 21460445]
30. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 2010; 66:213–21. [PubMed: 20124702]
31. Langer G, Cohen SX, Lamzin VS, Perrakis A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat Protoc. 2008; 3:1171–9. [PubMed: 18600222]
32. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–32. [PubMed: 15572765]
33. Urzhumtsev A, Afonine PV, Adams PD. TLS from fundamentals to practice. Crystallogr Rev. 2013; 19:230–270. [PubMed: 25249713]
34. Roche ED, Walsh CT. Dissection of the EntF condensation domain boundary and active site residues in nonribosomal peptide synthesis. Biochemistry. 2003; 42:1334–44. [PubMed: 12564937]
35. Luft JR, et al. A deliberate approach to screening for initial crystallization conditions of biological macromolecules. J Struct Biol. 2003; 142:170–9. [PubMed: 12718929]
36. Sundlov JA, Gulick AM. Structure determination of the functional domain interaction of a chimeric nonribosomal peptide synthetase from a challenging crystal with noncrystallographic translational symmetry. Acta Crystallogr D Biol Crystallogr. 2013; 69:1482–92. [PubMed: 23897471]
37. Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr D Biol Crystallogr. 2006; 62:1002–11. [PubMed: 16929101]
38. Ohi M, Li Y, Cheng Y, Walz T. Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy. Biol Proced Online. 2004; 6:23–34. [PubMed: 15103397]
39. Tang G, et al. EMAN2: an extensible image processing suite for electron microscopy. J Struct Biol. 2007; 157:38–46. [PubMed: 16859925]
40. Yang Z, Fang J, Chittuluru J, Asturias FJ, Penczek PA. Iterative stable alignment and clustering of 2D transmission electron microscope images. Structure. 2012; 20:237–47. [PubMed: 22325773]
41. Ikeuchi H, Meyer ME, Ding Y, Hiratake J, Richards NG. A critical electrostatic interaction mediates inhibitor recognition by human asparagine synthetase. Bioorg Med Chem. 2009; 17:6641–50. [PubMed: 19683931]
42. Rusnak F, Faraci WS, Walsh CT. Subcloning, expression, and purification of the enterobactin biosynthetic enzyme 2,3-dihydroxybenzoate-AMP ligase: demonstration of enzyme-bound (2,3-dihydroxybenzoyl)adenylate product. Biochemistry. 1989; 28:6827–35. [PubMed: 2531000]
43. Horswill AR, Escalante-Semerena JC. Characterization of the propionyl-CoA synthetase (PrpE) enzyme of Salmonella enterica: residue Lys592 is required for propionyl-AMP synthesis. Biochemistry. 2002; 41:2379–87. [PubMed: 11841231]

44. Reger AS, Carney JM, Gulick AM. Biochemical and Crystallographic Analysis of Substrate Binding and Conformational Changes in Acetyl-CoA Synthetase. Biochemistry. 2007; 46:6536–46. [PubMed: 17497934]

45. Wilson DJ, Aldrich CC. A continuous kinetic assay for adenylation enzyme activity and inhibition. Anal Biochem. 2010; 404:56–63. [PubMed: 20450872]

46. Reuter DC, McIntosh JE, Guinn AC, Madera AM. Synthesis of vinyl sulfonamides using the Horner reaction. Synthesis. 2003; 2003:2321–2324.

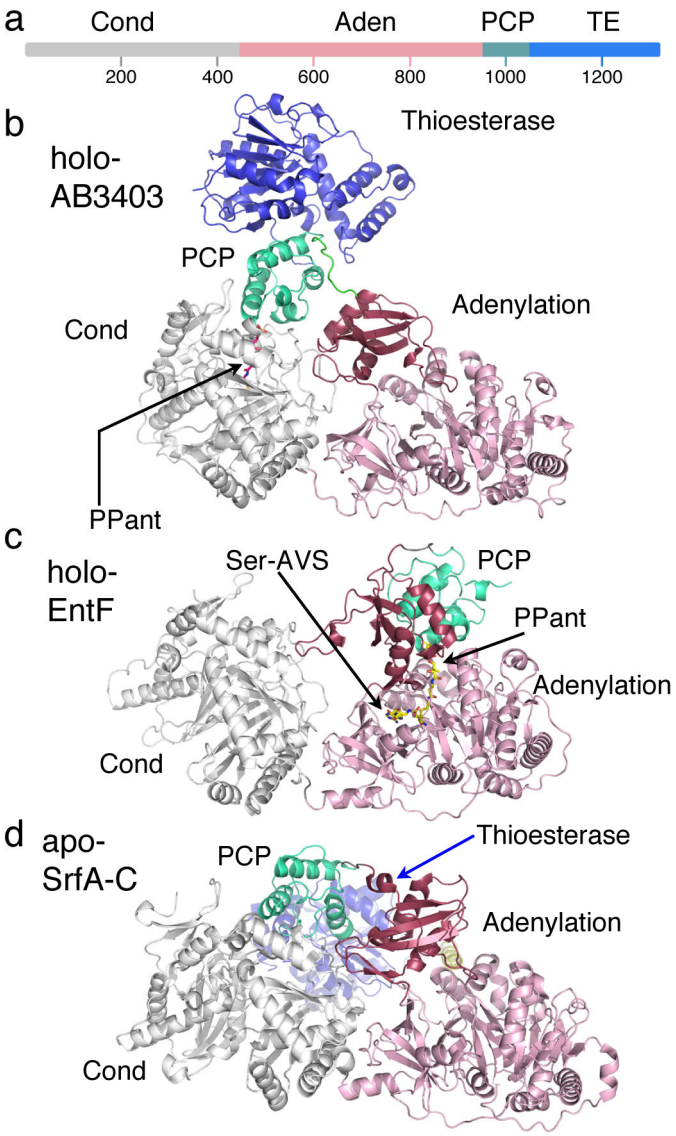


Figure 1. Ribbon diagrams of complete NRPS modules

(**a**) Domain architecture of three structurally characterized termination modules. The protein structures of (**b**) AB3403, (**c**) EntF, and (**d**) SrfA-C are colored with domains colored white (condensation), pink and red (adenylation domain N- and C-terminal subdomains), green-cyan (PCP), and blue (thioesterase). The phosphopantetheine moieties of AB3403 and EntF, and inhibitor Ser-AVS are highlighted.

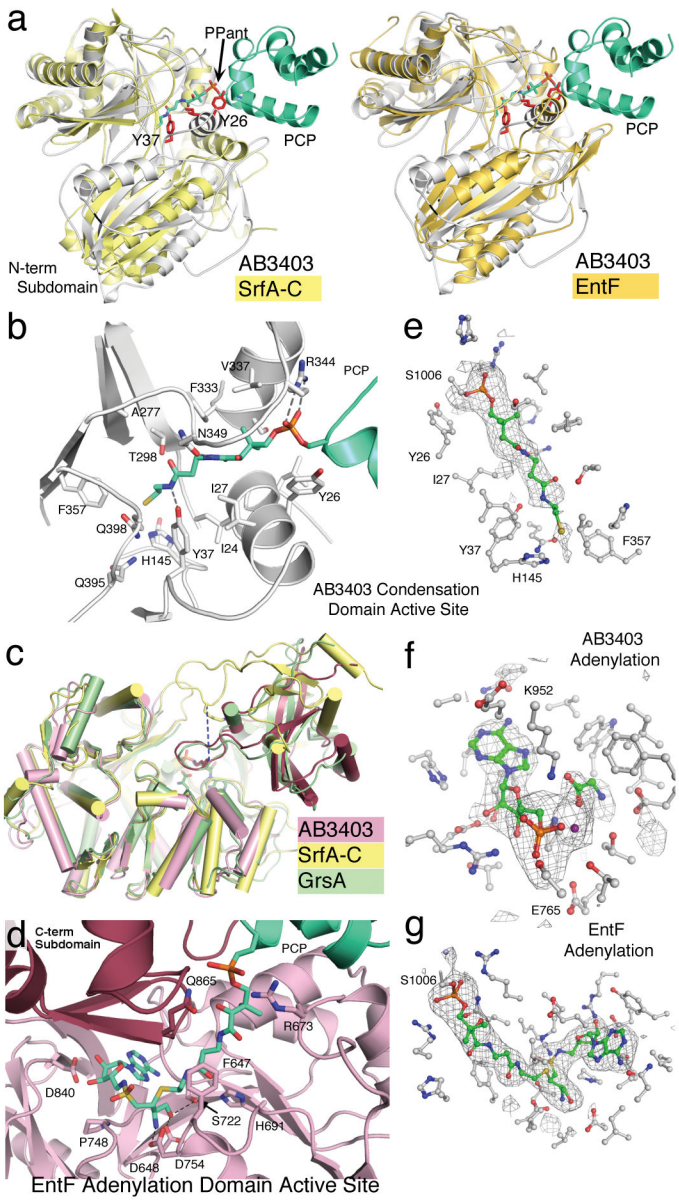


Figure 2. NRPS domain structures

The condensation domain of AB3403 (white) was aligned with (**a**) SrfA-C (yellow) and EntF (orange) on the basis of the condensation C-terminal subdomain. The AB3403 PCP is included. (**b**) The AB3403 condensation domain highlights residues that form the hydrophobic tunnel through which the pantetheine passes. (**c**) Superposition of adenylation domains of AB3403 (pink and maroon for N- and C-terminal subdomains), SrfA-C (yellow) and gramicidin synthetase, GrsA (cyan), with phenylalanine and AMP molecules of GrsA. The dotted line highlights the alternate position of the catalytic lysines of AB3403 and SrfA-C. (**d**) The EntF adenylation domain active site shows a covalent linkage from the pantetheine to the Ser-AVS inhibitor. Electron density calculated with coefficients of the form F_o-F_c generated prior to inclusion of ligands and contoured at 3σ, are shown for the (**e**) AB3403 condensation, (**f**) AB3403 adenylation, and (**g**) EntF adenylation domains.

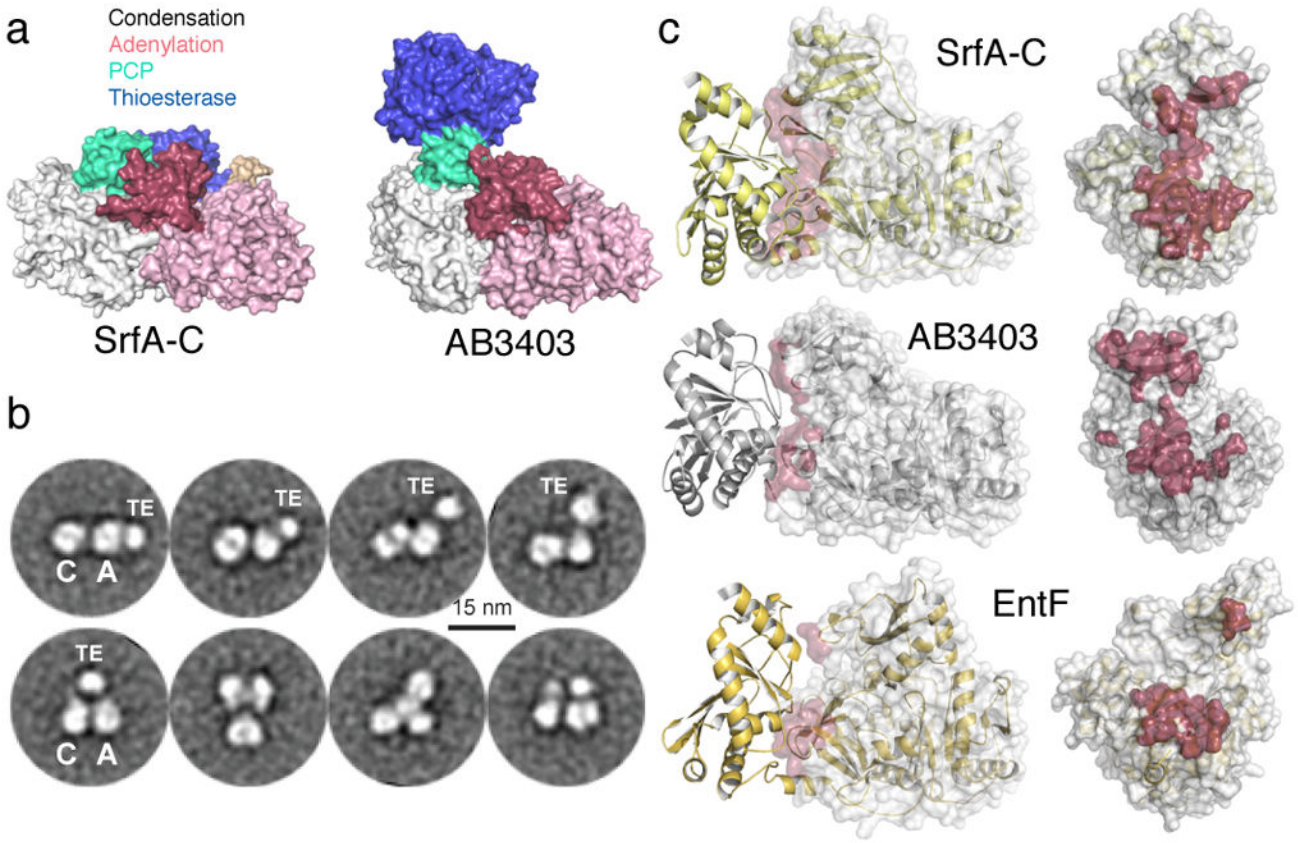


Figure 3. Conformational Dynamics in NRPS modules

(**A**) Alternate locations of the thioesterase domain SrfA-C and AB3403. (**b**) Representative EM class averages of EntF. The smaller TE domain is observed in various positions relative to the condensation-adenylation didomain. 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.

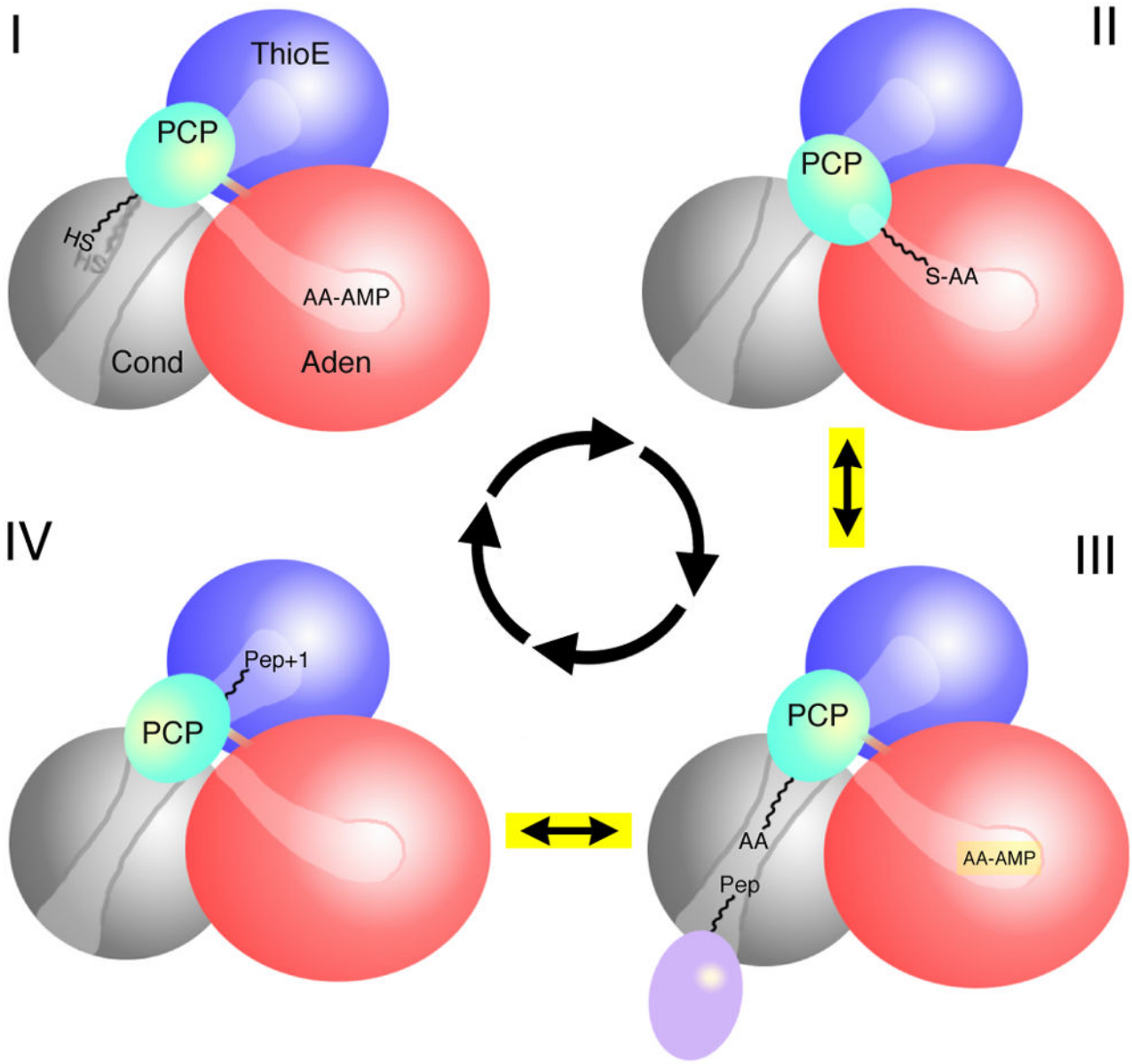


Figure 4. Dynamics of the NRPS cycle

The four stage catalytic cycle of an NRPS module. The revised NRPS structural cycle is highlighted in yellow showing that only three structural states are required.