Assembly line enzymology by multimodular nonribosomal peptide synthetases: the thioesterase domain of *E. coli* EntF catalyzes both elongation and cyclolactonization

Cathryn A Shaw-Reid[1], Neil L Kelleher[1], Heather C Losey[1], Amy M Gehring[1], Christian Berg[2] and Christopher T Walsh[1]

Background: EntF is a 142 kDa four domain (condensation–adenylation–peptidyl carrier protein–thioesterase) nonribosomal peptide synthetase (NRPS) enzyme that assembles the *Escherichia coli* *N*-acyl-serine trilactone siderophore enterobactin from serine, dihydroxybenzoate (DHB) and ATP with three other enzymes (EntB, EntD and EntE). To assess how EntF forms three ester linkages and cyclotrimerizes the covalent acyl enzyme DHB-Ser-S-PCP (peptidyl carrier protein) intermediate, we mutated residues of the proposed catalytic Ser–His–Asp triad of the thioesterase (TE) domain.

Results: The Ser1138→Cys mutant (k_{cat} decreased 1000-fold compared with wild-type EntF) releases both enterobactin (75%) and linear $(DHB\text{-}Ser)_2$ dimer (25%) as products. The His1271→Ala mutant (k_{cat} decreased 10,000-fold compared with wild-type EntF) releases only enterobactin, but accumulates both DHB-Ser-O-TE and $(DHB\text{-}Ser)_2$-O-TE acyl enzyme intermediates. Electrospray ionization and Fourier transform mass spectrometry of proteolytic digests were used to analyze the intermediates.

Conclusions: These results establish that the TE domain of EntF is both a cyclotrimerizing lactone synthetase and an elongation catalyst for ester-bond formation between covalently tethered DHB-Ser moieties, a new function for chain-termination TE domains found at the carboxyl termini of multimodular NRPSs and polyketide synthases.

Addresses: [1]Department of Biological Chemistry and Molecular Pharmacology, 240 Longwood Avenue, Harvard Medical School, Boston, MA 02115, USA. [2]Bruker Daltonics, Inc., Manning Parkway, Billerica, MA 01821, USA.

Correspondence: Christopher T Walsh
E-mail: walsh@walsh.med.harvard.edu

Key words: electrospray ionization, enterobactin, Fourier transform mass spectrometry, siderophore, thioesterase

Received: **10 February 1999**
Revisions requested: **5 March 1999**
Revisions received: **10 March 1999**
Accepted: **12 March 1999**

Published: **20 May 1999**

Chemistry & Biology June 1999, **6**:385–400
http://biomednet.com/elecref/1074552100600385

© Elsevier Science Ltd ISSN 1074-5521

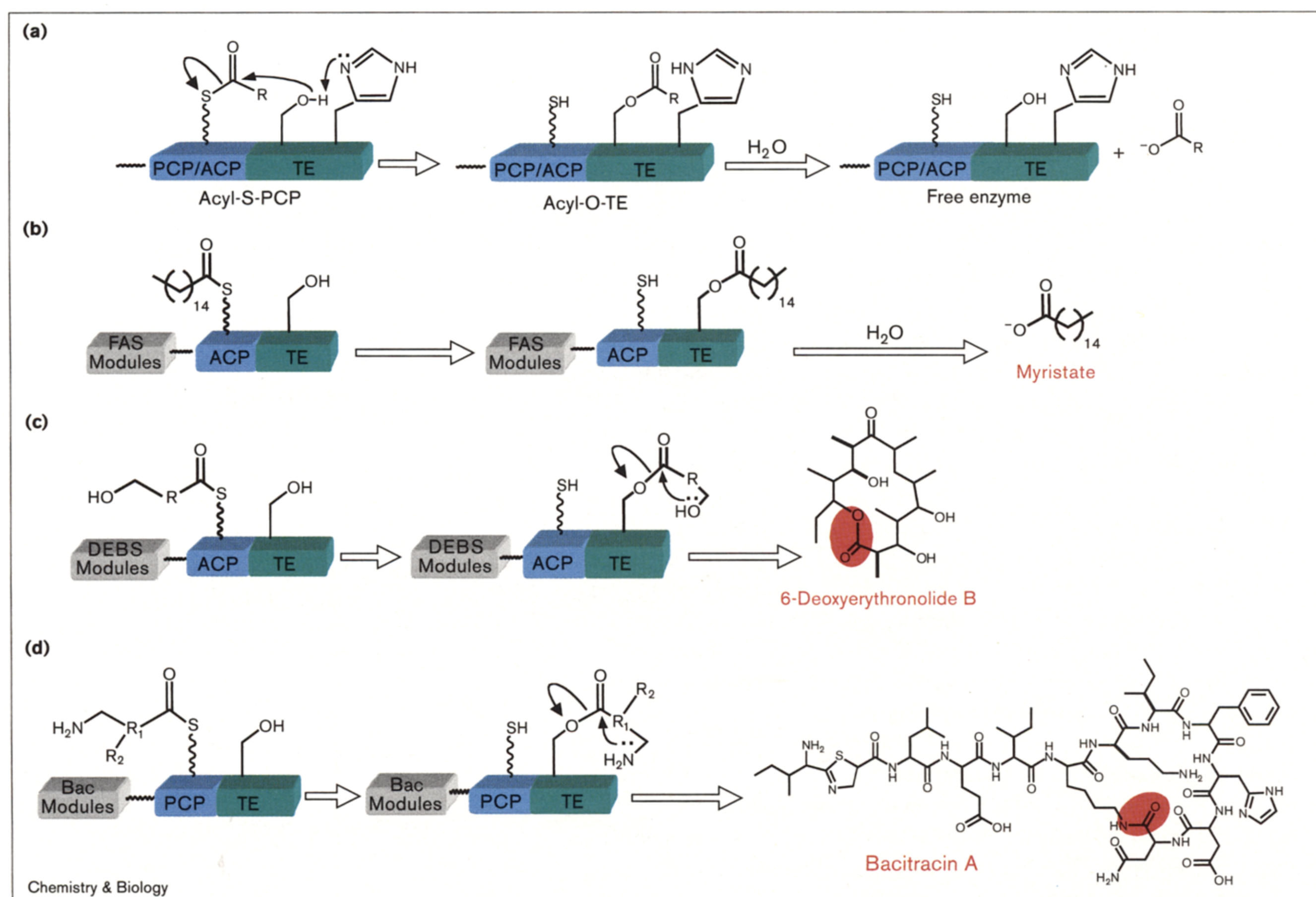
Introduction

All fatty acid synthases (FASs), polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs) operate with a common catalytic strategy in which the substrate chains are initiated and iteratively elongated as intermediates covalently tethered as acyl-S-enzyme species at carrier protein domains on the megasynthase catalysts [1–3]. This strategy first requires a priming of every acyl carrier protein (ACP) domain in FASs and PKSs and peptidyl carrier protein (PCP) domain in NRPSs by a family of selective phosphopantetheinyl transferases (PPTases) that introduce the 4′-phosphopantetheinyl (Ppant) moiety of Coenzyme A (CoA) onto a conserved serine residue in each carrier protein [4,5]. These holo carrier proteins thereby provide the thiol prosthetic groups on which substrate acyl chains can be assembled and translocated. Active holo synthases carry out the autoacylations that constitute both chain initiation and chain elongation steps as the growing fatty acyl, polyketide acyl and peptide acyl chains move from upstream to downstream carrier protein sites on the synthases. Finally, when the full-length natural product acyl chain reaches the most downstream ACP domain or PCP domain, the covalent acyl-S-ACP/PCP linkage must be broken in a chain termination step that regenerates the free enzyme and releases full-length product (Figure 1a).

To effect the acyl chain termination/release step most FASs, PKSs and NRPSs require acyl-thioesterase (TE) activity to cleave the full-length acyl-S-ACP/PCP nascent product. The multimodular NRPSs, PKSs and type I FASs have a 250–300 amino acid TE domain (~30 kDa) located as the most carboxy-terminal region, just downstream of the last ACP/PCP domain, for this catalytic acyl chain removal [1,3,6–8] (Figure 1b–d). In the multisubunit type I FASs there are also associated acyl-S-ACP thioesterases, which are separate proteins of 25–30 kDa and consequently have been the best characterized for structure–function analysis to date [1,9–12].

Determination of the X-ray crystal structure of LuxD, the 32 kDa myristoyl-S-ACP TE, from the photosynthetic bacterium *Vibrio harveyi* [13], suggests that fatty acyl TEs are members of the α/β structure superfamily that includes a number of lipases and acetylcholinesterases, with a proposed catalytic triad of Ser–His–Asp. Mutational analyses of LuxD and related TEs [11,14] have confirmed that the

Figure 1

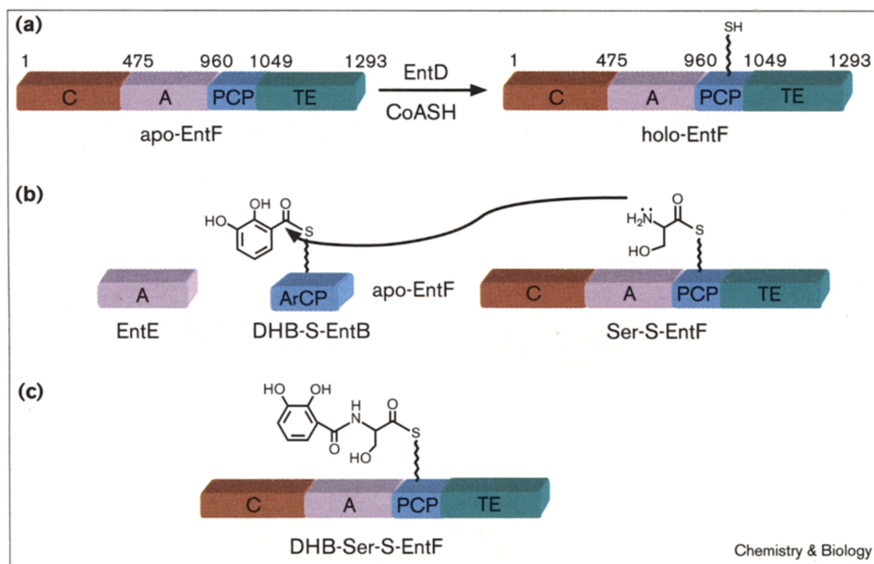


Acyl transfer catalysis by TE domains. **(a)** The generic scheme for termination by the TE domain as the active-site serine and histidine residues facilitate the transfer of an acyl group from the peptidyl carrier protein (PCP) or acyl carrier protein (ACP) to the conserved TE serine residue, followed by hydrolytic release of the free enzyme and final fatty acid, nonribosomal peptide synthetase (NRPS) or polyketide synthase (PKS) product. **(b)** Hydrolytic intermolecular release of a fatty acid product from a FAS. **(c)** Intramolecular capture to release a lactone ring, such as in the case of the PKS 6-deoxyerythronolide B, and **(d)** intramolecular capture to release a lactam ring such as in the case of the production of the NRPS bacitracin A. The 4′-phosphopantetheine cofactor tethered to the PCP/ACP domain is represented by a wavy line terminated by a thiol, and the sidechain of the TE active-site serine residue is represented by a solid line terminated by a hydroxyl moiety.

conserved serine in the G(Y/W/H)SxG signature motif is a catalytic nucleophile, and that the conserved histidine in the GxH motif is a catalytic base required for the formation of acyl-O-Ser-TE covalent intermediates during the hydrolysis reaction. Parallel studies with the ACP–TE double domain and the TE fragment from the carboxyl terminus of 6-deoxyerythronolide B synthase (DEBS) III have also validated catalytic competency and covalent acyl enzyme intermediacy [15–17]. Although there is a highly conserved aspartate located 27 residues downstream from the catalytic serine in the TE domains of NRPSs and the polyketide synthase DEBS [18], it is not clear if this TE domain harbors a catalytic Ser–His–Asp triad or the increasingly common catalytic Ser–His dyad [19].

Although FAS thioesterases and some NRPS TE domains may display purely hydrolytic function (Figure 1b), the TE domains at the carboxyl termini of PKSs (Figure 1c) and many NRPSs (Figure 1d) hypothetically reveal a broader acyl-transferase capacity, most notably the ability to direct internal capture by a sidechain nucleophile on the nascent product acyl chain. For polyketides, these are sidechain hydroxyl groups left in place by defective dehydratase domains in early elongation cycles. For example, in erythromycin biogenesis after six elongation cycles the DEBS III TE directs exclusive internal acyl transfer to the C-13 hydroxyl to yield the 14-member lactone ring of the deoxyerythronolide aglycone [7], prototypic for PKS-mediated macrolactonizations (Figure 1c). In NRPS action, directed macrocyclization is presumably mediated by the carboxyl-terminal TE domain either from hydroxyl sidechains derived from β-hydroxy amino acid sidechains (e.g. threonine in the streptogramin B family of depsipeptides, including pristinamycin and virginiamycin [18]) or

Figure 2



Participation of EntF domains in the initial priming reactions. **(a)** The PCP domain (residues 961–1049) is post-translationally modified at Ser1006 by the 4′-phosphopantetheinyl (Ppant) moiety from CoASH in a reaction catalyzed by the Ppant transferase EntD. **(b)** After the adenylation domains of EntE and EntF (A, residues 476–960) activate 2,3-dihydroxybenzoate (DHB) and serine, respectively, aminoacylation can proceed as the nucleophilic Ppant tether sequesters its substrate. The condensation domain of EntF (C, residues 1–475) subsequently facilitates amide-bond formation. **(c)** As a result, the DHB-Ser-S-PCP species is covalently bound to the PCP domain before it is transferred to the carboxy-terminal TE domain (residues 1050–1293).

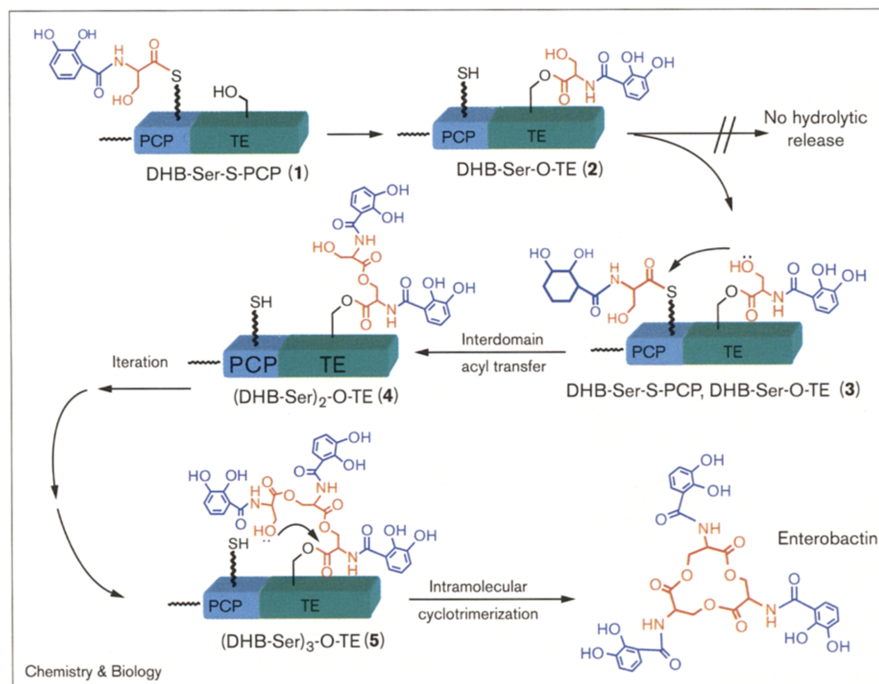
from β-hydroxy groups of *N*-capped fatty acid moieties (e.g. surfactins and iturins [6,20,21]). NRPS TEs might also catalyze internal capture during acyl chain release to sidechain amino groups to produce the macrolactams typical of the linkages in bacitracins (Figure 1d) [22]. An augmented understanding of TE domain specificity and recognition could therefore facilitate the deconvolution and/or reprogramming of the chain-termination steps in polyketide and nonribosomal peptide production.

One NRPS that is amenable to the mechanistic study of chain-termination strategy is the EntF protein responsible for orchestrating the last stages of assembly of the *Escherichia coli* iron-scavenging siderophore enterobactin [23]. The 142 kDa EntF has four domains (Figure 2a): condensation (C; residues 1–475) for amide-bond formation, adenylation (A; residues 476–960) for serine activation, PCP (residues 961–1049) for tethering serine, and the carboxy-terminal TE domain (residues 1050–1293). Ancillary proteins that participate with EntF to form the tris-*N*-dihydroxybenzoylated-serine trilactone enterobactin include EntD [4] (Figure 2a), the dedicated PPTase that primes both the apo-PCP domain of EntF and the apo-aryl carrier protein (ArCP) domain at the carboxyl terminus (100 residues) of the protein EntB [24]. The amino terminus (200 residues) of the bifunctional protein EntB contains isochorismate lyase catalytic activity, which is required at an early step in the enterobactin biosynthetic pathway [24,25]. The fourth protein required for enterobactin formation is EntE [23,26], a 2,3-dihydroxybenzoyl-AMP ligase that, in addition to activating dihydroxybenzoate (DHB), catalyzes its transfer to the holo-ArCP domain of EntB to yield DHB-S-EntB (Figure 2b). The DHB-S-EntB is the donor and Ser-S-EntF the nucleophilic acceptor in the amide-bond forming step catalyzed by EntF (via its condensation domain) to yield DHB-Ser-S-EntF [23], an acyl enzyme intermediate with the DHB-Ser tethered to the thiol of the phosphopantetheinylated PCP domain (Figure 2c).

When pure EntE, holo-EntB, holo-EntF, ATP, serine and DHB are mixed together, robust reconstitution of enterobactin synthetase activity occurs, with a k_{cat} value of 100–200 min^{-1} [23]. This reconstitutive turnover means that EntF can cyclotrimerize the monomeric DHB-Ser acyl enzyme intermediate (Figure 3, **1**) immediately before release by the TE domain. In addition to its presumptive catalytic role in lactonization to effect covalent acyl-S-PCP-EntF breakdown, the EntF TE domain is of interest because it might also play a role in chain elongation. We previously noted the dilemma that EntF has a single PCP domain and yet functions as a monomer [23], so it has no obvious way to generate the additional two covalent DHB-Ser-acyl-TE intermediates required for DHB-Ser dimeric and then DHB-Ser trimeric acyl enzyme species, unless both the PCP and TE domains of EntF can be covalently loaded simultaneously (Figure 3, **3**) and the elongating chain transferred onto the TE domain to yield **4** and **5** in Figure 3. This would be a novel role for a TE domain in an NRPS or PKS — chain elongation as well as termination.

Covalent catalysis by the TE domain of EntF has been postulated [23], but not detected previously. Mass spectrometry with 'soft' ionization techniques, such as electrospray ionization (ESI), allows large biomolecules to be analyzed [27]. The combination of ESI with Fourier transform mass spectrometry (FTMS) offers high resolving power (10^6) and mass accuracy (low ppm) even for large fragment ions or

Figure 3



Proposed mechanism for elongation, lactonization and termination of enterobactin by the EntF TE domain. The TE domain serves as a 'waiting room' for interdomain acyl transfers of covalent intermediates (species **1–5**) from the upstream peptidyl carrier protein and it is the putative location for intramolecular cyclotrimerization to release enterobactin.

peptides (>10 kDa) in complex mixtures resulting from protein degradation in gas or solution [28]. ESI/FTMS is utilized here to detect covalent intermediates bound to EntF during enterobactin biosynthesis. Unfractionated proteolytic digests are analyzed directly (no chromatography), and resolution of peptide isotopic peaks allows the use of stable isotopes to provide further structural evidence for covalently tethered biosynthetic intermediates.

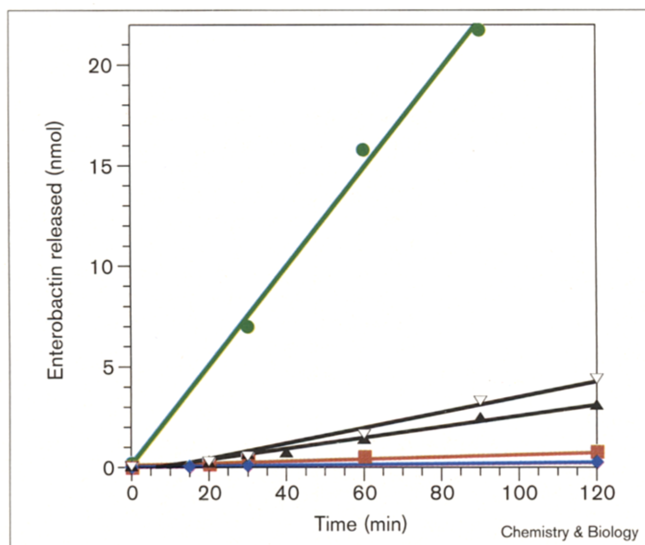
In this work we focus predominantly on the behavior of two TE domain mutants of EntF: Ser1138→Cys (S1138C) and His1271→Ala (H1271A). The turnover rate of these mutants is substantially decreased, allowing us to demonstrate conclusively that the linear DHB-Ser dimer is an intermediate in the reaction. During turnover with the S1138C mutant the dimer is prematurely released, whereas during turnover of H1271A we found the dimer and the DHB–Ser monomer covalently tethered to the active site of the TE domain as discrete acyl enzyme intermediates. Mutating the conserved Asp1165 reveals that this residue, the third in a putative catalytic triad, plays only a minor role in enterobactin formation.

Results

Preparation of EntF TE domain mutants

In previous studies we reported the mutation of the proposed active-site serine residue in the G(Y/W/H)SxG motif of the EntF TE domain to alanine, S1138A, and observed no retention of any detectable enterobactin synthetase activity [23], consistent with the blockage of the growing acyl chain transfer from the PCP to the TE domain. The nucleophilic serine is substituted by cysteine in only two NRPS TE domains known to date: the synthetases producing the siderophores mycobactin (MbtB [29]) and pyochelin (PchF [30]). Accordingly, construction of the S1138C mutant was pursued with the view that some degree of catalytic competency of TE domain function might be retained if the cysteine sidechain could function as a nucleophile. The conserved histidine found in most NRPS TE domains [6,18] is thought to function in general base catalysis for one or more of the acyl-transfer steps. In EntF, His1271 was that candidate residue and the H1271A mutant was constructed. The proposed role of the conserved acidic aspartate in the NRPS TE domain is to function as a possible hydrogen acceptor from the catalytic histidine base. In this work, Asp1165 in EntF was the putative third member of the triad chosen for mutation to test this hypothesis. All four EntF mutants (S1138C, H1271A, Asp1165→Ala, D1165A, and Asp1165→Ser, D1165S) were purified in comparable yields as previously described wild-type EntF [23]. Initial evaluation and confirmation of catalytic activities for all EntF mutants included activation of serine by the adenylation domain in serine-dependent [^{32}P]PP$_i$-ATP exchange assays [23], in which the mutants adenylated serine at the same rate (~300 min^{-1}) as wild-type EntF and post-translational Ppant priming of the apo-PCP domains in the presence of the PPTase EntD [4] (data not shown). When [^{3}H]CoASH was used in the PPTase precipitation assay, a saturation level of Ppant loading (60–70% of the anticipated stoichiometry) was

Figure 4



Rates of enterobactin released from wild-type and mutant EntF proteins. A radiolabel precipitation assay was used to quantitate the amount of [^{3}H]serine covalently bound to DHB and extracted into ethyl acetate as a measure of enterobactin formation. Wild-type EntF (10 nM, green circles) and EntF TE mutants (1 μM S1138C, red squares; 2.7 μM H1271A, blue diamonds; 50 nM D1165A, black closed triangles; and 450 nM D1165S, black open triangles) are shown.

achieved after 10–20 min in all mutant proteins, as demonstrated previously with wild-type EntF [4,23].

Catalytic characterization of the S1138C mutant

The S1138C mutant was assayed for full catalytic competency to make the three amide and ester bonds and release enterobactin in assays containing [^{3}H]serine, DHB, ATP, the PPTase EntD, the DHB-activating ligase EntE and the ArCP EntB. The initial assay for product formation was extraction of [^{3}H]serine radiolabel into ethyl acetate as the water-soluble serine substrate was converted to a hydrophobic DHB-derivatized product [23]. Under conditions in which wild-type EntF had a turnover rate of about 100 min^{-1}, the S1138C mutant retained a linear low level activity of 0.09 min^{-1} (over a 2 h incubation period; Figure 4 and Table 1). Further analysis by high-pressure liquid chromatography (HPLC; Figure 5a) revealed that in addition to the final cyclotrimeric product enterobactin there was an additional product released only in the S1138C mutant incubations. By comparison with retention times and coinjection of authentic samples [23] the new peak was identified as the linear dimeric ester of DHB-serine, and its identity was confirmed using mass spectrometry (theoretical relative molecular weight, M_r, was 464, and the experimental M_r was 463 under negative ion conditions; Figure 5b). It is likely that the extra species in the ESI mass spectrum with an M_r value of 240 (Figure 5b) corresponds to a

Table 1

Rates of enterobactin and (DHB-Ser)$_n$ release.

	Cyclized enterobactin (min^{-1})	DHB-Ser monomer (min^{-1})	(DHB-Ser)$_2$ dimer (min^{-1})	(DHB-Ser)$_3$ trimer (min^{-1})
Wild-type EntF	80–200	nd	nd	nd
S1138A*	nd	0.06	nd	nd
S1138C	0.09	nd	0.03	nd
H1271A	0.01	nd	nd	nd
D1165A	5.0	nd	nd	nd
D1165S	0.4–0.8	nd	nd	nd

*Datum reported in [23]. nd, not detectable.

DHB-Ser monomer byproduct that is formed under MS conditions (this species is also present in the mass spectrum of the purified synthetic linear dimeric ester; data not shown). The flux to the DHB-Ser-O-DHB-Ser dimer was approximately 25–33% (~0.03 min^{-1}; Table 1) of the flux towards enterobactin. The mutation of Ser1138 to cysteine in the TE domain therefore lowers the overall k_{cat} by 800–1000-fold and re-routes some of the flux to release an uncyclized ester dimer. Hydrolysis of the incomplete dimeric acyl enzyme intermediate (e.g. Figure 3, **4**) presumably competes with further elongation and cyclization to the cyclic ester trimer enterobactin.

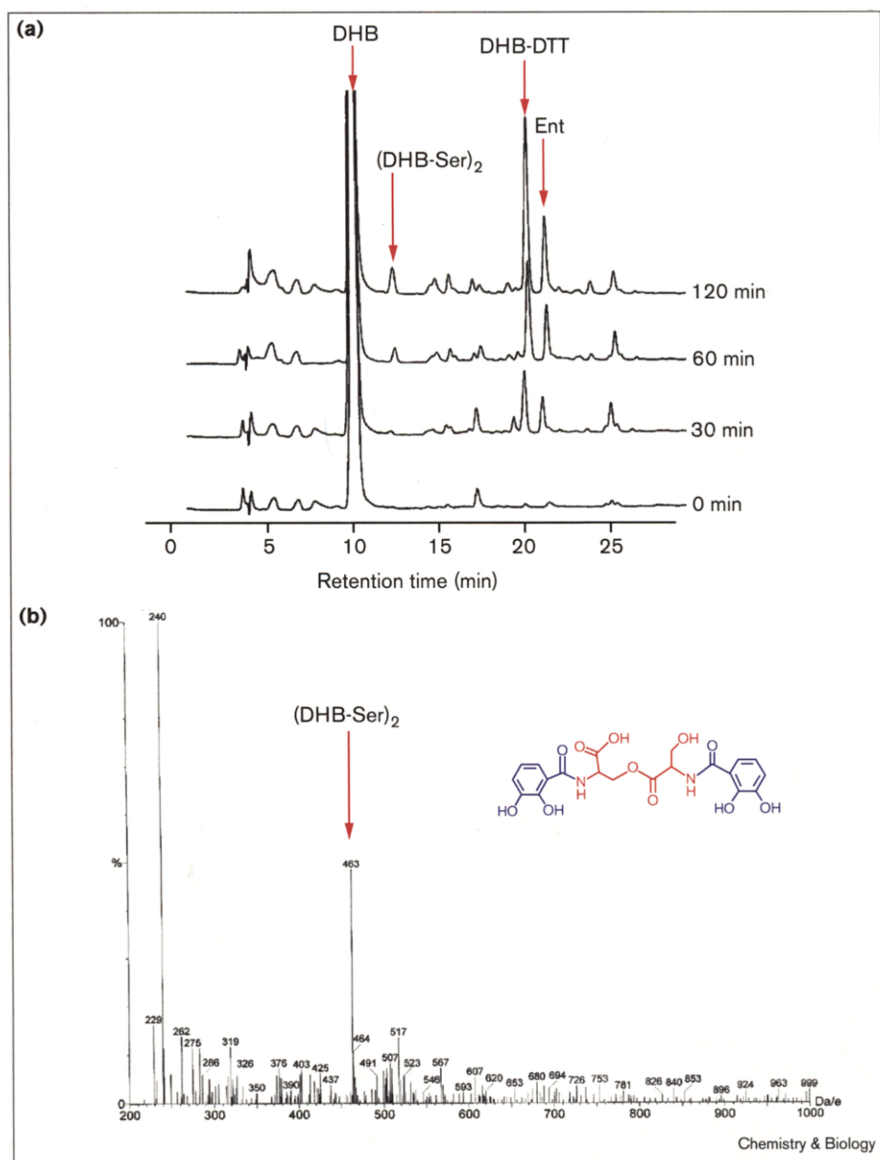
Enterobactin production by the H1271A mutant

When the H1271A mutant was examined for its ability to produce enterobactin in the radiolabel ethyl acetate extraction and HPLC assays, its turnover rate was down another order of magnitude compared with the S1138C mutant, with a k_{cat} ~0.01 min^{-1} (Figure 4 and Table 1) calculated from a signal clearly above background but yielding a rate ~8000–10,000-fold impaired relative to wild-type EntF. When the H1271A mutant was incubated with 100 mM imidazole, the enterobactin release rate increased tenfold (k_{cat} ~0.1 min^{-1} versus 100 min^{-1} for wild-type EntF; data not shown), indicating that the imidazole ring of H1271A can be replaced by imidazole binding from solution. These data support the role of His1271 as a general base in enterobactin TE catalysis. In HPLC experiments run with and without imidazole, the release of enterobactin (DHB-Ser)$_n$ intermediates (monomer, dimer and uncyclized trimer) were undetectable under conditions where 20% of the flux could have been visible.

Assessing the role of Asp1165 in enterobactin formation

The D1165A mutant was similar to wild-type EntF in that it released exclusively enterobactin at turnover rates within 20-fold of wild-type (k_{cat} ~5 min^{-1}, Figure 4 and Table 1), as shown by both radiolabeled ethyl acetate extraction and HPLC assays. Activity of the D1165S mutant was also evaluated because serine replaces the conserved aspartate in five known NRPS examples (MtbB [29], PchF [30], PvdS [31], YukK [32] and external SrfTE [33]). Surprisingly, the

Figure 5



Detection of dimeric $(DHB\text{-}Ser)_2$ released from the S1138C mutant. **(a)** HPLC traces demonstrate that both enterobactin (Ent) and the dimer $(DHB\text{-}Ser)_2$ intermediate are steadily released from S1138C and extracted into ethyl acetate during a 2 h incubation under full-turnover conditions. The identity of the DHB-dithiothreitol (DTT) adduct was confirmed by mass spectrometry (ESI/LCMS, data not shown). **(b)** Mass spectrometry (ESI/LCMS) confirmed the identification of the $(DHB\text{-}Ser)_2$ dimer ester released from the S1138C mutant with a relative mass of 463 in negative ion mode.

D1165S mutant has 200-fold lower activity (k_{cat} ~0.5 min^{-1}, Figure 4 and Table 1) than D1165A, so the serine mutant did not rescue turnover. These results suggest that Asp1165 is not essential for TE catalysis.

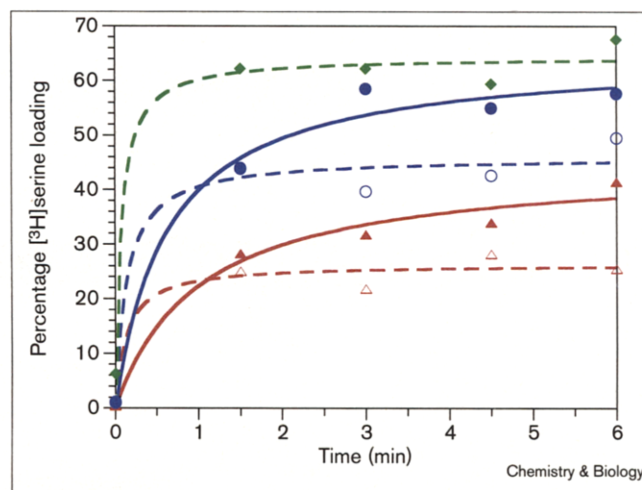
Covalent loading of [^{3}H]serine onto EntF wild-type and mutants

With pure preparations of wild-type and EntF proteins mutated in the proposed active site of the TE domain (S1138A, S1138C and H1271A), the levels of serine covalently incorporated could, in principle, be assessed using trichloroacetic acid (TCA)-precipitation assays followed by washing to remove physically trapped amino acids, a classic assay in the NRPS literature for detecting aminoacyl and peptidyl enzyme covalent intermediates [23,34]. We previously noted [23] that the high k_{cat} of wild-type EntF (~100–200 min^{-1}) leads to variable amounts of the released enterobactin product being trapped in TCA precipitates and gives such variable noise that signal cannot be reliably assessed in attempts to measure the covalently bound intermediates. In the absence of the cosubstrate DHB, [^{3}H]serine is loaded onto the holo-EntF (but not apo-EntF; data not shown) in a rapid aminoacylation of the Ppant thiol moiety in the PCP domain, allowing an estimate of the functional number of holo-PCP sites.

The S1138A mutant can be aminoacylated (as the PCP Ppant-thiol) but not transferred to the TE domain, even in the presence of DHB. As previously shown [23], the S1138A mutant consistently loads serine, both in the

presence and absence of DHB, with a fractional stoichiometry (~60% acylation) that is comparable with the calculated number of holo-PCP sites in wild-type EntF (without DHB). In contrast, the S1138C mutant aminoacylated to about 20% fractional stoichiometry (without DHB) and labeling was increased 10–15% (mol/mol) upon addition of DHB to the incubation (Figure 6). The H1271A mutant typically loaded ~40% of the anticipated stoichiometric amount of [^{3}H]serine in the absence of DHB, and consistently gave an additional 20% increase in stoichiometry of fractional acylation in the presence of DHB (Figure 6). The reasons for the low steady-state aminoacylation level by [^{3}H]serine in the S1138C mutant are unclear, but these results might reflect some problem in interdomainal (TE to PCP) communication that destabilizes the Ser-S-PCP thioester intermediate.

Figure 6



[^{3}H]Serine precipitation radioassay to measure stoichiometry of covalent loading of EntF wild-type and mutant proteins in the presence (solid line) and absence (dotted line) of DHB. Wild-type EntF (1 μM, green diamonds), the S1138C (4.8 μM, red triangles) and H1271A (2.4 μM, blue circles) mutants are shown.

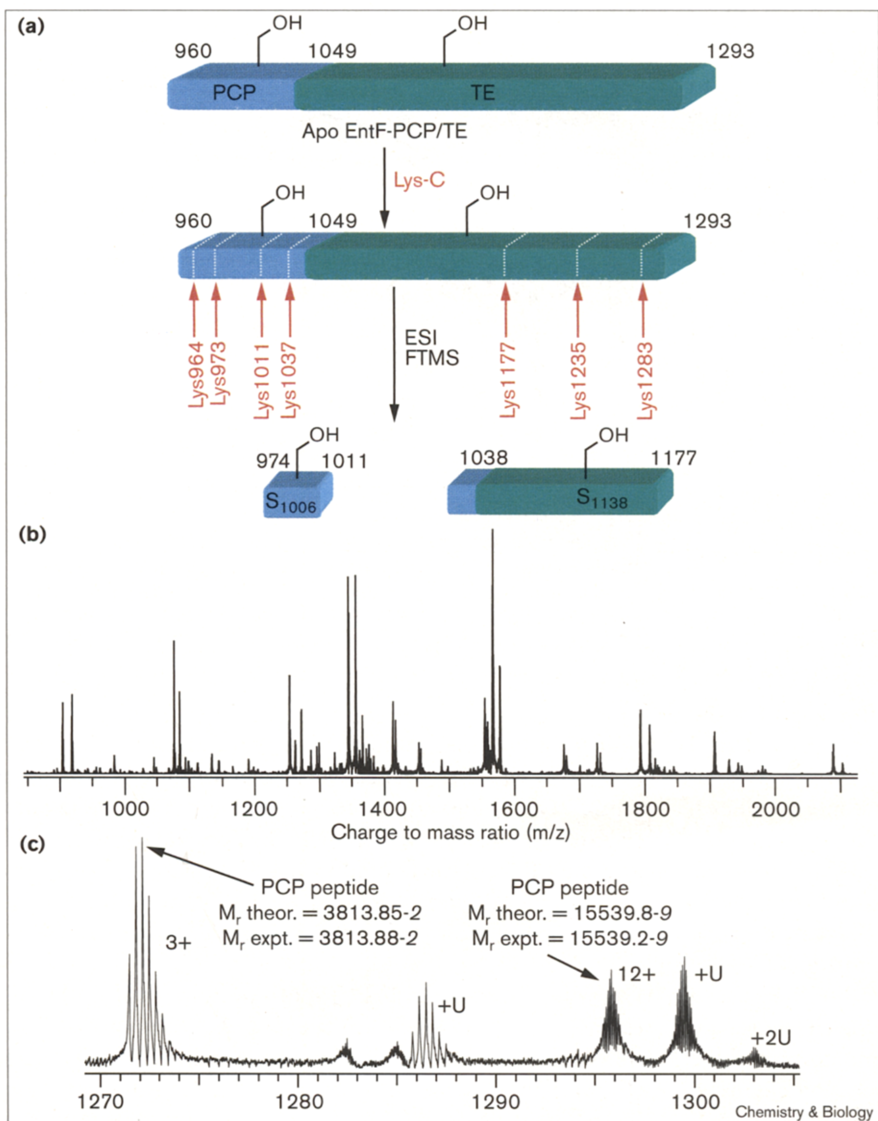
One might have anticipated that in the case of the low k_{cat} mutants, if Figure 3 were operational and the 10^3–10^4 drop in turnover rate were due to slow deacylation of various acyl-O-TE species (**3**, **4** or **5**), suprastoichiometric levels of radioactivity would accumulate in these EntF mutants, up to 2, 3 or even 4 equivalents respectively. We found that the [^{3}H]serine acylation stoichiometry was always less than one nominal equivalent, perhaps reflecting a collage of effects including rate-determining steps earlier than **4** or **5** (in Figure 3) and/or hydrolytic lability of various acyl-S-PCP, acyl-O-TE or acyl-S-TE covalent intermediates. Although the consistent increase in precipitable [^{3}H]serine found in the S1138C and H1271A mutants under turnover conditions (with DHB present) might just represent an increase in the DHB-Ser-S-PCP acyl enzyme amount over Ser-S-PCP acyl enzyme, this result could also demonstrate a signal for transfer of the DHB-Ser acyl group to the TE domain and reloading at the PCP site. To establish the identity of the acyl enzyme intermediate species (Figure 3, **1–5**) and to localize such acyl chains to the PCP domain and/or the TE domain, we turned to MS of EntF proteolytic fragments.

ESI/FTMS of the EntF-PCP–TE double domain

The first goal of the initial MS experiments was to identify EntF peptides containing the active-site serines of the PCP domain (Ser1006) and the TE domain (Ser1138). To detect active-site EntF peptides, we subcloned, expressed and purified the 343-residue EntF-PCP–TE protein, the carboxy-terminal third of full-length EntF (Leu960–Arg1293) followed by a two-residue linker and a carboxy-terminal hexahistidine tag. The predicted M_r of the two-domain EntF-PCP–TE peptide is 37496.0-*23* (where the mass difference between the most abundant isotopic peak and the monoisotopic peak is denoted in italics after each M_r value). Its ESI/FT mass spectrum yielded isotopic resolution and an M_r value of 37495.4-*23* (data not shown), within 1 Da of the expected mass. Fragmentation of the ions of this 37.5 kDa protein yielded 13 b-type and 7 y-type ions (fragment ions that contain the amino and carboxyl termini of EntF-PCP–TE, respectively) that further verified the entire DNA-derived protein sequence (MS data not shown). Incubation of the PCP–TE protein with CoASH and the PPTase EntD produced a +340.6 Da shift of ~70% of the material (data not shown), consistent with covalent phosphopantetheinylation of two thirds of the protein.

Digestion of the EntF-PCP–TE protein with endoproteinase Lys-C (Figure 7a) yielded 95 isotopic distributions representing 14 peptides from 1 to 28 kDa in size (Figure 7b). In this spectrum, the 38-residue peptide (corresponding to Ala974–Lys1011 in full-length EntF) containing the PCP consensus site for phosphopantetheinylation (Ser1006 in full-length EntF) was detected with an observed M_r value of 3813.88-*2* (Figure 7c), which confirms the theoretical value (3813.85-*2*) within 8 ppm. In addition, the 140-residue peptide (corresponding to Leu1038–Lys1177 in full-length EntF) harboring the active-site conserved serine (Ser1138 in full-length EntF) of the TE domain was observed (Figure 7b) with an M_r value of 15539.2-*9*, which is consistent with its theoretical value (–0.6 error). This TE domain peptide can now be used to indicate the covalent state of this portion of EntF during enterobactin biosynthesis during catalytic turnover. Both the Ala974–Lys1011 and the Leu1038–Lys1177 peptides were observed with covalent adducts of +43 Da (Figure 7c, '+U'). These species are consistent with carbamylation of lysine residues in aged urea solutions. Fresh urea in subsequent samples prevented further carbamylation events.

Figure 7



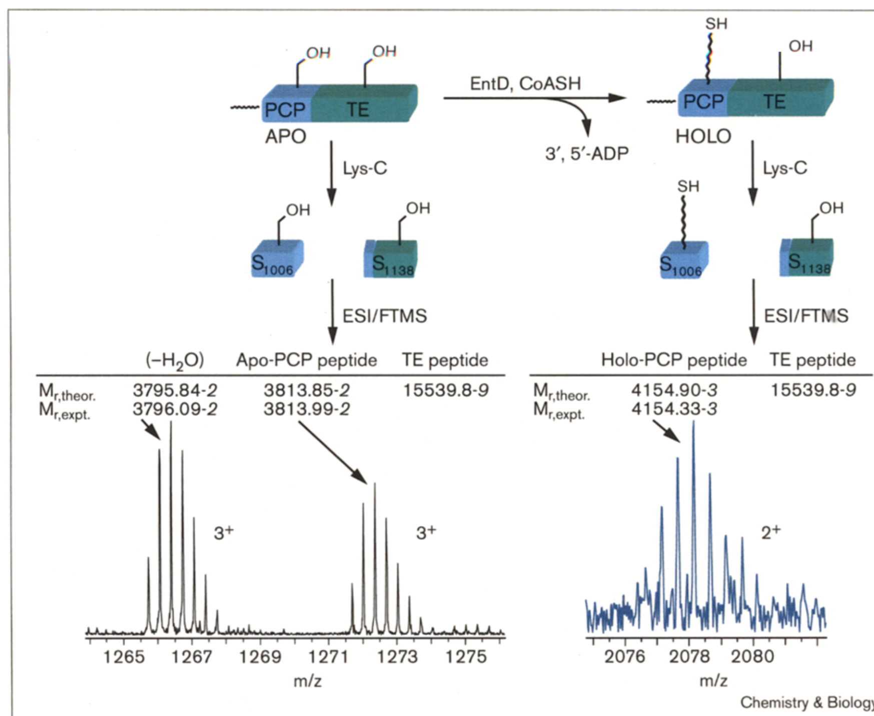
Proteolysis and ESI/FTMS of EntF-PCP–TE. **(a)** When the 37.5 kDa apo-EntF-PCP–TE protein is digested with endoproteinase Lys-C (cleavage sites indicated by red arrows), eight peptide fragments are generated. The peptide (Ala974–Lys1011) containing the PCP domain active-site serine (Ser1006) is found in a separate fragment to the peptide (Leu1038–Lys1177) containing the TE domain active site serine (Ser1138). **(b)** ESI/FT mass spectrum (64 scans) of the unfractionated Lys-C digest of apo-EntF-PCP–TE. **(c)** Expansion of the 1270–1305 *m/z* region; U, urea adduct (+43.0 Da).

When the holo-Ppant form of the EntF-PCP–TE fragment was analogously proteolyzed with Lys-C and analyzed using ESI/FTMS, a new isotopic distribution was observed at ~2078 m/z with an observed M_r value of 4154.33-*3*, which is within 0.6 of the predicted value (Figure 8 and Table 2). This value is consistent with addition of a Ppant moiety (+340.05 Da) to Ser1006 of the PCP domain and demonstrates that ESI/FTMS can monitor covalent alterations in this region. The signal-to-noise ratio is only 13:1, however, and addition of the Ppant moiety apparently reduces the charge state of the PCP domain (from 3^+ to 2^+). Also visible is the isotopic abundance for the apo-PCP peptide and an additional intense signal that probably corresponds to water loss with an M_r value 17.9 lower than the apo-PCP peptide (Figure 8). This water loss was not observed for the 100% apo-PCP peptide of Figure 7 and could therefore be due to β-elimination of the Ppant group in the gas phase. Variation of ESI source conditions did not prevent formation of this water loss isotopic distribution. Although the incomplete formation of holo enzyme and β-elimination of Ppant dramatically lowers the sensitivity for holo forms of the PCP peptide, the MS results of proteolyzed apo- and holo-EntF-PCP–TE (Figure 8) confirms that in this 37.5 kDa fragment of EntF one can unambiguously locate the active-site serine peptides in both the PCP and TE domains.

Lys-C digestion and ESI/FTMS analysis of full-length EntF to detect Ser1006 (PCP) and Ser1138 (TE) peptides

The Lys-C digest of the 1293-residue full-length EntF (142 kDa) yielded 150 isotopic distributions representing ~80 peptides of 2–40 kDa (Figure 9a). Still visible in this

Figure 8



Detection of the holo-PCP peptide in EntF-PCP–TE by ESI/FTMS. Priming of the PCP domain with 4′-phosphopantetheine (Ppant) occurs by the catalytic action of the Ppant transferase EntD. Lys-C digestion followed by mass spectrometry (64 scans) reveals a distinctive holo-PCP peptide (2+ ionization state) that has a higher mass (+340) than the apo-PCP peptide (3+ ionization state). This spectrum also contains an isotopic distribution corresponding to water loss (−17.9) from the apo-PCP peptide; this peak may reflect β-elimination of holo-PCP in the gas phase. These segments of the ESI/FT mass spectrum of apo- and holo-EntF-PCP–TE represent just the PCP peptides and their mass shifts. The TE peptides do not undergo covalent modification during phosphopantetheinylation.

more complex mixture are the isotopic distributions for the PCP peptide (Ala974–Lys1011; 3+ at 1272 m/z) and the TE peptide (Leu1038–Lys1177; 12+ at 1296 m/z; Figure 9b–c, Table 2). The experimental M_r value of the TE peptide is 2.6 lower than that calculated from its DNA-predicted sequence (15539.8-9). This shift might be due to nonspecific oxidation causing disulfide bond formation (-SH+HS- → -S–S- + $2H^+$; –2.0 Da). None of the 26 peptides predicted from exhaustive Lys-C digestion nor any other peptide from incomplete digestion fits this M_r value within 173.5. The stage was now set to probe full-length EntF mutants (e.g. H1271A) for acyl enzyme species during slow turnover.

ESI/FTMS analysis of EntF mutants to detect covalent intermediates

When the full-length catalytically competent wild-type EntF (k_{cat} ~100 min^{-1}), the S1138C mutant (k_{cat} ~0.09 min^{-1}) and the H1271A mutant (k_{cat} ~0.01 min^{-1}) proteins were incubated under turnover conditions for 1 h and subsequently proteolyzed with Lys-C for 1 h, the peptide mixtures could be probed for covalent linkages on the PCP peptide and/or the TE peptide using ESI/FTMS. Given the turnover numbers and the length of the incubations it was not anticipated that wild-type EntF would accrue significant amounts of acyl enzyme species (such as **1–5** in Figure 3), and none was detected. The S1138C mutant digests likewise provided no new spectral peaks showing covalent linkages to the Ser1006 or Ser1138 peptides.

The combination of Lys-C proteolysis and ESI/FTMS was useful for detecting multiple acyl enzyme species covalently bound to H1271A mutant, however. First, the H1271A mutation was verified by noting a shift of –66.0 Da in the Ala1236–Lys1282 peptide fragment containing His1271 (data not shown). When holo-H1271A was incubated with serine and ATP analogous to the stoichiometric-loading experiments, a new isotopic distribution was detected at 1415 m/z with a charge state of 3^+ and M_r value of 4240.95-*2* (Figure 9d, Table 2). This new

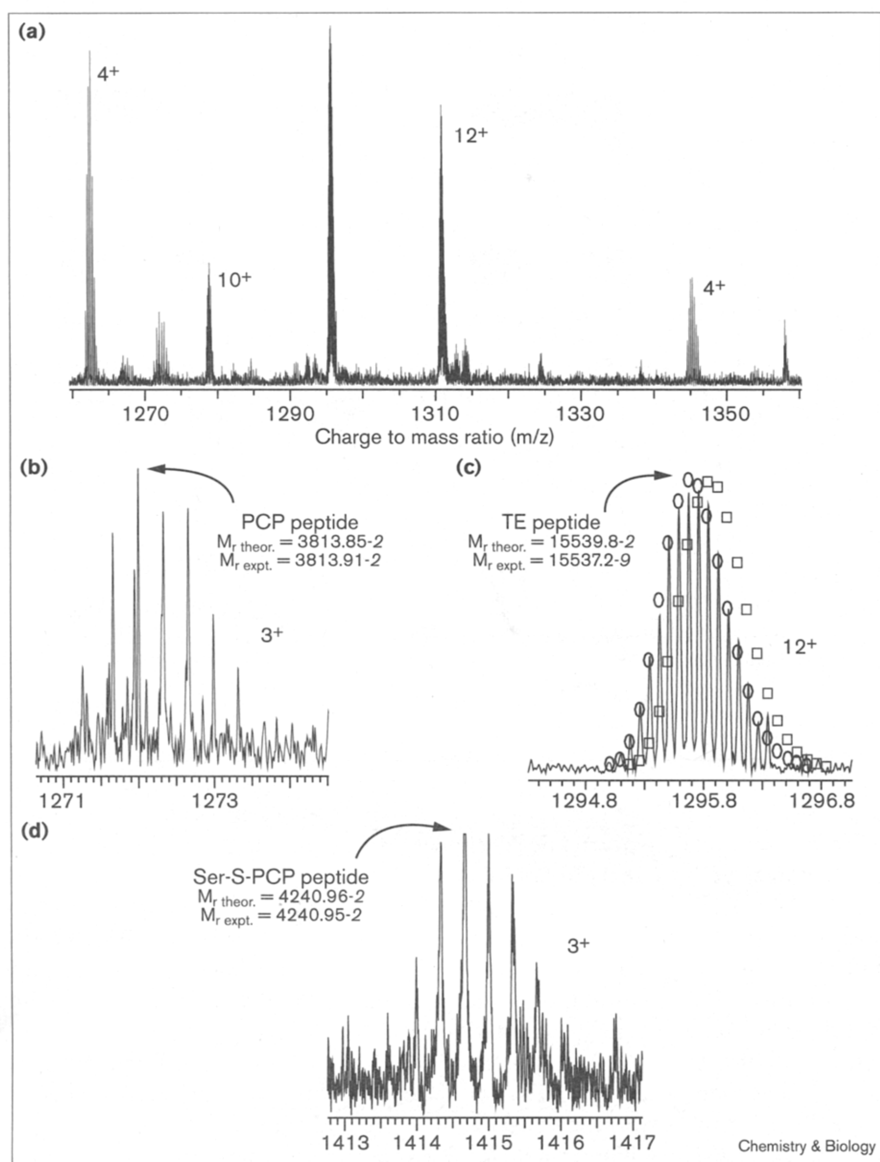
Table 2

Covalent acyl enzyme intermediates observed experimentally.

Acyl group tether	Mass increase*	M_r, expt PCP peptide	M_r, expt TE peptide
apo	0	38123.88-*2*	15539.8-9
holo (Ppant)	340	4154.33-*3*	nd
holo-Ser	427	4240.95-*2*	nd
DHB-Ser	223	nd	15762.4-9
$(DHB\text{-}Ser)_2$	446	nd	15985.2-9
$(DHB\text{-}D_3\text{-}Ser)_2$	452	nd	15991.0-9
$(DHB\text{-}Ser)_3$	669	nd	nd

*Values represent increase in theoretical mass above apo-peptide. nd, not detected.

Figure 9



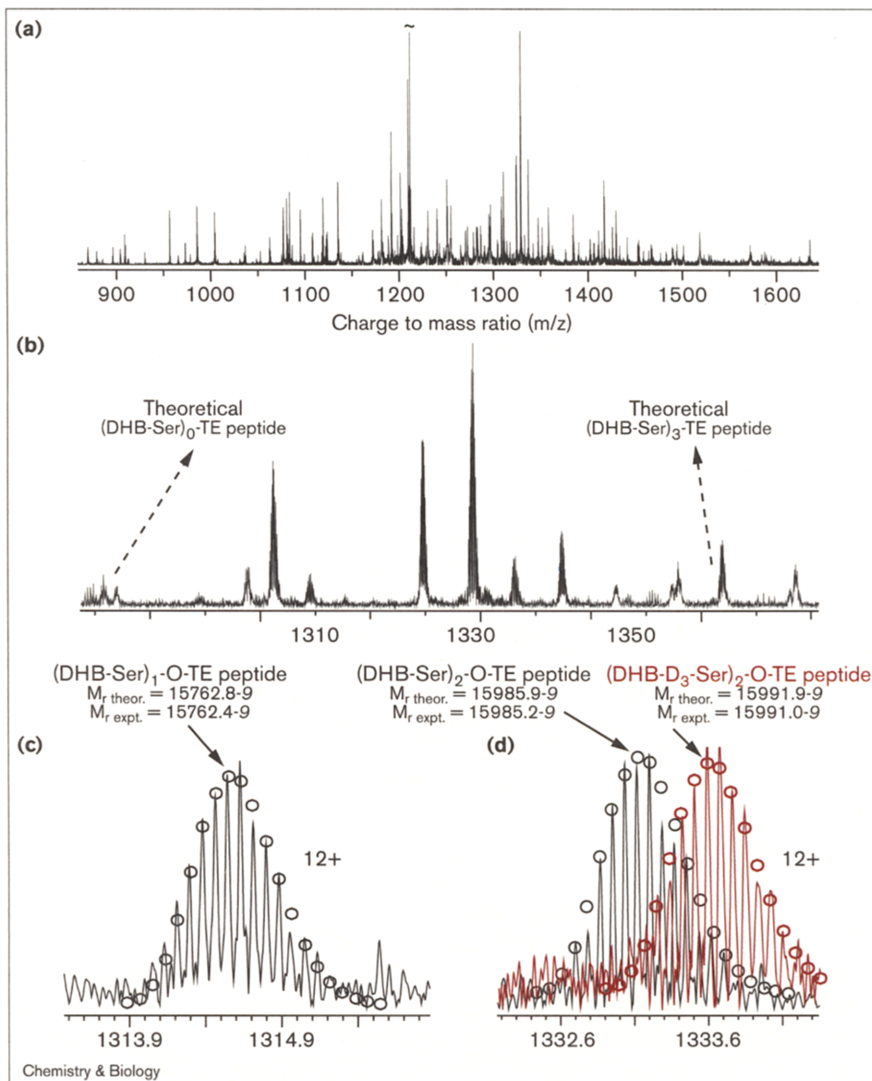
ESI/FTMS analysis of full-length EntF. **(a)** Partial ESI/FT mass spectrum (32 scans) of an unfractionated Lys-C digest of full length wild-type apo-EntF (142 kDa). **(b,c)** Expansions of the isotopic distributions for the PCP peptide (Ala974–Lys1011) and the TE peptide (Leu1038–Lys1177), respectively; open circles and squares, theoretical isotopic distributions for the TE peptide with and without one disulfide bond, respectively. **(d)** Isotopic distribution for the aminoacylated holo-Ser-PCP peptide of EntF-H1271A. This acyl-S-PCP peptide (3+ ionization state) has a mass shift +427 higher than the apo-PCP peptide.

species is 427.1 Da larger than the apo-PCP peptide and is consistent with the covalent attachment of serine to the Ppant moiety on the Ser1006 peptide to form Ser-S-PCP (theoretical M_r = 4240.96-*2*).

Under full turnover conditions and after Lys-C proteolysis, the ESI/FTMS spectrum for the H1271A mutant yielded 130 isotopic distributions from over 65 peptides of 2–41 kDa (Figure 10a), and two of these isotopic distributions were unique. Although the isotopic distribution for the TE peptide (12^+) was undetectable (Figure 10b, dashed arrow at left), two new 12^+ signals were visible, one at 15762.4-*9* and the second at 15985.2-*9* (Figure 10c,d and Table 2). These species are 222.7 Da and 445.6 Da higher in molecular weight than the unacylated TE peptide, and these values correspond to the molecular weight increments for monomeric DHB-Ser (+223.0 Da theoretically) and dimeric DHB-Ser-O-DHB-Ser (+446.0 Da theoretically) acyl groups tethered to the TE active-site serine residue.

Repeating the above experiment using α,β,β-trideuteroserine (+3 Da) in the H1271A mutant incubations yielded a similar mass spectrum (not shown) as that in Figure 10a. Again, no isotopic abundance was observed for the TE peptide either in unacylated form or with one or three DHB-Ser moieties bound. The isotopic distribution for the TE peptide with two bound DHB-Ser moieties was observed as 15991.0-*9* (12^+), however (Figure 10d and Table 2). This M_r value is 5.8 higher than that found in incubations with unlabeled serine as the H1271A mutant

Figure 10



Identification of enterobactin acyl enzyme intermediates on the EntF H1271A TE domain. **(a)** Full ESI/FT mass spectrum of an unfractionated Lys-C digest of the H1271A mutant after quenching enterobactin biosynthesis. **(b)** Expansion of the *m/z* range 1290–1360; arrows show theoretical positions of four possible forms of the TE peptide (12+ ions); dotted lines, not observed; solid lines, observed. **(c)** Expanded isotopic distributions for TE-bound intermediates with one DHB–Ser moiety covalently attached; open circles, theoretical isotopic distributions; arrow, most abundant isotopic peaks; 16 scans. **(d)** Comparison of expanded isotopic distributions observed for the dimeric $(DHB\text{-}Ser)_2$-TE intermediate formed using either unlabeled (H_2)-Ser (black) or deuterated (D_3)-Ser (red) as a substrate.

substrate, and it is consistent with the 6.0 Da increase expected for two D_3-DHB-Ser moieties bound to the TE domain (the D_6-version of **4** shown in Figure 3) as the dimeric acyl-O-Ser1138 peptide.

Discussion

The great majority of NRPS and PKS gene clusters sequenced to date reveal that these complexes are multimodular catalysts with specialized initiation, elongation and termination domains [2,3]. Except on rare occasions when mature acyl chains are liberated from the most downstream carrier proteins by reductive cleavage (e.g. fungal Lys2 catalysis [4]), or by diketopiperazine formation and release (e.g. lysergic acid assembly [35]), the 25–30 kDa TE domains are the most carboxy-terminal domains in the NRPS and PKS enzymes, consistent with a controlled, catalytic release function of the full-length acyl enzyme intermediate. TE domains are found in enzymes that act as hydrolases as well as cyclases, reflecting the ability to selectively transfer the mature natural-product acyl chain to either water as an intermolecular nucleophile or to an intramolecular nucleophilic sidechain within the acyl chain moiety. Intermolecular acyl chain transfer to water is the exclusive fate in FAS catalysis [1] and also occurs during the synthesis of the vancomycin family of glycopeptide antibiotics [36]. In contrast, the TE domain at the end of DEBS III carries out exclusive cyclization to the 14-member lactone 6-deoxyerythronolide B [7] and the carboxy-terminal TE domains of surfactin synthetase [6,20] or bacitracin synthetase [22] presumably catalyze cyclization reactions to make the lactone surfactin or the lactam bacitracin, respectively.

The structure of LuxD, the bacterial TE subunit responsible for myristoyl-S-ACP hydrolysis, has validated

homology predictions that TE domains are members of a serine hydrolase superfamily with a proposed Ser–His–Asp catalytic triad in which the serine sidechain is the catalytic nucleophile and the histidine serves as a general base. From the work reported here, Asp1165 of EntF is largely dispensable, as is the corresponding Asp236 in FAS TE II [11], so the active site might be more accurately represented by a catalytic dyad of Ser–His. The general mechanism for any TE domain, therefore, would be conversion of the acyl thioester linkage of the substrate (here the carrier protein domains acyl-S-ACP or peptidyl-S-PCP) to the oxoester acyl enzyme intermediate, acyl-O-TE, as a direct precursor for hydrolysis or cyclization steps (Figure 1). When the TE subunits associated with either the bacterial luciferase operon (LuxD) or a vertebrate FAS TE II, were analyzed after mutation of the active-site serine to cysteine, distinct kinetic outcomes were observed: the FAS TE II had the same turnover rate as wild-type TE II [11,12], indicating conservation of the rate-determining step, whereas turnover of the LuxD serine to cysteine mutant dropped 35–40-fold, suggesting that the ratio of enzyme acylation/deacylation rate constants had been altered [14]. Indeed the LuxD serine to cysteine mutant accumulated stoichiometric levels of acyl enzyme intermediate, which indicates selective deceleration of acyl enzyme hydrolytic deacylation [14].

When the comparable EntF mutant (S1138C) was analyzed, there was a 1000-fold drop in turnover, coincident with a drastic change in the rate-determining step. Along with the dramatic kinetic effect there was a change in the product composition such that 25–33% of the products was the linear ester dimer DHB-Ser-O-DHB-Ser, which has been postulated to be an intermediate (Figure 3, **4**) in an acyl enzyme cascade heading towards formation of a trimeric linear DHB-Ser-O-DHB-Ser-O-DHB-Ser-enzyme intermediate (Figure 3, **5**). The dimeric acyl-O-TE (**4**) would arise from **3** through intramolecular capture by the attack of the serine sidechain of the DHB-Ser moiety tethered at the TE domain and translocation of the upstream acyl group to form the ester bond (Figure 3). No release of dimer occurs during turnover of wild-type EntF at pH 7.5, although we have noted a very small amount of this dimer released in turnover at pH 8.8 [23]. The coupled kinetic and product effects in the S1138C mutant can be explained if elongation of the acyl dimeric $(DHB\text{-}Ser)_2$-S-enzyme (Figure 3, **4**) to trimeric $(DHB\text{-}Ser)_3$-S-enzyme (Figure 3, **5**) is selectively decelerated such that very slow forward progression makes the dimer susceptible to release by water capture (k_{hydrol} ~0.03 min^{-1}), a path that would be kinetically inconsequential in wild-type EntF (which has a k_{cat} ~100 min^{-1} for cyclotrimerization). The experiments with [^{3}H]serine acylation (Figure 6) demonstrate that even between the PCP and TE sites the total accumulation of acyl enzyme species (e.g. Figure 3, **1–5**) is still less than stoichiometric.

The EntF H1271A mutant does not release the DHB-Ser dimer, suggesting that it might be stuck at a different point in the acyl enzyme intermediate cascade. Perhaps His1271 is required for the slow water activation and release of the dimer. Furthermore, the k_{cat} of this mutant is decreased by another order of magnitude, at an anemic ~0.01 min^{-1} during enterobactin biosynthesis, and even when the histidine function is replaced by 100 mM imidazole in solution, the rate of enterobactin release from the H1271A mutant is improved only tenfold to 0.1 min^{-1}. These results corroborate the anticipated role of His1271 as a general base in transfer of the substrate acyl groups that are covalently translocated, although they do not pinpoint whether the rate decrease in the H1271A mutant results from defective interdomain acyl transfer (from DHB-Ser-S-PCP to DHB-Ser-O-TE, Figure 3, **1** to **2**), or from hydrolysis of any of the putative monomeric (**3**), dimeric (**4**) or trimeric (**5**) forms of $(DHB\text{-}Ser)_n$-O-TE.

We took advantage of the very slow turnover of the H1271A mutant to try to identify and localize these acyl enzyme intermediates on the TE domain. Recall that the [^{3}H]serine enzyme acylations did yield a DHB-dependent fractional increase in covalently attached serine species (Figure 6), but the amount was not suprastoichiometric. Because these experiments did not provide compelling evidence for acyl enzyme species detectable at both the PCP and TE peptides, or in dimer or trimer forms, we turned to mass spectrometric analysis.

The increasing usage of biological MS in the past decade has been driven by adaptation of existing mass analyzers (e.g., time-of-flight, quadrupole, and so on) to measure ions produced by 'soft' ionization methods such as ESI. Of the many ionization method/mass analyzer pairings, ESI with FTMS is useful because its high resolving power translates into resolution of peptide isotopic peaks. Such isotopic resolution is critical for measuring complex mixtures by ESI/FTMS without requiring chromatographic fractionation (Figures 7,9) and enables the use of stable isotopes for enhanced structural confidence even at 15 kDa (Figure 10). Direct analysis of unfractionated peptide mixtures using other instrumentation is possible, but requires digestion to much smaller peptides (<3–4 kDa) for high resolving power and mass accuracy. For example, a protease like trypsin would be required for degradation, and substantially more peptides (95 from exhaustive digestion of EntF) would be generated. The use of a more specific protease, such as Lys-C, yields simpler mixtures of larger peptides. The ability of ESI/FTMS to analyze such mixtures with high resolution but without chromatography makes high-throughput measurements commensurate with kinetic studies far more tractable.

Because EntF is a 142 kDa protein and could contain mixtures of at least six covalent acyl enzyme intermediates at

two distinct domains (Ser-S-PCP shown in Figure 2 and **1–5** in Figure 3), we opted for MS analysis of an unfractionated mixture of peptide fragments from the proteolytic digestion of EntF at acidic to neutral pH, in which the acyl oxoester and thioester enzyme linkages should survive. After exploration with the two domain EntF-PCP–TE fragment as a model it was possible to detect signature peptides containing either the Ser1006 of the PCP domain or the Ser1138 of the TE domain at the molecular weights predicted with cleavage by the Lys-C protease (Figure 7c). The apo and holo (Ppant) forms of the PCP peptide were readily distinguished and detected (Figure 8) as a further prelude to the full-length EntF analysis. In limited incubations of the H1271A mutant without DHB, the Ser-S-PCP peptide was clearly evident (Figure 9d), and during full turnover conditions MS results provided unambiguous evidence for acyl-O-TE intermediates (Figure 10). Most intriguingly, the incremental increase of 223 Da for the monomer DHB-Ser-O-TE species (Figure 3, **2**) was accompanied by a 445 Da increase exactly as expected for the dimeric $(DHB\text{-}Ser)_2$-O-TE acyl enzyme (Figure 3, **4**). Further validation was provided when deuterated serine was used as a substrate. The H1271A dimeric $(DHB\text{-}D_3\text{-}Ser)_2$-O-TE peptide had an additional six mass units, exactly as predicted for two D_3-serine residues in that acyl enzyme species (Figure 10).

These results demonstrate the power of MS to detect covalently derivatized enzymes. In this case the labile acyl-S-PCP (Ser1006) domain linkage was more difficult to detect than the more stable oxoester in the acyl-O-TE (Ser1138) domain species, and the bis monomeric DHB-Ser-enzyme (Figure 3, **3**) has not yet been detected using MS. Aggregate MS results validate the proposal of Figure 3, which demonstrates that the TE domain of EntF is not only the catalyst for product release but is also involved in acyl chain growth and elongation. EntF is an intriguing NRPS variant in that it catalyzes formation of both amide and ester bonds (three of each) between its substrates during monomer activation and chain growth. The amide-bond-forming step, joining DHB to the amino group of serine, occurs at the front end of EntF, using the separate protein DHB-S-EntB as the acyl donor and the aminoacylated PCP of EntF, Ser-S-PCP, as the nucleophilic acceptor to yield **1** as the immediate product from action of the EntF condensation domain. Until now, the thrice iterated ester-bond-forming step (in the serine trilactone scaffolding of enterobactin) had not been demonstrated experimentally. Given that EntF has only a single PCP domain that acts as a monomer [23], the most likely mechanism that could encompass three iterative steps is an alternative, novel use of the TE domain at carboxy-terminal end of EntF.

Transfer of the monomeric DHB-Ser acyl chain from the thiol of the PCP domain (**1**) to the sidechain CH_2OH of Ser1138 in the TE domain (**2**) was expected and validated directly here using the H1271A mutant, but the features that control the timing of acyl transfer from the PCP to the TE domain are still not understood. One presumes that in most PKS and NRPS enzymes, the rate of acyl chain release from the covalent acyl-O-TE domain would be variable. For those TE domains in which cyclization is the exclusive mechanism for chain release hydrolysis must be suppressed, and the acyl-O-TE variant must be stable enough to wait for capture by a hydroxyl or amide sidechain internal to the natural product acyl chain (e.g. the 13-OH in DEBS). In EntF the mono DHB–Ser-O-TE (**2**) must be relatively stable (oxoesters are generally less thermodynamically activated than thioesters) and serve as a waiting room while the free phosphopantetheinylated PCP domain is reacylated by seryl-AMP from the adenylation domain and *N*-capped by DHB-S-EntB through the condensation domain to produce the doubly monoacylated version of EntF (**3**), with one DHB–Ser chain in thioester linkage on PCP and the other DHB–Ser chain in oxoester linkage on the TE domain. At this point, catalysis of ester-bond formation could occur by attack of the sidechain of the seryl moiety on the DHB-Ser-O-TE onto the carbonyl of the thioester linkage of DHB-Ser-S-PCP to effect ester synthesis, chain translocation and generation of the dimeric DHB-Ser-O-DHB-Ser-O-TE (**4**). This dimer (**4**) is one of the acyl enzyme species that accumulates in assays using the H1271A mutant. Dimeric ester formation would require catalytic generation of the alkoxide form of the attacking nucleophilic serine in the DHB-Ser-O-TE sidechain, which might be one of the crucial roles of the TE active site catalytic machinery (still at least slowly achievable in H1271A). Hypothetically, the conversion of **3** to **4** could occur by retrograde chain transfer in which the nucleophilic sidechain of DHB-Ser during ester formation is on the PCP domain and not the TE domain, therefore yielding a transient $(DHB\text{-}Ser)_2$-S-PCP species that would then be transferred in the forward direction to the TE domain to produce **4**. The lack of evidence to support bidirectional chain transfer between the PCP and TE domains during elongation encourages us to present the simplest mechanism for elongation as depicted in Figure 3.

The second ester linkage (**5**) would result from iteration of the waiting room proposal. The phosphopantetheinylated PCP domain would be filled for a third time with a DHB-Ser acyl chain (not shown in Figure 3) while the $(DHB\text{-}Ser)_2$-O-TE is sequestered from hydrolysis. The serine sidechain alkoxide at the most distal of the DHB-Ser units tethered on the TE domain must again be generated to enable DHB-Ser translocation from the PCP domain to produce the last acyl enzyme intermediate (Figure 3, **5**). At this juncture before the phosphopantetheinylated PCP domain can fill for a fourth time and another DHB-Ser acyl group is translocated, generation of an alkoxide on the

CH_2OH sidechain of the most distal DHB-Ser of the trimer must occur in the TE domain, and this alkoxide must be in a conformation competent for intramolecular capture, cyclotrimerizing the 12-ring lactone formation, and facilitating enterobactin release. Among the steps that are selectively slowed in the H1271A mutant (which has a turnover 10,000-fold lower than wild-type EntF) must be the conversion of **4** to **5** because **4** accumulates in the H1271A mutant but not in the S1138C mutant.

It remains to be seen how portable the TE domain of EntF will be with this iterative esterification and trilactone cyclization activity for *N*-acylated serines. The TE domain of DEBS III is readily placed behind upstream ACP fragments in DEBS I or DEBS II, and this relocation effects lactonizing release of immature acyl chains [7,37–39]. In contrast, the ACP–TE or TE domain alone when excised from DEBS has hydrolase but not cyclase activity [16,17], so the upstream context of PKS modules might be important for acyl chain sequestration by a TE domain from unwanted hydrolysis. We also do not yet know to what extent TE domains are interchangeable (i.e. an NRPS TE domain swapped for a PKS TE domain or vice versa) because the basis of acyl group recognition and kinetic fates for acyl transfer have not yet been deciphered. Selective control of the routing and release of acyl chains in the termination phase of polyketide and nonribosomal peptide biosynthesis would be very desirable in combinatorial biosynthesis of novel natural products.

Significance

Enterobactin is a unique nonribosomally synthesized peptide that is assembled with three 2,3-dihydroxybenzoate (DHB) and L-serine residues linked by amide and ester bonds. Although it has been presumed that the EntF thioesterase (TE) domain participates with the peptidyl carrier protein (PCP) domain in three rounds of esterification to elongate the DHB-serine moieties during enterobactin biosynthesis, this study provides the first direct evidence of acyl enzyme intermediates covalently bound to the active-site serine residues in the PCP and TE domains.

When EntF TE active-site residues were mutagenized and probed for activity, the rates of enterobactin release dropped three and four orders of magnitude in the S1138C and H1271A mutants, respectively, compared with the wild-type k_{cat} of 100 min^{-1}. In addition, the S1138C mutant released the dimeric DHB-Ser-O-DHB-Ser, which indicates that the thioester is more susceptible to hydrolysis than the oxoester linkage in acyl-TE intermediates. Mutations in the conserved Asp1165 did not appreciably affect release rates, so the nucleophilic serine and histidine base in the EntF TE domain might represent a catalytic dyad.

Mass spectrometry (ESI/FTMS) was used to identify two acyl enzyme intermediates of the rate-impaired H1271A mutant after proteolysis: the monomeric DHB-Ser-O-TE and the dimeric $(DHB\text{-}Ser)_2$-O-TE. The detection of these covalent intermediates lends credence to the proposed mechanism that the TE domain serves as both a waiting room for acyl transfer from the PCP domain and an elongation catalyst for enterobactin cyclotrimerization. Increased understanding of TE mechanisms will facilitate future experiments designed to translocate TE domains for catalytic release of truncated peptides (precedented in the 6-deoxyerythronolide B polyketide synthase) and to exchange TE domains to investigate hydrolysis and cyclization fates in nonribosomally synthesized peptides.

Materials and methods

Construction of the EntF mutants and the EntF-PCP–TE two-domain fragment

The EntF site-directed mutants S1138C, H1271A, D1165A and D1165S were constructed using the splicing by overlap extension (SOE) method [40]. For each mutation, plasmid pMS22 [41] served as a template for the first round of PCR amplification using the two sets of primers listed in Table S1 (see the Supplementary material section). The two PCR-amplified DNA fragment products from round one were gel purified (Qiaquick Kit, Qiagen, Inc.) and used as the templates during a second round of amplification to yield the final DNA product containing the desired mutation. All final PCR products were gel purified and digested with the restriction enzymes listed in Table S1. The vector pMS22 was likewise digested with the corresponding enzymes from Table S1, with the exception of the construction of pMS22–S1138C where pMS22 was digested with *KpnI, SalI* and *StyI* as described previously [23]. Ligations were transformed into *E. coli* DH5α cells to propagate the plasmids.

EntF DNA encoding the PCP/TE peptide fragment (corresponding to EntF residues 960–1293) was amplified from the pMS22 template with one round of PCR using the primers shown in Table S1. The 1 kb PCR product and the vector pET22b (Novagen, Inc.) were digested with *NdeI* and *XhoI* and ligated. The resulting plasmid, pET22b–EntF960–1293, was designed to incorporate a carboxy-terminal hexahistidine tag to facilitate EntF-PCP–TE purification. Ligation mixtures were transformed into *E. coli* DH5α cells. All nucleotides in PCR-amplified fragments were confirmed by DNA sequencing (Dana Farber Molecular Biology Core Facility, Boston, MA).

Purification of enterobactin synthetase proteins

EntD [4], EntB [24], EntE [23], EntF [23] and the S1138A mutant were overproduced and purified as described previously [23]. For the overproduction and purification of the S1138C, H1271A, D1165A and D1165S mutant proteins, each overexpression plasmid was transformed into *E. coli* BL21(DE3), and the resulting strain was grown at 30°C to an optical density of 0.7 (2 l LB medium, 100 μg/ml ampicillin) and induced with 1 mM isopropyl β-D-thiogalactopyranoside (IPTG). Cells were harvested 4 h after induction by centrifugation, resuspended in buffer A (25 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 5 mM dithiothreitol) with 5% glycerol, and lysed by two passages through a French Pressure cell. Purification of the S1138C mutant was carried out by gel filtration and ion-exchange chromatography as described for the S1138A mutant [23]. The H1271A, D1165A and D1165S mutants were purified as described for wild-type EntF [23], but the 20–50% ammonium sulfate pellet was chosen for subsequent ion-exchange chromatography. Purity of the EntF mutant proteins was assessed by SDS–PAGE, and positive chromatography fractions were pooled, dialyzed against Buffer A with 10% glycerol and stored at –80°C.

The *E. coli* BL21(DE3)/pET22b–EntF960–1293 cultures (2 l LB medium, 100 μg/ml ampicillin) were likewise grown at 30°C to an

optical density of 0.4, induced with 1 mM IPTG, harvested and lysed in a French Pressure cell with 20 mM Tris–HCl (pH 7.9), 200 mM NaCl and 5 mM imidazole. Following lysis, the EntF-PCP–TE protein was purified by nickel-chelate chromatography according to the manufacturer's instructions (Novagen, Inc.). The purity of EntF-PCP–TE was assessed by SDS–PAGE, and positive chromatography fractions were pooled, dialyzed against 50 mM Tris-HCl (pH 8.0), 2 mM DTT and 5% glycerol and stored at −80°C.

Protein concentrations were determined using the calculated extinction coefficients for the absorbance of each protein at 280 nm [42]: 180,360 M^{-1} cm^{-1} for EntF mutants and 42,300 M^{-1} cm^{-1} for EntF-PCP–TE.

ATP–[^{32}P]PPi exchange assay
ATP–pyrophosphate exchange was assayed as described previously [23] using 3 mM L-serine with 10–12 nM apo-EntF wild-type and mutant proteins.

Radioassay for the phosphopantetheinylation of EntF
To quantify the ability of the EntF PCP domain to become phosphopantetheinylated by the PPTase EntD, [^{3}H]CoA was prepared as the disulfide [4] and incubated in a 100 µl reaction mixture containing 75 mM Tris–HCl (pH 7.5), 1 mM Tris-(2-carboxyethyl)phosphine (TCEP), 10 mM $MgCl_2$, 50 µM [^{3}H]CoA disulfide (74 Ci/mol), 222 nM EntD, and 380–440 nM EntF (wild-type, S1138C, H1271A, D1165A or D1165S). The phosphopantetheinylation of EntF-PCP–TE (5–20 µM) was assayed using 25 mM DTT instead of TCEP and slightly less EntD (198 nM). Reactions were incubated at 37°C, and over a 40 min time course samples were quenched at regular intervals with 800 µl of 10% trichloroacetate (TCA) and 2% bovine serum albumin. Precipitated protein was pelleted by centrifugation, washed 3 × with 800 µl TCA, dissolved in 150 µl formic acid and added to 3.5 ml scintillation fluid. The [^{3}H]Ppant covalently bound to the PCP domain of each protein was quantified by liquid scintillation counting.

Assay for covalent [^{3}H]serine incorporation into EntF PCP
Serine incorporation into holo carrier proteins was quantified with a TCA precipitation radioassay [23]. Reaction mixtures (100 µl final volume) included 75 mM Tris–HCl (pH 7.5), 1 mM DTT, 10 mM $MgCl_2$, 10 mM ATP, 500 µM CoASH, 222 nM EntD, 905 µM [^{3}H]L-serine (111 Ci/mol) (Dupont NEN) and EntF wild-type (1 µM), S1138A (2.3 µM), S1138C (2.4–4.8 µM) or H1271A (2.4 µM). Apo proteins were preincubated for 45 min at 37°C with EntD to initiate phosphopantetheinylation prior to addition of the radiolabeled serine and ATP. TCA precipitation and liquid scintiallation counting was performed as described above for the Ppant assay.

Ethyl acetate extraction radioassay and HPLC assays to identify and quantify enterobactin intermediates
Enterobactin, as well as (DHB-Ser)$_n$ monomer, dimer and uncyclized trimer intermediates were detected using the previously described radioassay and HPLC methods [23]. Reaction mixtures (100 µl final volume) were preincubated for 45 min at 37°C with 75 mM Tris-HCl (pH 7.5), 1 mM DTT, 10 mM $MgCl_2$, 500 µM CoASH and 222 nM EntD for phosphopantetheinylation of apo-EntB (11–18 µM) and apo-EntF (10 nM wild-type and 50–2700 nM mutants). Catalysis of enterobactin formation was then initiated with the addition of 10 mM ATP, 500 µM DHB, 905 µM [^{3}H]L-serine (111 Ci/mol; Dupont NEN) and 300 nM EntE. After a specified time, the reactions were quenched with 150 µl of 1 N HCl and extracted with 0.75 ml of ethyl acetate. In the radioassay, 0.5 ml of the ethyl acetate layer was removed and added to 3.5 ml of scintiallation fluid. The aggregate amount of (DHB-[^{3}H]Ser)$_n$ species released from the EntF protein was determined using liquid scintillation counting.

To identify and quantify the relative abundance of enterobactin intermediates, an HPLC assay was employed. The reaction mixtures were prepared as described above for the radioassay, except for the substitution of unlabeled 1 mM L-serine in place of radiolabeled serine. Quenching with HCl and ethyl acetate extraction was followed by concentration of the 0.5 ml layer containing (DHB-Ser)$_n$ intermediates by a rotary evaporator. The residue was resuspended in 250 µl of 30% acetonitrile/water and 50 µl samples were analyzed along with synthesized standards of enterobactin, monomer, dimer, and uncyclized trimer intermediates according to the HPLC methods described previously [23].

To further confirm the identity of the (DHB-Ser)$_2$ dimer intermediate product the reaction mixture was scaled up to 3 ml, quenched with 4.5 ml HCl, extracted with 7.5 ml ethyl acetate and dried with a rotary evaporator. The residue was resuspended in 300 ml 30% acetonitrile and concentrated under reduced pressure. The resulting sample was submitted for mass spectrometry (HPLC-MS) at the Mass Spectrometry Facility of the Department of Chemistry and Chemical Biology, Harvard University (Cambridge, MA). The chromatography was performed on a Hewlett Packard instrument (Model 1090) and mass spectrometry was carried out in negative ion mode using ESI–MS and atmospheric pressure chemical ionization (APCI–MS) (Micromass Platform II, Beverly, MA).

Electrospray ionization/fourier transform mass spectrometry (ESI/FTMS)
In preparation for ESI/FTMS, protein samples were first incubated in a reaction mixture (100 µl) containing 75 mM Tris–HCl (pH 7.5), 1 mM TCEP, 10 mM $MgCl_2$, 10 mM ATP, 500 µM CoASH, 222 nM EntD, 18.2 µM EntB, 300 nM EntE, 500 µM DHB, 3 mM L-serine (or deuterated L-serine-2,3,3-D_3, 98%, Cambridge Isotope Laboratories, Inc.) and EntF wild-type, mutant or EntF-PCP–TE proteins. The apo-, holo- or acyl-holo-proteins (88–370 µg) were then digested with endoproteinase Lys-C (Promega Corp.) at 37°C for 1 h. Reaction mixtures (175 µl) contained 25 mM sodium phosphate (pH 7.5), 1 mM EDTA, 2 M urea and a ratio of protein:Lys-C varying from 15:1 to 75:1. Proteolysis was terminated by lowering the pH to 5.5 with 4.3 N acetic acid.

Crude Lys-C digests were desalted using reversed-phase peptide traps (Michrom Bioresources, Auburn, CA) and eluted into 100 µl of acetonitrile, H_2O and acetic acid (78:20:2). A portion of this solution (~20 µl) was infused at 200 nl/min into a 50 µm inner diameter fused silica capillary terminated with an ESI tip (NewObjective, Cambridge, MA) grounded relative to the electrospray capillary which is held at −1.3 kV. The resulting ions were guided through the ESI source and four stages of differential pumping into the Infinity Cell ($<10^{-9}$ Torr) of a 4.7 Tesla APEX II Fourier transform MS (Bruker Daltonics, Billerica, MA) [43]. Theoretical isotopic distributions were generated using XMASS software or Isopro v3.0 [44] and fit to experimental data by least squares to assign the most abundant isotopic peak. Spectra were calibrated externally using bovine ubiquitin (8564.64-5 Da).

Supplementary material
Details of vector constructions are published with the online version of this paper.

Acknowledgements
This work was supported in part by National Institutes of Health (NIH) grant GM20011 (C.T.W.). C.A.S.R. is a National Science Foundation Minority Postdoctoral Research Fellow (DBI-9707847) and N.K. is supported by an NIH Postdoctoral Fellowship (F32 AI 10087-02). A.M.G. was a Howard Hughes Medical Institute Predoctoral Fellow and H.C.L. is a National Science Foundation Graduate Fellow.

References
1. Smith, S. (1994). The animal fatty acid synthase: one gene, one polypeptide, seven enzymes. *FASEB J.* **8**, 1248-1259.
2. Khosla, C. (1997). Harnessing the biosynthetic potential of modular polyketide synthases. *Chem. Rev.* **97**, 2577-2590.
3. Marahiel, M.A., Stachelhaus, T. & Mootz, H.D. (1997). Modular peptide synthetases involved in nonribosomal peptide synthesis. *Chem. Rev.* **97**, 2651-2673.
4. Lambalot, R.H., *et al.*, & Walsh, C.T. (1996). A new superfamily – the phosphopantetheinyl transferases. *Chem. Biol.* **3**, 923-936.
5. Stachelhaus, T., Huser, A. & Marahiel, M.A. (1996). Biochemical characterization of peptidyl carrier protein (PCP), the thiolation domain of multifunctional peptide synthetases. *Chem. Biol.* **3**, 913-921.

6. Schneider, A., & Marahiel, M.A. (1998). Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in non-ribosomal peptide biosynthesis in *Bacillus subtilis*. *Arch. Microbiol.* **169**, 404-410.
7. Cortes, J., Wiesmann, K.E.H., Roberts, G.A., Brown, M.J.B., Staunton, J. & Leadlay, P.F. (1995). Repositioning of a domain in a modular polyketide synthase to promote specific chain cleavage. *Science* **268**, 1487-1489.
8. Pazirandeh, M., Chirala, S.S. & Wakil, S.J. (1991). Site-directed mutagenesis studies on the recombinant thioesterase domain of chicken fatty acid synthase expressed in *Escherichia coli*. *J. Biol. Chem.* **266**, 20946-20952.
9. Witkowski, A., Naggert, J., Wessa, B. & Smith, S. (1991). A catalytic role for histidine 237 in rat mammary gland thioesterase II. *J. Biol. Chem.* **266**, 18514-18519.
10. Witkowski, A., Naggert, J., Witkowska, H.E., Randhawa, Z.I. & Smith, S. (1992). Utilization of an active serine 101 to cysteine mutant to demonstrate the proximity of the catalytic serine 101 and histidine 237 residues in thioesterase II. *J. Biol. Chem.* **267**, 18488-18492.
11. Tai, M.-H., Chirala, S.S. & Wakil, S.J. (1993). Roles of Ser101, Asp236, and His237 in catalysis of thioesterase II and of the C-terminal region of the enzyme in its interaction with fatty acid synthase. *Proc. Natl Acad. Sci. USA* **90**, 1852-1856.
12. Witkowski, A., Witkowska, H.E. & Smith, S. (1994). Reengineering the specificity of a serine active-site enzyme: two active-site mutations convert a hydrolase to a transferase. *J. Biol. Chem.* **269**, 379-383.
13. Lawson, D.M., *et al.*, & Derewenda, Z.S. (1994). Structure of a myristoyl-ACP-specific thioesterase from *Vibrio harveyi*. *Biochemistry* **33**, 9382-9388.
14. Li, J., Szittner, R., Derewenda, Z.S. & Meighen, E.A. (1996). Conversion of ser:114 to cyst:114 and the role of the active site nucleophile in acyl transfer by myristoyl-ACP thioesterase from *Vibrio harveyi*. *Biochemistry* **35**, 9967-9973.
15. Caffrey, P., Green, B., Packman, L.C., Rawlings, B.J., Staunton, J. & Leadlay, P.F. (1991). An acyl-carrier-protein-thioesterase domain from the 6-deoxyerythronolide B synthase of *Saccharopolyspora erythraea*. High-level production, purification and characterisation in *Escherichia coli*. *Eur. J. Biochem.* **195**, 823-830.
16. Aggarwal, R., Caffrey, P., Leadlay, P.F., Smith, C.J. & Staunton, J. (1995). The thioesterase of the erythromycin-producing polyketide synthase: Mechanistic studies *in vitro* to investigate its mode of action and substrate specificity. *J. Chem. Soc. Chem. Commun.*, 1519-1520.
17. Weissman, K.J., *et al.*, & Leadlay, P.F. (1998). The thioesterase of the erythromycin-producing polyketide synthase: Influence of acyl chain structure on the mode of release of substrate analogues from the acyl enzyme intermediates. *Angew. Chem. Int. Ed.* **37**, 1437-1440.
18. de Crecy-Lagard, V., *et al.*, & Blanc, V. (1997). Streptogramin B biosynthesis in *Streptomyces pristinaespiralis* and *Streptomyces virginiae*: Molecular characterization of the last structural peptide synthetase gene. *Antimicrob. Agents Chemother.* **41**, 1904-1909.
19. Paetzel, M. & Dalbey, R.E. (1997). Catalytic hydroxyl/amine dyads within serine proteases. *Trends Biochem. Sci.* **22**, 28-31.
20. Guenzi, E., Galli, G., Grgurina, I., Pace, E., Ferranti, P. & Grandi, G. (1998). Coordinate transcription and physical linkage of domains in surfactin synthetase are not essential for proper assembly and activity of the multienzyme complex. *J. Biol. Chem.* **273**, 14403-14410.
21. Feignier, C., Besson, F. & Michel, G. (1996). Characterization of iturin synthetase in the wild-type *Bacillus subtilis* strain producing iturin and in an inturin deficient mutant. *FEMS Microbiol. Lett.* **136**, 117-122.
22. Konz, D., Klens, A., Schorgendorfer, K. & Marahiel, M.A. (1997). The bacitracin biosynthesis operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three multi-modular peptide synthetases. *Chem. Biol.* **4**, 927-937.
23. Gehring, A.M., Mori, I. & Walsh, C.T. (1998). Reconstitution and characterization of the *Escherichia coli* enterobactin synthetase from EntB, EntE, and EntF. *Biochemistry* **37**, 2648-2659.
24. Gehring, A.M., Bradley, K.A. & Walsh, C.T. (1997). Enterobactin biosynthesis in *Escherichia coli*: isochorismate lyase (EntB) is a bifunctional enzyme that is phosphopantetheinylated by EntD and then acylated by EntE using ATP and 2,3-dihydroxybenzoate. *Biochemistry* **36**, 8495-8503.
25. Rusnak, F., Liu, J., Quinn, N., Berchtold, G.A. & Walsh, C.T. (1990). Subcloning of the enterobactin biosynthetic gene *entB*: expression, purification, characterization, and substrate specificity of isochorismatase. *Biochemistry* **29**, 1425-1435.
26. Rusnak, F., Faraci, W.S. & Walsh, C.T. (1989). 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.
27. Fenn, J.B., Mann, M., Meng, C. K. & Wong, S. F. & Whitehouse, C.M. (1990). Electrospray ionization-principles and practice. *Mass Spectrom. Rev.* **9**, 37-70.
28. McLafferty, F.W. (1994). High-resolution tandem FT mass spectrometry above 10 kDa. *Accounts Chem. Res.* **27**, 379-386.
29. Cole, S.T., *et. al.*, & Barrell, B.G. (1998). Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537-544.
30. Reimmann, C., Serino, L., Beyeler, M. & Haas, D. (1998). Dihydroaeruginoic acid synthetase and pyochelin synthetase, products of the *pchEF* genes, are induced by extracellular pyochelin in *Pseudomonas aeruginosa*. *Microbiology* **144**, 3135-3148.
31. Cunliffe, H.E., Merriman, T.R. & Lamont, I.L. (1995). Cloning and characterization of *pvdS*, a gene required for pyoverdine synthesis in *Pseudomonas aeruginosa*: PvdS is probably an alternative sigma factor. *J. Bacteriol.* **177**, 2744-2750.
32. Oudega, B., Vandenbol, M. & Koningstein G. (1997). A 12 kb nucleotide sequence containing the alanine dehydrogenase gene at 279 degrees on the *Bacillus subtilis* chromosome. *Microbiology* **143**, 1489-1491.
33. Cosmina, P., *et al.*, & van Sinderen, D. (1993). Sequence and analysis of the genetic locus responsible for surfactin synthesis in *Bacillus subtilis*. *Mol. Microbiol.* **8**, 821-831.
34. Gocht, M. & Marahiel, M.A. (1994). Analysis of core sequences in the D-phe activating domain of the multifunctional peptide synthetase TycA by site-directed mutagenesis. *J. Bacteriol.* **176**, 2654-2662.
35. Walzel, B., Riederer, B. & Keller, U. (1997). Mechanism of alkaloid cyclopeptide synthesis in the ergot fungus *Claviceps purpurea*. *Chem. Biol.* **4**, 223-230.
36. van Wageningen, A.M.A., *et al.*, & Solenberg, P.J. (1998). Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic. *Chem. Biol.* **5**, 155-162.
37. Kao, C.M., Luo, G., Katz, L., Cane, D.E. & Khosla, C. (1995). Manipulation of macrolide ring size by directed mutagenesis of a modular polyketide synthase. *J. Am. Chem. Soc.* **117**, 9105-9106.
38. Kao, C.M., Luo, G., Katz, L., Cane, D.E. & Khosla, C. (1996). Engineered biosynthesis of structurally diverse tetraketides by a trimodular polyketide synthase. *J. Am. Chem. Soc.* **118**, 9184-9185.
39. Weissman, K.J., *et al.*, & Leadlay, P.F. (1998). Evaluating precursor-directed biosynthesis towards novel erythromycins through *in vitro* studies on a bimodular polyketide synthase. *Chem. Biol.* **5**, 743-754.
40. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. & Pease, L.R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**, 51-59.
41. Reichert, J., Sakaitani, M. & Walsh, C.T. (1992). Characterization of EntF as a serine-activating enzyme. *Protein Sci.* **1**, 549-556.
42. Gill, S.C. & von Hippel, P.H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319-326.
43. Cassady, C.J., Wronka, J., Kruppa, G.H. & Laukien, F.H. (1994). Deprotonation of multiply protonated ubiquitin ions. *Rapid Comm. Mass Spectrom.* **8**, 394-400.
44. Senko, M.W., Beu, S.C. & McLafferty, F.W. (1995). Determination of monoisotopic masses and ion populations for large biomolecules from resolved isotopic distributions. *J. Am. Soc. Mass Spectrom.* **6**, 229-233.

Because *Chemistry & Biology* operates a 'Continuous Publication System' for Research Papers, this paper has been published via the internet before being printed. The paper can be accessed from http://biomednet.com/cbiology/cmb – for further information, see the explanation on the contents pages.