Published in final edited form as:

Biochemistry. 2017 October 10; 56(40): 5380–5390. doi:10.1021/acs.biochem.7b00517.

Characterization of the Functional Variance in MbtH-like Protein Interactions with a Nonribosomal Peptide Synthetase

Rebecca A. Schomer and Michael G. Thomas^{*,iD}

Department of Bacteriology, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States

Abstract

Many nonribosomal peptide synthetases (NRPSs) require MbtH-like proteins (MLPs) for solubility or for activation of amino acid substrate by the adenylation domain. MLPs are capable of functional crosstalk with noncognate NRPSs at varying levels. Using enterobactin biosynthesis in *Escherichia coli* as a model MLP-dependent NRPS system, we use *in vivo* and *in vitro* techniques to characterize how seven noncognate MLPs influence the function of the enterobactin NRPS EntF when the cognate MLP, YbdZ, is absent. Using a series of *in vitro* assays to analyze EntF solubility, adenylation, aminoacylation, and *in vitro* enterobactin production, we show that interactions between MLPs and NRPSs are multifaceted and more complex than previously appreciated. We separate MLP influence on solubility and function in a manner that shows altered solubility is not indicative of a functional MLP/NRPS pair. Although much of the functional variation among these noncognates can be explained by differences in EntF affinity for an MLP or the extent an MLP alters EntF L-Ser affinity, we demonstrate that MLPs can have a broader impact beyond solubility and adenylation. First, we show that a noncognate MLP can affect formation of L-Ser-S-EntF. Second, under *in vitro* conditions saturating for substrate and MLP, enterobactin production remains compromised in the absence of an appropriate MLP partner. These data suggest that we expand our investigations into how the MLPs influence NRPS enzymology. A more detailed understanding of these influences will be essential for downstream engineering of hybrid NRPS systems.

Graphical abstract

^{*}Corresponding Author. Phone: (608) 263-9075. michael.thomas@wisc.edu.

ORCID

Michael G. Thomas: 0000-0001-8699-711X

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.7b00517.

Analysis of purified MLPs (Figure S1), phylogenetic tree of MLPs (Figure S2), immunoblot of tagged MLPs (Figure S3), growth after induction of MXAN_3118 (Figure S4), copurification of PvdL A4PCP4/PA2412 (Figure S5), analysis of soluble MLPs *in vivo* (Figure S6), kinetic analyses of EntF (Figures S7 and S8), copurification of EntF/Atu3678 (Figure S9), LC-MS analysis of ENT biosynthetic intermediates (Figure S10 and Table S5), strains (Table S1), primer sequences (Table S2), plasmids (Table S3) and gBlocks (Table S4) (PDF)

The authors declare no competing financial interest.

Nonribosomal peptide synthetases (NRPSs) employ a repetitive, modularly organized enzymology to condense amino acids into the backbones of many natural products. NRPS enzymology typically requires three core catalytic domains: an adenylation (A) domain, a peptidyl carrier protein (PCP) domain, and a condensation (C) domain. The A domains recognize and activate the appropriate amino acid as an amino acyl-AMP intermediate and then covalently tether the amino acid onto partner PCP domains, forming a thioesterified aminoacyl-*S*-PCP intermediate. The C domains catalyze amide-bond formation between neighboring aminoacyl-*S*-PCP intermediates in a directional manner. We¹ and others^{2,3} determined that within these core catalytic domains, there are two classes of A domains: those that require an auxiliary protein belonging to the MbtH-like protein (MLP) superfamily, and those that are MLP-independent.

MLPs are small, bacterial proteins that are nearly always encoded within a nonribosomal peptide biosynthetic gene cluster. For example, MbtH, after which this protein superfamily is named, is encoded by the eighth gene in the NRPS-encoding mycobactin biosynthesis gene cluster in *Mycobacterium tuberculosis* H37Rv.⁴ Although the gene encoding MbtH was discovered nearly 20 years ago,⁴ the role MLPs play in NRPS enzymology has remained an enigma. Initial analyses of pyoverdine biosynthesis showed the MLP-encoding gene, PA2412, to be involved in the production of siderophore and required for *Pseudomonas aeruginosa* growth in iron-limited medium.⁵ Further genetic evidence for their essentiality in producing an associated nonribosomal peptide came from thorough genetic analyses of MLP-encoding genes in *Streptomyces coelicolor*.^{6,7} In these studies, it was shown that deletion of the MLP-encoding gene in a targeted gene cluster did not fully disrupt production of the associated nonribosomal peptide in a wild-type *S. coelicolor*. In contrast, when all the MLP-encoding genes in this bacterium were deleted, including the one associated with the metabolite of interest, metabolite production was abolished. These genetic studies provided not only evidence for the importance of MLPs in producing nonribosomal peptides, but also showed the ability of noncognate MLPs to functionally replace each other at some level.

The first evidence that MLPs bind and influence the function of NRPSs came from the finding that nikkomycin biosynthesis involved the function of a monomodular NRPS where the MLP along with the A and PCP domains are fused as a single peptide.^{8,9} This observation was overlooked until our group¹ and the Walsh³ and Heide² groups reported that MLPs influenced the solubility of the associated NRPSs, along with their function. One of more important insights from these data came from the analysis of EntF, an NRPS involved in enterobactin (ENT) biosynthesis in *Escherichia coli*.^{1,10–12} This NRPS can be

overproduced and purified with or without its MLP partner, YbdZ, providing a way to compare the kinetics of the enzyme purified under these two conditions. Our data showed that when EntF was coproduced with YbdZ, it not only copurified with the MLP and had increased solubility, but the enzyme also had a higher affinity for its amino acid substrate L-Ser. This was the first insight into how an MLP can directly influence NRPS enzymology. More recent work has shown that the presence of YbdZ also has a minor influence on A domain affinity for the cosubstrate ATP.¹³

Despite observations that MLPs are essential to produce specific nonribosomal peptides and that, in the absence of the MLP, the NRPS has reduced solubility and affinity for its substrate, a mechanism for how MLPs have these impacts remains elusive. A clear understanding of this mechanism is complicated by the finding that not all MLP/NRPS interactions are equal, even within the same nonribosomal peptide biosynthetic pathway. For example, we have shown that MbtH itself has different interactions with each of its NRPS partners involved in mycobactin biosynthesis.¹⁴ MbtF and MbtE are completely insoluble without coproduction with MbtH, but MbtB can be overproduced and purified without MbtH. Kinetic analysis of the two forms of MbtB found they had equivalent kinetics for amino acid activation; thus, the only observed impact of MbtH on MbtB is to enhance its solubility. These observations suggest there are three forms of MLP-dependent NRPSs: (1) NRPSs that are completely dependent upon an MLP for solubility and activity (e.g., MbtF, MbtE); (2) NRPSs that have increased solubility in the presence of an MLP but do not require the MLP for amino acid activation (e.g., MbtB); and (3) NRPSs that have increased solubility and superior kinetics for amino acid activation in the presence an MLP (e.g., EntF). It should also be noted that even if a biosynthetic gene cluster codes for an MLP, it does not necessarily mean all the A domains of the NRPS encoded by the same gene cluster are MLP-dependent. For example, we have shown that in the viomycin, capreomycin, and ENT biosynthetic pathways, at least one of the A domains of the associated NRPSs is MLP-independent even though the associated gene cluster codes for an MLP.¹

Several studies have attempted to gain insights into the mechanism and specificity of MLPs in NRPS enzymology using structural biology. The first MLP-A structure, solved by Heide and colleagues,¹⁵ was a didomain protein wherein the MLP was naturally fused to the N-terminus of the A domain (an *in cis* MLP). The structure of this protein, SlgN1, revealed that the MLP/NRPS interface does not fall near the substrate binding pocket, suggesting the impact of the MLP on A domain substrate recognition is likely at a distance. The interactions between the MLP and A domains were proposed to be driven by two highly conserved tryptophan residues of the MLP, Trp25 and Trp35 (SlgN1 amino acid residue numbering), on the hydrophobic surface of the MLP that form a pocket that accommodates Ala433 of the partner A domain.^{3,15} Using alignments of A domains characterized as MLP-dependent or MLP-independent, it is proposed that the presence of the alanine or proline residue at this position of the NRPS could be an indication of an MLP-dependent A domain.

15

Recently, Gulick and colleagues have reported structures of EntF with and without YbdZ, as well as with a noncognate MLP, PA2412, from *Pseudomonas aeruginosa*.^{13,16} These structures confirmed the location of the MLP/NRPS interface in the more common *in trans*

MLP-dependent NRPS is the same as that observed in the *in cis* SlgN1 structure. Furthermore, these structures showed the MLP is not located in a region of the NRPS that would interfere with any of the other domains of the NRPS (EntF domain structure is C-A-PCP-thioesterase).¹³ Surprisingly, a comparison of the structures of the A domains in EntF, EntF+YbdZ, and EntF+PA2412 structures showed that the MLP had little to no influence on the conformation of the A domain.¹³ The Schmeing group corroborated this work by reporting a structure of DhbF in complex with its partner MLP that also showed no significant conformational changes from a NRPS in the absence of an MLP; however, a decrease in the apparent molecular weight of DhbF A domain complexed with an MLP versus the A domain alone suggest conformational changes that are not captured by crystallization.¹⁷ Unfortunately, even with increasing amounts of structural, biochemical, and genetic data, the mechanism for how an MLP influences NRPS catalysis is still unknown.

Finally, using the structural data and accumulated biochemical characterization, Gulick and colleagues attempted to elucidate a method to identify MLP-dependence from protein sequence using the sequence marker GEx₁₀GYx₁₀FxA/P.¹³ The critical Ala/Pro residue of the A domain coordinated by the Trp pocket of the MLP lies just downstream of the conserved A6 motif (GEx₁₀GY) of A domains.¹⁸ However, a majority of A domains, regardless of whether they were MLP-dependent or not, have a sequence that is compatible with MLP-NRPS interaction. The only conclusion that can be made is that, if a bulky site chain is found in the A/P position, it may be MLP-independent, and the only clear evidence for MLP-dependence relies on biochemical analysis of the NRPS.

We are interested in understanding both the specificity of an NRPS for its cognate MLP and how the MLP influences NRPS enzymology. To this end, we have chosen to focus our efforts on the ENT NRPS system in *E. coli* based on its simplicity, the genetic and biochemical tools readily available for its analysis, and the availability of high resolution structures. The ENT NRPS is a two-module system that contains only one MLP-dependent A domain that is contained in EntF (Figure 1). Furthermore, its natural MLP, YbdZ, is the only MLP encoded in the genome. EntF can also be overproduced and purified without YbdZ,¹ enabling the use of *in vitro* studies to evaluate interactions between YbdZ and EntF, as well as interactions between EntF and noncognate MLPs. These *in vitro* experiments can be coupled to *in vivo* analyses since strains of *E. coli* that lack *ybdZ* cannot grow in iron-limited media (ILM).¹ Here we present the *in vivo* and *in vitro* analyses of ENT biosynthesis when this NRPS system contains or lacks its natural MLP, and how the solubility and function of this system change when noncognate MLPs replace YbdZ.

MATERIALS AND METHODS

Bacterial Strains and Plasmid Construction

Bacterial strains used in this study are listed in Table S1. BW27783 *ybdZ* was constructed using the temperature sensitive plasmid pMAK705-*ybdZ* as previously described.^{1,19} All expression plasmids were constructed using the primer incomplete polymerase extension (PIPE) cloning method.²⁰ Primers used to amplify vectors and inserts are listed in Supplementary Table 2. Genes coding for MbtH, PA2412, Atu3678, MXAN_3118,

BSU31959, CmnN, and VioN were cloned into pBAD33 with the ribosome binding site of pET37B.¹ These constructs control the expression of these genes by an arabinose-inducible promoter.²¹ These genes as well as those for YbdZ, EntB, and EntF were also cloned into pTEV5, an overexpression vector that fuses a tobacco etch virus (TEV) cleavage site in between the peptide and an N-terminal H6-tag.²² PA2412, Atu3878, BSU31959, and MXAN_3118 were cloned into pACYC-duet-1 for overexpression of an untagged protein. Additional plasmids are referenced in Table S3. We note that a recent reconstruction of the *Myxococcus xanthus* DK1622 genome has renamed the locus tags for this bacterium and MXAN_3118 has been renamed MXAN_RS15115. However, at the time of submission, there was no NCBI database entry for MXAN_RS15115. For this reason, we continue to refer to this MLP as MXAN_3118 in this work.

Construction of Phylogenetic Tree for MLP Proteins

Protein sequences labeled as MbtH-like proteins were compiled from the NCBI protein database. At the time of this study, there were approximately 9000 MLP homologues available in the NCBI protein database. This data set was clustered based on percent identity using cdhit,²³ and representative sequences were used for MLPs with 100% identity. Any sequences greater than 150 amino acids in length that remained in the data set were also eliminated unless they were identified as fused MLP-A didomains. This limited the data set to 5393 unique sequences, and the phylogenetic tree was constructed from the alignment of these unique sequences. Alignments were created by MAFFT,²⁴ and FastTree²⁵ was used to create the phylogenetic tree. The Interactive Tree of Life (iTOL) was used to visualize the tree.²⁶

Growth of BW27783 ybdZ in Iron-Limiting Media

E. coli strain BW27783 or BW27783 ybdZ was transformed with pBAD33 or pBAD33 containing ybdZ, mbtH, PA2412, Atu3678, MXAN_3118, BSU31959, cmnN, or vioN. For all growth curves, each strain was grown in triplicate in LB with 34 $\mu\text{g mL}^{-1}$ chloramphenicol for 8 h at 30 °C. Cells were washed in M9 minimal media (Teknova). Two hundred microliters of M9 minimal media containing 0.4% (v/v) glycerol, 34 $\mu\text{g mL}^{-1}$ chloramphenicol, and 25 μM ethylenediamine-*N,N'*-bis(2-hydroxyphenylacetic acid) (EDDHA) was inoculated with 1 μL of each washed culture. For induction of pBAD33 promoter, arabinose was also added to the media to a concentration of 0.001% (w/v). Cultures were grown in 96-well plate in a Biotek Eon plate reader using orbital shaking at 580 rpm and 37 °C. Absorbance at 600 nm was recorded every 30 min for 48 h. Each growth curve was done in triplicate and then averaged before being plotted using GraphPad Prism (ver. 6.0h).

Co-overproduction of EntF with MLPs for Solubility Assay

E. coli strain BL21(DE3) ybdZ::acc(IV) was transformed with pTEV-entF and pACYC-duet-1 or pACYC-duet-1 containing an MLP-encoding gene (ybdZ, mbtH, PA2412, MXAN_3118, Atu3678, BSU31959, vioN, or cmnN), selecting for growth in 100 $\mu\text{g mL}^{-1}$ ampicillin and 34 $\mu\text{g mL}^{-1}$ chloramphenicol. Each strain was grown to saturation in 10 mL LB with ampicillin and chloramphenicol and then subcultured into 1 L of the same media. These cultures were grown at 25 °C to an OD₆₀₀ of 0.425 before being transferred to 16 °C

for 1 h. Isopropyl β -D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 60 μ M, and cells were shaken at 200 rpm at 16 °C for 12 to 16 h. Cells were harvested by centrifugation for 10 min at 6640g, and 100 mg of cells was resuspended in 1 mL of Buffer A. Buffer A contained 300 mM NaCl, and 20 mM Tris-HCl, pH 8. The cells were lysed by sonication at 20% power for 2 min with 1 s on, 1 s off pulses (Fischer Scientific Sonic Dismembrator model 500). Cell-free extract was generated by centrifugation for 10 min at 16100 $\times$ g. A total of 75 μ g of protein in the cell-free extract was separated by SDS-PAGE (12.5% acrylamide) and visualized with Coomassie blue staining to visualize the amount of soluble EntF. Samples were also analyzed using Tris-Tricine-based SDS-PAGE (12% acrylamide) and Coomassie blue staining to visualize the co-overproduced MLPs.

Detection of in Vivo MLP Levels by T7 Tag Monoclonal Antibody

E. coli BW27749 *ybdZ* was transformed with pBAD33 containing each MLP-encoding gene expressing an MLP with a C-terminal T7 tag. Construction of each pBAD-MLP-T7 expression plasmid involved gBlocks (Integrated DNA Technologies) and restriction cloning (Table S4). PA2412, Atu3678, BSU31959, *cmnN*, and *vioN* were ligated into pBAD33 using the *Xba*I and *Pst*I restriction sites. The genes coding for YbdZ, MbtH, and MXAN_3118 were ligated into pBAD using *Xba*I and *Hind*III restriction sites. Each resulting expression strain was grown for 8 h in 2 mL LB with 34 μ g/mL chloramphenicol. Cells were washed in M9 minimal media (Teknova), and 100 μ L was subcultured into 3 mL of M9 minimal media supplemented with 0.4% glycerol and 34 μ g/mL chloramphenicol. When required, increased expression of the gene coding for an MLP was induced by the addition of 0.001% (w/v) arabinose when the culture reached an OD₆₀₀ of 0.5 and then incubated at 37 °C for 1 h before harvesting. Cell pellets were weighed and resuspended in 100 μ L 1 $\times$ Bugbuster (Novagen) per 100 mg of cell paste. Cultures were lysed for 10 min while gently rocking at room temperature. A total volume of 10 μ L of cell extract from each sample was separated by SDS-PAGE and transferred to PDVF membrane using a semidry apparatus for 30 min at 15 V. Before transfer, the SDS-PAGE gel and PDVF membrane were equilibrated in semidry transfer buffer containing 50 mM Tris-HCl, 40 mM glycine, 0.037% (w/v) SDS, 20% (v/v) MeOH, pH 8.3. The membrane was dried for 1 h at 37 °C before a 1 h incubation in PBS pH 7.4 with 2.5% (w/v) milk, 1% (v/v) Tween-20, and 1:5000 dilution of T7 monoclonal antibody-horseradish peroxidase conjugate. Membrane was washed twice for 5 min in PBS before exposure to SuperSignal West Pico chemiluminescent substrate (ThermoScientific).

Overproduction of Proteins for Purification

E. coli BL21(DE3) *ybdZ::acc(IV)* was used as the background strain for all overproductions in this study. All cultures were grown to saturation in LB supplemented with the appropriate drugs and 10 mL was subcultured into each liter of LB used for overproduction. Strains containing overexpression constructs for MLPs were grown in 3 L LB with 100 μ g mL⁻¹ ampicillin. The strain containing only pTEV5-*entF* was grown in total of 12 L LB supplemented with 100 μ g mL⁻¹ ampicillin. Strains for co-overproductions containing pTEV5-*entF* or pTEV5-*entB* with pSU20-*sfp*, pACYC-duet-1-*ybdZ*, or pACYC-duet-1-Atu3678 were grown in 6 L of LB supplemented with 100 μ g mL⁻¹ ampicillin and 34 μ g mL⁻¹ chloramphenicol. All strains were grown at 25 °C to an OD₆₀₀ of 0.425 before being transferred to 16 °C for 1 h. IPTG was added to a final concentration of 60 μ M and cells

were shaken at 200 rpm at 16 °C 12 to 16 h. Cells were harvested by centrifugation for 10 min at 6641g, and pellets were frozen at –20 °C before purification.

Purification of MLPs

MLPs were purified by nickel-affinity chromatography following a previously described protocol.²⁷ The 6x His-tag of each protein was then cleaved by TEV, and uncleaved protein was removed by a second round of nickel affinity chromatography as previously reported, but with slight modification.^{14,22} Briefly, after purification by nickel affinity, fractions containing the MLPs were dialyzed at 4 °C overnight with a 1:50 molar ratio of H6-TEV protease against Buffer B, containing 50 mM NaCl, 50 mM Tris-HCL, 5 mM DTT, pH 7.9, to cleave the H6 tag. The proteins were transferred to Buffer C, containing 25 mM NaCl, 50 mM Tris-HCl at pH 7.9, to remove DTT before the protein was passed through Ni-NTA resin for a second time. Resin was removed by centrifugation, and tagless protein was visualized by SDS-PAGE and Coomassie blue staining. Each protein was concentrated using a Centrprep centrifugal filter YM-3, flash frozen using liquid nitrogen, and stored at –80 °C. Only MLPs lacking an N-terminal his tag were used for *in vitro* studies. The concentrations of the MLPs were determined using the calculated molar extinction coefficients of the untagged proteins (YbdZ, 23 615 M⁻¹ cm⁻¹; MbtH, 17 990 M⁻¹ cm⁻¹; PA2412, 20 970 M⁻¹ cm⁻¹; MXAN_3118, 20 970 M⁻¹ cm⁻¹; Atu3678, 20 970 M⁻¹ cm⁻¹; BSU31959, 20 970 M⁻¹ cm⁻¹; CmnN, 19 480 M⁻¹ cm⁻¹; VioN, 19 480 M⁻¹ cm⁻¹). Purified MLPs are shown in Figure S1.

Purification of apo-Ent, holo-EntF, holo-EntB, apo-EntF/YbdZ, and apo-EntF/Atu3678

For all overproductions of EntF, nickel affinity chromatography and H6-tag cleavage by TEV protease were performed the same as the purification of the MLPs along, but with two additional purification steps. After the second pass through Ni-NTA resin, protein was loaded onto a 5 mL HiTrap Q Sepharose FastFlow column (New England BioLabs) and, using a low-pressure chromatography system, protein was eluted over a gradient from 100% Buffer C/0% Buffer D to 0% Buffer C/100% Buffer D over 80 min at a flow rate of 1 mL min⁻¹. Buffer D contained 500 mM NaCl, 50 mM Tris-HCl, pH 7.9. Fractions containing EntF were identified by SDS-PAGE and Coomassie blue staining. These fractions were pooled and concentrated to 1 mL using a Centrprep YM-30 and dialyzed in Buffer C at 4 °C overnight. Finally, the sample was loaded into a HiPrep 16/60 Sephracryl S-200 high-resolution column for size exclusion chromatography. Using the low-pressure chromatography system, protein was eluted with buffer C at a flow rate of 0.5 mL min⁻¹. Fractions containing the protein of interest were identified by SDS-PAGE and Coomassie blue staining. These fractions were pooled and concentrated in a Centrprep centrifugal filter YM-30 (Millipore). The samples were then flash frozen in liquid nitrogen and stored at –80 °C until use. The concentration of purified EntB, EntE, and EntF proteins were calculated using the calculated molar extinction coefficients. For EntF/YbdZ and EntF/Atu3678, the protein concentration was calculated using a calculated molar extinction coefficient of a fusion protein of EntF-MLP (EntF-YbdZ, 207 640 M⁻¹ cm⁻¹; EntF-Atu3678, 205 120 M⁻¹ cm⁻¹). Subsequent analysis of the EntF/MLP complex by sample dilution followed by SDS-PAGE, Coomassie blue staining, and comparison to a known

quantity of purified EntF confirmed the EntF concentration in the EntF/YbdZ and EntF/Atu3678 complexes.

Radiolabeled ATP/PP_i Exchange Assays

ATP/PP_i exchange assays were performed as previously described.^{1,14} Each 100 μ L assay contained 75 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 5 mM dithiothreitol (DTT), 3.5 mM ATP pH 7.5, 1 mM NaPP_i pH 7, [³²P] PP_i (0.9 Ci mol⁻¹, PerkinElmer), and varying concentrations of enzyme, MLP, and L-Ser. All assays were performed at room temperature in the linear range of enzyme concentration and less than 10% substrate to product conversion. To determine the kinetic parameters of EntF interaction with each MLP, the L-Ser concentration was held constant at 0.2 mM, while the concentration of MLP was varied. MLP concentrations were 0.05, 0.1, 0.25, 0.5, 0.75, 1.0, and 2.0 μ M. To determine the pseudo-first order kinetic parameters of L-Ser activation, the ATP concentration was held constant, while the L-Ser concentration was varied. For kinetics of L-Ser activation, MLP concentration was held at 5 $\times$ the k_d determined in this study, and the reaction was incubated in the reaction mixture for 10 min before the addition of amino acid substrate. For assays of EntF without an MLP and with CmnN or VioN, L-Ser concentrations used were 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.5, 2.0, 3.0, 4.0, 5.0, 6.0, 7.5, and 10 mM. For assays of EntF/YbdZ or EntF with YbdZ, MbtH, PA2412, MXAN_3118, or BSU31959, L-Ser concentrations used were 0.01, 0.025, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.5, 0.6, 0.8, and 1 mM. Kinetic parameters were determined using nonlinear regression analysis (GraphPad Prism ver. 6.0h). All kinetic studies were performed in the linear range of the assay for substrate to product conversion (<10% substrate to product conversion) and EntF concentration.

Formation of L-Ser-S-EntF

Formation of L-Ser-S-EntF was monitored using 30 μ L reaction mixtures containing 50 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 1 mM Tris(2-carboxyethylphosphine) (TCEP), 20 μ M EntF, 1 μ M Sfp, 500 μ M Coenzyme A, 5 mM ATP, pH 7 and 0.1 mM [¹⁴C]-L-Ser (100 mCi mmol⁻¹, PerkinElmer).²⁷ Reactions were incubated for 2 h at 37 °C to phosphopantetheinylate EntF before the addition of ATP, [¹⁴C]-L-Ser or MLP. Five times the k_d of EntF for each MLP determined by ATP-PP_i assays was used as the concentration of the MLP in this assay. After phosphopantetheinylation, ATP and MLP were added to the assay mixtures and incubated for 5 min at room temperature before the addition of [¹⁴C]-L-Ser. Reactions were stopped after 40 min by aliquoting 10 μ L of the reaction mixture into an equal volume of 2 $\times$ sample buffer containing 60 mM Tris-HCl, pH 6.8, 10% (w/v) SDS, 50% (v/v) glycerol, and 0.4% (w/v) bromophenol blue and 5 μ L was loaded onto an SDS-PAGE gel. Protein was visualized by Coomassie blue staining before the gel was dried and exposed to a phosphorimaging screen for 24 h. The gels were exposed to the phosphorimaging screen alongside a set of known [¹⁴C] standards to generate a standard curve for pixels to disintegrations per minute used to quantify [¹⁴C] L-Ser-S-EntF produced.

In Vitro Reconstitution of ENT Production

Production of ENT *in vitro* was based off of a previously published protocol¹² and monitored by chrome azurol S (CAS) liquid assays.²⁸ Each 200 μ L reaction mixture contained 75 mM Tris-HCl pH 7.5, 10 mM ATP, 5 μ M TCEP, 15 μ M MgCl₂, 300 nM EntE,

5 nM holo-EntF, 5 μ M holo-EntB and MLPs to a concentration 5 $\times$ the k_d determined by ATP-PP_i exchange assays in this study. Phosphopantetheinylation of EntF and EntB was achieved *in vivo* by co-overproduction of EntF and EntB with Sfp. EntE purification was previously published¹ and stored at -80°C until use. L-Ser and 2,3-dihydroxybenzoate (2,3-DHB) were added to final concentrations of 500 μ M and 10 mM, respectively, to initiate the reaction. Reactions lacking EntB were performed as a control. Assays were incubated at 37°C until 30 μ L of sample reaction was aliquoted into 120 μ L of CAS reagent to stop the reaction at 15 min intervals. CAS and reaction mixtures were incubated at room temperature for 5 min before the reciprocal absorbance at 630 nm was measured.

LC/MS Analysis of *in Vitro* ENT Biosynthesis Products

Reaction mixtures were prepared exactly as for the *in vitro* monitoring of ENT production by CAS reagent. Samples were incubated at 37°C for 60 min and quenched by the addition of equal volume 1 mM HCl. ENT and byproducts were removed by ethyl acetate extraction. A volume of 750 μ L ethyl acetate was added to the quenched reaction. From the ethyl acetate layer, 500 μ L was removed and concentrated by evaporation in an Eppendorf vacufuge. The remaining residue was resuspended in 50 μ L in 30% acetonitrile in H₂O, and 10 μ L of each sample was analyzed by ultrahigh-performance liquid chromatography (UHPLC) coupled with mass spectrometry (MS). The samples were separated on a ZORBAX Eclipse XDB-C18 column (Agilent, 2.1×150 mm with a 1.8 μ M particle size). Before every run, the system was equilibrated for 5 min at 0.5% (v/v) formic acid (FA) in 20% acetonitrile. A linear gradient of 0.5% FA in 20% to 100% acetonitrile was delivered over 15 min by a Vanquish UHPLC system (Thermo Scientific) with a flow rate of 0.2 mL min⁻¹. The gradient was followed by an isocratic step at 100% acetonitrile in 0.5% FA for 5 min. The UHPLC system was coupled to a Q Exactive hybrid quadrupole Orbitrap MS (Thermo Scientific). For electrospray ionization, the ion voltage was set at -3.5 kV in negative mode. The parent ion of the highest intensity was fragmented in the ion trap mass spectrometer. Data analysis was performed using Xcalibur (ThermoFisher Scientific) and Maven²⁹ software.

RESULTS AND DISCUSSION

Noncognate MLPs Can Functionally Replace YbdZ *In Vivo* to Varying Extents

We targeted seven MLPs from diverse sources to test their ability to functionally replace YbdZ in our model NRPS/MLP system. These MLPs map to several phylogenetically diverse lineages (Figure S2) and are associated with NRPS systems that produce distinct metabolites. The MLPs included in this study, besides YbdZ associated with ENT biosynthesis in *E. coli*,¹ include MbtH associated with mycobactin biosynthesis in *M. tuberculosis* str. H37Rv,⁴ PA2412 associated with pyoverdine biosynthesis in *P. aeruginosa* PAO1,⁵ Atu3678 associated with siderophore biosynthesis in *Agrobacterium fabrum* str. C58,³⁰ BSU31959 associated with corynebactin biosynthesis in *Bacillus subtilis* strain 168,³¹ MXAN_3118 an orphan MLP from *Myxococcus xanthus* DK1622,³² VioN associated with viomycin biosynthesis in *Streptomyces* sp. strain ATCC11863,³³ and CmnN associated with capreomycin biosynthesis in *Saccharothrix mutabilis* subsp. capreolus.³⁴

The gene coding for each MLP was cloned into a pBAD33 vector and transformed into the *E. coli* strain BW27749 *ybdZ*. We assessed whether the resulting strains were able to grow in iron-limited medium (Figure 2), which contained 25 μ M EDDHA as a competing iron chelator. The presence of MbtH or PA2412 restored growth of BW27749 *ybdZ* to a level comparable to that observed for YbdZ. In contrast, VioN or CmnN had similar growth characteristics to the empty-vector control. Strains carrying Atu3678, BSU31959, and MXAN_3118 had varying levels of growth in ILM, but all three MLPs failed to enable growth at a level similar to that observed when YbdZ was present.

One hypothesis for the observed differences in complementation is the varying levels of amino acid identity or similarity the noncognate MLPs have with YbdZ. On the basis of amino acid sequence alignments, this does not appear to be the case since the ability to complement for the loss of YbdZ does not correlate to similarity of the MLP to YbdZ (Figure 3a). In fact, despite having reduced or no ability to complement, BSU31959, CmnN, and VioN shared more amino acid identity with YbdZ than PA2412 and Atu3678 (Figure 3b), even though the latter were more efficient at replacing YbdZ.

Several structures of MLPs in complex with an NRPS have been published, allowing a more focused comparison of the MLP residues likely to interact with the NRPS.^{13,15,17} Residues involved in NRPS–MLP interaction in these structures were cross-referenced with the sequence alignments (Figure 3a). Among these residues, there is significant commonality between functional and nonfunctional MLPs. Despite this similarity at the interface, there are obvious functional differences among the MLPs resulting in the observed gradation of growth phenotypes. This incongruity between functionality and sequence similarity may imply that the overall topography of the MLP/NRPS interface, rather than a specific set of residues, allows for the MLP and NRPS to interact functionally *in vivo*. A compromised ability to complex correctly with a partner MLP due to topographical incompatibilities could lead to a variety of impacts on NRPS solubility or enzymology.

Varying in Vivo Levels of the MLPs Do Not Fully Explain Differential Complementation

Immunoblotting was used to investigate whether differences in the *in vivo* levels of the MLPs explained the varying degrees of complementation. A polyclonal antibody against YbdZ was only able to detect YbdZ, MbtH, and MXAN_3118; thus, it could not be used to compare the *in vivo* levels of all the MLPs. In an attempt to circumvent this issue, all of the pBAD33-MLP expression plasmids were reconstructed to produce an MLP with a C-terminal T7 tag. Commercially available monoclonal antibodies to the T7 tag were then used in immunoblotting experiments in an effort to detect the *in vivo* levels of the MLPT7 fusion proteins. Importantly, all of the MLP-T7-expression constructs resulted in complementation results consistent with those observed with the untagged versions shown in Figure 2.

We assessed the *in vivo* MLP-T7 levels when the constructs were grown in M9 minimal media containing glucose as the carbon source. We do not detect a growth defect in a *ybdZ* strain in this medium. Immunoblotting did not detect any MLP-T7 unless arabinose was added to increase the expression of the genes from the pBAD promoter. Under these media and expression conditions, we detected comparable levels of YbdZ-T7, MbtH-T7, MXAN_3118-T7, BSU31959-T7, and CmnNT7, with lower levels of VioN-T7 (Figure S3).

Surprisingly, we failed to detect any PA2412-T7 or Atu3678-T7 even though these constructs compensated for the loss of YbdZ. At this time, it is unclear why these proteins are not detectable, yet they are able complement for the loss of YbdZ.

These results also show that the MLPs that fail to fully compensate for the loss of YbdZ can be produced at higher levels in the cells than when relying on the background expression from the pBAD promoter. Reassessing complementation in ILM when increasing expression of the MLP genes found only MXAN_3118 gained more ability to compensate for the loss of YbdZ (Figure S4). Also, increased expression of *ybdZ* did not enhance the growth of the *ybdZ* strain, suggesting the background level expression of *ybdZ* from the pBAD promoter is efficient in providing enough of this MLP for ENT production under these growth conditions. Together, these data suggest the failure of BSU31959, CmnN, and VioN to fully replace YbdZ *in vivo* was unlikely to be due solely to low *in vivo* levels of these proteins.

MLPs Influence EntF Solubility to Varying Levels

We hypothesized that one explanation for the observed differences in complementation was a variation in the ability of an MLP to influence the solubility of EntF. This hypothesis was based on the consistent finding of increased solubility of MLP-dependent NRPSs when these enzymes are overproduced with a partner MLP.^{1-3,11,14,35} To investigate how the solubility of EntF was affected in the presence of different MLPs, we overproduced EntF in *E. coli* BL21(DE3) *ybdZ::aac(3)IV* in the presence or absence of each of the targeted MLPs and assessed the level of EntF solubility by SDS-PAGE and Coomassie blue staining of soluble protein in cell-free extracts (Figure 4). As we previously reported,¹ EntF solubility is significantly improved when co-overproduced with its cognate MLP, YbdZ. Here, we found that EntF solubility does not always correlate with the *in vivo* complementation data for a given noncognate MLP (Figure 2). For example, PA2412 is capable of restoring growth in BW27749 *ybdZ* in ILM to levels comparable to YbdZ (Figure 2), but fails to show enhanced EntF solubility when the two proteins were co-overproduced (Figure 4). Using this same PA2412 overexpression construct, we determined that PA2412 was required for solubility of three of the four adenylationpeptidyl carrier protein domains of PvdL from *P. aeruginosa* (Figure S5). Furthermore, each MLP could be detected in the soluble fraction of these samples at levels at least comparable to YbdZ (Figure S6). On the basis of those data, we concluded that the low solubility of EntF was not due to a failure of our PA2412 overexpression construct to overproduce a functional MLP; rather PA2412 does not increase the solubility of EntF under these conditions. In contrast to PA2412, neither CmnN nor VioN restored growth of *E. coli ybdZ* in ILM, but both resulted in significantly more EntF in the soluble fraction of the cell extracts when the NRPS was coproduced with these MLPs (Figure 4), although VioN may enable a lower level of soluble EntF compared to CmnN. The remaining MLPs (MbtH, MXAN_3118, Atu3678, and BSU31959) all enabled increased EntF solubility compared to the no-MLP control, with BSU31959 resulting in slightly lower EntF levels. These data indicate that the effects MLPs have on NRPS solubility and activity are separable. Unfortunately, repeated attempts to detect chromosomally encoded EntF using monoclonal antibodies to amino acid tags introduced into the coding region of either the 5' or 3' end of *entF* were not successful; thus, we were unable to assess EntF solubility under physiologically relevant conditions.

Defining the K_d for Binding of Noncognate MLPs to EntF by *in Vitro* Reconstitution of Complexes

To gain further insight into the observed differences in complementation and EntF solubility, we characterized the adenylation activity of the EntF and MLP complexes *in vitro*. We will refer to the different forms of EntF as either EntF (overproduced and purified independently of an MLP) or EntF/MLP (co-overproduced with an MLP). When an MLP is added to EntF, this form of the NRPS will be referred to as EntF+MLP. Consistent with prior reports by us¹ and Gulick and colleagues,¹³ our current study found that the impact YbdZ has on the activation of L-Ser is predominantly in enhancing the affinity of the enzyme for the amino acid, as noted by the approximately 10-fold lower K_m and a minimal impact on the k_{cat} (Table 1). We exploited the difference in K_m to determine the K_d for the binding of YbdZ to EntF. Briefly, we set the concentration of L-Ser to 200 μ M, which is approximately 10-fold below the K_m determined for EntF but near the K_m of EntF/YbdZ, to maximize the differences in the rate of L-Ser activation by the two forms of the NRPS. We then added increasing concentrations of purified YbdZ to the EntF reaction and monitored the increase in the reaction rate. Plotting the rate of the reaction versus YbdZ concentration enabled us to then use nonlinear regression analysis to define the K_d for the binding of YbdZ to EntF (Table 1). We then determined the kinetics of L-Ser activation of the reconstituted EntF +YbdZ complex when YbdZ was added at a concentration 5-fold higher than the K_d . These data showed that the kinetics of L-Ser activation by EntF+YbdZ were restored to levels comparable to EntF/YbdZ (Table 1). Since the kinetic parameters of EntF+YbdZ and EntF/YbdZ are comparable, we presumed the reconstituted EntF+YbdZ complex is relevant to the EntF/YbdZ complex formed *in vivo*. Therefore, we could use this reconstitution system to test the hypothesis that the observed differences in functional replacement of YbdZ by noncognate MLPs is due to variation in the binding affinities of these proteins for EntF. We note that repeated attempts to use alternative methods (differential scanning fluorimetry or isothermal titration calorimetry) to determine the K_d for binding of YbdZ to EntF gave uninterpretable results, necessitating the use of the approach outlined above.

Using the same method for determining the K_d for binding of YbdZ to EntF, we probed the interactions between EntF and the noncognate MLPs. We found that all the noncognate MLPs, with the exception of Atu3678, had an influence on the reaction rate of EntF and allowed us to define a K_d for each MLP (Table 1). As observed with the solubility studies, the differences in the K_d for binding of the MLPs to EntF did not correlate with the *in vivo* complementation studies. For example, CmnN and VioN both had K_d values similar to YbdZ, but neither could functionally replace YbdZ *in vivo* for ENT production. While the results were not consistent with the complementation studies, they were consistent with the observed influence on EntF solubility. The MLP with the highest K_d value (PA2412) resulted in the lowest levels of EntF solubility (Figure 4). The next two highest K_d values belonged to BSU31959 and VioN, both of which resulted in slightly reduced EntF solubility relative to YbdZ (Figure 4). Interestingly, the two MLPs with the lowest K_d values (MbtH and CmnN) both influenced EntF solubility but had opposite results in growth studies with MbtH fully complementing for the loss of YbdZ and CmnN failing to show any level of complementation.

Kinetics of L-Ser Activation by EntF + MLP Complexes

With the K_d values determined, we could address whether the EntF+MLP complexes have different kinetics of L-Ser activation, potentially explaining differences in *in vivo* complementation. Each MLP was added to a reaction at a concentration 5-fold higher than the respective K_d . The kinetics of L-Ser activation using standard ATP/PP_i exchange assays were then determined. As noted above, the addition of YbdZ to the reaction restored the kinetics of EntF-catalyzed L-Ser activation to levels comparable to those of the EntF/YbdZ complex (Table 1). Solely on the basis of the kinetic data, the noncognate MLPs fell into two groups. EntF+MbtH, EntF+PA2412, EntF+MXAN_3118, and EntF+BSU31959 all had k_{cat}/K_m values approximately half of that observed for EntF+YbdZ, with the decrease in these values attributable mainly to a doubling of the K_m of EntF for L-Ser (Table 1). In contrast, EntF+CmnN and EntF+VioN each showed kinetic parameters more similar to EntF lacking an MLP (Table 1). This division did not, however, correlate entirely with the *in vivo* complementation data. The failure of either CmnN or VioN to functionally replace YbdZ *in vivo* may be due solely to the poor affinity of the resulting NRPS+MLP complex for L-Ser (Table 1). Metabolomics has determined that when *E. coli* is growing on glycerol, as in the current complementation study, the intracellular concentration of L-Ser is approximately 150 μ M.³⁶ This is almost 10-fold below the K_m of EntF+CmnN and EntF+VioN for L-Ser. Interestingly, both CmnN and VioN enable EntF to be more soluble, but these protein–protein interactions presumably fail to hold EntF in the correct conformation to positively influence the K_m for L-Ser. Alternatively, MbtH and PA2412 were able to efficiently replace YbdZ and enable growth of BW27783 *ybdZ* in ILM. These results suggest a doubling of the K_m for L-Ser does not impact the enzymology enough to inhibit growth, yet BSU31959 poorly replaces YbdZ *in vivo* despite having kinetic parameters comparable to EntF+MbtH or EntF+PA2412 complexes (Table 1). This suggested that the reason for the failure of BSU31959 to fully replace YbdZ *in vivo* was due to something other than L-Ser activation.

The failure to detect any impact of Atu3678 addition on the activity observed with EntF suggests that a functional EntF+Atu3678 complex cannot be formed *in vitro*. We hypothesized this based on the findings that Atu3678 does complement for the loss of YbdZ in growth studies (Figure 2) and enables more soluble EntF to be detected in cell free extracts (Figure 4), but fails to impact EntF enzymology after overproduction and purification independent of EntF. To test this, we co-overproduced and purified an EntF/Atu3678 complex (Figure S9) and determined the kinetic parameters for L-Ser activation (Table 1). The kinetic parameters of EntF/Atu3678 were more similar to EntF/YbdZ than EntF, leading us to conclude that production of functional EntF+Atu3678 requires coproduction of these proteins. Since there is no way to control K_d effects using the copurified complex and due to our inability to reconstitute an EntF+Atu3678 interaction *in vitro*, this MLP was not characterized further.

Analysis of Aminoacylation by EntF+MLP Complexes

The observation that EntF+BSU31959 has L-Ser activation kinetics similar to other EntF+MLP complexes, but fails to functionally replace YbdZ *in vivo*, led us to investigate whether this MLP compromises EntF in aminoacylation of the PCP domain. To measure aminoacylation, we first phosphopantetheinylated the purified EntF *in vitro* using the

phosphopantetheinyltransferase, Sfp. Since the same EntF preparation was used in all reactions, it is reasonable to assume that observed differences between EntF+MLP complexes would not be due to differences in levels of phosphopantetheinylation. The phosphopantetheinylated EntF was incubated with ATP, [^{14}C]-L-Ser, and the targeted MLP for a set amount of time, and the amount of L-Ser-*S*-EntF formed was detected and quantified using SDS-PAGE followed by phosphorimaging (Figure 5). The concentration of [^{14}C]-L-Ser could only be added at levels comparable to the K_m of L-Ser for EntF+YbdZ (100 μM) because of limitations in the specific activity of commercially available [^{14}C]-L-Ser. We observed a maximal EntF-*S*-L-Ser level of approximately 50% in the presence or absence of YbdZ, consistent with prior EntF/YbdZ studies.³⁷ While we observed EntF requiring additional time to form the EntF-*S*-L-Ser intermediate compared to EntF+YbdZ, an accurate rate was difficult to determine because of the limitation in the sensitivity of the assay. Because of this, we determined a time at which the levels of EntF aminoacylation were similar between EntF and EntF+YbdZ and used an end-point assay to assess the level of aminoacylation by each EntF+MLP complex. This approach did not lend itself to determining the rate of EntF-*S*-L-Ser formation; rather it only evaluated whether an EntF+MLP complex forms the aminoacylated intermediate at a level than either EntF or EntF+YbdZ under these assay conditions.

EntF+MbtH, EntF+PA2412, EntF+MXAN_3118, and EntF+VioN showed equivalent aminoacylation levels to both EntF+YbdZ and EntF. In contrast, the EntF+BSU31959 complex failed to aminoacylate at a level comparable to EntF with or without the MLP. This suggests BSU31959 holds EntF in a conformation that negatively impacts EntF aminoacylation (Figure 5). These data offer an explanation for why BSU31959 cannot complement for the loss of YbdZ *in vivo* despite positively influencing both the solubility and kinetics of EntF adenylation (Figure 4 and Table 1). The EntF+CmnN complex is also moderately reduced in its ability to form the aminoacylated intermediate. These data are the first direct evidence of an MLP having an impact on a step in NRPS catalysis other than amino acid activation.

Influence of MLPs on the *in Vitro* Rate of ENT Production

Our data investigating A domain-associated activities suggested that MLPs may also influence functions besides aminoacyl-AMP and aminoacyl-*S*-PCP formation. To investigate this possibility, we reconstituted ENT biosynthesis *in vitro* to enable us to control for differences in the K_d for MLP interactions with EntF, EntF solubility, and affinities of EntF for L-Ser. We reconstituted ENT biosynthesis using prior studies by Walsh and colleagues as a guide.¹² Briefly, EntE, holo-EntB, holo-EntF, and saturating concentrations of the appropriate MLP were incubated with saturating concentrations of substrates, 2,3-dihydroxybenzoate (DHB), ATP, and L-Ser. In these reactions, the only limiting factor was how efficiently EntF functions, with or without an MLP, when integrated into the complete ENT biosynthetic pathway. To detect ENT production over time, the reaction was terminated by the addition of an aliquot of the reaction to a CAS reagent. The CAS reagent loses its absorption at 630 nm when the iron is removed by a chelator with a higher affinity for ferric iron²⁸ providing a simple colorimetric assay for detecting ENT biosynthesis.

A comparison of ENT biosynthesis by EntF+YbdZ versus EntF clearly showed a significant increase in ENT production when YbdZ was present (Figure 6). The failure of EntF to assemble ENT at a level comparable to EntF+YbdZ even at saturating L-Ser concentrations suggested EntF function was disrupted in more than just amino acid binding or L-Ser-*S*-EntF formation. Thus, our prior proposal that the reason a *ybdZ* strain of *E. coli* cannot grow in ILM is the decreased affinity of EntF for L-Ser is likely incomplete.¹ EntF+BSU31959, EntF+CmnN, and EntF+VioN also failed to synthesize ENT at a rate comparable to EntF+YbdZ even at high L-Ser concentrations (Figure 6). While the reduced rate of L-Ser-*S*-PCP formation by EntF+BSU31959 may explain its reduced rate of ENT biosynthesis, the reduced rate of ENT production by EntF+CmnN and EntF+VioN even at saturating levels of L-Ser suggests the NRPS is altered for more than just a reduced affinity for the amino acid substrate. One possible reason for the reduced rate of ENT production by EntF+CmnN and EntF+VioN is a reduced ability of these noncognate complexes to condense DHB from EntB onto L-ser-*S*-EntF. Initial studies of the ENT biosynthesis pathway indicate that the formation of (DHB-L-Ser)-*S*-EntF intermediates occurs rapidly; therefore, monitoring the rate of DHB loading would require rapid quench techniques.¹² Instead, we tested whether these complexes are reduced in their ability to function iteratively to generate the complete ENT cyclic trimer. We reasoned that if the enzyme complexes stall in the middle of ENT formation, these intermediates would be hydrolyzed from EntF and there would be a detectable buildup of intermediates in the reaction mixture. We used LC/MS to detect and quantify the relative amounts of ENT along with monomer (DHB-L-Ser), dimer (DHB-L-Ser),² and trimer (DHB-L-Ser)³ intermediates. The amounts of ENT produced by EntF, EntF+VioN, and EntF+CmnN were lower in comparison to the amount of ENT produced by EntF+YbdZ as expected (Figure 7). However, the levels of the biosynthetic intermediates were all similarly reduced as well (Figure S10, Table S5), suggesting the entire ENT biosynthetic machinery is functioning a lower level when these noncognate MLPs are present.

The remaining three EntF+MLP complexes biosynthesized ENT at a rate comparable to EntF+YbdZ.

CONCLUSIONS

The role MLPs play in NRPS enzymology has been an enigma since their initial discovery as a potential ORF in the ENT biosynthesis gene cluster¹⁰ and the subsequent naming of the protein superfamily after the eighth ORF in the mycobactin biosynthesis gene cluster.⁴ Results from the original *in vitro* studies¹⁻³ suggest that MLPs play a chaperone-like role in enabling proper folding of the NRPS, specifically impacting the A domain. While there are some MLPs that appear to function in a manner analogous to molecular chaperones, e.g., MbtH enabling enhanced solubility of MbtB without being required for catalysis,¹⁴ many MLPs do not dissociate from the NRPS once complexed with the NRPS, suggesting they play a more active role in NRPS enzymology.^{1,3,14}

Here, we aimed to gain additional insights into MLP function by studying noncognate NRPS +MLP complexes using the ENT NRPS system from *E. coli* (data summarized in Figure 8). We provide evidence that the enhanced solubility and catalysis that MLPs provide an NRPS are separable functions in a manner distinct from those observed with MbtB+MbtH. PA2412

is able to enhance EntF catalysis, but did not provide increased EntF solubility. In contrast, CmnN and VioN both provided increased EntF solubility, but both failed to positively influence EntF catalysis. How these differences occur is not evident at this time even when amino acid sequence comparisons, *in vivo* and *in vitro* data, and high-resolution structure of EntF+PA2412¹³ are combined. Directed evolution studies to evolve the NRPS+MLP interactions and evolve A domains to function independent of an MLP are ongoing and may provide more insight into these questions.

To date, all analyses of NRPS+MLP interactions have focused on the impact MLPs have on amino acid activation by A domains. Here we have provided the first evidence that improper NRPS+MLP interactions can negatively impact not only amino acid activation, but also aminoacyl-*S*-PCP formation (Figure 5), along with the overall function of the NRPS (Figure 7). When these data are combined with the solubility data, they suggest that simple assessments of NRPS+MLP interactions based solely on enhanced solubility or amino acid activation may not provide the necessary information to conclude a properly functioning NRPS+MLP interaction is occurring.

These conclusions raise important issues for the long-term goals of generating fully functional chimeric NRPSs. Improper NRPS+MLP pairings may negatively impact NRPS enzymology in unexpected ways as summarized in Figure 8. For example, the introduction of an NRPS module or A domain along with its partner MLP may not solve the MLP-dependence problem due to complications that arise from noncognate pairings from both native and introduced NRPS+MLP systems. Overcoming these obstacles will require a better understanding of MLP+NRPS specificity, the ability to convert an MLP-dependent NRPS to one that is MLP-independent, or the discovery of a universal MLP that functions with all NRPSs in a manner analogous to the universal phosphopantetheinyltransferase Sfp. On-going studies in our laboratory are addressing each of these research directions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Bradon MacDonald for technical assistance with phylogenetic modeling. Dr. Philipp Wiemann for invaluable technical assistance obtaining LC/MS data. Dr. Jade Wang for *Bacillus subtilis* 168. Dr. Lawrence Shimkets for *Myxococcus xanthus* DK1622. We thank Dr. Matthew D. McMahon for the analysis of PvdL and PA2412.

Funding

This work was funded by the National Institutes of Health grant GM100346.

ABBREVIATIONS

A	adenylation
PCP	peptidyl carrier protein
C	condensation

NRPS	nonribosomal peptide synthetase
MLP	MbtH-like protein
ENT	enterobactin
iTOL	interactive tree of life
EDDHA	ethylenediamine- <i>N,N'</i> -bis(2-hydroxyphenylacetic acid)
TCEP	tris(2-carboxyethylphosphine)
IPTG	isopropyl β -D-1-thiogalactopyranoside
PAGE	polyacrylamide gel electrophoresis
TEV	tobacco etch virus
DTT	dithiothreitol
FA	formic acid
CAS	chrome azurol S
UHPLC	ultrahigh performance liquid chromatography
LC/MS	liquid chromatography/ mass spectrometry

References

1. Felnagle EA, Barkei JJ, Park H, Podevels AM, McMahon MD, Drott DW, Thomas MG. MbtH-like proteins as integral components of bacterial nonribosomal peptide synthetases. *Biochemistry*. 2010; 49:8815–8817. [PubMed: 20845982]
2. Boll B, Taubitz T, Heide L. Role of MbtH-like proteins in the adenylation of tyrosine during aminocoumarin and vancomycin biosynthesis. *J. Biol. Chem.* 2011; 286:36281–36290. [PubMed: 21890635]
3. Zhang W, Heemstra JR, Walsh CT, Imker HJ. Activation of the pacidamycin PacL adenylation domain by MbtH-like proteins. *Biochemistry*. 2010; 49:9946–9947. [PubMed: 20964365]
4. Quadri LE, Sello J, Keating TA, Weinreb PH, Walsh CT. Identification of a *Mycobacterium tuberculosis* gene cluster encoding the biosynthetic enzymes for assembly of the virulence-conferring siderophore mycobactin. *Chem. Biol.* 1998; 5:631–645. [PubMed: 9831524]
5. Ochsner UA, Wilderman PJ, Vasil AI, Vasil ML. GeneChip @ expression analysis of the iron starvation response in *Pseudomonas aeruginosa*: identification of novel pyoverdine biosynthesis genes. *Mol. Microbiol.* 2002; 45:1277–1287. [PubMed: 12207696]
6. Wolpert M, Gust B, Kammerer B, Heide L. Effects of deletions of mbtH-like genes on chlorobiocin biosynthesis in *Streptomyces coelicolor*. *Microbiology*. 2007; 153:1413–1423. [PubMed: 17464055]
7. Lautru S, Oves-Costales D, Pernodet JL, Challis GL. MbtH-like protein-mediated cross-talk between non-ribosomal peptide antibiotic and siderophore biosynthetic pathways in *Streptomyces coelicolor* M145. *Microbiology*. 2007; 153:1405–1412. [PubMed: 17464054]
8. Chen H, Hubbard BK, O'Connor SE, Walsh CT. Formation of β -hydroxy histidine in the biosynthesis of nikkomycin antibiotics. *Chem. Biol.* 2002; 9:103–112. [PubMed: 11841943]
9. Lauer B, Russwurm R, Schwarz W, Kálmánchelyi A, Bruntner C, Rosemeier A, Bormann C. Molecular characterization of co-transcribed genes from *Streptomyces tendae* Tü901 involved in the

biosynthesis of the peptidyl moiety and assembly of the peptidyl nucleoside antibiotic nikkomycin. *Mol. Gen. Genet.* 2001; 264:662–673. [PubMed: 11212921]

10. Pettis GS, McIntosh MA. Molecular characterization of the *Escherichia coli* enterobactin cistron *entF* and coupled expression of *entF* and the *fes* gene. *J. Bacteriol.* 1987; 169:4154–4162. [PubMed: 3040679]
11. Reichert J, Sakaitani M, Walsh CT. Characterization of EntF as a serine-activating enzyme. *Protein Sci.* 1992; 1:549–556. [PubMed: 1338974]
12. Gehring AM, Mori I, Walsh CT. Reconstitution and characterization of the *Escherichia coli* enterobactin synthetase. *Biochemistry.* 1998; 37:2648–2659. [PubMed: 9485415]
13. Miller BR, Drake EJ, Shi C, Aldrich CC, Gulick AM. Structures of a nonribosomal peptide synthetase module bound to MbtH-like proteins support a highly dynamic domain architecture. *J. Biol. Chem.* 2016; 291:22559–22571. [PubMed: 27597544]
14. McMahon MD, Rush JS, Thomas MG. Analyses of MbtB, MbtE, and MbtF suggest revisions to the mycobactin biosynthesis pathway in *Mycobacterium tuberculosis*. *J. Bacteriol.* 2012; 194:2809–2818. [PubMed: 22447909]
15. Herbst DA, Boll B, Zocher G, Stehle T, Heide L. Structural basis of the interaction of MbtH-like proteins, putative regulators of nonribosomal peptide biosynthesis, with adenylating enzymes. *J. Biol. Chem.* 2013; 288:1991–2003. [PubMed: 23192349]
16. Drake EJ, Miller BR, Shi C, Tarrasch JT, Sundlov JA, Leigh Allen C, Skinotis G, Aldrich CC, Gulick AM. Structures of two distinct conformations of holo-non-ribosomal peptide synthetases. *Nature.* 2016; 529:235–238. [PubMed: 26762461]
17. Tarry MJ, Haque AS, Bui KH, Schmeing TM. X-Ray crystallography and electron microscopy of synthetase proteins reveal a flexible architecture. *Structure.* 2017; 25:783–793. [PubMed: 28434915]
18. Konz D, Marahiel MA. How do peptide synthetases generate structural diversity? *Chem. Biol.* 1999; 6:R39–R48. [PubMed: 10021423]
19. Hamilton CM, Aldea M, Washburn BK, Babitzke P, Kushner SR. New method for generating deletion and gene replacements in *Escherichia coli*. *J. Bacteriol.* 1989; 171:4617–4622. [PubMed: 2548993]
20. Klock HE, Koesema EJ, Knuth MW, Lesley SA. Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts. *Proteins: Struct., Funct., Genet.* 2008; 71:982–994. [PubMed: 18004753]
21. Guzman L, Belin D, Carson MJ, Beckwith J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose pBAD promoter. *J. Bacteriol.* 1995; 177:4121–4130. [PubMed: 7608087]
22. Rocco CJ, Dennison KL, Klenchin Va, Rayment I, Escalante-Semerena JC. Construction and use of new cloning vectors for the rapid isolation of recombinant proteins from *Escherichia coli*. *Plasmid.* 2008; 59:231–237. [PubMed: 18295882]
23. Huang Y, Niu B, Gao Y, Fu L, Li W. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics.* 2010; 26:680–682. [PubMed: 20053844]
24. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in performance and usability article fast track. *Mol. Biol. Evol.* 2013; 30:772–780. [PubMed: 23329690]
25. Price MN, Dehal PS, Arkin AP. FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLoS One.* 2010; 5:e9490. [PubMed: 20224823]
26. Letunic I, Bork P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics.* 2007; 23:127–128. [PubMed: 17050570]
27. Chan, YA., Thomas MG. *Complex Enzymes in Microbial Natural Product Biosynthesis, Part B: Polyketides, Aminocoumarins and Carbohydrates.* 1. Elsevier Inc.; 2009. Formation and characterization of acyl carrier protein–linked polyketide synthase extender units; p. 143–163.
28. Schwyn B, Neilands JB. Universal chemical assay for the detection and determination of siderophores. *Anal. Biochem.* 1987; 160:47–56. [PubMed: 2952030]
29. Melamud E, Vastag L, Rabinowitz JD. Metabolomic analysis and visualization engine for LC-MS data. *Anal. Chem.* 2010; 82:9818–9826. [PubMed: 21049934]

30. Rondon MR, Ballering KS, Thomas MG. Identification and analysis of a siderophore biosynthetic gene cluster from *Agrobacterium tumefaciens* C58. *Microbiology*. 2004; 150:3857–3866. [PubMed: 15528670]
31. May J, Wendrich TM, Marahiel MA. The *dhb* operon of *Bacillus subtilis* encodes the biosynthetic template for the catecholic siderophore 2,3-dihydroxybenzoate-glycine-threonine trimeric ester bacillibactin. *J. Biol. Chem.* 2001; 276:7209–7217. [PubMed: 11112781]
32. Müller S, Willett JW, Bahr SM, Darnell CL, Hummels KR, Dong CK, Vlamakis HC, Kirby JR. Draft genome sequence of *Myxococcus xanthus* wild-type strain DZ2, a model organism for predation and development. *Genome Announc.* 2013; 1:e00217–13. [PubMed: 23661486]
33. Thomas MG, Chan YA, Ozanick SG. Deciphering tuberactinomycin biosynthesis: isolation, sequencing, and annotation of the viomycin biosynthetic gene cluster. *Antimicrob. Agents Chemother.* 2003; 47:2823–2830. [PubMed: 12936980]
34. Felnagle EA, Rondon MR, Berti AD, Crosby HA, Thomas MG. Identification of the biosynthetic gene cluster and an additional gene for resistance to the antituberculosis drug capreomycin. *Appl. Environ. Microbiol.* 2007; 73:4162–4170. [PubMed: 17496129]
35. Heemstra JR, Walsh CT, Sattely ES. Enzymatic tailoring of ornithine in the biosynthesis of the *Rhizobium* cyclic trihydroxamate siderophore vicibactin. *J. Am. Chem. Soc.* 2009; 131:15317–15329. [PubMed: 19778043]
36. Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, Rabinowitz JD. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol.* 2009; 5:593–599. [PubMed: 19561621]
37. Shaw-Reid CA, Kelleher NL, Losey HC, Gehring AM, Berg C, Walsh CT. Assembly line enzymology by multimoldular nonribosomal peptide synthetases: the thioesterase domain of *E. coli* EntF catalyzes both elongation and cyclolactonization. *Chem. Biol.* 1999; 6:385–400. [PubMed: 10375542]

Figure 1.

Schematic of ENT NRPS domain structure. Light blue circles represent NRPS domains, the red circle represents YbdZ, the MLP of the ENT system. Abbreviations: A, adenylation; PCP, peptidyl carrier protein; C, condensation; Te, thioesterase.

Figure 2.

Analysis of complementation for *ybdZ* by various noncognate MLPs. Growth curves for strains containing pBAD33 with No MLP, *cmnN* or *vioN* are superimposed along the baseline. Data shown are means of triplicate, independent samples with standard deviation shaded.

Figure 3.
(A) MLP amino acid sequence alignments. Residues implicated in protein–protein interactions from structural studies of EntF/MLP complexes (pink) and didomain (MLP-adenylation domain) SlgN1 (green) are highlighted. (B) Amino acid identity (upper tier) and amino acid similarity (lower tier) comparisons for the MLPs included in this study.

Figure 4.

Analysis of EntF solubility when co-overproduced with MLPs. Equal amounts of total protein (75 μ g) were loaded for each sample on a SDS-PAGE (10% polyacrylamide) gel and stained with Coomassie blue. This gel is representative of multiple experimental replicates.

Figure 5. Formation of [^{14}C]-L-Seryl-*S*-EntF after 40 min of incubation at RT. EntF was phosphopantetheinylated *in vitro* prior to incubation with [^{14}C]-L-Ser and MLPs. Data shown are the means with standard deviation from triplicate assays. Asterisk indicates statistically significant difference ($p = 0.0194$) using an unpaired *t* test. ANOVA statistical analysis also indicated a statistically relevant result.

Figure 6.

In vitro ENT biosynthesis by various EntF+MLP complexes monitored by CAS absorbance. Assays were done in triplicate and represented as a mean with standard deviation.

Figure 7.

(A) Comparison of ENT detection by LC/MS from *in vitro* reconstitution reactions containing no MLP, YbdZ, CmnN, or VioN. LC/MS was run in triplicate and the area reported shown is the mean with standard deviation from triplicate samples. Values indicate the fold decrease in ENT produced compared to samples containing YbdZ. Statistical analysis by ANOVA indicates that ENT production in the presence of YbdZ is statistically distinct from samples containing no MLP, CmnN, and VioN ($p = 0.0006$).

Figure 8. Summary of results for each noncognate MLP interaction with EntF. This heat map summarizes the results from the various assays scored from worse than EntF to equivalent to EntF+YbdZ. n/a = not applicable.

Table 1Kinetics for EntF Copurified or Reconstituted with MLPs^a

protein complex:	$K_d(\text{MLP})$ (nM)	$K_m(\text{L-Ser})$ (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\text{M}^{-1} \text{s}^{-1}$)
EntF/YbdZ		149 $\pm$ 58	0.95 $\pm$ 0.1	6375
EntF		1029 $\pm$ 171	0.84 $\pm$ 0.05	816
EntF+YbdZ	186 $\pm$ 69	116 $\pm$ 22	0.88 $\pm$ 0.04	7858
EntF+MbtH	54.6 $\pm$ 8	229 $\pm$ 31	0.74 $\pm$ 0.04	3231
EntF+PA2412	358 $\pm$ 95	397 $\pm$ 94	1.21 $\pm$ 0.06	3047
EntF+MXAN_3118	166 $\pm$ 29	303 $\pm$ 42	1.37 $\pm$ 0.08	4521
EntF+BSU31959	310 $\pm$ 69	310 $\pm$ 45	1.14 $\pm$ 0.07	3677
EntF+CmnN	150 $\pm$ 7	1051 $\pm$ 244	0.35 $\pm$ 0.02	333
EntF+VioN	200 $\pm$ 63	1317 $\pm$ 119	1.20 $\pm$ 0.04	911
EntF/Atu3678		299.7 $\pm$ 85	1.17 $\pm$ 0.11	3913

^aData shown in graphical form in Figures S7 and S8.