Crystal Structure of a Complete Nonribosomal Peptide Synthetase Module Bound to Auxilliary MbtH-like Proteins

Bradley R. Miller[1,2], **Eric J. Drake**[1,2], **Ce Shi**[3], **Courtney C. Aldrich**[3], **and Andrew M. Gulick**[1,2]

[1,2]Hauptman-Woodward Medical Research Institute, and Department of Structural Biology, University at Buffalo. Buffalo, NY 14203 USA
[3]Center for Drug Design and Department of Medicinal Chemistry, University of Minnesota, Minneapolis MN 55455 USA
Running Title: *Structure of EntF bound to MbtH-Like Proteins*

Address correspondence to Andrew M. Gulick, Hauptman-Woodward Institute, 700 Ellicott St. Buffalo, NY 14203-1102, Phone (716) 898-8619, FAX (716) 898-8660 email, gulick@hwi.buffalo.edu

Supporting Information

Figure S1. Identity of the insertion residue frequency in adenylation domains S2
Table S1. List of MLPs that have been biochemically analyzed for MLP dependence S3
Table S3. Sequences and domain boundaries of contructs used herein. S4
Supplementary References S5

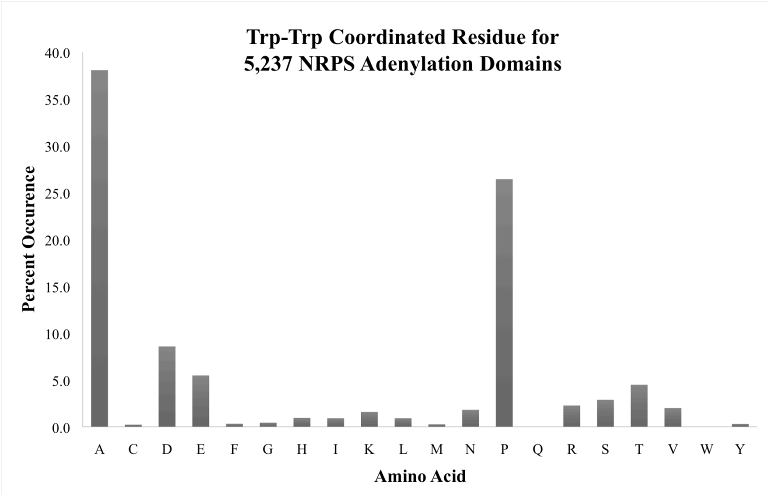

Figure S1. Identity of the "insertion residue" in the adenylation domains. A database of over 5200 adenylation domains sequences was analyzed and the residue that aligns with Ala826 of EntF was identified. The most common residues at this position are alanine and proline, which together are present in 66% of the sequences. The next two highest residues are glutamate and aspartate, which combined to fill 15% of the sequences. The other 16 amino acid residues account for the remaining 21% of the sequences at this site.

Table S1. MLPs that have been biochemically analyzed for MLP dependence

Adenylation Domain	MLP	Organism	Natural Product	Effect	Publication
CmnA-M1A	CmnN	*Saccharothrix mutabilis*	Capreomycin	Activity	(1)
CmnA-M2A	CmnN	*Saccharothrix mutabilis*	Capreomycin	Activity	(1)
CmnO	CmnN	*Saccharothrix mutabilis*	Capreomycin	Activity	(1)
NikP1	NikP1[3]	*Streptomyces tendae*	Nikkomycin	Activity	(2)
NocA-M1A	NocI	*Nocardia uniformis*	Nocardicin	Activity	(3)
NocB-M1A	NocI	*Nocardia uniformis*	Nocardicin	Activity	(3)
PacL	PacJ	*Streptomyces coeruleorubidus*	Pacidamycin 3	Activity	(4,5)
Pcza361.18	Orf1	*Amycolatopsis orientalis*	Vancomycin	Activity	(6)
PvdL-M2A	PA2412[1]	*Pseudomonas aeruginosa*	Pyoverdine	Activity	M.G. Thomas (*pers. comm.)*
SimH	SimY	*Streptomyces antibioticus*	Simocyclinone	Activity	(6)
SlgN1[1]	SlgN1[3]	*Streptomyces lydicus*	Streptolydigin	Activity	(7)
VbsS	VbsG	*Rhizobium etli*	Vicibactin	Activity	(8)
VioO	VioN	*Streptomyces* sp. ATCC-11861	Viomycin	Activity	(1)
EntF[1]	YbdZ[1]	*Escherichia coli*	Enterobactin	Enhanced	(1)
NovH	No MLP in Operon	*Streptomyces spheroides*	Novobiocin	Enhanced	(6)
Cgc18	SAMR0548	*Streptomyces ambofaciens*	Congocidine	Solubility	(9)
CloH	CloY	*Streptomyces roseochromogenes*	Clorobiocin	Solubility	(6)
GlbF	GlbE	*Burkholderia K481-B101*	Glidobactin A	Solubility	(10)
KtzH[2]	KtzJ	*Kutzneria sp 744*	Kutznerides	Solubility	(11)
TioK	TioT	*Micromonospora* sp.	Thiocoraline	Solubility	(12)
TioN[2]	TioT	*Micromonospora* sp.	Thiocoraline	Solubility	(13)
ABBFA_003403[1]	-	*Acinetobacter baumannii*	Unknown	Not Required[4]	(14)
CmnF	-	*Saccharothrix mutabilis capreolus*	Capreomycin	No Effect	(1)
CmnG	-	*Saccharothrix mutabilis capreolus*	Capreomycin	No Effect	(1)
EntE[1]	-	*Escherichia coli*	Enterobactin	No Effect	(1)
PA1221[1]	-	*Pseudomonas aeruginosa*	Unknown	No Effect	(15)
PheA[1]	-	*Bacillus brevis*	Gramicidin S	Not Required[4]	(16)

[1] Structure known

[2] Internal methylation domain within the adenylation domain

[3] MLP is attached with Adenylation domain as a single polypeptide

[4] No MLP in the genome of this organism has been identified

Table S2. Sequences and domain boundaries of contructs used herein.

Protein	Organism	GI	Accession	Construct	Start Residue	Stop Residue
EntF	*Escherichia coli* (K-12 subst W3110)	85674707	BAE76341.1	Wild Type	1	1293
NikP1	*Streptomyces tendae*	10639825	CAC11137.1	Wild Type	1	677
				Adenylation - PCP	76	677
				P449A	1	677
				MLP alone	1	75
PvdL	*Pseudomonas aeruginosa* PAO1	9948469	AAG05812.1	Module 2 Adenylation	1107	1631
YbdZ	*Escherichia coli* (K-12 substr W3110)	85674706	BAE76340.1	Wild Type	1	72
PA2412	*Pseudomonas aeruginosa* PAO1	15597608	NP_251102.1	Wild Type	1	72

Supplementary References

1. Felnagle, E. A., Barkei, J. J., Park, H., Podevels, A. M., McMahon, M. D., Drott, D. W., and Thomas, M. G. (2010) MbtH-like proteins as integral components of bacterial nonribosomal peptide synthetases. *Biochemistry* **49**, 8815-8817
2. Chen, H., Hubbard, B. K., O'Connor, S. E., and Walsh, C. T. (2002) Formation of beta-hydroxy histidine in the biosynthesis of nikkomycin antibiotics. *Chem Biol* **9**, 103-112
3. Davidsen, J. M., Bartley, D. M., and Townsend, C. A. (2013) Non-ribosomal propeptide precursor in nocardicin A biosynthesis predicted from adenylation domain specificity dependent on the MbtH family protein NocI. *J Am Chem Soc* **135**, 1749-1759
4. Zhang, W., Heemstra, J. R., Jr., Walsh, C. T., and Imker, H. J. (2010) Activation of the pacidamycin PacL adenylation domain by MbtH-like proteins. *Biochemistry* **49**, 9946-9947
5. Zhang, W., Ostash, B., and Walsh, C. T. (2010) Identification of the biosynthetic gene cluster for the pacidamycin group of peptidyl nucleoside antibiotics. *Proc Natl Acad Sci U S A* **107**, 16828-16833
6. Boll, B., Taubitz, T., and Heide, L. (2011) Role of MbtH-like proteins in the adenylation of tyrosine during aminocoumarin and vancomycin biosynthesis. *J Biol Chem* **286**, 36281-36290
7. Herbst, D. A., Boll, B., Zocher, G., Stehle, T., and Heide, L. (2013) Structural Basis of the Interaction of MbtH-like Proteins, Putative Regulators of Nonribosomal Peptide Biosynthesis, with Adenylating Enzymes. *Journal of Biological Chemistry* **288**, 1991-2003
8. Heemstra, J. R., Jr., Walsh, C. T., and Sattely, E. S. (2009) Enzymatic tailoring of ornithine in the biosynthesis of the Rhizobium cyclic trihydroxamate siderophore vicibactin. *J Am Chem Soc* **131**, 15317-15329
9. Al-Mestarihi, A. H., Garzan, A., Kim, J. M., and Garneau-Tsodikova, S. (2015) Enzymatic Evidence for a Revised Congocidine Biosynthetic Pathway. *Chembiochem* **16**, 1307-1313
10. Imker, H. J., Krahn, D., Clerc, J., Kaiser, M., and Walsh, C. T. (2010) N-acylation during glidobactin biosynthesis by the tridomain nonribosomal peptide synthetase module GlbF. *Chem Biol* **17**, 1077-1083
11. Zolova, O. E., and Garneau-Tsodikova, S. (2014) KtzJ-dependent serine activation and O-methylation by KtzH for kutznerides biosynthesis. *J Antibiot (Tokyo)* **67**, 59-64
12. Mady, A. S., Zolova, O. E., Millan, M. A., Villamizar, G., de la Calle, F., Lombo, F., and Garneau-Tsodikova, S. (2011) Characterization of TioQ, a type II thioesterase from the thiocoraline biosynthetic cluster. *Mol Biosyst* **7**, 1999-2011
13. Al-Mestarihi, A. H., Villamizar, G., Fernandez, J., Zolova, O. E., Lombo, F., and Garneau-Tsodikova, S. (2014) Adenylation and S-methylation of cysteine by the bifunctional enzyme TioN in thiocoraline biosynthesis. *J Am Chem Soc* **136**, 17350-17354
14. Drake, E. J., Miller, B. R., Shi, C., Tarrasch, J. T., Sundlov, J. A., Allen, C. L., Skiniotis, G., Aldrich, C. C., and Gulick, A. M. (2016) Structures of two distinct conformations of holo-non-ribosomal peptide synthetases. *Nature* **529**, 235-238
15. Mitchell, C. A., Shi, C., Aldrich, C. C., and Gulick, A. M. (2012) Structure of PA1221, a Nonribosomal Peptide Synthetase Containing Adenylation and Peptidyl Carrier Protein Domains. *Biochemistry* **51**, 3252-3263
16. Conti, E., Stachelhaus, T., Marahiel, M. A., and Brick, P. (1997) Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S. *Embo Journal* **16**, 4174-4183