



Published in final edited form as:

*Nat Chem Biol.* 2015 August ; 11(8): 551–552. doi:10.1038/nchembio.1862.

## Leading the Way to RiPPs

**A. James Link**

Departments of Chemical and Biological Engineering and Molecular Biology, Princeton University, Princeton, NJ 08544

A. James Link: ajlink@princeton.edu

### Abstract

Many peptide-based natural products require a leader peptide to reach their final modified form, but identifying general rules for leader peptide interactions have been stymied by the diversity of these molecules. Two papers reporting crystallographic and bioinformatic analysis of these systems now reveal a structurally-conserved domain that mediates leader peptide binding.

---

Ribosomally-synthesized and posttranslationally modified peptides (RiPPs) are a diverse family of natural products, many of which exhibit therapeutically-relevant activities.<sup>1</sup> RiPPs produced by bacteria run the gamut from small molecules such as microcin C7<sup>2</sup> and pyrroloquinoline quinone<sup>3</sup> to the proteusins,<sup>4</sup> molecules made up of >40 amino acids with dozens of posttranslational modifications. One feature that connects these otherwise disparate molecules is the presence of a peptide sequence, the leader peptide, which is proteolytically removed either during or after the maturation of the natural product.<sup>5</sup> These leader peptides, however, vary tremendously in both length and sequence composition (Figure 1A). Thus, understanding the role of these leader peptides in determining the final product has been difficult, and mostly tackled on a system-by-system basis. In this issue, two papers reverse that trend, using structural and bioinformatic approaches to provide general insights into the role of leader peptides in RiPP biosynthesis<sup>6,7</sup>.

All RiPPs start out as linear chains of amino acids, referred to as the precursor peptide. This is in contrast to natural products made by polyketide synthases (PKS) or nonribosomal peptide synthetases (NRPS), which are assembled one monomer at a time. The RiPP precursor peptide can be further broken down into the leader peptide, which is eventually cleaved off, and the core peptide, which becomes the natural product via the action of one or more maturation enzymes. While it has been assumed that the leader peptide is a substrate recognition element that directs the maturation of RiPPs in some way, only recently have crystal structures emerged that show the interaction of a RiPP precursor peptide with a maturation enzyme. The structure of the microcin C7 biosynthetic enzyme MccB bound to its precursor peptide substrate MccA has been determined.<sup>2</sup> More recently, a structure of the nisin precursor NisA leader peptide in complex with the lanthipeptide biosynthetic enzyme NisB was reported.<sup>8</sup> These structures provided the first glimpses of leader peptide interactions with maturation enzymes.

Koehnke *et al.* have added to this list by solving the structure of LynD, an ATP-dependent heterocyclase involved in cyanobactin biosynthesis, with bound leader peptide from an artificial cyanobactin precursor PatE'.<sup>6</sup> The LynD structure was solved with bound peptide

substrate as well as bound ATP and analogs. The structure of these complexes revealed that substrate recognition (via leader peptide binding) and the heterocyclase activity are located in two distinct domains of the protein. This modularity is reminiscent of the MccB and NisB structures. The functional form of LynD is a dimer in which the leader peptide binding domain of one chain is in close proximity to the catalytic domain of the other chain. Using these structural insights, Koehnke *et al.* also engineered a version of the heterocyclase in which the leader peptide is provided in *cis* via its fusion to the N-terminus of LynD.<sup>6</sup> This enzyme has the advantage of being able to process “leader peptide-free” cyanobactin precursors, reducing the size of the required substrate from >50 aa to as little as 12 aa.

Burkhart *et al.* realized that the leader peptide binding domains of LynD, MccB, and NisB were structurally similar (Figure 1B), even though sequence identity was low, prompting the authors to develop a bioinformatic method to identify leader peptide-binding domains in other RiPP gene clusters.<sup>7</sup> These authors show that such domains are widespread among many different classes of RiPP gene clusters, suggesting that the leader peptide binding domain is a defining feature of RiPP biosynthesis. With the observation of similar leader peptide-binding domains across three different RiPPs families, Burkhart *et al.* proposed the name RiPP precursor protein recognition element, or RRE, for such domains. Using the sequence alignment tool HHpred, the authors found that the NisB and LynD leader peptide binding domains exhibited structural similarity to the protein PqqD. This is yet another connection to RiPP biosynthesis as PqqD has recently been shown to bind the putative peptide precursor to pyrroloquinoline quinone, PqqA.<sup>3</sup> The RRE was found in a total of 11 different classes of RiPP gene clusters, over half of the currently known RiPP classes. The interaction between leader peptide and RRE was demonstrated using fluorescence polarization binding assays and mutagenesis for three additional classes of RiPPs: linear azoline-containing peptides, thiopeptides, and lasso peptides.

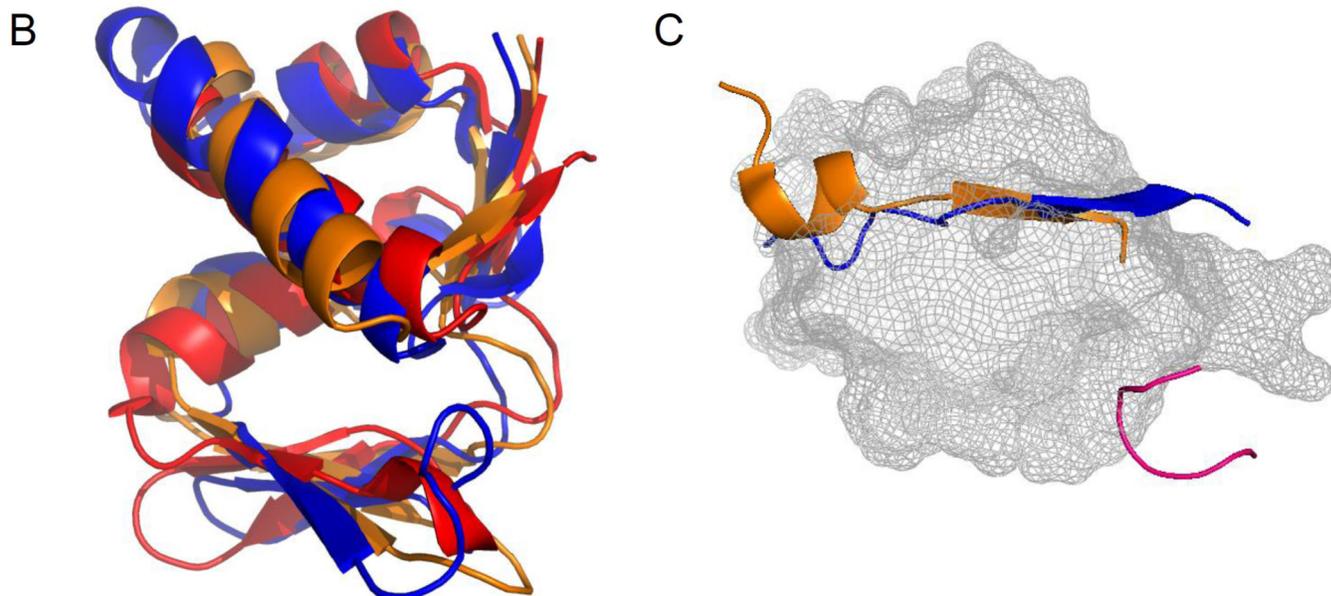
Collectively, these papers show that leader peptide recognition by PqqD-like domains is well-conserved across RiPP biosynthetic pathways. In many ways, this is an extraordinary observation: a single ~80 aa protein domain is involved in coordinating the biosynthesis of more than ten different classes of natural products, each with their own idiosyncrasies in biosynthetic logic. The only commonality amongst these natural products is their origin as linear chains of amino acids produced on the ribosome. RRE domains engage RiPP precursor peptides long and short, polar and hydrophobic. The RRE binds these peptides in different poses (Figure 1C), perhaps explaining how this simple domain can interact with so many different sequences. RREs from different RiPP classes have low sequence homology, so it is interesting to think about how such a domain evolved. Phylogenetic analyses of the RRE may reveal as yet unappreciated connections between different RiPP classes. Given their robustness in binding peptide substrates, it is also of interest to search for PqqD homologs in contexts beyond RiPP biosynthesis. Such searches may reveal further functions for what seems to be a highly capable protein.

## References

1. Arnison PG, et al. Natural Product Reports. 2013; 30:108–160. [PubMed: 23165928]
2. Regni CA, et al. EMBO Journal. 2009; 28:1953–1964. [PubMed: 19494832]

3. Latham JA, Iavarone AT, Barr I, Juthani PV, Klinman JP. *Journal of Biological Chemistry*. 2015
4. Freeman MF, et al. *Science*. 2012; 338:387–390. [PubMed: 22983711]
5. Oman TJ, van der Donk WA. *Nature Chemical Biology*. 2010; 6:9–18. [PubMed: 20016494]
6. Koehnke J, et al. *Nature Chemical Biology*. 2015
7. Burkhart BJ, Hudson GA, Dunbar KL, Mitchell DA. *Nature Chemical Biology*. 2015
8. Ortega MA, et al. *Nature*. 2015; 517:509–512. [PubMed: 25363770]

<b>A</b>	Microcin C7	MccA	<u>MRTGNAN</u>
	Lanthipeptide	NisA	<u>MSTKDFNLDLVSVSKKDSGASPR</u>
	Cyanobactin	PatE'	<u>MDKKNILPQQGQPVIRLTAGQLSSQLAELSEEAL</u>
	PQQ	PqqA	MWKKPAFIDLRLGLEVTLYISNR
	LAP	BalhA1	MEQKKILDIKLTETGKINYAHPDD
	Lasso peptide	StmA	MSAYEIPTLTRIGKFKDVTK



**Figure 1.**

Binding of diverse RiPP precursor peptides to their biosynthetic enzymes is mediated by a conserved domain related to the protein PqqD. A: Leader peptides of six different RiPP classes show diversity in leader peptide length and sequence composition. For microcin C7 and PQQ, the entire precursor protein is shown. The underlined segments in MccA, NisA, and PatE' are portions of the peptide resolved in crystal structures of the precursors with their cognate biosynthetic enzymes. PQQ: pyrroloquinoline quinone, LAP: linear azoline. B: Superposition of the leader peptide-binding domains from MccB (red, PDB file 3H9G), NisB (blue, PDB file 4WD9), and LynD (orange, PDB file 4V1T). C: Comparison of leader peptide binding poses. MccB, NisB, and LynD are aligned as in part B, but only LynD is shown as gray mesh. The precursor peptides MccA (magenta), NisA (blue), and PatE' (orange) are shown. The  $\beta$ -strands of NisA and PatE' align with the  $\beta$ -strands of their cognate PqqD homolog.