Published in final edited form as:

Chem Commun (Camb). 2018 August 21; 54(65): 9007–9010. doi:10.1039/c8cc04411b.

Elucidation of the Roles of Conserved Residues in the Biosynthesis of the Lasso Peptide Paeninodin

Julian D. Hegemann^{a,b}, Christopher J. Schwalen^{b,c}, Douglas A. Mitchell^{b,c,†}, and Wilfred A. van der Donk^{a,b,c,†}

^aHoward Hughes Medical Institute, University of Illinois at Urbana-Champaign, 600 S. Mathews Ave, Urbana, Illinois 61801, United States

^bDepartment of Chemistry, University of Illinois at Urbana-Champaign, 600 S. Mathews Ave, Urbana, Illinois 61801, United States

^cCarl R. Woese Institute for Genomic Biology, University of Illinois, Urbana, IL 61801

Abstract

Substrate binding assays, *in vitro* proteolytic processing assays, and heterologous lasso peptide production were used to investigate the roles of conserved precursor peptide residues during paeninodin maturation. Specifically, we delineate which residues are important for substrate recognition, proteolysis, and lasso peptide macrocyclization.

Graphical abstract

Lasso peptides are a class of ribosomally-synthesized and posttranslationally-modified peptides (RiPPs).^{1–3} Their defining feature is a unique lariat knot topology (Figure 1a): a macrolactam, formed between the N-terminal α -amino group and the side chain of an Asp or Glu residue at position 7–9, that is threaded by the linear C-terminal tail.^{1–4} This fold is maintained by bulky side chains positioned above and below the ring sterically trapping the tail inside. Lasso peptide biosynthesis usually requires three to four gene products (Figure 1);^{1–3, 5–10} the precursor peptide A consists of an N-terminal leader and a C-terminal core

[†]Corresponding authors: vddonk@illinois.edu and douglasm@illinois.edu.

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

Conflicts of interest

The authors declare no competing financial interest.

region. The leader region is required for enzymatic recognition, while the core region is converted into the mature lasso peptide. The B protein is a protease that separates the leader and core regions and the C protein catalyzes ATP-dependent macrolactam formation of the released core peptide. The B proteins often contain two domains, an N-terminal RiPP recognition element (RRE) of <100 amino acids and a larger C-terminal domain that contains the catalytic residues of the cysteine protease (~120–150 amino acids);^{5, 7, 10} in some systems these domains are separate proteins.^{6, 9, 11–16} Several conserved precursor peptide residues have been identified to be important for lasso peptide production, mostly by mutational analysis employing *Escherichia coli*-based heterologous production platforms and sequence alignments (see ESI Table S1 for an overview of the results of these reports).^{2, 4, 5, 8, 9, 12, 23–27}

These previous studies have revealed the following: (i) The position and identity (Asp vs. Glu) of the macrolactam acceptor residue cannot be varied.^{4, 8, 12, 24, 25, 27} (ii) All characterized lasso peptide biosynthetic machineries exhibit a strong residue identity preference for position 1 of the core peptide and substitutions are usually tolerated poorly or not at all.^{8, 12, 23–25, 27} (iii) Exchange of the universal Thr–2 residue always negatively affects lasso peptide processing, although substitutions with structurally similar residues are somewhat tolerated. The extent by which these exchanges affect lasso peptide production is dependent on the specific biosynthetic machinery.^{4, 8, 12, 23–27} (iv) In biosynthetic gene clusters with discretely encoded RRE proteins and proteases, a YxxP motif is strongly conserved in the N-terminal part of the leader peptide.^{6, 9, 11–14, 16} While not investigated experimentally thus far, this strong correlation suggests a role for this motif as putative substrate recognition sequence. Conversely, this motif is absent in precursors from clusters with a fused RRE-protease protein.^{9, 23}

Despite the bioinformatic and experimentally demonstrated importance of the aforementioned residues, the step(s) for which each of these residues are needed is unresolved (with exception of the ring forming Asp/Glu residues, which need to be activated in an ATP-dependent manner by direct action of the C protein⁸). In this work, we employed the recently described paeninodin gene cluster¹² (Figure 1c) to study the role of these residues by using a combination of heterologous expression, *in vitro* proteolytic processing, and fluorescence polarization (FP) binding assays. In previous studies, both the *in vitro* binding of the paeninodin precursor PadeA to the RRE protein PadeB1 and cleavage of PadeA by the leader peptidase PadeB2 was observed.⁵ While PadeC is intractable due to solubility issues, a heterologous *E. coli*-based production system of paeninodin is available,^{5, 12} which allowed us to assay variants for lasso peptide production. In this study, we investigated the roles of Ala1, Thr–2, Pro–12, and Tyr–15 (Figure 1d).

After co-expression of PadeA or a variant thereof with PadeB1, PadeB2, and PadeC in M9 minimal medium, the cell pellets were extracted with MeOH and subsequently analyzed by high performance liquid chromatography – mass spectrometry (HPLC-MS). UV absorbance peaks corresponding to the produced lasso peptides were integrated for quantification (Figure 2a). This experiment showed that exchange of each of the chosen residues negatively affects lasso peptide production. These effects are moderate for T–2A and P–12A and more pronounced for A1G, Y–15A, and the P–12A/Y–15A double substitution. The results for the

T-2A and A1G variants are in agreement with previous studies,¹² while the P-12A, Y-15A and P-12A/Y-15A variants provide evidence for the predicted importance of the YxxP motif. After determining that these residues play a role in lasso peptide maturation, we generated an expression plasmid allowing isolation of His₆-MBP-PadeA (MBP = maltose-binding-protein), featuring a tobacco etch virus protease (TEV) cleavage site for release of the PadeA precursor (leaving an N-terminal Ser in the process). We also changed the N-terminal Met of PadeA to a Cys residue, which allows selective modification of this position. Subsequent purification of wild-type (WT) PadeA and variants thereof was accomplished via HPLC. These peptides were then used for *in vitro* cleavage assays. In these assays, the RRE protein PadeB1 is needed for recognition of PadeA and mediating the interaction with the leader peptidase PadeB2.⁵ Whereas the A1G and P-12A variants were processed almost as efficiently as the WT substrate, substitutions of Thr-2 and Tyr-15 strongly decreased the extent of cleavage (Figure 2b).

Finally, we assessed if these substitutions affect recognition of the precursor peptide by the RRE. First, the Cys residue introduced at the N-terminus of PadeA was labeled with fluorescein-5-maleimide. The resulting fluorescent peptide (Fl-PadeA) was purified and used to assess binding with MBP-tagged PadeB1 via fluorescence polarization. The MBP tag was added to the RRE protein to enhance apparent FP upon Fl-PadeA binding due to its significantly higher molecular weight relative to the substrate⁷ (M(His₆-PadeB1) = 12.5 kDa versus M(His₆-MBP-PadeB1) = 56.0 kDa) while still allowing both PadeA recognition and interaction with His₆-MBP-PadeB2 (ESI Figure S2). Under the assay conditions, a dissociation constant K_d of 69 ± 2 nM was determined (ESI Figure S3). To compare the binding affinities of the PadeA variants, an FP competition assay was established (ESI Figure 4). The observed inhibition constant K_i for WT PadeA ($K_i = 137 \pm 4$ nM) was comparable to the PadeA variants A1G ($K_i = 127 \pm 5$ nM) and T-2A ($K_i = 152 \pm 6$ nM), whereas the K_i of the P-12A variant ($K_i = 476 \pm 28$ nM) was increased more than 3-fold. A more than 8-fold higher K_i was observed for the Y-15A variant ($K_i = 1176 \pm 74$ nM) and the P-12A/Y-15A double exchange ($K_i = 1344 \pm 56$ nM) led to an almost 10-fold increased K_i . These data clearly show that only substitutions in the YxxP motif affect RRE binding, while exchanges closer to or at the start of the core peptide do not adversely affect RRE binding. To corroborate this observation, an additional peptide was generated that consisted of residues -19 to -7 of the leader region of PadeA. The K_i obtained with this PadeA(-19 to -7) peptide ($K_i = 119 \pm 4$ nM) was again comparable to the K_i of the full length WT peptide, showing that this segment of the precursor peptide contained all residues important for RRE recognition.

A side by side comparison of all data presented in this study (Table 1) reveals the importance of each conserved residue in the various steps of lasso peptide maturation. Exchange of Ala1 strongly decreases lasso peptide production *in vivo* (of all residues investigated, substitution of Ala1 affected lasso production most severely), but shows little to no effect on recognition by the RRE or cleavage by PadeB2. Although our evidence is indirect, we hypothesize that position 1 of the core region is likely involved in substrate recognition by the lasso macrocyclase PadeC. Conversely, exchange of Thr-2 only modestly affects overall lasso peptide production in *E. coli* and does not affect recognition by PadeB1. At the same time, this exchange drastically decreases the rate of PadeB2-mediated leader

peptide proteolysis. On first glance, it is surprising to see a much stronger effect on the proteolysis rate than lasso peptide production *in vivo*. Perhaps the tight binding to PadeB1 protects the precursor peptide against nonspecific proteolysis in the cytosol. Finally, the P-12A, Y-15A and P-12A/Y-15A exchanges of the conserved YxxP motif show significant effects in all assays performed. Substitution of Pro-12 with Ala moderately affects *in vivo* lasso production, *in vitro* cleavage and binding to the RRE. A much stronger effect is observed when Tyr-15 is replaced with Ala. Moreover, the P-12A/Y-15A double substitution shows an effect that seems to be the additive of the single residue exchanges. This behavior suggests that reduced affinity for the RRE reduces both the *in vitro* cleavage rate and the overall amount of lasso peptide produced *in vivo*. This notion is supported by previous investigations, where PadeB1 variants with lower affinity for PadeA binding caused lower proteolysis rates when used in assays with His₆-MBP-PadeB2.⁵ Hence, our results clearly show that the role of the YxxP motif lies in enabling the interaction between precursor peptide and RRE protein. This conclusion is corroborated by the observation that PadeA(-19 to -7) shows binding to PadeB1 comparable to WT PadeA, indicating that the absence of Thr-2 and Ala1 did not affect recognition by the RRE. The data also suggest that other elements contribute to affinity to the RRE, as the P-12A/Y-15A double exchange does not completely abolish binding by PadeB1, which may be interactions with other residues in this region and/or the peptide backbone. In a recent study, crosslinking experiments were carried out between the LarB1 RRE protein and the LarA precursor peptide that suggested an additional interaction of LarB1 and the core peptide region.¹⁰ While our experiments do not rule out that such an interaction exists, our data suggest that this interaction does not contribute to the overall PadeA-PadeB1 affinity.

Based on our findings, we propose a more detailed biosynthesis mechanism for lasso peptides from gene clusters with discrete RRE and protease encoding genes (Figure 3). In general, the observed effects on lasso peptide production of the variants tested in this study are in line with previous studies of lasso peptide maturation by mutational analysis of *E. coli*-based heterologous production systems (ESI Table S1).^{4, 12, 23–27} They also agree with qualitative *in vitro* experiments performed with the microcin J25 biosynthetic enzymes.⁸ This prior study is the only published example of *in vitro* data with an RRE-protease fusion protein (McjB). No cleavage of a T-2A precursor peptide variant was observed and while small amounts of a G1A variant were cleaved, no core peptide cyclization was detected.⁸ However, these assays were carried out with the entire biosynthetic machinery (McjB and McjC) present as no activity was observed when one of the enzymes was omitted.⁸ Therefore, these experiments cannot distinguish lack of precursor binding from absence of enzymatic activity. Conversely, our experimental design allowed us to study the effect of each substitution on every step of the maturation process. Future studies may indicate, whether this separation of function between binding and catalysis also holds true for biosynthetic machineries with RRE-protease fusions, which lack the YxxP motif in their substrates.^{9, 23}

Our experiments additionally revealed that the leader peptidase PadeB2 can tolerate exchanges at position 1, but is much more sensitive to the replacement of Thr-2. At the same time, the importance of the conserved YxxP motif was demonstrated for binding to a discrete lasso peptide RRE protein and only the N-terminal part of PadeA is needed for

PadeB1 recognition. These findings suggest that each of the conserved residues and motifs in PadeA are recognized by a distinct part of the maturation machinery and that there is likely no segment that is recognized by more than one protein. An extensive genome-mining study⁹ for lasso peptides recently demonstrated that clusters with discrete RRE proteins are prevalent in actinobacteria and firmicutes, and our findings are likely applicable to these biosynthetic machineries. Conversely, the RRE-protease fusion is more common in proteobacteria.⁹ Our findings potentially also provide insights into enzymatic recognition in these systems, as it was shown that in some precursors Gly-8, which is also often found in actinobacterial precursors,⁹ is important for efficient lasso peptide processing.²³ Taken together with our observation that PadeA(-19 to -7) shows WT level binding to PadeB1, we hypothesize that in general for lasso peptide precursors, the recognition sequence might be located in this segment.

We note that while point mutations of conserved residues in lasso peptide precursors might drastically decrease efficiency of either binding, proteolysis or macrocyclization, they do not necessarily completely abolish any of these activities, thereby emphasizing the complex interplay of proteins and precursor peptides during lasso peptide maturation.

In conclusion, the combination of *in vivo* production, *in vitro* proteolytic processing and *in vitro* binding experiments enabled elucidation of the roles of each conserved residue in PadeA during biosynthesis of the paeninodin lasso peptide and thus to better understand how nature produces these interesting and unique natural products.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by grants from the Howard Hughes Medical Institute (to W.A.V.), the David and Lucile Packard Fellowship for Science and Engineering (to D.A.M.), the Deutsche Forschungsgemeinschaft (DFG Research Fellowship 309199717 to J.D.H.), and the Chemistry-Biology Interface Training Program (GM070421 to C.J.S.). A Bruker UltrafleXtreme MALDI TOF/TOF mass spectrometer was bought in part with a grant from the National Institutes of Health (S10 RR027109 A). We thank Prof. Mohamed Marahiel for providing some of the plasmids used in this study.

Notes and references

1. Maksimov MO, Pan SJ, James Link A. Nat Prod Rep. 2012; 29:996. [PubMed: 22833149]
2. Hegemann JD, Zimmermann M, Xie X, Marahiel MA. Acc Chem Res. 2015; 48:1909. [PubMed: 26079760]
3. Hegemann JD, Marahiel MA. Cyclic Peptides: From Bioorganic Synthesis to Applications. Koehnke J, Naismith J, van der Donk WA, editors The Royal Society of Chemistry; 2018. 206
4. Hegemann JD, Zimmermann M, Zhu S, Steuber H, Harms K, Xie X, Marahiel MA. Angew Chem Int Ed. 2014; 53:2230.
5. Zhu S, Fage CD, Hegemann JD, Mielcarek A, Yan D, Linne U, Marahiel MA. Sci Rep. 2016; 6:35604. [PubMed: 27752134]
6. Li Y, Ducasse R, Zirah S, Blond A, Goulard C, Lescop E, Giraud C, Hartke A, Guittet E, Pernodet JL, Rebuffat S. ACS Chem Biol. 2015; 10:2641. [PubMed: 26343290]

7. Burkhart BJ, Hudson GA, Dunbar KL, Mitchell DA. Nat Chem Biol. 2015; 11:564. [PubMed: 26167873]
8. Yan KP, Li Y, Zirah S, Goulard C, Knappe TA, Marahiel MA, Rebuffat S. Chembiochem. 2012; 13:1046. [PubMed: 22488892]
9. Tietz JI, Schwalen CJ, Patel PS, Maxson T, Blair PM, Tai HC, Zakai UI, Mitchell DA. Nat Chem Biol. 2017; 13:470. [PubMed: 28244986]
10. Cheung WL, Chen MY, Maksimov MO, Link AJ. ACS Cent Sci. 2016; 2:702. [PubMed: 27800552]
11. Zong C, Cheung-Lee WL, Elashal HE, Raj M, Link AJ. Chem Commun. 2018; 54:1339.
12. Zhu S, Hegemann JD, Fage CD, Zimmermann M, Xie X, Linne U, Marahiel MA. J Biol Chem. 2016; 291:13662. [PubMed: 27151214]
13. Metelev M, Tietz JI, Melby JO, Blair PM, Zhu L, Livnat I, Severinov K, Mitchell DA. Chem Biol. 2015; 22:241. [PubMed: 25601074]
14. Elsayed SS, Trusch F, Deng H, Raab A, Prokes I, Busarakam K, Asenjo JA, Andrews BA, van West P, Bull AT, Goodfellow M, Yi Y, Ebel R, Jaspars M, Rateb ME. J Org Chem. 2015; 80:10252. [PubMed: 26402731]
15. Gavrish E, Sit CS, Cao S, Kandror O, Spoering A, Peoples A, Ling L, Fetterman A, Hughes D, Bissell A, Torrey H, Akopian T, Mueller A, Epstein S, Goldberg A, Clardy J, Lewis K. Chem Biol. 2014; 21:509. [PubMed: 24684906]
16. Inokoshi J, Matsuhama M, Miyake M, Ikeda H, Tomoda H. Appl Microbiol Biotechnol. 2012; 95:451. [PubMed: 22388571]
17. Maksimov MO, Link AJ. J Am Chem Soc. 2013; 135:12038. [PubMed: 23862624]
18. Hegemann JD, Zimmermann M, Zhu S, Klug D, Marahiel MA. Biopolymers. 2013; 100:527–542. [PubMed: 23897438]
19. Maksimov MO, Pelczer I, Link AJ. Proc Natl Acad Sci. 2012; 109:15223. [PubMed: 22949633]
20. Fage CD, Hegemann JD, Nebel AJ, Steinbach RM, Zhu S, Linne U, Harms K, Bange G, Marahiel MA. Angew Chem Int Ed. 2016; 55:12717.
21. Maksimov MO, Koos JD, Zong C, Lisko B, Link AJ. J Biol Chem. 2015; 290:30806. [PubMed: 26534965]
22. Choudhury HG, Tong Z, Mathavan I, Li Y, Iwata S, Zirah S, Rebuffat S, van Veen HW, Beis K. Proc Natl Acad Sci. 2014; 111:9145. [PubMed: 24920594]
23. Zimmermann M, Hegemann JD, Xie X, Marahiel MA. Chem Sci. 2014; 5:4032.
24. Zimmermann M, Hegemann JD, Xie X, Marahiel MA. Chem Biol. 2013; 20:558. [PubMed: 23601645]
25. Hegemann JD, Zimmermann M, Xie X, Marahiel MA. J Am Chem Soc. 2013; 135:210. [PubMed: 23214991]
26. Pan SJ, Rajniak J, Maksimov MO, Link AJ. Chem Commun. 2012; 48:1880.
27. Knappe TA, Linne U, Robbel L, Marahiel MA. Chem Biol. 2009; 16:1290. [PubMed: 20064439]

Figure 1.

(a) Schematic depiction of the lasso peptide topology. (b) The biosynthetic gene cluster of the proteobacterial lasso peptide capistrin as example of a cluster with a fused RRE-protease protein and (c) the paeninodin biosynthetic gene cluster. In proximity to the open reading frames essential for lasso peptide biosynthesis, genes encoding ABC-transporters, tailoring enzymes and/or pairs of TonB-dependent receptors and dedicated lasso peptide isopeptidases are often found.^{9, 17–22} (d) Sequence of the PadeA precursor peptide. The macrolactam in paeninodin is formed between Ala1 and Asp9. Leader and core peptide regions are indicated and residues investigated in this study are marked by colors that will be used throughout the manuscript.

Figure 2.

(a) Relative heterologous production levels of paeninodin using WT PadeA and variants thereof. (b) *In vitro* cleavage of PadeA (or variants thereof) in the presence of His₆-PadeB1 and His₆-MBP-PadeB2. The observed proteolytic turnover of PadeA and each of its variants is shown. For representative LC traces for each cleavage reaction see ESI Figure S1. All experiments were performed in triplicates with the error bars shown the standard deviation of the mean.

Figure 3.

(a) Graphical summary highlighting the parts of PadeA that are needed for each step of lasso peptide biosynthesis. (b) Lasso peptide maturation mechanism for paeninodin and related lasso peptides as suggested by our findings. First, the PadeA precursor peptide is recognized and bound by PadeB1. Then, PadeB1 mediates interaction with the leader peptidase PadeB2, which in turn cleaves the precursor peptide. The free core peptide then interacts with PadeC and is cyclized into the mature lasso peptide in an ATP-dependent manner. It is still unknown, if any direct interaction between the PadeB1/PadeB2 complex and PadeC is needed during the maturation process, if PadeC already interacts with the full length precursor prior to cleavage and how PadeC interacts with the core peptide to accomplish the threaded lasso peptide fold. Our experiments do not rule out the possibility that the enzymatic machinery forms a complex under physiological conditions and that complex formation is independent of the presence of substrate peptide.

Table 1

Overview of the data obtained in this study from the heterologous production system, the *in vitro* cleavage and the *in vitro* binding assays. Experiments were performed in triplicates and errors shown are standard deviation of the mean.

PadeA variant	<i>E. coli</i> production level	<i>in vitro</i> leader peptide cleavage ^a	K _i PadeA:MBP-PadeB1
WT	100% ± 32%	56% ± 4%	137 ± 4 nM
A1G	13% ± 4%	43% ± 4%	127 ± 5 nM
T-2A	36% ± 13%	6% ± 1%	152 ± 6 nM
P-12A	46% ± 26%	39% ± 4%	476 ± 28 nM
Y-15A	25% ± 5%	15% ± 1%	1176 ± 74 nM
P-12A/Y-15A	24% ± 10%	7% ± 1%	1344 ± 56 nM
-19 to -7	<i>_b</i>	<i>_b</i>	120 ± 4 nM

^aSee ESI for details;

^bNot applicable.