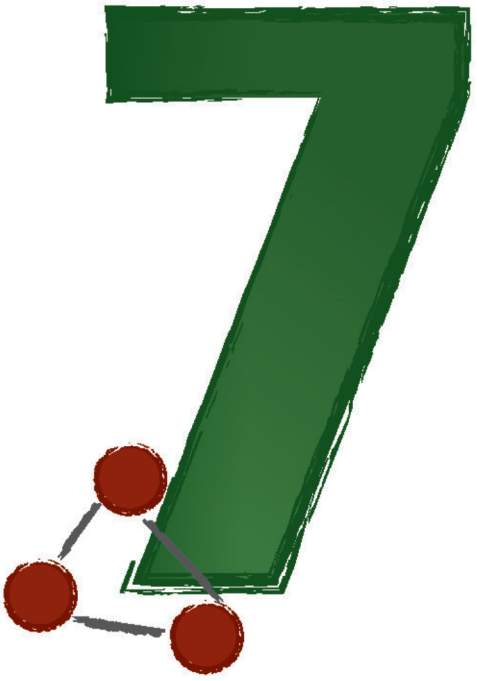
Universiteit Leiden
The Netherlands

Expanding the chemical space of antibiotics produced by Paenibacillus and Streptomyces

Machushynets, N.V.



Chapter 7

Discovery and derivatization of tridecaptin antibiotics with altered host specificity and enhanced bioactivity

Nataliia V. Machushynets, Karol Al Ayed, Barbara R. Terlouw, Chao Du, Ned P. Buijs, Joost Willemse, Somayah S. Elsayed, Julian Schill, Vincent Trebosc, Michel Pieren, Francesca M. Alexander, Stephen A. Cochrane, Mark R. Liles, Marnix H. Medema, Nathaniel I. Martin and Gilles P. van Wezel

This chapter was published as:
Machushynets N.V., Ayed K.A., Terlouw B.R., Du C., Buijs N.P., Willemse J., Elsayed S.S., Schill J., Trebosc V., Pieren V., Alexander F.M., Cochrane S.A., Liles M.R., Medema M.H, Martin N.I. & van Wezel G.P.(2024). Discovery and derivatization of tridecaptin antibiotics with altered host specificity and enhanced bioactivity. *ACS Chem Bio.*

Abstract

The prevalence of multidrug-resistant (MDR) pathogens combined with a decline in antibiotic discovery presents a major challenge for health care. To refill the discovery pipelines, we need to find new ways to uncover new chemical entities. Here, we report the global genome mining-guided discovery of new lipopeptide antibiotics tridecaptin A_5 and tridecaptin D, which exhibit unusual bioactivities within their class. The change in antibacterial spectrum of Oct-TriA_5 was explained solely by a Phe to Trp substitution as compared to Oct-TriA_1, while Oct-TriD contained 6 substitutions. Metabolomics analysis of producer *Paenibacillus* sp. JJ-21 validated the predicted amino acid sequence of tridecaptin A_5. Screening of tridecaptin analogs substituted at position 9 identified Oct-His9 as a potent congener with exceptional efficacy against *Pseudomonas aeruginosa* and reduced hemolytic and cytotoxic properties. Our work highlights the promise of tridecaptin analogs to combat MDR pathogens.

Introduction

The overuse of antibiotics over the past many decades has contributed to the rapid emergence and spread of antimicrobial resistance. Combined with the decline in the number of new clinically approved antibacterial drugs, infections caused by resistant bacteria are frequently difficult to treat (Wright, 2015). The ever-increasing prevalence of multidrug-resistant (MDR) bacteria is recognized by the World Health Organization (WHO) as an imminent threat to human health, particularly the Gram-negative critical priority pathogens *Acinetobacter baumannii, Pseudomonas aeruginosa* and *Enterobacteriaceae* (WHO, 2017). Thus, there is an urgent need for novel antimicrobial agents, and the challenge lies in finding the undiscovered gems that may form the basis for our future medicines (Newman & Cragg, 2007, Davies & Davies, 2010). Traditional high-throughput screening is becoming less attractive due to the issue of dereplication, rediscovering compounds that had already been identified before (Kolter & van Wezel, 2016, Lewis, 2013). However, next-generation sequencing has revealed that microbial genomes still harbor a huge underexplored biosynthetic potential of bacteria (Medema *et al.*, 2015, Gavriilidou *et al.*, 2022).

Polymyxins are a promising class of antibiotics that are used as a last line of therapy against MDR pathogens (Chiu *et al.*, 2022), and are primarily produced by *Paenibacillus* species. With resistance against polymyxins inevitably increasing, we urgently need to search for potent alternatives. *Bacillus* and *Paenibacillus* produce a range of other classes of lipopeptides that act on MDR Gram-negative pathogens, including brevicidines, laterocidines, relacidines, paenibacterins and tridecaptins (Al Ayed *et al.*, 2023, Cochrane & Vederas, 2016, Olishevska *et al.*, 2019, Raaijmakers *et al.*, 2010, Bann *et al.*, 2021). Of these, tridecaptins stand out as they possess a mechanism of action that is distinct from the aforementioned lipopeptides. In addition to interacting with lipopolysaccharides (LPS), tridecaptins also bind to Gram-negative lipid II and in doing so cause the disruption of the proton motive force (Cochrane *et al.*, 2015, Buijs *et al.*, 2023, Medeiros-Silva *et al.*, 2019, Cochrane *et al.*, 2016). This unique dual mechanism of action reduces the likelihood of cross-resistance (Jangra *et al.*, 2020). Since the discovery of tridecaptin A in 1978, multiple tridecaptin variants with modifications at either the N-terminal fatty acid moiety or in the amino acid sequence have been reported. Tridecaptin A (Shoji *et al.*, 1978), tridecaptin B (Cochrane *et al.*, 2015), tridecaptin C (Shoji J. *et al.*, 1979) and tridecaptin M (Jangra *et al.*, 2019b) are predominantly active against Gram-negative bacteria, while the recently discovered tridecaptin G (da Costa *et al.*, 2022) has a broad-spectrum activity. Studies on natural tridecaptins include research on their biosynthesis, structure-activity relationships, their mechanism of action and their synergy with other antibiotics (Bann *et al.*, 2021). Structure-activity studies formed the basis for drug development studies to enhance their potency (Ballantine *et al.*, 2019, Cochrane *et al.*, 2014), expand their bioactivity spectrum (Wang *et al.*, 2022a), or increase their stability (Ballantine *et al.*, 2018).

In this regard, approaches that combine genome mining with variation in culturing conditions has proven to be a valuable way to achieve differential synthesis of NPs, followed by the metabolic profiling-based identification of the bioactivity of interest (O'Leary *et al.*, 2016). However, a major challenge is to find the appropriate chemical triggers or ecological cues to elicit the production of cryptic antibiotics (Zhu *et al.*, 2014a, Xu *et al.*, 2019, van der Meij *et al.*, 2017, van Bergeijk *et al.*, 2020). An alternative way to investigate the potential of novel scaffolds predicted by genome mining is via organic synthesis or chemoenzymatic total synthesis (Wang *et al.*, 2022a, Wang *et al.*, 2022b, Wang *et al.*, 2022c). Genome mining tools such as antiSMASH (Blin *et al.*, 2023) allow interrogation of microbial genomes for the presence of biosynthetic gene clusters (BGCs) that specify the biosynthesis of natural products and, subsequently, predict the types of compounds that are derived from them, such as polyketides (Ray & Moore, 2016), ribosomally synthesized and post-translationally modified peptides (RiPPs) (Montalban-Lopez *et al.*, 2021), terpenes (Avalos *et al.*, 2022) or nonribosomal peptides (NRPs) (Süssmuth & Mainz, 2017). In the case of nonribosomal peptide synthetases (NRPS), NRPS modules and the amino acids predicted to be incorporated by the A domain in each module can be predicted using NRPS-specific prediction algorithms, such as NRPSpredictor2 (Röttig *et al.*, 2011) or SANDPUMA (Chevrette *et al.*, 2017) or machine learning techniques trained on a set of A domains with known specificities (Mongia *et al.*, 2023). With the power of modern synthetic organic chemistry and the increasing accuracy of natural product structure prediction algorithms, it is increasingly possible to chemically synthesize new bioactive molecules based on BGC sequences (Wang *et al.*, 2022b).

Here, we report a global genome-mining approach that led to the discovery of novel tridecaptins. Bioinformatics analysis of 785 bacterial genomes identified novel tridecaptins tridecaptin A_5 and tridecaptin D. Unexpectedly, two synthetic analogs Oct-TriD and Oct-Tri$A_{5,}$ efficiently killed both Gram-negative and Gram-positive bacteria. The increased bioactivity spectrum of tridecaptin Oct-TriA_5 compared to Oct-TriA_1 also correlated with hemolytic and cytotoxic activity and could be attributed to a single amino acid substitution at position 9. In contrast to Oct-TriA_1, the broadened antimicrobial spectrum of Oct-TriA_5 and Oct-TriD is caused by its increased membrane disruptive capacity on Gram-positive pathogens. Subsequent screening of other amino acids at position 9 led to the discovery of the potent tridecaptin analog Oct-TriHis9 with reduced hemolytic and cytotoxic properties and potent activity against *P. aeruginosa*.

Results and Discussion

Large-scale network analysis and bioinformatic prediction of novel tridecaptin BGCs

Paenibacillus sp. JJ-21 is a gifted natural product producer isolated from corn rhizosphere and identified as an antibacterial metabolite producer based on its antagonistic potential against polymyxin-resistant *Escherichia coli* ATCC 25922 harboring plasmid-mediated *mcr-1* resistance gene. The *Paenibacillus* sp. JJ-21 genome was sequenced using the PacBio platform. Assembly of the PacBio reads with Falcon (version 1.8.1) (Wang *et al.*, 2016) resulted in a single contig of 6.2 Mb (GenBank accession number: CP132974). The *Paenibacillus* sp. JJ-21 genome was analyzed using antiSMASH 6.0.1 (Blin *et al.*, 2021) to obtain an overview of the predicted BGCs. AntiSMASH predicted a total of 18 BGCs, including three for NRPSs (Figure 1A). These NRPS BGCs were predicted to encode the biosynthesis of fusaricidin, tridecaptin and polymyxin BGCs, respectively, based on their high similarity (identical gene order and >75% nucleotide sequence identity) to characterized clusters (listed in MIBiG) (Terlouw *et al.*, 2023). An in silico analysis of the A domain substrate specificity of these BGCs was conducted with the software tool PARAS (v0.0.4, available at https://github.com/BTheDragonMaster/paras) to predict the amino acid composition of peptide scaffolds (Table S1). Surprisingly, the amino acid sequence predicted from the tridecaptin BGC differed from the well-known tridecaptin A_1 and contained Trp instead of Phe in position 9 (Table S2).

To gain insight into the chemical space of NRPSs and specifically tridecaptin BGCs, we bioinformatically analyzed 785 complete genomes from *Paenibacillus* spp. for BGCs encoding NRPSs. The genome of *Paenibacillus* sp. JJ-21 was included as a reference genome in the data set. To visualize the diversity, distribution, and NRPS novelty, a BiG-SCAPE (Navarro-Muñoz *et al.*, 2020) sequence similarity network (SSN) was then constructed, which consisted of 4,367 NRPS BGCs forming 1609 gene cluster families (GCFs) (Figure 1B).

Genetic variation of BGCs within GCFs is often directly associated with structural differences between their molecular products, and even small chemical variations can lead to different biological activities (Medema *et al.*, 2014). We also visualized the modular architecture of NRPS assembly lines, revealing the order and the number of modules for each NRPS synthase that, in turn, helped to determine the class of the molecules they encode. The BiG-SCAPE sequence similarity network highlighted various known classes of NRPS BGCs, such as polymyxins (three NRPSs, 10 modules), tridecaptins (two NRPSs, 13 modules), paenilipoheptins (three NRPSs, seven modules), fusaricidins (one NRPS, six modules), paenibacterins (four NRPSs, 13 modules), octapeptins (three NRPS, eight modules), bacitracins (three NRPS, 12 modules) and cilagicins (three NRPS, 12 modules), but about twenty percent of the detected BGCs could not be linked to known compounds (Terlouw *et al.*, 2023).

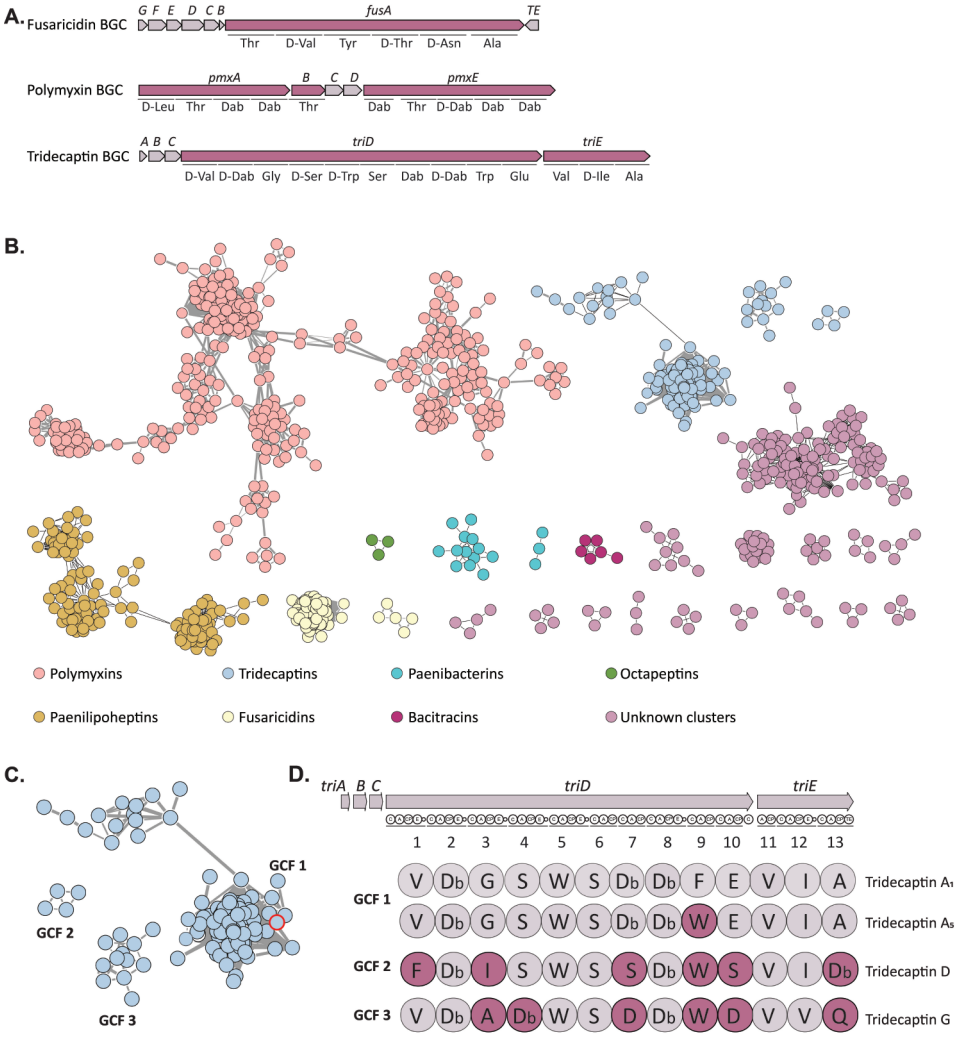


Figure 1. ***Paenibacillus*** **sp. JJ-21 BGC sequence similarity network and tridecaptin BGC analysis. A.** Organization of the NRPS BGCs predicted in the *Paenibacillus* sp. JJ-21 genome that encode the biosynthesis of lipopeptides. **B**. BiG-SCAPE sequence similarity network (SSN) (c 0.25) containing validated NRPS BGCs of *Paenibacillus* spp. visualized in Cytoscape. Each node represents one NRPS BGC identified by antiSMASH. Singletons and single links are not shown. **C.** Enlarged gene cluster families (GCFs) of tridecaptins BGCs. Red circle indicates the tridecaptin BGC detected in the genome of *Paenibacillus* sp. JJ-21. **D.** Scheme of the tridecaptin BGC and amino acid sequences of tridecaptin A_1, new tridecaptin A_5, tridecaptin D and tridecaptin G. The sequences of their peptide moieties were predicted from the active site sequences of their A domains. Note, the peptide sequences of tridecaptin A_5 and the well-known tridecaptin A_1 differ exclusively in one amino acid at position 9. Db: 2,4-diaminobutyric acid.

Surprisingly, three GCFs contained BGCs with 13 NRPS modules and were predicted as tridecaptin BGCs based on the antiSMASH (Blin *et al.*, 2021) and BiG-SCAPE (Navarro-Muñoz *et al.*, 2020) results (Figure 1C). The assembly line of tridecaptins is subdivided into two peptide synthetases, namely TriD and TriE, consisting of ten and three modules, respectively

(Lohans *et al.*, 2014). Notably, substrate specificity predictions by PARAS indicated that the tridecaptin BGCs within GCF 1 not only encoded the previously characterized tridecaptin A_1, but also a new analog, designated tridecaptin A_5, which should incorporate Trp in position 9, and was detected in the genome of *Paenibacillus* sp. JJ-21 (Figure 1D, Table S3). Additionally, we discovered another new structurally distinct analog of tridecaptins, which we designated tridecaptin D, and is predicted to be encoded by the tridecaptin BGCs of GCF 2. GCF 3 represented BGCs corresponding to the recently published tridecaptin G (da Costa *et al.*, 2022). Tridecaptin D differs from tridecaptin A_1 at six amino acid positions, with five positions of the molecule featuring amino acids that are unique among all known tridecaptins identified so far, namely Phe1, Ile3, Ser7, Ser10 and Dab13. We used antiSMASH (Blin *et al.*, 2021) to predict which modules in the BGC for tridecaptin D contained epimerization domains, *i.e.* domains that catalyze the conversion from L- to D-amino acids. Modules 1–5 and 8 of TriD and module 12 of TriE, which incorporate Phe, Dab, Ile, Ser, Trp, Dab, and Ile, respectively, were all predicted to contain epimerization domains (Table S3).

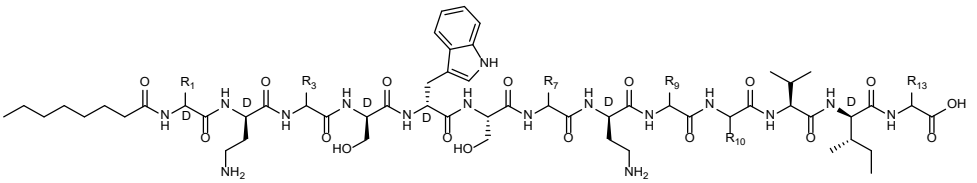
Identification of tridecaptin A_5 in vivo using metabolomics

To validate the bioinformatic prediction of the amino acid sequence of tridecaptin A_5, the lipopeptide was detected in the cultures of *Paenibacillus* sp. JJ-21 and analyzed by LC-MS/MS. For this, *Paenibacillus* sp. JJ-21 was fermented in liquid TSB media, biomass was collected by centrifugation and, subsequently, extracted with acidified 70% IPA. LC-MS/MS analysis identified fusaricidins and polymyxins as the major specialized metabolites. Considering that all known tridecaptins contain several positively charged diaminobutyric acids (Dab), we searched for the MS/MS spectra that contained a *b* ion with the *m/z* 101.0709. Besides polymyxins that also contained multiple Dab residues, we detected a compound with an *m/z* 539.9712, corresponding to an $[M+3H]^{3+}$ ion. On the basis of the biosynthetic features derived from the genome sequence of *Paenibacillus* sp. JJ-21, we initiated *de novo* sequencing of the compound with the *m/z* 539.9712 from the *y* and *b* ion series (Figure S1). The MS/MS spectrum showed a low-mass ion region containing a Trp immonium ion with an *m/z* 159.0912, as well as internal fragment ions. Fragment analysis of the *y* and *b* ions yielded an amino acid sequence that matched that predicted for tridecaptin A_5, which is FA-D-Val–D-Dab–Gly–D-Ser–D-Trp–L-Ser–L-Dab–D-Dab–L-Trp–L-Glu–L-Val–D-Ile–L-Ala. The MS/MS spectrum revealed the presence of a hydroxy-containing C11 fatty acid in tridecaptin A_5, similar to the previously reported 3-hydroxy-methyldecanoic fatty acids found in tridecaptin A_3 and tridecaptin A_4 (Lohans *et al.*, 2014). The MS/MS studies did not allow us to discriminate between Leu, Ile and *allo*-Ile in position 12 of the tridecaptin sequence. Here, the assignment was performed according to the structure prediction from our genome mining studies and comparison with the literature data (Lohans *et al.*, 2014). Additionally, the MS results would not allow optical isomers to be distinguished. Therefore, the configurations

of the residues of tridecaptin A_5 were predicted from the domain organization of the modules along the assembly lines.

Chemical synthesis and antimicrobial testing of tridecaptin analogs

Because of the low levels of production of tridecaptin A_5 by *Paenibacillus* sp. JJ-21, we decided to produce synthetic analogs of the peptides by solid phase peptide synthesis (SPPS). The added value of this approach is that it also allows the synthesis and direct comparison of analogs, chosen based on the substrate specificity and stereochemical predictions generated from the analysis of the tridecaptin BGCs. Previously, it was demonstrated that a synthetic analog of tridecaptin A_1 ($TriA_1$) wherein the N-terminal acyl moiety was replaced with octanoic acid retained the full activity of the natural product against ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *A. baumannii*, *P. aeruginosa*, and *Enterobacter* species) (Cochrane *et al.*, 2014). For this reason, in the present studies the chiral lipid tail of the tridecaptins was also replaced with the octanoyl chain, which is more accessible than natural lipid and enables larger quantities of peptide to be obtained for biological studies. Linear synthetic analogs of tridecaptin A_1, tridecaptin A_5 and tridecaptin D were synthesized and designated as Oct-$TriA_1$ (**1**), Oct-$TriA_5$ (**2**) and Oct-TriD (**3**), respectively (Table S4, Figure 2, Figures S2–S4).



		R_1	R_3	R_7	R_9	R_{10}	R_{13}
Oct-$TriA_1$	(1)	D-Val	Gly	Dab	Phe	Glu	Ala
Oct-$TriA_5$	(2)	D-Val	Gly	Dab	Trp	Glu	Ala
Oct-TriD	(3)	D-Phe	D-Ile	Ser	Trp	Ser	Dab

Figure 2. Structures of synthetic tridecaptin A variants Oct-$TriA_1$ (1), Oct-$TriA_5$ (2) and Oct-TriD (3).

Subsequently, Oct-$TriA_1$, Oct-$TriA_5$ and Oct-TriD were assayed for activity against the ESKAPE pathogens. MIC values were determined using broth-dilution assays (Table 1). Oct-$TriA_1$ displayed potent activity against most of the Gram-negative test-strains, but was less active against Gram-positive bacteria, similar to known tridecaptins (i.e., tridecaptins A, B, C, and M) (Cochrane *et al.*, 2014). Surprisingly, Oct-TriD showed significant bioactivity against the Gram-positive species *S. aureus, E. faecalis and E. faecium*, with an MIC between 4–8 μg/mL, in addition to moderate activity against Gram-negative pathogens. To the best of our knowledge, Oct-TriD is the first tridecaptin analog that is predominantly active against Gram-positive bacteria. Moreover, Oct-$TriA_5$ showed broad-spectrum bioactivity, similar to

the recently published synthetic analog syn-CNRL5 and natural tridecaptin G (Wang *et al.*, 2022a, da Costa *et al.*, 2022). Specifically, Oct-TriA$_5$ exhibited activity against Gram-negative ESKAPE pathogens, with MICs ranging from 2 to 16 μg/mL. Notably, a single amino acid substitution at position 9 from Phe in Oct-TriA$_1$ to Trp in Oct-TriA$_5$ changed the bioactivity profile of Oct-TriA$_5$ from being limited to Gram-negative bacteria to broad-spectrum activity.

In addition, tridecaptins were tested for their hemolytic and cytotoxic activity (Table 1). Oct-TriA$_5$ and Oct-TriD showed increased hemolytic and cytotoxic activity compared to Oct-TriA$_1$, suggesting that the altered spectrum of activity of Oct-TriA$_5$ and Oct-TriD is due to unspecific membrane lysis properties and emphasizing the importance of the amino acid in position 9 for selectivity towards antibacterial activity.

Mode of action on Gram-positive and Gram-negative bacteria

Tridecaptin A$_1$ exerts its bactericidal effect on Gram-negative bacteria by binding to lipid II on the surface of the inner membrane and by disrupting the proton motive force (Cochrane *et al.*, 2016). Tridecaptin A$_1$ binds to both Gram-positive and Gram-negative lipid II, although it has a much higher affinity for the Gram-negative analog (Cochrane *et al.*, 2014, Bann *et al.*, 2019). To assess if Oct-TriA$_5$ analogs bind to lipid II, we performed in vitro lipid II antagonization assays. Gram-positive lipid II, containing lysine at position 3 of the pentapeptide, was prepared by total chemical synthesis (Dong *et al.*, 2018, Karak *et al.*, 2024). The bioactivity of Oct-TriA$_5$ against *S. aureus* USA300 (MRSA) was evident with a MIC of 8 μg/mL. However, the addition of lipid II significantly reduced the efficacy of Oct-TriA$_5$, as the growth of *S. aureus* USA300 (MRSA) was not inhibited at the concentration of 8x MIC (Figure S5). This indicates that Oct-TriA$_5$ binds to lipid II of Gram-positive bacteria.

To monitor the effect of tridecaptins on the membrane permeability of Gram-positive and Gram-negative bacteria we performed a LIVE/DEAD staining assay with SYTO 9 and propidium iodide with *S. aureus* ATCC 29213 and *E.coli* ATCC 25922, as published previously (Tenconi *et al.*, 2018). Microscopic analysis demonstrated that the vast majority of *S. aureus* cells, when exposed to Oct-TriA$_5$ and Oct-TriD, exhibited red staining comparable to the positive control (Figure 3).

Table 1. In vitro minimum inhibitory concentrations (MICs) of Oct-TriA$_1$, Oct-TriA$_5$ and Oct-TriD[a]

Strain		Species	MIC			
			Oct-TriA$_1$ (1)	Oct-TriA$_5$ (2)	Oct-TriD(3)	Colistin
BV18	UAMS-1625	*S. aureus*	32	8	16	> 64
BV249	ATCC 29213	*S. aureus*	32	8	8	> 64
BV1402	1126387	*S. aureus*	32	8	16	> 64
BV29	ATCC 29212	*E. faecalis*	32	8	8	> 64
BV99	ATCC 25922	*E. faecalis*	32	8	8	> 64
BV144	IHMA 890472	*E. faecium*	16	8	4	> 64
BV159	ATCC 51559	*E. faecium*	32	8	4	> 64
BV94	IHMA 777621	*A. baumannii*	8	4	8	64
BV160	ATCC 17978	*A. baumannii*	8	4	16	1
BV374	HUMC1	*A. baumannii*	16	8	16	1
BV23	ATCC 25922	*E. coli*	2	2	32	1
BV1475	1220120	*E. coli*	2	4	16	0.5
BV1469	1214245	*E. coli*	2	4	16	32
BV306	ATCC 13883	*K. pneumoniae*	4	4	8	1
BV1445	1227947	*K. pneumoniae*	4	4	> 64	0.5
BV1447	1228586	*K. pneumoniae*	4	4	64	64
BV34	PAO1	*P. aeruginosa*	64	16	> 64	2
BV1551	1226072	*P. aeruginosa*	8	8	32	0.5
BV1544	1218019	*P. aeruginosa*	4	4	32	> 64
Hemolysis EC50 [μg/mL]			102.2	6.9	2.2	n/a
IC50 on HepG2 w/o FCS [μg/mL]			109.6	11.4	6.2	n/a

[a] MIC values reported in units of μg/mL

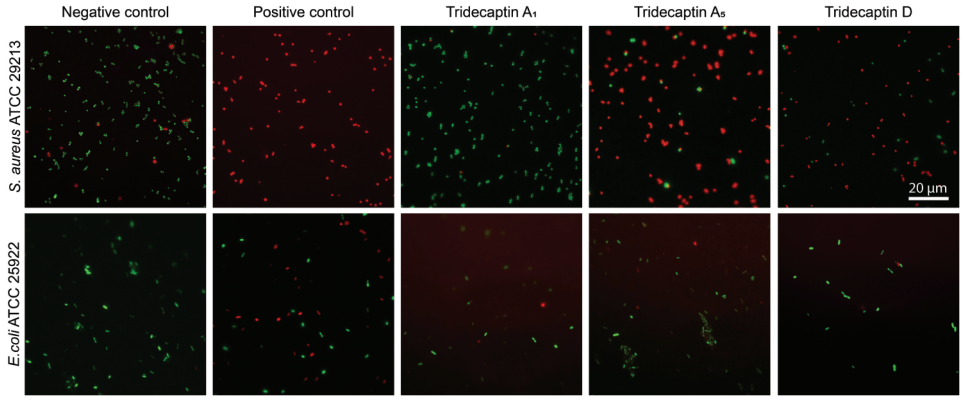


Figure 3. LIVE/DEAD staining of S. *aureus* ATCC 29213 and *E. coli* ATCC 25922 cells in the control conditions and after being exposed to nisin/polymyxin B or tridecaptin analogs Oct-TriA$_1$, Oct-TriA$_5$ and Oct-TriD. Green color indicates the cells with intact membrane, while cell with compromised membrane are stained in red. Note, after the treatment with Oct-TriA$_5$ and Oct-TriD *S. aureus* cells were stained red indicating the membrane disruption.

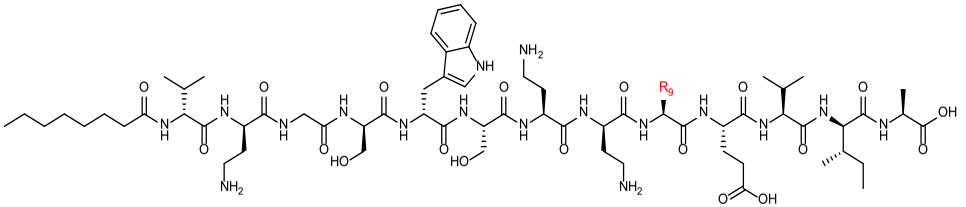
S. aureus cells incubated with Oct-TriA$_5$ and Oct-TriD had lysed, indicative of nonspecific membrane disruption, while those incubated with Oct-TriA$_1$ were still alive, suggesting that Oct-TriA$_1$ did not have any influence on the membrane permeability. Treatment of *E. coli* cells with Oct-TriA$_1$, Oct-TriA$_5$ or Oct-TriD did affect the membrane integrity, while after the same treatment with polymyxin (4 µg/mL; 2 × MIC) resulted in death of half of the *E.coli* cells. These data strongly suggest that Oct-TriA$_5$ and Oct-TriD act on Gram-positive bacteria by disrupting the cellular membrane. This is most likely explained by the presence of Trp9 in Oct-TriA$_5$ instead of Phe9 in Oct-TriA$_1$, that made the Oct-TriA$_5$ more hydrophobic and, therefore, increased its membrane disruption capacity.

Amino acid substitutions in tridecaptin A at position 9 impacting bioactivity and cytotoxicity

Considering the striking difference in the spectrum of activity between tridecaptin A$_1$ (active against Gram-negatives) and A$_5$ (active against both Gram-positives and Gram-negatives), we decided to look into the impact of changes in position 9 on bioactivity. To do so, we synthesized a series of analogs wherein residue 9 was changed into either Gly, Ala, Val, Ile, Ser, Tyr, Glu, His or Dab, and compared these new variants to Oct-TriA$_1$ and Oct-TriA$_5$ for bioactivity, as well as its hemolytic and cytotoxic activity (Table S4, Figures S6–S14).

The synthetic analogs exhibited significant variation in their MIC values against the pathogens tested (Table 2). Hemolytic and cytotoxic activities also differed significantly, which provided further proof of concept that indeed the residue in position 9 plays a critical role in the hemolytic and cytotoxic activity of tridecaptin A (Table 2).

Table 2. In vitro minimum inhibitory concentrations (MICs) of synthetic tridecaptin analogs [a].



Strain		MIC									
		Oct-Gly9 (4)	Oct-Ala9 (5)	Oct-Val9 (6)	Oct-Ile9(7)	Oct-Ser9 (8)	Oct-Tyr9 (9)	Oct-Glu9 (10)	Oct-His9 (11)	Oct-Dab9(12)	Colistin
UAMS-1625	*S. aureus*	> 64	> 64	> 64	> 64	> 64	> 64	> 64	> 64	> 64	> 64
ATCC 29213	*S. aureus*	> 64	> 64	> 64	64	> 64	> 64	> 64	> 64	> 64	> 64
1126387	*S. aureus*	> 64	> 64	> 64	> 64	> 64	> 64	> 64	> 64	> 64	> 64
ATCC 29212	*E. faecalis*	> 64	> 64	64	64	> 64	64	> 64	> 64	> 64	> 64
ATCC 25922	*E. faecalis*	> 64	> 64	> 64	64	> 64	64	> 64	> 64	> 64	> 64
IHMA 890472	*E. faecium*	> 64	> 64	> 64	32	> 64	32	> 64	> 64	> 64	> 64
ATCC 51559	*E. faecium*	> 64	> 64	64	64	> 64	64	> 64	> 64	> 64	> 64
IHMA 777621	*A. baumannii*	> 64	> 64	32	16	> 64	32	> 64	> 64	> 64	64
ATCC 17978	*A. baumannii*	> 64	> 64	32	16	> 64	32	> 64	> 64	> 64	1
HUMC1	*A. baumannii*	> 64	> 64	32	16	> 64	32	> 64	> 64	> 64	1
ATCC 25922	*E. coli*	16	16	4	2	32	2	> 64	16	> 64	1
1220120	*E. coli*	16	8	4	2	32	4	64	8	> 64	0.5
1214245	*E. coli*	16	16	4	2	64	2	> 64	16	> 64	32
ATCC 13883	*K. pneumoniae*	16	8	2	2	64	4	> 64	16	> 64	1
1227947	*K. pneumoniae*	32	16	8	4	> 64	4	> 64	32	> 64	0.5

1228586	*K. pneumoniae*	8	8	8	4	8	4	> 64	8	> 64	64
PAO1	*P. aeruginosa*	64	64	64	64	> 64	64	> 64	32	> 64	2
1226072	*P. aeruginosa*	32	32	16	16	32	4	> 64	4	8	0.5
1218019	*P. aeruginosa*	> 64	> 64	16	16	32	4	> 64	2	> 64	> 64
Hemolysis EC50 [μg/mL]		>128	>128	>128	>128	>128	65.6	>128	>128	>128	n/a
IC50 on HepG2 w/o FCS [μg/mL]		>128	>128	>128	110.2	>128	64	>128	>128	>128	>128

[a] MIC values reported in units of μg/mL

Of particular interest was Oct-His9, which efficiently inhibited colistin-resistant *P. aeruginosa* 1218019, with an MIC of 2 μg/mL. Furthermore, Oct-His9 showed lower hemolytic and cytotoxic activity compared to Oct-TriA_1. Conversely, synthetic analogs Oct-Gly9, Oct-Ala9, Oct-Ser9, Oct-Glu9 and Oct-Dab9 showed no antibacterial activity against Gram-positive bacteria at a concentration of 64 μg/mL and were moderately active against Gram-negative pathogens, with MICs ranging from 8 to 64 μg/mL. Similar to Oct-TriA_5 and Oct-TriD, variants Oct-Tyr9 and Oct-Ile9 showed moderate bioactivity against Gram-positives, and high cytotoxic and hemolytic activity. Synthetic analogs with Val and Ile at position 9 shared bioactivity, cytotoxic and hemolytic profiles with Oct-TriA_1. Taken together, our data highlight position 9 as a key residue for the spectrum and degree of bioactivity, as well as the cytotoxicity of tridecaptin A.

We here report the discovery of new lipopeptide antibiotics tridecaptin A_5 and tridecaptin D. Bioinformatics analysis of 785 complete *Paenibacillus* genomes revealed the BGC for tridecaptin A_5 in the genome of *Paenibacillus* sp. JJ-21. Thorough examination of the adenylation domains encoded by the tridecaptin A_5 BGC allowed us to predict the amino acid sequence of the peptide scaffold; this was then matched to the compound produced by *Paenibacillus* sp. JJ-21. Subsequent comparison of the tridecaptin A_5 BGC with that of tridecaptin A_1 revealed that they differed exclusively in module 9 of TriD. Based on the PARAS substrate specificity predictions we synthesized tridecaptin analogs Oct-TriA_5, Oct-TriD and reference Oct-TriA_1 to evaluate their antibacterial activity against a panel of ESKAPE pathogens. Surprisingly, while tridecaptins are primarily active against Gram-negative bacteria, Oct-TriA_5 possessed significant activity against both Gram-positive and Gram-negative bacteria, while Oct-TriD was primarily active against Gram-positive pathogens. Particularly striking is that a single amino acid substitution at position 9, from Phe in Oct-TriA_1 to Trp in Oct-TriA_5, changes the bioactivity profile from Gram-negative to broad-spectrum activity. Through the screening of analogs substituted at position 9, we subsequently identified the highly potent tridecaptin, Oct-His9, that exhibits effective activity against *P. aeruginosa* and, at the same time, possesses reduced hemolytic and cytotoxic properties. The observed decrease in cytotoxicity and the promising antibacterial properties, provide strong grounds for further investigation and development of these analogs as alternative antibiotics for combating infectious diseases associated with antibiotic-resistant bacterial pathogens.

Materials and Methods

Global genome mining of NRPSs

Genomes of all *Paenibacillus* spp. available from RefSeq (release 213) (O'Leary *et al.*, 2016) were downloaded from the NCBI FTP site. All genomes with more than 400 contigs were considered as low-quality assemblies and were removed from the collection. All genomes were analyzed using AntiSMASH (version 6.0.1) (Blin *et al.*, 2021) to obtain BGC predictions. These predictions were then used as input into BiG-SCAPE (version 1.1.4) (Navarro-Muñoz *et al.*, 2020), for the creation of a sequence similarity network, with distance matrix cutoff set to 0.25. The resulting full network was visualized by Cytoscape (3.9.1) (Shannon *et al.*, 2003). To predict the specificity of the adenylation domains of NRPS BGCs in this study, we used the software package PARAS (v0.0.4, available at https://github.com/BTheDragonMaster/paras). PARAS is an adenylation domain predictor that uses structure-guided sequence alignments to extract the active site prior to making a prediction.

Genome sequencing, assembly and annotation of *Paenibacillus* sp. JJ-21

Paenibacillus sp. JJ-21 was grown in tryptic soy broth (TSB) at 30 °C and 220 rpm for 24 h. DNA was extracted from *Paenibacillus* sp. JJ-21 as described (Kieser, 2000). DNA quality was verified by agarose gel electrophoresis. PacBio sequencing and assembly was performed by Novogene (UK). Generally, libraries were prepared using SMRTbell template prep kit (PacBio, USA) according to manufacturer instructions. Sequencing was then performed using PacBio Sequel platform in continuous long reads mode. Assembly was done using Falcon (version 1.8.1) (Chin *et al.*, 2016). BGCs in this genome were annotated using AntiSMASH (version 6.0.1) (Blin *et al.*, 2021).

Growth of *Paenibacillus* sp. JJ-21 and specialized metabolites extraction

Paenibacillus sp. JJ-21 was obtained from the Auburn University Plant-Associated Microbial strain collection and had previously been isolated from the root surface of a field-grown corn plant (*Zea mays*) grown in Dunbar, Nebraska, and maintained as a viable cryostock in a -80 °C freezer. *Paenibacillus* sp. JJ-21 was grown at 30 °C on tryptic soy agar (TSA) for 72 h and two to three colonies were inoculated into TSB and incubated overnight at 30 °C. This inoculum (1%) was transferred to 1 L Erlenmeyer flasks containing 400 mL of sterile TSB and fermented at 30 °C while shaking at 220 rpm for 96 h. Cells were collected by centrifugation (5000 × *g*, 30 min, 4 °C) and extracted for 6 h with 70% isopropyl alcohol (IPA) supplemented with 0.1% (v/v) formic acid (FA). The crude extracts were clarified by centrifugation, then the solvent was evaporated under reduced pressure and reconstituted in 50% MeOH.

LC-MS/MS analysis

For LC-MS analyses, extracts were dissolved in 50% MeOH to a final concentration of 2 mg/mL, and 1 µL was injected into Shimadzu Nexera X2 UHPLC system coupled to a Shimadzu 9030 QTOF mass spectrometer, and data acquisition was performed as previously described (van Bergeijk *et al.*, 2022). Briefly, all the samples were analyzed in positive polarity, using data-dependent acquisition mode. In this regard, full scan MS spectra (m/z 100–1700, scan rate 10 Hz, ID enabled) were followed by two data-dependent MS/MS spectra (m/z 100–1700, scan rate 10 Hz, ID disabled) for the two most intense ions per scan. The ions were fragmented using collision-induced dissociation (CID) with fixed collision energy (CE 20 eV) and excluded for 1 s before being re-selected for fragmentation. The parameters used for the ESI source were: interface voltage 4 kV, interface temperature 300 °C, nebulizing gas flow 3 L/min, and drying gas flow 10 L/min.

HRMS analyses of the synthetic peptides were performed on a Thermo Scientific Dionex UltiMate 3000 HPLC system with a Phenomenex Kinetex C_{18} (2.1 x 150 mm, 2.6 µm) column at 35 °C and equipped with a diode array detector. The following solvent system, at a flow rate of 0.3 mL/min, was used: solvent A, 0.1 % formic acid in water; solvent B, 0.1% formic acid in acetonitrile. Gradient elution was as follows: 95:5 (A/B) for 1 min, 95:5 to 5:95 (A/B) over 9 min, 5:95 to 2:98 (A/B) over 1 min, 2:98 (A/B) for 1 min, then reversion back to 95:5 (A/B) over 2 min, 95:5 (A/B) for 1 min. This system was connected to a Bruker micrOTOF-Q II mass spectrometer (ESI ionization) calibrated internally with sodium formate.

Peptide synthesis

Fmoc-L-Dab(Boc)-OH, Fmoc-D-Dab(Boc)-OH, Fmoc-D-Ile-OH, Fmoc-allo-Ile-OH and 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU) were purchased from Combi-Blocks. All other Fmoc-amino acids, the Fmoc-Ala-Wang resin and Fmoc-Glu(OtBu)-Wang resin were purchased from P3 BioSystems. 2-Chlorotrityl chloride resin (CTC) was purchased from Iris Biotech. Octanoic acid was purchased from Alfa Aesar. ((1H-Benzo[d][1,2,3]triazol-1-yl)oxy)tris(dimethylamino) phosphonium hexafluorophosphate (BOP), N,N-Diisopropylcarbodiimide (DIC) and triisopropylsilane (TIPS) were purchased from Manchester Organics. Diisopropylethylamine (DIPEA), piperidine, trifluoroacetic acid (TFA) and dimethyl sulfoxide (DMSO) were purchased from Carl Roth. Dichloromethane (CH_2Cl_2) and petroleum ether were purchased from VWR Chemicals. Acetonitrile (MeCN), dimethylformamide (DMF) and methyl tertiary-butyl ether (MTBE) were purchased from Biosolve.

Resin swelling

The resin was swollen in 10 mL of DMF for 300 s prior to the first coupling.

Automated coupling protocol

Step	Function	Duration/Temperature
1	Initial Deprotection N-terminus	15 s at 60°C then 30 s at 70°C
2	Deprotection N-terminus	15 s at 60°C then 180 s at 70°C
3	Wash (DMF)	RT
4	Wash (DMF)	RT
5	Wash (DMF)	RT
6	Coupling amino acid	15 s at 60°C then 300 s at 70°C

After coupling of the final residue on the synthesizer, the resin was washed with DCM, filtered and treated with 3 mL of TFA : TIPS : H_2O (95 : 2.5 : 2.5, v/v) for 90 min. The reaction mixture was filtered through cotton, the filtrate was precipitated in MTBE : petroleum ether (1 : 1, v/v) and centrifuged (4500 rpm, 5 min). The pellet was then resuspended in MTBE : petroleum ether (1 : 1, v/v) and centrifuged again (4500 rpm, 5 min). Finally the pellet containing the crude lipopeptide was dissolved in tBuOH : H_2O (1 : 1, v/v) and lyophilized overnight. The crude mixtures were subsequently purified by RP-HPLC. Fractions were assessed by HPLC and LC-MS and product containing fractions were pooled, frozen and lyophilized to yield the pure lipopeptides in >95% purity (determined by HPLC).

General procedure for manual solid phase peptide synthesis

The peptides were made on a 0.25 mmol scale on either preloaded Wang resin or 2-Chlorotrityl chloride (CTC) resin. The syntheses of Oct-TriA$_1$ and Oct-TriA$_5$ were performed on preloaded Fmoc-Ala-Wang resin (0.29 mmol/g loading). The synthesis of Oct-TriD was performed by loading Fmoc-Dab(Boc)-OH on CTC resin. Resin loading was determined to be 0.50 mmol/g. All couplings with the exception of Fmoc-d-*allo*-Ile-OH and Fmoc-d-Ile-OH were performed using 4 eq. of amino acid or fatty acid, benzotriazol-1-yloxytris(dimethylamino) phosphonium hexafluorophosphate (BOP) (4 eq.) and N,N-diisopropylethylamine (DiPEA) (8 eq.) in 10 mL of DMF for 1 h at RT, under nitrogen flow. Fmoc-d-*allo*-Ile and Fmoc-d-Ile-OH were coupled by treating the resin with 2 eq. of the amino acid, 2 eq. of BOP and 4 eq. of DiPEA in 10 mL of DMF overnight at RT. Fmoc group removal was performed by treating the resin with 10 mL of piperidine : DMF (1 : 4, v/v) for 5 min and then again for 15 min. Final sidechain deprotection and cleavage from the resin was carried out by treating the resin with 10 mL of TFA : TIPS : H_2O (95 : 2.5 : 2.5, v/v) for 90 min. The reaction mixture was filtered through cotton, the filtrate was precipitated in MTBE : petroleum ether (1 : 1, v/v) and centrifuged (4500 rpm, 5 min). The pellet was then resuspended in MTBE : petroleum ether (1 : 1, v/v) and centrifuged again (4500 rpm, 5 min). Finally the pellet containing the crude lipopeptide was dissolved in tBuOH:H_2O (1 : 1, v/v) and lyophilized overnight. The crude mixtures were subsequently purified by RP-HPLC. Fractions were assessed by HPLC and LC-

MS and product containing fractions were pooled, frozen and lyophilized to yield the pure lipopeptides in >95% purity (determined by HPLC).

General procedure for automated solid phase peptide synthesis

Position 9 analogs were made on a 0.05 mmol scale on preloaded Fmoc-Ala-CTC resin (0.68 mmol/g) using a CEM Liberty Blue automated peptide synthesizer with microwave irradiation. Couplings were performed at 0.125 M concentration using 5 eq. of amino acid, 5 eq. of 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate (HATU) and 10 eq. of diisopropylethylamine (DIPEA). Fmoc group removal was performed using piperidine : DMF (1 : 4, v/v). Cleavage from resin along with global deprotection and subsequent purification by RP-HPLC was performed as described above for the manually synthesized peptides.

Prep RP-HPLC purification

Peptides were purified using a BESTA-Technik system with a Dr. Maisch Reprosil Gold 120 C_{18} column (25 × 250 mm, 10 µm) and equipped with a ECOM Flash UV detector monitoring at 214 nm and 254 nm. The following solvent system, at a flow rate of 12 mL/min, was used: solvent A, 0.1 % TFA in water/acetonitrile 95/5; solvent B, 0.1 % TFA in water/acetonitrile 5/95. Gradient elution was as follows: 100:0 (A/B) for 5 min, 100:0 to 50:50 (A/B) over 50 min, 50:50 to 0:100 (A/B) for 3min, then reversion back to 100:0 (A/B) over 1 min, 100:0 (A/B) for 5 min.

Analytical RP-HPLC

Shimadzu Prominence-i LC-2030 system with a Dr. Maisch ReproSil Gold 120 C_{18} column (4.6 × 250 mm, 5 µm) at 30 °C and equipped with a UV detector monitoring 214 nm and 254 nm. The following solvent system, at a flow rate of 1 mL/min, was used: solvent A, 0.1 % TFA in water/acetonitrile 95/5; solvent B, 0.1 % TFA in water/acetonitrile 5/95. Method A: Gradient elution was as follows: 100:0 (A/B) for 2 min, 100:0 to 0:100 (A/B) over 23 min, 0:100 (A/B) for 1 min, 0:100 (A/B) then reversion back to 100:0 (A/B) over 1 min, 100:0 (A/B) for 3 min. Method B: Gradient elution was as follows: 100:0 (A/B) for 2 min, 100:0 to 50:50 (A/B) over 45 min, 50:50 (A/B) to 0:100 (A/B) over 1 min, 0:100 (A/B) for 6 min then reversion back to 100:0 (A/B) over 1min, 100:0 (A/B) for 5min. Method C: Gradient elution was as follows: 100:0 (A/B) for 2 min, 100:0 to 0:100 (A/B) over 53 min, 0:100 (A/B) for 1 min, 0:100 (A/B) then reversion back to 100:0 (A/B) over 1 min, 100:0 (A/B) for 4 min.

Antimicrobial testing

All minimum inhibitory concentrations (MICs) were determined according to Clinical and Standards Laboratory Institute (CLSI) guidelines. In brief, agar plates were inoculated from

glycerol stocks and incubated overnight at 35 °C. Peptides DMSO stocks were dispensed into 96-well round bottom plates using a TECAN D300e dispenser. Colistin sulfate (Sigma-Aldrich) was dissolved in water and serially diluted in cation-adjusted Mueller Hinton broth (10 µL per well). The final concentrations of the test compounds were 0.06 to 64 µg/mL. Inocula were prepared by direct colony suspension in NaCl at 0.5 MacFarland which was further diluted 200-fold in cation-adjusted Mueller Hinton broth for a target inoculum of 2 x 10^5 colony-forming units (CFUs)/mL. 100 µL of the prepared inoculum was added to each well of the prepared plates containing the test compounds. Plates were sealed with parafilm and incubated for 20 hours at 35 °C. Images of all plates were recorded and MIC was assessed visually.

Hemolysis assays

Peptides, dissolved in DMSO, were dispensed into 96-well round bottom microplates using a TECAN D300e dispenser and subsequently diluted with 100 µL PBS. PBS or 2 % Triton X-100 were used as negative and positive controls, respectively. Rabbit red blood cells (BioConcept) were diluted in PBS to a final concentration of 2 % and 100 µL were added to each well of the prepared 96-well plates. Final concentration of the test compounds were 0.25 to 128 µg/mL. Plates were incubated at 37 °C for 1 h and afterwards subjected to centrifugation at 1000 x *g* for 5 min. Then 30 µL of supernatant was transferred to a new round bottom plate and absorbance was measured at 405 nm (TECAN Infinite F200). EC50s were calculated after blank subtraction and normalization to the Triton control.

Cytotoxicity determination

HepG2 cells were grown in EMEM (Sigma) supplemented with 2 mM glutamine and 10 % fetal bovine serum (Fisher) at 37 °C, 5 % CO_2. 20'000 cells per well were seeded into clear tissue culture treated 96 well plates and incubated for 24 hours. Peptides, dissolved in DMSO, were dispensed into 96 deep well plates (TECAN D300e dispenser) and diluted to a final concentration of 1 to 128 µg/mL using fresh medium without serum. The next day, the old media was removed from the cells and 200 µL of the prepared compound dilutions were added per well. Plates were incubated for another 24 hours before assessing the cell viability using the CellTiter-Glo kit (Promega) according to the manufacturers protocol. IC50s were calculated after normalization to untreated control.

Lipid II antagonization assay

The minimum inhibitory concentrations (MICs) of Oct-TriA_1, Oct-TriA_5 and vancomycin against *S. aureus* USA300 (MRSA) were determined according to Clinical and Standards Laboratory Institute (CLSI) guidelines. Gram-positive lipid II, containing lysine at position 3

of the pentapeptide, was prepared by total chemical synthesis (Dong *et al.*, 2018, Karak *et al.*, 2024). Lipid II in 1 : 1 chloroform : methanol (v/v) was added to a polypropylene 96-well plate (5-fold molar excess compared to test antibiotics) and the organic solvent was allowed to evaporate. Oct-TriA$_5$ and vancomycin in MHB (50 µL, 16xMIC) were added to the wells with the 5-fold molar excess of pure Gram-positive lipid II in triplicates and to the control wells without lipid II. Single colony of *S. aureus* USA300 (MRSA) from a fresh blood agar plate was suspended in TSB and grown to an OD_{600} of 0.5. The bacterial culture then was diluted in MHB with 0.002% polysorbate-80 to reach 10^6 CFU/mL and 50 µL were mixed with test compounds to achieve the final concentration of 8 x MIC for all tested compounds. The plate was incubated at 37 °C for 18 h with constant shaking (600 rpm) and subsequently inspected for visible bacterial growth. To visualize the viability of the indicator strain, resazurin sodium salt solution (Merck, Dorset, UK) was added to the cultures to achieve the final concentration of 0.0015% (w/v) and incubated at room temperature for 1 h.

LIVE/DEAD staining and confocal microscopy

This assay was performed according to a previously described procedure (Li *et al.*, 2020, Zacchetti *et al.*, 2018). Briefly, overnight cultures of *E. coli* ATCC 25922 and *S. aureus* ATCC 29213 were diluted to an OD_{600} of 0.2 in MHB and mixed with 1 × MIC of Oct-TriA$_1$, Oct-TriA$_5$ and Oct-TriD. Nisin and polymyxin B were used as a positive controls at a concentration of two-fold MIC. At the same time, green fluorescent nucleic acid stain SYTO 9 (excitation, 450–490 nm; emission, 500–550 nm) and red fluorescent nucleic acid stain propidium iodide (excitation, 574–599 nm; emission, 612–682 nm) (LIVE/DEAD Baclight Bacterial Viability Kit, Invitrogen) were added to the above cell suspensions. After mixing and incubating for 15 min in the dark at room temperature, cells were briefly sedimented via centrifugation and resuspended in fresh MHB. Then the cell suspensions were loaded on 1.5% agarose pads and analyzed by Zeiss Imager M2 microscope. Confocal images were obtained at 4 random locations for each sample and visualized using FIJI version 1.51H (Schindelin *et al.*, 2012).

Acknowledgments

The work was supported by the NACTAR program of The Netherlands Organization for Scientific Research (NWO), Grant 16440 to G.P.vW., M.H.M. and N.I.M.

Supplementary information

Table S1. A-domain specificity analysis of NRPS BGCs from the genome of *Paenibacillus* sp. JJ-21.

Domain	Residues in the binding pocket	Amino acid prediction	Probability
Fusaricidin A1	DFWNIGMVH	Threonine	0.716357143
Fusaricidin A2	DAFWLGCTF	Valine	0.795809524
Fusaricidin A3	DASTLAGVC	Tyrosine	0.78
Fusaricidin A4	DFWNIGMVH	Threonine	0.716357143
Fusaricidin A5	DLTKIGEVG	Asparagine	0.96
Fusaricidin A6	DFPNFCIVY	Alanine	0.77
Polymyxin A1	DVGEISSID	2,4-diaminobutyric acid	1
Polymyxin A2	DFWNIGMVH	Threonine	0.716357143
Polymyxin A3	DVGEISSID	2,4-diaminobutyric acid	1
Polymyxin A4	DVGEISSID	2,4-diaminobutyric acid	0.99
Polymyxin A5	DVGEISSID	2,4-diaminobutyric acid	0.99
Polymyxin A6	DAWIVGAIV	Leucine	0.716357143
Polymyxin A7	DFWNIGMVH	Threonine	0.18
Polymyxin A8	DVGEISAID	2,4-diaminobutyric acid	0.716357143
Polymyxin A9	DVGEISAID	2,4-diaminobutyric acid	0.99
Polymyxin A10	DFWNIGMVH	Threonine	0.99
Tridecaptin A1	DAFWLGGTF	Valine	0.661904762
Tridecaptin A2	DVGEISSID	2,4-diaminobutyric acid	0.99
Tridecaptin A3	DILQMGMVW	Glycine	0.98

Tridecaptin A4	DVWHFSLVD	Serine	0.99
Tridecaptin A5	DAWAFAGVA	Tryptophan	0.63
Tridecaptin A6	DVWHFSLVD	Serine	0.99
Tridecaptin A7	DVGEISSID	2,4-diaminobutyric acid	0.99
Tridecaptin A8	DVGEISSID	2,4-diaminobutyric acid	0.99
Tridecaptin A9	DAWAFAGVA	Tryptophan	0.63
Tridecaptin A10	DAKDLGVVD	Glutamic acid	0.66
Tridecaptin A11	DAFWLGGTF	Valine	0.621571429
Tridecaptin A12	DAFFLGITF	Isoleucine	0.815
Tridecaptin A13	DVFWLGGTF	Alanine	0.72

Table S2. A-domain specificity comparison of reference tridecaptin A_1 BGC and tridecaptin cluster from the genome of *Paenibacillus* sp. JJ-21.

Domain	Residues in the binding pocket		Amino acid prediction	
	Tridecaptin A_1 reference BGC BGC0000449	Tridecaptin A_5 BGC of *Paenibacillus* sp. JJ-21	Tridecaptin A_1 reference BGC BGC0000449	Tridecaptin A_5 BGC of *Paenibacillus* sp. JJ-21
TriA1	DAFWLGGTF	DAFWLGGTF	Valine/Glycine	Valine
TriA2	DVGEISSID	DVGEISSID	2,4-diaminobutyric acid	2,4-diaminobutyric acid
TriA3	DILQMGMVW	DILQMGMVW	Glycine	Glycine
TriA4	DVWHFSLVD	DVWHFSLVD	Serine	Serine
Tri A5	DAWAFAGVA	DAWAFAGVA	Tryptophan /Phenylalanine	Tryptophan
TriA6	DVWHFSLVD	DVWHFSLVD	Serine	Serine
TriA7	DVGEISSID	DVGEISSID	2,4-diaminobutyric acid	2,4-diaminobutyric acid
TriA8	DVGEISSID	DVGEISSID	2,4-diaminobutyric acid	2,4-diaminobutyric acid
TriA9	DAWTFAGVA	DAWAFAGVA	Phenylalanine /Valine/ Isoleucine	Tryptophan
TriA10	DAKDLGVVD	DAKDLGVVD	Glutamic acid	Glutamic acid
TriA11	DAFWLGGTF	DAFWLGGTF	Valine	Valine
TriA12	DAFFLGITF	DAFFLGITF	Isoleucine	Isoleucine
TriA13	DVFWLGGTF	DVFWLGGTF	Alanine	Alanine

Table S3. Predicted amino-acid specificities of the adenylation domains present in representative BGCs of tridecaptin gene cluster families (GCFs) in a BiG-SCAPE sequence similarity network.

Accession number	GCF	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	Compound
NZ_CP092831.1. region009	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Phe	Glu	Val	D-Ile	Ala	Tridecaptin A_1
NZ_CP048793.1. region008	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Phe	Glu	Val	D-Ile	Ala	Tridecaptin A_1
NZ_CP061172.1. region008	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Phe	Glu	Val	D-Ile	Ala	Tridecaptin A_1
NZ_CP073683.1. region012	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Phe	Glu	Val	D-Ile	Ala	Tridecaptin A_1
NZ_CP097778.1. region009	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Phe	Glu	Val	D-Ile	Ala	Tridecaptin A_1
NZ_JTHP01000234.1. region001	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Phe	Glu	Val	D-Ile	Ala	Tridecaptin A_1
NZ_CP023711.1. region001	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_CP025957.1. region010	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	NA	Val	D-Ile	Ala	Tridecaptin A_5
NZ_CP084033.1. region008	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_CP086373.1. region009	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NC_023037.2. region009	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_CP011420.1. region010	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	NA	Tridecaptin A_5
NZ_CP011512.1. region010	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_CP015423.1. region015	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_CP017968.3. region011	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5

NZ_CP073682.1. region011	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_CP097767.1. region001	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_FOYG01000003.1. region001	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_JAFIBS010000001.1. region012	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_JMLR01000007.1. region003	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_NOLA01000018.1. region001	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_OXKC02000014.1. region001	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_POVS01000007.1 .region001	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
*Paenibacillus*_sp._MBT _JJ-21. region011	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_AMQU01000029.1. region001	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_CP040829.1. region014	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_CP049598.1. region019	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_CP097766.1. region009	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_CP097771.1. region013	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_JAACZY010000021.1. region001	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_JAAMNS010000007.1. region001	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_JNCB01000011.1. region006	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_JTHO01000008.1. region001	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5

NZ_JWJJ01000007.1. region001	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NZ_QVPU01000010.1. region001	1	D-Val	D-Dab	Gly	D-Ser	D-Trp	Ser	Dab	D-Dab	Trp	Glu	Val	D-Ile	Ala	Tridecaptin A_5
NC_015690.1. region011	2	D-Phe	D-Dab	D-Ile	D-Ser	D-Trp	Ser	Ser	D-Dab	Trp	Ser	Val	D-Ile	Dab	Tridecaptin D
NC_016935.1. region011	2	D-Phe	D-Dab	D-Ile	D-Ser	D-Trp	Ser	Ser	D-Dab	Trp	Ser	Val	D-Ile	Dab	Tridecaptin D
NC_017672.3. region011	2	D-Phe	D-Dab	D-Ile	D-Ser	D-Trp	Ser	Ser	D-Dab	Trp	Ser	Val	D-Ile	Dab	Tridecaptin D
NZ_JAM-AVM010000004.1. region001	3	D-Val	D-Dab	D-Ala	D-Dab	D-Trp	Ser	Asp	D-Dab	Trp	Asp	Val	D-Val	Gln	Tridecaptin G

Table S4. HPLC and HRMS analysis of peptides.

Peptide	Name	Chemical Formula	Calcd. Exact mass	Mass found	Calcd.	Overall yield [%]
1	Oct-TriA$_1$	$C_{72}H_{113}N_{17}O_{19}$	1519,8399	760,9278 $[M+2H]^{2+}$	760,9272	19
2	Oct-TriA$_5$	$C_{74}H_{114}N_{18}O_{19}$	1558,8508	780,4332 $[M+2H]^{2+}$	780,4327	23
3	Oct-TriD	$C_{80}H_{120}N_{18}O_{19}$	1636,8977	819,4565 $[M+2H]^{2+}$	819,4562	51
4	Oct-Gly9	$C_{65}H_{107}N_{17}O_{19}$	1429,7929	715,9046 $[M+2H]^{2+}$	715,9038	10
5	Oct-Ala9	$C_{66}H_{109}N_{17}O_{19}$	1443,8086	722,9120 $[M+2H]^{2+}$	722,9116	11
6	Oct-Val9	$C_{68}H_{113}N_{17}O_{19}$	1471,8399	736,9278 $[M+2H]^{2+}$	736,9272	11
7	Oct-Ile9	$C_{69}H_{115}N_{17}O_{19}$	1485,8555	743,9357 $[M+2H]^{2+}$	743,9351	16
8	Oct-Ser9	$C_{66}H_{109}N_{17}O_{20}$	1459,8035	730,9098 $[M+2H]^{2+}$	730,9090	16
9	Oct-Tyr9	$C_{72}H_{113}N_{17}O_{20}$	1535,8348	768,9256 $[M+2H]^{2+}$	768,9247	16
10	Oct-Glu9	$C_{68}H_{111}N_{17}O_{21}$	1501,8140	751,9150 $[M+2H]^{2+}$	751,9143	12
11	Oct-His9	$C_{69}H_{111}N_{19}O_{19}$	1509,8304	755,9232 $[M+2H]^{2+}$	755,9225	16
12	Oct-Dab9	$C_{67}H_{112}N_{18}O_{19}$	1472,8351	737,4254 $[M+2H]^{2+}$	737,4249	12

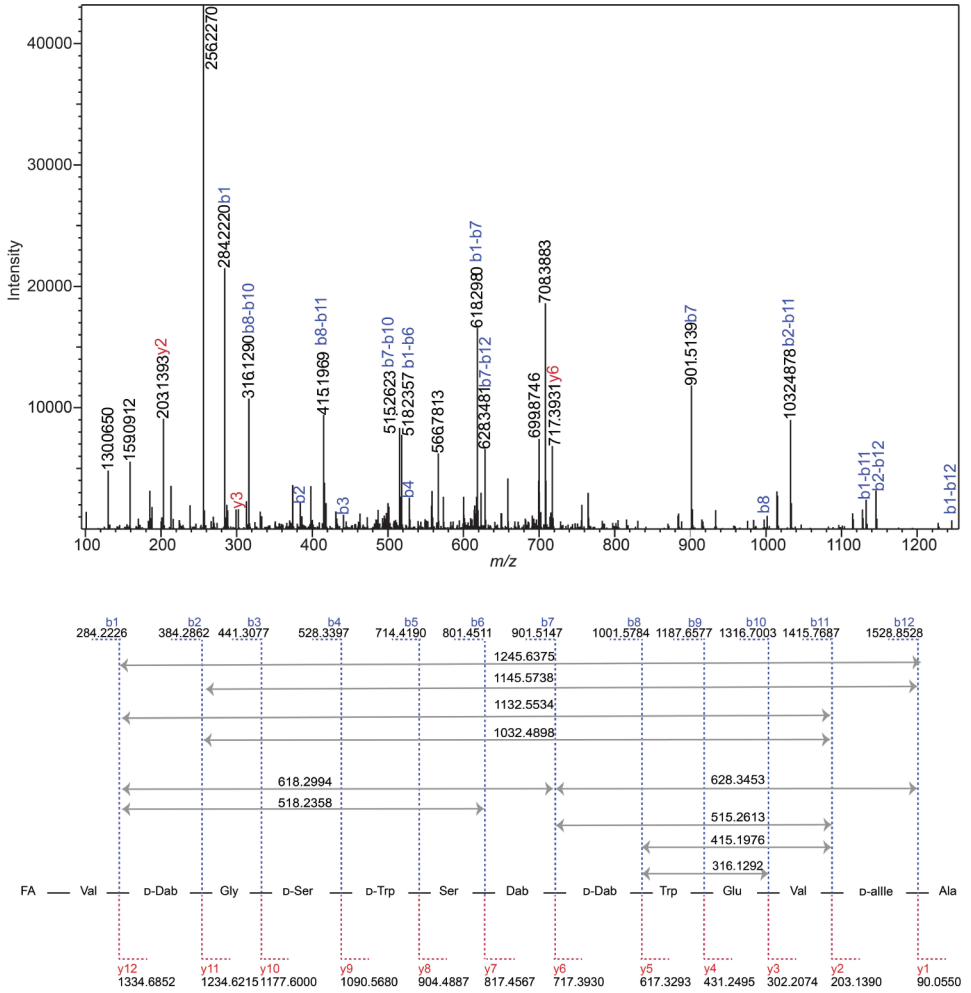


Figure S1. MS/MS spectrum of tridecaptin A_5 produced by of *Paenibacillus* sp. JJ-21 (precursor ion $[M+3H]^{3+}$ *m/z* 539.9712). The assignment of the sequence of amino acid residues at the top of the spectrum is based on the mass differences between the consecutive *y* and *b* ions.

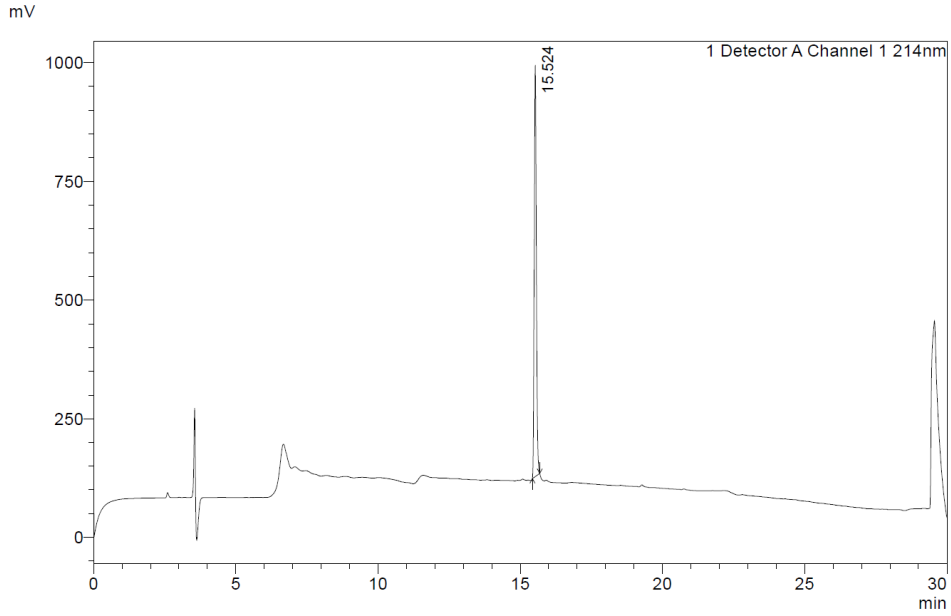


Figure S2. HPLC trace showing the reinjection of purified Oct-TriA$_1$ (1). The peptide eluted as a single peak at 15.524 min using the HPLC method A.

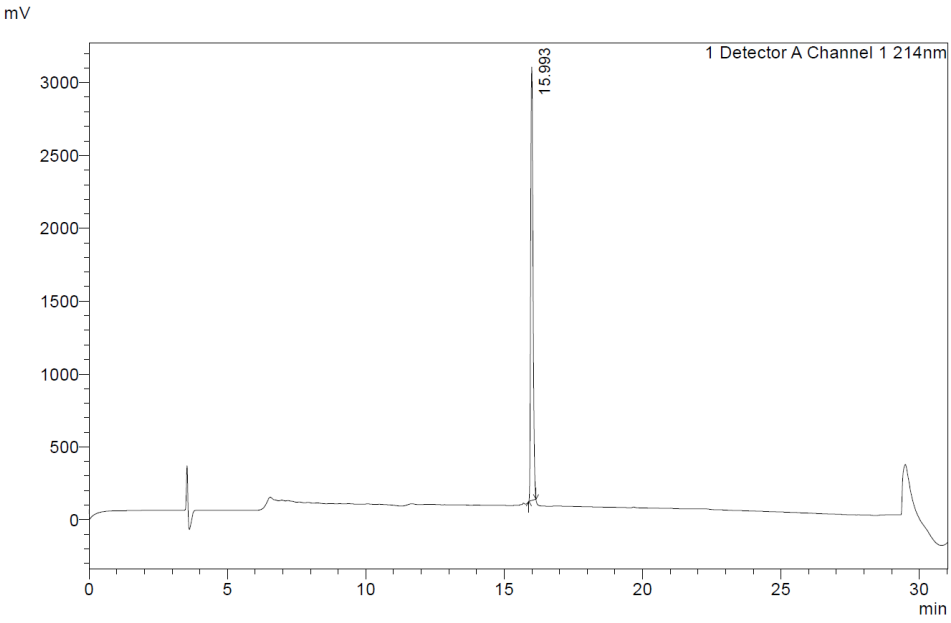


Figure S3. HPLC trace showing the reinjection of purified Oct-TriA5 (2). The peptide eluted as a single peak at 15.993 min using the HPLC method A.

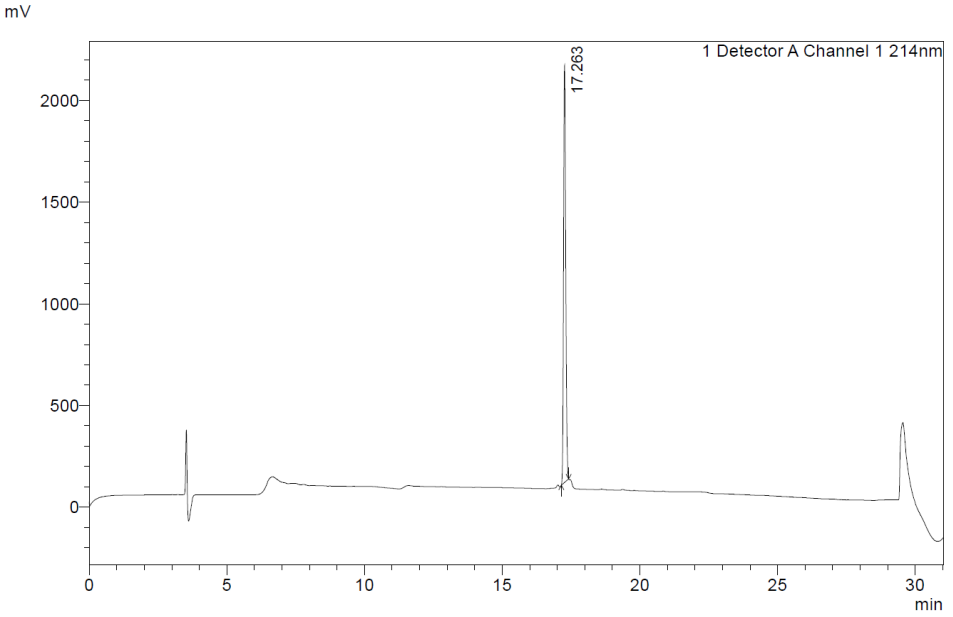


Figure S4. HPLC trace showing the reinjection of purified Oct-TriD (3). The peptide eluted as a single peak at 17.263 min using the HPLC method A.

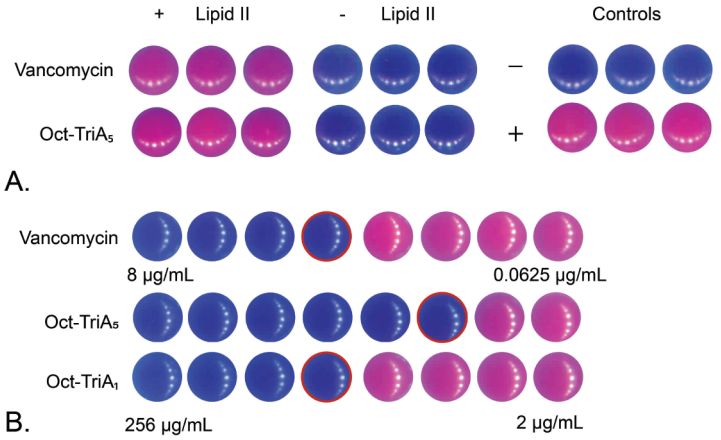


Figure S5. A. Gram-positive lipid II binding assays with *S. aureus* USA300 (MRSA) to show lipid II binding of Oct-TriA$_5$. Vancomycin was used as a positive control for lipid II binding. Resazurin was used to visualize the viability of the indicator strain (pink colour indicates growth and blue means inhibition of growth). Note, that addition of lipid II significantly reduced the efficacy of Oct-TriA$_5$ and vancomycin, as the growth of *S. aureus* USA300 (MRSA) was not inhibited at the concentration of 8x MIC. This indicates that Oct-TriA$_5$ binds to lipid II of Gram-positive bacteria. **B.** MIC assays for Oct-TriA$_1$, Oct-TriA$_5$ and vancomycin against *S. aureus* USA300. The wells with the MIC of the test compounds are highlighted with red edge.

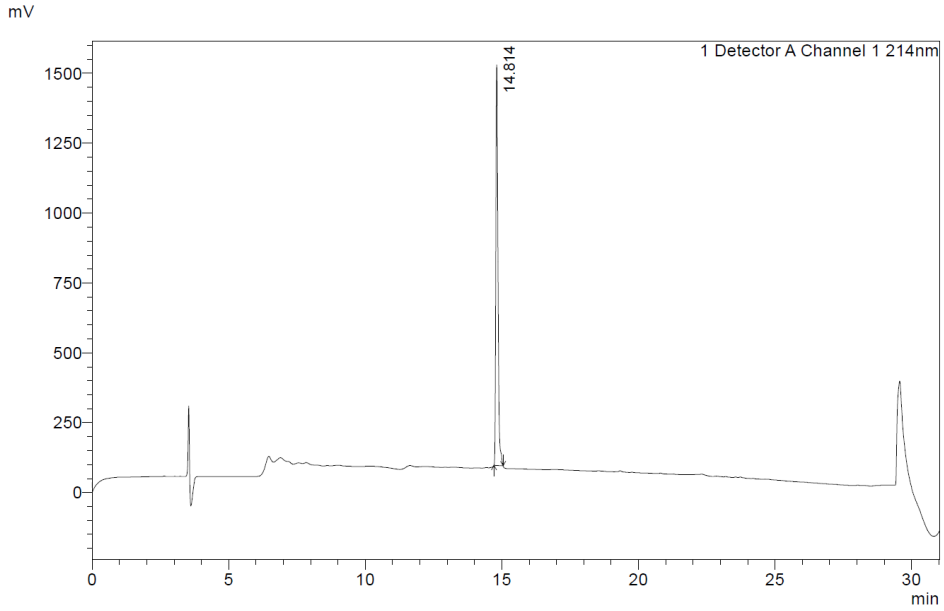


Figure S6. HPLC trace showing the reinjection of purified Oct-Gly9 (4). The peptide eluted as a single peak at 14.814 min using the HPLC method A.

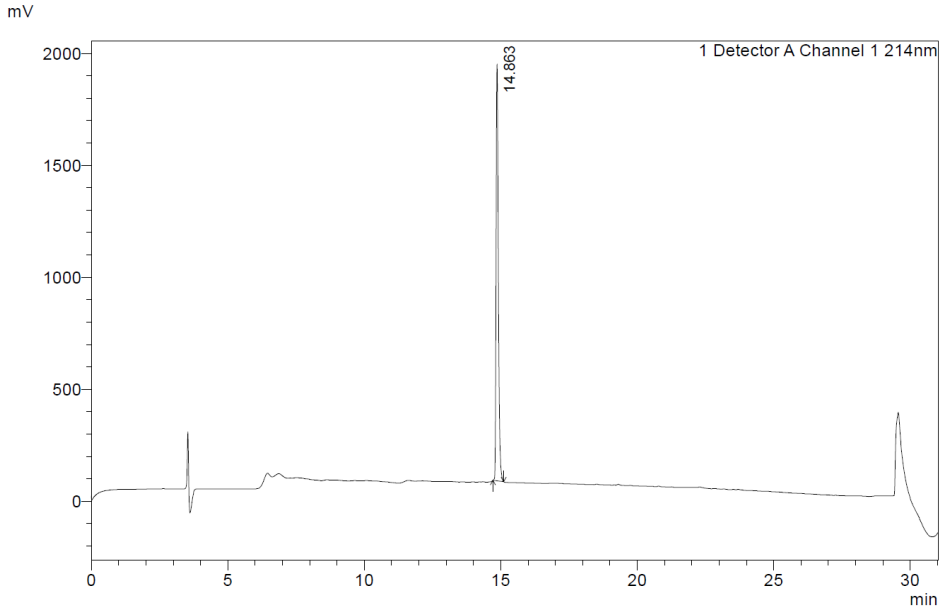


Figure S7. HPLC trace showing the reinjection of purified Oct-Ala9 (5). The peptide eluted as a single peak at 14.863 min using the HPLC method A.

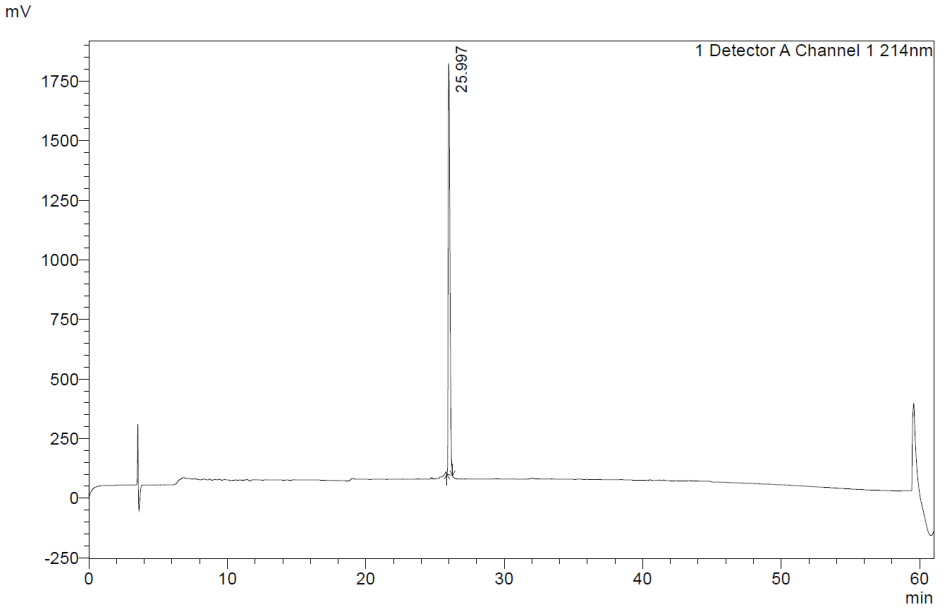


Figure S8. HPLC trace showing the reinjection of purified Oct-Val9 (6). The peptide eluted as a single peak at 25.997 min using the HPLC method C.

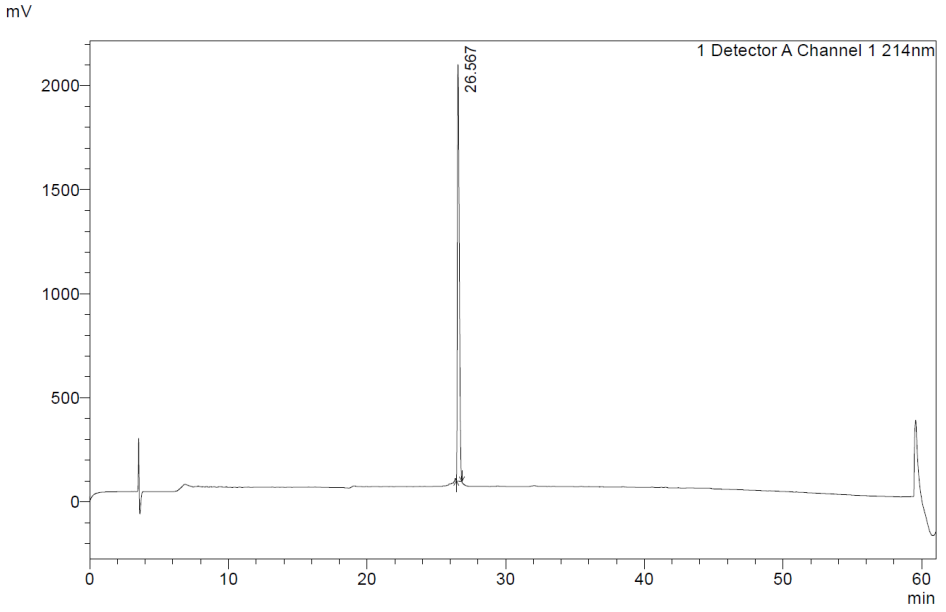


Figure S9. HPLC trace showing the reinjection of purified Oct-Ile9 (7). The peptide eluted as a single peak at 26.567 min using the HPLC method C.

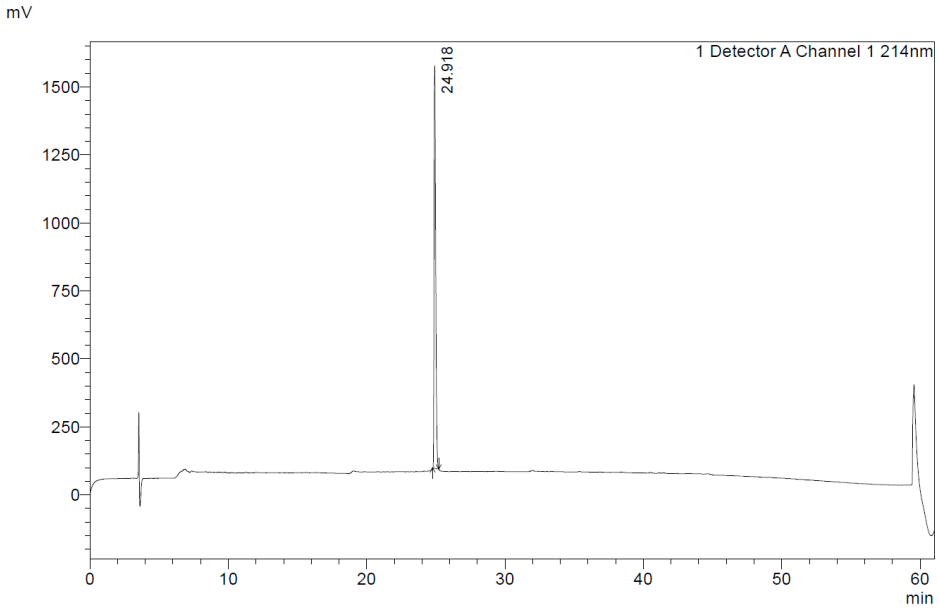


Figure S10. HPLC trace showing the reinjection of purified Oct-Ser9 (8). The peptide eluted as a single peak at 24.918 min using the HPLC method C.

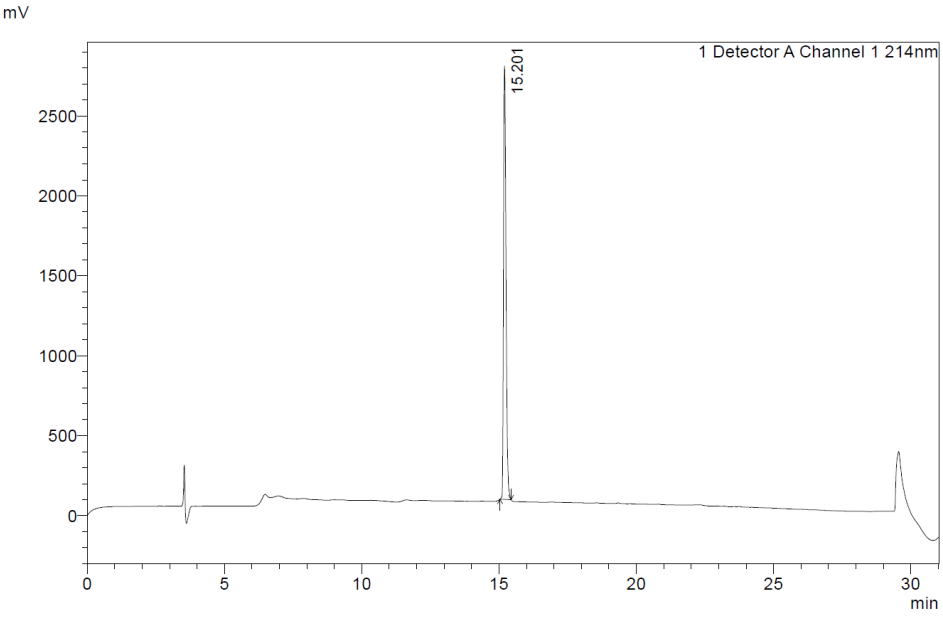


Figure S11. HPLC trace showing the reinjection of purified Oct-Tyr9 (9). The peptide eluted as a single peak at 15.201 min using the HPLC method A.

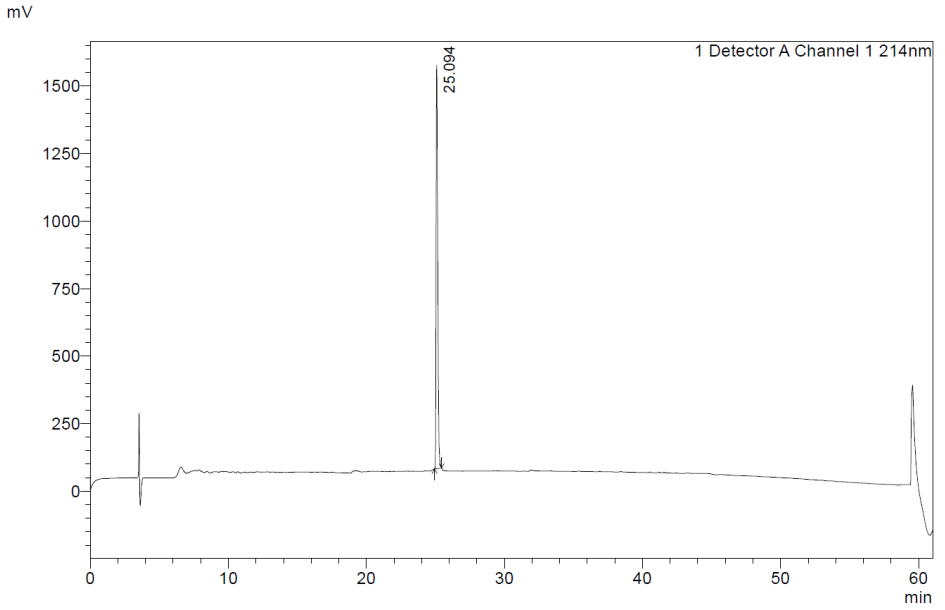


Figure S12. HPLC trace showing the reinjection of purified Oct-Glu9 (10). The peptide eluted as a single peak at 25.094 min using the HPLC method C.

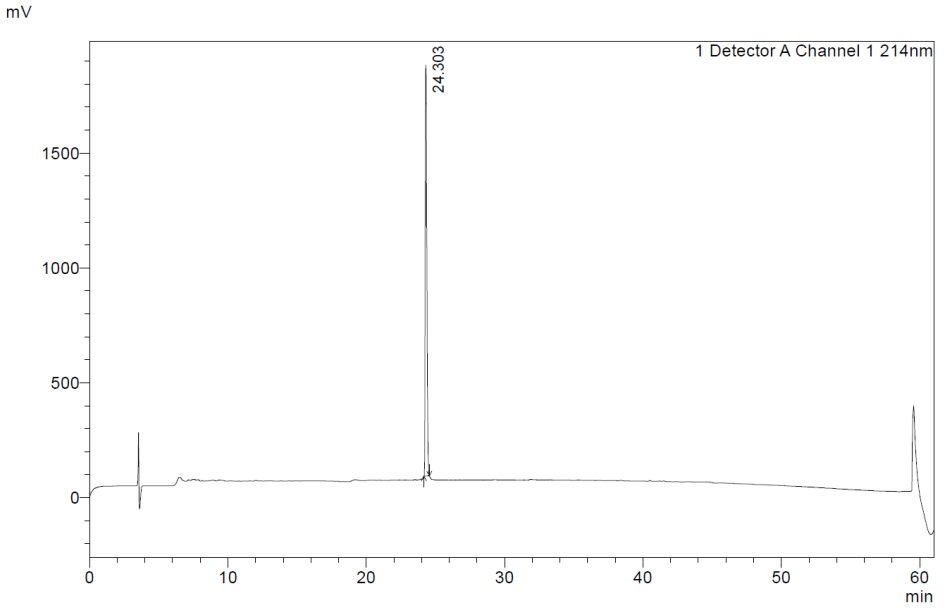


Figure S13. HPLC trace showing the reinjection of purified Oct-His9 (11). The peptide eluted as a single peak at 24.303 min using the HPLC method C.

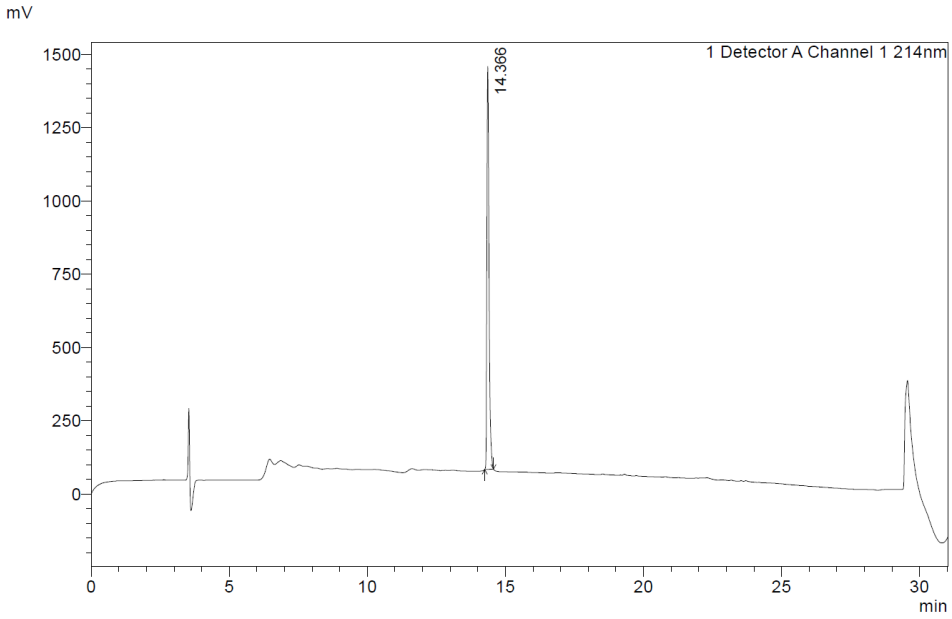


Figure S14. HPLC trace showing the reinjection of purified Oct-Dab9 (12). The peptide eluted as a single peak at 14.366 min using the HPLC method A.