

 Open access • Posted Content • DOI:10.1101/2020.05.06.080655

Synthetic Zippers as an Enabling Tool for Engineering of Non-Ribosomal Peptide Synthetases — [Source link](#)

Kenan A. J. Bozhueyuek, Jonas Watzel, Nadya Abbood, Helge B. Bode

Institutions: Goethe University Frankfurt

Published on: 06 May 2020 - bioRxiv (Cold Spring Harbor Laboratory)

Related papers:

- [Synthetic Zippers as an Enabling Tool for Engineering of Non-Ribosomal Peptide Synthetases*](#).
- [Tailored peptide synthetases, a method for their production and the use thereof](#)
- [Artificial non-ribosomal peptide synthetases](#)
- [Modification and de novo design of non-ribosomal peptide synthetases \(NRPS\) using specific assembly points within condensation domains](#)
- [Modification and de novo design of non-ribosomal peptide synthetases using specific assembly points within condensation domains.](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/synthetic-zippers-as-an-enabling-tool-for-engineering-of-non-25gm8s4a40>

1 **Synthetic Zippers as an Enabling Tool for Engineering of Non-Ribosomal** 2 **Peptide Synthetases**

3 Kenan A. J. Bozhueyuek^{1,†}, Jonas Watzel^{1,†}, Nadya Abbood^{1,†}, Helge B. Bode^{1,2,3,*}

4

5 1 Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt,
6 60438, Frankfurt am Main, Germany.

7 2 Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt,
8 60438, Frankfurt am Main, Germany.

9 3 Senckenberg Gesellschaft für Naturforschung, 60325, Frankfurt am Main, Germany

10 † equal contribution

11

12 * Corresponding author

13

14 **Abstract**

15 Non-ribosomal peptide synthetases (NRPSs) are the origin of a wide range of natural
16 products, including many clinically used drugs. Engineering of these often giant
17 biosynthetic machineries to produce novel non-ribosomal peptides (NRPs) at high titre
18 is an ongoing challenge. Here we describe a strategy to functionally combine NRPS
19 fragments of Gram-negative and -positive origin, synthesising novel peptides at titres
20 up to 290 mg l⁻¹. Extending from the recently introduced definition of eXchange Units
21 (XUs), we inserted synthetic zippers (SZs) to split single protein NRPSs into up to three
22 independently expressed and translated polypeptide chains. These synthetic type of
23 NRPS (type S) enables easier access to engineering, overcomes cloning limitations,
24 and provides a simple and rapid approach to building peptide libraries via the
25 combination of different NRPS subunits.

26

27

28 **One Sentence Summary:**

29 Divide and Conquer: A molecular tool kit to reprogram the biosynthesis of non-
30 ribosomal peptides.

31

32

33 Introduction

34 Non-ribosomal peptide synthetases (NRPSs) are multifunctional enzymes, producing
35 a broad range of structural diverse and valuable compounds with diverse applications
36 in medicine and agriculture (1) making them key targets for bioengineering. The
37 structural diversity of non-ribosomal peptides (NRPs) arises from the assembly line
38 architecture of their biosynthesis. According to their biosynthetic logic, known NRPS
39 systems are classified into three groups, linear (type A), iterative (type B), and
40 nonlinear NRPSs (type C) (2). Type A NRPSs are composed of sequential catalytically
41 active domains organised in modules, each responsible for the incorporation and
42 modification of one specific amino acid (AA). The catalytic activity of a canonical
43 module is based upon the orchestrated interplay of an adenylation (A) domain for AA
44 selection and activation, a condensation (C) domain to catalyse peptide bond
45 formation, and a thiolation/ peptidyl-carrier protein (T) onto which the AAs or
46 intermediates are covalently tethered (3). In addition, tailoring domains, including
47 epimerization (E), methylation, and oxidation domains can be part of a module, or a
48 heterocyclization (Cy) domain instead of a C-domain can be present. Finally, most
49 NRPS termination modules harbour a TE-domain, usually responsible for the release
50 of linear, cyclic or branched cyclic peptides (4).

51 Type A NRPSs (Fig. 1a) follow the collinearity rule, *i.e.* the number of NRPS modules
52 corresponds directly to the number of monomers incorporated into the associated
53 product, and the arrangement of the modules directly follows the peptides primary
54 sequence (5). Whereas in *in cis* type A NRPSs all modules are arranged on a single
55 polypeptide chain (*e.g.* ACV-synthetase (6)), *in trans* assembly-lines comprise a
56 number of individual proteins (Daptomycin-synthetase (7)). Mutual protein-protein
57 interactions of the latter are mediated by specialized C- (donor) and N-terminally
58 (acceptor) attached ~30 AAs long α -helical structural elements, so called
59 communication mediating (COM) or docking domains (DDs) (8). DDs typically are
60 located in between two modules and only interact with weak affinities (4-25 μ M) (9–
61 13), but are crucial to ensure biosynthesis of the desired product(s) (8, 11, 14). Despite
62 recent progress on applying DD substitutions to program new assembly lines, in most
63 cases structural information is lacking to effectively apply DDs for general engineering
64 purposes (11, 15, 16).

65 Although early engineering attempts, including the exchange of DDs, the targeted
66 modification of the A-domains substrate specificity conferring AA residues, and the
67 substitution of domains as well as whole modules, gave mixed results, several notable
68 advances have been published recently (16–18). To give but one example, we
69 comprehensively analysed structural data as well as inter-domain linkers in NRPSs to
70 define novel fusion sites and to provide guidelines for exchanging A-T-C units, denoted
71 as eXchange Units (XUs), as opposed to canonical modules (C-A-T) (19). By
72 combining XUs from 15 NRPSs *in cis*, it was possible to reconstitute naturally available
73 peptides, peptide derivatives, and to generate new-to-nature peptides *de novo* in high
74 yields.

75 Herein, starting from our recently published XU concept, we explored the ability of
76 synthetic zippers (SZs (20)) to manipulate collinear type A NRPSs by introducing
77 artificial *in trans* regulation. SZs interact with high affinity (KD<10 nM) via a coiled-coil
78 structural motif, enabling the specific association of two proteins. Such a strategy not
79 only would allow creating a synthetic type of *in trans* regulated mega-synthetases
80 (type S), by combining NRPSs with high-affinity SZs (20) (Fig. 1a), but to overcome
81 cloning and protein size limitations associated with heterologous NRP production.

82 Results

83 In depth structural analysis of the crystallised termination module SrfA-C (PDB-ID:
84 2VSQ) suggested splitting NRPSs in between consecutive XUs at the previously
85 defined W]-[NATE motif of the conformationally flexible C-A linker (21–23) region. As
86 already known (21, 22), this splicing position bears several advantages. Of particular
87 importance is that it keeps intact the short (~ 10 AAs) α -helical structure at the C-
88 terminus of the resulting truncated protein (**subunit 1**) – as in wild type (WT) NRPSs
89 this helical structure not only regulates the C-A distance throughout the catalytic cycle
90 (21), but also associates with the A-domains hydrophobic protein surface (23).

91 Attempts of *in silico* creating NRPS domains connected via SZs, composed of ~40
92 amino acids (AAs), were unsuccessful. Nevertheless, careful revision of available
93 structural data indicated that ~10 AAs from the unstructured N-terminus of **subunits 2**
94 must be removed to meet the distance-criteria set out by the WT C-A inter-domain
95 linker to ensure correct C-A di-domain contacts before SZs N-terminally could be
96 introduced (Fig. 1b). After perusing characterized SZ pairs (20), to begin with we chose
97 the SZ pair 17 & 18 (Fig. 1c & d).

98 **Proof of Concept**

99 To assess the general suitability of SZ-pairs to *in trans* connect two NRPS proteins
100 and mediate biosynthetically functional protein-protein interface interactions, we
101 targeted the xenotetrapeptide (**1**) producing NRPS (XtpS; Fig. S1) from the Gram-
102 negative entomopathogenic bacterium *Xenorhabdus nematophila* HGB081 (24). We
103 decided to split XtpS into two subunits in between XUs 2 and 3 and four artificial two
104 component type S NRPS (Fig. 2a) were constructed and heterologously produced in
105 *E. coli* DH10B::*mtaA* (25) – either with SZs fused to both subunits (NRPS-1: subunit 1-
106 SZ17, SZ18-subunit 2); only fused to subunit 1 (NRPS-2: subunit 1-SZ17, subunit 2)
107 or subunit 2 (NRPS-3: subunit 1, SZ18-subunit 2), and without SZs (NRPS-4: subunit
108 1, subunit 2).

109 NRPS-2 and NRPS-4 showed no detectable peptide production, whereas NRPS-1
110 lead to the production of **1** with ~30% (28 mg l⁻¹) yield compared to WT XtpS (Fig. 2a,
111 Fig. S2), confirming that SZs indeed can be used to functionally mediate new-to-nature
112 *in trans* regulation of NRP biosynthesis. Interestingly, NRPS-3 with SZ18 fused to
113 subunit 2, but lacking SZ17 on subunit 1, showed moderate yields of **1**. Despite lacking
114 SZ17, the C-terminus of XtpS subunit 1 is forming a Leucine rich α -helical structure
115 (PDB-ID: 2VSQ) that might be able to interact with SZ18 of subunit 2 and mediate an
116 impaired but catalytically active C-A interface (21–23, 26).

117 Additionally, SZ17:18 were used to split the GameXPepptide A-D (**2-5**) producing NRPS
118 (GxpS (27, 28)) and the recombinant thiazole-peptide (**6**) producing NRPS (RtpS (29)).
119 Whereas GxpS originates from the Gram-negative bacterium *Photorhabdus*
120 *luminescens* TTO1, RtpS was constructed previously (29) from building blocks (BBs)
121 of Gram-positive origin (using NRPSs for the production of bacitracin (30) and surfactin
122 (31)). Both resulting type S NRPSs (Fig. 2b) showed good to very good titres of desired
123 peptides. NRPS-5 produced **2** (Fig. S3) with yields of ~64 % (4.9 mg l⁻¹) compared to
124 WT GxpS and NRPS-6 produced **6** (Fig. S4) at WT RtpS level (~20 mg l⁻¹).

125 All product structures and yields were confirmed by tandem mass spectrometry
126 (MS/MS) analysis and comparison of the retention times with synthetic standards.

127 **Creating Synonymous Chimeras**

128 To explore the recombination potential of chimeric type S NRPSs, initially we co-
129 expressed non-cognate subunits from NRPS-1 and -5 (Fig. 3a). Both, NRPS-1

130 subunit 1 and NRPS-5 subunit 1, largely possess synonymous A-domain (Val, Leu)
131 and C-domain (hydrophobic AAs) specificities, preventing potential upstream C- and
132 downstream TE-domain substrate specificity issues (22).

133 NRPS-7 produced **1** at same titres as NRPS-1 with 25.2 % (~23 mg l⁻¹) yield compared
134 to WT XtpS (Fig. S5). NRPS-8 showed 4- to 5-fold increased productivity compared to
135 NRPS-5 and even exceeded WT GxpS in yields of **2** and **4** with 170 % (~9 mg l⁻¹) and
136 80 %, respectively (Fig. S6). Like observed in our previous work (22, 29) the formal
137 exchange of the promiscuous XU1 from GxpS (for Val/Leu) against the Val-specific
138 XU1 from XtpS led to the exclusive production of **2** and **4**, without synthesis of **3** and **5**
139 observed in the original GxpS. Increased peptide yields of NRPS-8 compared to its
140 WT GxpS counterpart currently cannot be explained but were described before (29).

141 **Reprogramming Type S NRPS**

142 Major drawbacks of past reprogramming attempts have been (I) the incompatibility of
143 bacterial NRPS BBs from Gram-positive and -negative origin (22, 29) as well as (II) the
144 C-domains' specificity rule (22); both (I + II) severely limiting combinatorial space. With
145 tools to functionally split NRPSs at hand, we explored if type S NRPS have the potential
146 to overcome these biosynthetic bottlenecks.

147 To address (I), subunits of Gram-negative (NRPS-1 & -5) and -positive (NRPS-6) origin
148 were co-expressed (Fig. 3b) and culture extracts analysed via HPLC-MS (Fig. S7-9).
149 Three out of four resulting NRPSs (NRPS-9 – 12) not only showed detectable, but very
150 good peptide titres up to ~290 mg l⁻¹ (NRPS-9, **10**). Yet, all active type S NRPSs
151 showed unexpected peptide production profiles. For instance, NRPS-10 only produced
152 trace levels of the expected linear penta-peptide (**8**, linear vLL/L; D-AAAs in italics and
153 lowercase throughout this work) but produced yields of ~236 mg l⁻¹ and ~95 mg l⁻¹ of
154 linear vLL/ (**10**) and /LL/ (**11**), respectively. NRPS-9 and -10 not only constitute two
155 polypeptide chains linked via SZs, but the termination-module also interacts *in trans*
156 via natural DDs with subunit 2. Due to observed very high catalytic activities of NRPS-
157 9 and -10, along with SZs interacting with much higher affinities than DDs, 'auto-
158 catalytic' offloading might be catalysed by the E-domain present at the C-termini of
159 respective subunits 2 as it was recently reported by non-cognate termination domains
160 of chimeric NRPSs utilising internal C-domains (22).

161 When NRPS-11 and -12 were compared, the latter did not produce any detectable
162 peptide amounts, likely being an issue of TE-domain substrate specificity of XtpS (*c.f.*
163 Fig. S10), NRPS-11 produced the linear thiazole containing peptide IC*//L (**12**; ~53 mg
164 l⁻¹; Fig. S9). In its natural NRPS context as well as *in vitro*, the A3-domain of GxpS
165 prefers Phe over Leu (29). In case of NRPS-11, the terminal C-domain of subunit 1,
166 expecting Leu at its acceptor site, either prevents the incorporation of Phe due to its
167 gatekeeping activity or rather fine tunes the downstream A-domain specificity. Similar
168 effects of engineered NRPSs, exhibiting chimeric C-A interfaces (22, 32) or C-domains
169 (29), have been described. Aforementioned results (NRPS-9 – 12, Fig. 3b) represent
170 the first successful *in trans* as well as the first efficient strategy to recombine Gram-
171 positive and -negative NRPS BBs, producing peptides at an industrially relevant scale.

172 Tackling (II), subunits of NRPSs with non-synonymous specificities were co-expressed
173 with subunit 2 of NRPS-1 as well as NRPS-5, either respecting the C-domains'
174 specificity rule (NRPS-13 and 14, Fig. 3c) or not (NRPS-15 and 16, Fig. 3d; NRPS-17
175 – 19, Fig. S10). As subunits 1, XUs1-2 of NRPSs producing ambactin (AmbS (25)),
176 szentiamide (SzeS (22)), bicornutin (BicA (33)), and xenolindicin (XldS (25)) were
177 used.

178 Both type S NRPS (NRPS-13: AmbS subunit 1 + GxpS subunit 2; and NRPS-14: SzeS
179 subunit 1 + GxpS subunit 2) respecting C-domains' acceptor site specificities (Phe)
180 produced the desired derivatives (Fig. 3c). NRPS-13 produced **13** and **14** with yields
181 of ~218 and ~46 mg l⁻¹, respectively (Fig. S11). In addition, peptide yields of **15** (~24
182 mg l⁻¹) and **16** (~4 mg l⁻¹) from NRPS-14 were even higher compared to a homologous
183 *in cis* NRPS (**15**: ~15 mg l⁻¹; **16**: ~2 mg l⁻¹) constructed from the same BBs in an earlier
184 study (22) (Fig. S12 & S13). Peptides **13/14** and **15/16**, respectively, only differ in Leu
185 or Phe at position 3 from the relaxed substrate specificity of XU3 from GxpS (27, 28).

186 NRPS-15 (BicA subunit 1 + GxpS subunit 2) and -16 (XldS subunit 1 + GxpS subunit
187 2), not complying with the XUs' specificity rules (Fig. 3d), produced peptides **17** (~40
188 mg l⁻¹, Fig. S14) and **18-21** (0.1 – 5.5 mg l⁻¹, Fig. S15), respectively. The latter peptides
189 (**18-21**) only differ in the *N*-terminal acyl starter unit, originating from the *E. coli* fatty-
190 acid pool, as also observed in the original xenolindicins (25). Especially NRPS-15 was
191 expected to be inactive, as previous studies have shown that the BicA C3-domains
192 acceptor site is highly specific for Arg and cannot process Phe or Leu when covalently
193 fused to subunit 2 (29). This might indicate that splitting *in cis* NRPSs in between C-

194 and A-domains potentially decreases C-domains' acceptor site specificity by
195 introducing more geometric flexibility and minimizing potentially restrictive effects on
196 A-domain movements (32) as supported by a recently published study suggesting that
197 C-domains indeed do not exhibit intrinsic substrate specificities (34).

198 In contrast, all type S NRPS (NRPS-17 – 19, and NRPS-12) sharing subunit 2 from
199 XtpS (NRPS-1), also not complying with the C-domain specificity rule (22, 29), failed
200 to produce detectable amounts of any peptide (Fig. S10). The reason for this might be
201 the TE-domains high specificity for peptide length and amino acid composition.

202 **Unpaired Activity of GxpS Subunit 2**

203 All type S NRPS split in between C-A domains and sharing GxpS subunit 2 (NRPS-5,
204 -8, -11, and NRPS-13 – 16) showed an unexpected behaviour, producing a range of
205 tripeptides (**33/34** and **35/36**) at high titre up to 86 mg l⁻¹ related to the unpaired activity
206 of GxpS subunit 2 (Fig. S16). Due to the promiscuous GxpS A3-domain, **33/34** and
207 **35/36** differ from each other at position one, either carrying Phenylalanine or Leucine.
208 In addition, **33** and **35** show a *D-D-L* configuration, whereas **34** and **36** have a *L-D-L*
209 configuration.

210 **Optimization and Extending Functionality**

211 To explore the optimization potential of SZs we not only successfully applied parallel
212 interacting (Fig. 1c) SZs 19 & 18 (Fig. 4a & S17a, NRPS-20), but also optimized
213 SZ17:18 interactions. Thus, to introduce more spatial freedom potentially enhancing
214 flexible domain-domain interactions, synthetic stretches of Gly-Ser (GS) varying in
215 length of 4-10 AAs were introduced in between the C-terminus of XtpS (NRPS-1)
216 subunit 1 and SZ17 (Fig. 4b). All resulting chimeric NRPSs (NRPS-21 - 23) showed
217 ~3x increased yields of **1** (Fig. S17b) compared to NRPS-1, raising titres back at WT
218 level (~220 mg l⁻¹) – and therefore indicating that introduced GS linkers had the desired
219 effect.

220 To further explore and extend functionality of type S NRPSs, we again targeted XtpS
221 introducing SZs within the T-C (NRPS-24, Fig. S18a) and A-T linker-regions (NRPS-
222 25, Fig. S18b). Both, NRPS-24 and -25 (Fig. 4c & S17c) synthesised **1** with titres at
223 ~86 % (174 mg l⁻¹) compared to WT XtpS level. While catalytic activity of NRPS-24
224 was expected, as the introduced SZs are mimicking natural DDs (35), NRPS-25,
225 showing an unusual split between A2-T2, truly represents a type S NRPS. The A-T

226 linker sequence, consisting of ~15 AAs, represents the shortest inter-domain linker in
227 the context of NRPS elongation modules, which is conformationally also the most
228 flexible one. Structural insights indicate that T and A_{sub} domains adopt alternative
229 conformations to shuffle reaction intermediates among catalytic domains (23). Thus, it
230 was expected that the additional rigidity, inserted by the structured α -helical AA
231 stretches, would result in loss of function. Structures of large constructs of the linear
232 gramicidin synthesising NRPS (LgrA) (35) show a very high structural flexibility,
233 potentially bringing closely together domains that are far apart in protein sequence and
234 therefore facilitating synthetic cycles with inserted tailoring domains, unusual domain
235 arrangements like A-C-T (36), module skipping (37) and presumably also SZs.

236 More dipartite type S NRPSs (NRPS-44 - 48), split in between and within modules are
237 depicted in Fig. S19 and S20.

238 **Tripartite NRPSs**

239 SZs mediated reprogramming of NRPSs, other than ‘simply’ splitting *in cis* NRPSs in
240 two, makes it necessary to express the proteins in three parts – *i.e.* being able to target
241 one specific position of the synthesised peptide. Therefore, again starting with XtpS,
242 we created three orthogonal interaction networks (Fig. 1c) by introducing the anti-
243 parallel and parallel interacting SZ pairs SZ17:18 and SZ1:2 in between XU2-3 and
244 XU3-4 (NRPS-26), in between module 2-3 and module 3-4 (NRPS-27), as well as
245 within the A-T linker regions of module 2 and 3 (NRPS-28), respectively (Fig. 5).

246 All resulting tripartite type S NRPSs (NRPS-26 – 28, Fig. 5 & S17d) produced **1** with
247 good yields of 17 – 70% (24 – 104 mg l⁻¹) compared to WT XtpS but with decreased
248 yields compared to their dipartite counterparts (NRPS-21, -24, & -25).

249 In order to demonstrate the potential of artificial *in trans* NRPSs, we designed and
250 cloned a small library of type S BBs (Fig. 6a), placing SZ17:18 and SZ1:2 within the A-
251 T linker regions to perform co-expression experiments (Fig. 6) in a quick plug-and-play
252 manner. The A-T linker region was targeted because in this case C-domain specificities
253 presumably do not represent a limitation of recombination and because T-C-A tri-
254 domains as a catalytically active unit to reprogram NRPSs are under underrepresented
255 (38, 39).

256 In brief, the created plasmid library, expressing 11 different type S NRPS BBs from
257 XtpS, GxpS, SzeS, XldS, and the gargantuanin producing synthetase (GarS). Overall

258 15 (NRPS-28 – 42) from 22 co-expressions of three plasmids each yielded detectable
259 amounts (0.1 – 38 mg l⁻¹) of 16 different peptides, 11 of which were new (Fig. 6, S21 –
260 S38). Despite the method's general simplicity, the overall efficacy or recombination
261 potential of T-C-A units compared to XUs appears to be slightly more restricted. For
262 example, neither co-expression of all type S BBs to reconstitute SzeS, nor any
263 combination involving the Ser and Thr specifying BBs from XldS and GarS, yielded any
264 detectable peptide, respectively. These results probably indicate an incompatibility of
265 formed chimeric A-T interfaces or substrate incompatibilities at the respective C-
266 domains donor site. Yet, in light of previous results (40, 41) concerning C-domain
267 specificities, the latter seems to be unlikely.

268 **Conclusion**

269 Recently the successful application of SZs to replace naturally present DDs in
270 polyketide synthases (PKS) as a tool to create chimeric PKSs was published (42). Here
271 we reported the use of high-affinity SZs to split native single protein NRPSs into two
272 and three individual components. Generating artificially *in trans* regulated assembly
273 lines not only represents a new NRPS architecture, referred to as type S, but emerges
274 to be highly productive with yields comparable to WT levels. Especially the efficient
275 combination of NRPS BBs of Gram-negative and -positive origin greatly expands the
276 combinatorial space for reprogramming pharmaceutically relevant entities or creating
277 diverse NRP libraries.

278 Observed accumulation of side products may be a drawback, especially for industrial
279 production purposes, but also indicates the high productivity potential of type S NRPS.
280 For example, aggregated peptide yields of tetra- and penta-peptides produced by
281 NRPS-10 sums up to ~330 mg l⁻¹, whereas the WT counterparts in total “only”
282 produced ~12 mg l⁻¹ (GxpS) and ~22 mg l⁻¹ (RtpS), respectively. Increased total peptide
283 yields, *i.e.* for type S NRPS split in between XUs, might be due to higher catalytic
284 activities of C- and A-domains involved in forming the SZs mediated chimeric C-A
285 interface, as geometric restrictions and feedback mechanisms of the catalytic cycle
286 may have been suppressed.

287 Although it was not possible to insert SZs in between the *N*- and *C*-lobe of C-domains
288 (Fig. S19c), following our recently published XUC concept (29), in principle it was
289 possible to introduce SZs in between any di-domain (A-T, C-A, T-C). Having SZs at
290 hand, not only peptide libraries quickly can be constructed with high success rates, but

291 now it also should be conceivable to combine different biosynthetic pathways *in situ*,
292 by introducing SZs at the genomic level – *i.e.* applying CRISPR/Cas9 based genetic
293 engineering. We are convinced that further research into this direction, like elucidating
294 structures of SZ connected NRPS domains, eventually will bring up even more
295 versatile artificial DDs as it is already suggested by NRPS-21 – 23 constructed with an
296 synthetic (GS)_x C-A linker region.

297

298 **Acknowledgements**

299 This work was supported by the LOEWE research cluster MegaSyn funded by the state
300 of Hesse and an ERC advanced grant (grant agreement number 835108).

301

302 **Competing interests**

303 Goethe University filed a patent application for SZ technology in NRPSs. The patent is
304 currently pending.

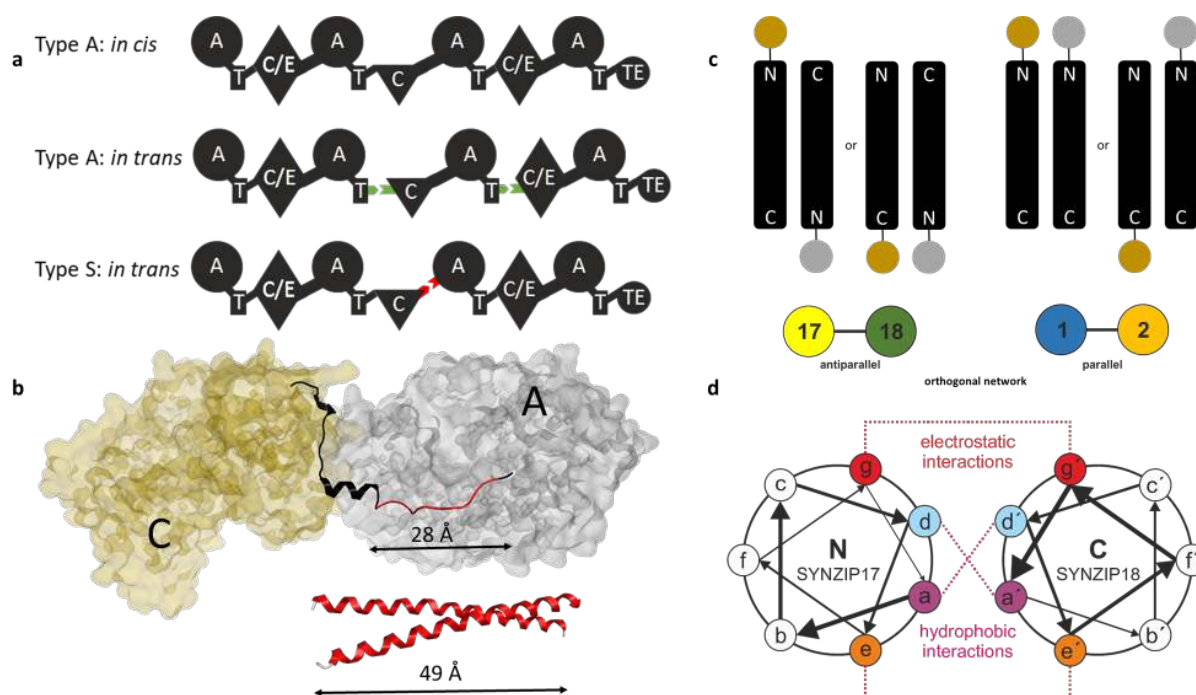
305

306 References

- 307 1. R. D. Süssmuth, A. Mainz, Nonribosomal Peptide Synthesis-Principles and
308 Prospects. *Angew. Chem., Int. Ed.* **56**, 3770–3821 (2017).
- 309 2. H. D. Mootz, D. Schwarzer, M. A. Marahiel, Ways of assembling complex natural
310 products on modular nonribosomal peptide synthetases. *ChemBioChem.* **3**, 490
311 (2002).
- 312 3. S. A. Sieber, M. A. Marahiel, Molecular mechanisms underlying nonribosomal
313 peptide synthesis: approaches to new antibiotics. *Chem. Rev.* **105**, 715–738
314 (2005).
- 315 4. F. Kopp, M. A. Marahiel, Macrocyclization strategies in polyketide and
316 nonribosomal peptide biosynthesis. *Nat. Prod. Rep.* **24**, 735–749 (2007).
- 317 5. K. A. Bozhüyük, J. Micklefield, B. Wilkinson, Engineering enzymatic assembly
318 lines to produce new antibiotics. *Curr. Opin. Microbiol.* **51**, 88–96 (2019).
- 319 6. M. F. Byford, J. E. Baldwin, C.-Y. Shiau, C. J. Schofield, The mechanism of ACV
320 synthetase. *Chem. Rev.* **97**, 2631–2650 (1997).
- 321 7. R. H. Baltz, V. Miao, S. K. Wrigley, Natural products to drugs: daptomycin and
322 related lipopeptide antibiotics. *Nat. Prod. Rep.* **22**, 717–741 (2005).
- 323 8. M. Hahn, T. Stachelhaus, Selective interaction between nonribosomal peptide
324 synthetases is facilitated by short communication-mediating domains. *Proc. Natl.
325 Acad. Sci. U. S. A.* **101**, 15585–15590 (2004).
- 326 9. J. R. Whicher *et al.*, Cyanobacterial polyketide synthase docking domains: a tool
327 for engineering natural product biosynthesis. *Chem. Biol.* **20**, 1340–1351 (2013).
- 328 10. J. Dorival *et al.*, Characterization of intersubunit communication in the
329 virginiamycin trans-acyl transferase polyketide synthase. *J. Am. Chem. Soc.* **138**,
330 4155–4167 (2016).
- 331 11. C. Hacker *et al.*, Structure-based redesign of docking domain interactions
332 modulates the product spectrum of a rhabdopeptide-synthesizing NRPS. *Nat.
333 Commun.* **9**, 4366.
- 334 12. T. J. Buchholz *et al.*, Structural basis for binding specificity between subclasses of
335 modular polyketide synthase docking domains. *ACS Chem. Biol.* **4**, 41–52 (2009).
- 336 13. J. Watzel, C. Hacker, E. Duchardt-Ferner, H. B. Bode, J. Wöhnert, A new docking
337 domain type in the Peptide-Antimicrobial-Xenorhabdus peptide producing
338 nonribosomal peptide synthetase from *Xenorhabdus bovienii*. *ACS Chem. Biol.*
339 (2020).
- 340 14. C. Kegler, H. B. Bode, Artificial splitting of a non-ribosomal peptide synthetase by
341 inserting natural docking domains. *Angew. Chem., Int. Ed.* (2020).
- 342 15. X. Cai *et al.*, Entomopathogenic bacteria use multiple mechanisms for bioactive
343 peptide library design. *Nat. Chem.* **9**, 379 (2017).
- 344 16. X. Cai, L. Zhao, H. B. Bode, Reprogramming promiscuous nonribosomal peptide
345 synthetases for production of specific peptides. *Org. Lett.* **21**, 2116–2120 (2019).
- 346 17. H. Kries, D. L. Niquille, D. Hilvert, A subdomain swap strategy for reengineering
347 nonribosomal peptides. *Chem. Biol.* **22**, 640–648 (2015).
- 348 18. K. Zhang *et al.*, Engineering the substrate specificity of the DhbE adenylation
349 domain by yeast cell surface display. *Chem. Biol.* **20**, 92–101 (2013).
- 350 19. H. D. Mootz, M. A. Marahiel, Design and application of multimodular peptide
351 synthetases. *Curr. Opin. Biotechnol.* **10**, 341–348 (1999).
- 352 20. K. E. Thompson, C. J. Bashor, W. A. Lim, A. E. Keating, SYNZIP protein
353 interaction toolbox: *in vitro* and *in vivo* specifications of heterospecific coiled-coil
354 interaction domains. *ACS Synth. Biol.* **1**, 118–129 (2012).

- 355 21.E. J. Drake *et al.*, Structures of two distinct conformations of holo-non-ribosomal
356 peptide synthetases. *Nature*. **529**, 235 (2016).
- 357 22.K. A. J. Bozhüyük *et al.*, *De novo* design and engineering of non-ribosomal
358 peptide synthetases. *Nat. Chem.* **10**, 275 (2018).
- 359 23.A. Tanovic, S. A. Samel, L.-O. Essen, M. A. Marahiel, Crystal structure of the
360 termination module of a nonribosomal peptide synthetase. *Science*. **321**, 659–663
361 (2008).
- 362 24.C. Kegler *et al.*, Rapid determination of the amino acid configuration of
363 xenotetrapeptide. *ChemBioChem*. **15**, 826–828 (2014).
- 364 25.O. Schimming, F. Fleischhacker, F. I. Nollmann, H. B. Bode, Yeast homologous
365 recombination cloning leading to the novel peptides ambactin and xenolindicin.
366 *ChemBioChem*. **15**, 1290–1294 (2014).
- 367 26.K. Bloudoff, T. M. Schmeing, Structural and functional aspects of the
368 nonribosomal peptide synthetase condensation domain superfamily: discovery,
369 dissection and diversity. *Biochim. Biophys. Acta, Proteins Proteomics*. **1865**,
370 1587–1604 (2017).
- 371 27.H. B. Bode *et al.*, Determination of the absolute configuration of peptide natural
372 products by using stable isotope labeling and mass spectrometry. *Chemistry*. **18**,
373 2342–2348 (2012).
- 374 28.F. I. Nollmann *et al.*, Insect-specific production of new GameXPeptides in
375 *Photobacterium luminescens* TTO1, widespread natural products in
376 entomopathogenic bacteria. *ChemBioChem*. **16**, 205–208 (2015).
- 377 29.K. A. J. Bozhüyük *et al.*, Modification and *de novo* design of non-ribosomal
378 peptide synthetases using specific assembly points within condensation domains.
379 *Nat. Chem.* **11**, 653–661 (2019).
- 380 30.D. Konz, A. Klens, K. Schörgendorfer, M. A. Marahiel, The bacitracin biosynthesis
381 operon of *Bacillus licheniformis* ATCC 10716: molecular characterization of three
382 multi-modular peptide synthetases. *Chem. Biol.* **4**, 927–937 (1997).
- 383 31.P. Cosmina *et al.*, Sequence and analysis of the genetic locus responsible for
384 surfactin synthesis in *Bacillus subtilis*. *Mol. Microbiol.* **8**, 821–831 (1993).
- 385 32.S. Meyer *et al.*, Biochemical dissection of the natural diversification of microcystin
386 provides lessons for synthetic biology of NRPS. *Cell Chem. Biol.* **23**, 462–471
387 (2016).
- 388 33.S. W. Fuchs *et al.*, Neutral loss fragmentation pattern based screening for
389 arginine-rich natural products in *Xenorhabdus* and *Photobacterium*. *Anal. Chem.*
390 **84**, 6948–6955 (2012).
- 391 34.M. J. Calcott, J. G. Owen, D. F. Ackerley,
392 <https://www.biorxiv.org/content/10.1101/2020.02.28.970632v1> (2020).
- 393 35.J. M. Reimer *et al.*, Structures of a dimodular nonribosomal peptide synthetase
394 reveal conformational flexibility. *Science*. **366** (2019).
- 395 36.T. A. Keating, C. G. Marshall, C. T. Walsh, Reconstitution and characterization of
396 the *Vibrio cholerae* vibriobactin synthetase from VibB, VibE, VibF, and VibH.
397 *Biochemistry*. **39**, 15522–15530 (2000).
- 398 37.F. Yan *et al.*, Synthetic biology approaches and combinatorial biosynthesis
399 towards heterologous lipopeptide production. *Chem. Sci.* **9**, 7510–7519 (2018).
- 400 38.T. Duerfahrt, S. Doekel, T. Sonke, P. J. L. M. Quaedflieg, M. A. Marahiel,
401 Construction of hybrid peptide synthetases for the production of alpha-l-aspartyl-l-
402 phenylalanine, a precursor for the high-intensity sweetener aspartame. *Eur. J.*
403 *Biochem.* **270**, 4555–4563 (2003).
- 404 39.M. J. Calcott, D. F. Ackerley, Portability of the thiolation domain in recombinant
405 pyoverdine non-ribosomal peptide synthetases. *BMC Microbiol.* **15**, 162 (2015).

- 406 40.P. J. Belshaw, C. T. Walsh, T. Stachelhaus, Aminoacyl-CoAs as probes of
407 condensation domain selectivity in nonribosomal peptide synthesis. *Science*. **284**,
408 486–489 (1999).
- 409 41.S. A. Samel, G. Schoenafinger, T. A. Knappe, M. A. Marahiel, L.-O. Essen,
410 Structural and functional insights into a peptide bond-forming bidomain from a
411 nonribosomal peptide synthetase. *Structure*. **15**, 781–792 (2007).
- 412 42.M. Klaus, A. D. D'Souza, A. Nivina, C. Khosla, M. Grninger, Engineering of
413 chimeric polyketide synthases using SYNZIP docking domains. *ACS Chem. Biol.*
414 **14**, 426–433 (2019).
415



416

417 **Figure 1.** Introduction to SZ mediated in trans protein-protein interaction. **(a)**

418 Schematic overview of type A and type S NRPSs. Natural DDs (*in trans* type A NRPS)

419 are shown in green and artificial SZs in red (*in trans* type S NRPS). For domain

420 assignment the following symbols are used: A, adenylation domain, large circles; T,

421 thiolation domain, rectangle; C, condensation domain, triangle; C/E, dual

422 condensation/epimerization domain, diamond; TE, thioesterase domain, small circle.

423 **(b)** Top: excised C-A di-domain and linker region (ribbon representation) from the SrfA-

424 C termination module (PDB-ID: 2VSQ). Removed area of the C-A linker region to

425 introduce SZs is highlighted red. Bottom: modelled 41 AAs comprising SZ pair. **(c)** Top:

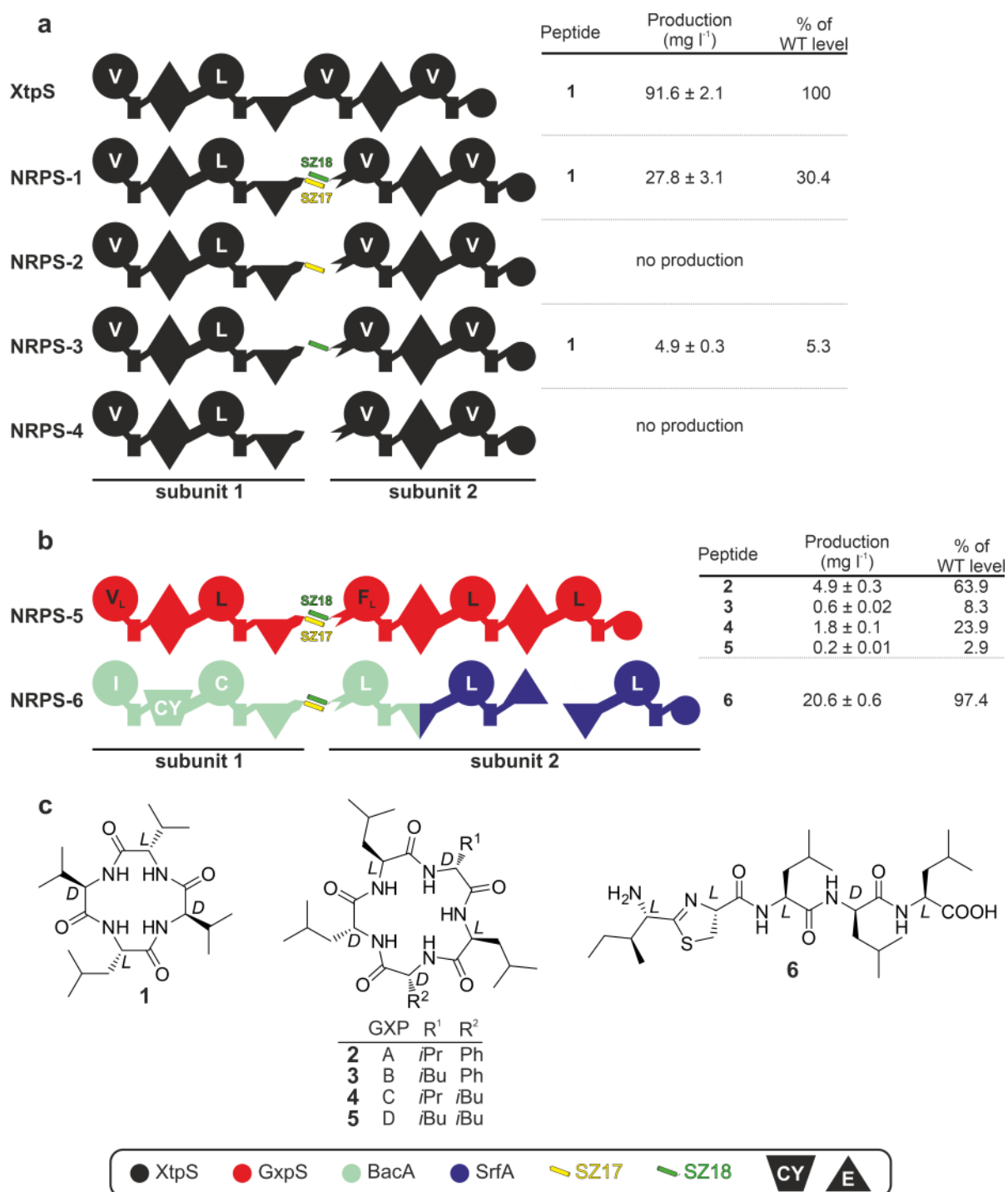
426 antiparallel (left) and parallel (right) interacting hetero-specific SZs. Bottom: SZ17:18

427 and SZ1:2 are forming an orthogonal interaction network. **(d)** SZ interactions. SZ17:18

428 are predicted to be electrostatic complementary at adjacent interfacial e and g

429 positions.

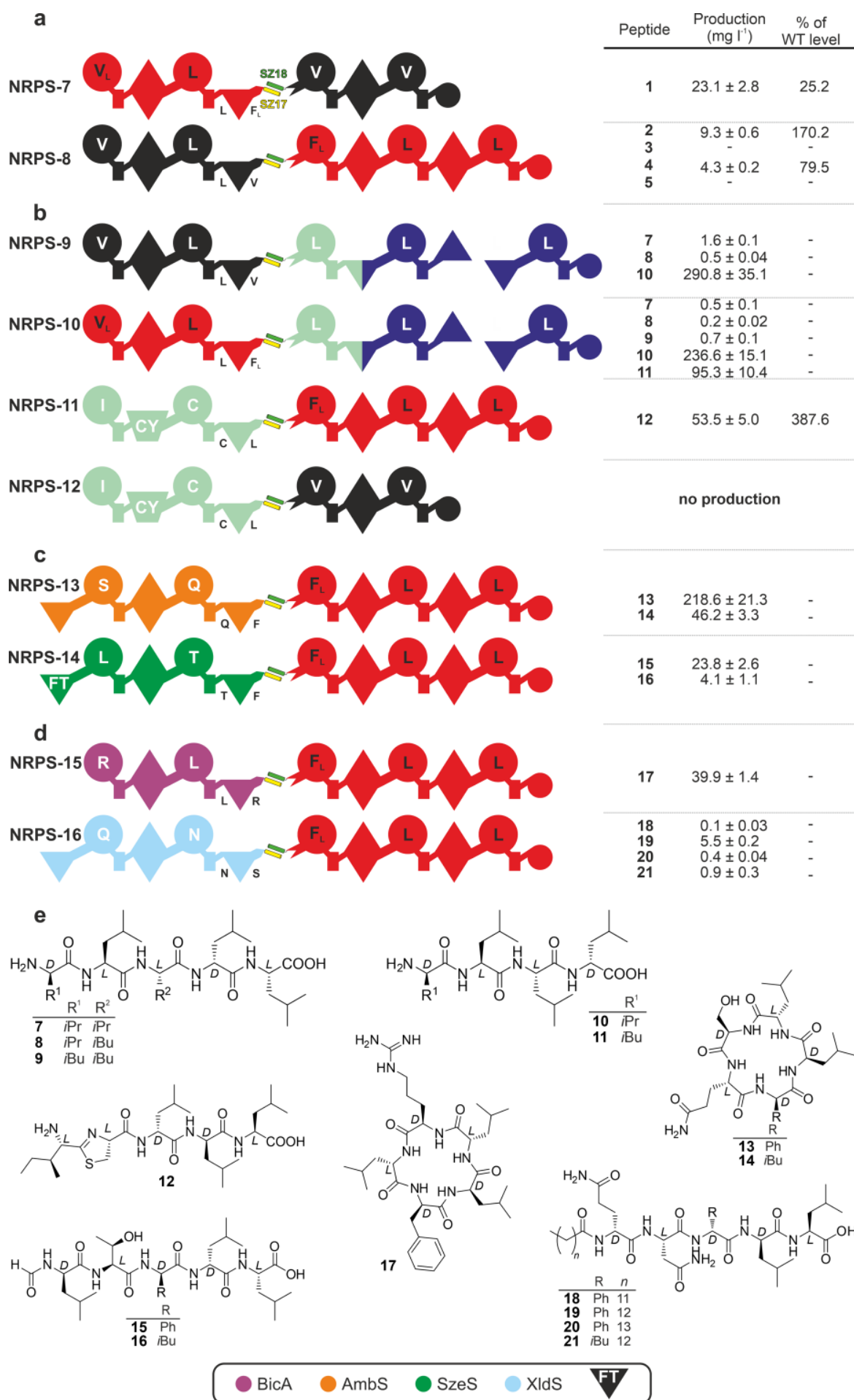
430



431

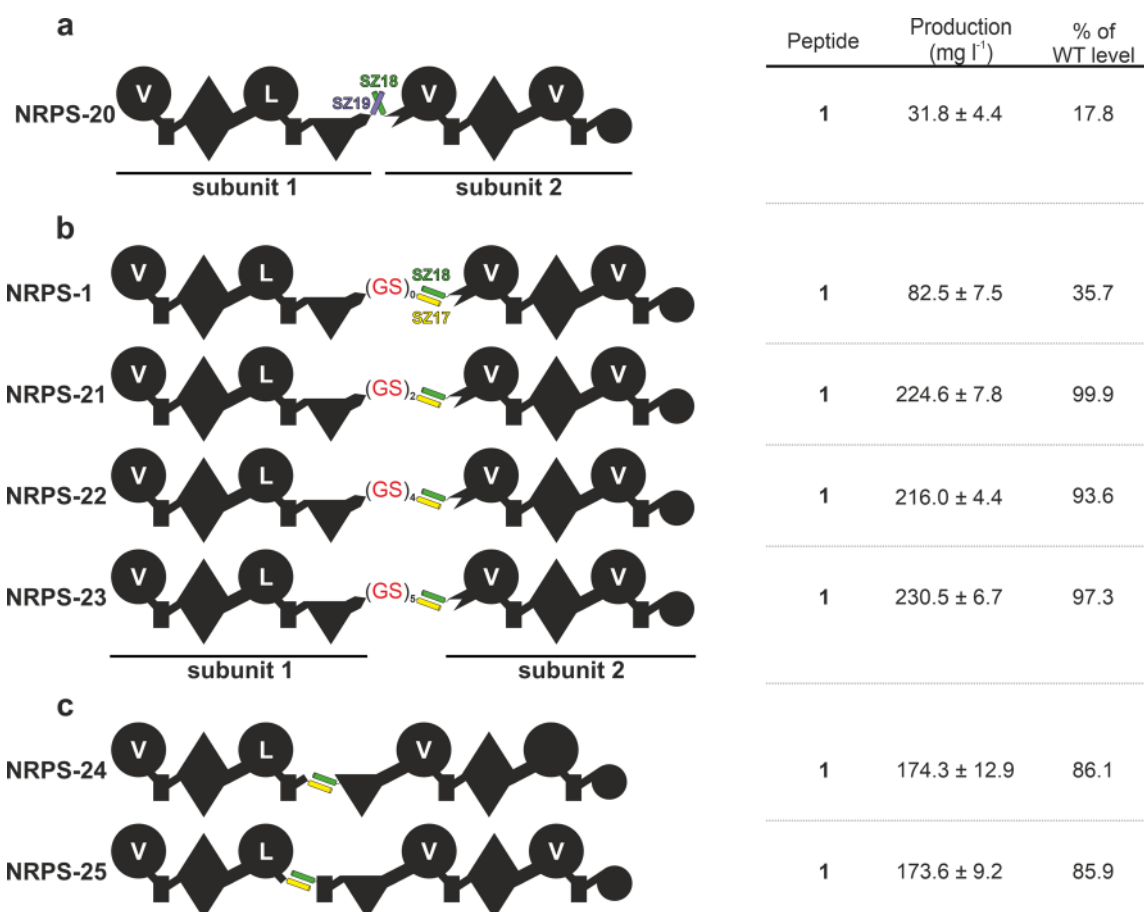
432

433 **Figure 2. (a)** Type S NRPS-1 – 4, as well as corresponding peptide yields obtained
 434 from triplicate experiments. **(b)** Type S NRPS-5 and -6, where GxpS and RtpS are split
 435 in two subunits. **(c)** Structures of 1–6 produced from NRPS-1 to
 436 NRPS-6 expressed in *E. coli*. See Fig. 1 for assignment of the domain symbols; further
 437 symbols: CY, heterocyclization domain; E, epimerization domain. Boxed are the colour
 438 coded NRPSs used as building blocks and the used SZ pairs.



440 **Figure 3. (a)** Type S NRPSs using building blocks from Gram-negative bacteria
441 **(NRPS-7 and NRPS-8)**, **(b)** from Gram-negative and -positive bacteria **(NRPS-9 –**
442 **NRPS-12)**, **(c)** that consider the specificity of the C domain acceptor site **(NRPS-13**
443 **and NRPS-14)** and **(d)** that do not consider the specificity of the C domain acceptor
444 site **(NRPS-15 and NRPS-16)**. **(f)** The structures of **7–21** produced from **NRPS-9** to
445 **NRPS-16** expressed in *E. coli*. See Fig. 1 and 2 for assignment of the domain symbols;
446 further symbols: FT, formyl-transferase domain.

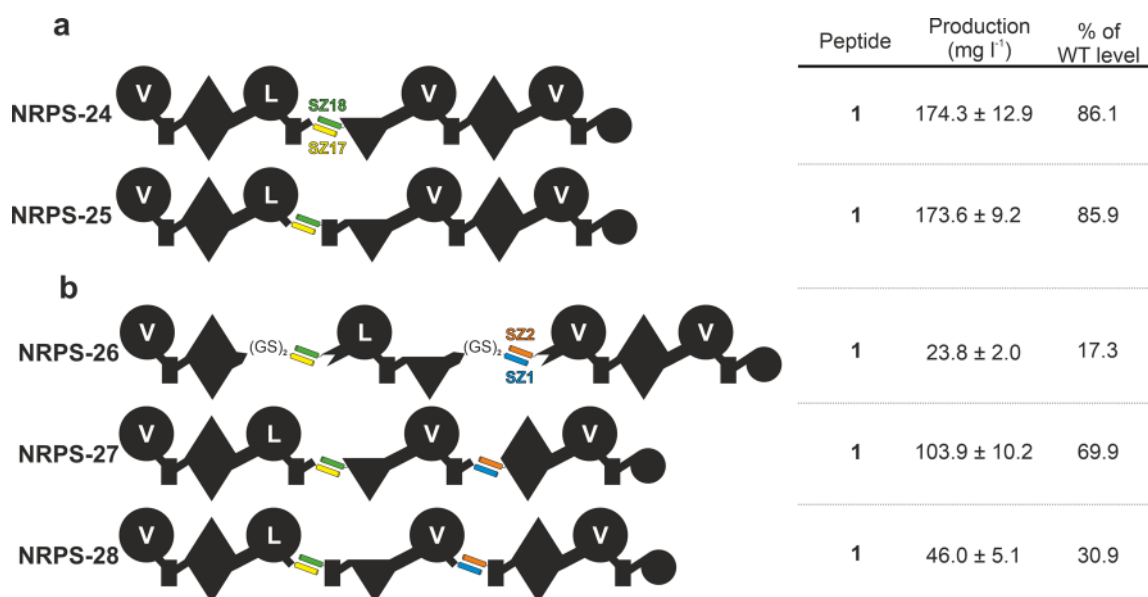
447



448

449 **Figure 4. (a)** Type S XtpS using parallel interacting SZ18:19, **(b)** (GS)_x-elongated C-A
 450 linker sequences (**NRPS-21 – NRPS-23**), **(c)** using different split positions in between
 451 the T-C (**NRPS-24**) and A-T domains (**NRPS-25**). See Fig. 1 for assignment of the
 452 domain symbols.

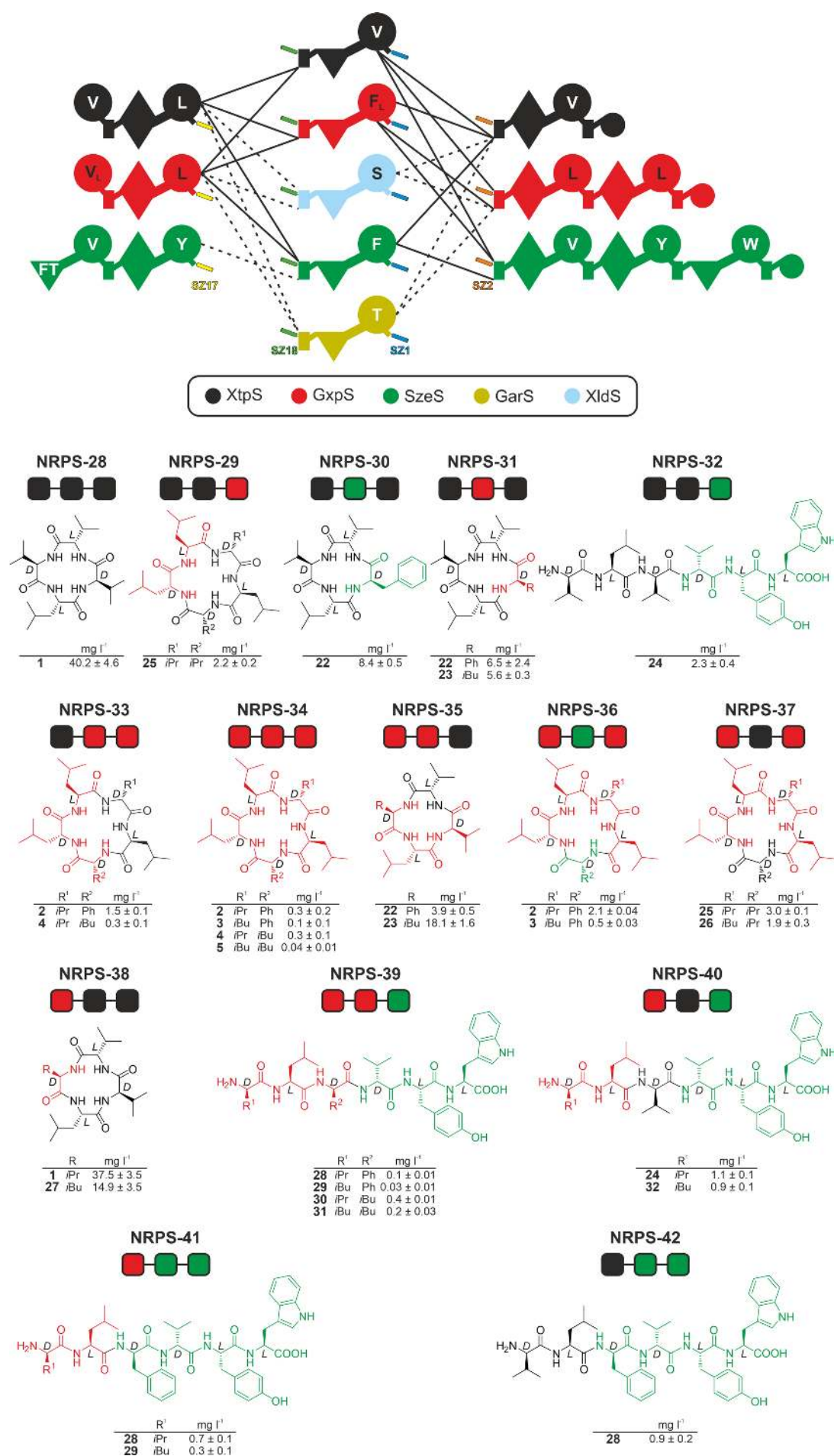
453



454

455 **Figure 5.** Schematic representation of three subunit type S XtpS using the SZ1:2 and
 456 SZ17:18 pairs splitting in between the C-A (NRPS-26), T-C (NRPS-27) and A-T linker
 457 region (NRPS-28). See Fig. 1 for assignment of the domain symbols.

458



460 **Figure 6.** Schematic representation of recombinant type S NRPSs based on three
461 subunit combinations (**NRPS-28 – NRPS-42**) using the A-T linker region as split
462 position leading to a structurally diverse peptide library of **22–32** all expressed in
463 *E. coli*. See Fig. 1 and 3 for assignment of the domain symbols.