

ARTICLE

DOI: 10.1038/s41467-018-06712-1

OPEN

Structure-based redesign of docking domain interactions modulates the product spectrum of a rhabdopeptide-synthesizing NRPS

Carolin Hacker¹, Xiaofeng Cai², Carsten Kögler², Lei Zhao², A. Katharina Weickhmann¹, Jan Philip Wurm^{1,3}, Helge B. Bode^{2,4}  & Jens Wöhnert¹

Several peptides in clinical use are derived from non-ribosomal peptide synthetases (NRPS). In these systems multiple NRPS subunits interact with each other in a specific linear order mediated by specific docking domains (DDs), whose structures are not known yet, to synthesize well-defined peptide products. In contrast to classical NRPSs, single-module NRPS subunits responsible for the generation of rhabdopeptide/xenortide-like peptides (RXP) can act in different order depending on subunit stoichiometry thereby producing peptide libraries. To define the basis for their unusual interaction patterns, we determine the structures of all N-terminal DDs (^NDDs) as well as of an ^NDD-^CDD complex and characterize all putative DD interactions thermodynamically for such a system. Key amino acid residues for DD interactions are identified that upon their exchange change the DD affinity and result in predictable changes in peptide production. Recognition rules for DD interactions are identified that also operate in other megasynthase complexes.

¹Institute of Molecular Biosciences and Center for Biomolecular Magnetic Resonance (BMRZ), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany. ²Molecular Biotechnology, Department of Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany. ³Institute of Biophysics and Physical Biochemistry, University of Regensburg, 93053 Regensburg, Germany. ⁴Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany. These authors contributed equally: Carolin Hacker, Xiaofeng Cai. Correspondence and requests for materials should be addressed to H.B.B. (email: h.bode@bio.uni-frankfurt.de) or to J.Wöh. (email: woehnert@bio.uni-frankfurt.de)

Non-ribosomal peptides are a large family of structurally diverse and pharmacologically useful natural products with broad biological activities. Prominent examples are the antibiotic daptomycin¹ or the immunosuppressant cyclosporine A². They are assembled by multifunctional enzyme complexes called non-ribosomal peptide synthetases (NRPSs) that are organized in a modular fashion. Each module activates and modifies a specific amino acid (aa) that is then subsequently elongated with an aa activated and modified by the next module thereby generating peptides with their length depending on the number of modules used. A typical NRPS elongation module consists of an adenylation (A) domain for activation of a specific aa as aminoacyl adenylate, a condensation (C) domain for peptide bond formation, and a thiolation (T) domain. In the T domains, the aas are covalently attached as reactive thioesters to a phosphopantetheine arm for transfer to the next module^{3,4}. In classical NRPSs, different subunits (i.e., individual NRPS proteins often containing multiple modules on a single protein chain) selectively interact with each other non-covalently in a strictly defined order following the collinearity rule and give rise to the synthesis of peptides with defined sequences. Non-covalent interactions between NRPS subunits are mediated by specialized N- and C-terminal docking domains (DDs). Stachelhaus and co-workers have demonstrated that for the NRPS systems synthesizing tyrocidin and surfactin matching pairs of short DD or COM (communication-mediating) domains at the C-terminus of the peptidyl-donating NRPS (^CDD, ~25 aas) and the N-terminus of the accepting NRPS (^NDD, ~14 aas) play a decisive role in defining the order of interactions between subunits *in vitro* and *in vivo* by swapping COM domains between different subunits⁵. They also constructed a “universal COM system” *in vitro* by a comparison of COM domains in the tyrocidin A and surfactin-like NRPSs, which led to enzyme crosstalk between different biosynthetic systems that promoted the combinatorial biosynthesis of different peptides^{6,7}. The structures of these DDs have not been elucidated so far but interactions between DDs have been mapped based on photocrosslinking experiments⁸. However, for polyketide synthase (PKS) systems that are also organized as modular megasynthases, three structurally different types of DD

pairs have been described so far and two structures for ^NDDs without a bound ^CDD have been solved for NRPS/PKS hybrid systems^{9–14}.

Classic NRPSs are often multimodular—each protein subunit normally consists of the processing modules for multiple aas arranged in a linear fashion. Furthermore, the different protein subunits (i.e., multiple NRPS subunits representing the complete assembly line) interact with each other in a strictly defined linear order yielding a single peptide product with a sequence that faithfully reproduces the linear order of the processing modules along the protein chains. A novel class of NRPSs where up to three single-module NRPS subunits produce complex libraries of rhabdopeptide/xenortide-like peptides (RXPs) has been recently described from entomopathogenic bacteria of the genera *Xenorhabdus* and *Photorhabdus*¹⁵. The RXP product spectra of different strains differ mainly in peptide length and sequence. In order to synthesize such a range of related products by a single NRPS system, an iterative use of certain NRPS modules or the activity of multiple copies of these modules is apparently required. The observed product spectrum also depends on the stoichiometry of NRPS modules acting in elongation and termination (Supplementary Fig. 1)¹⁵. This suggests that the subunits of these NRPSs interact with each other not in a well-defined linear order but stochastically. Furthermore, the individual single-module subunits of these NRPSs contain putative N- and C-terminal DDs (^NCDD and ^CDD, respectively) that differ in length from and show no sequence homology to the DDs in the well-characterized tyrocidin and surfactin producing NRPSs. The Kj12ABC NRPS system from *Xenorhabdus stockiae* KJ12.1 consisting of the three proteins Kj12A, Kj12B, and Kj12C (Fig. 1) is an example for an RXP-synthesizing NRPS. The peptide spectrum produced by this NRPS includes peptides with 2–8 valine and *N*-methyl valine residues in varying orders. The A domains in Kj12A and Kj12B are both specific for Val. However, module Kj12B includes an additional methyltransferase domain that further modifies valine to *N*-methyl valine. While Kj12A and Kj12B are single-module NRPS subunits, Kj12C is a stand-alone C domain that transfers the final peptide chain bound to the phosphopantetheine arm of the T domain in Kj12B to a free amine, mostly phenylethylamine

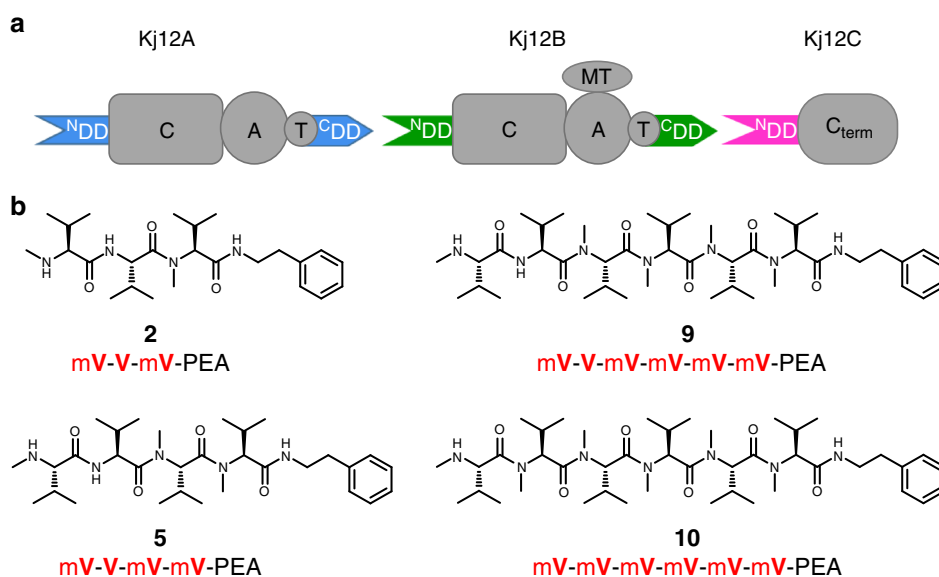


Fig. 1 RXP NRPSs (Kj12ABC) from *Xenorhabdus stockiae* KJ12.1 and selected RXPs found in this strain. **a** Overview of the domain organization of KJ12ABC (C condensation, A adenylation, MT methyltransferase, T thiolation, C_{term} terminal condensation domain, ^NDD N-terminal docking domain, ^CDD C-terminal docking domain). **b** Structures of selected RXPs derived from the KJ12ABC system, showing differences in size and methylation patterns. The simplified structure nomenclature used within other figures is also shown (V Val, mV *N*-methylated Val, PEA phenylethylamine)

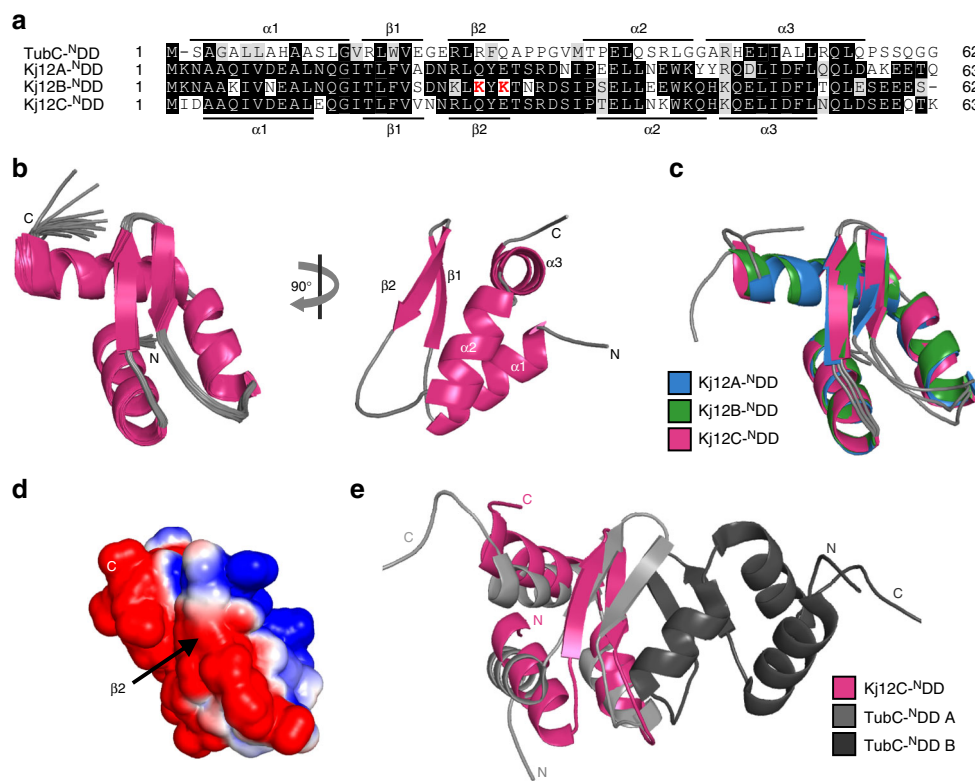


Fig. 2 N-terminal docking domains have the same three-dimensional structure. **a** Structure-based sequence alignment of the TubC^{-NDD} and the N-terminal docking domains of KJ12ABC. Identical residues are highlighted with dark gray boxes and residues with similar chemical properties are shown in light gray boxes. Key residues for DD interactions are shown in red. The secondary structure based on structural information for TubC^{-NDD} and KJ12C^{-NDD} is indicated above and below the sequence. **b** Solution structure bundle of the 19 lowest energy conformers and the regularized mean structure for KJ12C^{-NDD}. **c** Overlay of cartoon representations of the energy minimized mean structures of KJ12A^{-NDD} (blue), KJ12B^{-NDD} (green), and KJ12C^{-NDD} (magenta). **d** Electrostatic surface potentials of KJ12C^{-NDD} mapped on the solvent-accessible surface in the same orientation as in **b** (left) with negatively charged surface areas colored in red, positively charged areas coloured in blue, and white areas corresponding to hydrophobic surfaces. **e** Overlay of the energy-minimized mean structure of KJ12C^{-NDD} with one monomer of the TubC^{-NDD} dimer

(PEA) (Fig. 1a, Supplementary Fig. 1)¹⁵. However, the sequences of the RXP products are biased suggesting that some subunits in this system are used preferentially. These preferences might reflect differences in the interactions between subunits as mediated by their DDs. The structural and thermodynamic basis for the differential DD interactions in these unusual NRPS systems and in NRPS systems in general is not clear. Therefore, we characterize DD interactions in these systems in detail and compare them to DD interactions from classical NRPS and other NRPS/PKS systems. We determine the structures of all ^{NDD}s in the three protein NRPS system KJ12ABC from *X. stockiae* KJ12.1¹⁵. We also characterize the thermodynamic basis for the interaction of these ^{NDD}s with the two ^{CDD}s present in this system and solve the structure of one ^{NDD}-^{CDD} complex. The structural information for the ^{NDD}/^{CDD} interaction allows us to derive a set of simple recognition rules for this type of DD interactions as well as the targeted reprogramming of selected DDs via rationally designed aa exchanges leading to differences in the produced peptides. The type of ^{NDD}/^{CDD} interaction observed for this NRPS system is also found in other megasynthase systems from widely divergent classes of bacteria.

Results

Structure determination of ^{NDD}s in RXP-type NRPS subunits.

Structural and thermodynamic information about DD interactions in single-module NRPS systems that use their subunits in a nonlinear fashion has previously not been available. The RXP

NRPS system from *X. stockiae* KJ12.1 as a model system¹⁵ consists of the three proteins KJ12A, KJ12B, and KJ12C (Fig. 1a). Bioinformatic analysis of the DDs in this and related RXP NRPSs revealed that all three proteins contain an ^{NDD}, and a distinct ^{CDD} was found in KJ12A and KJ12B but not in the termination module KJ12C which catalyzes the reaction of the peptide chain with a terminal amine (Supplementary Fig. 2). The three ^{NDD}s are ~65 aas long with >70% sequence identity among them (Fig. 2a and Supplementary Fig. 2) and a predicted mixed α/β -secondary structure. Thus these ^{NDD}s are much longer and structurally more complex than the previously identified very short and probably unstructured ^{NDD}s in the tyrocidin- and surfactin-producing NRPS systems from *Bacillus*^{5,6}. Furthermore, all three ^{NDD}s showed low sequence homology (<25% identity) to a structurally characterized ^{NDD} from the TubC subunit of the tubulysin-synthesizing PKS (TubC^{-NDD}) from *Angiocooccus disciformis*¹³ that was shown to be a homodimer as well as to a monomeric ^{NDD} of subunit B of the epothilone-synthesizing NRPS-PKS system (EpoB^{-NDD}) crystallized in its native context as a covalent fusion with the cyclization domain of EpoB¹⁴. The two ^{CDD}s of the modules KJ12A and KJ12B were predicted to be rather short (~20 aas) and unstructured (Supplementary Fig. 2). We therefore decided to determine the structures of all three ^{NDD}s excised from the KJ12ABC RXP NRPS system. According to gel filtration in combination with size exclusion chromatography (SEC)-multi-angle static light scattering, all three ^{NDD}s were monomeric in solution in contrast to what was observed for the dimeric TubC^{-NDD} (Supplementary Fig. 3). Solution-state

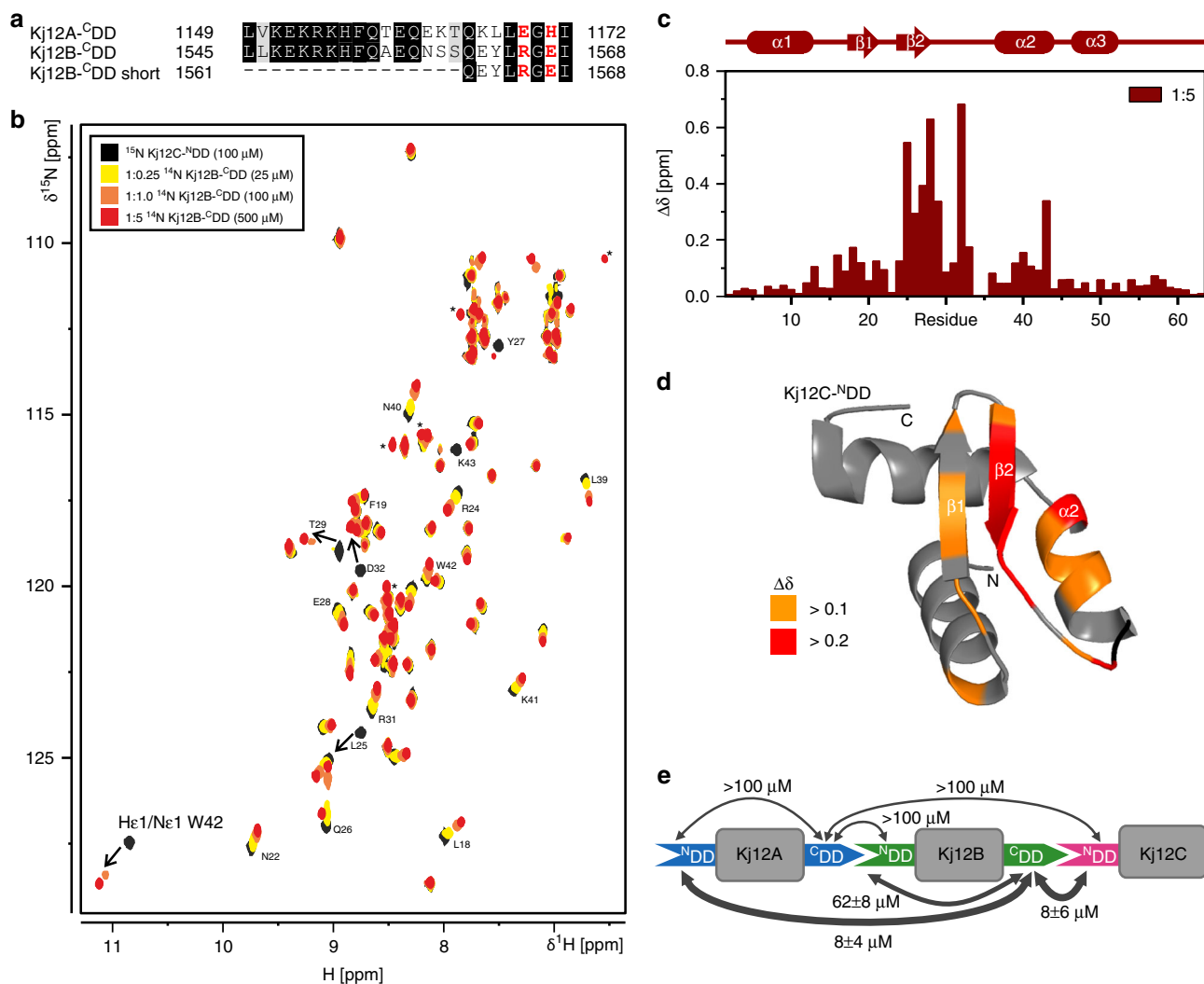


Fig. 3 Docking domain interaction. **a** Sequence alignment of Kj12A-^CDD and Kj12B-^CDD used in this study. Identical residues are highlighted with dark gray boxes and residues with similar chemical properties are shown in light gray boxes. Key residues for DD interactions are shown in red. **b** Overlay of the ¹H,¹⁵N-HSQC spectrum of 100 μM ¹⁵N-labeled Kj12C-^NDD in the absence (black) and presence of increasing amounts of unlabeled Kj12B-^CDD. The molar ratios of the two docking domains are 1:0.25 (yellow), 1:1 (orange), 1:2.5 (dark orange), and 1:5 (red). **c** Histogram of chemical shift changes vs. sequence for Kj12C-^NDD upon addition of a five-fold molar excess of unlabelled Kj12B-^CDD with the secondary structure depicted above. **d** Chemical shift changes for Kj12C-^NDD upon addition of unlabeled Kj12B-^CDD mapped onto the structure of Kj12C-^NDD. **e** K_d values for all DD interactions in the Kj12ABC NRPS as measured by ITC

nuclear magnetic resonance (NMR) using band-selective excitation short-transient (BEST)-transverse relaxation optimized spectroscopy (TROSY)-based pulse sequences and non-uniform sampling rapidly yielded complete NMR resonance assignments for all three ^NDDs of Kj12ABC. The backbone chemical shift derived from the secondary structures of all three ^NDDs revealed the presence of three α -helices and two β -strands in the order of $\alpha 1$ - $\beta 1$ - $\beta 2$ - $\alpha 2$ - $\alpha 3$ (Fig. 2a). The location of the secondary structure elements along the sequence is very similar between the three ^NDDs as well as to those of the dimeric TubC-^NDD and the monomeric EpoB-^NDD^{13,14}. The NMR solution structures of the three ^NDDs from Kj12ABC were solved at very high resolution (backbone root mean square deviation (RMSD) of 0.1–0.2 Å for ordered residues). A complete list of structural statistics according to the recommendations of the NMR-VTF can be found in Supplementary Table 1¹⁶. The solution structure ensemble of the 19 lowest energy structures calculated with CYANA and an energy-minimized representative mean structure for Kj12C-^NDD is shown in Fig. 2b. A comparison of the structural ensembles and

the mean structures for all three Kj12-^NDDs are shown in Supplementary Fig. 4. In all three ^NDDs, $\beta 1$ and $\beta 2$ form an anti-parallel β -hairpin. Helices $\alpha 1$ and $\alpha 2$ are packed against each other in an antiparallel fashion. They also pack together against one side of the β -hairpin. Helix $\alpha 3$ is separated only by a very short loop (aa 45) from $\alpha 2$ and a sharp kink is introduced in the protein backbone. Thus, $\alpha 3$ crosses the β -hairpin at a $\sim 90^\circ$ angle. Despite the very similar three-dimensional (3D) structures (RMSDs range from 0.8 to 0.9) of the three ^NDDs (Fig. 2c), their electrostatic surface potentials differ significantly, which could have an effect on their binding affinities for the ^CDDs (Fig. 2d and Supplementary Fig. 4). A significant charge difference is found between Kj12B-^NDD and the other two ^NDDs on the solvent-exposed side of strand $\beta 2$ (aa 24–28) where for instance E28 in Kj12A-^NDD and Kj12C-^NDD is replaced by a lysine in Kj12B-^NDD (Supplementary Fig. 4d). The topology of the three RXP-^NDDs is already known from the TubC-^NDD structure of *A. disciformis* (Fig. 2e)¹³, but the relative positioning of the secondary structure elements is different between the dimeric

TubC-^NDD and the RXP NRPS-^NDDs (C α RMSD of 4.7 Å). In contrast, the C α RMSD is only 1.3 Å (Supplementary Fig. 4f) between the Kj12C-^NDD and the monomeric EpoB-^NDD.

Interactions with the ^CDDs. To further investigate the DD interactions, NMR titration experiments were carried out. All three ¹⁵N-labeled ^NDD proteins were titrated with the two unlabeled (¹⁴N) ^CDD peptides (Fig. 3a). In Fig. 3b, the titration of Kj12C-^NDD with Kj12B-^CDD as followed in ¹H,¹⁵N-heteronuclear single quantum correlation (HSQC) experiments is shown as an example. The NMR data for all other titrations are shown in Supplementary Fig. 5. In all six titration experiments, gradual chemical shift changes and/or peak broadening during the stepwise addition of the ^CDD peptides were observed. This suggests that all three ^NDDs interacted with both the Kj12A- and the Kj12B-^CDD, which is in agreement with the observed product spectrum of this NRPS, and form ^NDD/^CDD complexes in the fast-to-intermediate exchange regime on the NMR timescale. The chemical shift changes during the titrations were quantified and mapped to the structures of the three ^NDDs in order to identify the ^CDD-binding sites. A histogram of chemical shift changes vs. sequence for the titration of the Kj12C-^NDD with the Kj12B-^CDD is shown in Fig. 3c. A mapping of these shift changes on the cartoon representation of the Kj12C-^NDD structure is shown in Fig. 3d (also see Supplementary Figs. 6 and 7 for the data of all six titrations). Unexpectedly, in all six titration experiments, the largest chemical shift changes were always observed for the aas in the β -hairpin and in particular strand β 2 and helix α 2 of the three ^NDDs. This suggests that both the Kj12A- and Kj12B-^CDD bind to all three ^NDDs at the same sites.

The pairwise interactions between the three ^NDDs and the two ^CDDs were further characterized thermodynamically using isothermal titration (ITC) experiments (Supplementary Fig. 8 and Supplementary Table 2). The obtained K_d values for all interactions (Fig. 3e) are in good agreement with the fast-to-intermediate exchange observed in the NMR titrations. The highest affinity interactions with K_d values of $8 \pm 4 \mu\text{M}$ and $8 \pm 6 \mu\text{M}$ were observed between the Kj12B-^CDD and the ^NDDs of Kj12A and Kj12C, respectively. The Kj12B-^NDD binds Kj12B-^CDD with an \sim 8-fold higher K_d of $62 \pm 8 \mu\text{M}$. Kj12A-^CDD (Fig. 3e) binds only weakly to Kj12C-^NDD (\sim 100 μM), whereas its interaction with the other two ^NDDs is too weak to be reliably quantified (Supplementary Fig. 8) The reported affinities for DD pairs in other megasynthases were found to be in a similar range as those observed here^{11,13}.

Solution structure of a DD complex. Structural information about the interaction of the type of ^NDD found in the Kj12 RXP NRPS cluster so far is limited to ^CDD peptide titration experiments for the dimeric TubC-^NDD. There the interaction surface on the dimeric ^NDD was identified but no complex structure was obtained¹³. Furthermore, our NMR titration experiments identify a ^CDD-binding site on the monomeric ^NDDs that is part of the dimer interface in the TubC-^NDD. Thus the structural basis for the ^NDD/^CDD interactions in the Kj12 RXP NRPS is apparently different from what was observed for TubC. For these reasons and in order to understand the structural basis for the widely different ^NDD/^CDD affinities, we solved the structure of Kj12C-^NDD/Kj12B-^CDD, the highest affinity DD complex from the Kj12 RXP NRPS. This complex is, however, still in fast-to-intermediate exchange on the NMR timescale and therefore the collection of a large enough number of intermolecular nuclear Overhauser effects (NOEs) is difficult. To overcome this problem, we designed a covalently linked ^NDD-^CDD complex with flexible

glycine-serine (GS) linkers of different lengths to increase the local ^CDD concentration at the ^NDD. To verify that our artificially linked DD pair interacts in *cis* with the same binding mode as the isolated domains in *trans*, we compared the titration end point ¹H,¹⁵N-HSQC spectrum of the separate domains with the ¹H,¹⁵N-HSQC spectrum of the linked constructs (Supplementary Figs. 9 and 10). For this purpose, we screened different constructs with different linker length (6, 9, and 12 residues) and different domain order (^NDD-linker-^CDD or ^CDD-linker-^NDD). The construct with the longest linker (12 residues) and an ^NDD-linker-^CDD arrangement (Fig. 4a) was the best mimic for the natural ^NDD-^CDD complex and therefore we solved the structure of this fusion protein (Fig. 4b). To our knowledge, this structure represents the first high-resolution structure of an NRPS DD pair. Surprisingly, the ^NDD-^CDD interaction involves only the last five C-terminal aas of the ^CDD. These aas form an additional β -strand, β 3, which interacts in an antiparallel orientation with β 2 of the ^NDD β -sheet as well as with parts of helix α 2 (Fig. 4c and Supplementary Fig. 11). The α -helix formed by the first nine aas of the ^CDD (aa 1543–1555 of Kj12B) does not interact with the ^NDD (Fig. 4b) and the observed secondary structure is transient, as evidenced by NMR and circular dichroism experiments (Supplementary Fig. 12). An overlay with the structure of the isolated ^NDD shows that the ^NDD does not change its conformation in complex with the ^CDD (Supplementary Fig. 13).

A detailed view of the intermolecular backbone hydrogen bonding interactions between β -strands β 2 and β 3 from the ^NDD and the ^CDD, respectively, is shown in Supplementary Fig. 14. Interestingly, the C-terminal end of the β -sheet of Kj12B ^CDD (β 3) is highly twisted toward helix α 2 (Fig. 4c). Thereby, the side chain of the last aa (I1568) is buried in a hydrophobic pocket consisting of the surrounding side chains of helix α 2 of the ^NDD (Supplementary Fig. 11). The interaction of the ^CDD with this helix explains the large chemical shift changes (Supplementary Fig. 6) observed for aa 39–43 of the ^NDD during the titration experiments. Additionally, the side chain of the first aa of β 3 (L1564) is located in a hydrophobic pocket built by hydrophobic sidechains from β 2 and from the loop between β 2 and α 2 (Supplementary Fig. 11). The signals for the backbone amide groups of the loop residues are either not observable (I34) or have very low intensities in the free ^NDD (D32, S33) indicative of conformational exchange. In complex with the ^CDD, these signals have significantly higher intensities. Thus loop2 of the ^NDD is stabilized upon ^CDD binding. Y27 of the ^NDD packs tightly against the small side chain of G1566 in the ^CDD. The intermolecular interaction between β 2 of the ^NDD and β 3 of the ^CDD is also stabilized by two salt bridges between side chains involving R24 of β 2 and E1567 of β 3 as well as E28 of β 2 and R1565 of β 3, respectively (Fig. 4c).

Taken together, ^NDD/^CDD complex formation is apparently dependent upon the formation of an intermolecular β -sheet stabilized by two salt bridges and the burial of two large hydrophobic side chains. The involvement of only the five C-terminal aas but not the remainder of the ^CDD in complex formation is also supported by steady-state ¹H-¹⁵N heteronuclear Overhauser effect (hetNOE) measurements. As expected, low hetNOE values for residues in the GS linker and in the majority of the ^CDD including the residues of the transient α -helix formed by aa 1543–1555 demonstrate that these aa residues are highly flexible in solution. Only the C-terminal aas of the ^CDD that are part of β 3 have the same high hetNOE values as the conformationally rigid residues of the ^NDD (Supplementary Fig. 12). To further verify that the ^NDD-^CDD interaction is exclusively based on this β -sheet interaction, we repeated the titration experiments with a shortened Kj12B-^CDD peptide that

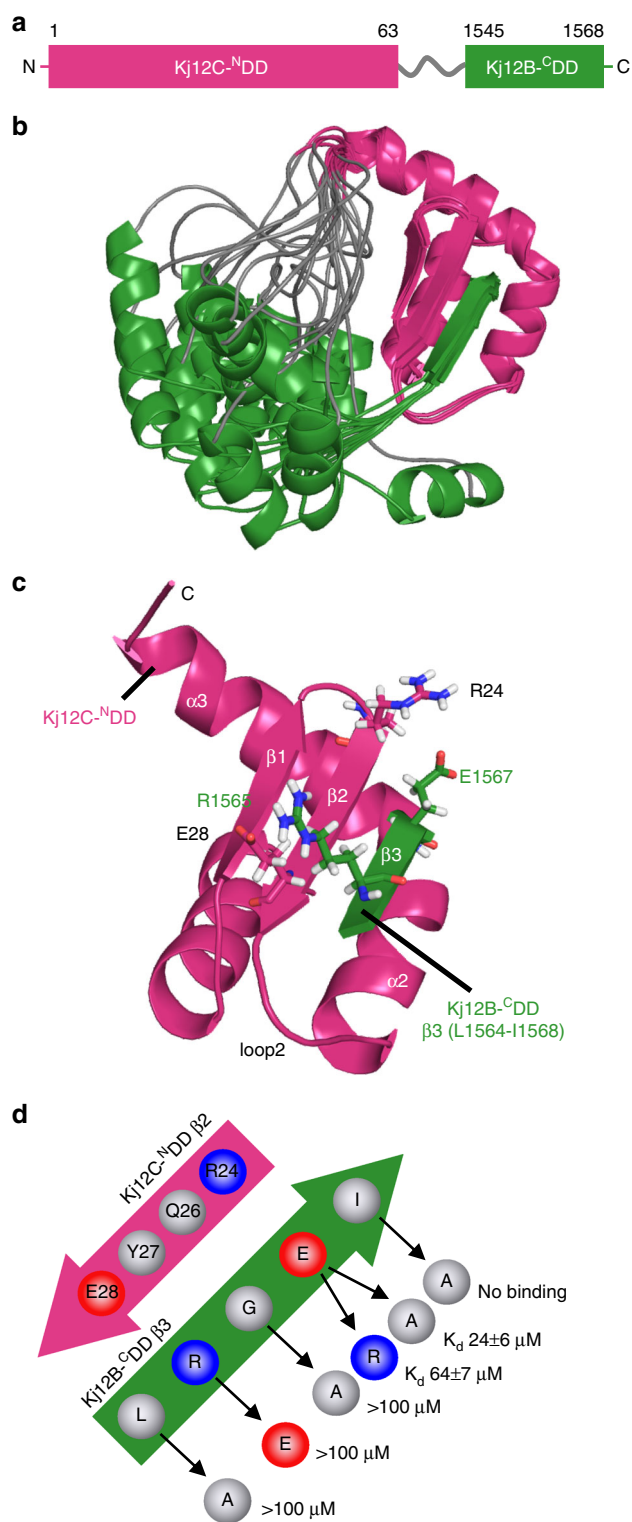


Fig. 4 Structure of $^{\text{NDD}}\text{-}^{\text{CDD}}$ complex. **a** Schematic representation of the N- and C-terminal docking domain linker construct used in this study with Kj12C- $^{\text{NDD}}$ in magenta, the 12 amino acid long Gly-Ser linker in gray, and Kj12B- $^{\text{CDD}}$ (amino acids 1545–1568 from Kj12B corresponding to residues 75–99 in the complex construct Kj12C- $^{\text{NDD}}$ -12xGS-Kj12B- $^{\text{CDD}}$) in green. **b** Solution structure bundle of Kj12C- $^{\text{NDD}}$ -Kj12B $^{\text{CDD}}$ linker construct with colour coding as in **a**. **c** Detailed view of the $^{\text{NDD}}\text{-}^{\text{CDD}}$ interaction of β -sheet 2 of Kj12C- $^{\text{NDD}}$ (magenta) with the last 5 amino acids of Kj12B- $^{\text{CDD}}$ (green). The charged residues forming salt bridges between the two docking domains are shown in stick representation. **d** Schematic representation of the “recognition rules” for the interaction of β 2 of Kj12C- $^{\text{NDD}}$ (magenta) and β 3 of Kj12B- $^{\text{CDD}}$ (green) in the complex. Positively and negatively charged residues are shown as blue and red circles, respectively, and hydrophobic residues as white circles. K_d values were determined by ITC titration experiments with synthetic peptides of Kj12B- $^{\text{CDD}}_{\text{short}}$ (Supplementary Table 2 and 4) carrying individual variations of all five residues of β 3

peptide were synthesized (Figure 4d, Supplementary Table 4, Supplementary Fig. 16) and tested. Replacement of either of the two large hydrophobic side chains L1564 and I1567 by the smaller alanine led to a significant loss of affinity (Supplementary Fig. 17 and Supplementary Table 2). Increasing the size of the side chain of G1565 that stacks against Y27 by replacement with alanine also leads to a decrease in binding affinity. Importantly, breaking of the salt bridges between E28 and R1565 and between R24 and E1567, respectively, significantly lowers the affinities between the Kj12C- $^{\text{NDD}}$ and the $^{\text{CDD}}$ peptide. In particular, the salt bridge between E28 and R1565 seems to contribute strongly to the binding affinity.

In this respect, it is interesting to note that in the Kj12B- $^{\text{NDD}}$ the equivalent of E28 is replaced by lysine (Fig. 2a) explaining the lower affinity measured for its interaction with the native Kj12B- $^{\text{CDD}}$ peptide. In contrast, a Kj12B- $^{\text{CDD}}$ peptide R1565E that should restore salt bridge formation with K28 in Kj12B- $^{\text{NDD}}$ binds with a much higher affinity to the Kj12B- $^{\text{NDD}}$ ($13 \pm 1 \mu\text{M}$) (Supplementary Table 2). Furthermore, these data rationalize why the Kj12A- $^{\text{CDD}}$ binds rather weakly to all three $^{\text{NDD}}$ s. At the positions corresponding to E1567 and R1565 in the Kj12B- $^{\text{CDD}}$, the Kj12A- $^{\text{CDD}}$ contains a histidine and a glutamate residue, respectively (Fig. 3a). Thus the formation of the two intermolecular salt bridges across the intermolecular β -sheet is weakened or prevented in the interactions involving the Kj12A- $^{\text{CDD}}$. In contrast, a Kj12A- $^{\text{CDD}}$ peptide with an E1169R and a H1171E double mutant supporting the formation of both salt bridges binds with high affinity ($3.5 \pm 0.1 \mu\text{M}$) to the Kj12C- $^{\text{NDD}}$ (Fig. 5a, Supplementary Table 2).

We also tested whether the $^{\text{NDD}}\text{-}^{\text{CDD}}$ interaction is influenced by adjacent domains in the respective modules. To this end, we recombinantly produced the Kj12B- $^{\text{CDD}}$ fused to its preceding T domain (Kj12B-Tdom- $^{\text{CDD}}$) and the full-length Kj12C module consisting of the $^{\text{NDD}}$ and the entire condensation domain (Kj12C- $^{\text{NDD}}$ - C_{term}) (Supplementary Fig. 18). ITC experiments showed that the presence of the Kj12B T domain does not influence the affinity of the Kj12B $^{\text{CDD}}$ for the Kj12C $^{\text{NDD}}$ and that the full-length Kj12C module binds with the same affinity to the Kj12B $^{\text{CDD}}$ peptide as the isolated Kj12C $^{\text{NDD}}$ (Supplementary Fig. 18 and Supplementary Table 2). The Kj12C module without the $^{\text{NDD}}$ shows no binding to the Kj12B $^{\text{CDD}}$ peptide. Thus our results suggest that the tuning of intermolecular DD interactions by modulating the formation of the two intermolecular salt bridges in the context of full-length modules could influence the interactions between entire NRPS modules and thereby the frequency of intermodular substrate transfer and the resulting product spectrum in this NRPS system.

comprises only the β -sheet (eight C-terminal residues from Kj12B- $^{\text{CDD}}$, aa 1561–1568, Fig. 3a). The end points of both NMR titration experiments (Kj12C- $^{\text{NDD}}$ with Kj12B- $^{\text{CDD}}$ and Kj12B- $^{\text{CDD}}_{\text{short}}$, respectively, in identical concentrations and ratios) overlapped perfectly (Supplementary Fig. 15) and ITC measurements confirmed that the short $^{\text{CDD}}$ peptide binds with a very similar affinity ($15 \pm 3 \mu\text{M}$) compared to the original longer $^{\text{CDD}}$ peptide (Supplementary Fig. 15).

In order to test the relative importance of the observed intermolecular interactions, variants of the Kj12B- $^{\text{CDD}}_{\text{short}}$

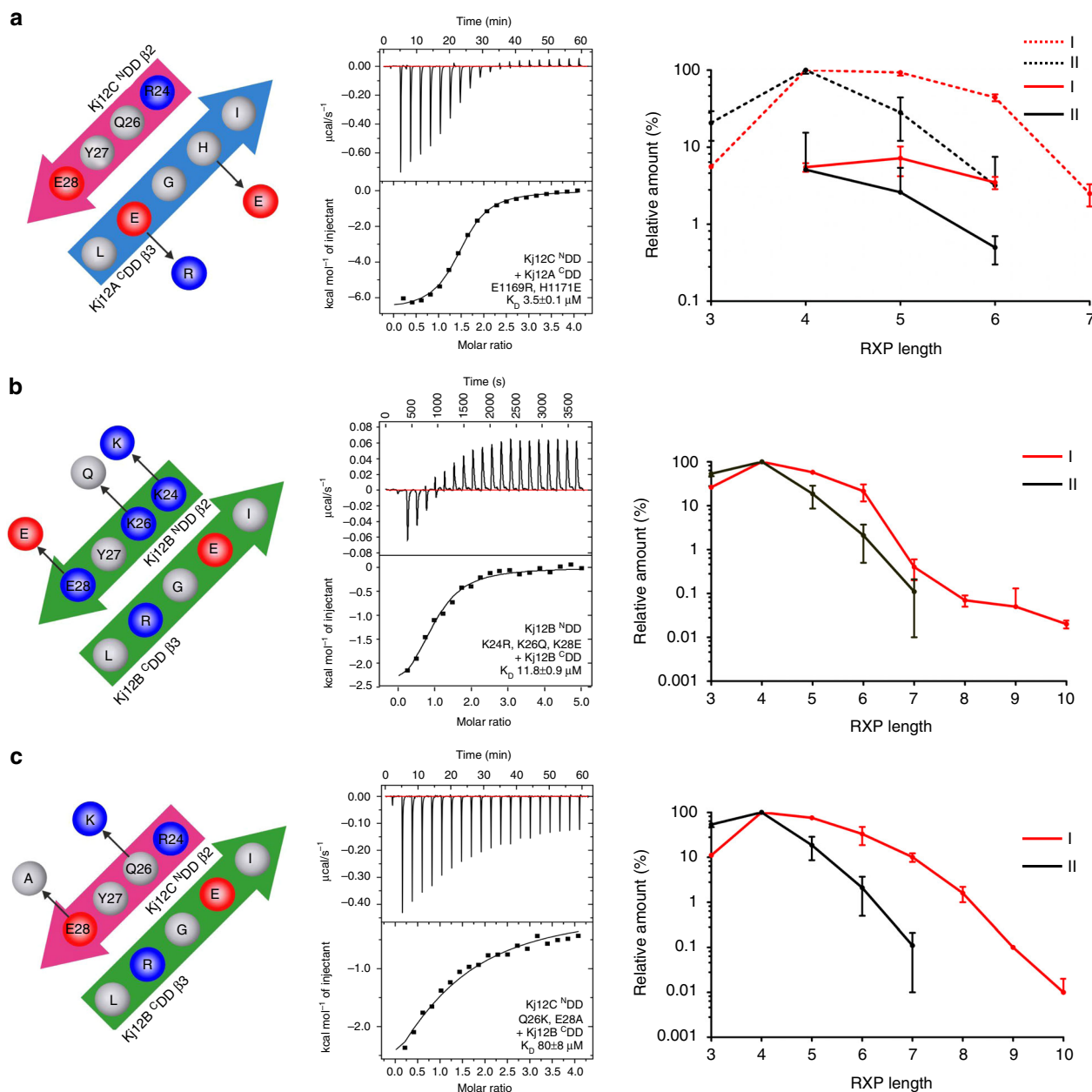


Fig. 5 Optimization of ^CDD or ^NDDs in Kji2ABC system for the production of longer RXP. Additionally, the ITC thermograms and the derived binding curves for titrations between optimized ^CDD or ^NDDs variants are shown. **a** Kji2A-^CDD was optimized by two amino acid exchanges, E1169R and H1171E, on Kji2A-^CDD. Co-expression of natural Kji2B with optimized Kji2C led to increased production of longer RXP. Red lines (I) represent RXP production in the modified system, black lines (II) represent RXP production in the natural Kji2BC. Solid lines indicate fully methylated Val (mV) RXP, dashed lines indicate RXP containing only one non-methylated Val. **b** Optimization of Kji2B-^NDD for tighter interaction with Kji2B-^CDD via three amino acid exchanges, K26Q, K24R and K28E, on Kji2B-^NDD. A red line (I) represents RXP production in the optimized system, a black line (II) represents RXP production in the natural Kji2BC. Only fully methylated RXP are shown. **c** In addition to amino acid exchanges in **b**, Kji2C-^NDD was additionally modified via two amino acid exchanges, Q26K and E28A, to reduce the interaction with Kji2B-^CDD and allowing better interaction between Kji2B-^NDD and Kji2B-^CDD. No RXP were detected after co-expression of natural Kji2B and modified Kji2C-^NDD probably due to very weak affinities. A red line (I) represents RXP production in the optimized system, a black line (II) represents RXP production in natural Kji2BC. Only fully methylated RXP are shown. x Axis, numbers of amino acid residues in RXP (RXP length). y Axis, production of the corresponding RXP relative to the most abundant derivative set to 100%

DD reprogramming results in peptide diversification in vivo.

In order to verify the functional importance of the identified key residues for DD interactions in the context of the entire Kji2ABC system in vivo, we systematically changed selected positions in the DDs and analyzed the RXP profile in *E. coli* expressing the complete Kji2ABC NRPS system with either the parent or the modified DDs. As a starting point, in Kji2A the ^CDD was changed to E1169R and H1171E in order to increase the affinity

between Kji2A-^CDD and Kji2C-^NDD (Fig. 5a) in analogy to the in vitro peptide-binding experiments described above (Supplementary Fig. 19d, Supplementary Table 2). Indeed, the optimized Kji2A variant was able to interact with Kji2C resulting in the formation of RXP 1, V-PEA not detected in the native Kji2AC system (Supplementary Fig. 20). When the optimized Kji2A was combined with native Kji2BC, an increase of longer RXP was observed (Fig. 5a and Supplementary Fig. 21) that was also

observed for an artificial system where the aa specificity of Kj12B was changed from V to L allowing an easier differentiation of the activities of Kj12A and Kj12B (Supplementary Fig. 22). Since it was previously shown that peptide length depends on the protein stoichiometry between Kj12B (elongation) and Kj12C (termination)¹³, this likely results from more efficient binding between Kj12A and Kj12C. Thus, less Kj12C is available for peptide termination via binding to Kj12B.

A similar production of longer peptides with a chain length of up to ten aas was observed in a Kj12BC system when the ^NDD of Kj12B- was optimized via K28E, K26Q, and K24R exchange for a higher affinity toward Kj12B-^CDD (Fig. 5b and Supplementary Fig. 23). The higher affinity of these Kj12B-^NDD mutants toward the Kj12B-^CDD was confirmed in vitro in ITC measurements (Supplementary Fig. 19). Aa exchange of E28A and Q26K in the ^NDD of Kj12C designed to weaken the interaction between the Kj12B-^CDD and the Kj12C-^NDD and thereby to reduce termination also resulted in longer peptides of up to nine aas in a Kj12BC system (Supplementary Fig. 24). When the optimized sequences for Kj12B-^NDD (K28E, K26Q, and K28E) and Kj12C-^NDD (E28A and Q26K) were combined, even more long-chain peptides were produced (Fig. 5c and Supplementary Fig. 24).

Prediction of DD interactions in other megasynthase systems.

In order to investigate whether the DD recognition mode of the RXP-NRPS system observed here also occurs in other systems, a BLASTP search using Kj12C-^NDD and TubC-^NDD as query sequences was carried out. In *Xenorhabdus* strains, besides highly related RXP-NRPS systems, a new family of NRPS related to the taxllaid lipopeptide-producing NRPS TxLAB¹⁷ was identified. Here DDs connecting NRPS subunits carrying C-terminal epimerization (E) to N-terminal condensation (C) domains (Supplementary Fig. 2) apparently interact in a fashion similar to that observed in Kj12ABC.

Moreover, DD pairs with a very similar putative interaction mode were identified in *Pseudomonas*, *Janthinobacterium*, *Paenibacillus*, *Acidobacterium*, *Mycolicibacterium*, cyanobacteria, and myxobacteria (Supplementary Fig. 25, Supplementary Table 3).

In different myxobacteria and cyanobacteria, the ^CDD-^NDD pair was found to connect NRPS or PKS (...-T-^CDD) with NRPS (^NDD-C-...) in different PKS/NRPS and NRPS systems (e.g., melithiazol¹⁸, tubulylin, micropeptin¹⁹, and cyanopeptolin²⁰). In *Acidobacterium*, the DD pair was found as part of a large PKS/NRPS system where it connects a PKS with an NRPS subunit and in the Gram-positive *Mycolicibacterium* where it connects a stand-alone thioesterase (TE) domain with an NRPS. In *Pseudomonas*, it is part of a conserved but not further characterized PKS/NRPS system of which the produced natural product is not known yet. There it connects an NRPS carrying a C-terminal oxidoreductase domain with an NRPS carrying an N-terminal C domain (Supplementary Table 3). A DD pair highly similar to that in Kj12ABC is also present in the PKS/NRPS system responsible for romidepsin (FK228, Istodax®) production in *Chromobacterium violaceum*, where it connects the PKS (DepC) with the NRPS (DepD) part²¹. Romidepsin is a known histone deacetylase inhibitor used as anticancer agent in cutaneous and peripheral T cell lymphomas²².

In all these systems, putative interactions between ^NDDs and ^CDDs dependent on salt bridges and hydrophobic interactions similar to the Kj12ABC system could be predicted (Supplementary Fig. 25). However, in some examples as in *Janthinobacterium* (uncharacterized NRPS system), the polarity of the salt bridges is reversed. In other cases, a salt bridge is replaced by a hydrogen bond or by a hydrophobic interaction.

Discussion

Engineering of NRPS or PKS systems for the production of derivatives or even new natural products requires not only efficient modification of catalytic domains^{23,24} but also modification of protein–protein interactions within megasynthases that comprise multiple subunits, as is the case for most systems that incorporate >5 building blocks.

Here we have analyzed the DD-mediated interactions required for the production of rhabdopeptides using RXP-type NRPS subunits from *Xenorhabdus* KJ12.1 as a model system. We demonstrated that aa exchanges in the respective DDs not only results in a predicted shift in protein affinity in vitro but also in the production of different peptides in vivo.

From an overlay of the TubC-^NDD structure with that of Kj12C-^NDD, it is obvious that the ^CDD-binding sites must be very different in the two systems (Supplementary Fig. 4e). This is due to the dimeric structure of the TubC-^NDD where the β -hairpin forms part of the dimerization surface (Fig. 2e) leaving no space for interaction with a ^CDD as observed in Kj12ABC. However, from a brief analysis of different biosynthetic gene clusters it was obvious that several other megasynthases apparently use ^NDD/^CDD interactions that are structurally similar to what we found for Kj12ABC (Supplementary Figs. 2 and 25) and conform to similar recognition rules. The identified DD pairs link different NRPS and PKS enzymes in several different bacterial genera and taxa and thus might be a useful tool for megasynthase engineering in the future.

An understanding of the structural principles governing megasynthase DD interactions forms a basis for future engineering approaches of such systems since it might allow the combination of megasynthase subunits from different biosynthetic pathways that have been individually optimized in a combinatorial approach. Also the splitting of large megasynthase modules that are too difficult to engineer as intact units into smaller functional units might be possible using well-studied DDs. Additionally, DD-engineered crosstalk of megasynthase subunits from different biosynthesis pathways could even increase the chemical diversity of currently known natural products, while the DD specificity code described here allows for fast identification of specific protein–protein interactions and thus might help to elucidate biosynthesis pathways for systems that are not collinear.

However, it should be mentioned that the DDs described in this work and also found in the taxllaid-producing NRPS can link E and C domains (...-E-^CDD to ^NDD-C-...) as it is also the case for the DDs described in the tyrocidine- and surfactin-producing NRPS in *Bacillus*^{5,6}. However, despite their functional similarity the structural basis for their interaction in *Bacillus* and *Xenorhabdus* is very different due to differences in DD length, folding, and aa composition suggesting that these interaction domains might indeed be orthogonal and thus represent interesting tools for synthetic biology of NRPS and beyond. From our brief bioinformatics analysis of DDs in other NRPS clusters, it is evident that, while the RXP type of DD interactions is also found in other systems, additional DD types exist for NRPS and other assembly lines that require further structural and biochemical analysis enabling their future use in NRPS engineering or understanding the basic principles of these megasynthase pathways.

Methods

General molecular biology. Cultivation of *Xenorhabdus* and *Escherichia coli* strains (Supplementary Table 5) was performed in LB medium following standard protocols¹⁵. Procedures, such as plasmid DNA preparation, transformation, restriction digestion, and DNA gel electrophoresis, were adapted from standard protocols²⁵. Isolation of genomic DNA was carried out according to the

manufacturer's instructions (QIAGEN). Phusion High-Fidelity DNA-Polymerase (Thermo Scientific) was used for PCR amplifications. PCR primers used in this study are listed in Supplementary Table 6. All the plasmids (Supplementary Table 7) generated in this study were constructed via Gibson assembly²⁶. The basic cloning was performed in *E. coli* DH10B MtaA.

Expression and purification of DDs and DD complexes. For structure determination, ^NDDs from *X. stockiae* KJ12ABC as well as the covalently linked ^NDD-^CDD complexes and ^NDD mutants were heterologously expressed in *E. coli* BL21-Gold(DE3) under the control of a T7 promoter. The coding sequences were cloned into a modified pET11a vector containing an N-terminal His₆-SUMO tag. DNA fragments encoding KJ12C-^NDD and KJ12B-^CDD were linked with a 12 aa long GS linker in between. The resulting constructs were grown in uniformly ¹⁵N and ¹³C M9 minimal media containing 1 g L⁻¹ ¹⁵NH₄Cl (Cambridge Isotope Laboratories) or 1 g L⁻¹ ¹⁵NH₄Cl and 2.5 g L⁻¹ ¹³C₆-D-glucose (Cambridge Isotope Laboratories) and 100 mg mL⁻¹ ampicillin. For ITC measurements, proteins were expressed in *E. coli* BL21-Gold (DE3) using LB medium. Protein expression was induced at an OD₆₀₀ of 0.8 with 1 mM IPTG overnight at 25 °C. After expression, cells were lysed by sonication in lysis buffer containing 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, one spatula tip of Dnase, RNase (Roche), and protease inhibitor (Roche). The lysate was cleared by centrifugation (30 min, 8000 × g, 4 °C) and the supernatant was passed through a HisTrap HP column (GE Healthcare) using HisTrap-buffer (50 mM sodium phosphate buffer at pH 8 with 30 mM NaCl and 500 mM imidazole for elution). The His₆-SUMO tag was cleaved off by Ulp1 protease treatment in dialysis buffer (50 mM Tris/HCl, pH 8, 50 mM NaCl) and removed by a second purification step with the HisTrap HP column using HisTrap-buffer. All three ^NDDs and the ^NDD-^CDD complexes were further purified via ion exchange chromatography (IEX) with a Q-Sepharose anion exchange column (HiPrep 16/10 column, GE Healthcare, IEX binding buffer: 50 mM Tris/HCl, pH 8, 50 mM NaCl, IEX elution buffer 50 mM Tris/HCl, pH 8, 1 M NaCl) followed by SEC on a HiPrep 16/60 Sephacryl S-100 High Resolution column (GE Healthcare) in SEC buffer and 50 mM sodium phosphate buffer at pH 6.5 with 100 mM NaCl.

The coding sequences for KJ12C-^NDD with the terminal condensation domain (KJ12C-^NDD-Cterm) and a construct lacking the ^NDD (KJ12C-Cterm) were cloned into a pCOLA vector with a kanamycin resistance gene and a C-terminal His₆-tag with a TEV cleavage site. The coding sequence for the thiolation domain of KJ12B together with the ^CDD (KJ12B-Tdom-^CDD, aa 1459–1568) was cloned into a modified pET11a vector containing an N-terminal His₆-SUMO tag. Constructs were heterologously expressed in *E. coli* BL21-Gold(DE3) using LB medium for cultivation. Protein expression was induced at an OD₆₀₀ of 0.8 with 1 mM IPTG overnight at 20 °C. After expression, cells were lysed by sonication in lysis buffer containing 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, one spatula tip of Dnase, RNase (Roche), and protease inhibitor (Roche). The lysate was cleared by centrifugation (30 min, 8000 × g, 4 °C) and the supernatant was passed through a HisTrap HP column (GE Healthcare). KJ12B-Tdom-^CDD was further purified via Q-Sepharose anion exchange (HiPrep 16/10 column, GE Healthcare) using IEX buffers followed by SEC on a HiPrep 16/60 Sephacryl S-100 High Resolution column (GE Healthcare) using 50 mM Tris/HCl, pH 8 with 100 mM NaCl. The His₆-tag was cleaved off by TEV protease treatment in dialysis buffer and removed by a second purification step with the HisTrap HP column. SEC for KJ12C-^NDD-Cterm and KJ12C-Cterm constructs was performed on a HiPrep 16/60 Sephacryl S-200 High Resolution column (GE Healthcare) using SEC buffer with 2 mM β-Mercaptoethanol.

Construction of plasmids encoding modified NRPS subunits. All constructs for the production of modules containing DDs with exchanged aas (1–3 exchanges) were designed using oligonucleotides carrying the required mutations. For example, for the production of the KJ12A-^CDD carrying E1169R and H1171E, Gibson assembly of the required PCR fragments into the vector pCOLA-ara-tac1 was used and similar approaches were used also for the production of the plasmids encoding KJ12B modules with an ^NDD with single (K28E or K26Q), double (K28E and K24R or K26Q and K28E) and triple (K28E, K24R, and K26Q) aa exchanges, as well as KJ12C-^NDD with single (E28K or E28A) or double (Q26K and E28A) aa exchanges to weaken the interaction with other ^CDDs.

Heterologous production of RXPs in *E. coli* DH10B MtaA. To compare the metabolic profiles after modifications of ^NDDs and/or ^CDDs, *E. coli* DH10B MtaA strain was separately transformed with a mixture of different plasmid pairs carrying full-length modules with optimized KJ12A-^CDD, KJ12B-^NDD, modified KJ12C-^NDD, or those carrying the original DDs. The resulting strains were individually inoculated into 5 mL of liquid LB media supplemented with appropriate antibiotics (kanamycin, 50 µg/mL; chloramphenicol, 34 µg/mL and spectinomycin, 50 µg/mL) and cultivated at 30 °C overnight with shaking at 200 rpm. The next day, 100 µL of overnight culture was transferred into 10 mL fresh LB medium with corresponding antibiotics, 1 mM PEA, 2% (v/v) of Amberlite XAD-16 resin (Sigma-Aldrich), 0.1% of L-arabinose for inducing the expression of RXP-NRPSs, and followed by growing the culture at 30 °C, 1 day, 200 rpm for production of RXPs.

Culture extraction. The bacterial cell pellets and XAD beads were collected after centrifugation and resuspended in 10 mL of methanol. XAD beads were washed with methanol by inverting for 1 h and followed by separating from methanol through filter paper. The resulting methanol extracts were evaporated to dryness and dissolved in 1 mL of methanol.

High-performance liquid chromatography-mass spectrometry (HPLC-MS)

analysis. The methanol extracts obtained above (1 mL) were centrifuged at 17,000 × g for 20 min. Twenty µL of crude extracts were diluted in 180 µL methanol before analysis, 5 µL of which was injected and analyzed by electrospray ionization–HPLC-MS by a Dionex UltiMate 3000 system coupled to a Bruker AmaZon X mass spectrometer with an ACQUITY UPLC™ BEH C18 column (130 Å, 2.1 × 100 mm², 1.7 µm particle size, Waters GmbH) at a flow rate of 0.6 mL min⁻¹ using acetonitrile (ACN) and water containing 0.1% formic acid (v/v) in a gradient ranging from 5 to 95% of ACN over 16 min. BPC spectra for RXPs were recorded in positive ion mode with the range from 80 to 1600 m/z and ultraviolet (UV) at 200–600 nm.

Chemical synthesis of short peptides.

For a schematic overview, see Supplementary Fig. 16 showing the synthesis of QEYARGEI as an example. Step a was loading of the first aa (Ile) on the 2-chlorotriethyl chloride (2-CTC) resin. A solution of Fmoc-Ile-OH (212 mg, 0.6 mmol, 3 eq.) and *N,N*-diisopropylethylamine (DIPEA, 0.31 mL, 1.8 mmol, 9 eq.) in 4 mL dry dichloromethane (DCM) was placed in a plastic reactor vessel filled with 2-CTC resin (125 mg, 0.2 mmol, 1.0 eq.). The resulting mixture was incubated at room temperature overnight. The remaining free binding sites were capped upon incubating twice with a mixture of 80% DCM, 15% CH₃OH, and 5% DIPEA for 10 min at room temperature. The resin was washed several times with dimethylformamide (DMF), CH₃OH, and DCM and treated with 20% piperidine in *N*-methylpyrrolidone (NMP, 3 × 10 min) to remove the Fmoc-protecting group. The combined filtrates were used to determine the actual loading of the resin at λ_{301 nm}. Afterwards, the resin was washed with DCM and dried. Step b is solid-phase peptide synthesis. The linear sequence was synthesized on the preloaded Ile-2-CTC resin on a 25-µmol scale with a Syro Wave peptide synthesizer by using standard Fmoc/*t*-Bu chemistry. The resin was placed in a plastic reactor vessel with a Teflon frit and an amount of 6 eq. of aa derivatives (Fmoc-Glu(OtBu)-OH, Fmoc-Gly-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Ala-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Gln(Trt)-OH, 0.2 M) was activated in situ at room temperature with 6 eq. of *O*-(6-chlorobenzotriazol-1-yl)-*N,N,N'*-tetramethyluronium hexafluorophosphate (0.6 M) in DMF in the presence of 12 eq. DIPEA (2.4 M) in NMP for 50 min. Fmoc-protecting groups were removed with a solution of 40% piperidine in NMP for 5 min and the deprotection step was repeated for another 10 min with 20% piperidine in NMP. After each coupling and deprotection step, the resin was washed with NMP. After the addition of the final residue, the resin was washed with NMP, DMF, and DCM and dried. Step c is cleavage of peptide from the resin. A total of 1 mL 95% trifluoroacetic acid and 2.5% triisopropylsilane in water was added to the peptidyl resin (25 µmol) and the mixture was agitated for at least 3 h at room temperature. The resin was removed by filtration and the crude peptide was precipitated in cold diethyl ether/petroether (2:1). After centrifugation (4000 rpm, 10 min, 4 °C) and careful decantation of the supernatant, the crude peptide was washed with diethyl ether twice, centrifuged, and dried. The peptide was further purified by preparative HPLC system and confirmed by HPLC/MS and high-resolution MS (QEYARGEI, mass calc. 483.2380 [M + 2 H]²⁺, found 483.2382 [M + 2 H]²⁺).

NMR spectroscopy. For NMR measurements, the DDs (0.5–1 mM) were prepared in 50 mM sodium phosphate buffer pH 6.5, 100 mM NaCl, and 10% D₂O. For NMR titration experiments, a protein concentration of 100 µM was used. NMR spectra were acquired at 20 °C on Bruker AVANCE III 600, 700, 800, and 950 MHz spectrometers equipped with cryogenic triple resonance probes. The proton chemical shifts were internally referenced to 2,2-dimethyl-2-silapentane-5-sulfonic acid and the heteronuclear ¹³C and ¹⁵N chemical shifts were indirectly referenced with the appropriate conversion factors²⁷. The standard set of triple resonance experiments (HNCO, HN(CO)CA, HNCACB) was used for the backbone resonance assignments of the ^NDD of KJ12C²⁸. For the other ^NDDs and the ^NDD-^CDD linker constructs, BEST-TROSY versions of the triple resonance spectra were used together with non-uniform sampling²⁹. Shaped proton pulses with a bandwidth of 5.0 ppm centered at 8.5 ppm were used. The delay between scans was set to 0.3 s in all experiments. Non-uniform sampling was employed in all BEST-TROSY 3D experiments, where the percentage of data points of the full t₁/t₂ grid actually acquired varied between 25 and 30 with T₂ weighting of 0.05 s in the ¹⁵N dimension and 0.01 s in the ¹³C dimension, respectively. Processing of 3D data sets with non-uniform sampling was carried out with the multidimensional decomposition (mdd) algorithm provided by TopSpin 3.5 using default parameters. For side chain resonance assignment, 3D HBHA(CO)NH, (H)CCH-TOCSY, and H(C)CH-TOCSY experiments were used. All spectra were recorded and processed using Bruker TopSpinTM 3.5 and analyzed using the programs CARA³⁰ and CcpNmr Analysis³¹. [¹H]-¹⁵N-hetNOE experiments³² were carried out for the ¹⁵N-labeled ^NDD of KJ12C and the ¹⁵N-labeled ^NDD-^CDD linked complex construct and recorded using standard Bruker pulse sequences. Experiments were run twice in an

interleaved fashion with and without proton saturation during the recovery delay. Peak intensity differences were obtained using Bruker TopSpin™ 3.5 for peak integration and the peak intensity ratio was calculated as $I = I_x/I_0$. For titration experiments with NMR, C DD peptides (Kj12A- C DD, YLVKEKRKHQFE-QEKTQKLLFGHI; and Kj12B- C DD, LLKEKRKHQAEQNSSQEYLRGEI), both from JPT Peptides Technologies GmbH, or the short peptide variants were lyophilized twice and dissolved in water to measure the peptide concentrations with UV-vis spectroscopy. Therefore, the Kj12A- C DD peptide contains a non-native N-terminal tyrosine. For titration experiments, 1H , ^{15}N HSQCs or 1H , ^{15}N BEST-TROSY-HSQCs were recorded after each successive addition of unlabeled C DD (25–500 μ M) to a 100 μ M ^{15}N N DD protein sample in the identical buffer. To evaluate NMR titration experiments, the chemical shifts were determined using the peak picking function of CcpNmr Analysis³¹. The chemical shift differences were calculated using the following function³³:

$$\Delta\delta = \sqrt{\Delta\delta_H^2 + \left(\frac{\Delta\delta_N}{6.5}\right)^2} \quad (1)$$

Structure calculation. ^{15}N -nuclear Overhauser spectroscopy (NOESY)-HSQC, ^{13}C -NOESY-HSQC (aliphatic carbons), and ^{13}C -NOESY-HSQC (aromatic carbons) experiments in H_2O with mixing times of 250 ms were used to obtain distance restraints. The TALOS-N server was used to generate torsion angle restraints³⁴ based on backbone H, N, Ca, C β , and CO chemical shifts and used for all residues with ^{15}N hetNOE values >0.5 . Peak picking and NOE assignment was performed with the ATNOS/CANDID module in UNIO³⁵ in combination with CYANA³⁶ using the 3D NOESY spectra. To correct falsely picked artifacts, the peak lists were reviewed manually and corrected. Distance restraints were obtained using the automated NOE assignment and structure calculation protocol available in CYANA³⁶. An assignment of $>93\%$ of the observable NOESY crosspeaks for all NOESY spectra was achieved. Restrained energy refinement with OPALp³⁷ and the AMBER94 force field³⁸ of the 20 structures with the lowest target function was carried out. With the CYANA “regularized”-macro, a single representative structure was obtained³⁹. This representative mean structure as well as the 19 conformers with the lowest CYANA target function were used for the structure validation with the Protein Structure Validation Software suite1.5⁴⁰ restricted to residues with hetNOE values >0.6 . Electrostatic surface potential calculations were conducted with the PDB2PQR web server⁴¹ using the PARSE force field and visualized with the APBS plug-in⁴² for PyMOL with a threshold for electrostatic potential shading from -1 kT/e to $+1$ kT/e (k = Boltzmann’s constant, T = absolute temperature, and e = electron charge (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC)). All figures of structures were prepared with PyMOL.

ITC calorimetry. ITC measurements were performed at 20 °C in 50 mM sodium phosphate buffer, pH 6.5, and 100 mM NaCl using a MicroCal iTC200 (Malvern Instruments) or a Nano ITC Low Volume (TA Instruments) calorimeter. In all, 50 μ M of the N DDs were titrated with 1 mM of the C DD peptides and their variants. In order to measure the affinity of the Kj12B T-domain- C DD for the Kj12C N DD, a 50 mM solution of the Kj12B T-domain- C DD construct was placed in the cell and titrated with a 1 mM solution of the Kj12C N DD in the syringe due to the lower solubility of the Kj12B T-domain- C DD construct. NMR experiments showed that Kj12C N DD does not aggregate at a concentration of 1 mM and remains well folded. ITC experiments started with an initial waiting time of 120 s. The first injection of 0.2 μ l was followed by 19 serial injections of 2 μ l, separated by an interval of 180 s. For each experiment, the reference power was set to 11 μ cal $^{-1}$, stirring speed to 750 rpm (75 rpm for the Nano ITC), and the high feedback mode was selected. Three independent titrations were performed for each combination of DDs. The thermograms were processed using Origin7.0 (OriginLab) or NanoAnalyze Data Analysis 3.7.5 (TA Instruments) assuming a one site binding model. For the all ITC measurements where the protein was saturated with peptide during the final titration steps and a clear plateau was reached in the titrations, this plateau was used for baseline correction. In titrations of the Kj12B- N DD with Kj12B- C DD and Kj12B- C DD E1567R, respectively, full saturation was not reached. In these cases, the C DD peptides were titrated into buffer and the resulting thermograms were subtracted from those of the actual titration. c -Values were calculated by using Eq. (2):

$$c = nK_a[M]_T \quad (2)$$

where K_a is the binding constant, $[M]_T$ the total macromolecular concentration in the cell, and n is the stoichiometry of interaction⁴³. Reliable binding constants were determined for ITC data with c -values >1 and are listed in Supplementary Table 2. For ITC data with c -values <1 , the determination of binding constants is not reliable⁴⁴. If there is apparent binding in the ITC curves and in NMR titration experiments, we labeled these interactions only qualitatively as weak binding.

Alignments. Sequence alignments of C DDs were carried out with the multiple sequence alignment tool Clustal Omega⁴⁵. Owing to the low sequence identity between RXP N DDs and the tubulysin DD (TubC- N DD), a structural alignment was done with the Espresso mode of the T-Coffee alignment web server⁴⁶. All alignments were visualized with Boxshade.

CD measurements. CD measurements were carried out on a Jasco-815 CD spectrometer (Jasco, Gross-Umstadt, Germany) with 1 mm quartz cuvettes. In all, 50 μ M peptide samples were dissolved in 10 mM sodium phosphate buffer at pH 6.5 with 25 mM sodium chloride. Baseline corrections were performed automatically and automatic averaging of three measurements was performed. Spectra were recorded at 293 K in a spectral range between 190 and 290 nm with 1 nm scanning intervals, 5 nm bandwidth and 50 nm min $^{-1}$ scanning speed.

Data availability

All structures have been deposited in the Protein Data Bank under the ID codes 6EWS (Kj12A- N DD), 6EWT (Kj12B- N DD), 6EWU (Kj12C- N DD), and 6EWV (Kj12- N DD-Kj12B- C DD). Other data are available from the corresponding authors upon reasonable request.

Received: 20 December 2017 Accepted: 21 September 2018

Published online: 19 October 2018

References

- Baltz, R. H., Miao, V. & Wrigley, S. K. Natural products to drugs: daptomycin and related lipopeptide antibiotics. *Nat. Prod. Rep.* **22**, 717–725 (2005).
- Lawen, A. Biosynthesis of cyclosporins and other natural peptidyl prolyl cis/trans isomerase inhibitors. *Biochem. Biophys. Acta* **1850**, 2111–2120 (2015).
- Fischbach, M. A. & Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev.* **106**, 3468–3496 (2006).
- Süssmuth, R. D. & Mainz, A. Nonribosomal peptide synthesis-principles and prospects. *Angew. Chem. Int. Ed. Engl.* **56**, 3770–3821 (2017).
- Hahn, M. & Stachelhaus, T. Selective interaction between nonribosomal peptide synthetases is facilitated by short communication-mediating domains. *Proc. Natl Acad. Sci. USA* **101**, 15585–15590 (2004).
- Hahn, M. & Stachelhaus, T. Harnessing the potential of communication-mediating domains for the biocombinatorial synthesis of nonribosomal peptides. *Proc. Natl Acad. Sci. USA* **103**, 275–280 (2006).
- Chiocchini, C., Linne, U. & Stachelhaus, T. In vivo biocombinatorial synthesis of lipopeptides by COM domain-mediated reprogramming of the surfactin biosynthetic complex. *Chem. Biol.* **13**, 899–908 (2006).
- Dehling, E. et al. Mapping of the communication-mediating interface in nonribosomal peptide synthetases using a genetically encoded photocrosslinker supports an upside-down helix-hand motif. *J. Mol. Biol.* **428**, 4345–4360 (2016).
- Broadhurst, R. W., Nietlispach, D., Wheatcroft, M. P., Leadlay, P. F. & Weissman, K. J. The structure of docking domains in modular polyketide synthases. *Chem. Biol.* **10**, 723–731 (2003).
- Orival, J. et al. Characterization of intersubunit communication in the virginiamycin trans-acyl transferase polyketide synthase. *J. Am. Chem. Soc.* **138**, 4155–4167 (2016).
- Buchholz, T. J. et al. Structural basis for binding specificity between subclasses of modular polyketide synthase docking domains. *ACS Chem. Biol.* **4**, 41–52 (2009).
- Whicher, J. R. et al. Cyanobacterial polyketide synthase docking domains: a tool for engineering natural product biosynthesis. *Chem. Biol.* **20**, 1340–1351 (2013).
- Richter, C. D., Nietlispach, D., Broadhurst, R. W. & Weissman, K. J. Multienzyme docking in hybrid megasynthetases. *Nat. Chem. Biol.* **4**, 75–81 (2008).
- Dowling, D. P. et al. Structural elements of an NRPS cyclization domain and its intermodule docking domain. *Proc. Natl Acad. Sci. USA* **113**, 12432–12437 (2016).
- Cai, X. et al. Entomopathogenic bacteria use multiple mechanisms for bioactive peptide library design. *Nat. Chem.* **9**, 379–386 (2017).
- Montelione, G. T. et al. Recommendations of the wwPDB NMR Validation Task Force. *Structure* **21**, 1563–1570 (2013).
- Kronenwerth, M. et al. Characterisation of taxillalids A-G; natural products from *Xenorhabdus indica*. *Chem. Eur. J.* **20**, 17478–17487 (2014).
- Weinig, S., Hecht, H.-J., Mahmud, T. & Müller, R. Melithiazol biosynthesis: further insights into myxobacterial PKS/NRPS systems and evidence for a new subclass of methyl transferases. *Chem. Biol.* **10**, 939–952 (2003).

19. Nishizawa, T. et al. Characterization of the locus of genes encoding enzymes producing heptadepsipeptide micropeptin in the unicellular cyanobacterium *Microcystis*. *J. Biochem.* **149**, 475–485 (2011).
20. Rounge, T. B., Rohrlack, T., Tooming-Klunderud, A., Kristensen, T. & Jakobsen, K. S. Comparison of cyanopeptolin genes in *Planktothrix*, *Microcystis*, and *Anabaena* Strains: evidence for independent evolution within each genus. *Appl. Environ. Microbiol.* **73**, 7322–7330 (2007).
21. Cheng, Y. Q., Yang, M. & Matter, A. M. Characterization of a gene cluster responsible for the biosynthesis of anticancer agent FK228 in *Chromobacterium violaceum* No. 968. *Appl. Environ. Microbiol.* **73**, 3460–3469 (2007).
22. Kitagaki, J., Shi, G., Miyauchi, S., Murakami, S. & Yang, Y. Cyclic depsipeptides as potential cancer therapeutics. *Anti-Cancer Drugs* **26**, 259–271 (2015).
23. Kries, H. Biosynthetic engineering of nonribosomal peptide synthetases. *J. Pept. Sci.* **22**, 564–570 (2016).
24. Bozhüyük, K. A. J. et al. De novo design and engineering of non-ribosomal peptide synthetases. *Nat. Chem.* **10**, 275–281 (2018).
25. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).
26. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
27. Markley, J. L. et al. Recommendations for the presentation of NMR structures of proteins and nucleic acids. IUPAC-IUBMB-IUPAB Inter-Union Task Group on the Standardization of Data Bases of Protein and Nucleic Acid Structures Determined by NMR Spectroscopy. *J. Biomol. NMR* **12**, 1–23 (1998).
28. Sattler, M. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**, 93–158 (1999).
29. Favier, A. & Brutscher, B. Recovering lost magnetization: polarization enhancement in biomolecular NMR. *J. Biomol. NMR* **49**, 9–15 (2010).
30. Keller, R. L. J. *The Computer Aided Resonance Assignment Tutorial* (CANTINA Verlag, Goldau, Switzerland, 2004).
31. Vranken, W. F. et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
32. Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. A heteronuclear correlation experiment for simultaneous determination of ¹⁵N longitudinal decay and chemical exchange rates of systems in slow equilibrium. *J. Biomol. NMR* **4**, 727–734 (1994).
33. Mulder, F. A., Schipper, D., Bott, R. & Boelens, R. Altered flexibility in the substrate-binding site of related native and engineered high-alkaline *Bacillus subtilis*. *J. Mol. Biol.* **292**, 111–123 (1999).
34. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).
35. Herrmann, T., Güntert, P. & Wüthrich, K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* **319**, 209–227 (2002).
36. Güntert, P. Automated structure determination from NMR spectra. *Eur. Biophys. J.* **38**, 129–143 (2008).
37. Koradi, R., Billeter, M. & Güntert, P. Point-centered domain decomposition for parallel molecular dynamics simulation. *Comput. Phys. Commun.* **124**, 139–147 (2000).
38. Ponder, J. W. & Case, D. A. Force fields for protein simulations. *Adv. Protein Chem.* **66**, 27–85 (2003).
39. Gottstein, D., Kirchner, D. K. & Güntert, P. Simultaneous single-structure and bundle representation of protein NMR structures in torsion angle space. *J. Biomol. NMR* **52**, 351–364 (2012).
40. Bhattacharya, A., Tejero, R. & Montelione, G. T. Evaluating protein structures determined by structural genomics consortia. *Proteins* **66**, 778–795 (2006).
41. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, W665–W667 (2004).
42. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA* **98**, 10037–10041 (2001).
43. Wiseman, T., Williston, S., Brandts, J. F. & Lin, L. N. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **179**, 131–137 (1989).
44. Velázquez-Campoy, A., Ohtaka, H., Nezami, A., Muzammil, S. & Freire, E. Isothermal titration calorimetry. *Curr. Protoc. Cell Biol.* **Chapter 17**, Unit 17.8 (2004).
45. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 1–6 (2011).
46. Notredame, C., Higgins, D. G. & Heringa, J. T-coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217 (2000).

Acknowledgements

This work was supported by the LOEWE program (Landes-Offensive zur Entwicklung wissenschaftlich-ökonomischer Exzellenz) of the state of Hesse and was conducted within the framework of the MegaSyn Research Cluster in the laboratories of H.B.B. and J.W. Additionally, work in the Bode laboratory was supported by an ERC Starting Grant (grant agreement number 311477). All NMR experiments were carried out at the Center for Biomolecular Magnetic Resonance (BMRZ) at Goethe-University Frankfurt, which is supported by the state of Hesse. We are grateful to Dr. Frank Löhr and Dr. Christian Richter for their support in setting up BEST-TROSY-based experiments and to Dr. Elke Duchardt-Ferner for helpful discussions. The authors acknowledge the Deutsche Forschungsgemeinschaft for funding of the Impact II qTof mass spectrometer (INST 161/810–1).

Author contributions

H.B.B. and J.W. designed the experiments. C.H. performed all NMR and ITC experiments and elucidated the DD structures. J.P.W. analyzed NMR data. Constructs for the production of DDs were provided by C.K. and X.C. X.C. carried out all in vivo experiments, including strain construction and natural product quantification. Chemical synthesis of ¹³C-DD peptides was performed by L.Z. A.K.W. conducted ITC experiments. C.H., X.C., J.W., and H.B.B. wrote the paper with input from all authors.

Additional information

Supplementary Information accompanies this paper at <https://doi.org/10.1038/s41467-018-06712-1>.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018