



# NMR resonance assignments for a docking domain pair with an attached thiolation domain from the PAX peptide-producing NRPS from *Xenorhabdus cabanillasii*

Jonas Watzel<sup>1</sup> · Sepas Sarawi<sup>1,2</sup> · Elke Duchardt-Ferner<sup>2</sup> · Helge B. Bode<sup>1,3,4,5</sup> · Jens Wöhnert<sup>2</sup>

Received: 12 December 2020 / Accepted: 16 February 2021 / Published online: 5 March 2021  
© The Author(s) 2021

## Abstract

Non-ribosomal peptide synthetases (NRPSs) are large multienzyme machineries. They synthesize numerous important natural products starting from amino acids. For peptide synthesis functionally specialized NRPS modules interact in a defined manner. Individual modules are either located on a single or on multiple different polypeptide chains. The “peptide-antimicrobial-*Xenorhabdus*” (PAX) peptide producing NRPS PaxS from *Xenorhabdus* bacteria consists of the three proteins PaxA, PaxB and PaxC. Different docking domains (DDs) located at the N-termini of PaxB and PaxC and at the C-termini of PaxA and PaxB mediate specific non-covalent interactions between them. The N-terminal docking domains precede condensation domains while the C-terminal docking domains follow thiolation domains. The binding specificity of individual DDs is important for the correct assembly of multi-protein NRPS systems. In many multi-protein NRPS systems the docking domains are sufficient to mediate the necessary interactions between individual protein chains. However, it remains unclear if this is a general feature for all types of structurally different docking domains or if the neighboring domains in some cases support the function of the docking domains. Here, we report the <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N NMR resonance assignments for a C-terminal di-domain construct containing a thiolation (T) domain followed by a C-terminal docking domain (<sup>C</sup>DD) from PaxA and for its binding partner – the N-terminal docking domain (<sup>N</sup>DD) from PaxB from the Gram-negative entomopathogenic bacterium *Xenorhabdus cabanillasii* JM26 in their free states and for a 1:1 complex formed by the two proteins. These NMR resonance assignments will facilitate further structural and dynamic studies of this protein complex.

**Keywords** NMR assignments · Peptide-antimicrobial-*Xenorhabdus* (PAX) peptide · Non-ribosomal peptide synthetase (NRPS) · Docking domain · Thiolation domain

✉ Jonas Watzel  
watzel@bio.uni-frankfurt.de

✉ Jens Wöhnert  
woehnert@bio.uni-frankfurt.de

<sup>1</sup> Molecular Biotechnology, Institute of Molecular Biosciences, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

<sup>2</sup> Institute of Molecular Biosciences and Center for Biomolecular Magnetic Resonance (BMRZ), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

<sup>3</sup> Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

<sup>4</sup> Senckenberg Gesellschaft Für Naturforschung, 60325 Frankfurt am Main, Germany

<sup>5</sup> Department of Natural Products in Organismic Interactions, Max-Planck-Institute for Terrestrial Microbiology, 35043 Marburg, Germany

## Biological context

Non-ribosomal peptide synthetases (NRPS) use amino acids as building blocks for the synthesis of complex natural products. Each amino acid is incorporated into the final product by an individual specialized functional module of the NRPS. In turn, each individual module contains a number of enzymatically distinct catalytic domains. A typical NRPS elongation module contains at least an adenylation (A), a thiolation (T) and a condensation (C) domain. The adenylation (A) domain uses ATP to activate a specific amino acid as aminoacyl adenylate and then transfers the activated amino acid to the neighbouring thiolation (T) domain. The T domain in its *holo* form contains a phosphopantetheinyl (Ppant) moiety derived from coenzyme A covalently bound to a conserved serine side chain. The activated amino acid reacts with the terminal thiol group

of the Ppant moiety to form a thioester. The T domain is followed by a condensation domain (C) which catalyses the formation of peptide bonds between the amino acid bound to the T domain of its own module and an amino acid or a peptide chain bound to the T domain of the downstream module. Additional domains with enzymatic activities such as methylation or epimerization can be included in individual modules and may further modify the peptide product. Finally, the nascent peptide is released by a thioesterase (TE) domain localized at the C-terminus of a specialized termination module. TE domains can release either linear, cyclic or branched cyclic peptides (Süssmuth and Mainz 2017).

The individual functional modules necessary for the step-wise incorporation of each amino acid into the final product can be located either on a single long protein chain or on multiple proteins. When all modules are arranged on a single protein chain their linear order directly predicts (Mootz et al. 2002) the order of amino acid building blocks in the final product. In multi-protein NRPS systems specific non-covalent interactions between the individual protein chains determine the functional assembly of the NRPS complex and thereby the order of amino acids in the synthesized peptide. Hahn and Stachelhaus (Hahn and Stachelhaus 2004) were first in identifying short amino acid sequences at the N- and C-termini of the individual protein chains in a multi-protein NRPS that were capable of mediating specific non-covalent interactions between these protein chains. They referred to these sequences as “communication-mediating (COM) domains”. Functionally similar structural elements, which were named “docking domains (DD)” (Broadhurst et al. 2003), were identified in polyketide synthase (PKS) complexes which possess a multi-modular architecture similar to NRPS systems. The term “docking domain” is now more commonly used for domains that enable non-covalent specific interactions between individual protein chains in both PKS and NRPS complexes. Multiple structurally diverse families of DD architectures have been described so far (Broadhurst et al. 2003; Buchholz et al. 2009; Dorival et al. 2016; Hacker et al. 2018; Watzel et al. 2020; Whicher et al. 2013). Many DD structural families are dominated by  $\alpha$ -helical secondary structure elements (Buchholz et al. 2009; Watzel et al. 2020; Whicher et al. 2013). The binding affinities even for specific interactions between DDs are relatively weak with typical dissociation constants for the DD complexes in the range between 5  $\mu$ M and 25  $\mu$ M (Dorival et al. 2016; Hacker et al. 2018; Watzel et al. 2020; Whicher et al. 2013). Nevertheless, for many types of DD interactions it has been experimentally demonstrated that the DDs act independently from the other functional domains in an NRPS or PKS in order to mediate the non-covalent interactions between protein chains needed for the functional assembly of functional megasynthase complexes (Dorival

et al. 2016; Hacker et al. 2018). However, it is not yet clear if this is true for all types of docking domains.

The PaxS NRPS found in Gram-negative entomopathogenic bacteria from the genus *Xenorhabdus* is a prototypical example for a multi-protein assembly line. It consists of the three proteins PaxA, PaxB and PaxC. The three proteins interact with each other non-covalently in a unidirectional manner where PaxA is bound by PaxB and PaxB is bound by PaxC. The DD pair mediating the specific interaction between the C-terminus of PaxB and the N-terminus of PaxC has been characterized functionally, structurally and biophysically in previous work (Kegler and Bode 2020; Watzel et al. 2020). A structural basis for the specific interaction between the C-terminus of PaxA and the N-terminus of PaxB has not yet been established. Notably, the putative C-terminal docking domain (<sup>C</sup>DD) in PaxA was predicted to be very short in comparison to other structurally characterized types of docking domains. Thus, in this case the T domain directly preceding the <sup>C</sup>DD of PaxA might play a supporting role in establishing a specific non-covalent interaction with the N-terminal docking domain (<sup>N</sup>DD) of PaxB. As a prerequisite for investigating a putative role of the T domain in this DD interaction and for establishing a structural basis for the specific interaction between PaxA and PaxB in the PaxS NRPS from *Xenorhabdus cabanillasii* JM26 we present here the NMR resonance assignments for a PaxA T<sub>1</sub>-<sup>C</sup>DD di-domain construct and for the PaxB <sup>N</sup>DD in their free states and in a 1:1 non-covalent complex. The T<sub>1</sub>-<sup>C</sup>DD (amino acids 981–1084 of PaxA) comprises 104 amino acids (12 kDa), whereas the <sup>N</sup>DD (amino acids 1–30 of PaxB) contains only 30 amino acids (3.6 kDa).

## Methods and experiments

### Cloning, expression and purification

The DNA coding sequences for PaxA T<sub>1</sub>-<sup>C</sup>DD and PaxB <sup>N</sup>DD were generated by PCR amplification using the genomic DNA from *Xenorhabdus cabanillasii* JM26 as the template. All protein sequences referred to in this work are based on the UniProt Archive (UniParc) entries for PaxA (UPI0003E57C57) and PaxB (UPI000C04EDD1) based on a genome assembly for *Xenorhabdus cabanillasii* JM26 produced in our group (NCBI: ASM263290v1; GenBank: NJGH00000000 (Tobias et al. 2017)). For protein concentration measurements a codon for a tyrosine residue was added downstream of the PaxB <sup>N</sup>DD coding sequence via the used primer. The inserts were cloned into a modified pET-11a vector (Hacker et al. 2018) containing the sequence for an N-terminal hexahistidine (His<sub>6</sub>)-tag followed by a SUMO (SMT3)-tag which also functions as a cleavage site for the ULP1 protease (Malakhov et al. 2004). The PaxA T<sub>1</sub>-<sup>C</sup>DD

was expressed in its *apo* form by using *Escherichia coli* BL21(DE3) $\Delta$ *entD* (Owen et al. 2012) cells preventing post-translational modification of the T domain by the covalent addition of a phosphopantetheinyl moiety, whereas the PaxB<sup>NDD</sup> was expressed in *E. coli* BL21(DE3) Gold (Agilent Technologies/Stratagene) cells.

Protein expression was induced at OD<sub>600</sub> ~ 0.7 with 1 mM IPTG at 20 °C for ~ 18 h in media supplemented with ampicillin (100 µg/ml). The expression of uniformly <sup>15</sup>N- and <sup>15</sup>N, <sup>13</sup>C-labelled proteins was achieved by using M9 minimal medium supplemented with 1 g/l <sup>15</sup>NH<sub>4</sub>Cl or 1 g/l <sup>15</sup>NH<sub>4</sub>Cl and 2.5 g/l <sup>13</sup>C<sub>6</sub>-D-glucose (Cambridge Isotope Laboratories). For the stereospecific NMR assignment of  $\gamma^{1/2}$  and  $\delta^{1/2}$  methyl groups of valine and leucine, the PaxA T<sub>1</sub>-<sup>C</sup>DD as well as the PaxB<sup>NDD</sup> were expressed as fractionally <sup>13</sup>C-labelled proteins in M9 minimal medium containing a mixture of 0.25 g L<sup>-1</sup> <sup>13</sup>C<sub>6</sub>-D-glucose and 2.25 g L<sup>-1</sup> unlabelled glucose as the sole carbon source (Neri et al. 1989).

Cell lysis of *E. coli* cells was accomplished by sonication in a buffer containing 50 mM sodium phosphate, pH 8.0, 300 mM NaCl, 1 mM EDTA, 10 mM MgCl<sub>2</sub>, 2 mM  $\beta$ -mercaptoethanol, Benzonase (Merck) and cOmplete protease inhibitor (Roche). The cell debris was cleared from the lysate by centrifugation (8000  $\times$  g, 4 °C, 30 min) and the supernatant was run through a HisTrap HP column (GE Healthcare). With the help of a recombinantly produced His<sub>6</sub>-tagged ULP1 protease the His<sub>6</sub>-tagged SUMO (SMT3)-tag was cleaved off from the recombinant fusion proteins, leading to native target protein sequences. Both His<sub>6</sub>-tagged ULP1 and the SUMO-tag were separated from the protein of interest by a second immobilized metal ion affinity chromatography using a HisTrap HP column, followed by gel filtration chromatography with a HiPrep 16/60 Sephacryl S-100 high resolution column (GE Healthcare). For NMR measurements, the samples of the individual proteins (300 µM protein concentration) and the PaxA T<sub>1</sub>-<sup>C</sup>DD:PaxB<sup>NDD</sup>/PaxB<sup>NDD</sup>:PaxA T<sub>1</sub>-<sup>C</sup>DD complexes (300 µM protein A concentration: 360 µM protein B concentration) were prepared in a buffer containing 50 mM sodium phosphate, pH 6.5, 100 mM NaCl, 2 mM  $\beta$ -mercaptoethanol and 5% (v/v) D<sub>2</sub>O.

## NMR spectroscopy

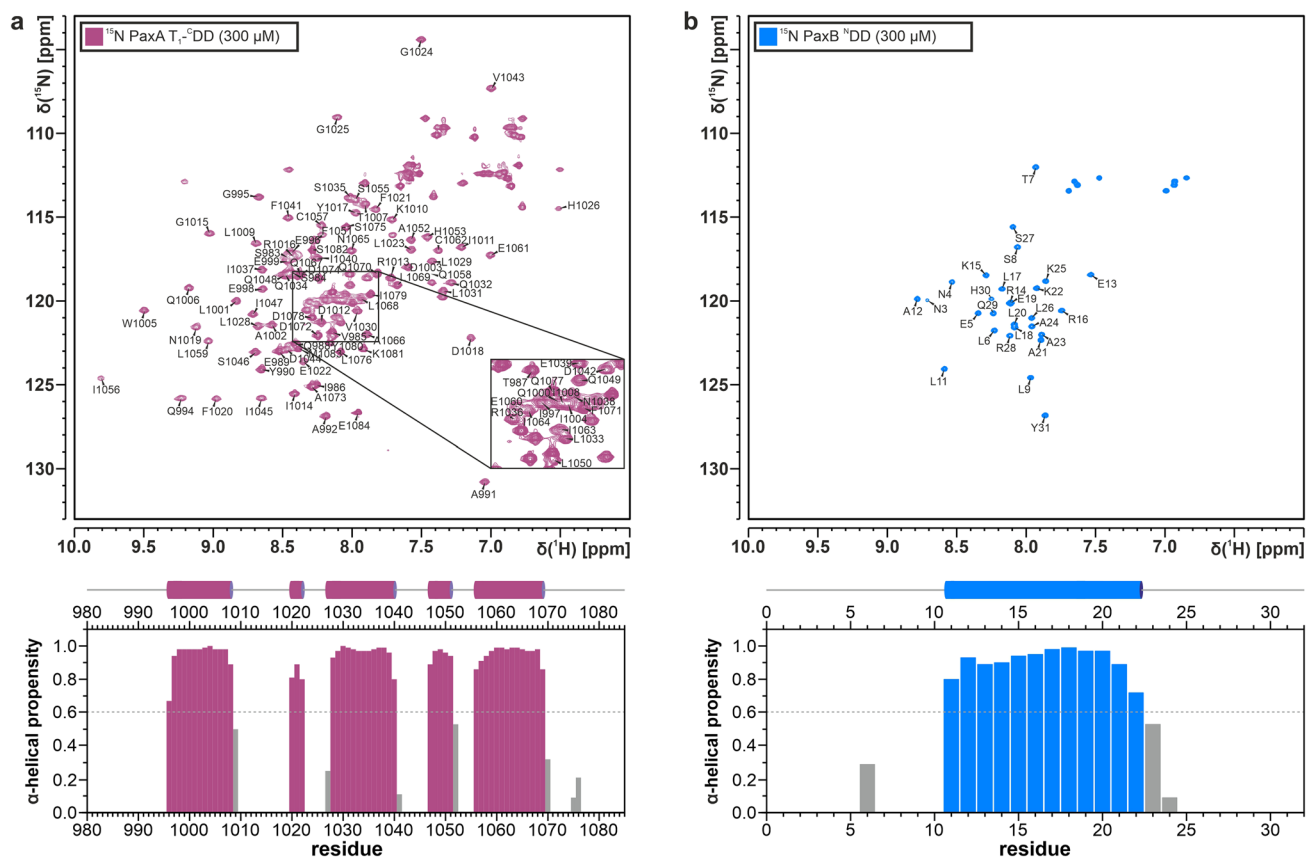
NMR experiments were recorded at 293 K on Bruker AVANCE III HD 600, 700 and 800 MHz spectrometers, each of them equipped with 5 mm cryogenic triple resonance probes. <sup>1</sup>H chemical shifts were internally referenced to DSS, whereas the heteronuclear <sup>13</sup>C and <sup>15</sup>N chemical shifts were indirectly referenced with the appropriate conversion factors (Markley et al. 1998). Spectra were processed using TOPSPIN 3.6.2 (Bruker) and analysed with CARA (Keller 2004). The secondary structure of the unbound and

bound PaxA T<sub>1</sub>-<sup>C</sup>DD and PaxB<sup>NDD</sup> was derived from TALOS-N (Shen and Bax 2013) based on the chemical shift assignments.

For the backbone assignment of the free PaxA T<sub>1</sub>-<sup>C</sup>DD a uniformly <sup>13</sup>C, <sup>15</sup>N-labelled sample was used and the following triple resonance experiments were recorded: HNCO, HNCA, HNCACB, HBHA(CO)NH, CBCA(CO)NH (Sattler 1999). The backbone and side chain assignments of the unbound <sup>13</sup>C, <sup>15</sup>N-labelled PaxB<sup>NDD</sup> were derived from 3D HNCO, HNCACB, HBHA(CO)NH, H(CCO)NH (mixing time 12 ms) and (H)C(CO)NH (mixing time 12 ms) experiments (Sattler 1999). The backbone resonances of <sup>13</sup>C, <sup>15</sup>N-labelled PaxA T<sub>1</sub>-<sup>C</sup>DD/PaxB<sup>NDD</sup> in complex with a 1.2-fold excess of unlabelled PaxB<sup>NDD</sup>/PaxA T<sub>1</sub>-<sup>C</sup>DD were assigned on the basis of 3D HNCO, HNCA, HNCACB, and CBCA(CO)NH experiments (Sattler 1999). Side chain assignments were obtained from 3D HBHA(CO)NH, H(C)CH/(H)CCH-TOCSY (mixing times 12 ms) (Sattler 1999) and H(C)CH-COSY (Bax et al. 1990) experiments. Stereospecific assignments of valine  $\gamma^{1/2}$  and leucine  $\delta^{1/2}$  methyl groups of the bound PaxA T<sub>1</sub>-<sup>C</sup>DD and free and bound PaxB<sup>NDD</sup> were determined in <sup>1</sup>H, <sup>13</sup>C-HSQC experiments with a resolution in the <sup>13</sup>C dimension of ~ 28 Hz, ~ 26 Hz and ~ 23 Hz, respectively. This allowed the unambiguous discrimination between the signals for the  $\gamma^2/\delta^2$  CH<sub>3</sub> groups of valine and leucine which appear as singlets in the <sup>13</sup>C-dimension and the signals of the  $\gamma^1/\delta^1$  CH<sub>3</sub> groups which appear as doublets separated by the <sup>1</sup>J<sub>13C,13C</sub> coupling constant of ~ 33 Hz (Neri et al. 1989).

## Assignment and data deposition

The protein construct PaxA T<sub>1</sub>-<sup>C</sup>DD contains amino acids 981 to 1084 of the PaxA protein from *Xenorhabdus cabanillasii* JM26. It has a molecular weight of 12 kDa and consists of 104 amino acid residues. Most of the backbone amid signals are well dispersed in the <sup>1</sup>H, <sup>15</sup>N-HSQC spectrum of the unbound PaxA T<sub>1</sub>-<sup>C</sup>DD as is typical for a well-folded protein. The <sup>1</sup>H, <sup>15</sup>N-HSQC is expected to contain 101 backbone amide signals since there is no observable amide signal for the N-terminal residue D981 and there are two proline residues (P993, P1054) present in the sequence. However, only 99 backbone amide signals were observed and assigned (99/101, 98.0%; Fig. 1a, top). No backbone amide signals were detectable for residues H982 and S1027 most likely due to fast exchange of the respective amide protons with the solvent or due to conformational exchange. S1027 is the residue that would be posttranslationally modified by addition of the Ppant-arm in the native environment. Furthermore, 99% of all C $\alpha$  (103/104), 98% of all C $\beta$  (98/100), 95% of all CO (99/104) and H $\alpha$  (99/104) and 94% of all H $\beta$  (94/100) chemical shifts were assigned. The backbone chemical shifts were used to derive the putative secondary structure of the



**Fig. 1** (a, top)  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectra and assigned backbone amide signals for the free PaxA  $\text{T}_1$ - $^{\text{CDD}}$  and (b, top) the free PaxB  $^{\text{NDD}}$ . The spectra were recorded at 293 K on a 600 MHz Bruker Avance III spectrometer. The assigned NMR signals are labelled with their

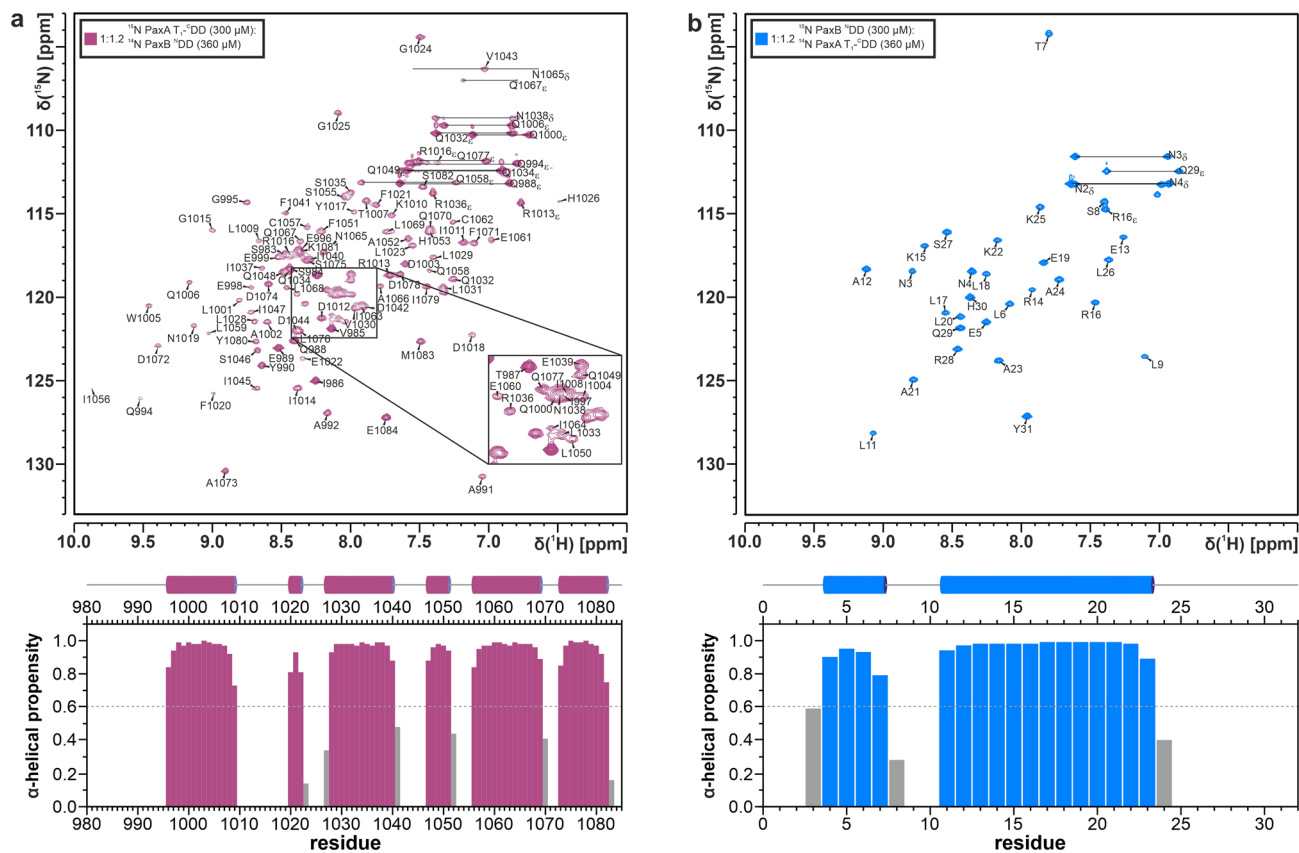
respective residue names and numbers and the central spectral region of the PaxA  $\text{T}_1$ - $^{\text{CDD}}$  with increased peak overlap is enlarged for a better overview (box). (a, b, bottom) TALOS-N based secondary structure analysis

free PaxA  $\text{T}_1$ - $^{\text{CDD}}$ . Based on these chemical shifts TALOS-N (Shen and Bax 2013) identified five amino acid stretches with high  $\alpha$ -helical propensity (Fig. 1a, bottom). This secondary structure fits well to that of previously described T domain structures, which typically feature a four-helix bundle (Weber et al. 2000). Compared to the canonical four-helix bundle T domain fold the PaxA  $\text{T}_1$  domain contains an additional very short (3 residues)  $\alpha$ -helix in the linker region between the canonical helices  $\alpha 1$  and  $\alpha 2$ . Such an additional helix is also present at this position in some other carrier protein structures (Lohman et al. 2014). According to TALOS-N the C-terminus including the predicted  $^{\text{CDD}}$  of the PaxA  $\text{T}_1$ - $^{\text{CDD}}$  construct is unstructured.

The PaxB  $^{\text{NDD}}$  construct contains amino acids 1–30 of the PaxB protein followed by a non-native tyrosine at the C-terminus and has a molecular weight of 3.6 kDa. Surprisingly, the backbone amide signals of the free PaxB  $^{\text{NDD}}$  are well dispersed in the  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectrum (Fig. 1b, top) indicating a folded protein. Overall, 27 backbone amide signals were assigned since the sequence contains one proline residue (P10) and no backbone amide signals were

observable for M1, N2 and the non-native Y31. Additionally, 93% of all CO (28/30) and all  $\text{C}\alpha$  (30/30),  $\text{C}\beta$  (30/30),  $\text{H}\alpha$  (30/30) and  $\text{H}\beta$  (30/30) chemical shifts were assigned. The overall side chain assignments were completed to 99.2% for aliphatic protons and to 99.0% for aliphatic carbons. No aromatic side chain NMR signals were assigned since the PaxB  $^{\text{NDD}}$  contains only a single native aromatic residue (H30) and the non-native Y31. The TALOS-N analysis of the secondary structure suggests the presence of a single continuous  $\alpha$ -helix including residues L11 to K22 in the free PaxB  $^{\text{NDD}}$  (Fig. 1b, bottom). The  $^{13}\text{C}\gamma$  and  $^{13}\text{C}\beta$  chemical shifts of proline P10 are in agreement with a *trans* conformation for this residue (Schubert et al. 2002). In addition, all valine  $\gamma^{1/2}$  and leucine  $\delta^{1/2}$  methyl groups of the unbound PaxB  $^{\text{NDD}}$  were stereospecifically assigned.

PaxA  $\text{T}_1$ - $^{\text{CDD}}$  and PaxB  $^{\text{NDD}}$  form a stable 1:1 complex according to analytical gel filtration and NMR titration experiments which is in slow exchange on the NMR time scale. Assignments for both proteins in their bound states were made using samples containing one binding partner in  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled form and a 1.2-fold excess of



**Fig. 2** (a, top)  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC spectra and assigned backbone amide signals for the bound PaxA  $\text{T}_1$ - $\text{CDD}$  and (b, top) the bound PaxB  $\text{NDD}$ . The spectra were recorded at 293 K on a 600 MHz Bruker Avance III spectrometer. The assigned NMR signals are labelled with their respective residue names and numbers. The central spectral region of the PaxA  $\text{T}_1$ - $\text{CDD}$  with increased peak overlap is enlarged

for a better overview (box). Assigned sidechain  $\text{NH}_2$  signals are indicated by horizontal bars and labelled with the respective residue names and numbers. The labelled  $\epsilon$  imino groups of R1013, R1016 and R1036 in the PaxA  $\text{T}_1$ - $\text{CDD}$  and of R16 in the PaxB  $\text{NDD}$  are folded into the spectrum. (a, b, bottom) TALOS-N based secondary structure analysis

the unlabelled binding partner. The backbone resonances of PaxA  $\text{T}_1$ - $\text{CDD}$  bound to the PaxB  $\text{NDD}$  could be assigned with the same degree of completeness as for the free protein. All expected backbone amide signals except those for H982 and S1027 (99/101, 98.0%; Fig. 2a, top) and all  $\text{C}\alpha$  (104/104), 99% of all  $\text{C}\beta$  (99/100), 95% of all CO (99/104), all  $\text{H}\alpha$  (104/104) and 99% of all  $\text{H}\beta$  (99/100) chemical shifts were assigned. The overall side chain assignments were completed to 98.7% for the remaining aliphatic protons and to 98.2% for aliphatic carbons. All valine  $\gamma^{1/2}$  and leucine  $\delta^{1/2}$  methyl groups were stereospecifically assigned. In addition, 56.3% of all aromatic side chain proton-bound carbon and carbon-bound proton signals were assigned.

The analysis of the backbone chemical shifts of the PaxA  $\text{T}_1$ - $\text{CDD}$  with TALOS-N in its bound state suggests that its secondary structure is very similar to that in its free state but that there is now an additional  $\alpha$ -helix present at the very C-terminus suggesting that the predicted docking domain becomes structured upon binding (Fig. 2a, bottom).

The backbone assignment of the bound PaxB  $\text{NDD}$  (Fig. 2b, top) includes 27 amide backbone signals, 93% of all CO (28/30) and all  $\text{C}\alpha$  (30/30),  $\text{C}\beta$  (30/30),  $\text{H}\alpha$  (30/30) and  $\text{H}\beta$  (30/30) chemical shifts. The overall side chain assignments were completed to 99.2% for aliphatic protons and to 99.0% for aliphatic carbons. All valine  $\gamma^{1/2}$  and leucine  $\delta^{1/2}$  methyl groups were stereospecifically assigned. In addition, 50% of all aromatic side chain proton-bound carbon and carbon-bound proton signals were assigned. The TALOS-N derived secondary structure for the bound PaxB  $\text{NDD}$  indicates the presence of two  $\alpha$ -helices (Fig. 2b, bottom). The proline residue P10 remains in the *trans* conformation in the bound state according to its  $^{13}\text{C}\gamma$  and  $^{13}\text{C}\beta$  chemical shifts (Schubert et al. 2002).

The assigned chemical shifts have been deposited in the BMRB under the accession numbers 50,594, 34,576 and 34,575 for the free PaxA  $\text{T}_1$ - $\text{CDD}$  construct, the free PaxB  $\text{NDD}$  and the complex, respectively.

**Acknowledgements** All NMR measurements were carried out at the Center for Biomolecular Magnetic Resonance (BMRZ) at the Goethe University Frankfurt which is generously supported by the State of Hesse. This work was further supported by the LOEWE program 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. Work in the Bode lab is also supported by an ERC Advanced Grant (grant agreement number 835108).

**Funding** Open Access funding enabled and organized by Projekt DEAL. The funding was supported by H2020 European Research Council (Grant No. 835108) and the LOEWE program of the state of Hesse (Grant No. MegaSyn Research Cluster).

**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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Bax A, Clore GM, Driscoll PC, Gronenborn AM, Ikura M, Kay LE (1990) Practical aspects of proton-carbon-carbon-proton three-dimensional correlation spectroscopy of  $^{13}\text{C}$ -labeled proteins. *J Magn Reson* 87:620–627
- Broadhurst RW, Nieltispach D, Wheatcroft MP, Leadlay PF, Weissman KJ (2003) The structure of docking domains in modular polyketide synthases. *Chem Biol* 10:723–731
- Buchholz TJ, Geders TW, Bartley FE, Reynolds KA, Smith JL, Sherman DH (2009) Structural basis for binding specificity between subclasses of modular polyketide synthase docking domains. *ACS Chem Biol* 4:41–52
- Dorival J, Annaïval T, Risser F, Collin S, Roblin P, Jacob C, Gruez A, Chagot B, Weissman KJ (2016) Characterization of intersubunit communication in the virginiamycin trans-acyl transferase polyketide synthase. *J Am Chem Soc* 138:4155–4167
- Hacker C, Cai X, Kegler C, Zhao L, Weickhmann AK, Wurm JP, Bode HB, Wöhnert J (2018) Structure-based redesign of docking domain interactions modulates the product spectrum of a rhabdo-peptide-synthesizing NRPS. *Nat Commun* 9:4366
- Hahn M, Stachelhaus T (2004) Selective interaction between nonribosomal peptide synthetases is facilitated by short communication-mediating domains. *Proc Natl Acad Sci U S A* 101:15585–15590
- Kegler C, Bode HB (2020) Artificial splitting of a non-ribosomal peptide synthetase by inserting natural docking domains. *Angew Chem Int Ed Engl* 59:13463–13467
- Keller R (2004) The computer aided resonance assignment tutorial, 1st edn. Cantina Verlag, Goldau
- Lohman JR, Ma M, Cuff ME, Bigelow L, Bearden J, Babnigg G, Joachimiak A, Phillips GN, Shen B (2014) The crystal structure of BlmI as a model for nonribosomal peptide synthetase peptidyl carrier proteins. *Proteins* 82:1210–1218
- Malakhov MP, Mattern MR, Malakhova OA, Drinker M, Weeks SD, Butt TR (2004) SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J Struct Func Genom* 5:75–86
- Markley JL, Bax A, Arata Y, Hilbers CW, Kaptein R, Sykes BD, Wright PE, Wüthrich K (1998) 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
- Mootz HD, Schwarzer D, Marahiel MA (2002) Ways of assembling complex natural products on modular nonribosomal peptide synthetases. *ChemBioChem* 3:490–504
- Neri D, Szyperski T, Otting G, Senn H, Wüthrich K (1989) Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional  $^{13}\text{C}$  labeling. *Biochemistry* 28:7510–7516
- Owen JG, Robins KJ, Parachin NS, Ackerley DF (2012) A functional screen for recovery of 4'-phosphopantetheinyl transferase and associated natural product biosynthesis genes from metagenome libraries. *Environ Microbiol* 14:1198–1209
- Sattler M (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog Nucl Magn Reson Spectrosc* 34:93–158
- Schubert M, Labudde D, Oschkinat H, Schmieder P (2002) A software tool for the prediction of Xaa-Pro peptide bond conformations in proteins based on  $^{13}\text{C}$  chemical shift statistics. *J Biomol NMR* 24:149–154
- Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* 56:227–241
- Süssmuth RD, Mainz A (2017) Nonribosomal peptide synthesis – principles and prospects. *Angew Chem Int Ed Engl* 56:3770–3821
- Tobias NJ, Wolff H, Djahanschiri B, Grundmann F, Kronenwerth M, Shi Y-M, Simonyi S, Grün P, Shapiro-Ilan D, Pidot SJ, Stinear TP, Ebersberger I, Bode HB (2017) Natural product diversity associated with the nematode symbionts *Photorhabdus* and *Xenorhabdus*. *Nat Microbiol* 2:1676–1685
- Watzel J, Hacker C, Duchardt-Ferner E, Bode HB, Wöhnert J (2020) A new docking domain type in the peptide-antimicrobial-*Xenorhabdus* peptide producing nonribosomal peptide synthetase from *Xenorhabdus bovienii*. *ACS Chem Biol* 15:982–989
- Weber T, Baumgartner R, Renner C, Marahiel MA, Holak TA (2000) Solution structure of PCP, a prototype for the peptidyl carrier domains of modular peptide synthetases. *Structure* 8:407–418
- Whicher JR, Smaga SS, Hansen DA, Brown WC, Gerwick WH, Sherman DH, Smith JL (2013) Cyanobacterial polyketide synthase docking domains: a tool for engineering natural product biosynthesis. *Chem Biol* 20:1340–1351

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