

**Supplementary Information: Structural basis for chain release from the enacyloxin polyketide synthase**

Simone Kosol,<sup>1,†</sup> Angelo Gallo,<sup>1,†</sup> Daniel Griffiths,<sup>1,†</sup> Timothy R. Valentic,<sup>2,†</sup> Joleen Masschelein,<sup>1</sup> Matthew Jenner,<sup>1,3</sup> Emmanuel L.C. de los Santos,<sup>3</sup> Lucio Manzi,<sup>4</sup> Paulina K. Sydor,<sup>1</sup> Dean Rea,<sup>5</sup> Shanshan Zhou,<sup>1</sup> Vilmos Fülöp,<sup>5</sup> Neil J. Oldham,<sup>4</sup> Shiou-Chuan Tsai,<sup>2</sup> Gregory L. Challis,<sup>1,3,6\*</sup> and Józef R. Lewandowski<sup>1\*</sup>

<sup>1</sup>Department of Chemistry, University of Warwick, Coventry CV4 7AL, UK

<sup>2</sup>Departments of Molecular Biology and Biochemistry, Chemistry, and Pharmaceutical Sciences, University of California, Irvine, CA 92697, USA

<sup>3</sup>Warwick Integrative Synthetic Biology Centre, University of Warwick, Coventry CV4 7AL, UK

<sup>4</sup>School of Chemistry, University of Nottingham, Nottingham NG7 2RD, UK

<sup>5</sup>School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK

<sup>6</sup>Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC 3800, Australia

<sup>†</sup>These authors contributed equally to this work.

\*Authors to whom the correspondence should be addressed. Email: [j.r.lewandowski@warwick.ac.uk](mailto:j.r.lewandowski@warwick.ac.uk) and [g.l.challis@warwick.ac.uk](mailto:g.l.challis@warwick.ac.uk)

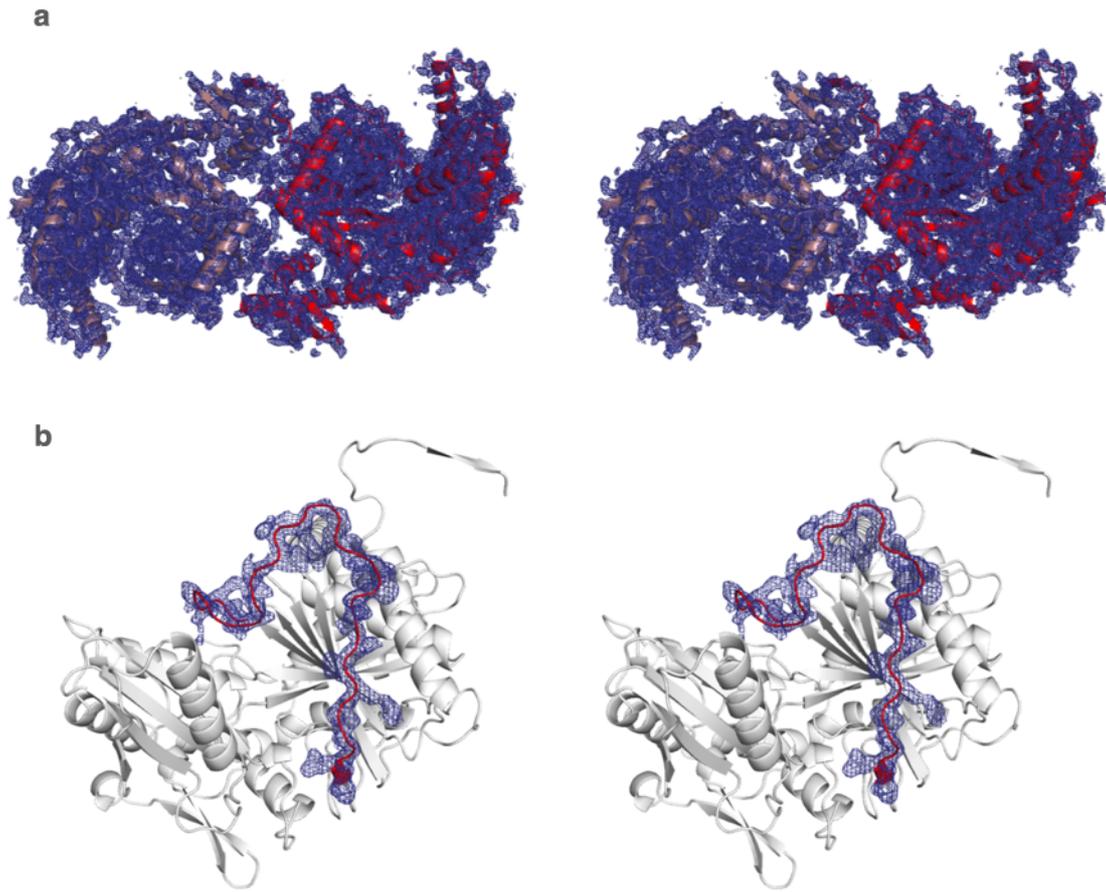
## Contents

### Figures

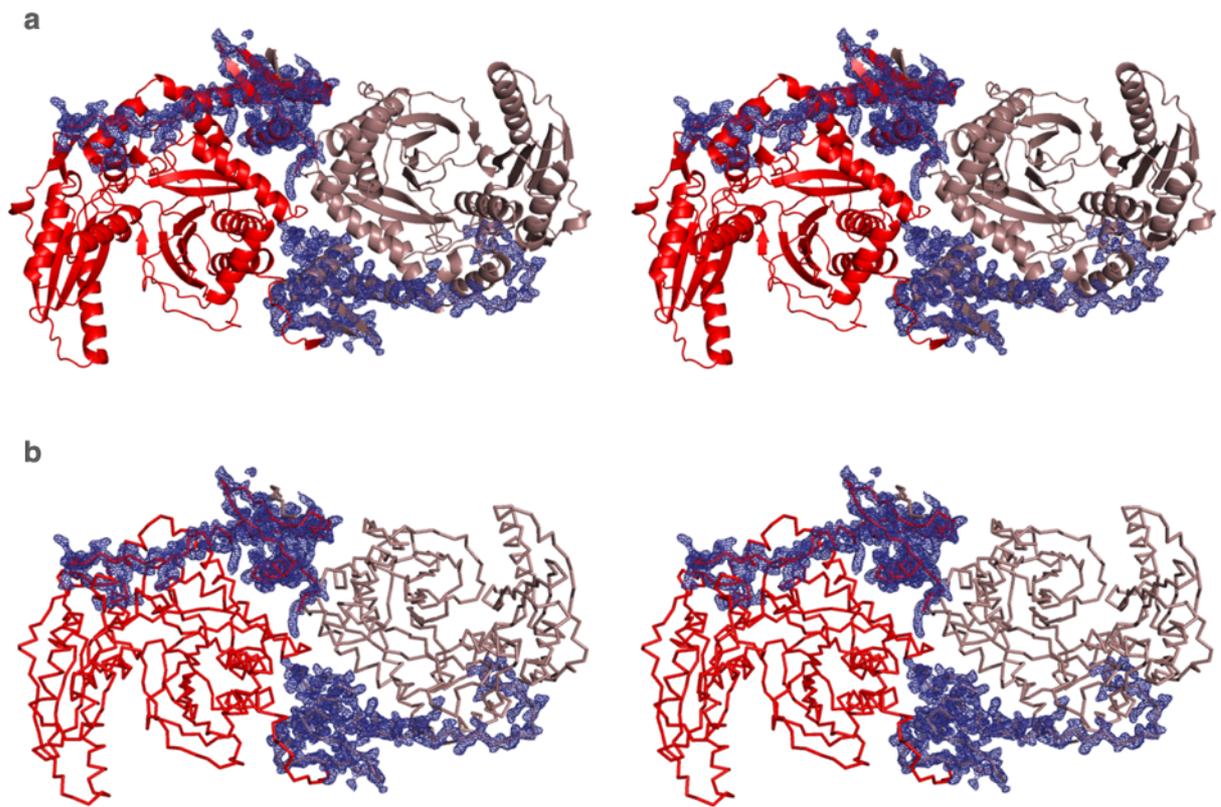
Supplementary Figure 1	Stereo views of Bamb_5915 and the latch region between $\beta$ -strands 12 & 13. ....	5
Supplementary Figure 2	Stereo views of Bamb_5915 highlighting its N-terminal $\beta$ HD domain. ....	6
Supplementary Figure 3	Stereo views of the active site of the Bamb_5915 C domain. ....	7
Supplementary Figure 4	Secondary structure of Bamb_5915. ....	7
Supplementary Figure 5	The crystal structure of Bamb_5915. ....	8
Supplementary Figure 6	Location of the donor channel in condensation domains. ....	9
Supplementary Figure 7	A structural comparison of Bamb_5915 to other C domain structures. ....	9
Supplementary Figure 8	Location and state of acceptor channel in condensation domains. ....	10
Supplementary Figure 9	Superimposition of Bamb_5915 with the most similar available structure, SgcC5. ...	11
Supplementary Figure 10	Comparison of known $\beta$ HD domain structures. ....	12
Supplementary Figure 11	Superimposition of Bamb_5915 with the multiple crystal forms of EpoBdd. ....	13
Supplementary Figure 12	A comparison of the Bamb_5917 PCP domain with other PCP domains. ....	14
Supplementary Figure 13	Solution NMR dynamics data for the apo and holo-Bamb_5917 PCP domain. ....	15
Supplementary Figure 14	Dynamics and order for Bamb_5917. ....	16
Supplementary Figure 15	2D $^1\text{H}$ - $^{15}\text{N}$ HSQC correlation spectra of the Bamb_5917 PCP domain in solution. .	16
Supplementary Figure 16	2D $^1\text{H}$ - $^{15}\text{N}$ HSQC correlation spectra of the holo-Bamb_5917 PCP domain in isolation and in the presence of the Bamb_5915 $\beta$ HD domain in solution. ....	17
Supplementary Figure 17	Residues with greater than average local apparent $K_a$ s in NMR titrations of the holo-Bamb_5917 PCP domain with the excised Bamb_5915 $\beta$ HD domain in solution with buried side chains. ....	17
Supplementary Figure 18	Solution NMR $^{15}\text{N}$ relaxation data for the holo-Bamb_5917 PCP domain in the presence of Bamb_5915. ....	18
Supplementary Figure 19	$^{15}\text{N}$ CEST profiles for the holo-Bamb_5917 PCP domain in the presence of Bamb_5915. ....	19
Supplementary Figure 20	Titrations profiles for the peaks disappearing in the NMR titration of the $^{15}\text{N}$ -labelled holo-Bamb_5917 PCP domain with Bamb_5915 in solution as a result of a complex formation. ....	20
Supplementary Figure 21	Titrations profiles for the peaks disappearing in the NMR titration of the $^{15}\text{N}$ -labelled holo-Bamb_5917 PCP domain with Bamb_5915 $\beta$ HD domain in solution as a result of a complex formation. .	21
Supplementary Figure 22	Titration profiles for the peaks appearing as a result of complex formation in the NMR spectra of the <i>holo</i> -Bamb_5917 PCP domain mixed with the Bamb_5915 $\beta$ HD domain in solution. ....	22
Supplementary Figure 23	Biolayer interferometry (BLI) data for binding of the Bamb_5917 PCP domain to Bamb_5915. ....	23
Supplementary Figure 24	Carbene footprinting of the holo-Bamb_5917 PCP domain in complex with Bamb_5915. ....	24
Supplementary Figure 25	Differential carbene labelling for tryptic and AspN fragments of Bamb_5915. ....	25
Supplementary Figure 26	Combined mapping of peptides from carbene footprinting on Bamb_5915 in complex with the holo-Bamb_5917 PCP domain. ....	25
Supplementary Figure 27	Interaction interface of TubCdd. ....	26
Supplementary Figure 28	Overlay of 2D $^1\text{H}$ - $^{15}\text{N}$ correlation spectra of the holo-Bamb_5917 PCP domain in a sedimented complex with Bamb_5915 in the solid state and in solution. ....	27
Supplementary Figure 29	Standard 100 ns molecular dynamics (MD) simulations of Bamb_5915. ....	28
Supplementary Figure 30	Solvent accessibility of the C-terminal part of the “latch” region in Bamb_5915 during aMD simulations. ....	29
Supplementary Figure 31	Solvent accessibility of $\alpha$ -helix 6 in Bamb_5915 during aMD simulations. ....	30
Supplementary Figure 32	Sum of solvent accessible area for residues in $\alpha$ -helix 10 and 12 in Bamb_5915 during aMD simulations. ....	31
Supplementary Figure 33	Correlation between conformations of domains with the CAT fold versus the state of the acceptor and donor channel. ....	32
Supplementary Figure 34	Examples of docking simulations for the Bamb_5917 PCP domain with Bamb_5915 performed in HADDOCK. ....	33
Supplementary Figure 35	Clusters of best solutions in docking simulations of the Bamb_5917 PCP domain:Bamb_5915 complex performed in HADDOCK using conformers from 500 ns aMD of the Bamb_5917 PCP domain sampled every 100 ns as starting points. ....	34
Supplementary Figure 36	Comparison of the orientation of PCP domains docked at the donor side of C domains. ....	35
Supplementary Figure 37	Interactions of residues from the latch with the phosphor diester group of the Ppart arm. ....	35

Supplementary Figure 38   a, Area under peaks due to N-acetyl-AHCCA in LC-MS chromatograms as a function of the molar ratio of the Bamb_5917 SLiM to the Bamb_5917 PCP domain, and the Bamb_5915 $\beta$ H <sub>2</sub> O domain to Bamb_5915 in the competition assays using the Bamb_5917 SLiM peptide (a) and the excised Bamb_5915 $\beta$ H <sub>2</sub> O domain (b).....	36
Supplementary Figure 39   Multiple sequence alignment for $\beta$ H <sub>2</sub> O domains appended to the N-terminus of dual C/E domains.....	36
Supplementary Figure 40   Sequence alignments of PCP domains. ....	37
Supplementary Figure 41   Sequence alignment of $\beta$ H <sub>2</sub> O domains. ....	38
Supplementary Figure 42   Titration profiles for the peaks disappearing in the NMR titration of the <sup>15</sup> N-labelled holo-Bamb_5917 PCP domain with the Sven_0517 $\beta$ H <sub>2</sub> O domain in solution. ....	39
Supplementary Figure 43   Local apparent association constants, K <sub>a</sub> , obtained from titrations of the holo-Bamb_5917 PCP domain with the Sven_0517 $\beta$ H <sub>2</sub> O domain. ....	40
Supplementary Movie 1   Fly through the solvent channel in the X-ray crystal structure of Bamb_5915.....	41
Supplementary Movie 2   1 us aMD simulation of Bamb_5915.....	42
Supplementary Movie 3   Fly through the solvent channel after 0.528 us of aMD simulations of Bamb_5915 ...	43
Supplementary Movie 4   The 1st mode from Principal Component Analysis of the 1 us aMD simulations of Bamb_5915.....	44
Supplementary Movie 5   The 1st mode from Principal Component Analysis of the 0.5 us aMD simulations of the holo-Bamb_5917 PCP domain.....	45
<b>Materials and Methods</b> .....	46
Expression construct cloning and site-directed mutagenesis.....	46
Supplementary Table 1 Protein constructs. ....	46
Expression and purification of proteins.....	47
Isotopic labelling and NMR sample preparation.....	47
In vitro conversion of the apo-Bamb_5917 PCP domain to its holo form.....	47
Bamb_5915 purification and sample preparation for crystallisation.....	48
Overproduction and purification of SeMet-labelled Bamb_5915.....	48
Crystallization of Bamb_5915.....	48
X-ray data collection, processing and structure determination.....	48
Supplementary Table 2 SeMet-labelled and unlabelled Bamb_5915 crystallographic statistics.....	49
NMR spectroscopy experiments and structure calculation.....	49
Supplementary Table 3 NMR and refinement statistics for the apo-Bamb_5917 PCP domain.....	50
Solution NMR dynamics.....	51
Supplementary Table 4 Parameters used in NMR <sup>15</sup> N dynamics measurements.....	51
Solution NMR titration experiments.....	52
Supplementary Table 5 Samples used in NMR titration experiments.....	52
NMR titration data fitting.....	52
<sup>15</sup> N CEST NMR Spectroscopy on Bamb_5917 PCP domain:10% Bamb_5915 complex.....	53
<sup>15</sup> N CEST data analysis.....	54
Sample preparation for solid state NMR analysis of the Bamb_5917 PCP domain:Bamb_5915 complex.....	54
Solid state NMR spectroscopy.....	54
Supplementary Table 6 Solid state NMR experiments.....	55
BioLayer Interferometry (BLI) measurements.....	55
Carbene footprinting.....	55
Supplementary Table 7 Statistical analysis of tryptic digest for the Bamb_5917 PCP domain alone and in the presence of Bamb_5915.....	56
Supplementary Table 8 Statistical analysis of AspN digest for the Bamb_5917 PCP domain alone and in the presence of Bamb_5915.....	57
Supplementary Table 9 Statistical analysis of tryptic digest for Bamb_5915 alone and in the presence of Bamb_5917 PCP domain.....	57
Supplementary Table 10 Statistical analysis of AspN digest for Bamb_5915 alone and in the presence of Bamb_5917 PCP domain.....	58
Protein docking simulations.....	58
In vitro acetylation assays with Bamb_5915, AHCCA and acetylated holo-PCP domains.....	58
Supplementary Table 11 Elution conditions used for LC-MS analyses of small molecules.....	59
UHPLC-ESI-Q-TOF-MS analysis of proteins and small molecules.....	59
Supplementary Table 12 Elution conditions used for LC-MS analyses of purified recombinant proteins.....	59
Dynamics predictions from sequence.....	59

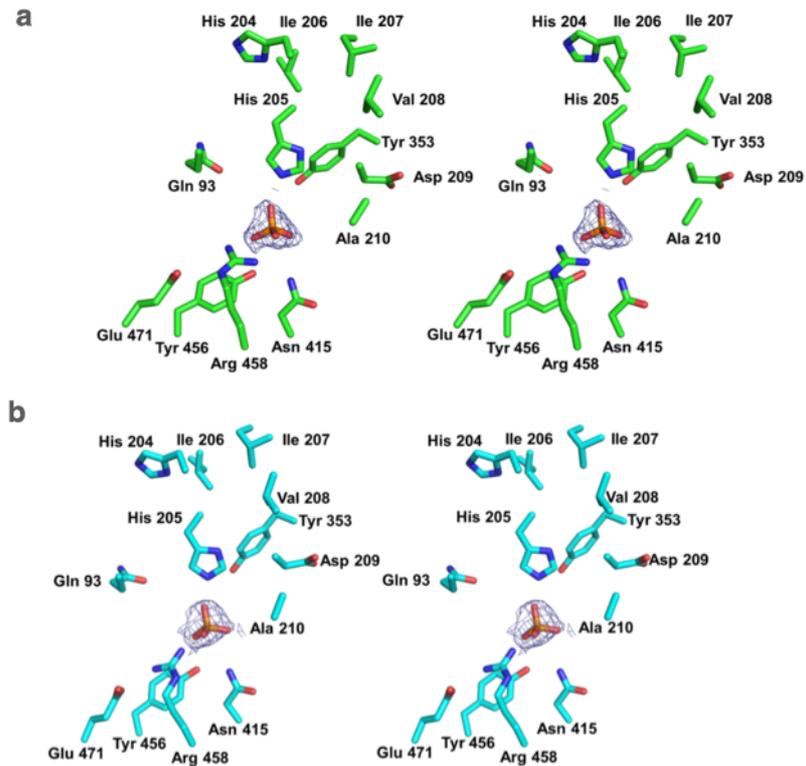
Molecular dynamic simulations.....	60
Construction of HMMs.....	60
Supplementary Table 13 Amino acid sequences that were used to build an initial $\beta$ HD domain HMM.....	60
Supplementary Table 14 Systems with identified SLiM- $\beta$ HD domain interfaces in MIBiG database. ....	61
Supplementary Table 15 Examples of systems with known metabolites and identified SLiM- $\beta$ HD domain boundaries that are not included in the MIBiG database at the time of writing of the manuscript.....	62
Genebank database search .....	62
References.....	63



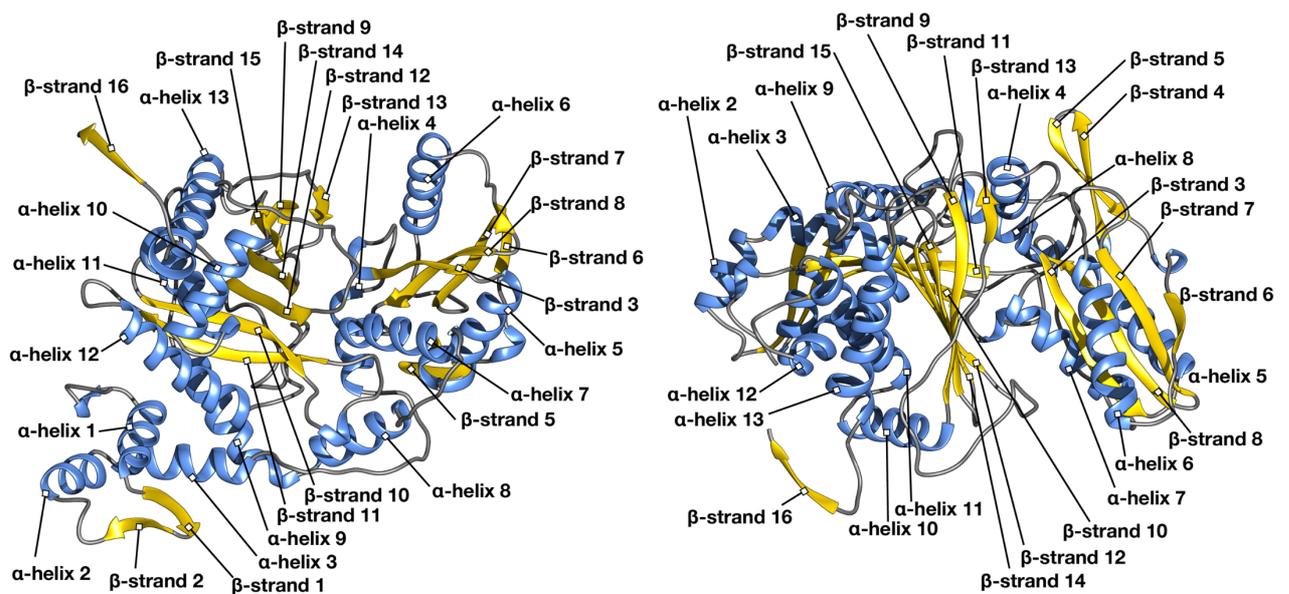
**Supplementary Figure 1 | Stereo views of Bamb\_5915 and the latch region between  $\beta$ -strands 12 & 13. a,** The overall architecture of the Bamb\_5915 with individual monomers coloured red and deep violet. **b,** A snapshot of the latch region, highlighted in red, of Bamb\_5915-monomer A.  $2F_o-F_c$  maps are contoured to  $1.0 \sigma$  at  $2.00 \text{ \AA}$ .



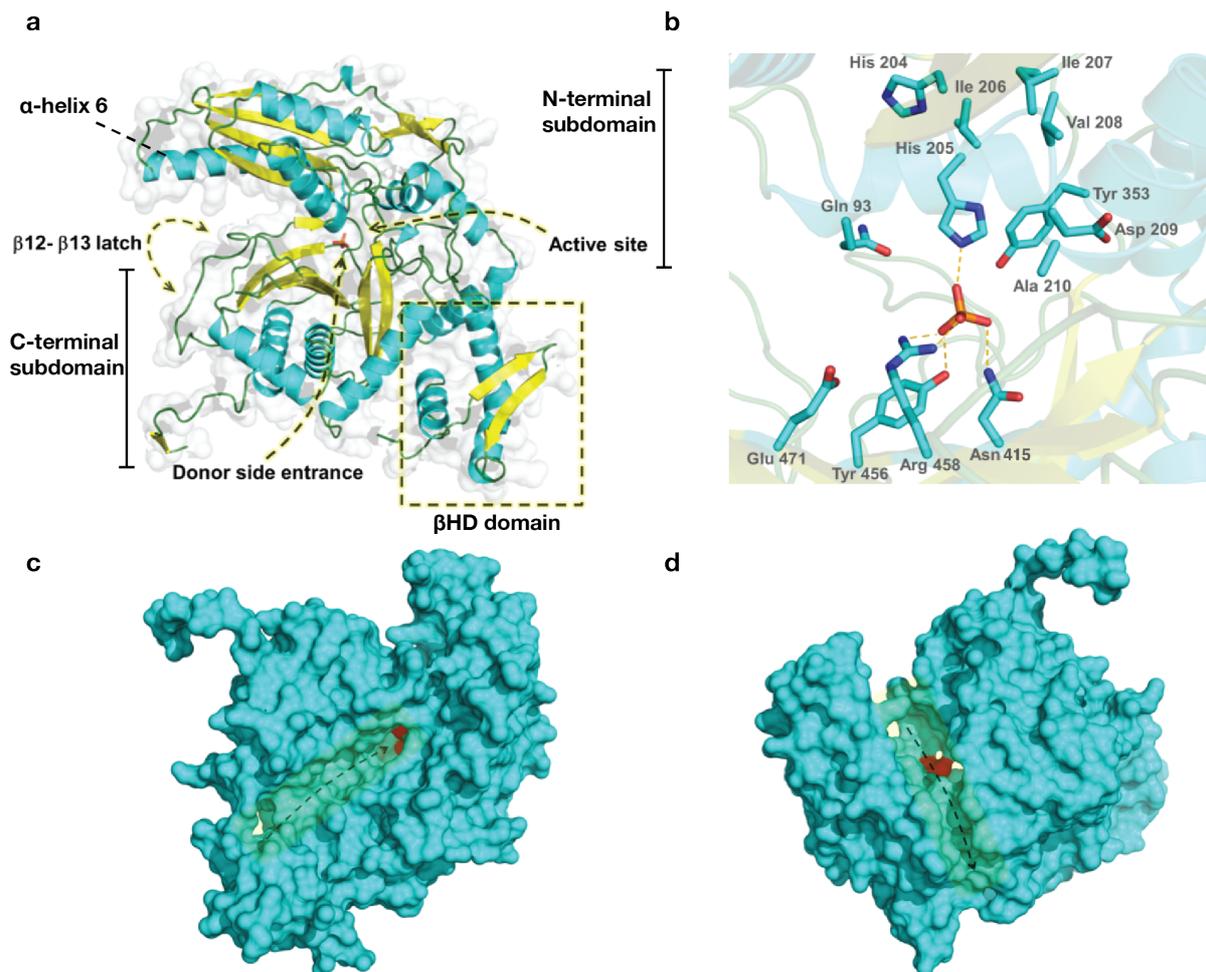
**Supplementary Figure 2 | Stereo views of Bamb\_5915 highlighting its N-terminal  $\beta$ -hairpin docking ( $\beta$ HD) domain shown in **a** cartoon and **b** ribbon representations. The individual monomers are coloured red and deep violet with  $2F_o-F_c$  maps contoured to  $1.0 \sigma$  at  $2.00 \text{ \AA}$ .**



**Supplementary Figure 3 | Stereo views of the active site of the Bamb\_5915 C domain with phosphate bound.** **a**, The active site of Bamb\_5915 monomer A and **b**, monomer B, coloured blue and green, respectively, with  $2F_o - F_c$  SA omit maps contoured to  $1.0 \sigma$  at  $2.00 \text{ \AA}$ .

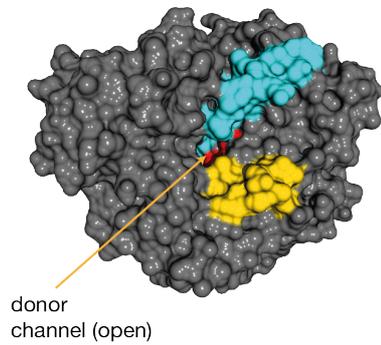


**Supplementary Figure 4 | Secondary structure of Bamb\_5915.** Secondary structure elements are labelled according to the convention used throughout the manuscript.

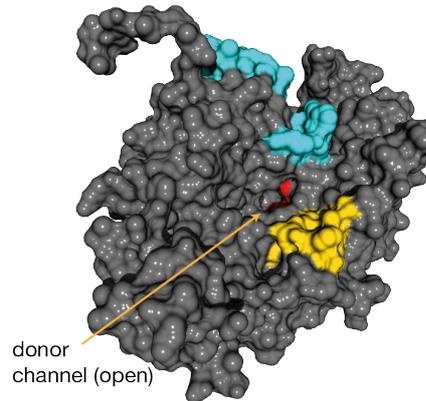


**Supplementary Figure 5 | The crystal structure of Bamb\_5915.** **a**, Tertiary structure of Bamb\_5915 with helices coloured in cyan and  $\beta$ -sheets coloured in yellow. **b**, The active site of Bamb\_5915. The canonical C domain active site residues are sequentially and structurally conserved. **c**, Surface representation of Bamb\_5915 coloured cyan with the active site residue His 205 shown in red, highlighting the front entrance of the substrate binding tunnel. **d**, A snap shot of the back entrance of the substrate binding tunnel with Bamb\_5915 coloured in cyan and active site His 205 highlighted in red.

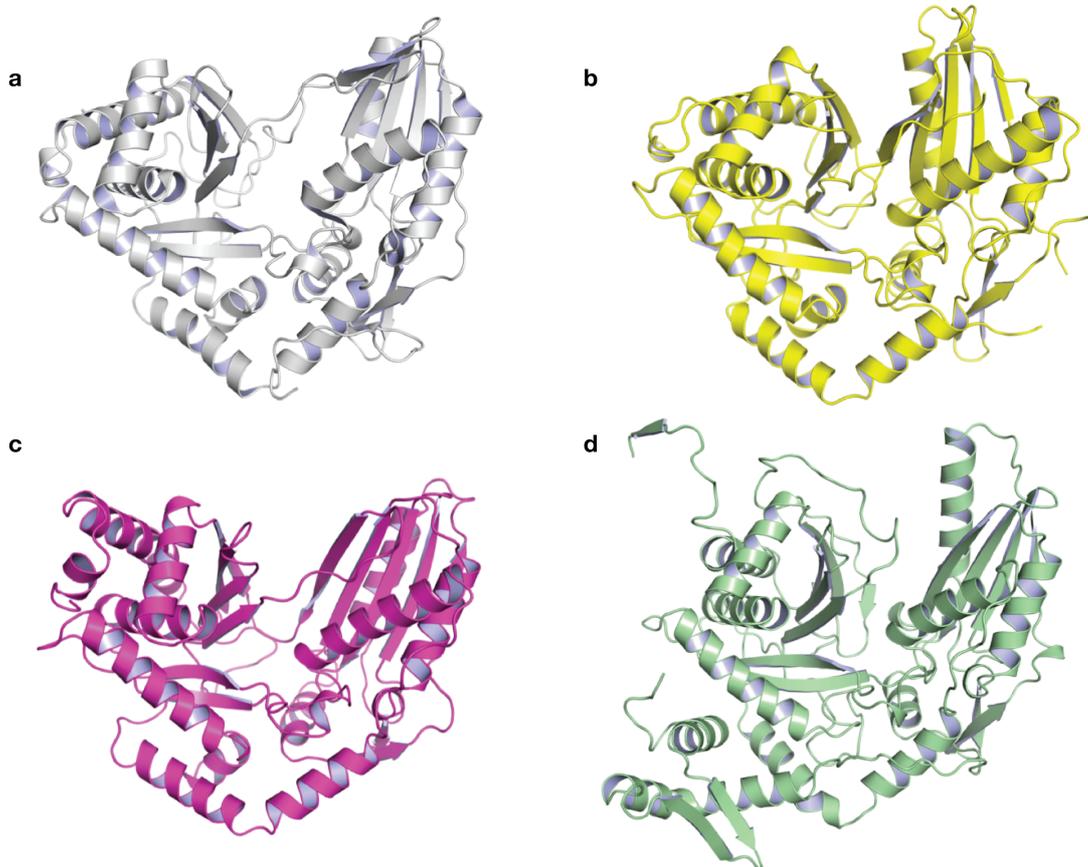
**a** CDA-C1 (PDB 4jn3)



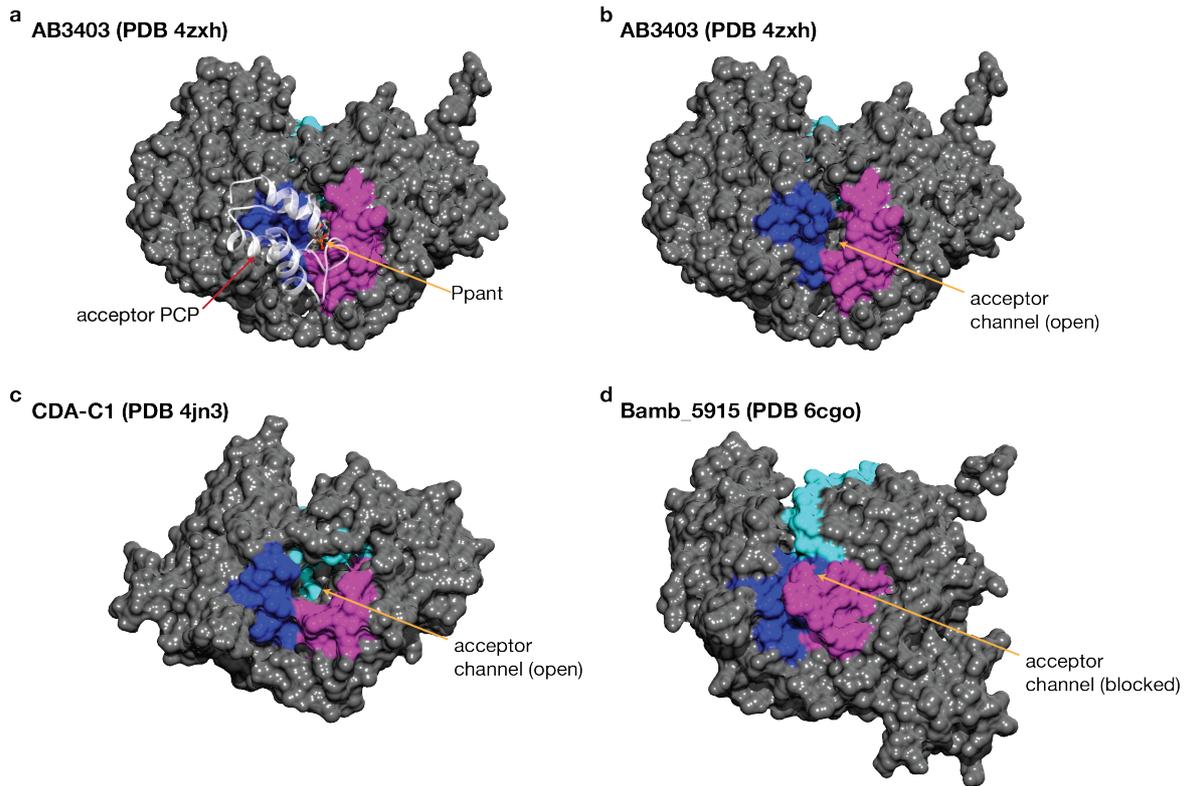
**b** Bamb\_5915 (PDB 6cgo)



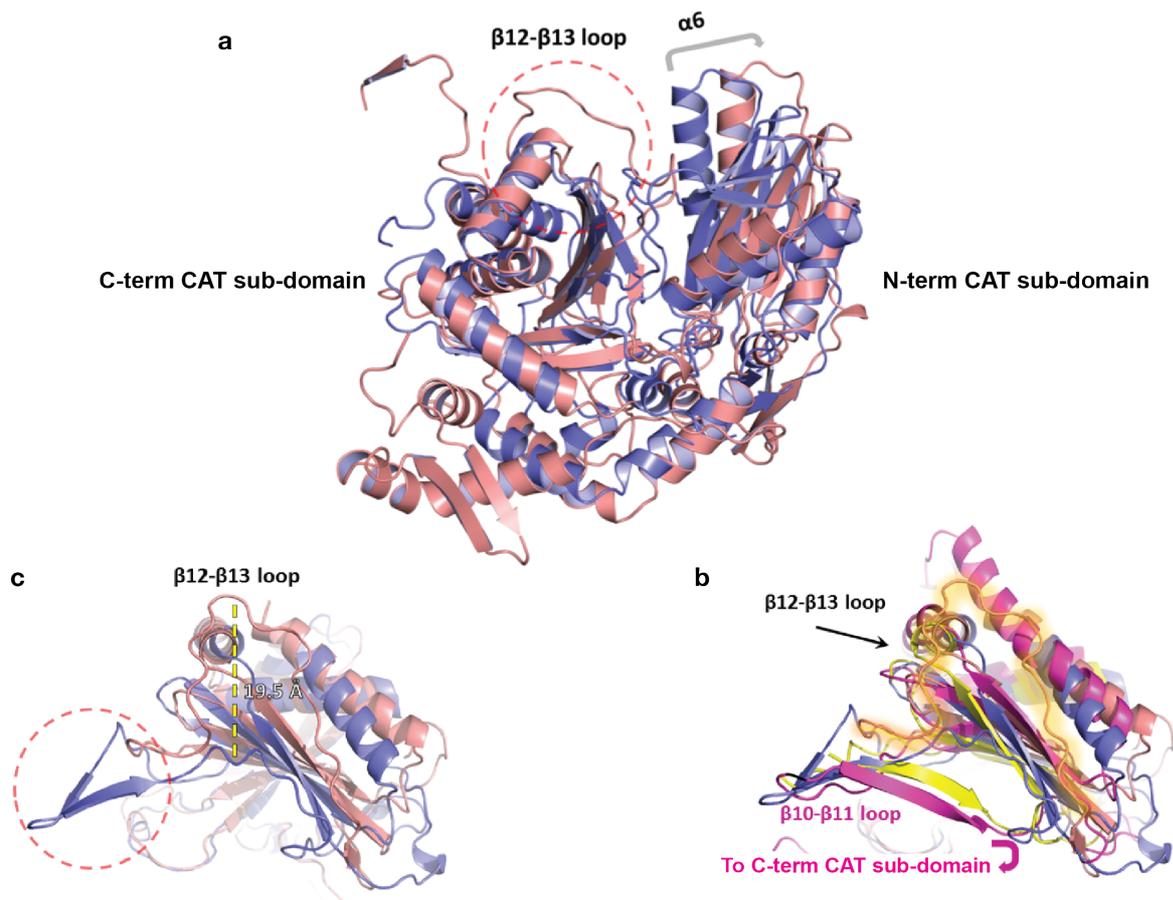
**Supplementary Figure 6 | Location of the donor channel in condensation domains. a**, A view of the donor channel in the CDA-C1 domain<sup>1</sup>(PDB code 4JN3). **b**, A view of the donor channel in the X-ray structure of Bamb\_5915 (PDB code 6CGO). The “latch” is coloured in cyan, “floor loop” in yellow and the residues in the active site motif in red.



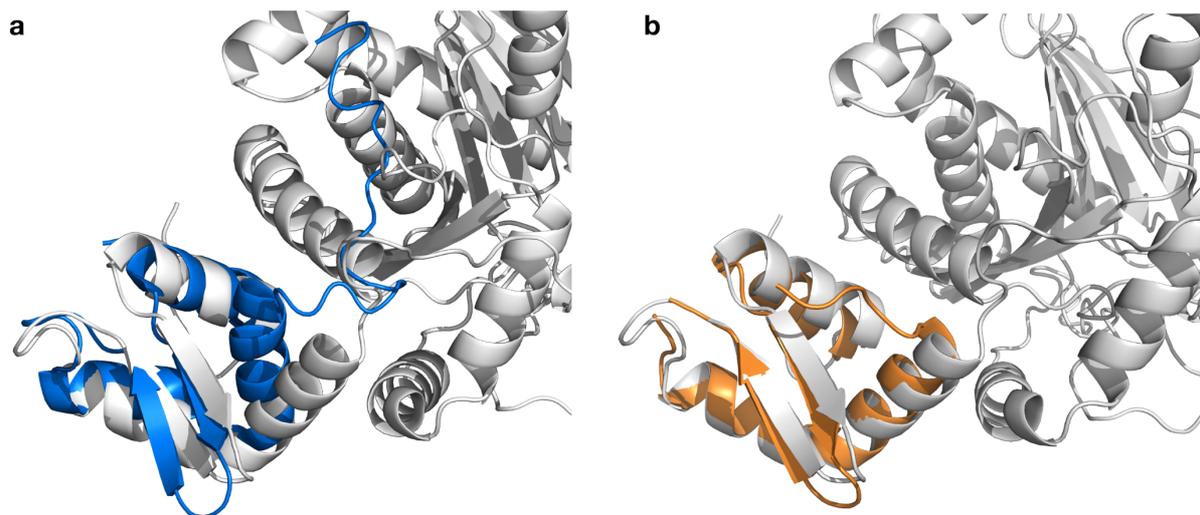
**Supplementary Figure 7 | A structural comparison of Bamb\_5915 to other C domain structures. a**, TycC5-PCP di-domain (PDB code: 2JGP) with PCP domain removed for clarity. **b**, CDA C1 domain<sup>1</sup> (PDB code: 4JN5). **c**, VibH<sup>2</sup> (PDB code: 1L5A). **d**, Bamb\_5915 (PDB code: 6CGO).



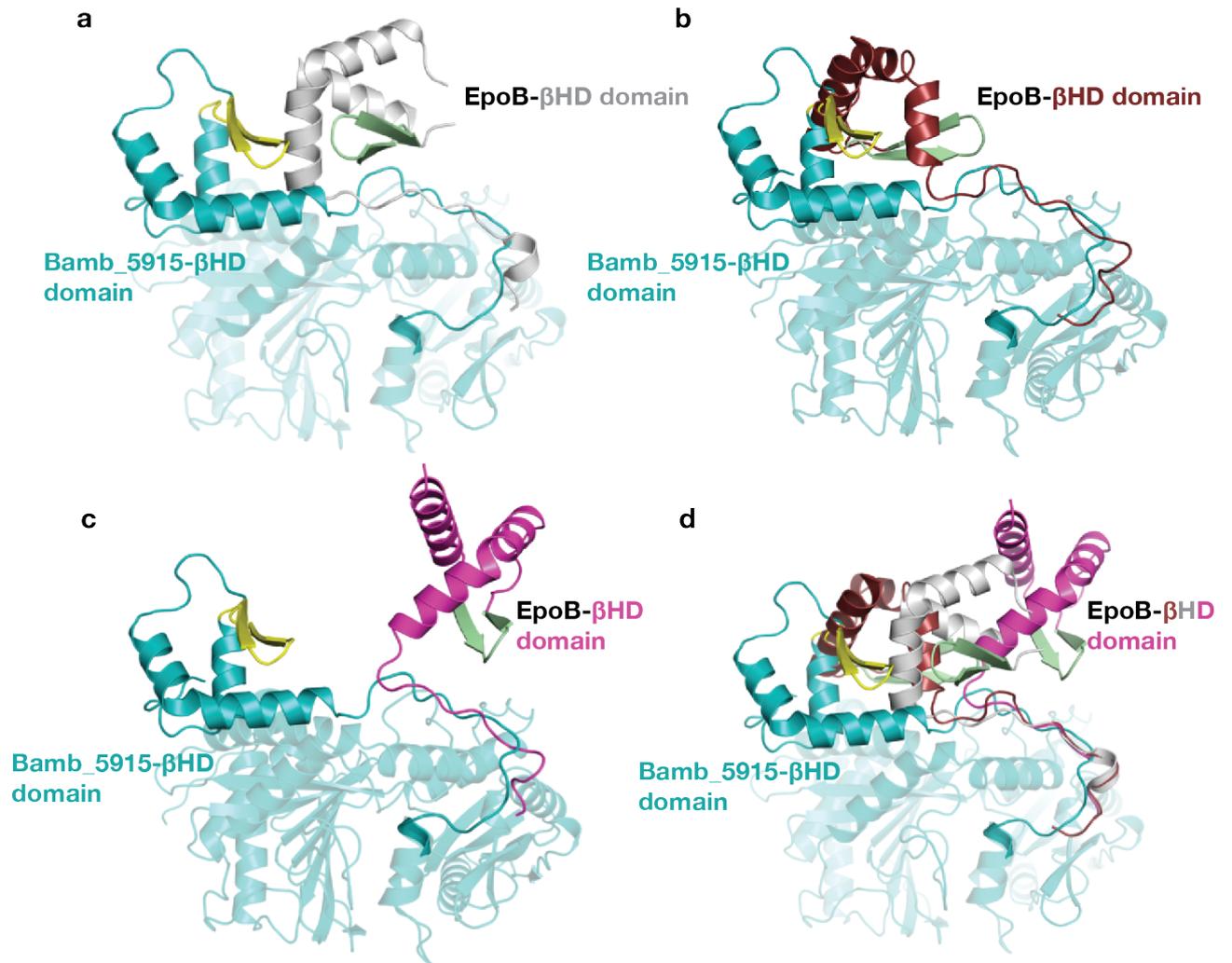
**Supplementary Figure 8 | Location and state of acceptor channel in condensation domains.** **a**, A view of the acceptor channel with docked acceptor side PCP and Ppant inserted into the channel in the X-ray structure of the AB3403 C domain<sup>3</sup> (PDB code 4ZXH). **b**, A view of the acceptor channel in the AB3403 C domain with the PCP domain removed. **c**, A view of the acceptor channel in the CDA-C1 domain<sup>1</sup> (PDB code 4JN3). **d**, The acceptor channel is blocked in the X-ray structure of Bamb\_5915 (PDB code 6cgo). The “latch” is coloured cyan,  $\alpha$ -helix 4 is blue and the loop between  $\beta$ -strand 12 and  $\alpha$ -helix 9 is magenta in Bamb\_5915. Equivalent fragments in the other C domains are coloured analogously.



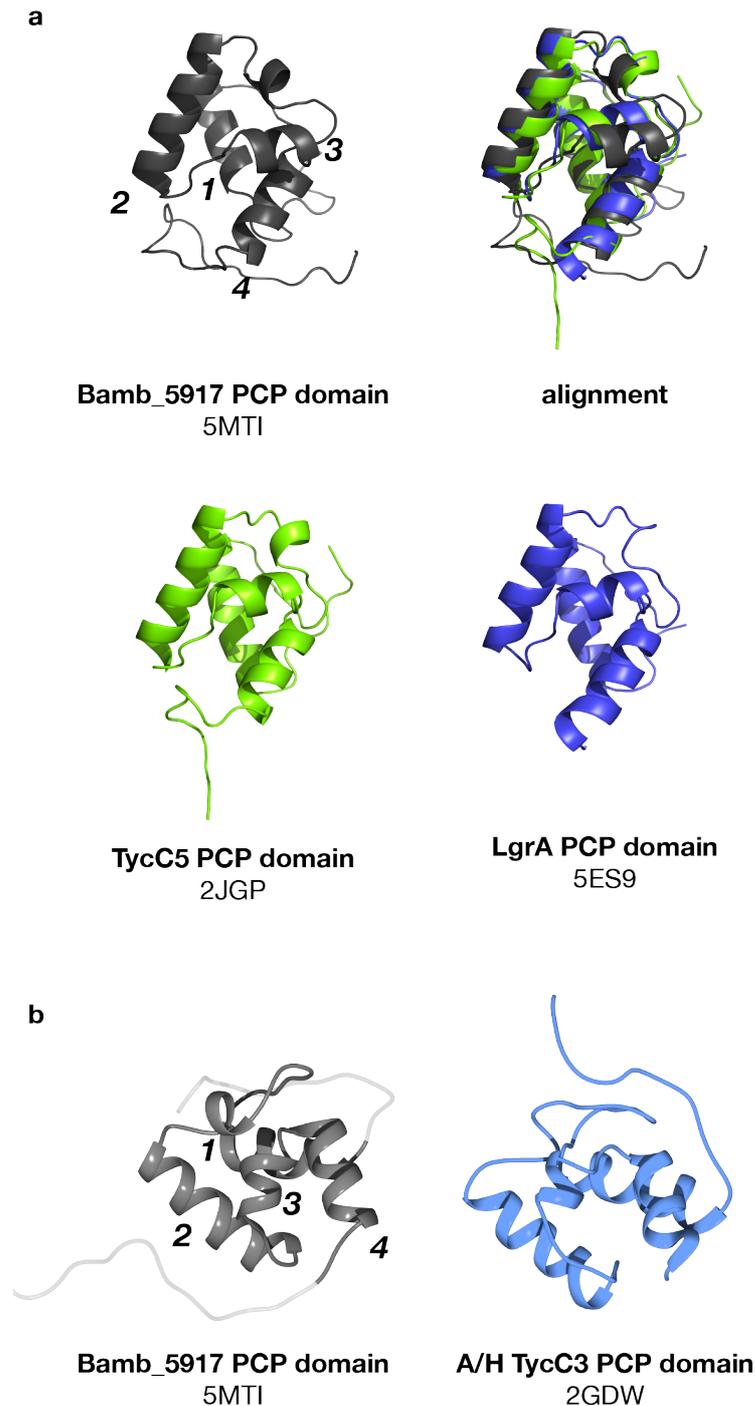
**Supplementary Figure 9 | Superimposition of Bamb\_5915 (pink) with the most similar available structure, SgcC5 (blue), as determined by Dali<sup>4</sup> server (C <sup>$\alpha$</sup>  RMSD of 3.7 Å, Z score 28.5). a, Structural overlay of Bamb\_5915 (salmon) with SgcC5<sup>5</sup> (blue) emphasizing structural differences in the loop between  $\beta$ -strands 12 and 13 and  $\alpha$ -helix 6. b, The loop between  $\beta$ -strands 12-13 of Bamb\_5915 is shifted 19.5 Å upwards with respect to SgcC5 (PDB code: 4ZNM). c, In VibH<sup>2</sup> (PDB code:1L5A), the CDA-C1 domain<sup>1</sup> (PDB code: 4JN5), and SgcC5, the long loop between  $\beta$ -strands 12-13 found in Bamb\_5915 is absent and instead is replaced by a 10-amino acid linker that links  $\beta$ -strand 10 of the C-term CAT sub-domain to  $\beta$ -strand 11 of the N-term CAT sub-domain. VibH, the CDA-C1 domain, and SgcC5 are coloured magenta, yellow, and blue, respectively.**



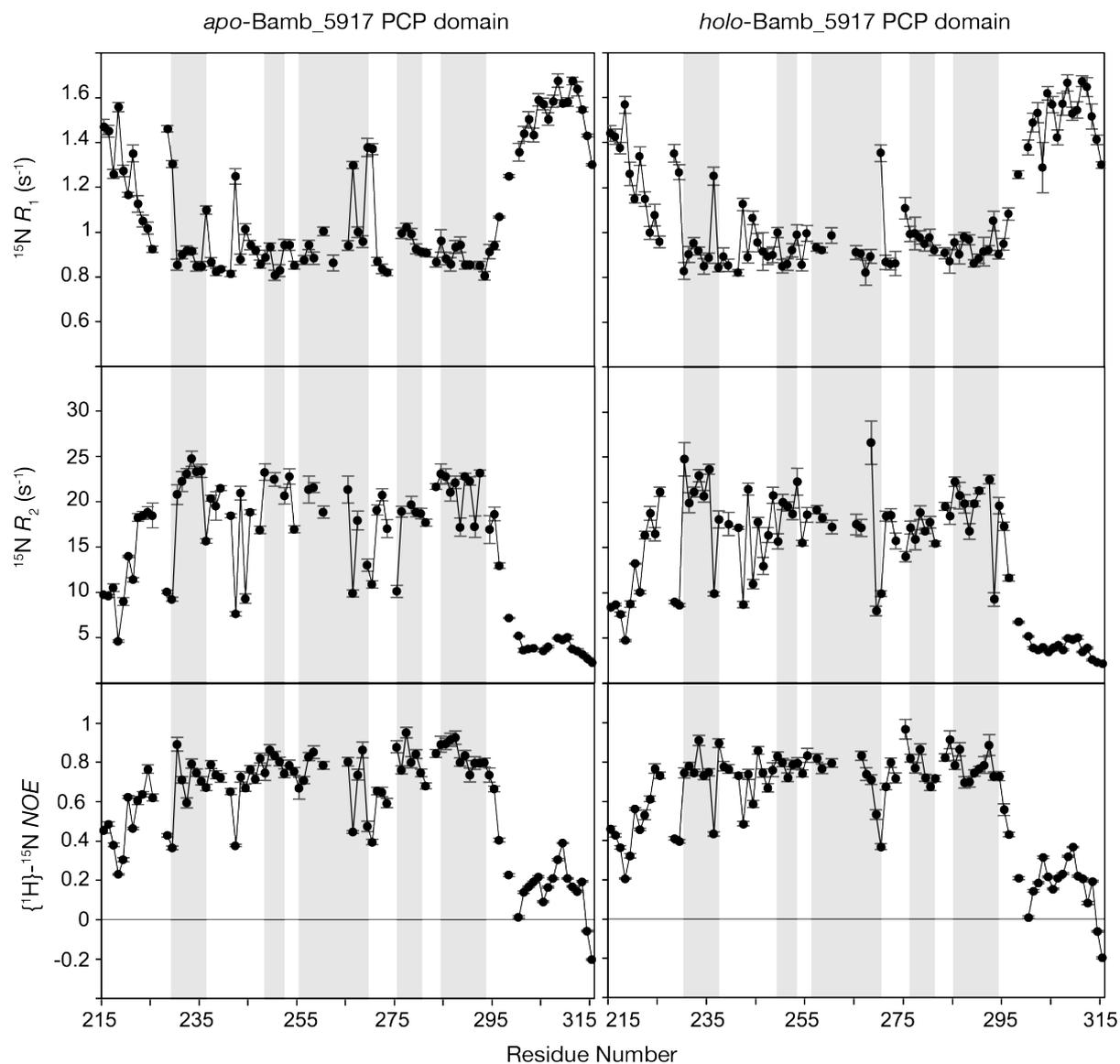
**Supplementary Figure 10 | Comparison of known  $\beta$ HD domain structures.** **a**, Structural alignment of the  $\beta$ HD domain from the crystal structure of Bamb\_5915 (grey) and solution NMR structure of TubCdd from the tubulysin PKS/NRPS<sup>6</sup> (blue). **b**, Structural alignment of the  $\beta$ HD domain from the crystal structure of Bamb\_5915 (grey) and the crystal structure of EpoBcy<sup>7</sup> (orange).



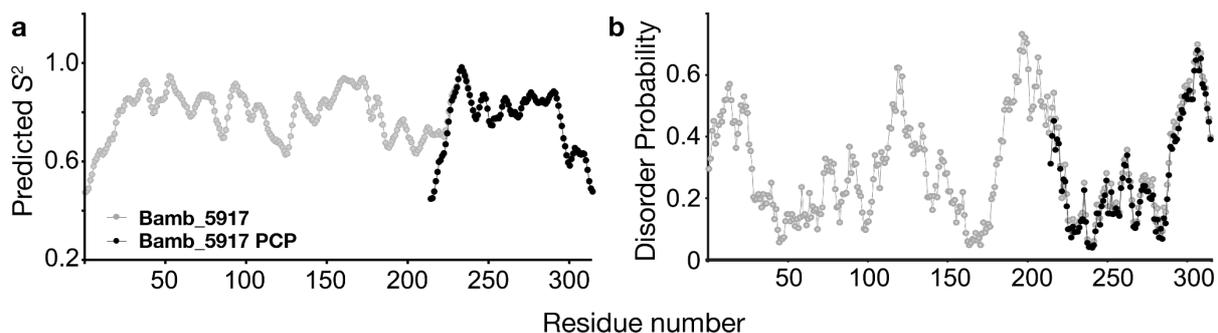
**Supplementary Figure 11 | Superimposition of Bamb\_5915 (cyan) with the multiple crystal forms of EpoBdd (a-c). d, a, Overlay of residues 2-71 from the monoclinic form of EpoBdd<sup>7</sup> (PDB code 5T7Z) with Bamb\_5915. b-c, superimposition of Bamb\_5915 with residues 1-76 of chain A (b) and residues 1-73 of chain B (c) from the rhombohedral form of EpoBdd (PDB code 5T81). (d) Overlay of Bamb\_5915 with all three crystal forms of EpoBdd.**



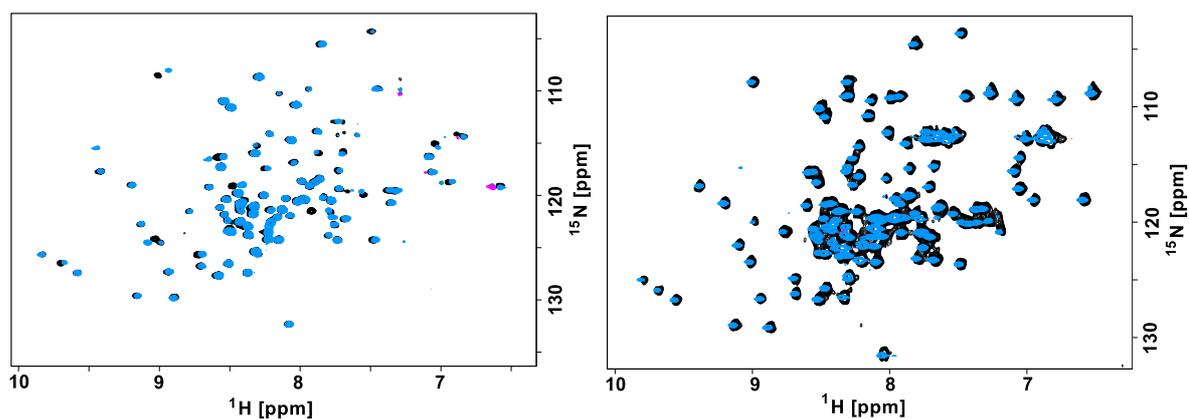
**Supplementary Figure 12 | A structural comparison of the BamB\_5917 PCP domain with other PCP domains.** **a**, Comparison to the closest matches from Dali<sup>4</sup> structural alignment. The BamB\_5917 PCP domain (PDB code: 5MTI) shares the highest degree of structural similarity with the PCP domains in the TycC5-6 PCP-C di-domain<sup>8</sup> (PDB code 2JGP; blue) and the LgrA initiation module<sup>9</sup> (PDB code 5ES9; red) with Z score 7.1.  $\alpha$ -Helix 1 of the BamB\_5917 PCP domain (numbers of the helices are indicated in bold italic font) is shorter by one turn.  $\alpha$ -Helix 3 is displaced further away from  $\alpha$ -helix 2 and closer to  $\alpha$ -helix 4 and the single helical turn in loop 1 (A238-D254) is positioned further away from  $\alpha$ -helix 1. **b**, Comparison of the BamB\_5917 PCP domain to the A/H state of the TycC3 PCP domain<sup>10</sup> (PDB code: 2GDW).



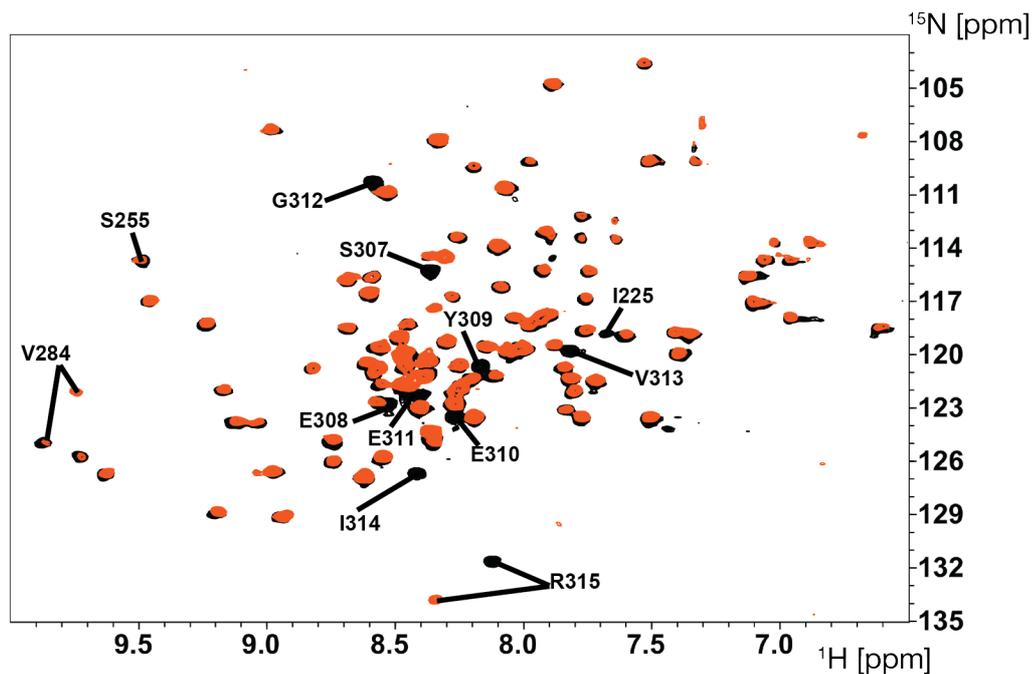
**Supplementary Figure 13 | Solution NMR dynamics data for the *apo* and *holo*-Bamb\_5917 PCP domain.**  $^{15}\text{N } R_1$ ,  $^{15}\text{N } R_2$  and  $\{^1\text{H}\}-^{15}\text{N}$  NOE of *apo* and *holo*-Bamb\_5917 PCP domain illustrate the high ps-ns flexibility of the C-terminus. Grey strips show the location of helical elements. Data recorded at 700 MHz  $^1\text{H}$  Larmor frequency and 280 K sample temperature. Error bars represent  $\pm$  standard deviation of the measurements.



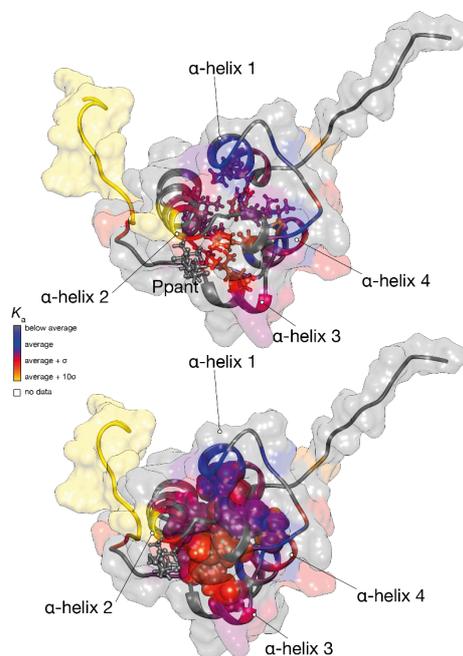
**Supplementary Figure 14 | Dynamics and order for Bamb\_5917.** **a**, Dynamine<sup>11</sup>  $S^2$  order parameter predictions. **b**, IUpred<sup>12</sup> disorder probability predictions. The predictions for full-length Bamb\_5917 (grey) and the Bamb\_5917 PCP domain (black) suggest high flexibility and intrinsic disorder at the C-terminus and linker region between the KS domain remnant (indicated with “?” in Fig. 1a) and the PCP domain.



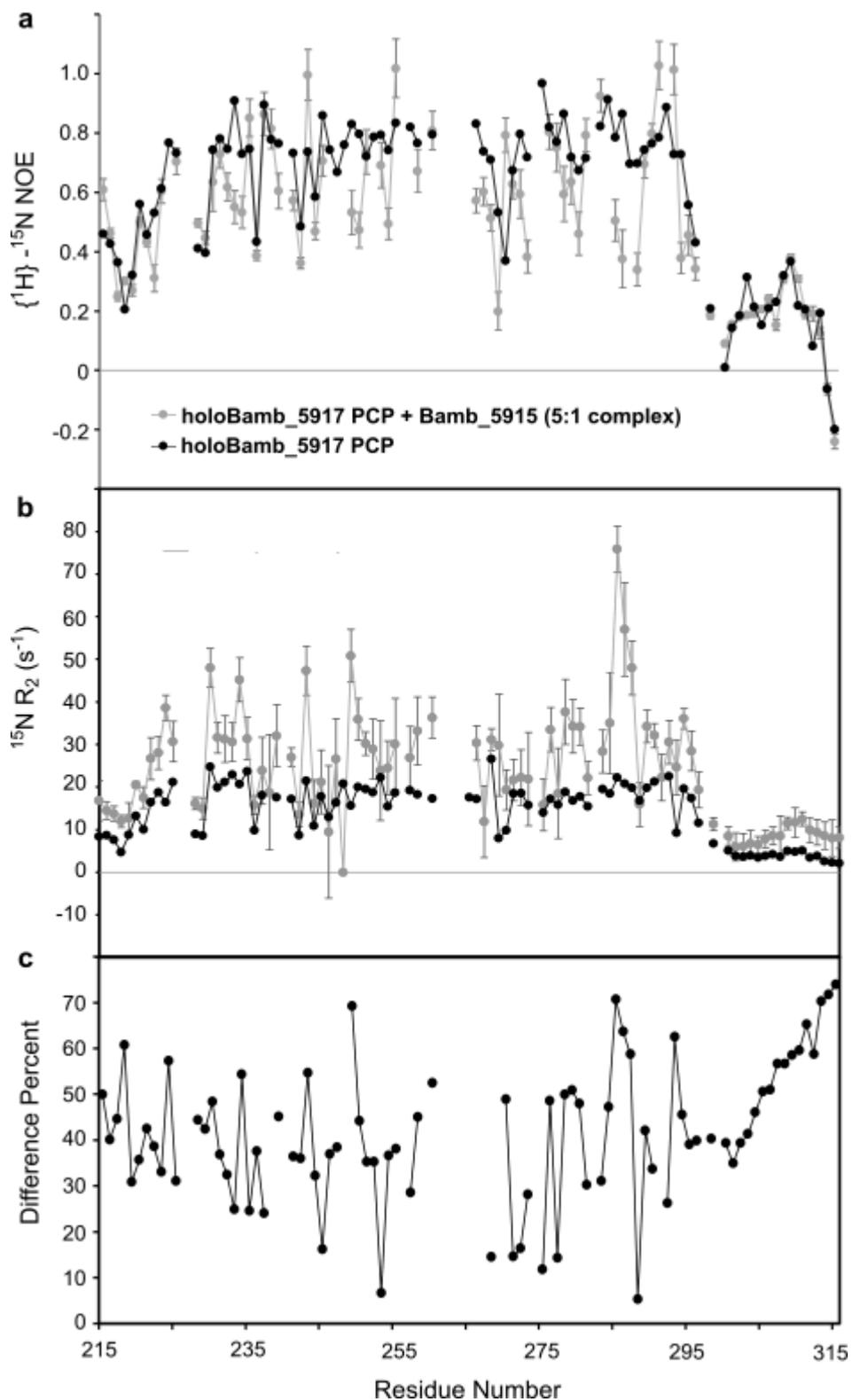
**Supplementary Figure 15 | 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC correlation spectra of the Bamb\_5917 PCP domain in solution.** **a**, Conversion of the *apo*-Bamb\_5917 PCP domain (black) to its *holo* form (blue) by addition of Sfp, CoA and  $\text{Mg}_2\text{Cl}$ . **b**, Spectra of a sample prepared from the freshly purified His-tagged *apo*-Bamb\_5917 PCP domain (black) and the freeze dried *apo*-Bamb\_5917 PCP domain (blue).



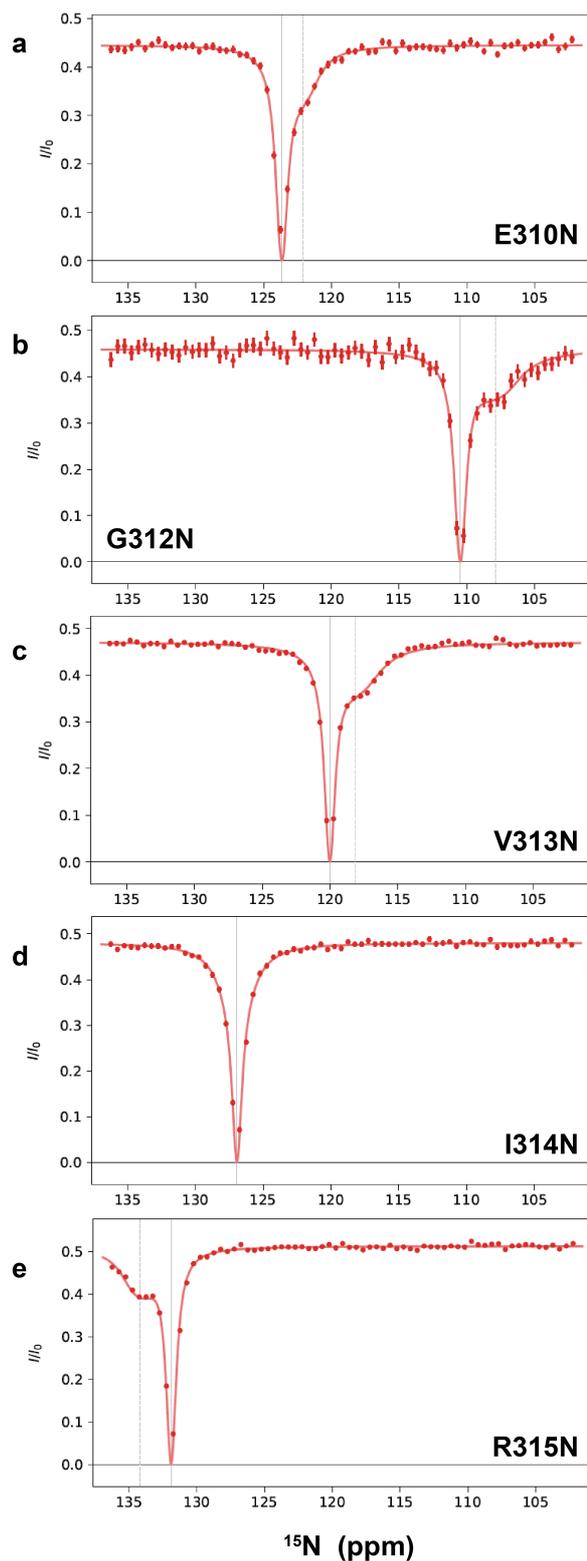
**Supplementary Figure 16 | 2D  $^1\text{H}$ - $^{15}\text{N}$  HSQC correlation spectra of the *holo*-Bamb\_5917 PCP domain in isolation (black) and in the presence of the Bamb\_5915  $\beta$ HD domain (orange-red) in solution.** The gradual disappearance of the peaks for the free form of Bamb\_5917 PCP domain and appearance of new peaks for the bound form of Bamb\_5917 PCP domain indicate slow to intermediate exchange regime. The assignments for the bound form were made with the help of the chemical shift information from a  $^{15}\text{N}$  CEST experiment (Fig. 19).



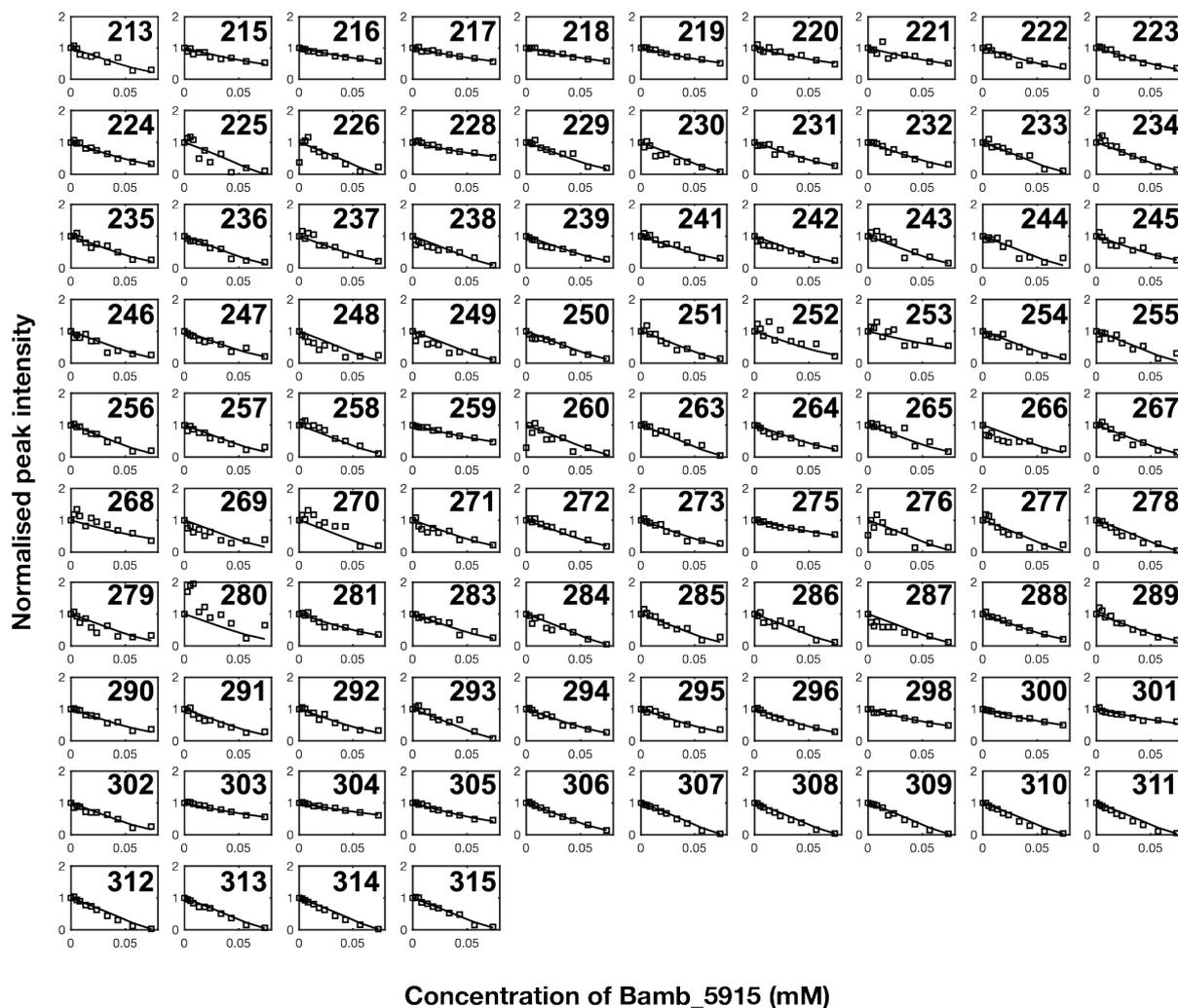
**Supplementary Figure 17 | Residues with greater than average local apparent  $K_{\text{a}}$ s in NMR titrations of the *holo*-Bamb\_5917 PCP domain with the excised Bamb\_5915  $\beta$ HD domain in solution with buried side chains.** Because the residues of these side chains make direct contact with each other, the above average local apparent  $K_{\text{a}}$ s are more likely to be due to a conformational change upon binding rather than a direct effect of binding.



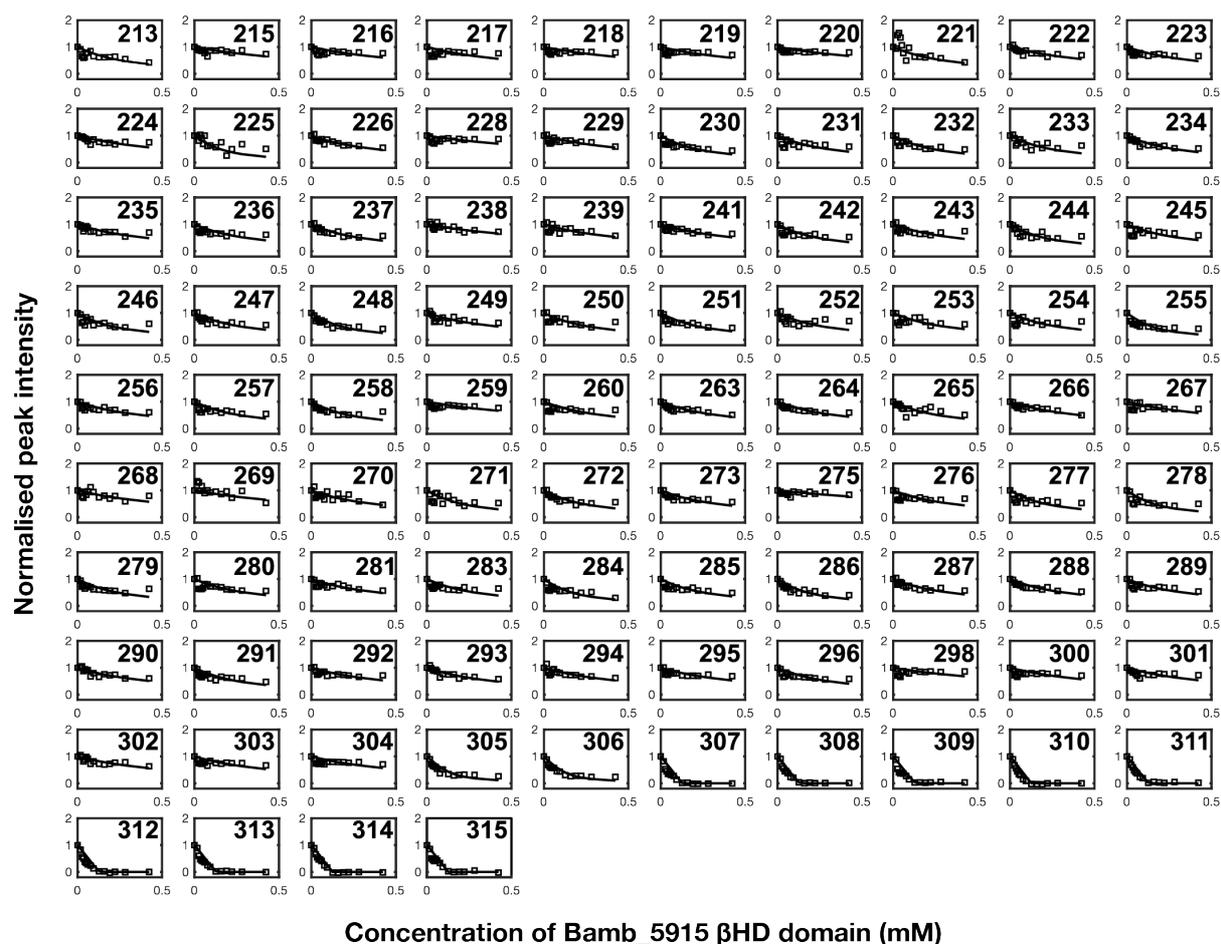
**Supplementary Figure 18 | Solution NMR  $^{15}\text{N}$  relaxation data for the *holo*-Bamb\_5917 PCP domain in the presence of Bamb\_5915. **a**, Comparison of  $\{^1\text{H}\}-^{15}\text{N}$  NOEs for the *holo*-Bamb\_5917 PCP domain in isolation (black) and in the presence of Bamb\_5915 (~5:1 ratio; grey). **b**, Comparison of  $^{15}\text{N}$   $R_2$  values for the *holo*-Bamb\_5917 PCP domain in isolation (black) and in the presence of Bamb\_5915 (~5:1 ratio). **c**, Percent difference between  $^{15}\text{N}$   $R_2$  values for the *holo*-Bamb\_5917 PCP domain in the presence of Bamb\_5915 (~5:1 ratio) and in isolation normalized to the value in the presence of Bamb\_5915. Data were recorded at 700 MHz  $^1\text{H}$  Larmor frequency and 280 K sample temperature. Error bars represent  $\pm$  standard deviation of the measurements.**



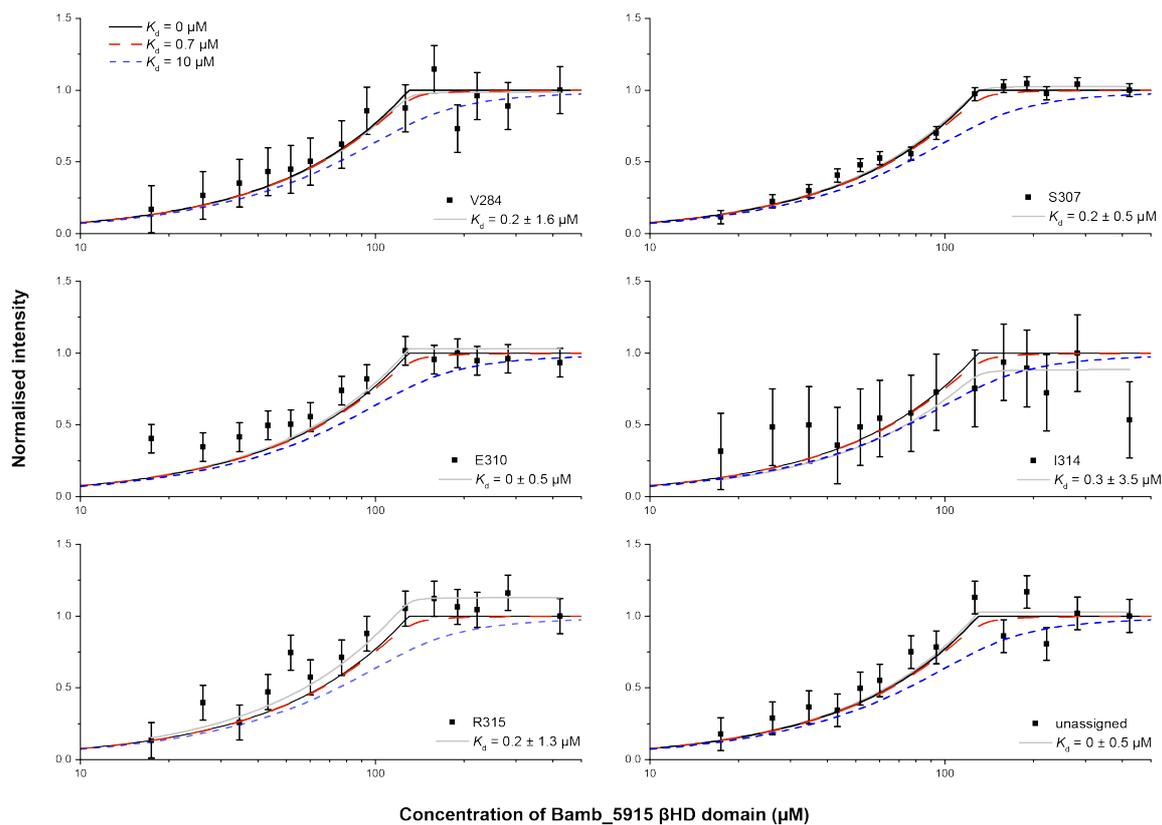
**Supplementary Figure 19 |  $^{15}\text{N}$  CEST profiles for the *holo*-Bamb\_5917 PCP domain in the presence of Bamb\_5915.** Experimental Chemical Exchange Saturation Transfer (CEST)<sup>13</sup> profiles of the normalized peak intensities obtained for selected residues (E310, G312, V313, and R315) in the Bamb\_5917 PCP domain using a  $B_1$  saturating field of 25 Hz. Experimental data are shown as small full red circles. Red fitted curves are shown as solid lines in all panels. The red lines correspond to a fit to a two-site exchange model by analysing simultaneously data for selected residues using ChemEx<sup>14</sup>. All CEST experiments were recorded in the presence of 10% of unlabelled Bamb\_5915. Data were recorded at 700 MHz  $^1\text{H}$  Larmor frequency and 288 K sample temperature.



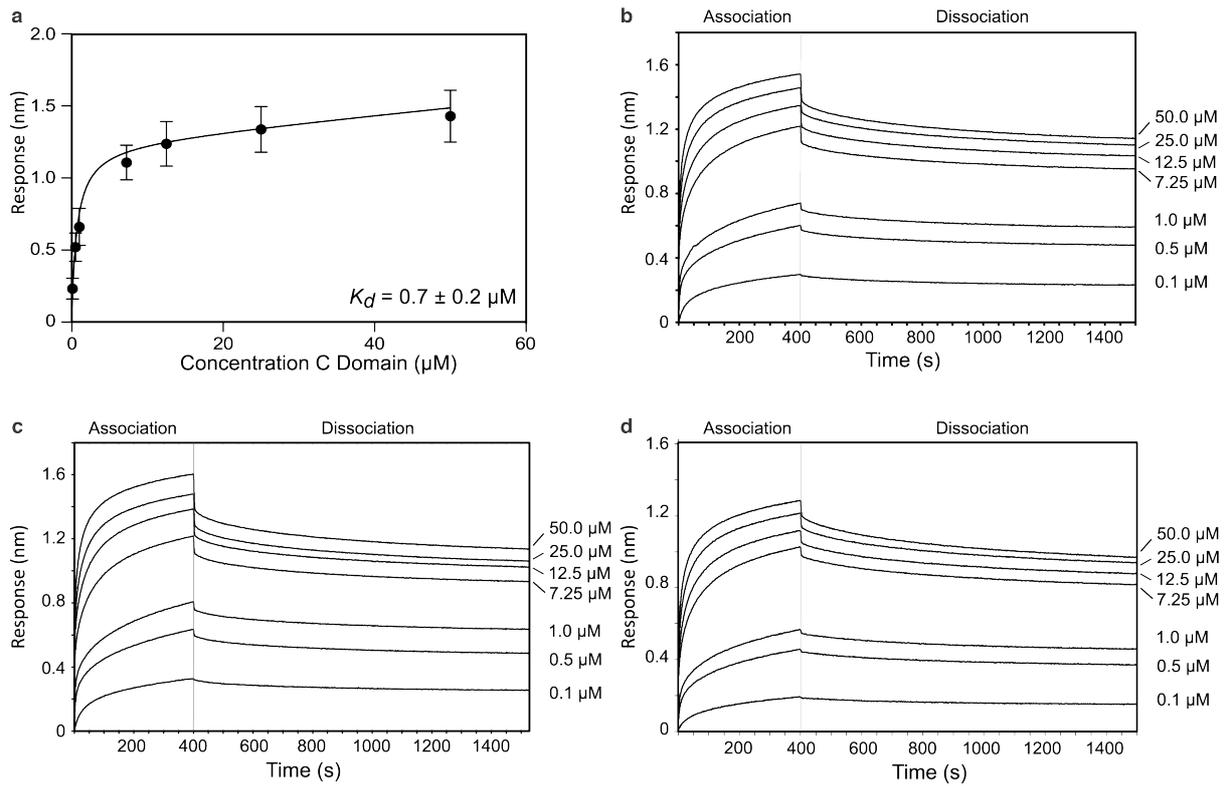
**Supplementary Figure 20 |** Titrations profiles for the peaks disappearing in the NMR titration of the  $^{15}\text{N}$ -labelled *holo*-Bamb\_5917 PCP domain with Bamb\_5915 in solution as a result of a complex formation. The squares represent peak volumes normalised to the peak volume of the *holo*-Bamb\_5917 PCP domain before addition of Bamb\_5915 and the lines represent the best fit curves to equation 3 (see Supplementary Section *NMR titration data fitting*). The residue numbers for the Bamb\_5917 PCP domain are given in the top right corner of each panel.



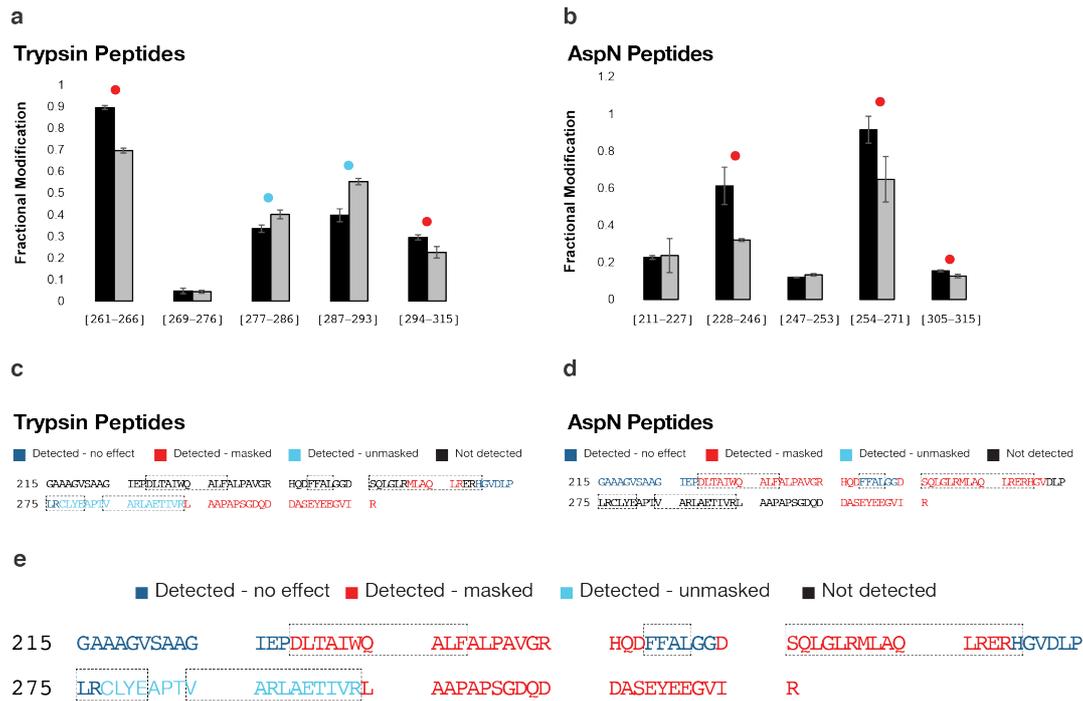
**Supplementary Figure 21 |** Titrations profiles for the peaks disappearing in the NMR titration of the  $^{15}\text{N}$ -labelled *holo*-Bamb\_5917 PCP domain with Bamb\_5915  $\beta$ HD domain in solution as a result of a complex formation. The squares represent peak volumes normalised to the peak volume of the *holo*-Bamb\_5917 PCP domain before addition of the Bamb\_5915  $\beta$ HD domain and the lines represent the best fit curves to equation 3 (see Supplementary Section *NMR titration data fitting*). The residue numbers for the Bamb\_5917 PCP domain are given in the top right corner of each panel.



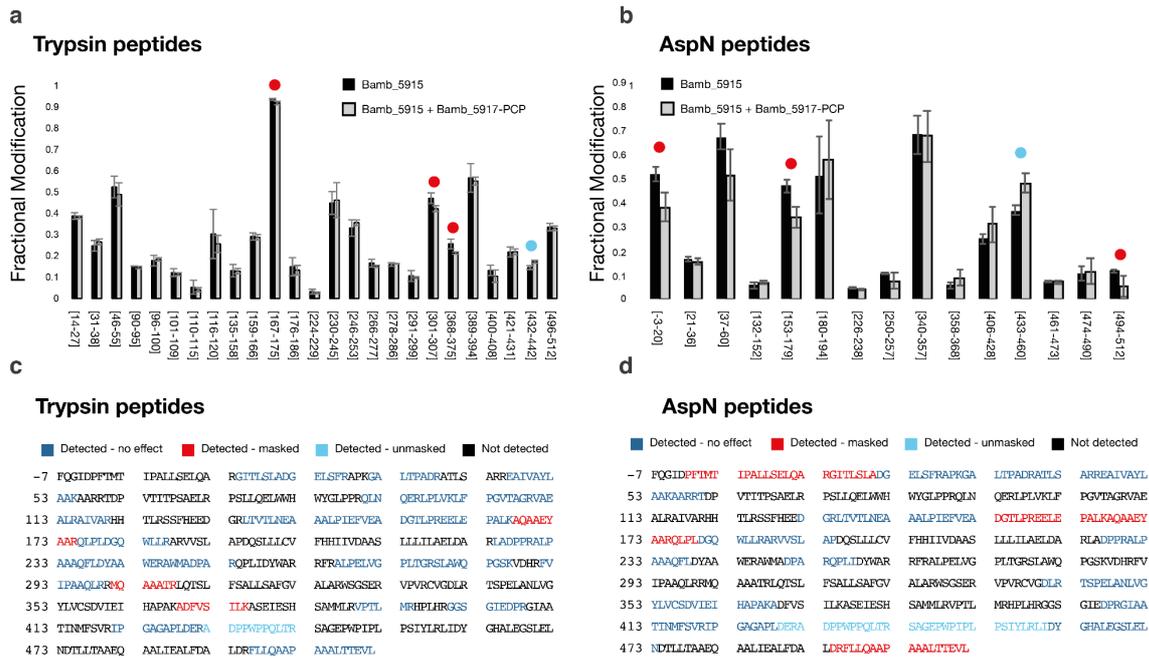
**Supplementary Figure 22 | Titration profiles for the peaks appearing as a result of complex formation in the NMR spectra of the *holo*-Bamb\_5917 PCP domain mixed with the Bamb\_5915  $\beta$ HD domain in solution.** The squares represent normalized experimental peak volumes. The red dashed lines represent trends simulated for  $K_d = 0.7$   $\mu$ M, which is the  $K_d$  measured by BLI. The solid black lines represent the simulated trends in the limit of  $K_d = 0$ . The dashed blue lines represent the trends for  $K_d = 10$   $\mu$ M. The solid grey lines represent the best fit curves to equation 4. The data are overall consistent with  $< 10$   $\mu$ M  $K_d$ . Error bars correspond to  $\pm$  mean squared deviation of the noise in the spectra.



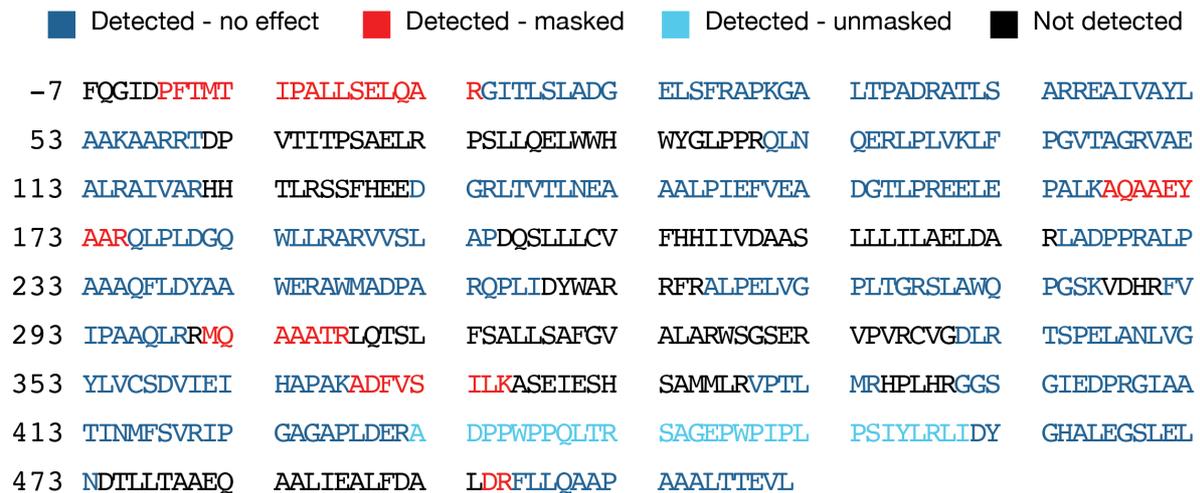
**Supplementary Figure 23 | Biolayer interferometry (BLI) data for binding of the Bamb\_5917 PCP domain to Bamb\_5915.** **a**, Steady state data were fitted to a one-site binding model to estimate the  $K_d$ . Association and dissociation measurements of the interaction between the immobilized biotinylated Bamb\_5917 PCP domain and free Bamb\_5915 in seven different concentrations was performed in triplicate (**b**, **c**, **d**). Error bars correspond to  $\pm$  standard deviation of the measurements.



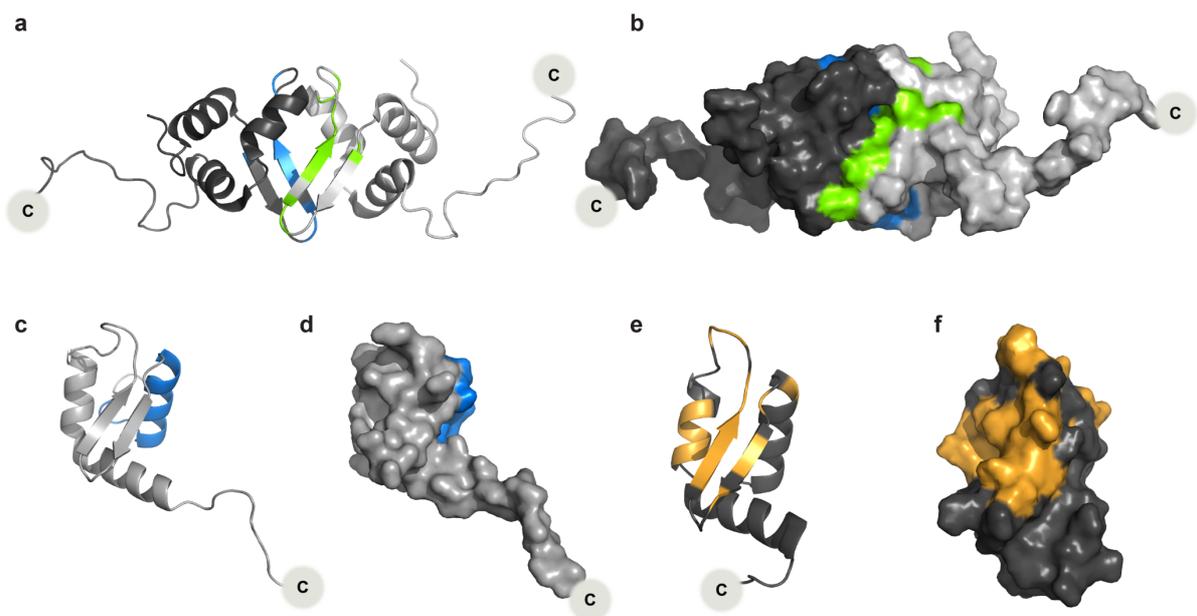
**Supplementary Figure 24 | Carbene footprinting<sup>13</sup> of the *holo*-Bamb\_5917 PCP domain in complex with Bamb\_5915.** **a-b**, Differential carbene labelling observed for tryptic and AspN fragments of the *holo*-Bamb\_5917 PCP domain in isolation and in complex with Bamb\_5915. **a**, Fraction of the *holo*-Bamb\_5917 PCP domain tryptic fragments labelled by the carbene in the absence (black) and presence (grey) of Bamb\_5915. **b**, Fraction of the *holo*-Bamb\_5917 PCP domain AspN fragments labelled by the carbene in the absence (black) and presence (grey) of Bamb\_5915. Fragments that showed a significantly different level of labelling are highlighted with red (masked) and cyan (unmasked) dots. Error bars are  $\pm$  standard deviation for three measurements (three independent replicates). **c-d**, Mapping of the proteolytic fragments from trypsin and AspN digestions onto the sequence of the Bamb\_5917 PCP domain. **e**, Combined mapping of peptides from **c-d**. Masked regions, which acquire less labelling and thus have reduced solvent accessibility as a result of binding, are highlighted in red. Regions which are unmasked and thus acquire additional labelling (and therefore are more solvent exposed) are highlighted in cyan. Peptides that were detected but with unchanged levels of labelling are highlighted in blue. The data obtained by combining the trypsin and AspN digestions results in 100% sequence coverage. Note that changes in solvent accessibility as a result of binding may be due to direct interaction or conformational change as a result of binding. The dotted lines in **c-e** outline the location of helices in the structure.



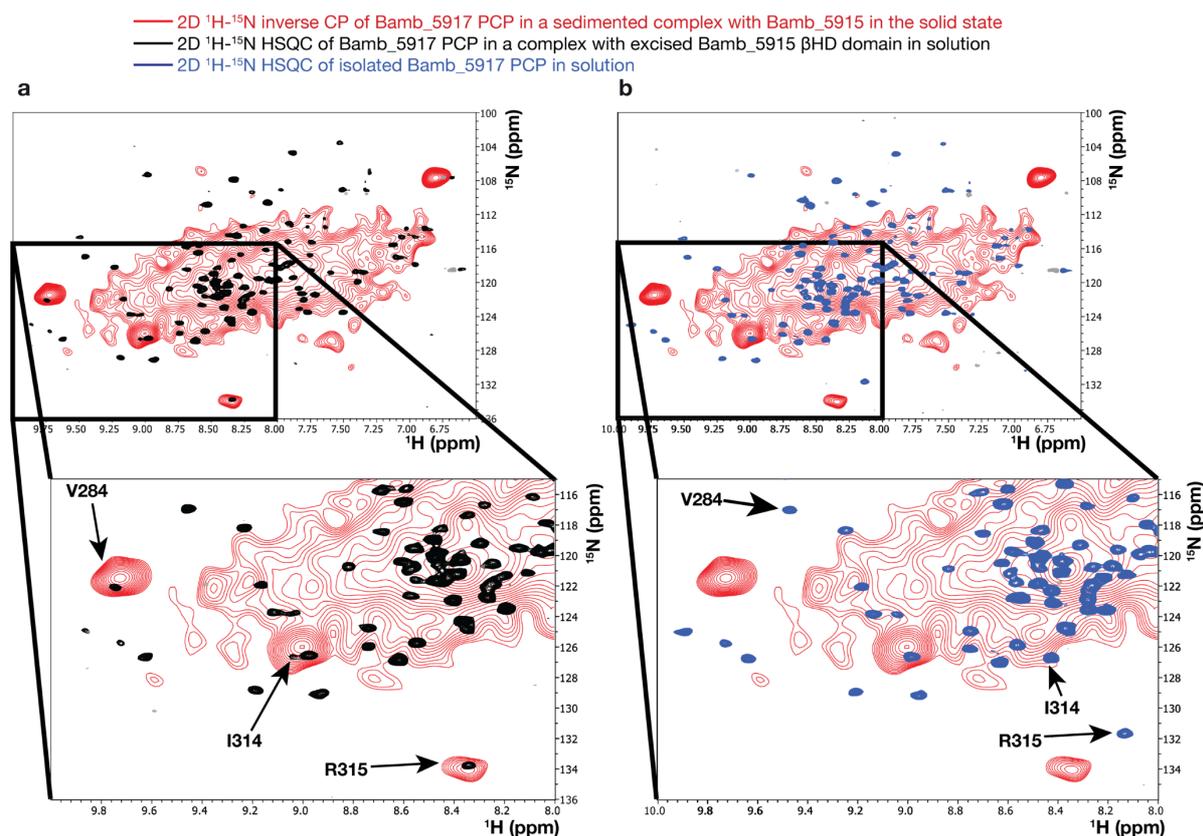
**Supplementary Figure 25 | Differential carbene labelling for tryptic and AspN fragments of Bamb\_5915.** **a**, Fraction of Bamb\_5915 tryptic fragments labelled by the carbene in the absence (black) and presence (grey) of the *holo*-Bamb\_5917 PCP domain. **b**, Fraction of Bamb\_5915 AspN fragments labelled by the carbene in the absence (black) and presence (grey) of the *holo*-Bamb\_5917 PCP domain. Residues in fragments that showed a significantly different level of labelling are highlighted with red (masked) and cyan (unmasked) dots. Error bars are  $\pm$  standard deviation for three measurements (three independent replicates). **c-d**, The proteolytic fragments from trypsin and AspN digestions (a-b) mapped onto the sequence of Bamb\_5915 (with part of the His-tag sequence omitted).



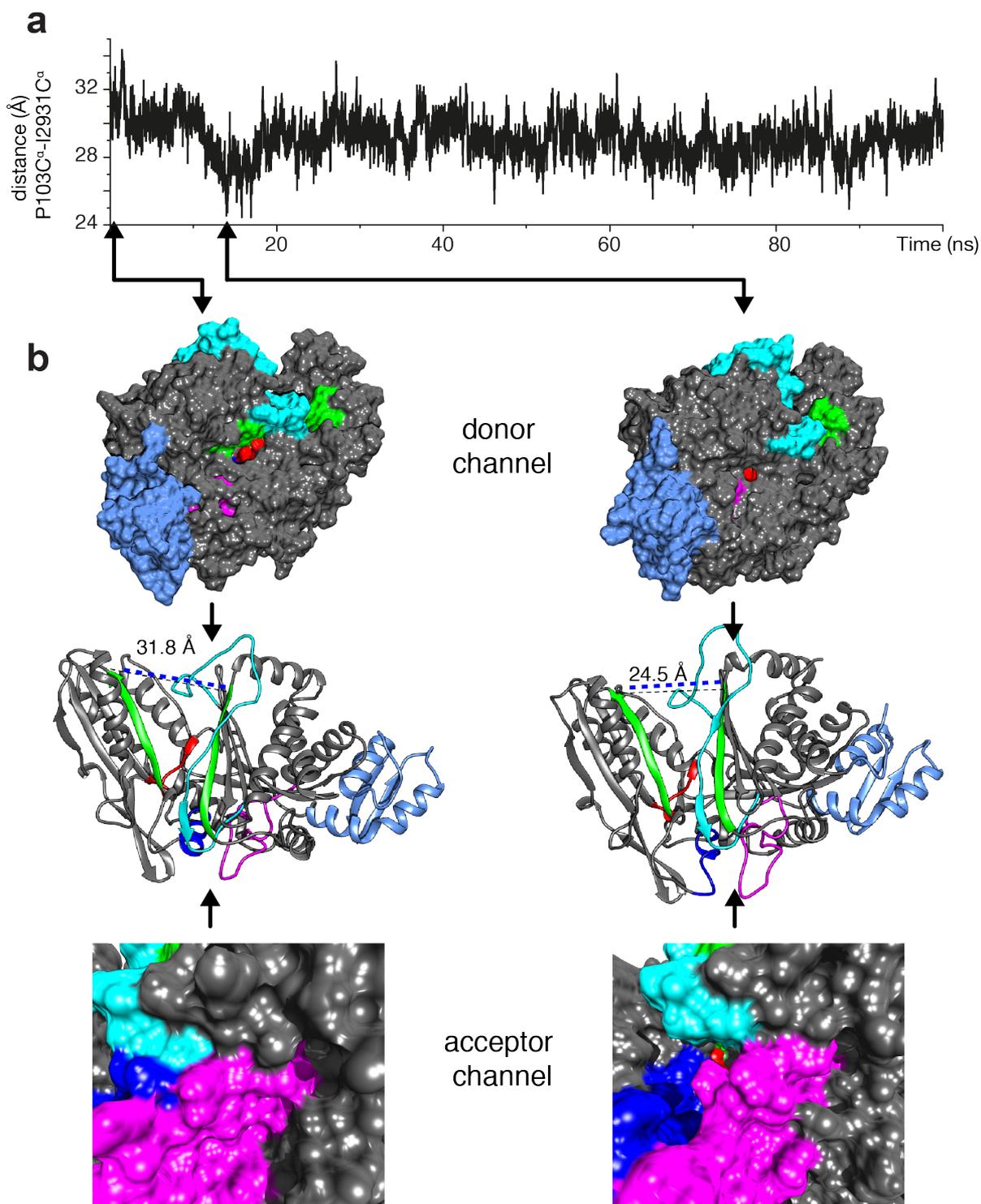
**Supplementary Figure 26 | Combined mapping of peptides from carbene footprinting on Bamb\_5915 in complex with the *holo*-Bamb\_5917 PCP domain.** Masked regions, which acquire less labelling and thus have reduced solvent accessibility as a result of binding, are highlighted in red. Regions which are unmasked and thus acquire additional labelling (and therefore are more solvent exposed) are highlighted in cyan. Peptides that were detected but with unchanged levels of labelling are highlighted in blue. The data obtained by combining trypsin and AspN digestions (see Fig 25) provide 69% sequence coverage. Note that changes in solvent accessibility as a result of binding may be due to direct interaction or conformational change as a result of binding.



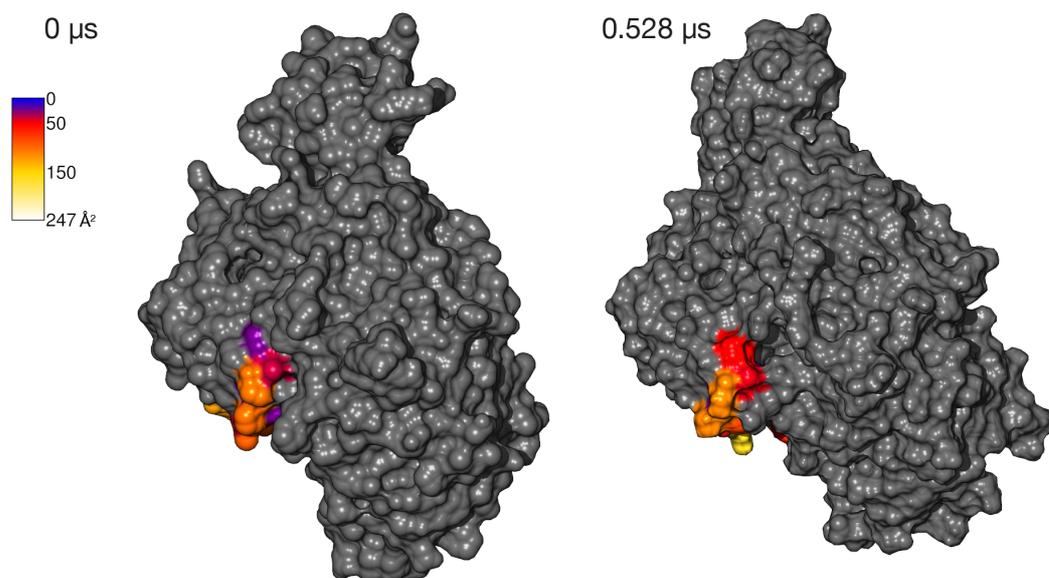
**Supplementary Figure 27 | Interaction interface of TubCdd.** **a**, Cartoon rendering of the TubCdd dimer<sup>15</sup>. **b**, Surface rendering of the TubCdd dimer. The amino acid residues identified in NMR titrations of the mono-dimer TubCdd with TubBdd from the tubulysin PKS/NRPS<sup>15</sup> are highlighted in green and blue on each monomer. **c-d**, The Bamb\_5915  $\beta$ HD domain with the region masked in the carbene footprinting experiments indicated in blue. **e-f**, Structure of the  $\beta$ HD domain homologue Kj12C-<sup>N</sup>DD (PDB code: 6EWU) from the rhabdopeptide NRPS<sup>16</sup>. The amino acid residues identified in NMR titrations with Kj12B-<sup>C</sup>DD are highlighted in orange.



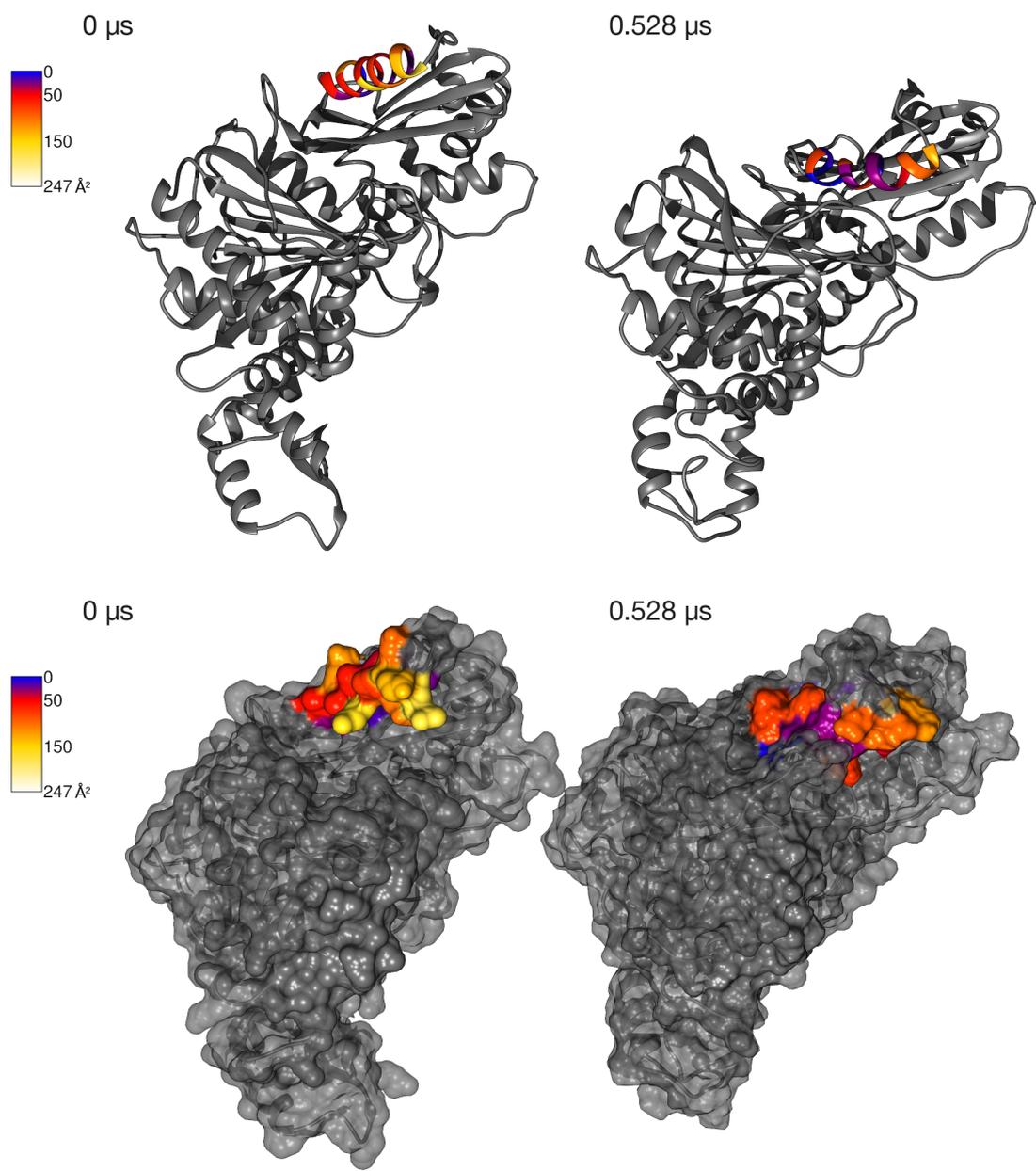
**Supplementary Figure 28 | Overlay of 2D  $^1\text{H}$ - $^{15}\text{N}$  correlation spectra of the *holo*-Bamb\_5917 PCP domain in a sedimented complex with Bamb\_5915 in the solid state (red) and in solution (black & blue). a, Overlay with *holo*-Bamb\_5917 PCP domain in complex with the excised  $\beta\text{HD}$  domain in solution (black). b, Overlay with the isolated *holo*-Bamb\_5917 PCP domain in solution (blue). The  $^1\text{H}$ - $^{15}\text{N}$  cross peaks for the V284 (located at the terminus of  $\alpha$ -helix 4), I314 and R315 (located at the C-terminus of the SLiM) in the complex of the *holo*-Bamb\_5917 PCP domain with the Bamb\_5915  $\beta\text{HD}$  domain appear in essentially the same positions in the complex with full length Bamb\_5915, which indicates preservation of the interactions. The identities of the indicated resonances were confirmed in a series of assignment spectra in the solid state. The solid-state NMR spectrum (red) was obtained using a  $^1\text{H}$ - $^{15}\text{N}$  cross-polarisation based experiment, which uses dipolar coupling for polarization transfer and which requires that the observed sites are largely immobilized. The solution  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum was obtained at 700 MHz  $^1\text{H}$  Larmor frequency on a sample of the  $[\text{U}-^{13}\text{C}, ^{15}\text{N}]\text{holo}$ -Bamb\_5917 PCP domain in the presence and absence of the unlabelled Bamb\_5915  $\beta\text{HD}$  domain. The solid state  $^1\text{H}$ - $^{15}\text{N}$  CP-based spectrum was obtained at 100 kHz magic angle spinning at 700 MHz  $^1\text{H}$  Larmor frequency on the  $[\text{U}-^2\text{H}, ^{13}\text{C}, ^{15}\text{N}]\text{holo}$ -Bamb\_5917 PCP domain in complex with natural abundance full-length Bamb\_5915 sedimented from a solution where water was 100%  $\text{H}_2\text{O}$ .**



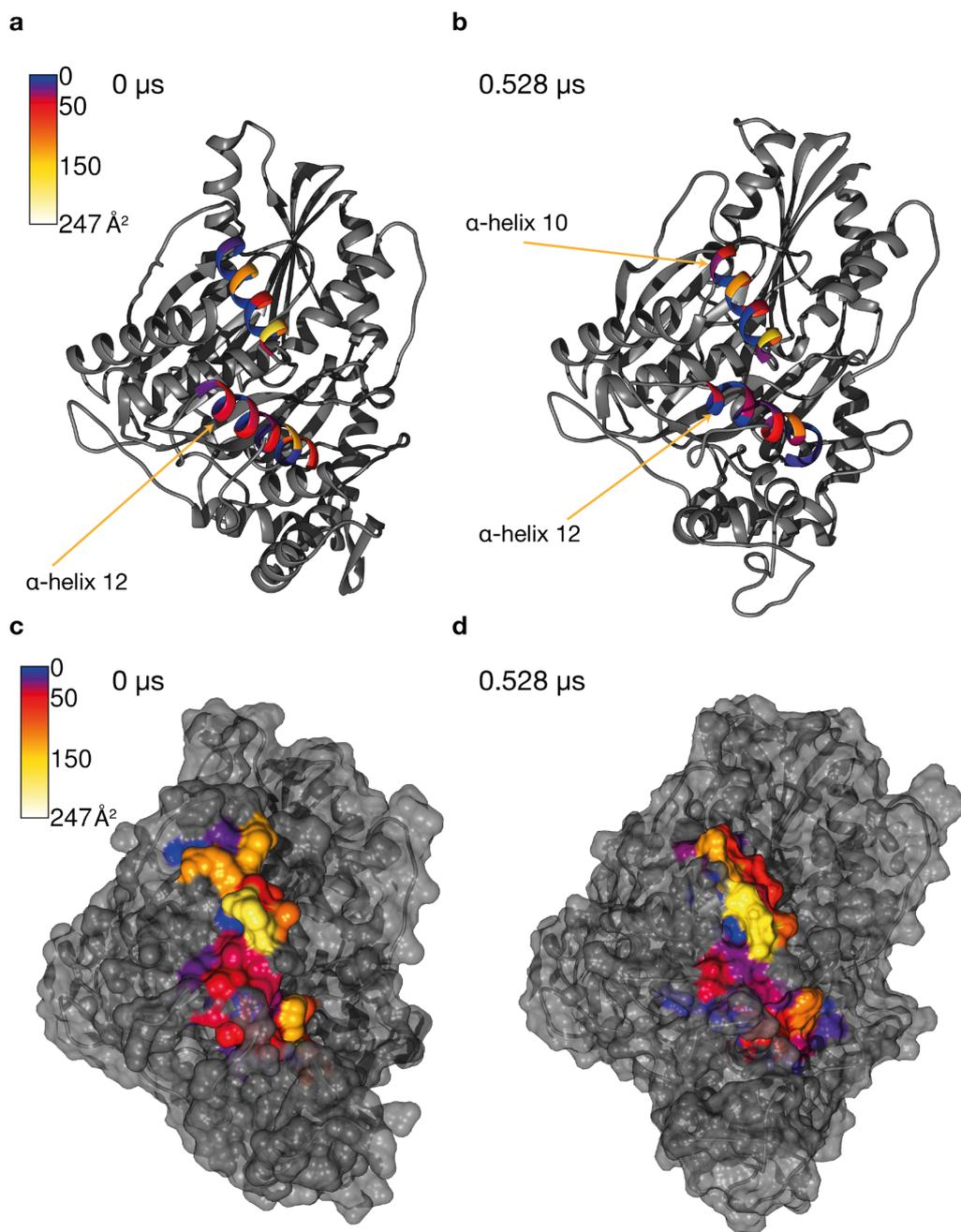
**Supplementary Figure 29 | Standard 100 ns molecular dynamics (MD) simulations of Bamb\_5915. | a,** Evolution of the distance between P103C<sup>α</sup> and I293C<sup>α</sup> during 100 ns MD simulations. This distance is used as an indication of the degree of association between the C- and N-terminal subdomains. **b,** Conformations of Bamb\_5915 at different points of the MD simulation (indicated with arrows). Views of the donor channel, acceptor channel and a side view highlighting the V-conformation are shown. In the crystal structure (left)  $\alpha$ -helix 4 (blue) and parts of the loops between  $\beta$ -strands 13 and 14 (cyan) and  $\beta$ -strand 12 and  $\alpha$ -helix 9 (magenta) block the acceptor channel but the donor channel is open providing access to the active site (conserved active site residues are indicated in red). To aid the reader  $\beta$ -strands 3 and 9 are coloured green.



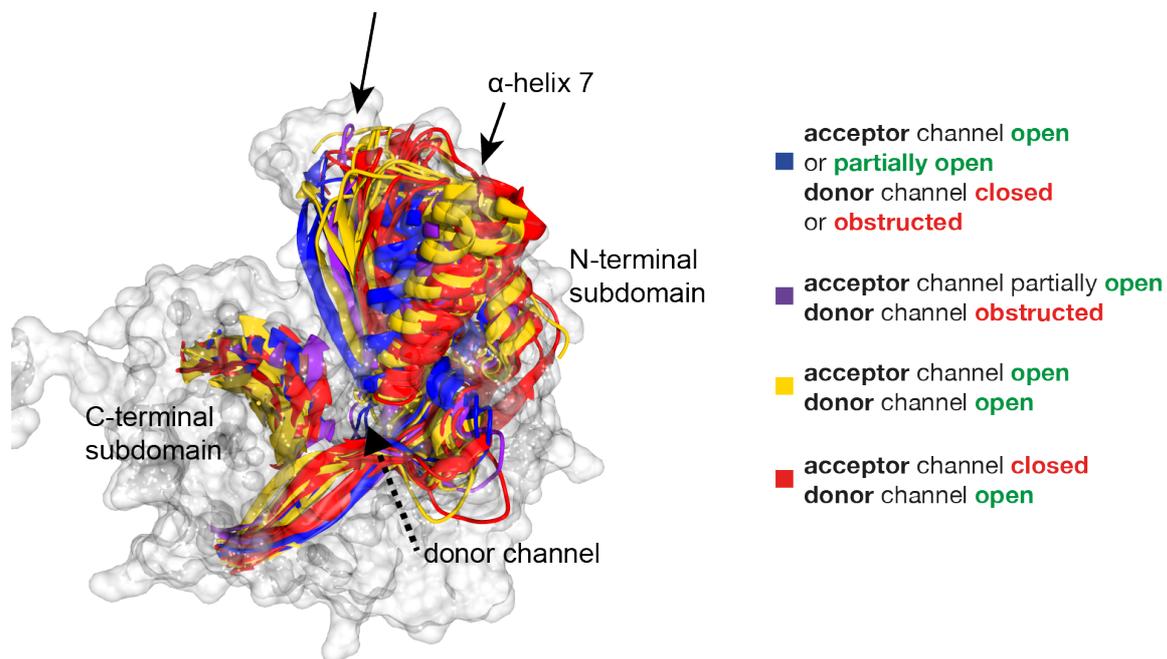
**Supplementary Figure 30 | Solvent accessibility of the C-terminal part of the “latch” region in Bamb\_5915 during aMD simulations.** The figure shows total solvent accessible area per residue in Bamb\_5915 during aMD at 0 μs (left) and 0.528 μs (right). The region becomes more solvent accessible as a result of opening of the acceptor channel.



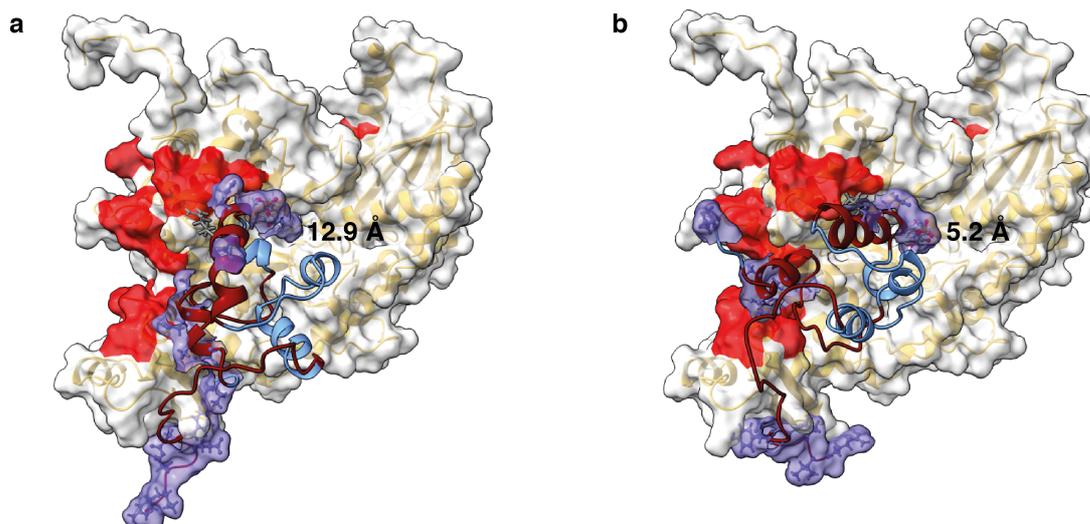
**Supplementary Figure 31 | Solvent accessibility of  $\alpha$ -helix 6 in Bamb\_5915 during aMD simulations.** The figure shows total solvent accessible area per residue in Bamb\_5915 during aMD at 0  $\mu\text{s}$  (left) and 0.528  $\mu\text{s}$  (right). The highlighted  $\alpha$ -helix 6 becomes more buried as a result of conformational change in Bamb\_5915. Carbene footprinting data indicate masking of part of  $\alpha$ -helix 6 as a result of the binding of the Bamb\_5917 PCP domain to Bamb\_5915, which suggests that upon binding Bamb\_5915 undergoes a conformational change resulting in  $\alpha$ -helix 6 becoming less solvent accessible.



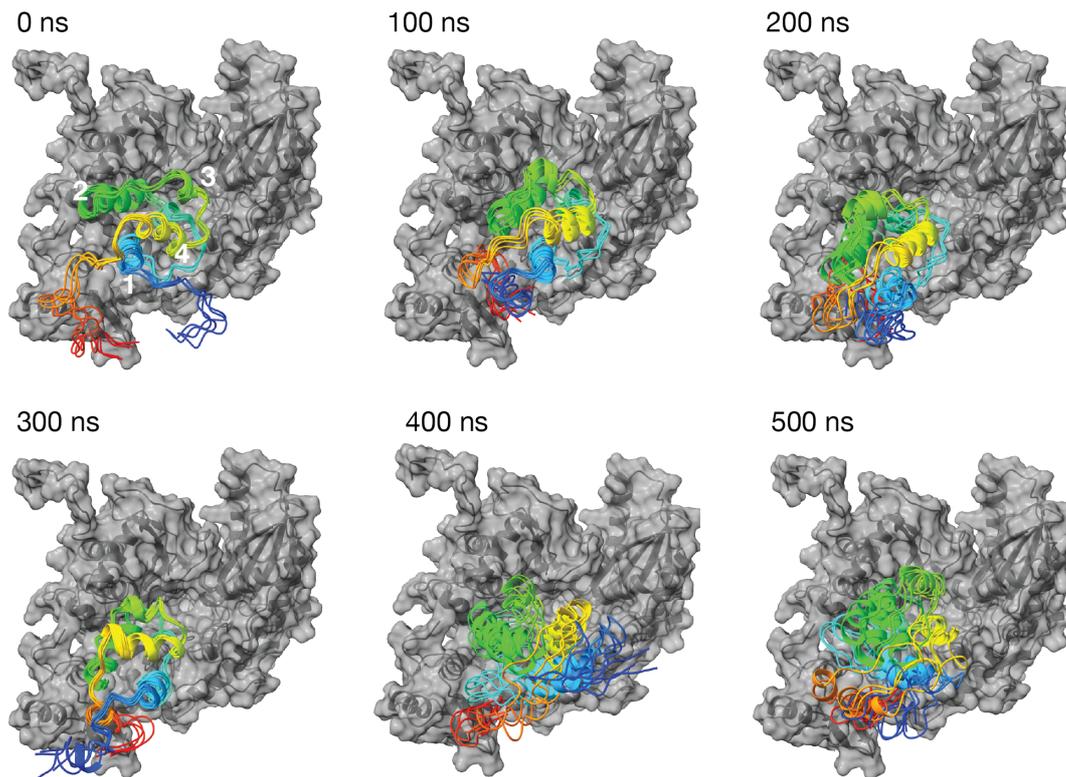
**Supplementary Figure 32 | Sum of solvent accessible area for residues in  $\alpha$ -helix 10 and 12 in Bamb\_5915 during aMD simulations.** The figure shows total solvent accessible area per residue in Bamb\_5915 during aMD at  $0 \mu\text{s}$  (left) and  $0.528 \mu\text{s}$  (right).  $\alpha$ -helix 12 becomes more buried as a result of conformational change in Bamb\_5915. At the same time there are no major changes in solvent accessibility of  $\alpha$ -helix 10. In addition to binding of the Bamb\_5917 PCP domain, the burying of the N-terminal end of  $\alpha$ -helix 12 could contribute to the masking in that region observed in the carbene footprinting experiments.



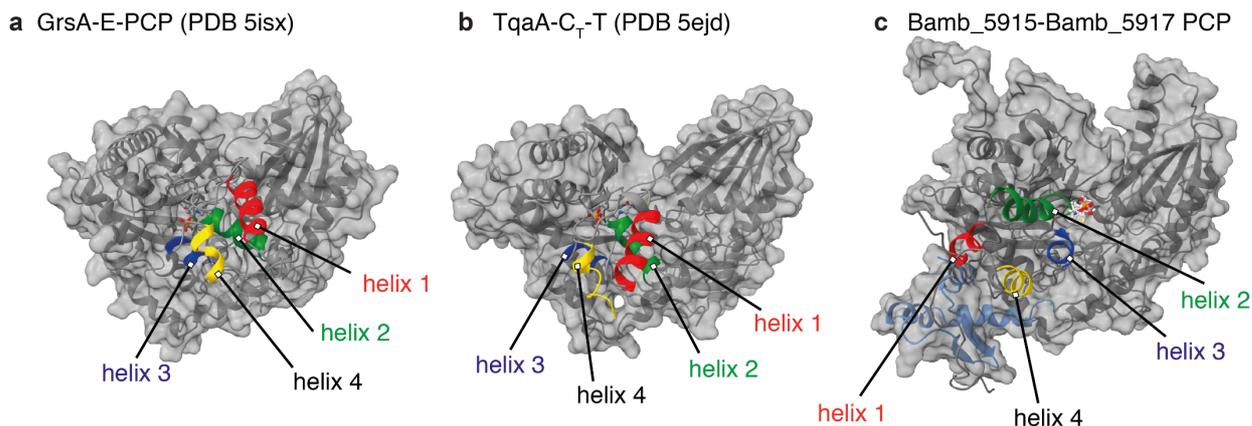
**Supplementary Figure 33 | Correlation between conformations of domains with the CAT fold versus the state of the acceptor and donor channel.** A structural alignment was performed in UCSF Chimera<sup>17</sup> and the structures colour coded according to the state of the channels leading to the active site. Structures with a fully or partially open acceptor channel and a closed or obstructed donor channel are coloured blue. The structure with a partially open acceptor channel and an obstructed donor channel is coloured purple. Structures with both the acceptor and donor channels open are coloured yellow. Structures with a closed acceptor channel and an open donor channel are coloured red. The structural alignment was performed using MatchMaker<sup>18</sup>. First pairwise sequence alignment was performed using a global Needleman-Wunsch algorithm and then the alignment was used to superimpose the structures by fitting the aligned residue pairs. The fitting uses one point per residue, which corresponds to the C $^{\alpha}$  site. The structures with similar states of the channels leading to the active site cluster in the same ranges of conformations between the C-terminal and N-terminal subdomains. The change in position of the N-terminal subdomain relative to the C-terminal subdomain can be visualised by considering the position of the  $\beta$ -sheet and  $\alpha$ -helix 7 relative to the N-terminal subdomain. An arrow highlights the location of the  $\beta$ -sheet in the C-terminal subdomain. For clarity only the  $\beta$ -sheet and  $\alpha$ -helix 7 are shown in cartoon representation, with the remainder of Bamb\_5915 rendered as a semi-transparent surface. The structures considered include the fungal terminal-condensation-like C<sub>T</sub> domain of TqaA(5ejd)<sup>19</sup>, the X domain from the teicoplanin pathway (4tx3)<sup>20</sup>, epimerization domains: GrsA-E1 from the gramicidin pathway (5isx)<sup>21</sup>, TycA-E1 from the tyrocidine pathway (2xhg)<sup>22</sup>, condensation domains: Bamb\_5915 from the enacyloxin pathway (6cgo), SgcC5 from the C-1027 pathway (4zxw)<sup>22</sup>, TycC5 from the tyrocidine pathway (2jgp)<sup>23</sup>, EntFC from the enterobactin pathway (5t3d)<sup>24</sup>, VibH from the vibriobactin pathway (115a)<sup>24</sup>, AB3403C (4zxh)<sup>24</sup>, SrfA-C (2vsq)<sup>25</sup>, CDA-C1 (4jn3)<sup>25</sup>, BmdB-Cy2 (5t3e)<sup>26</sup> and EpoB-Cy (5t81 & 5t7z)<sup>7</sup>.



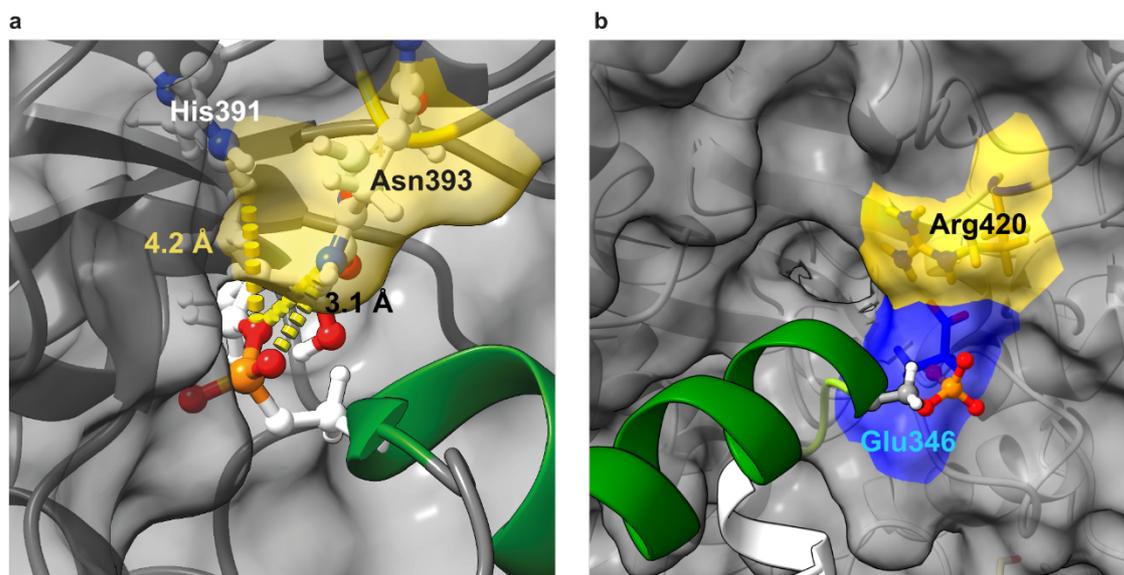
**Supplementary Figure 34 | Examples of docking simulations for the Bamb\_5917 PCP domain with Bamb\_5915 performed in HADDOCK<sup>27</sup>.** **a**, Model of the complex obtained by docking the NMR structure of the Bamb\_5917 PCP domain with the X-ray structure of Bamb\_5915. **b**, Model of the complex obtained by docking the structure of the Bamb\_5917 PCP domain from a frame of aMD at 180 ns with the X-ray structure of Bamb\_5915. The distance between the Ppant thiol group and the N<sup>62</sup> of the H205 residue in the active site is given next to the Ppant. In general, when the conformational changes observed in the aMD are considered, the PCP domain docks closer to the active site. The regions masked on the surface of Bamb\_5915 in the carbene footprinting experiments are red. The regions of the Bamb\_5917 PCP domain masked in the carbene footprinting experiments are maroon. The residues identified as interacting in the NMR titrations are blue and their side chains are displayed.



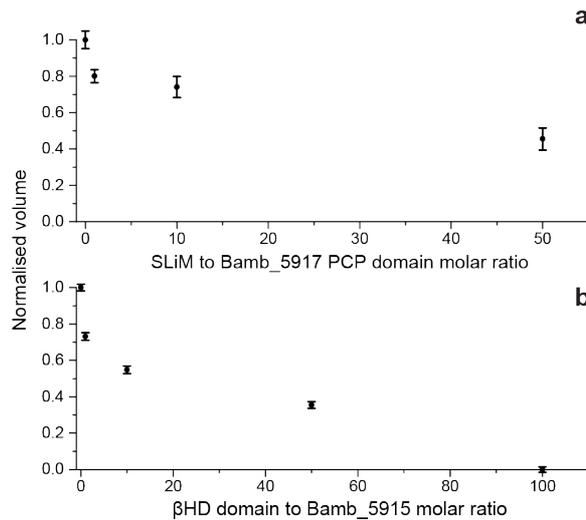
**Supplementary Figure 35 | Clusters of best solutions in docking simulations of the Bamb\_5917 PCP domain:Bamb\_5915 complex performed in HADDOCK using conformers from 500 ns aMD of the Bamb\_5917 PCP domain sampled every 100 ns as starting points.** The time shown in the top left corner of each image is the number of ns after initiation of the aMD simulation for which PCP conformers were selected for further simulation using HADDOCK.  $\alpha$ -helix 1 is coloured blue,  $\alpha$ -helix 2 is coloured green,  $\alpha$ -helix 3 is coloured yellow-green,  $\alpha$ -helix 4 is coloured yellow and the SLiM is coloured red. The helix numbering is also indicated in the first panel. Residues 215-228 and 294-315 in the Bamb\_5917 PCP domain were defined as fully flexible in the HADDOCK simulations.



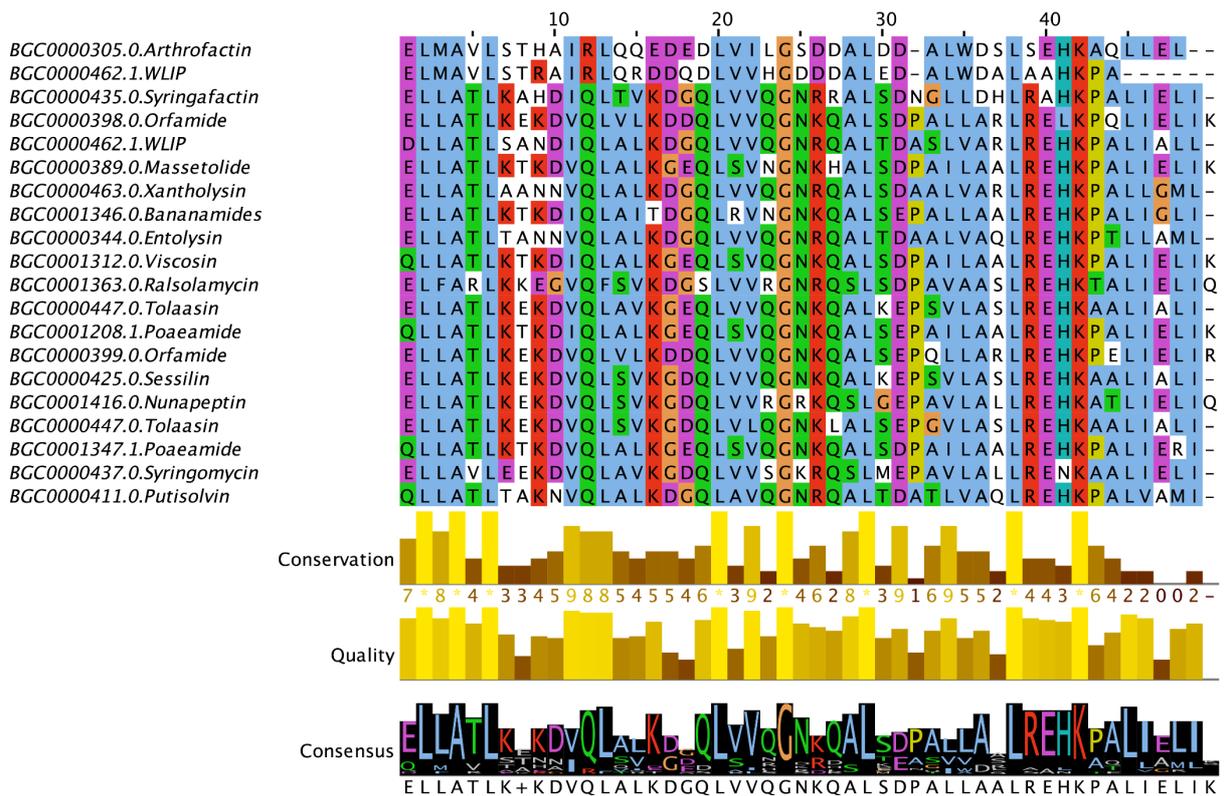
**Supplementary Figure 36** | Comparison of the orientation of PCP domains docked at the donor side of C domains. **a**, GrsA-E-PCP (PDB code: 5isx)<sup>21</sup>. **b**, TqaA-C<sub>T</sub>-T (PDB code: 5ejd)<sup>27</sup>. **c**, Bamb\_5915-Bamb\_5917 PCP domain. Helices 1, 2, 3 and 4 are coloured red, green, blue and yellow, respectively.



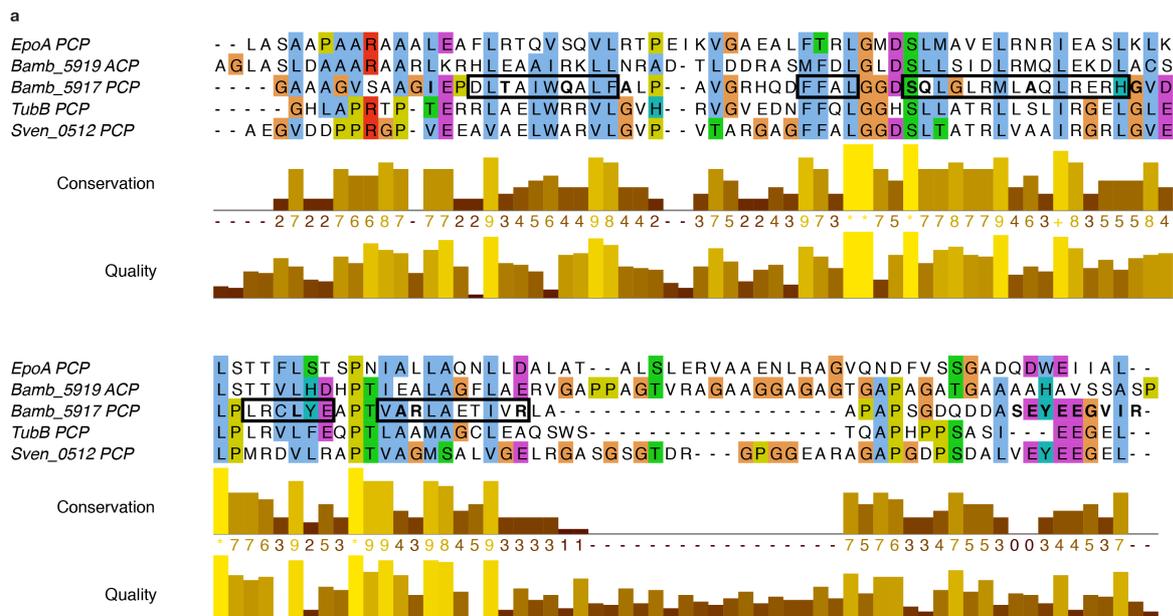
**Supplementary Figure 37** | Interactions of residues from the latch with the phosphor diester group of the Ppant arm. **a**, In TqaA C<sub>T</sub>-T His and Asn residues from the latch (the surface of these residues is highlighted in yellow) interact (dashed yellow lines) with the phosphodiester (orange-red)<sup>21</sup> of the Ppant arm. **b**, Structural alignment of the Bamb\_5917 PCP domain-Bamb\_5915 complex with a frame taken at 650 ns in the aMD simulation of Bamb\_5915, where the latch has formed two hydrogen bonds with β-strand 3. The side chain of the R420 residue (the atoms (except nitrogen – in blue) and surface are highlighted in yellow) interacts with the E346 residue (atoms and surface in blue). In the model of the complex, the position of the phosphodiester group of the Ppant arm (orange-red) is similar to that of the side chain of the E346 residue. The phosphodiester group may therefore interact with the Arg420 residue of the latch instead of the Glu346 residue.



**Supplementary Figure 38** | a, Area under peaks due to *N*-acetyl-AHCCA in LC-MS chromatograms as a function of the molar ratio of the Bamb\_5917 SLiM to the Bamb\_5917 PCP domain, and the Bamb\_5915  $\beta$ HD domain to Bamb\_5915 in the competition assays using the Bamb\_5917 SLiM peptide (a) and the excised Bamb\_5915  $\beta$ HD domain (b). Error bars represent standard deviation of the volumes over three independent replicates. The volumes are normalized to the peak volume at 0 molar ratio.



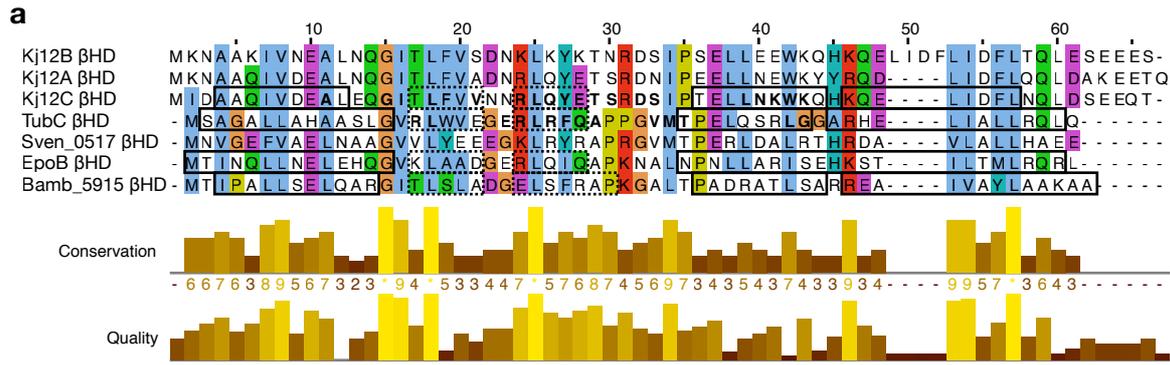
**Supplementary Figure 39** | Multiple sequence alignment for  $\beta$ HD domains appended to the N-terminus of dual C/E domains. Sequences from the systems included in the MIBiG<sup>28</sup> database with E-val < 10<sup>-10</sup> are shown. Sequences were aligned using MUSCLE<sup>29</sup> and visualized using Jalview<sup>30</sup>.



**b**

	Percent identity				
EpoA PCP	100.0	31.7	21.0	29.9	20.9
Bamb_5919 ACP	31.7	100.0	23.8	26.8	23.1
Bamb_5917 PCP	21.0	23.8	100.0	32.6	37.4
TubB PCP	29.9	26.8	32.6	100.0	47.4
Sven_0512 PCP	20.9	23.1	37.4	47.4	100.0

**Supplementary Figure 40 | Sequence alignments of PCP domains.** **a**, Sequence alignments of PCP domains with SLiMs from the tubulyisin (TubB PCP)<sup>15</sup>, watasemycin (Sven\_0512 PCP), enacyloxin (Bamb\_5917 PCP) and epothilone (EpoA PCP) assembly lines. For comparison the sequence of the Bamb\_5919 ACP domain is also included. Black rectangles highlight the helices in the Bamb\_5917 PCP domain. Residues with local apparent  $K_a$  values larger than the average plus the standard deviation, obtained from titrations of the *holo*-Bamb\_5917 PCP domain with Bamb\_5915, are in bold. **b**, Identity matrix calculated using MUSCLE<sup>29</sup>.

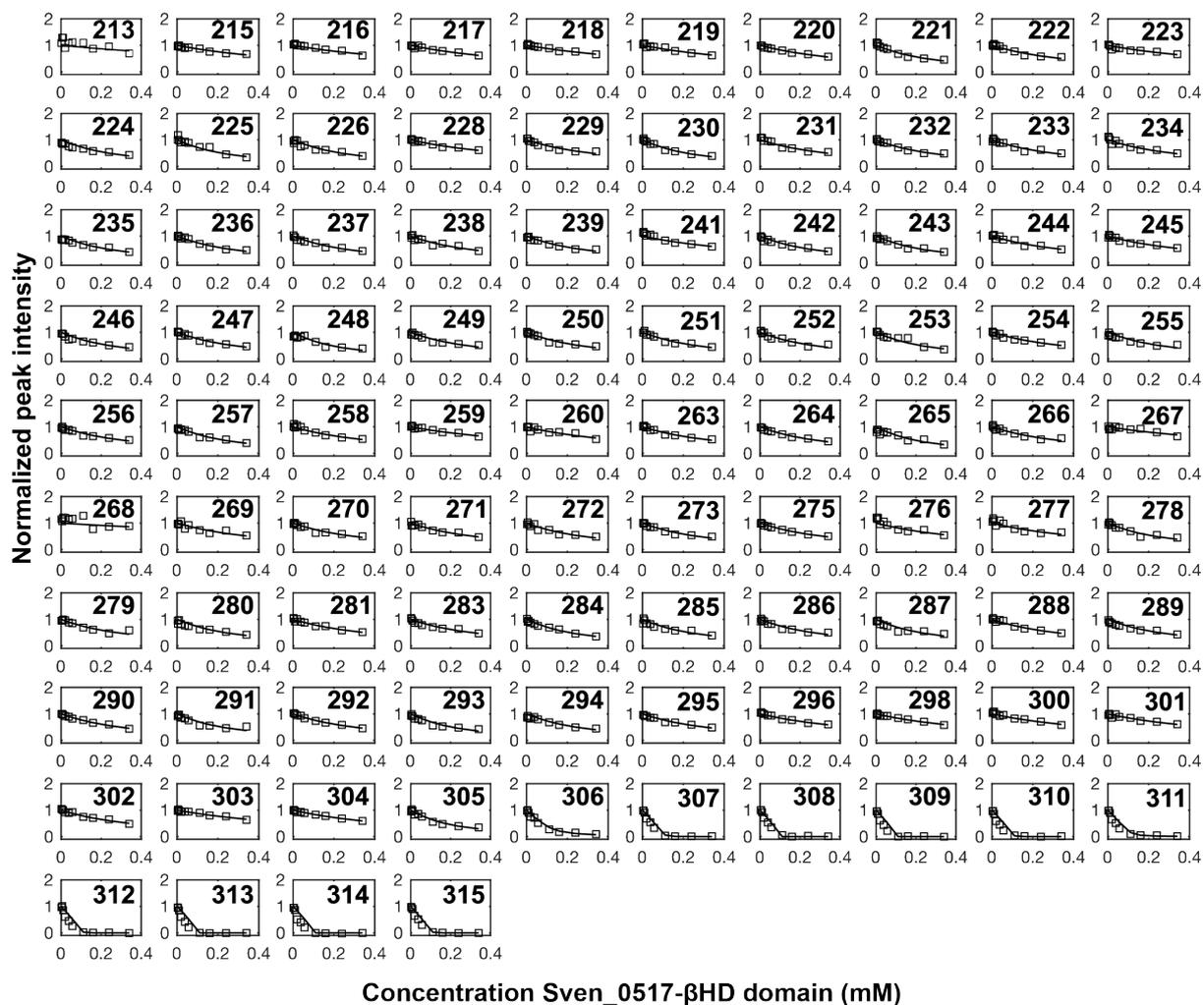


**c**

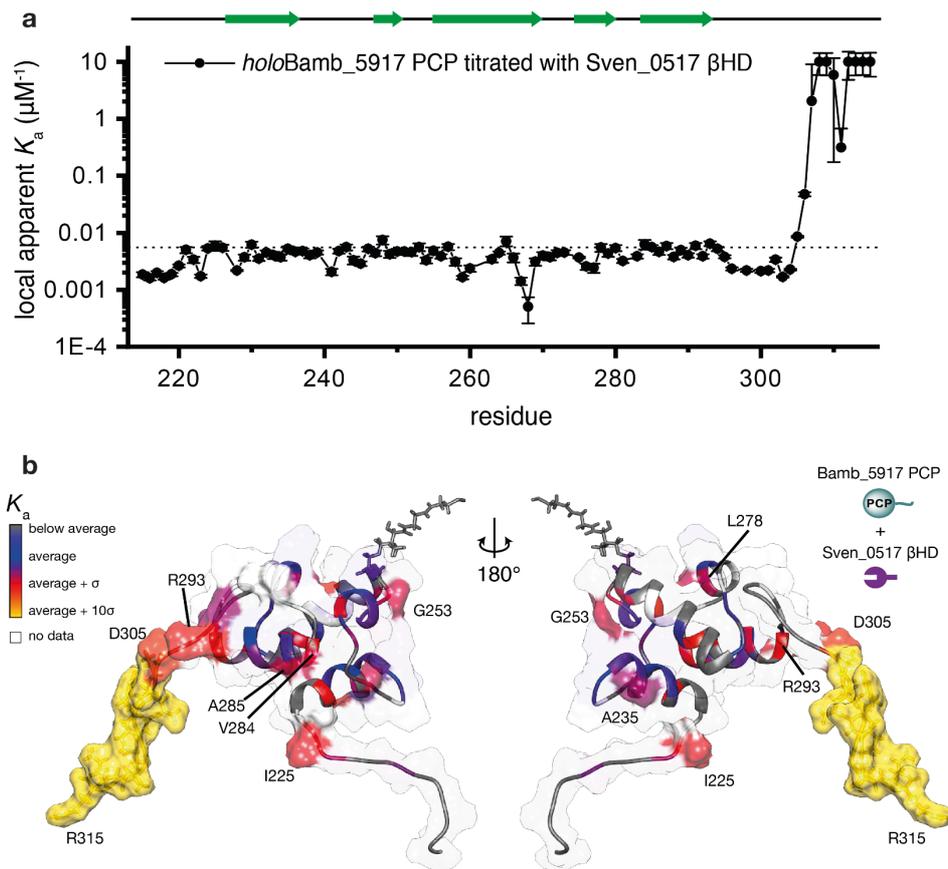
Percent identity

Kj12B βHD	100.0	69.4	71.0	25.0	25.0	21.4	14.3
Kj12A βHD	69.4	100.0	74.2	26.8	21.4	21.4	17.9
Kj12C βHD	71.0	74.2	100.0	26.8	19.6	25.0	12.5
TubC βHD	25.0	26.8	26.8	100.0	39.3	33.9	30.4
Sven_0517 βHD	25.0	21.4	19.6	39.3	100.0	25.0	<b>35.7</b>
EpoB βHD	21.4	21.4	25.0	33.9	25.0	100.0	35.7
Bamb_5915 βHD	14.3	17.9	12.5	30.4	<b>35.7</b>	35.7	100.0

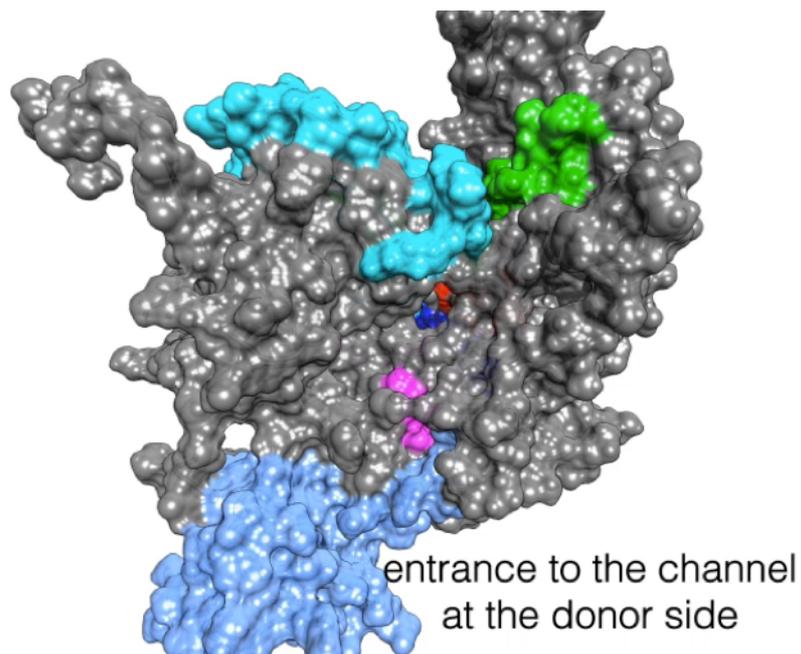
**Supplementary Figure 41 | Sequence alignment of βHD domains.** **a**, Sequence alignments of βHD domains from the tubulysin, watasemycin, enacyloxin, epothilone, and rhabdopeptide/xenortide assembly lines. Black rectangles highlight helices in systems with known structures. Solid line rectangles outline β-strands for systems with known structures. Interacting residues are in bold. **b**, Phylogenetic tree. **c**, Identity matrix. The alignment was performed using MUSCLE<sup>29</sup>.



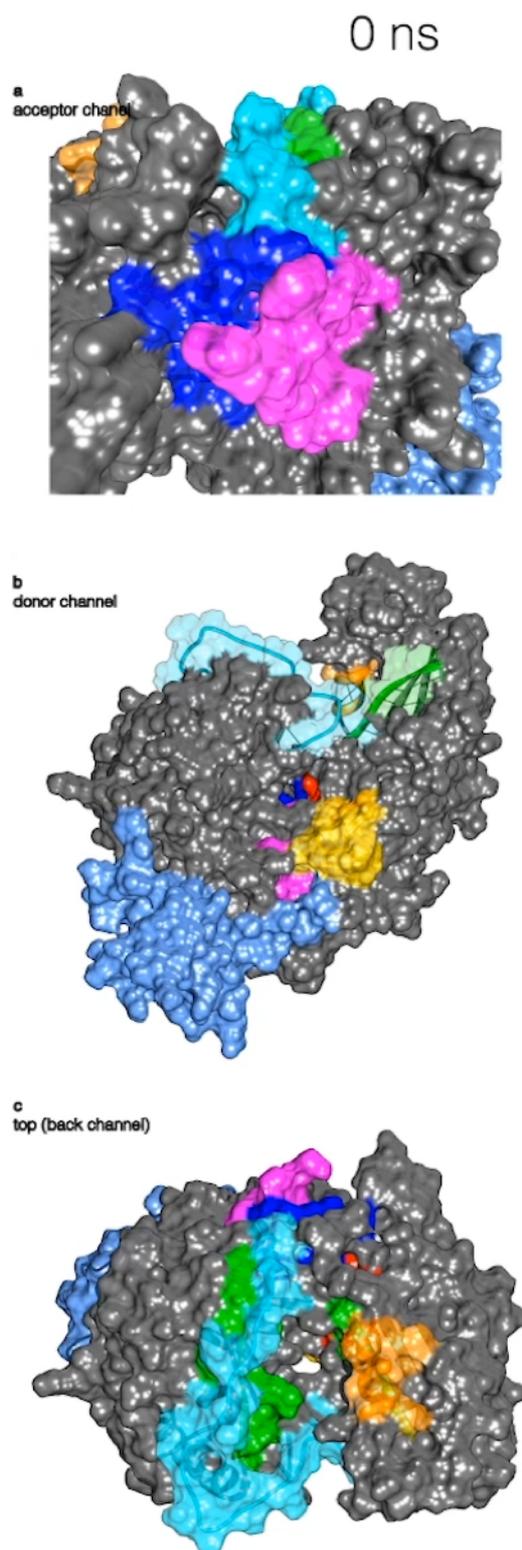
Supplementary Figure 42 | Titration profiles for the peaks disappearing in the NMR titration of the  $^{15}\text{N}$ -labelled *holo*-BamB\_5917 PCP domain with the Sven\_0517  $\beta$ HD domain in solution resulting from complex formation. The squares represent experimental volumes normalised to peak volume in the absence of the Sven\_0517  $\beta$ HD domain and the lines represent the best fit curves to equation 3 (see Supplementary Section *NMR titration data fitting*). Residue numbers for the BamB\_5917 PCP domain are given in the top right corner of each panel.



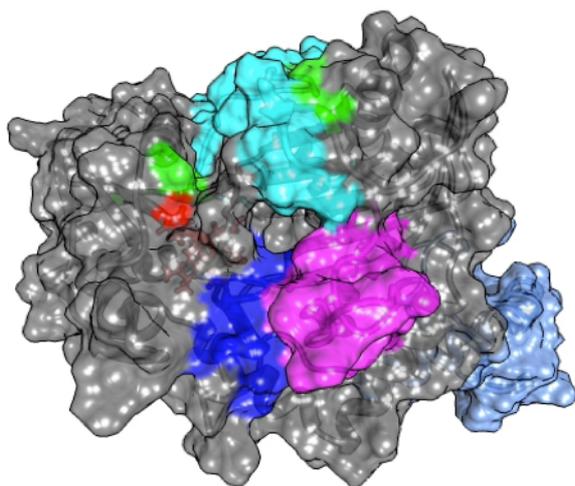
**Supplementary Figure 43 | Local apparent association constants,  $K_a$ , obtained from titrations of the *holoBamb\_5917* PCP domain with the *Sven\_0517*  $\beta$ HD domain<sup>30</sup>.** **a**, Local apparent  $K_a$  plotted as a function of residue number. The horizontal lines indicate thresholds for the average value plus one standard deviation. For the calculation of the average and standard deviations the residues with values larger than 3 standard deviations from the average were removed iteratively. Error bars represent two standard deviations from 1000 iterations in Monte Carlo error propagation. **b**, Higher than average local apparent  $K_a$  projected onto the structure of the *Bamb\_5917* PCP domain. The overall pattern of the residues with above average  $K_a$  is similar to that found in the titration with the *Bamb\_5915*  $\beta$ HD domain. The strongest interaction with the *Sven\_0517*  $\beta$ HD domain involves the C-terminal SLiM of the *Bamb\_5917* PCP domain. The observed pattern suggests that the *Bamb\_5917* PCP domain undergoes a similar conformational change upon binding to that seen for interaction with the *Bamb\_5915*  $\beta$ HD domain.



**Supplementary Movie 1 | Fly through the solvent channel in the X-ray crystal structure of Bamb\_5915.** The different secondary structure elements that contribute to blocking the acceptor channel in the X-ray crystal structure of Bamb\_5915 are blue, magenta and cyan. The latch is cyan. Conserved active site residues are red. The  $\beta$ HD domain is light blue.  $\beta$ -Strands 3 and 9 are green. The movie is provided as a separate file: *Supplementary\_Movie1.mp4*.

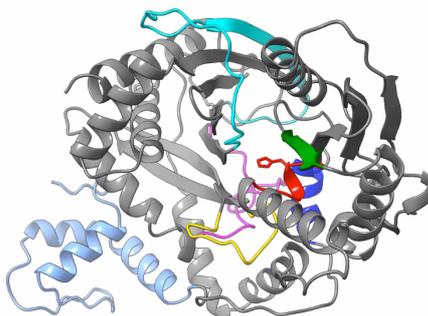


**Supplementary Movie 2 | 1  $\mu$ s aMD simulation of BamB\_5915.** **a**, The acceptor channel. **b**, The donor channel. **c**, Top view showing the back channel and the V-conformation. The residues blocking the acceptor channel are coloured blue, magenta and cyan. The latch and floor loop are cyan and yellow, respectively.  $\beta$ -Strands 3 and 9 are green. The  $\beta$ HD domain is light blue. The conserved active site residues are red.  $\alpha$ -helix 6, one of the sites masked in the carbene footprinting experiments upon binding of BamB\_5917 PCP domain but away from the expected binding interface, is coloured orange. The time point in the aMD simulation is shown at the top of panel **a**. The animation is rendered using every 100<sup>th</sup> frame from the aMD simulation. The movie is provided as a separate file: *Supplementary\_Movie2.mp4*.

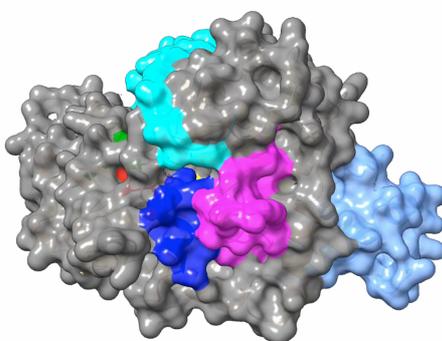


**Supplementary Movie 3 | Fly through the solvent channel after 0.528  $\mu$ s of aMD simulations of BamB\_5915 (frame 264199 - highlighted in Fig. 3) |** The different residues blocking the acceptor channel in the X-ray crystal structure of BamB\_5915 are blue, magenta and cyan. The latch is cyan. Conserved active site residues are red. The  $\beta$ HD domain is light blue.  $\beta$ -Strands 3 and 9 are green. The movie is provided as a separate file: *Supplementary\_Movie3.mp4*.

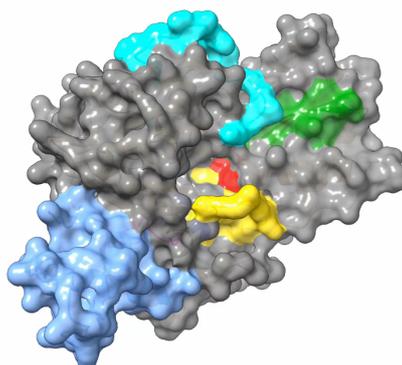
a



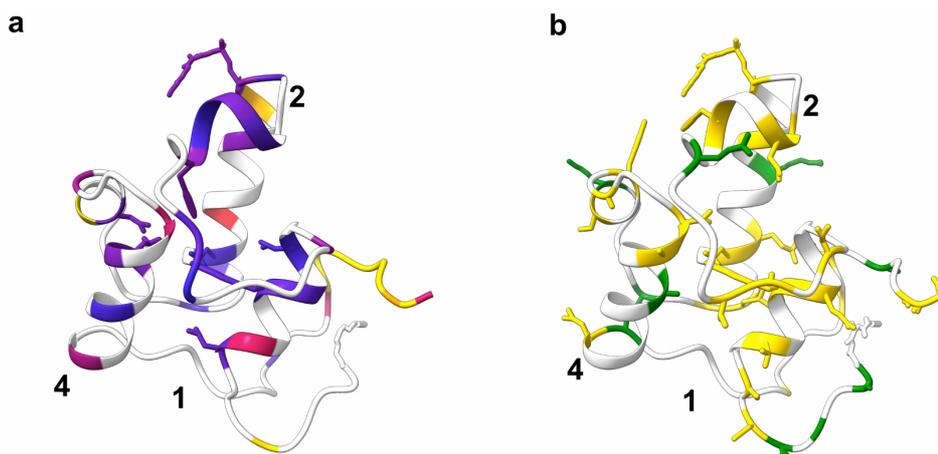
b



c



**Supplementary Movie 4 | The 1<sup>st</sup> mode from Principal Component Analysis of the 1  $\mu$ s aMD simulations of BamB\_5915.** The conformational change involves donation of part of the latch (cyan) to the N-terminal subdomain of BamB\_5915, in addition to anticorrelated modulation of the opening of the acceptor and donor side channels to the active site. As the acceptor channel opens up the donor channel becomes progressively more restricted. The main fragments contributing to the modulation of the acceptor side channel opening are the latch (cyan),  $\alpha$ -helix 4 (blue) and the loop between  $\beta$ -strand 12 and  $\alpha$ -helix 9 (magenta). The main contributor to the modulation of the donor side channel is the floor loop (yellow). The conserved active site residues are red. The movie is provided as a separate file: *Supplementary\_Movie4.mp4*.



**Supplementary Movie 5 | The 1<sup>st</sup> mode from Principal Component Analysis of the 0.5  $\mu$ s aMD simulations of the *holo*-Bamb\_5917 PCP domain.** The largest conformational change involves rotation and shifting of  $\alpha$ -helix 1 and a large amplitude motion of loop 1 leading to “uncovering” of the N-terminal end of  $\alpha$ -helices 3 and 4. In panel **a** the red residues are those most affected in solution NMR titrations of the *holo*-Bamb\_5917 PCP domain with Bamb\_5915. In panel **b** the yellow residues are those for which C $^{\alpha}$  chemical shifts change by >1 ppm between the Bamb\_5917 PCP domain free in solution and the Bamb\_5917 PCP domain in a sedimented complex with Bamb\_5915. The movie is provided as a separate file: *Supplementary\_Movie5.mp4*.

## Materials and Methods

### *Expression construct cloning and site-directed mutagenesis*

The plasmid constructs for expression of Bamb\_5915 and Bamb\_5917 PCP were generated using the Champion pET151 TOPO cloning kit as described in the accompanying manuscript. A plasmid for expression of the Bamb\_5915  $\beta$ HD domain (amino acids 1-67) was created by introduction of a stop codon in the pET151-Bamb\_5915 expression construct using the Q5 site-directed mutagenesis kit (New England Biolabs) with the primers 5'-ATCACGCCCTAGGCCGAGCTG-3' and 5'-CGTACCGGATCGGTGCG-3'. The presence of the desired mutation was confirmed by Sanger sequencing (GATC Biotech).

The gene encoding Sven\_0512 PCP was amplified by PCR using the forward primer: CACCGCGGAGGGCGTGGACGAT and the reverse primer: TCACAGCTCCCCCTCCTCGTAC and cloned into pET151 (Invitrogen) following the manufacturer's instructions to give the plasmid construct. The gene encoding Sven\_0517  $\beta$ HD domain was amplified by PCR using the forward primer: CACCATGAATGTCTGGGAGTTCGTGG and the reverse primer: TCAGTCGGCCTCGGTGGC and cloned into pET151 (Invitrogen) following the manufacturer's instructions to give the plasmid construct.

**Supplementary Table 1 Protein constructs used in this study.** Residues belonging to then-terminal His-tag are in italics. Underlined residues were removed by cleavage with the TEV protease for NMR experiments.

Plasmid	Product
pET151-Bamb_5917 PCP domain	<i>MHHHHHHGKPIP</i> <u><i>NLLGLDSTENLYFQGIDPFTGAAAGVSAAGIEPDLT</i></u> <i>AIWQALFALPAVGRHQDF</i> <u><i>FALGGDSQLGLRMLAQLRERHGVDLPLRC</i></u> <i>LYEAPTVARLAETIVRLAAPAPSGDQDDASEYEEGVIR</i>
pET151-Bamb_5915	<i>MHHHHHHGKPIP</i> <u><i>NLLGLDSTENLYFQGIDPFTMTIPALLSELQARGITL</i></u> <i>SLADGELSFRA</i> <u><i>PKGALTPADRATLSARREAIVAYLAAKAARRTDPVTI</i></u> <i>TPSAELRPSLLQELW</i> <u><i>WHWYGLPPRQLNQERLPLVKLFPVGTAGRVAE</i></u> <i>ALRAIVARHHTLRSS</i> <u><i>FHEEDGRLTVTLNEAAALPIEFVEADGTLPREEL</i></u> <i>EPALKAQAAEYAARQL</i> <u><i>PLDGQWLLRARVVS LAPDQSLLLCVFHIIIV</i></u> <i>DAASLLLILAELDARLADPPRAL</i> <u><i>PAAAQFLDYAA WERAWMADPARQ</i></u> <i>PLIDYWARRFRALPEL</i> <u><i>VGPLTGRSLAWQPGSKVDHRFVIPAAQLRRM</i></u> <i>QAAATRLQTSLFSALLSAFGVALARW</i> <u><i>SGSERVPVRCVGD LRTSPELA</i></u> <i>NLVGYLVCSDVIEIHAPAKADFVSILKASEIESHSAMMLRVPTLMRHP</i> <i>LHRGGSGIEDPRGIAATINMFSVRIPGAGAPLDERADPPWPPQLTRSAG</i> <i>EPWPIPLPSIYLR</i> <u><i>LIDYGH</i></u> <u><i>ALEGSLELNDTLLTAAEQAA</i></u> <u><i>LIEALFDALDR</i></u> <i>FLLQAAPAAAPLTTEVL</i>
pET151-Bamb_5915- $\beta$ HD domain	<i>MHHHHHHGKPIP</i> <u><i>NLLGLDSTENLYFQGIDPFTMTIPALLSELQARGITL</i></u> <i>SLADGELSFRA</i> <u><i>PKGALTPADRATLSARREAIVAYLAAKAARRTDPVTI</i></u> <i>TP</i>
pET151-Sven_0512 PCP domain	<i>MHHHHHHGKPIP</i> <u><i>NLLGLDSTENLYFQGIDPFTAEGVDDPPRGPVEEAV</i></u> <i>AELWARVLGVPVTARGAGFFALGGDSL</i> <u><i>TATRLVAAIRGRLGVELPMR</i></u> <i>DVLRAPT</i> <u><i>VAGMSALVGELRGASGSGTDRGPGGEARAGAPGDPSDAL</i></u> <i>VEYEEGEL</i>
pET151-Sven_0517 $\beta$ HD domain	<i>MHHHHHHGKPIP</i> <u><i>NLLGLDSTENLYFQGIDPFTMNVGEFVAELNAAGV</i></u> <i>VLYEEEGKLR</i> <u><i>YRAPRGVMTPERLDALRTHRDAVLALLHAE EYGPATE</i></u> <i>AD</i>

### *Expression and purification of proteins*

*E. coli* BL21 Star (DE3) or C43 (DE3) cells were transformed with the expression constructs and grown in LB medium supplemented with ampicillin (100 µg/mL) at 37 °C with shaking at 180 rpm until an optical density at 600 nm of 0.6 was reached. Protein expression was then induced by addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1.0 mM and the cultures were incubated overnight at 18 °C with shaking at 180 rpm.

For purification of the His<sub>6</sub>-tagged proteins, the *E. coli* cells were harvested by centrifugation (5000 x g, 15 min, 4 °C) and cell pellets were resuspended in 20 mL of lysis buffer 1 (20 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10% glycerol) for Bamb\_5915, buffer 2 (20 mM Tris-HCl (pH 8.5), 300 mM NaCl, 20 mM imidazole) for the Bamb\_5917 or Sven\_0512 PCP domains, or buffer 3 (50 mM potassium phosphate, 150 mM NaCl, pH 7.4) for the Bamb\_5915 and Sven\_0517 βHD domains. The cells were lysed by addition of lysozyme (1 mg/mL) and sonication before centrifugation to clear the lysate (15000 x g, 45 min, 4 °C). The His<sub>6</sub>-tagged proteins were then purified using an imidazole gradient on a 5 mL HiTrap<sup>®</sup> Chelating Column (GE Healthcare) on an AKTAPure. Fractions containing the protein were concentrated and where necessary further purified on a HiLoad Superdex pg75 (GE Healthcare) gel filtration column (running buffer for Bamb\_5915: 20 mM Tris (pH 8.0), 150 mM NaCl, 10% glycerol, and for the Bamb\_5917 PCP domain: 50 mM potassium phosphate (pH 7.4, 150 mM NaCl)). Protein concentrations were determined by measuring the absorbance at 280 nm on a NanoDrop spectrophotometer (Thermo Scientific) and using calculated extinction coefficients: 8480 M<sup>-1</sup>cm<sup>-1</sup> (Bamb\_5917 PCP domain), 74035 M<sup>-1</sup>cm<sup>-1</sup> (Bamb\_5915), 2980 M<sup>-1</sup>cm<sup>-1</sup> (Bamb\_5915-βHD domain), 5960 M<sup>-1</sup>cm<sup>-1</sup> (Sven\_0517-βHD domain), 8480 M<sup>-1</sup>cm<sup>-1</sup> (Sven\_0512 PCP domain). Protein fractions were concentrated, aliquoted, flash frozen and stored at -80 °C until used.

### *Isotopic labelling and NMR sample preparation*

[U-<sup>13</sup>C,<sup>15</sup>N]-labelled Bamb\_5917 PCP domain was obtained from *E. coli* BL21(DE3) cells that were first grown in 2 L LB at 37 °C to an optical density >1.0. The cells were pelleted by centrifugation at 4000 x g for 15 min and washed with phosphate-buffered saline. After centrifuging again, the cells were re-suspended in 1 L M9 medium supplemented containing 1 g of <sup>15</sup>NH<sub>4</sub>Cl and 2 g of [U-<sup>13</sup>C]-glucose. After incubation for 2 h at 37 °C, expression was induced with 1 mM IPTG and the culture was incubated at 18 °C overnight. [U-<sup>15</sup>N]-labelled Bamb\_5917 PCP domain was produced analogously with natural abundance glucose instead of [U-<sup>13</sup>C]-glucose in 1 L M9, and [U-<sup>2</sup>D,<sup>13</sup>C,<sup>15</sup>N]-labelled Bamb\_5917 PCP domain was produced in M9 prepared with D<sub>2</sub>O. Subsequent purification steps were conducted as described, except after gel filtration the protein was buffer exchanged into ddH<sub>2</sub>O using a PD10 desalting column (GE Healthcare) before freeze drying. The lyophilized powder was stored at -20 °C until used.

Isotope labelled PCP samples for assignment and dynamics studies by NMR spectroscopy were prepared by dissolving the freeze dried [U-<sup>15</sup>N,<sup>13</sup>C] or [U-<sup>15</sup>N]-labelled Bamb\_5917 PCP domains in 50 mM potassium phosphate, 200 mM NaCl, pH 6.5, 10% D<sub>2</sub>O to obtain a protein concentration of 450 µM. For interaction studies, the [U-<sup>15</sup>N]-Bamb\_5917 PCP domain was dissolved in 50 mM potassium phosphate, 150 mM NaCl, pH 7.4, 10% D<sub>2</sub>O to obtain 80 µM or 100 µM samples. To remove the ~3.5 kDa His-tag, the Bamb\_5917 PCP domain was incubated with a five-fold excess of TEV-GFP<sup>31</sup> at 37 °C for 2 h. TEV-GFP was removed by Ni-affinity chromatography and the Bamb\_5917 PCP domain was buffer exchanged using a PD10 column before concentrating to 200 µL total volume.

### *In vitro conversion of the apo-Bamb\_5917 PCP domain to its holo form*

The Bamb\_5917 PCP domain was converted to its *holo* form by using the promiscuous phosphopantetheinyl transferase (PPTase) Sfp from *B. subtilis*<sup>32</sup>. The *apo*-PCP domain (400 µM) was combined with MgCl<sub>2</sub> (10 mM), Sfp (10 µM) and CoA (1 mM) in a total volume of 500 µL. Reactions were carried out for two hours at room temperature in buffer containing 50 mM potassium phosphate (pH 7.4) and 150 mM NaCl. For acetyl transfer assays, excess acetyl-CoA was removed using Amicon

Ultra centrifugal filters with a 5 kDa molecular weight cut off (MWCO) membrane (Millipore). Successful phosphopantetheinylation was verified by UHPLC-ESI-Q-TOF-MS analysis.

#### *Bamb\_5915 purification and sample preparation for crystallisation*

Bamb\_5915 was overproduced as described above, but using *E.coli* BL21(DE3) cells in the presence of 50 µg/mL carbenicillin. The cleared cell lysate was incubated with 5 mL of nickel NTA resin (Bio-Rad) pre-equilibrated with lysis buffer (50 mM Tris·HCl pH 8.0, 20 mM imidazole, 300 mM NaCl and 10% glycerol) and gently stirred for 1 hour at 4 °C. The cell lysate was then transferred to a fritted gravity flow column and Bamb\_5915 was eluted by adding lysis buffer containing increasing quantities of imidazole. Individual fractions were analyzed by SDS PAGE for purity prior to being collected and dialyzed against Q column buffer (50 mM Tris·HCl pH 8.0 and 1 mM DTT) at 4 °C for 16 hours. Post dialysis, Bamb\_5915 was concentrated to a final volume of 5 mL using a 30,000 Da MWCO centrifugal filter (Millipore) and passed through a High Trap Q column (GE Healthcare). Bamb\_5915 was eluted using a linear gradient of 0-1 M NaCl in Q column buffer and dialyzed against 50 mM Tris·HCl pH 8.0 and 1 mM DTT. Post Q column treatment, Bamb\_5915 was concentrated to 5-10 mg/mL and was injected onto a Superdex 200 column pre-equilibrated with 50 mM Tris·HCl pH 8.0 and 1 mM DTT. Protein concentrations were determined by Bradford assay (Bio-Rad). Fractions containing the highest purity (≥90 %) were pooled, concentrated to 10 mg/mL and flash frozen in liquid nitrogen and stored at -80 °C.

#### *Overproduction and purification of SeMet-labelled Bamb\_5915*

To obtain SeMet-labelled Bamb\_5915, LB cultures supplemented with 50 µg/mL carbenicillin (4 x 1 L) were grown at 37 °C and 180 rpm until  $A_{600}$  reached 0.6. The cells were pelleted (5500 x g for 10 minutes at 4 °C) and washed three times by suspension in 50 mL M9 medium, before being evenly transferred to (4 x 1 L) flasks containing M9 minimal medium supplemented with 50 µg/L carbenicillin and the following additives: lysine, phenylalanine, and threonine (100 mg each); valine, isoleucine, and leucine (50 mg each); and L-selenomethionine (80 mg). The resulting cultures were grown overnight for 16-18 hours at 18 °C after addition of 0.4 mM IPTG. Cells were collected via centrifugation (5500 x g for 12 minutes), resuspended in 50 mM Tris·HCl pH 8.0, pelleted and stored at -80 °C. The purification of SeMet-labelled Bamb\_5915 followed the procedure described for unlabelled Bamb\_5915. The incorporation of selenium into Bamb\_5915 was verified using MALDI-TOF mass spectrometry.

#### *Crystallization of Bamb\_5915*

Frozen aliquots of either unlabelled or SeMet-labelled Bamb\_5915 were thawed on ice prior to setting up crystallisation screens. Crystals were obtained using the hanging drop vapor diffusion method by mixing 2 µL of native Bamb\_5915 at 10 mg/mL with 2 µL of 0.2 M di-ammonium phosphate and 20% PEG 3,350 at 25 °C. Large rectangular crystals of unlabelled Bamb\_5915 grew between 7 and 14 days. SeMet-labelled Bamb\_5915 crystals were obtained by mixing 2 µL of SeMet-labelled protein at 7-8 mg/mL with 0.17-0.2 M di-ammonium phosphate and 22-25 % PEG 3,350 at 25 °C. SeMet-labelled Bamb\_5915 crystals grew after a month in the dark. Both SeMet-labelled and unlabelled Bamb\_5915 crystals were harvested, washed in well solution (20-25% PEG 3,350 and 0.17-0.2 M di-ammonium phosphate), and flash frozen in liquid nitrogen prior to data collection.

#### *X-ray Data collection, processing and structure determination*

Datasets for SeMet-labelled crystals of Bamb\_5915 were collected at the SSRL beamline 12-2. The MAD data were collected at 0.9157, 0.9793 and 0.9796 Å wavelengths corresponding to the remote, peak and inflection, respectively. For MAD data collection, two individual data sets for peak, remote and inflection were collected, consisting of frames 1-138 (Set 1) and 139-360 (Set 2). The individual

data sets were processed in the P21 space group using Mosflm and scaled together using Aimless (CCP4 suites)<sup>16,17</sup>. Phasing was accomplished using PHENIX Autosol<sup>18</sup>. A preliminary model was constructed using Autobuild. This was further improved through repetitive rounds of refinement and model rebuilding using PHENIX Refine and COOT<sup>19,20</sup>. Data for unlabelled Bamb\_5915 were collected at the ALS beamline 8.2.1 and were refined to 2.0 Å. The data were processed and scaled using Mosflm<sup>16</sup>. The higher resolution structure of the unlabelled protein was determined using molecular replacement (Phaser) using the lower resolution model of the SeMet-labelled Bamb\_5915 as a search model<sup>21</sup>. The unlabelled Bamb\_5915 structure was refined using a combination of PHENIX Refine and COOT<sup>19,20</sup>.

**Supplementary Table 2 SeMet-labelled and unlabelled Bamb\_5915 crystallographic statistics**

	<b>Native Bamb_5915 PDB:6CGO</b>	<b>SeMet Bamb_5915</b>		
<b>Data collection</b>				
Space group	P 1 21 1	P 1 21 1		
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	103.63, 58.77, 107	103.71, 58.9484, 107.741		
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 104.96, 90	90, 105.304, 90		
		<i>Peak</i>	<i>Inflection</i>	<i>Remote</i>
Wavelength (Å)	0.9999	0.9793	0.9796	0.9157
Resolution (Å)	64.12-2.0 (2.11-2.0)	83.99-2.65 (2.77-2.65)	83.99-2.65 (2.77-2.65)	83.99-2.65 (2.77-2.65)
<i>R</i> <sub>merge</sub>	0.107 (0.332)	0.135 (0.476)	0.153 (0.507)	0.119 (0.352)
<i>I</i> / $\sigma$ <i>I</i>	6.8 (2.9)	9.6 (3.9)	8.2 (3.3)	10.00 (4.1)
Completeness (%)	99.9 (100)	95.7 (96.2)	95.9 (97.0)	95.1 (93.8)
Redundancy	3.6 (3.7)	5.9 (5.8)	6.0 (5.9)	5.6 (5.3)
<i>CC1/2</i>	0.986 (0.878)	0.988(0.839)	0.985 (0.820)	0.990 (0.893)
<b>Refinement</b>				
Resolution (Å)	51.70-2.0 (2.05-2.00)	52.07-2.65 (2.745-2.65)		
No. reflections	84036 (8402)	35048 (3425)		
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.1803/0.2150 (0.2093/0.2715)	0.1975/0.2406 (0.2495/0.3193)		
No. atoms				
Protein	7955	7943		
Phosphate ions	10	10		
Water	1068	539		
<i>B</i> -factors				
Protein	32.96	29.70		
Phosphate ions	25.23	36.55		
Water	37.27	28.62		
R.m.s deviations				
Bond lengths (Å)	0.012	0.012		
Bond angles (°)	1.45	1.60		

#### *NMR spectroscopy experiments and structure calculation*

All solution NMR experiments used for resonance assignments of the *apo*- and *holo*-Bamb\_5917 PCP domain and for structure calculations of the *apo*-Bamb\_5917 PCP domain were recorded on Bruker AVANCE II 700 spectrometer, equipped with a TCI cryoprobe, on 0.6 mM <sup>13</sup>C, <sup>15</sup>N-labelled or <sup>15</sup>N-labelled samples in 50 mM phosphate buffer, pH 7.4, containing 10% (v/v) D<sub>2</sub>O, 150 mM NaCl and DSS (for internal referencing) at 288 K. For the assignment of the backbone and side-chain resonances of the *apo*-Bamb\_5917 PCP domain 2-D <sup>15</sup>N-HSQC, 2-D <sup>1</sup>H-<sup>1</sup>H NOESY, 3-D BEST-HNCO, BEST-CBCAcoNH, BEST-HNCACB, BEST-HNCA, BEST-HNcoCA, HBHAcoNH, hCCH-TOCSY, NOESY-<sup>15</sup>N-HSQC, and NOESY-<sup>13</sup>C-HSQC data were collected<sup>33-41</sup>. For the assignment of the backbone of the *holo*-Bamb\_5917 PCP domain, a series of 2-D <sup>15</sup>N-HSQC were collected during the course of the phosphopantetheinylation. Topspin 3.2 was used to process NMR data and CARA and CCPN were used for data analysis and sequential assignment. The overall completeness of chemical shift assignments for the *apo*- and *holo*-Bamb\_5917 PCP domain was 81%.

Backbone dihedral angle ( $\phi$  and  $\psi$ ) constraints were derived from the chemical shift values for the backbone atoms of the *apo*-Bamb\_5917 PCP domain using TALOS<sup>+</sup><sup>42</sup>. The peaks in 2-D NOESY, 3-D NOESY-<sup>15</sup>N and 3-D NOESY-<sup>13</sup>C spectra of the *apo*-Bamb\_5917 PCP domain were picked automatically using UNIO<sup>43-45</sup>, incorporating ATNOS/CANDID, and were used to generate distance constraints and calculate initial structural models using CYANA-2.1<sup>46</sup>. The programs ATNOS/CANDID/CYANA-2.1 use the amino acid sequence, the chemical shift lists, three NOESY spectra and  $\phi$  and  $\psi$  dihedral angle constraints as inputs. After an initial structure calculation, all assignments of the three NOESY spectra were manually inspected to pick peaks and modify assignments, where necessary.

The structure calculations were carried out using 625 distance constraints, 112 dihedral angles and 28 hydrogen bonds. The 20 conformers of the *apo*-Bamb\_5917 PCP domain with the lowest target function values were subjected to restrained energy minimization and molecular dynamics simulations in explicit solvent with AMBER 14.0<sup>47</sup>. NOEs and torsion angle constraints were applied with force constants of 50 kcal mol<sup>-1</sup>Å<sup>-2</sup> and 32 kcal mol<sup>-1</sup> rad<sup>-2</sup>, respectively. The quality of the structures was then evaluated using the Protein Structure Validation Software suite<sup>48-53</sup> and validated using the PDB validation server. The Ramachandran plot for selected residues with regular secondary structure (13-25,33-80) from Procheck shows that 93.1% of residues are in the most favoured regions, 6.9% are in the additionally allowed regions and no residues are in the generously allowed and disallowed regions. Further analysis of the Ramachandran plot using the Richardson Lab's Molprobity<sup>54</sup> shows that 97.8% of residues are in the most favoured regions, while 2.2% are in the allowed regions and no residues are in the disallowed regions.

**Supplementary Table 3** NMR and refinement statistics for the *apo*-Bamb\_5917 PCP domain

<i>Apo</i> -Bamb_5917 PCP domain	
<b>NMR distance and dihedral constraints</b>	
Distance constraints	
Total NOE	<b>603</b>
Intra-residue	130
Inter-residue	473
Sequential ( $ i-j  = 1$ )	212
Medium-range ( $ i-j  < 4$ )	179
Long-range ( $ i-j  > 5$ )	82
Intermolecular	0
Hydrogen bonds	28
Total dihedral angle restraints	<b>112</b>
Phi	66
Psi	66
<b>Structure statistics</b>	
Violations (mean and s.d.)	0.00201 ± 0.00879
Distance constraints (Å)	0.01 Å
Dihedral angle constraints (°)	0.43°
Max. dihedral angle violation (°)	4.90°
Max. distance constraint violation (Å)	0.33 Å
Deviations from idealized geometry	
Bond lengths (Å)	0.014 Å
Bond angles (°)	2.1°
Improper (°)	4.90°
Average pairwise r.m.s.d.* ** (Å) (13-25,33-80)	
Backbone	0.79 ± 0.15
Heavy	1.32 ± 0.22

\*Pairwise r.m.s.d. was calculated among 20 refined structures

The coordinates for the *apo*-Bamb\_5917 PCP domain have been deposited in the PDB (accession no.: 5MTI). Resonance assignments for the *apo*- and *holo*-Bamb\_5917 PCP domain have been deposited in

BMRB<sup>55</sup> (accession codes 34085 and 27304, respectively).

Dali structural alignments show that the Bamb\_5917 PCP domain is the most similar to PCP domains in the TycC5-6 PCP-C didomain and the LgrA chain initiation module but with  $\alpha$ -helix 1 shorter by one turn and slightly different positioning of  $\alpha$ -helix 3 and loop 1 (A238-D254) (Supplementary Fig. 12). Overall, the secondary chemical shifts, lack of medium- and long-range NOEs and relaxation data correspond well with IUpred predictions of a disordered C-terminus (Supplementary Figs 13 and 14).

### Solution NMR dynamics

[U-<sup>15</sup>N]-labelled Bamb\_5917 PCP domain samples for dynamics studies by NMR spectroscopy were prepared by dissolving the freeze-dried [U-<sup>15</sup>N]-Bamb\_5917 PCP domain in 50 mM potassium phosphate, 200 mM NaCl, pH 6.5, 10% D<sub>2</sub>O to obtain a protein concentration of 450  $\mu$ M.

For dynamics measurements of the [U-<sup>15</sup>N]-Bamb\_5917 PCP domain in complex with Bamb\_5915, 70  $\mu$ M Bamb\_5915 was added to 370  $\mu$ M [U-<sup>15</sup>N]-Bamb\_5917 PCP domain in measurement buffer. All spectra were acquired in a 3 mm tube at 280K on a Bruker Avance II 700 MHz spectrometer equipped with a TCI cryoprobe. TopSpin 3.2 was used to process all spectra. Dynamics and titration data was analyzed with Ccpnmr. <sup>15</sup>N longitudinal ( $R_1$ ), <sup>15</sup>N transverse ( $R_2$ ) relaxation and <sup>15</sup>N-<sup>1</sup>H heteronuclear NOE data were obtained using standard methods and fitted using the Ccpnmr rate analysis tool and implemented in the software error analysis. Relaxation delays from 0 to 2000 ms and 317 ms were employed for  $R_1$  and  $R_2$  measurements (see Table SI 4) of the *apo*- and *holo*-Bamb\_5917 PCP domain. The <sup>1</sup>H-<sup>15</sup>N heteronuclear NOEs were obtained for <sup>1</sup>H-saturated and unsaturated spectra (5 s saturation time).

Supplementary Table 4 Parameters used in NMR <sup>15</sup>N dynamics measurements

	Experiment	SW and Aq		# of scans	Inter-scan delays (s)	Size of FID <sup>1</sup> H	Size of FID <sup>15</sup> N	Time Points
<i>apo</i> Bamb_5917 PCP	<sup>15</sup> N $R_1$	<sup>1</sup> H 8400 Hz 121.6 ms	<sup>15</sup> N 2500 Hz 50.1 ms	16	4	2048	256	0.025; 0.05; 0.10; 0.20; 0.30; 0.40; 0.50; 0.80; 1.1; 2.0 (s)
	{ <sup>1</sup> H}- <sup>15</sup> N NOE	<sup>1</sup> H 8400 Hz 121.6 ms	<sup>15</sup> N 2500 Hz 100.2 ms	32	5	2048	512	-
	<sup>15</sup> N $R_2$	<sup>1</sup> H 8400 Hz 121.6 ms	<sup>15</sup> N 2500 Hz 50.1 ms	16	2.5	2048	256	16.96; 33.92; 50.88; 67.84; 84.80; 101.76; 135.68; 169.6; 203.52; 254.4; 339.2 (ms)
<i>holo</i> Bamb_5917 PCP	<sup>15</sup> N $R_1$	<sup>1</sup> H 8400 Hz 121.6 ms	<sup>15</sup> N 2500 Hz 50.1 ms	16	4	2048	256	0.025; 0.05; 0.075; 0.10 0.20; 0.30; 0.40; 0.50; 0.80; 1.1; 2.0 (s)
	{ <sup>1</sup> H}- <sup>15</sup> N NOE	<sup>1</sup> H 8400 Hz 121.6 ms	<sup>15</sup> N 2500 Hz 100.2 ms	32	5	2048	512	-
	<sup>15</sup> N $R_2$	<sup>1</sup> H 8400 Hz 121.6 ms	<sup>15</sup> N 2500 Hz 50.1 ms	16	2.5	2048	256	16.96; 33.92; 50.88; 67.84; 84.80; 101.76; 135.68; 169.6; 254.4; 339.2 (ms)

Correlation times of  $12.0 \pm 0.2$  ns for the *apo*-Bamb\_5917 PCP domain and  $11.4 \pm 0.2$  ns for the *holo*-Bamb\_5917 PCP domain were estimated from  $R_2/R_1$  ratios of structured residues. These values compare well with the average correlation time calculated for the 20 lowest energy structures of the *apo*-

Bamb\_5917 PCP domain using HYDRONMR<sup>56</sup> (12.8 ns), and are consistent with a monomeric state of the protein, consisting of a compact folded core with disordered termini (the effective correlation time is longer than for a compact globular protein of this size).

### *Solution NMR titration experiments*

In the NMR titrations<sup>57</sup> of the Bamb\_5917 holo-PCP domain with both full-length Bamb\_5915 and its excised  $\beta$ HD domain the system is in slow to intermediate exchange<sup>57</sup>: several signals gradually disappear and new signals (from the bound Bamb\_5917) appear as the concentrations of the titrants are increased (see Supplementary Fig. 16).

For titration studies, the freeze dried [U-<sup>15</sup>N]-*apo*-Bamb\_5917 PCP domain was dissolved in buffer (50 mM potassium phosphate (pH 7.4) and 150 mM NaCl) and treated with GFP-TEV protease<sup>31</sup> to remove the His<sub>6</sub>-tag as described above. After a reverse Ni affinity purification step, the removal of the tag was confirmed by UHPLC-ESI-Q-TOF-MS analysis. To obtain the [U-<sup>15</sup>N]-*holo*-Bamb\_5917 PCP domain, the sample was treated with Sfp as described above. The resulting tag-free protein was concentrated in 5000 Da MWCO spin filters to a final concentration of 80 or 130  $\mu$ M.

All titration spectra were acquired in a 3 mm tube at 288K on a Bruker Avance II 700 MHz spectrometer equipped with a TCI cryoprobe. Bamb\_5915 or the excised Bamb\_5915  $\beta$ HD domain were added stepwise (see Supplementary Table 5 for concentrations) and 2-D <sup>15</sup>N-HSQC or <sup>1</sup>H-<sup>15</sup>N BEST-TROSY-HSQC spectra were acquired at each titration step. The data were processed and analysed using TopSpin 3.5 and Ccpnmr and fitted using Matlab with in-house scripts.

**Supplementary Table 5 Samples used in NMR titration experiments**

<sup>15</sup> N labelled protein	Natural abundance titrant concentrations
<i>holo</i> -Bamb_5917 PCP domain; c=80 $\mu$ M	<b>Bamb_5915</b> c=0; 2.7; 5.5; 8.2; 13.6; 19.1; 24.5; 35.0; 45.8; 61.4; 82.0 ( $\mu$ M)
<i>holo</i> -Bamb_5917 PCP domain; c=130 $\mu$ M	<b>Bamb_5915 <math>\beta</math>HD domain</b> c=0; 17.4; 26.0; 34.6; 43.2; 51.7; 60.2; 77.0; 93.7; 126.5; 158.7; 190.3; 221.3; 281.6; 423.1
<i>holo</i> -Bamb_5917 PCP domain with His <sub>6</sub> tag; c=180 $\mu$ M	<b>Bamb_5915</b> c=0; 100.0; 127.3; 140.0; 152.2; 175.0 ( $\mu$ M)
<i>apo</i> -Bamb_5917 PCP domain; c=80 $\mu$ M	<b>Bamb_5915</b> c=0; 2.7; 5.5; 8.2; 13.6; 19.1; 24.5; 35.0; 45.8; 61.4; 82.0 ( $\mu$ M)
<i>holo</i> -Bamb_5917 PCP domain; c=125 $\mu$ M	<b>Sven_0517 <math>\beta</math>HD domain</b> c=0; 2.5; 5.0; 9.9; 19.6; 38.5; 56.6; 111.1; 159.7; 239.5; 339.9 ( $\mu$ M)

### *NMR titration data fitting*

Changes in signal intensities as a function of titrant concentrations in solution NMR titrations of the Bamb\_5917 PCP domain were fitted in Matlab to obtain apparent dissociation constants ( $K_d$ ) for each residue. Note that both binding and conformational changes upon binding can contribute to this parameter so the values obtained are only approximate (even though they appear to correspond reasonably well to the bulk measurements of  $K_d$  from BLI).

**Equation 1** defines the dissociation of the Bamb\_5917 PCP domain (P) and Bamb\_5915 (L) from Bamb\_5917 PCP domain:Bamb\_5915 complex (PL). This is a valid description because the affinity of

each residue in the PCP domain for the C domain is examined individually. The sum of the concentrations of the bound and unbound states gives the total concentrations of L and P (**Equation 2**). By combining **Equations 1 and 2**, and including a multiplier constant, A, which takes into account error in the integral of the first data point, **Equation 3** can be used to determine an approximate value of the apparent dissociation constant ( $K_d$ ) for each residue of the Bamb\_5917 PCP domain in the Bamb\_5917 PCPdomain:Bamb\_5915 complex. The decay in integral value for each signal as a function of Bamb\_5915/Bamb\_5915  $\beta$ HD domain concentration was fitted to **Equation 3**. The integral as a function of the Bamb\_5915  $\beta$ HD domain concentration for the peaks appearing in the titration of the Bamb\_5917 PCP domain with the Bamb\_5915  $\beta$ HD domain was fitted to **Equation 4**. Errors were calculated using Monte Carlo error propagation with the error bars set to two standard deviations over 1000 iterations. To identify the residues, which are the most affected by binding a threshold was defined as the average apparent  $K_a$  ( $K_a=1/ K_d$ ) plus one standard deviation. To obtain the average and standard deviations used to define the threshold the residues with values of  $K_a$  greater than the average plus three standard deviations were iteratively removed from the calculation.

$$\text{Equation 1} \quad K_d = \frac{1}{K_a} = \frac{[P][L]}{[PL]}$$

$$\text{Equation 2} \quad [L]_t = [L] + [PL] \text{ and } [P]_t = [P] + [PL]$$

$$\text{Equation 3} \quad [P] = \frac{\left(-\frac{1}{K_a} - [L]_t + [P]_t + \sqrt{\frac{4}{K_a}[P]_t + \left(\frac{1}{K_a} + [L]_t - [P]_t\right)^2}\right)}{2[P]_t}$$

$$\text{Equation 4} \quad [PL] = \frac{\left(\frac{1}{K_a} + [L]_t + [P]_t - \sqrt{-4[L]_t[P]_t + \left(-\frac{1}{K_a} - [L]_t - [P]_t\right)^2}\right)}{2[P]_t}$$

#### *<sup>15</sup>N CEST NMR Spectroscopy on Bamb\_5917 PCP domain:10% Bamb\_5915 complex*

Chemical Exchange Saturation Transfer (CEST)<sup>58</sup> NMR experiments on 150  $\mu$ M [<sup>15</sup>N]<sub>apo</sub>-Bamb\_5917 PCP domain (50 mM potassium phosphate, 150 mM NaCl, 10% D<sub>2</sub>O; pH 7.4) in the presence of 15  $\mu$ M Bamb\_5915 were performed at 15°C on a 700 MHz AVANCE II Bruker spectrometer equipped with a TCI cryoprobe. The field strength for the CW pulses, applied during the CEST element, generated an effective CEST field strength of 25 Hz. Pseudo-3-D CEST data sets containing 70 2-D planes were generated. This pseudo-3-D CEST comprised a series of 2-D spectra with <sup>15</sup>N offsets ranging between 102 and 136 ppm obtained in increments of 0.5 ppm for  $\nu_1$  fields of 25 Hz. The CEST period,  $T_{\text{ex}}$ , was set to 0.4 s; each 2-D plane was recorded with 32 transients per FID. The recycle delay was 1 s, and 2048 / 124 complex points in  $t_2$  /  $t_1$ , respectively, were recorded.<sup>41</sup>

We observed the presence of a minor state, due to the bound form, in profiles for several residues in the SLiM. We selected a subset of good quality profiles displaying a minor state in CEST (residues E310, G312, V313, and R315, Supplementary Figure 19) and analysed them jointly using a two-state exchange model in ChemEx<sup>30</sup>.  $k_{\text{ex}}$  and  $p_{\text{bound}}$  were fitted as global parameters, while the chemical shift difference,  $\Delta\omega$ , between the free and bound Bamb\_5917 PCP domain,  $R_{2\text{free}}$  and  $R_{2\text{bound}}$ , were fitted on a per-residue basis. The best fit value for population of bound the Bamb\_5917 PCP domain,  $p_{\text{bound}}$ , was 8.7%, which is in a good agreement with the expected values according to the amount of complex present in solution. The global exchange rate obtained was  $0.58 \pm 0.02 \text{ s}^{-1}$ .

The chemical shift values for residues in the bound form of the Bamb\_5917 PCP domain, wherever a minor state was observed, provided the basis for assignment of the peaks appearing in the titration of Bamb\_5917 PCP domain with the excised  $\beta$ HD domain.

### *<sup>15</sup>N CEST data analysis*

NMR spectra were processed with Topspin 3.5pl7 and analyzed using CARA.<sup>25</sup> Peak intensities were extracted in CARA. CEST profiles were analyzed with ChemEx<sup>41</sup>. Fits of <sup>15</sup>N CEST profiles were performed as for the regular analysis of <sup>15</sup>N CEST data sets<sup>41</sup>.

The <sup>15</sup>N chemical shift of the bound form of the Bamb\_5917 PCP domain obtained from CEST profiles were used to assign peaks appearing in the 2D <sup>1</sup>H-<sup>15</sup>N correlation spectra in the titrations of the Bamb\_5917 PCP domain with the βHD domain.

### *Sample preparation for solid state NMR analysis of the Bamb\_5917 PCP domain:Bamb\_5915 complex*

[U-<sup>2</sup>D,<sup>13</sup>C,<sup>15</sup>N]-*holo*-Bamb\_5917 PCP domain was prepared as detailed above for the [U-<sup>15</sup>N]-*holo*-PCP domain at a concentration of 100 μM in 450 μL of 50 mM potassium phosphate, 200 mM NaCl, pH 6.5. After addition of 70 μL of Bamb\_5915 (c = 500 μM in 20 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10% glycerol), the sample was incubated for 30 min at room temperature. The complex was centrifuged for 2 min at 10 000 x g to remove any precipitated protein before sedimenting the sample at 700 000 x g for 20 h (Beckmann Optima Max-XP). After sedimentation the supernatant was removed and the gel-like protein complex was packed manually in 1.3 mm or 0.7 mm rotors (Bruker). As an internal standard, 0.5 μL of supernatant containing 2 % (w/v) DSS were added to the sample. [U-<sup>1</sup>H,<sup>13</sup>C,<sup>15</sup>N]-Bamb\_5917 PCP domain samples and *apo*-Bamb\_5917 PCP domain samples were prepared analogously.

### *Solid-state NMR spectroscopy*

All solid-state NMR experiments used for backbone resonance assignment of the protein complex between the *holo*-Bamb\_5917 PCP domain and Bamb\_5915 were recorded on a Bruker Ascend 700 WB (wide bore) spectrometer equipped with AVANCE III HD console and using a 0.7 mm Bruker probe in HCN configuration at 100 kHz spinning frequency MAS (Supplementary Table 6). The experiments were performed on [U-<sup>13</sup>C,<sup>15</sup>N]-labelled or [U-<sup>2</sup>H,<sup>13</sup>C,<sup>15</sup>N]-labelled *holo*-Bamb\_5917 PCP domain in sedimented complex with natural abundance Bamb\_5915 (50 mM phosphate buffer, pH 7.4, 150 mM NaCl and DSS). Sample temperatures were about 300K (as indicated by the chemical shift of water) with active cooling using a Bruker BCU-X cooling unit. Topspin 3.5 was used to process NMR data and CARA for data analysis and resonances assignment.

For experiment carried out at 100 kHz MAS frequency, the hard 90° pulses were set to 2.0 μs at 125 kHz nutation frequency for <sup>1</sup>H, 3 μs at 83.3 kHz nutation frequency for <sup>15</sup>N, and 2.5 μs at 100 kHz nutation frequency for <sup>13</sup>C. Double quantum (DQ) <sup>1</sup>H-<sup>13</sup>C cross-polarization (CP) was achieved using a contact time of 0.2 ms during which the <sup>1</sup>H rf-field amplitude was ramped linearly from 85 to 100% corresponding to ~90 kHz, and a constant <sup>13</sup>C rf-field amplitude at ~10 kHz. DQ <sup>1</sup>H-<sup>15</sup>N CP was achieved using a contact time of 0.6 ms (<sup>1</sup>H→<sup>15</sup>N) and 0.4 ms (<sup>15</sup>N→<sup>1</sup>H), during which the <sup>1</sup>H rf-field amplitude was ramped linearly from 85 to 100% corresponding to ~90 kHz, and a constant <sup>15</sup>N rf-field amplitude of ~10 kHz. DQ <sup>13</sup>C-<sup>15</sup>N CP was accomplished using a contact time of 10 ms, during which the <sup>13</sup>C rf-field amplitude was held constant at approximately 40 kHz, and the <sup>15</sup>N rf-field amplitude was ramped linearly from 85 to 100% of approximately nominally 10 kHz. <sup>13</sup>C-<sup>13</sup>C and <sup>1</sup>H-<sup>1</sup>H homonuclear mixing was accomplished using RFDR. The RFDR mixing (1.6 ms for <sup>13</sup>C-<sup>13</sup>C and 0.64 ms for <sup>1</sup>H-<sup>1</sup>H) with 5 μs (<sup>13</sup>C) and 4 μs (<sup>1</sup>H) π pulses. <sup>13</sup>C-<sup>13</sup>C TOCSY mixing (~17.4 ms) was achieved using DIPSI-3 scheme at 25 kHz <sup>13</sup>C nutation frequency. The 10 kHz WALTZ-64 sequence was used for <sup>1</sup>H decoupling during t1 and for <sup>15</sup>N decoupling during <sup>1</sup>H acquisition. The MISSISSIPPI sequence was used to suppress the water signal, with a 12.5 kHz <sup>1</sup>H irradiation for 100 ms. All the spectra were acquired using the States-TPPI procedure. The experiments were acquired with a single 2 s recycle delay.

The <sup>1</sup>H-<sup>13</sup>C HSQC pulse sequence was performed using nonselective pulses. The hard 90° pulses were set to 2.0 μs at 125 kHz <sup>1</sup>H nutation frequency, and 2.5 μs at 100 kHz <sup>13</sup>C nutation frequency. Refocused INEPT delays were set to 1.8 ms and 0.8 ms respectively. The WALTZ-64 sequence was used for <sup>1</sup>H

heteronuclear decoupling during t1 and for  $^{15}\text{N}$  decoupling during  $^1\text{H}$  acquisition using ~10 and ~20 kHz nutation frequencies, respectively.

The *de novo* assignment of the signals in a sedimented 70 kDa complex of *holo*-Bamb\_5917 PCP domain with Bamb\_5915 was challenging due to limited resolution and sensitivity even at 100 kHz magic angle spinning. However, by comparing the signals in 3D NMR spectra to those observed for the Bamb\_5917 PCP domain in solution, it was possible to obtain a tentative set of self-consistent of assignments. To achieve this, we exploited the fact that  $\text{C}^\alpha$  resonances are highly sensitive to local conformation but relatively insensitive to binding. For the residues that remain in the same local conformation in the PCP domain alone and in the complex the chemical shifts should thus be approximately the same (Fig. 2).

**Supplementary Table 6 Solid state NMR experiments**

Experiments	# of Scans	Aq F1 (ms)	Points F1	SW F1 (ppm)	Points F2	SW F2 (ppm)	Points F3	SW F3 (ppm)
$^{13}\text{C}$ - $^1\text{H}$	128	30	668	16	224	80	–	–
$^{15}\text{N}$ - $^1\text{H}$	448	30	668	16	84	42	–	–
$^1\text{H}$ - $^{13}\text{C}$ HSQC	64	40	890	16	422	80	–	–
$^{13}\text{C}\alpha$ - $^{15}\text{N}$ - $^1\text{H}$	224	25	1386	40	72	30	32	36
$^{13}\text{CO}$ - $^{15}\text{N}$ - $^1\text{H}$	512	25	1386	40	24	13	28	35
$^{13}\text{C}$ - $^{13}\text{C}$ - $^1\text{H}$ -TOCSY	64	26	654	18	38	70	38	70
$^{13}\text{C}$ - $^{13}\text{C}$ - $^1\text{H}$ -RFDR	16	30	668	16	112	80	112	80
$^1\text{H}$ - $^1\text{H}$ - $^{15}\text{N}$ - $^1\text{H}$ -RFDR	32	30	1258	30	84	48	66	12

#### *BioLayer Interferometry (BLI) measurements*

The presence of only one lysine residue in the His<sub>6</sub>-tag was exploited for biotinylation of the Bamb\_5917 PCP domain using the EZ-link<sup>TM</sup> Sulfo-NHS-Biotin cross-linking reagent (ThermoFisher Scientific), which reacts selectively with primary amines. After incubation for 30 min at room temperature, non-reacted cross-linking agent was removed on ZEBRA spin desalting columns (7K MWCO, ThermoFisher Scientific) before loading Streptavidin biosensors (Pall ForteBio) with the biotinylated *apo*-Bamb\_5917 PCP domain at a concentration of 70 nM. The binding of Bamb\_5915 in concentrations of 0.1, 0.5, 1.0, 7.25, 12.5, 25.0, and 50.0  $\mu\text{M}$  to the immobilized biotinylated *apo*-Bamb\_5917 PCP domain was monitored on an Octet<sup>®</sup> RED96 System (Pall ForteBio) in triplicate. The sensor responses were fitted using the Octet<sup>®</sup> Software and GraphPad Prism7. Affinity constants were extracted from steady-state fittings using the equation  $B_{\text{max}} * X / (X + K_d)$ , where X corresponds to the Bamb\_5915 concentration and B is the measured response in nm.

#### *Carbene footprinting*

Carbene footprinting entails exposure of the isolated proteins and their corresponding complexes to a highly reactive carbene and subsequent high resolution UHPLC-MS analyses to compare the extent to which proteolytic peptides derived from the two experiments are labelled. Using two proteases with differing cleavage specificities (trypsin and AspN) the sequence coverage for the *holo*-PCP domain of Bamb\_5917 and Bamb\_5915 was 100% and 69%, respectively (Supplementary Figs 24-26; raw data in Supplementary File 2). By calculating the fraction of single carbene labelling for each peptide resulting from the proteolytic digests, the solvent accessibility of residues within each peptide can be inferred.

Photolysis experiments were carried out using Bamb\_5915 (50  $\mu$ M) and/or the Bamb5917 PCP domain (100  $\mu$ M) in 20 mM Tris, 100 mM NaCl, 10 mM aryl diazirine, pH 7.4 (total volume 20  $\mu$ L). The mixture was left to equilibrate for 5 minutes at room temperature before 6  $\mu$ L aliquots were placed in crystal clear vials (Fisher Scientific UK) and snap-frozen in liquid nitrogen. The labelling reaction was initiated by photolysis of the mixture using the third harmonic of a Nd:YLF laser (Spectra Physics, repetition frequency 1000 Hz, pulse energy 125  $\mu$ J) at a wavelength of 349 nm. The frozen samples were irradiated for 16 s. All experiments were performed in triplicate. Following irradiation, samples were reduced (DTT, 10 mM in 10 mM ammonium bicarbonate), alkylated (Iodoacetamide, 55 mM in 10 mM ammonium bicarbonate) and incubated at 37 °C with AspN or trypsin overnight (1:20 protease/protein ratio in 10 mM ammonium bicarbonate). The digestion was quenched by adjusting the pH to below 6 by adding formic acid to 2% final concentration.

The analysis of the digests was carried out on a Bruker MaXis II ESI-Q-TOF-MS connected to a Dionex 3000 RS UHPLC fitted with an ACE C18 RP column (100 x 2.1 mm, 5  $\mu$ m, 30 °C). The column was eluted with a linear gradient of 5–100% MeCN containing 0.1% formic acid over 40 min. The mass spectrometer was operated in positive ion mode with a scan range of 200–3000  $m/z$ . Source conditions were: end plate offset at –500 V; capillary at –4500 V; nebulizer gas ( $N_2$ ) at 1.6 bar; dry gas ( $N_2$ ) at 8 L min<sup>-1</sup>; dry temperature at 180 °C. Ion transfer conditions were: ion funnel RF at 200 Vpp; multiple RF at 200 Vpp; quadrupole low mass at 55  $m/z$ ; collision energy at 5.0 eV; collision RF at 600 Vpp; ion cooler RF at 50–350 Vpp; transfer time at 121 s; pre-pulse storage time at 1  $\mu$ s. Calibration was performed with 1 mM sodium formate through a loop injection of 20  $\mu$ L at the start of each run.

A previously described method was used to quantitate the fraction of each peptide modified<sup>59</sup>. Briefly, the chromatograms for each singly-labelled and unlabelled peptide were extracted within a range of  $\pm 0.1$   $m/z$  and the spectrum for each peak was manually inspected to ensure the sampling of the correct ion only. The peptide fractional modification was calculated using **Equation 5**.

$$\text{Equation 5} \quad P = \frac{A_{\text{labelled}}}{A_{\text{labelled}} + A_{\text{unlabelled}}}$$

Where  $A_{\text{labelled}}$  and  $A_{\text{unlabelled}}$  correspond respectively to the peak area of each labeled and unlabeled peptide. Differences in the extent of labeling between peptides were considered significant when the  $p$ -value obtained from a Student  $t$ -test was  $< 0.05$ .

**Supplementary Table 7 Statistical analysis of tryptic digest for the Bamb\_5917 PCP domain alone and in the presence of Bamb\_5915**

Peptide	Bamb_5917 PCP domain alone (AspN)		Bamb_5917 PCP domain + Bamb_5915 (AspN)		$p$ -value
	Mean Fractional Labelling	Std. Dev.	Mean Fractional Labelling	Std. Dev.	
[211-227]	0.227	0.010	0.237	0.092	0.863572
[228-246]	0.613	0.101	0.320	0.007	0.007385
[247-253]	0.120	0.002	0.133	0.007	0.028772
[254-271]	0.915	0.073	0.647	0.123	0.031503
[305-315]	0.155	0.005	0.127	0.010	0.012659

**Supplementary Table 8 Statistical analysis of AspN digest for the Bamb\_5917 PCP domain alone and in the presence of Bamb\_5915**

Peptide	Bamb_5917 PCP domain alone (Trypsin)		Bamb_5917 PCP domain+ Bamb_5915 (Trypsin)		p-value
	Mean Fractional Labelling	Std. Dev.	Mean Fractional Labelling	Std. Dev.	
[261-266]	0.896	0.010	0.697	0.011	0.00002
[269-276]	0.047	0.013	0.044	0.007	0.677824
[277-286]	0.336	0.017	0.402	0.020	0.012848
[287-293]	0.398	0.030	0.553	0.015	0.001273
[294-315]	0.295	0.012	0.226	0.027	0.015068

**Supplementary Table 9 Statistical analysis of tryptic digest for Bamb\_5915 alone and in the presence of Bamb\_5917 PCP domain**

Peptide	Bamb_5915 alone (Trypsin)		Bamb_5915 + Bamb_5917 PCP domain (Trypsin)		p-value
	Mean Fractional Labelling	Std. Dev.	Mean Fractional Labelling	Std. Dev.	
[14-27]	0.388	0.015	0.383	0.006	0.636327
[31-38]	0.248	0.025	0.267	0.013	0.306741
[46-55]	0.525	0.051	0.490	0.054	0.464438
[90-95]	0.153	0.001	0.146	0.006	0.106615
[96-100]	0.178	0.025	0.184	0.008	0.701231
[101-109]	0.122	0.018	0.115	0.007	0.562138
[110-115]	0.053	0.033	0.041	0.010	0.61052
[116-120]	0.304	0.116	0.257	0.041	0.65605
[135-158]	0.131	0.029	0.129	0.012	0.910586
[159-166]	0.292	0.017	0.289	0.012	0.786162
[167-175]	0.937	0.004	0.920	0.009	0.038236
[176-186]	0.150	0.043	0.134	0.022	0.597248
[224-229]	0.032	0.012	0.028	0.004	0.550744
[230-245]	0.450	0.054	0.463	0.082	0.831621
[246-253]	0.332	0.039	0.357	0.013	0.348601
[266-277]	0.166	0.018	0.152	0.005	0.251043
[278-286]	0.160	0.009	0.164	0.002	0.485756
[291-299]	0.107	0.026	0.099	0.004	0.636432
[301-307]	0.472	0.025	0.423	0.014	0.040118
[368-375]	0.257	0.023	0.214	0.006	0.036107
[389-394]	0.568	0.067	0.552	0.019	0.722774
[400-406]	0.132	0.024	0.105	0.029	0.284526
[421-431]	0.219	0.024	0.220	0.014	0.936599
[432-442]	0.146	0.009	0.174	0.007	0.012767
[496-512]	0.336	0.018	0.330	0.011	0.641389

**Supplementary Table 10 Statistical analysis of AspN digest for Bamb\_5915 alone and in the presence of Bamb\_5917 PCP domain**

Peptide	Bamb_5915 alone (AspN)		Bamb_5915 + Bamb_5917 PCP domain (AspN)		p-value
	Mean Fractional Labelling	Std. Dev.	Mean Fractional Labelling	Std. Dev.	
[-3-20]	0.521	0.028	0.381	0.060	180305
[21-36]	0.167	0.010	0.156	0.012	0.280668
[37-60]	0.676	0.056	0.515	0.107	0.08243
[132-152]	0.060	0.014	0.069	0.006	0.398344
[153-179]	0.471	0.028	0.341	0.043	0.012313
[180-194]	0.516	0.159	0.578	0.162	0.656685
[226-238]	0.047	0.004	0.040	0.003	0.104074
[250-257]	0.108	0.004	0.076	0.033	0.172771
[340-357]	0.686	0.080	0.677	0.105	0.913283
[358-368]	0.059	0.014	0.090	0.034	0.229068
[406-428]	0.252	0.019	0.311	0.074	0.254223
[433-460]	0.370	0.022	0.483	0.040	0.012914
[461-473]	0.074	0.006	0.071	0.009	0.64808
[474-490]	0.107	0.029	0.116	0.053	0.790698
[494-512]	0.119	0.007	0.054	0.042	0.055629

#### *Protein docking simulations*

A series of docking simulations for the Bamb\_5917 PCP domain:Bamb\_5915 complex were performed using High Ambiguity Driven protein-protein DOCKing (HADDOCK). The NMR structure of the Bamb\_5917 PCP domain, the X-ray structure of Bamb\_5915 and frames from the aMD simulations of the two proteins sampled every 50 ns were used as starting points for the docking simulations. In addition, the overall position of the disordered C-terminus was varied in the starting structures of the Bamb\_5917 PCP domain. Simulations were performed with both the *apo*- and *holo*-Bamb\_5917 PCP domain. Residues 225, 230, 231, 234, 238, 255, 256, 260, 263, 309, 310, 311, 312, 313, 314, 315 in the PCP domain were defined as active. Residues 9, 11, 12, 13, 16, 18, 19, 20, 21, 23, 24, 25, 26, 27, 302, 303, 306, 307, 375 were defined as active in the Bamb\_5915. Only surface residues from the masked peptides were chosen, with the regions for which masking can be explained through condensation domain conformational changes being excluded. For the  $\beta$ -hairpin on the  $\beta$ HD domain all the surface residues were defined as active. A distance of  $<5 \text{ \AA}$  was defined as an unambiguous constraint for the Ppant thiol and N $\epsilon$ 2 of the His205 residue of Bamb\_5915. Finally, an effective distance of  $<10 \text{ \AA}$  was defined as an unambiguous constraint for the distance between the R315 residue of the Bamb\_5917 PCP domain and the D21 residue of Bamb\_5915.

#### *In vitro acetylation assays with Bamb\_5915, AHCCA and acetylated holo-PCP domains*

AHCCA acetylation assays were carried out by combining acetylated *holo*-PCP domains (200  $\mu\text{M}$ ) in separate reactions with Bamb\_5915 (20  $\mu\text{M}$ ) and AHCCA (1 mM), in a total volume of 100  $\mu\text{L}$ . Reactions were performed in buffer containing 20 mM Tris-HCl (pH 7.5) and 100 mM NaCl. For competition experiments, the excised  $\beta$ HD domain or the Bamb\_5917/Sven\_0512 SLiMs (purchased from Insight Biotechnologies, 98% purity) were added to the reaction in different molar ratios. Control reactions were performed in the absence of Bamb\_5915. Following incubation for 1 h at 30  $^{\circ}\text{C}$ , the reactions were stopped by the addition of two volumes of methanol. Precipitates were removed by centrifugation (13000 g, 1 min) and the supernatants were analyzed by UHPLC-ESI-QTOF-MS using the elution profile in Supplementary Table 11. All reactions were performed in triplicate.

**Supplementary Table 11 Elution conditions used for LC-MS analyses of small molecules**

Time (min)	H <sub>2</sub> O/0.1 % formic acid (%)	MeOH/0.1 % formic acid (%)	Flow rate (mL/min)
0	95	5	0.2
5	95	5	0.2
18	0	100	0.2
22	0	100	0.2
25	95	5	0.2
35	95	5	0.2

*UHPLC-ESI-Q-TOF-MS analysis of proteins and small molecules*

UHPLC-ESI-Q-TOF-MS analyses of small molecules were carried out using a Dionex UltiMate 3000 UHPLC connected to a Zorbax Eclipse Plus column (C18, 100 × 2.1 mm, 1.8 mm) coupled to a Bruker maXis Impact mass spectrometer. Mobile phases consisted of water (A) and acetonitrile (B), each supplemented with 0.1% formic acid. A gradient of 5% B to 100% B was employed at a flow rate of 0.2 mL/min (Supplementary Table 11). The mass spectrometer was operated in positive ion mode with a scan range of 50-3000 *m/z*. Source conditions were: end plate offset at -500 V; capillary at -4500 V; nebulizer gas (N<sub>2</sub>) at 1.6 bar; dry gas (N<sub>2</sub>) at 8 L min<sup>-1</sup>; dry temperature at 180 °C. Ion transfer conditions were: ion funnel RF at 200 Vpp; multiple RF at 200 Vpp; quadrupole low mass at 55 *m/z*; collision energy at 5.0 eV; collision RF at 600 Vpp; ion cooler RF at 50–350 Vpp; transfer time at 121 ms; pre-pulse storage time at 1 ms. Calibration was performed with 1 mM sodium formate through a loop injection of 20 mL at the start of each run.

UHPLC-ESI-Q-TOF-MS analyses of intact proteins were conducted using a Dionex UltiMate 3000 RS UHPLC connected to an ACE 3 C4-300 reverse phase column (Advanced Chromatography Technologies, Aberdeen, UK; 100 × 2.1 mm) coupled to a Bruker maXis II mass spectrometer. The elution profile is given in Supplementary Table 12. Source conditions were: capillary at -4500 V, ion polarity positive, time-of-flight tube at 9860 V and detector at 3500 V.

**Supplementary Table 12 Elution conditions used for LC-MS analyses of purified recombinant proteins**

Time (min)	H <sub>2</sub> O/0.1 % formic acid (%)	MeOH/0.1% formic acid (%)	Flow Rate (mL/min)
0	90	10	0.1
8	90	10	0.1
20	0	100	0.1
25	0	100	0.1
30	90	10	0.1

*Dynamics predictions from sequence*

Order parameters,  $S^2$ , for the Bamb\_5917 PCP domain were predicted using the program Dynamine<sup>11</sup>. The IUpred server<sup>12</sup> was employed to predict intrinsic disorder.

### Molecular dynamic simulations

Initial coordinates for simulation of Bamb\_5915 were prepared using AMBER. Ligands and ions, including the phosphate ion bound in the active site in the crystal structure were removed. Molecular dynamics (MD) simulations were performed using the AMBER ff98SB force field.<sup>60</sup> The resulting structure was solvated with TIP3BOX water, such that no atom belonging to Bamb\_5915 was less than 10 Å from any box edge, using the TLEAP program available on the WENMR and WEST-LIFE grid/cloud resources<sup>61</sup>.

MD heating, equilibration, and production steps were performed using the AMBER 14 software available on the WENMR and WEST-LIFE grid/cloud resources. Simulation used the SHAKE<sup>62</sup> algorithm to constrain all protein bonds involving a hydrogen atom; a 2.0 fs time-step was used in these simulations<sup>63</sup>. Long-range electrostatics were calculated using the Particle Mesh Ewald (PME) method with a 12.0 Å cutoff. PME was used for nonbonded interactions. In all simulations, the Langevin thermostat ( $\gamma = 2.0 \text{ ps}^{-1}$ ) was used to maintain temperature control<sup>64</sup>. The solvated protein was equilibrated by carrying out a short minimization, 50 ps of heating and 50 ps of density equilibration with weak restraints on the protein followed by 500 ps of constant pressure equilibration at 300K. More specifically, after a two-step minimization process, in which solvent molecules were allowed to relax before the entire system was minimized, the system was slowly heated to 300K over 0.1 ns in a canonical ensemble (NVT) simulation, then equilibrated for 2 ns by performing isothermal-isobaric (NPT) simulations at 300K using a Berendsen barostat.

Accelerated molecular dynamics<sup>65-67</sup> (aMD) simulation was performed using CUDA-accelerated Amber 12. aMD simulations enhanced the sampling of rare events and high energy states. One independent production of a 1  $\mu\text{s}$  aMD simulation for Bamb\_5915 was performed on a local computational cluster as follows: aMD parameters a) EthreshP:  $E(\text{tot}) = -193186 \text{ kcal mol}^{-1}$ ; b) alphaP:  $\text{Alpha}(\text{tot}) = 10773 \text{ kcal mol}^{-1}$ ; c) EthreshD:  $E(\text{dih}) = 7446 \text{ kcal mol}^{-1}$ ; d) alphaD:  $\text{Alpha}(\text{dih}) = 406.4 \text{ kcal mol}^{-1}$ . Simulations frames were written every 1 ps for analysis. Long-range electrostatics were calculated using the Particle Mesh Ewald (PME) method with a 10.0 Å cutoff. MD trajectories were analyzed using the CPPTRAJ<sup>68</sup> module of AmberTools 17, Chimera was used throughout the course to prepare and visualize the structures. All images of protein structures depicted herein were rendered using Chimera<sup>17</sup> or ChimeraX<sup>69</sup>.

### Construction of HMMs

The initial  $\beta$ HD domain HMM was built from a multiple protein sequence alignment using Clustal Omega<sup>70</sup> of the  $\beta$ HD domain sequences from the tubulysin (TubC), epothilone (EpoB), and enacyloxin (Bamb\_5915) biosynthetic gene clusters (Supplementary Table 13), using hmmbuild<sup>71</sup>.

**Supplementary Table 13 Amino acid sequences that were used to build an initial  $\beta$ HD domain HMM**

Specialised metabolite	GeneBank	Sequence
enacyloxin	ABI91460.1	IPALLSELQARGITLSLADGELSFRAPKGALTPADRATLSARREAIIVAYLAAKA
epothilone	AAF62881.1	AINQLLNELEHQGVKLAADGERLQIQAPKNALNPPLLARISEHKSTILTMLRQRL
tubulysin	CAF05648.1	SAGALLAHAASLGVRLWVEGERLRFQAPPVMTPELQSRLLGGARHELIALLRQLQ

The resulting HMM was used to search the coding sequences of the MiBiG<sup>28</sup> database for more hits using hmmsearch<sup>71</sup> (min Eval= $10^{-5}$ ). This resulted in 51 HMM hits, the sequences for these hits were aligned and used to construct an expanded hmm ( $\beta$ HD domain expanded).  $\beta$ HD domain expanded was used to search the MiBiG database again (min Eval= $10^{-5}$ ). This resulted in a total of 86 hits. We excluded 3 of these hits from the analysis as their upstream genes were not included in the MiBiG database so we could not identify a putative SLiM partner. The hits with Eval  $>10^{-5}$  were also examined manually for presence of SLiM- $\beta$ HD domain interfaces. This resulted in 85 hits from biosynthetic pathways for 58 metabolites (Supplementary Table 14).

**Supplementary Table 14 Systems with identified SLiM-βHD domain interfaces in MIBiG database.** <sup>1</sup>CL – condensation domain, H – halogenase domain, Cy – heterocyclisation domain, C/E – dual epimerisation and condensation domain.

Specialised metabolite	Organism	GeneBank SLiM	Locus tag SLiM	GeneBank βHD	Locus tag βHD	Next Domain
aeruginoside	<i>Planktothrix agardhii</i> NIVA-CYA 126	CAM59601.1	aerA	CAJ21198.2	aerB	<sup>1</sup> CL
aeruginosin	<i>Microcystis aeruginosa</i> NIES-98	ACM68682.1	AerA	ACM68683.1	aerJ	H
aeruginosin	<i>Microcystis aeruginosa</i> NIES-98	ACM68682.1	AerA	ACM68684.1	aerB	<sup>1</sup> CL
ajudazol	<i>Chondromyces crocatus</i>	CAQ18838.1	ajuK	CAQ18839.1	ajuL	<sup>1</sup> CL
ajudazol	<i>Chondromyces crocatus</i>	CAQ18830.1	ajuC	CAQ18831.1	ajuD	Cy
althiomycin	<i>Myxococcus xanthus</i>	CCA29202.1	alma	CCA29203.1	almB	<sup>1</sup> CL
anabaenopeptin	<i>Anabaena</i> sp. 90	ACZ55944.1	aptC	ACZ55946.1	aptD	<sup>1</sup> CL
anabaenopeptin	<i>Planktothrix agardhii</i> NIVA-CYA 126/8	ABV79987.1	apnC	ABV79988.1	apnD	<sup>1</sup> CL
antalid	<i>Polyangium spumosum</i>	ANI24099.1	antA	ANI24100.1	antB	<sup>1</sup> CL
arthrofactin	<i>Pseudomonas</i> sp. MIS38	BAC67534.2	arfA	BAC67535.1	arfB	C/E
arthrofactin	<i>Pseudomonas</i> sp. MIS38	BAC67535.1	arfB	BAC67536.1	arfC	C/E
bacitracin	<i>Bacillus licheniformis</i>	AAC06346.1	bacA	AAC06347.1	bacB	<sup>1</sup> CL
bananamides	<i>Pseudomonas fluorescens</i>	AOA33122.1	banB	AOA33123.1	banC	C/E
barbamide	<i>Lyngbya majuscula</i>	AAN32980.1	barF	AAN32981.1	barG	<sup>1</sup> CL
bleomycin	<i>Streptomyces verticillus</i>	AAG02357.1	blmVIII	AAG02358.1	blmVII	<sup>1</sup> CL
burkholdac	<i>Burkholderia thailandensis</i> E264	ABC36687.1	BTH_I2365	ABC36785.1	BTH_I2364	<sup>1</sup> CL
chondramide	<i>Chondromyces crocatus</i>	CAJ46689.1	cmdA	CAJ46691.1	cmdC	<sup>1</sup> CL
coelibactin	<i>Streptomyces coelicolor</i> A3(2)	CAC17499.1	SC4C2.17	CAC17500.1	SC4C2.18	Cy
cryptophycin	<i>Nostoc</i> sp. ATCC 53789	ABM21570.1	crpB	ABM21571.1	crpC	<sup>1</sup> CL
cyanopeptin	<i>Planktothrix agardhii</i> NIVA-CYA 116	ABI26078.1	ociB	ABI26079.1	ociC	<sup>1</sup> CL
cyanopeptolin	<i>Microcystis</i> sp. NIVA-CYA 172/5	AAZ03552.1	mcnC	AAZ03554.1	mcnE	<sup>1</sup> CL
cyanopeptolin	<i>Microcystis</i> sp. NIVA-CYA 172/5	AAZ03552.1	mcnC	AAZ03553.1	mcnD	H
cystothiazole A	<i>Cystobacter fuscus</i>	AAW03329.1	ctaF	AAW03330.1	ctaG	<sup>1</sup> CL
cystothiazole A	<i>Cystobacter fuscus</i>	AAW03325.1	ctaB	AAW03326.1	ctaC	Cy
cystothiazole A	<i>Cystobacter fuscus</i>	AAW03326.1	ctaC	AAW03327.1	ctaD	Cy
didemnin	<i>Tistrella mobilis</i> KA081020-065	AFK57213.1	didB	AFK57214.1	didC	<sup>1</sup> CL
dkxanthene	<i>Stigmatella aurantiaca</i> DW4/3-1	CAQ34917.1	dkxF	CAQ34918.1	dkxG	Cy
dkxanthene	<i>Stigmatella aurantiaca</i> DW4/3-1	CAQ34920.1	dkxI	CAQ34921.1	dkxJ	<sup>1</sup> CL
enacyloxin	<i>Burkholderia ambifaria</i> AMMD	ABI91461.1	bamb_5917	ABI91460.1	bamb_5915	<sup>1</sup> CL
entolysin	<i>Pseudomonas entomophila</i> L48	CAK15815.1	PSEEN3045	CAK15814.1	PSEEN3044	C/E
entolysin	<i>Pseudomonas entomophila</i> L48	CAK15814.1	PSEEN3044	CAK15815.1	PSEEN3045	C/E
epothilone	<i>Sorangium cellulosum</i>	AAF26919.1	epoA	AAF26925.1	epoP	Cy
epothilone	<i>Sorangium cellulosum</i>	AAF62880.1	epoA	AAF62881.1	epoB	Cy
epothilone	<i>Sorangium cellulosum</i> strain KYC3013	ADB12488.1	epoA	ADB12489.1	epoB	Cy
epothilone	<i>Sorangium cellulosum</i> So0157-2	ACB46192.1	epoA	ACB46193.1	epoP	Cy
FK228/romidepsin	<i>Chromobacterium violaceum</i>	ABP57747.1	DepC	ABP57748.1	depD	<sup>1</sup> CL
hectochlorin	<i>Lyngbya majuscula</i>	AAY42397.1	hctE	AAY42398.1	hctF	<sup>1</sup> CL
hectochlorin	<i>Lyngbya majuscula</i>	AAY42396.1	hctD	AAY42397.1	hctE	<sup>1</sup> CL
jamaicamide	<i>Lyngbya majuscula</i>	AAS98785.1	jamN	AAS98786.1	jamO	<sup>1</sup> CL
leupyrrin	<i>Sorangium cellulosum</i>	ADZ24998.1	leuD	ADZ24999.1	leuE	<sup>1</sup> CL
leupyrrin	<i>Sorangium cellulosum</i>	ADZ24993.1	leu9	ADZ24995.1	leuA	Cy
massetolide A	<i>Pseudomonas fluorescens</i> SS101	ABH06368.2	massB	ABH06369.2	massC	C/E
massetolide A	<i>Pseudomonas fluorescens</i> SS101	ABH06367.1	massA <sup>a</sup>	ABH06368.2	massB	C/E
melithiazol	<i>Melittangium lichenicola</i>	CAD89773.1	melB	CAD89774.1	melC	Cy
melithiazol	<i>Melittangium lichenicola</i>	CAD89774.1	melC	CAD89775.1	melD	Cy
melithiazol	<i>Melittangium lichenicola</i>	CAD89777.1	melF	CAD89778.1	melG	<sup>1</sup> CL
microsclerodermins	<i>Jahnella</i> sp. MSr9139	AHB82071.1	mseH	AHB82072.1	mseI	<sup>1</sup> CL
microsclerodermins	<i>Sorangium cellulosum</i>	AHB82058.1	mseH	AHB82059.1	mseI	<sup>1</sup> CL
myxalamid	<i>Stigmatella aurantiaca</i>	AAK57185.1/	mxAB	AAK57184.1	myxA	<sup>1</sup> CL
myxalamid	<i>Stigmatella aurantiaca</i>	AAK57186.1	/1mxAB2	AAK57184.1	myxA	<sup>1</sup> CL
myxothiazol	<i>Stigmatella aurantiaca</i>	AAF19811.1	mtaC	AAF19812.1	mtaD	Cy
myxothiazol	<i>Stigmatella aurantiaca</i>	AAF19814.1	mtaF	AAF19815.1	mtaG	<sup>1</sup> CL
nannocystin	<i>Nannocystis</i> sp. MB1016	ALD82525.1	ncyE	ALD82526.1	ncyF	<sup>1</sup> CL
nannocystin	<i>Nannocystis</i> sp. MB1016	ALD82523.1	ncyC	ALD82525.1	ncyE	<sup>1</sup> CL
nostopeptolide	<i>Nostoc</i> sp. GSV224	AAF15892.2	nosB	AAF17280.1	nosC	<sup>1</sup> CL
nostophycin	<i>Nostoc</i> sp. 152	AEU11006.1	nnpB	AEU11003.1	nnpC	<sup>1</sup> CL
nunapeptin	<i>Pseudomonas</i> sp. In5	KPN93063.1	AL066_08505	KPN93064.1	AL066_08510	C/E
orfamide	<i>Pseudomonas</i> sp. CMR12a	AFH75329.1	ofaB	AFH75330.1	ofaC	C/E
orfamide	<i>Pseudomonas</i> sp. CMR12a	AAV91420.2	ofaB/PFL_214 6	AAV91421.3	ofaC/PFL_2147	C/E
orfamide	<i>Pseudomonas</i> sp. CMR12a	AFH75328.1	ofaA	AFH75329.1	ofaB	C/E

phenalamide	<i>Myxococcus stipitatus</i> DSM 14675	AGC45620.1	MYSTI_04322	AGC45618.1	MYSTI_04320	<sup>1</sup> C <sub>L</sub>
poaeamide	<i>Pseudomonas poae</i> RE*1-1-14	AGE25479.1	H045_07035	AGE25480.1	H045_07040	C/E
poaeamide B	<i>Pseudomonas synxantha</i>	ANQ43348.1	ANQ43348.1	ANQ43349.1	ANQ43349.1	C/E
putisolvin	<i>Pseudomonas putida</i>	ABW17376.1	psoB	ABW17377.1	psoC	C/E
putisolvin	<i>Pseudomonas putida</i>	ABW17375.1	psoA	ABW17376.1	psoB	C/E
puwainaphycins	<i>Cylindrospermum alatosporum</i> CCALA 988	AIW82282.1	puwE	AIW82283.1	puwF	<sup>1</sup> C <sub>L</sub>
pyochelin	<i>Pseudomonas aeruginosa</i> PAO1	AAC83656.1	pchE	AAC83657.1	pchF	Cy
ralsolamycin	<i>Ralstonia solanacearum</i> GM11000	CAD17792.1	RSp0641	CAD17793.1	RSp0642	C/E
rhabdopeptides	<i>Xenorhabdus nematophila</i> ATCC 19061	CBJ90288.1	XNC1_2229	CBJ90287.1	XNC1_2228	<sup>1</sup> C <sub>L</sub>
rhabdopeptides	<i>Xenorhabdus nematophila</i> ATCC 19061	CBJ90287.1	XNC1_2228	CBJ90288.1	XNC1_2229	<sup>1</sup> C <sub>L</sub>
rhabdopeptides	<i>Xenorhabdus nematophila</i> ATCC 19061	CBJ90288.1	XNC1_2229	CBJ90289.1	XNC1_2230	<sup>1</sup> C <sub>L</sub>
sessilin	<i>Pseudomonas</i> sp. CMR12a	AFH75320.1	sesA	AFH75321.1	sesB	C/E
spiruchostatin	<i>Pseudomonas</i> sp. Q71576	AFR69333.1	spiC1	AFR69334.1	spiDE1	<sup>1</sup> C <sub>L</sub>
spumigin	<i>Nodularia spumigena</i> CCY9414	AHJ31216.1	NSP_49240	AHJ31215.1	NSP_49230	<sup>1</sup> C <sub>L</sub>
syringafactin	<i>Pseudomonas syringae</i>	AAO56328.1	syfA	AAO56329.1	syfB	C/E
syringomycin	<i>Pseudomonas syringae</i> pv. <i>syringae</i> B728a	AAY37653.1	Psyr_2614	AAY37654.1	Psyr_2615	C/E
syringopeptin	<i>Pseudomonas syringae</i> pv. <i>syringae</i>	AAF99707.2	sypA	AAO72424.1	sypB	C/E
tallysomyacin	<i>Streptoalloteichus hindustanus</i>	ABL74938.1	tImVIII	ABL74939.1	tImVII	<sup>1</sup> C <sub>L</sub>
tolaasin	<i>Pseudomonas costantinii</i>	CCJ67637.1	taaB	CCJ67638.1	taaC	C/E
tolaasin	<i>Pseudomonas costantinii</i>	CCJ67636.1	taaA	CCJ67637.1	taaB	C/E
tubulysin	<i>Cystobacter</i> sp. SBCb004	ADH04678.1	tubB	ADH04679.1	tubC	<sup>1</sup> C <sub>L</sub>
tubulysin	<i>Archangium disciforme</i>	CAF05647.1	tubB	CAF05648.1	tubC	<sup>1</sup> C <sub>L</sub>
viscosin	<i>Pseudomonas fluorescens</i> SBW25	CAY48788.1	viscB/PFLU_2 552	CAY48789.1	viscC/PFLU_255 3	C/E
WLIP	<i>Pseudomonas putida</i>	AFJ23825.1	wlpB	AFJ23826.1	wlpC	C/E
xantholysin	<i>Pseudomonas putida</i>	AGM14933.1	xtlB	AGM14934.1	xtlC	C/E
xenortide	<i>Xenorhabdus nematophila</i>	CBJ90359.1	XNC1_2300	CBJ90358.1	XNC1_2299	<sup>1</sup> C <sub>L</sub>
zorbamycin	<i>Streptomyces flavoviridis</i>	ACG60781.1	zImVIII	ACG60777.1	zImVIIa	<sup>1</sup> C <sub>L</sub>

<sup>a</sup>massA is not included in the MiBiG massetolide A cluster.

### Supplementary Table 15 Examples of systems with known metabolites and identified SLiM-βHD domain boundaries that are not included in the MiBiG database at the time of writing of the manuscript

Specialised metabolite	Organism	GeneBank SLiM	L. tag SLiM	GeneBank βHD	Locus tag βHD	Next Domain
anabaenopeptilide	<i>Anabaena</i> sp. 90	CAC01604.1	adpB	CAC01605.1	adpC	H
columbamide	<i>Moorea bouillonii</i> PNG	AKQ09583.1	ColF	AKQ09584.1	ColG	<sup>1</sup> C <sub>L</sub>
equibactin	<i>Streptococcus equi</i>	CRV28838.1	eqbE	CRV28842.1	eqbG	Cy
microginin	<i>Microcystis aeruginosa</i>	CCI09459.1	MicA	CCI09457.1	MicD	<sup>1</sup> C <sub>L</sub>
microginin	<i>Microcystis aeruginosa</i>	CCI09459.1	MicA	CCI09458.1	MicC	<sup>1</sup> C <sub>L</sub>
watasemycin	<i>Streptomyces</i> sp. TP-A0597	CCA53799.1	Sven_0512	CCA53804.1	Sven_0517	Cy

To determine the likely interaction partner of the βHD domains, we assumed collinearity and looked at the gene upstream of the gene encoding it, checking that it was in the same direction. To further analyse these βHD domain-SLiM pairs, we annotated them with the cluster-specific NRPS/PKS HMMs included in antiSMASH<sup>72</sup>. We analysed the domains downstream of the βHD domain and upstream of the putative SLiM partner.

In putative SLiM-containing proteins, the C-terminal region was defined as all amino acids from the C-terminus that did not conform to any other domain, where boundaries were set by NRPS/PKS HMMs. In addition, we have identified likely disordered residues in such C termini using disorder prediction by IUPRED<sup>12</sup>. The python libraries in clusterTools<sup>73</sup> were used to perform the domain annotation and extraction, identify the putative interaction partners, and extract the disordered regions.

#### Genebank database search

We performed the same analysis as described above on databases of coding sequences from the NCBI Genbank (Bacterial and Environmental Division) and Genomes (Refseq: Archea, Bacteria, Protozoa reference and representative genomes) constructed using clusterTools. Using the Eval = 10<sup>-5</sup> cutoff we obtained 557 hits for a βHD domain in front of an <sup>1</sup>C<sub>L</sub> domain, 547 hits for a βHD domain in front of a

Cy domain, 183 hits for a  $\beta$ HD domain in front of a C/E domain, 22 for a  $\beta$ HD in front of a halogenase, 14 for a  $\beta$ HD domain in front of a TR domain, 11 for a  $\beta$ HD domain in front of a TE domain, 1 for a  $\beta$ HD domain in front of an nMT domain and 83 for a  $\beta$ HD domain in front of other proteins. The hits are listed in the accompanying Excel file.

## References

1. Bloudoff, K., Rodionov, D. & Schmeing, T. M. Crystal structures of the first condensation domain of CDA synthetase suggest conformational changes during the synthetic cycle of nonribosomal peptide synthetases. *J. Mol. Biol.* **425**, 3137–3150 (2013).
2. Keating, T. A., Marshall, C. G., Walsh, C. T. & Keating, A. E. The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains. *Nat. Struct. Biol.* **9**, 522–526 (2002).
3. Drake, E. J. *et al.* Structures of two distinct conformations of holo-non-ribosomal peptide synthetases. *Nature* **529**, 235–238 (2016).
4. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. *Nucleic Acids Res.* **38**, W545–W549 (2010).
5. Chang, C.-Y. *et al.* Structural Insights into the Free-Standing Condensation Enzyme SgcC5 Catalyzing Ester-Bond Formation in the Biosynthesis of the Eneidyne Antitumor Antibiotic C-1027. *Biochemistry* **57**, 3278–3288 (2018).
6. Richter, C. D., Nietlispach, D., Broadhurst, R. W. & Weissman, K. J. Multienzyme docking in hybrid megasynthetases. *Nat. Chem. Biol.* **4**, 75–81 (2008).
7. Dowling, D. P. *et al.* Structural elements of an NRPS cyclization domain and its intermodule docking domain. *Proc. Natl. Acad. Sci.* **113**, 12432–12437 (2016).
8. Samel, S. A., Schoenafinger, G., Knappe, T. A., Marahiel, M. A. & Essen, L.-O. Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase. *Structure* **15**, 781–792 (2007).
9. Reimer, J. M., Aloise, M. N., Harrison, P. M. & Schmeing, T. M. Synthetic cycle of the initiation module of a formylating nonribosomal peptide synthetase. *Nature* **529**, 239–42 (2016).
10. Koglin, A. *et al.* Conformational switches modulate protein interactions in peptide antibiotic synthetases. *Science* **312**, 273–6 (2006).
11. Cilia, E., Pancsa, R., Tompa, P., Lenaerts, T. & Vranken, W. F. The DynaMine webserver: predicting protein dynamics from sequence. *Nucleic Acids Res.* **42**, W264–W270 (2014).
12. Dosztanyi, Z., Csizmok, V., Tompa, P. & Simon, I. IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21**, 3433–3434 (2005).
13. Bouvignies, G. & Kay, L. E. Measurement of Proton Chemical Shifts in Invisible States of Slowly Exchanging Protein Systems by Chemical Exchange Saturation Transfer. *J. Phys. Chem. B* **116**, 14311–14317 (2012).
14. Valluruapalli, P., Bouvignies, G. & Kay, L. E. Studying ‘invisible’ excited protein states in slow exchange with a major state conformation. *J. Am. Chem. Soc.* **134**, 8148–8161 (2012).
15. Richter, C. D., Nietlispach, D., Broadhurst, R. W. & Weissman, K. J. Multienzyme docking in hybrid megasynthetases. *Nat. Chem. Biol.* **4**, 75–81 (2008).
16. Hacker, C. *et al.* Structure-based redesign of docking domain interactions modulates the product spectrum of a rhabdopeptide-synthesizing NRPS. *Nat. Commun.* **9**, 4366 (2018).
17. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
18. Meng, E. C., Pettersen, E. F., Couch, G. S., Huang, C. C. & Ferrin, T. E. Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC Bioinformatics* **7**, 1–10 (2006).
19. Zhang, J. *et al.* Structural basis of nonribosomal peptide macrocyclization in fungi. *Nat. Chem. Biol.* **12**, 1001–1003 (2016).
20. Haslinger, K., Peschke, M., Brieke, C., Maximowitsch, E. & Cryle, M. J. X-domain of peptide synthetases recruits oxygenases crucial for glycopeptide biosynthesis. *Nature* **521**, 105–109

- (2015).
21. Chen, W. H., Li, K., Guntaka, N. S. & Bruner, S. D. Interdomain and Intermodule Organization in Epimerization Domain Containing Nonribosomal Peptide Synthetases. *ACS Chem. Biol.* **11**, 2293–2303 (2016).
  22. Samel, S. A., Czodrowski, P. & Essen, L. O. Structure of the epimerization domain of tyrocidine synthetase A. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **70**, 1442–1452 (2014).
  23. Samel, S. A., Schoenafinger, G., Knappe, T. A., Marahiel, M. A. & Essen, L.-O. Structural and Functional Insights into a Peptide Bond-Forming Bidomain from a Nonribosomal Peptide Synthetase. *Structure* **15**, 781–792 (2007).
  24. Drake, E. J. *et al.* Structures of two distinct conformations of holo-non-ribosomal peptide synthetases. *Nature* **529**, 235–8 (2016).
  25. Tanović, A., Samel, S. A., Essen, L.-O. & Marahiel, M. A. Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science* **321**, 659–63 (2008).
  26. Bloudoff, K., Fage, C. D., Marahiel, M. A. & Schmeing, T. M. Structural and mutational analysis of the nonribosomal peptide synthetase heterocyclization domain provides insight into catalysis. *Proc. Natl. Acad. Sci.* **114**, 95–100 (2017).
  27. Dowling, D. P. *et al.* Structural elements of an NRPS cyclization domain and its intermodule docking domain. *Proc. Natl. Acad. Sci.* **113**, 12432–12437 (2016).
  28. Van Zundert, G. C. P. *et al.* The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J. Mol. Biol.* **428**, 720–725 (2016).
  29. Medema, M. H. *et al.* Minimum Information about a Biosynthetic Gene cluster. *Nat. Chem. Biol.* **11**, 625–631 (2015).
  30. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
  31. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
  32. Wu, X. *et al.* A Novel Method for High-Level Production of TEV Protease by Superfolder GFP Tag. *J. Biomed. Biotechnol.* **2009**, 1–8 (2009).
  33. Jenner, M. *et al.* An unusual Burkholderia gladioli double chain-initiating nonribosomal peptide synthetase assembles ‘fungal’ icosalide antibiotics. *Chem. Sci.* (2019). doi:10.1039/C8SC04897E
  34. Lescop, E., Schanda, P. & Brutscher, B. A set of BEST triple-resonance experiments for time-optimized protein resonance assignment. *J. Magn. Reson.* **187**, 163–169 (2007).
  35. Grzesiek, S. & Bax, A. Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson.* **96**, 432–440 (1992).
  36. Schleucher, J., Sattler, M. & Griesinger, C. Coherence Selection by Gradients without Signal Attenuation: Application to the Three-Dimensional HNC0 Experiment. *Angew. Chemie Int. Ed. English* **32**, 1489–1491 (1993).
  37. Kay, L. E., Xu, G. Y. & Yamazaki, T. Enhanced-Sensitivity Triple-Resonance Spectroscopy with Minimal H<sub>2</sub>O Saturation. *Journal of Magnetic Resonance, Series A* **109**, 129–133 (1994).
  38. Zhang, O., Kay, L. E., Olivier, J. P. & Forman-Kay, J. D. Backbone <sup>1</sup>H and <sup>15</sup>N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR* **4**, 845–858 (1994).
  39. Palmer, A. G., Cavanagh, J., Wright, P. E. & Rance, M. Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy. *J. Magn. Reson.* **93**, 151–170 (1991).
  40. Kay, L. E., Keifer, P. & Saarinen, T. Pure Absorption Gradient Enhanced Heteronuclear Single Quantum Correlation Spectroscopy with Improved Sensitivity. *Journal of the American Chemical Society* **114**, 10663–10665 (1992).
  41. Schleucher, J. *et al.* A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J. Biomol. NMR* **4**, 301–306 (1994).
  42. Schanda, P., Van Melckebeke, H. & Brutscher, B. Speeding Up Three-Dimensional Protein NMR Experiments to a Few Minutes. *J. Am. Chem. Soc.* **128**, 9042–9043 (2006).

43. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–23 (2009).
44. Volk, J., Herrmann, T. & Wüthrich, K. Automated sequence-specific protein NMR assignment using the memetic algorithm MATCH. *J. Biomol. NMR* **41**, 127–138 (2008).
45. Fiorito, F., Herrmann, T., Damberger, F. F. & Wüthrich, K. Automated amino acid side-chain NMR assignment of proteins using <sup>13</sup>C- and <sup>15</sup>N-resolved 3D [1H,1H]-NOESY. *J. Biomol. NMR* **42**, 23–33 (2008).
46. Herrmann, T., Güntert, P. & Wüthrich, K. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J. Biomol. NMR* **24**, 171–89 (2002).
47. Güntert, P. & Buchner, L. Combined automated NOE assignment and structure calculation with CYANA. *J. Biomol. NMR* **62**, 453–471 (2015).
48. D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R. E., Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. K., T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K.M. Merz, F. Paesani, D.R. Roe, A. Roitberg, C. S., R. Salomon-Ferrer, G. Seabra, C.L. Simmerling, W. Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. & Kollman, W. and P. A. AMBER 14. *AMBER 14* (2014).
49. Bhattacharya, A., Tejero, R. & Montelione, G. T. Evaluating protein structures determined by structural genomics consortia. *Proteins Struct. Funct. Bioinforma.* **66**, 778–795 (2006).
50. Huang, Y. J., Powers, R. & Montelione, G. T. Protein NMR Recall, Precision, and F -measure Scores (RPF Scores): Structure Quality Assessment Measures Based on Information Retrieval Statistics. *J. Am. Chem. Soc.* **127**, 1665–1674 (2005).
51. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
52. Lovell, S. C. *et al.* Structure validation by C $\alpha$  geometry:  $\phi, \psi$  and C $\beta$  deviation. *Proteins Struct. Funct. Bioinforma.* **50**, 437–450 (2003).
53. Lüthy, R., Bowie, J. U. & Eisenberg, D. Assessment of protein models with three-dimensional profiles. *Nature* **356**, 83–85 (1992).
54. Sippl, M. J. Recognition of errors in three-dimensional structures of proteins. *Proteins Struct. Funct. Genet.* **17**, 355–362 (1993).
55. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 12–21 (2010).
56. Ulrich, E. L. *et al.* BioMagResBank. *Nucleic Acids Res.* **36**, D402-8 (2008).
57. García de la Torre, J., Huertas, M. L. & Carrasco, B. HYDRONMR: Prediction of NMR Relaxation of Globular Proteins from Atomic-Level Structures and Hydrodynamic Calculations. *J. Magn. Reson.* **147**, 138–146 (2000).
58. Manzi, L. *et al.* Carbene footprinting accurately maps binding sites in protein–ligand and protein–protein interactions. *Nat. Commun.* **7**, 13288 (2016).
59. Ponder, J. W. & Case, D. A. Force fields for protein simulations. *Adv. Protein Chem.* **66**, 27–85 (2003).
60. Bertini, I., Case, D. A., Ferella, L., Giachetti, A. & Rosato, A. A Grid-enabled web portal for NMR structure refinement with AMBER. *Bioinformatics* **27**, 2384–2390 (2011).
61. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. . Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).
62. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
63. Loncharich, R. J., Brooks, B. R. & Pastor, R. W. Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-acetylalanine-N'-methylamide. *Biopolymers* **32**, 523–535 (1992).
64. Hamelberg, D., Mongan, J. & McCammon, J. A. Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *J. Chem. Phys.* **120**, 11919–

- 11929 (2004).
65. Pierce, L. C. T., Salomon-Ferrer, R., Augusto F. De Oliveira, C., McCammon, J. A. & Walker, R. C. Routine access to millisecond time scale events with accelerated molecular dynamics. *J. Chem. Theory Comput.* **8**, 2997–3002 (2012).
  66. Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh ewald. *J. Chem. Theory Comput.* **9**, 3878–3888 (2013).
  67. Roe, D. R. & Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).
  68. Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
  69. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539–539 (2014).
  70. Eddy, S. R. Accelerated Profile HMM Searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).
  71. Blin, K. *et al.* antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **45**, W36–W41 (2017).
  72. Santos, E. L. de los & Challis, G. L. clusterTools: Proximity Searches For Functional Elements To Identify Putative Biosynthetic Gene Clusters. *bioRxiv* 119214 (2017). doi:10.1101/119214