Science Advances
AAAS

Supplementary Materials for

The structure of a polyketide synthase bimodule core

Yves U. Tittes *et al.*

Corresponding author: Timm Maier, timm.maier@unibas.ch

Sci. Adv. **8**, eabo6918 (2022)
DOI: 10.1126/sciadv.abo6918

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

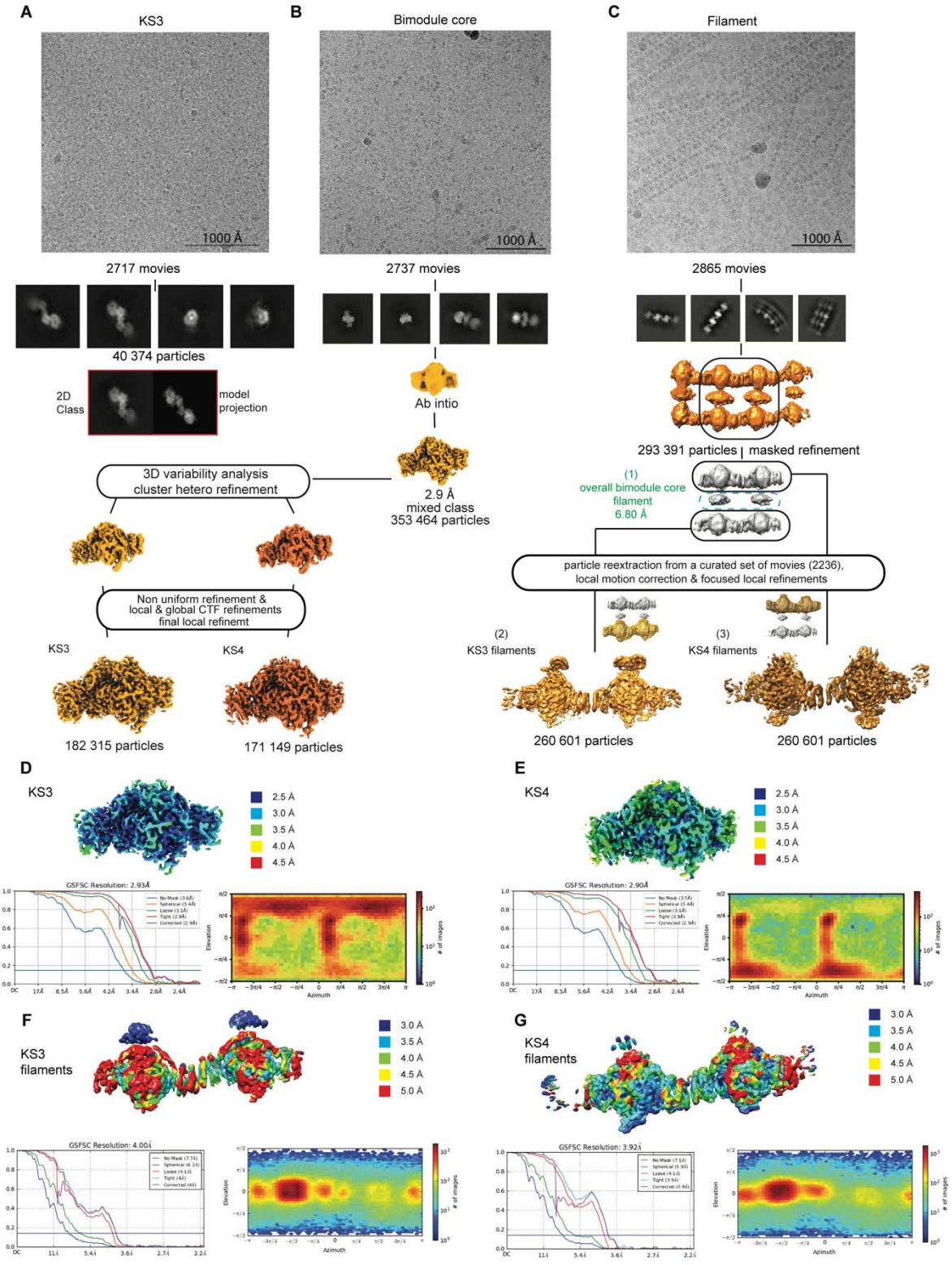

Fig. S1.

Schematic representation of cryoEM data processing. A total of three datasets were collected, one of the recombinantly expressed isolated KS3 (A), one of individual bimodule core dimers (B) and one of bimodule core filaments (C). **(A)** Data set for individual KS3

domains. 2D class averages reveal dimerization via the LINKS interface of the LD instead of the canonical KS dimer interface, as shown by comparison of a 2D class to a projection based on a model of LINKS-bridged KS3 monomers. Particle number, size and variability precluded high-resolution 3D reconstruction. **(B)** Data set for individual K3DAK4 bimodule cores. The overall bimodule core is highly flexible, but 3D variability analysis-based sorting into particle sets representing either KS3 (EMDB-14795) or KS4 (EMDB-14793) provided individual reconstructions for each KS to 2.9 Å resolution. Validation of sorting based on map analysis is provided in Figure S3. **(C)** Analysis of the K3DAK4 filament data set. To prevent misalignment due to pseudosymmetry a box-size large enough to fit four adjacent bimodule cores was chosen in the initial particle picking and reconstruction steps. The intermediate resolution overall bimodule core filament map (EMDB-14945) (1) was used for the rigid body refinement of the overall model. Focused refinements on pairs of KS dimers connected via LINKS interaction lead to improved local maps for pairs of KS3 (2) and KS4 (3) dimers at resolutions of 4.0 Å and 3.9 Å, respectively. Masks used in local refinement of pairs of KS3 or KS4 dimers are shown and the region excised from the overall map (1) for representing DH3 domains in the hybrid model in Fig. 2 is schematically indicated by a blue dashed line. **(D-G)** Local resolution maps, FSC plots and viewing direction distribution plots for individual reconstructions as indicated.

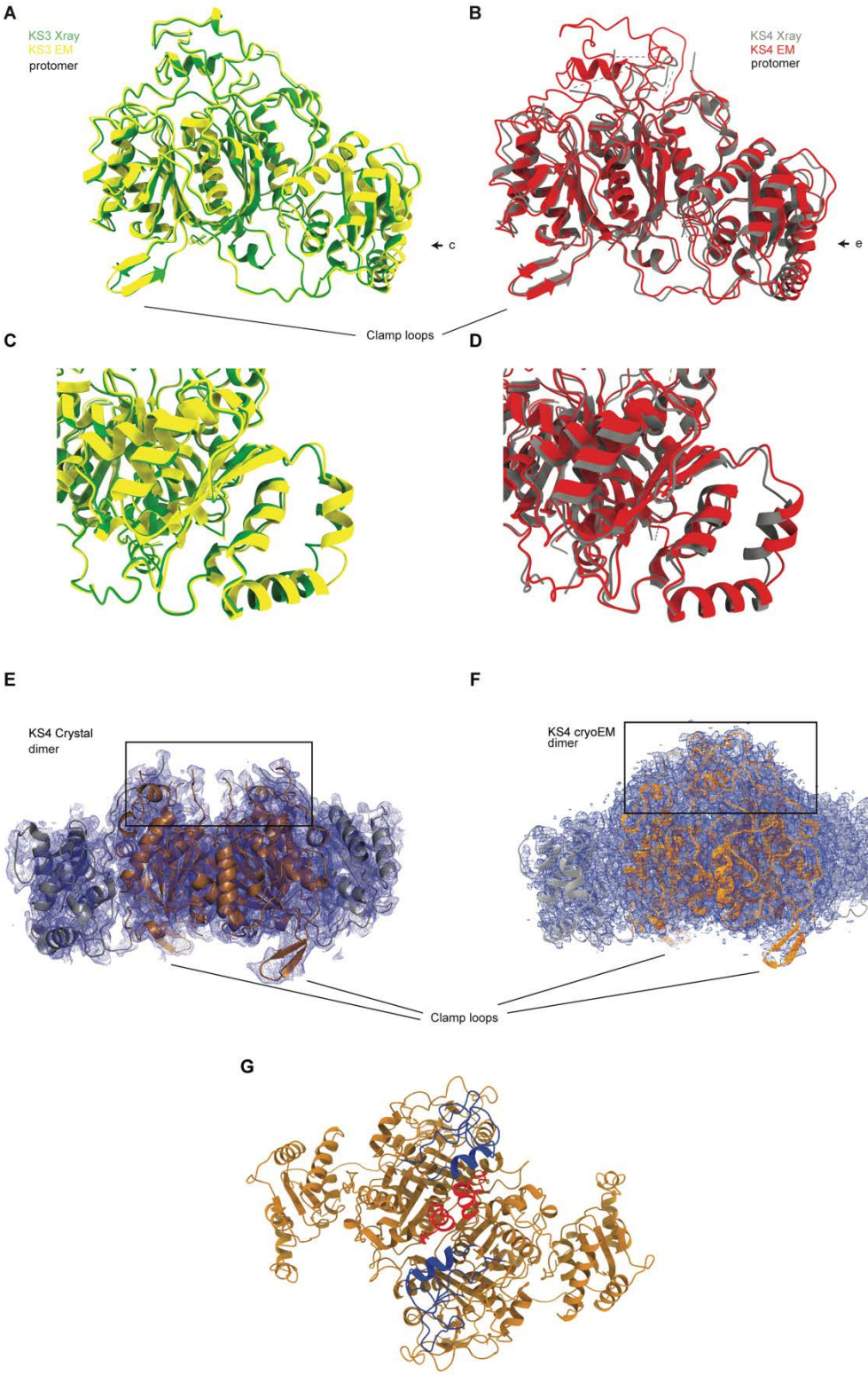
A
KS3 Xray
KS3 EM
protomer
c
B
KS4 Xray
KS4 EM
protomer
e
Clamp loops
C
D
E
KS4 Crystal
dimer
F
KS4 cryoEM
dimer
Clamp loops
G

Fig. S2.

Analysis of KS3 and KS4 domain structures. (A-F) Comparison of models of KSs derived from X-ray crystallographic analysis of the isolated domains and cryoEM analysis of KS domains in the context of individual bimodule cores for KS3 (A,C; yellow: cryoEM, green: crystallographic) and KS4 (B,D; red: cryoEM, grey: crystallographic) **(E,F,G)**, A central region around the active site (indicated by rectangles in (E,F) and red or blue color in (G)) is disordered in crystallographic analysis and not visualized in the crystallographic map at a contour level of 1.7 sigma (E), but is visualized in the cryoEM-derived map at contour level 0.26 in arbitrary units (F). (G) Cartoon style overview representation of KS4 based on cryoEM analysis. Regions shown in blue and red are only visualized in the cryoEM map (F), but not in the crystallographic map (E). The helix colored in red is a 20 AA extension of KS4 relative to KS3.

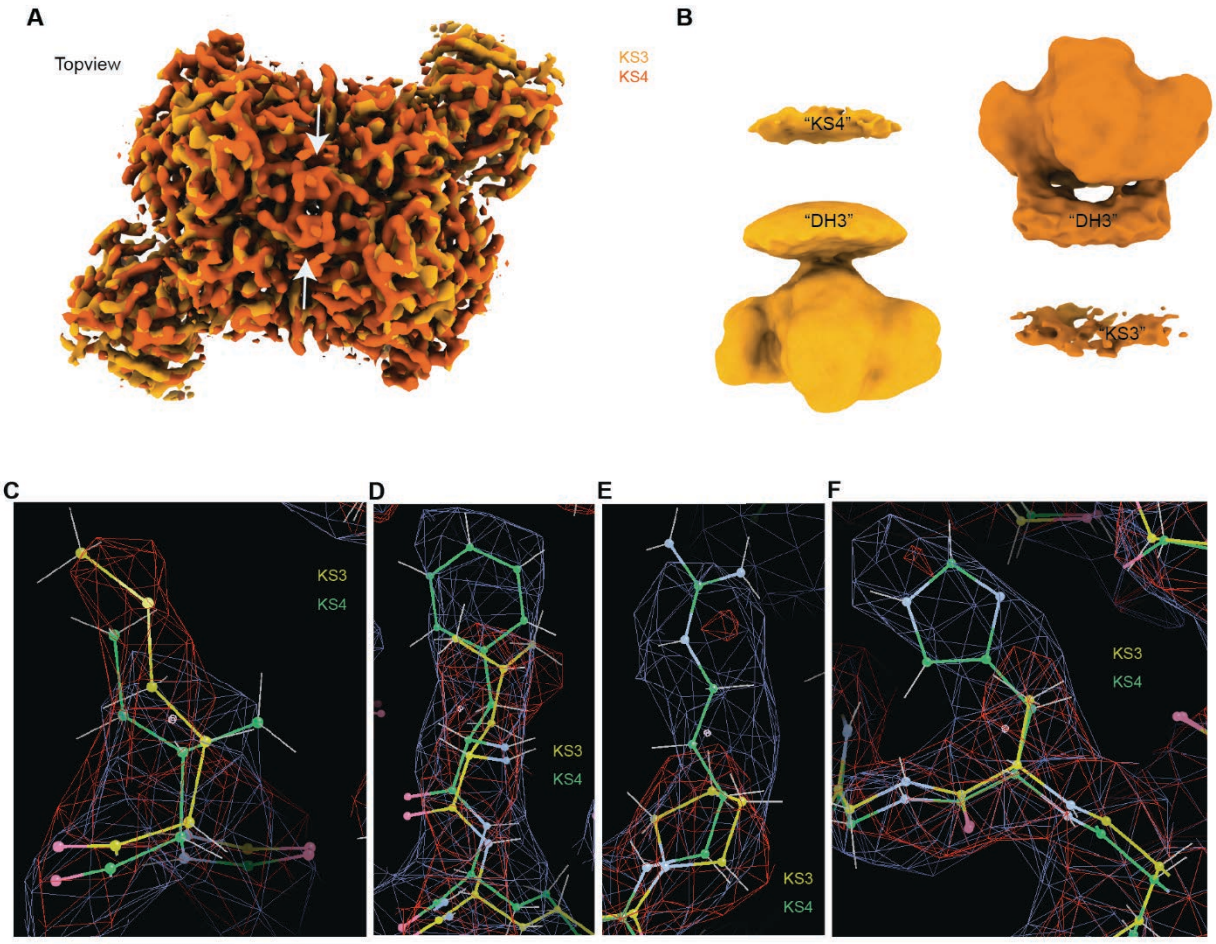

Fig. S3.

Comparison of KS3 and KS4 cryoEM maps and models. (A) Overlay of the individual maps for the closely-related KS3 and KS4 domains shows an overall highly similar structure. Differences are apparent around the center region close to the C-terminus above the active sites, where a loop (indicated by an arrow) in KS4 is around 20 AA longer than in KS3. In **(B)** B-factor softened (600 $Å^2$) maps for KS3 and KS4 reconstructions reveal residual low-resolution densities presumably representing the additional domains of the bimodule core in agreement with their relative positioning to each KS. **(C-F)** Individual examples of differences in sequence between KS3 and KS4 reflected in the respective cryoEM maps, confirming successful sorting into particles representing either KS3 or KS4. Model and map for KS3 in yellow and red; for KS4 in green and blue respectively.

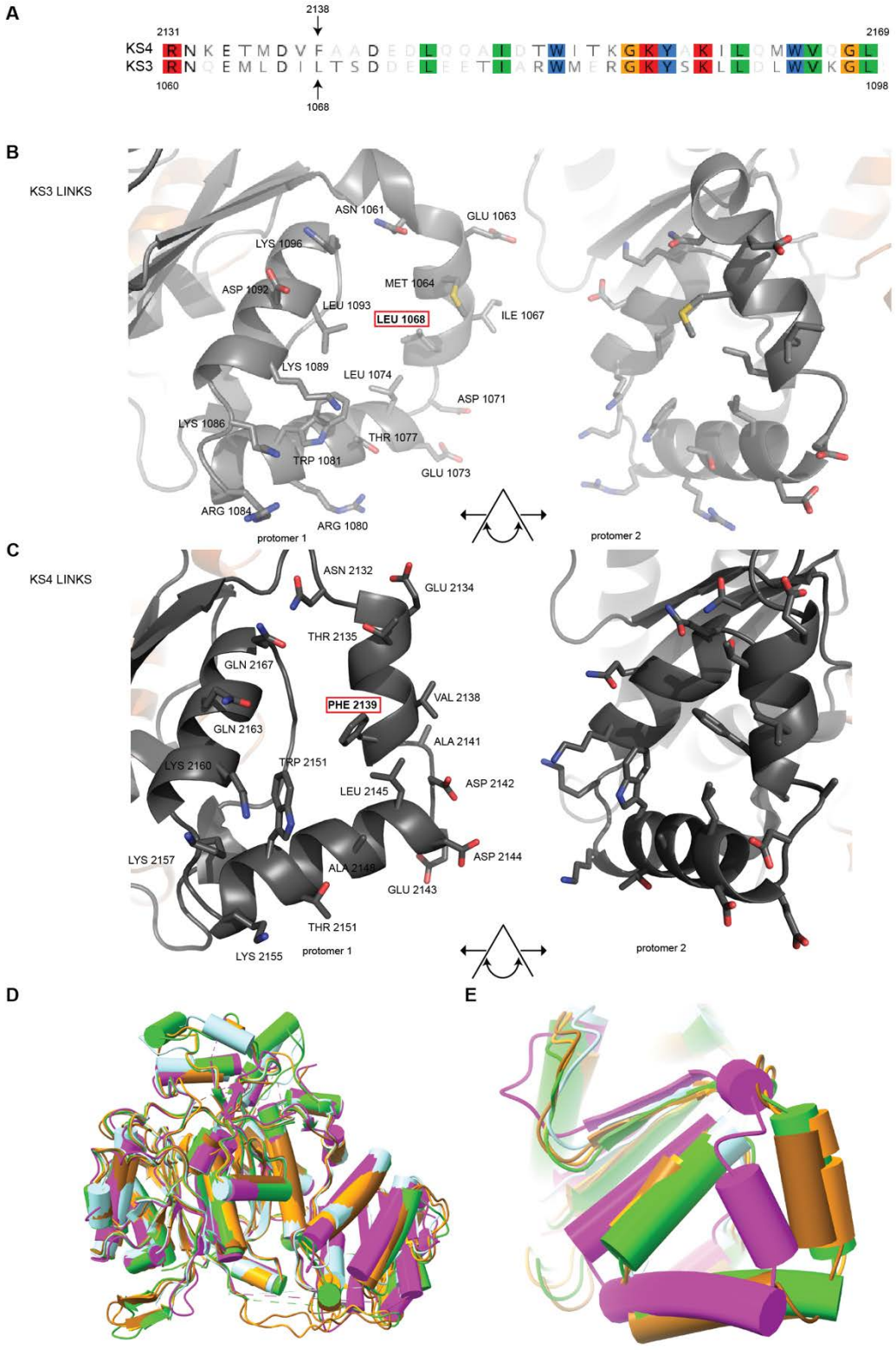
A
2138
2131
2169
KS4
KS3
1060
1068
1098
B
KS3 LINKS
ASN 1061
GLU 1063
LYS 1096
MET 1064
ASP 1092
LEU 1093
LEU 1068
ILE 1067
LYS 1089
LEU 1074
ASP 1071
LYS 1086
THR 1077
TRP 1081
GLU 1073
ARG 1084
ARG 1080
protomer 1
protomer 2
C
KS4 LINKS
ASN 2132
GLU 2134
THR 2135
GLN 2167
PHE 2139
VAL 2138
GLN 2163
ALA 2141
LYS 2160
TRP 2151
ASP 2142
LEU 2145
LYS 2157
ASP 2144
GLU 2143
THR 2151
LYS 2155
protomer 1
protomer 2
D
E

Fig. S4.

Structural analysis of KS3 and KS4 LINKS interfaces. (A) Sequence alignment of the three-alpha helical LINKS motifs in KS3 and KS4. **(B,C)** LINKS interfaces of KS3 (B) and KS4 (C) as observed in crystal structures. For visualization, the interfaces are opened up by rotation and translation. The overall three-alpha helical fold is similar for KS3 and KS4, but variations of individual residues such as LEU1068 in KS3 matching PHE2139 in KS4 (indicated also in (A)) ensure that steric clashes prevent mispairing between the KS domains of module 3 and 4. **(D,E)** Structural superimposition of the KS domains KS3 (light orange) and KS4 (dark orange) (reported here), isolated KS2 from the same BGC11 (green, PDB 4Z37), KS9 of the bacillaene cluster *in Bacillus amyloliquifaciens* (cyan, PDB 6MHK) and KS6 of the bacillaene cluster from *Bacillus subtilis* (purple, PDB 5ERF), the overall RMSDs (Cα) are below 1.1 Å. (D) Side view and (E) view onto the LINKS interface of superimposed KS domains.

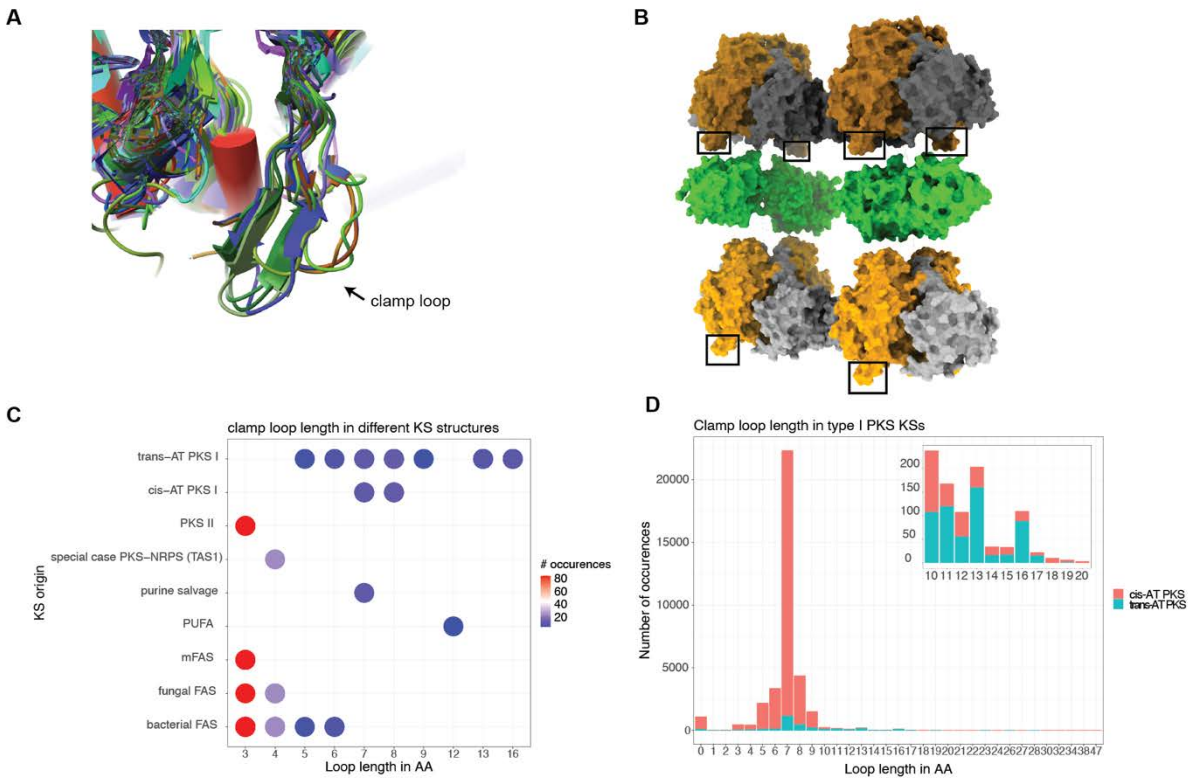

Fig. S5.

Analysis of clamp loops in KS domains. (A) Zoom-in view onto the clamp loop region of structurally characterized KSs (listed in Data S1) superimposed on KS3. The clamp loop is positioned between residues 920 and 936 in KS3, locating it just before the core β-strand 13. **(B)** Localization of the KS clamp loops in the assembled K3DAK4 filament (two clamp loops for KS3 are obscured from vision by other parts of the molecule). **(C)** Visualization of clamp loop lengths for different classes of KS domains derived from the structural comparison shown in (A); only KS domains of *trans*-AT PKSs exhibit long clamp loops (13/16 residues), but short clamp loops also occur in *trans*-AT PKS. In type II PKS, mammalian and fungal fatty acid synthases, the clamp loop is completely absent. Most *cis*-AT PKS KSs feature a shorter clamp loop, similar to that of short clamp loops in *trans*-AT PKS KSs. **(D)** Analysis of clamp loops across a set of ~36'500 type I PKS KS sequences aligned using ClustalOmega (*49*). There is a distinct preference towards loops of 7 AAs length in both *cis*-and *trans*-AT KSs. Longer loops are less common in both systems; however, they are much more common in *trans*-AT PKSs than in *cis*-AT PKSs. Sequences and their classification as *cis*- or *trans*-AT are derived from "Antismash" (*48*) and have not been checked individually for misclassifications. The insert shows a zoom of the 10-20 AA loop-length region.

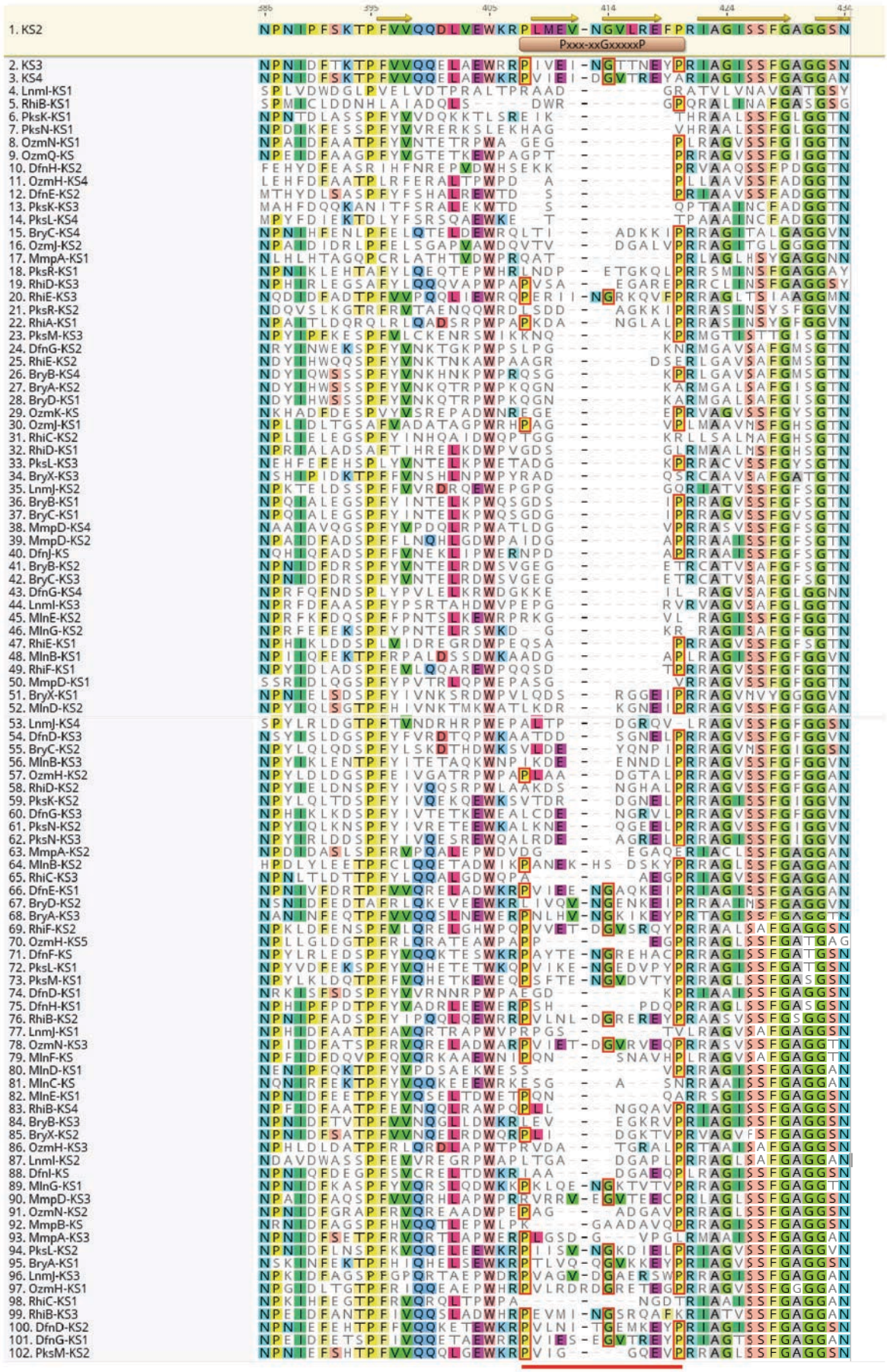
1. KS2
Pxxx-xxGxxxxxP
2. KS3
3. KS4
4. LnmI-KS1
5. RhiB-KS1
6. PksK-KS1
7. PksN-KS1
8. OzmN-KS1
9. OzmQ-KS
10. DfnH-KS2
11. OzmH-KS4
12. DfnE-KS2
13. PksK-KS3
14. PksL-KS4
15. BryC-KS4
16. OzmJ-KS2
17. MmpA-KS1
18. PksR-KS1
19. RhiD-KS3
20. RhiE-KS3
21. PksR-KS2
22. RhiA-KS1
23. PksM-KS3
24. DfnG-KS2
25. RhiE-KS2
26. BryB-KS4
27. BryA-KS2
28. BryD-KS1
29. OzmK-KS
30. OzmJ-KS1
31. RhiC-KS2
32. RhiD-KS1
33. PksL-KS3
34. BryX-KS3
35. LnmJ-KS2
36. BryB-KS1
37. BryC-KS1
38. MmpD-KS4
39. MmpD-KS2
40. DfnJ-KS
41. BryB-KS2
42. BryC-KS3
43. DfnG-KS4
44. LnmI-KS3
45. MlnE-KS2
46. MlnG-KS2
47. RhiE-KS1
48. MlnB-KS1
49. RhiF-KS1
50. MmpD-KS1
51. BryX-KS1
52. MlnD-KS2
53. LnmJ-KS4
54. DfnD-KS3
55. BryC-KS2
56. MlnB-KS3
57. OzmH-KS2
58. RhiD-KS2
59. PksK-KS2
60. DfnG-KS3
61. PksN-KS2
62. PksN-KS3
63. MmpA-KS2
64. MlnB-KS2
65. RhiC-KS3
66. DfnE-KS1
67. BryD-KS2
68. BryA-KS3
69. RhiF-KS2
70. OzmH-KS5
71. DfnF-KS
72. PksL-KS1
73. PksM-KS1
74. DfnD-KS1
75. DfnH-KS1
76. RhiB-KS2
77. LnmJ-KS1
78. OzmN-KS3
79. MlnF-KS
80. MlnD-KS1
81. MlnC-KS
82. MlnE-KS1
83. RhiB-KS4
84. BryB-KS3
85. BryX-KS2
86. OzmH-KS3
87. LnmI-KS2
88. DfnI-KS
89. MlnG-KS1
90. MmpD-KS3
91. OzmN-KS2
92. MmpB-KS
93. MmpA-KS3
94. PksL-KS2
95. BryA-KS1
96. LnmJ-KS3
97. OzmH-KS1
98. RhiC-KS1
99. RhiB-KS3
100. DfnD-KS2
101. DfnG-KS1
102. PksM-KS2

Fig. S6.

Sequence comparison of the clamp loop region across a subset of KS domains. KS domains with long clamp loops (13-16AA) from *trans*-AT PKSs show a recurring PxxxxxGxxxxP motif, presumably linked to ß-hairpin formation.

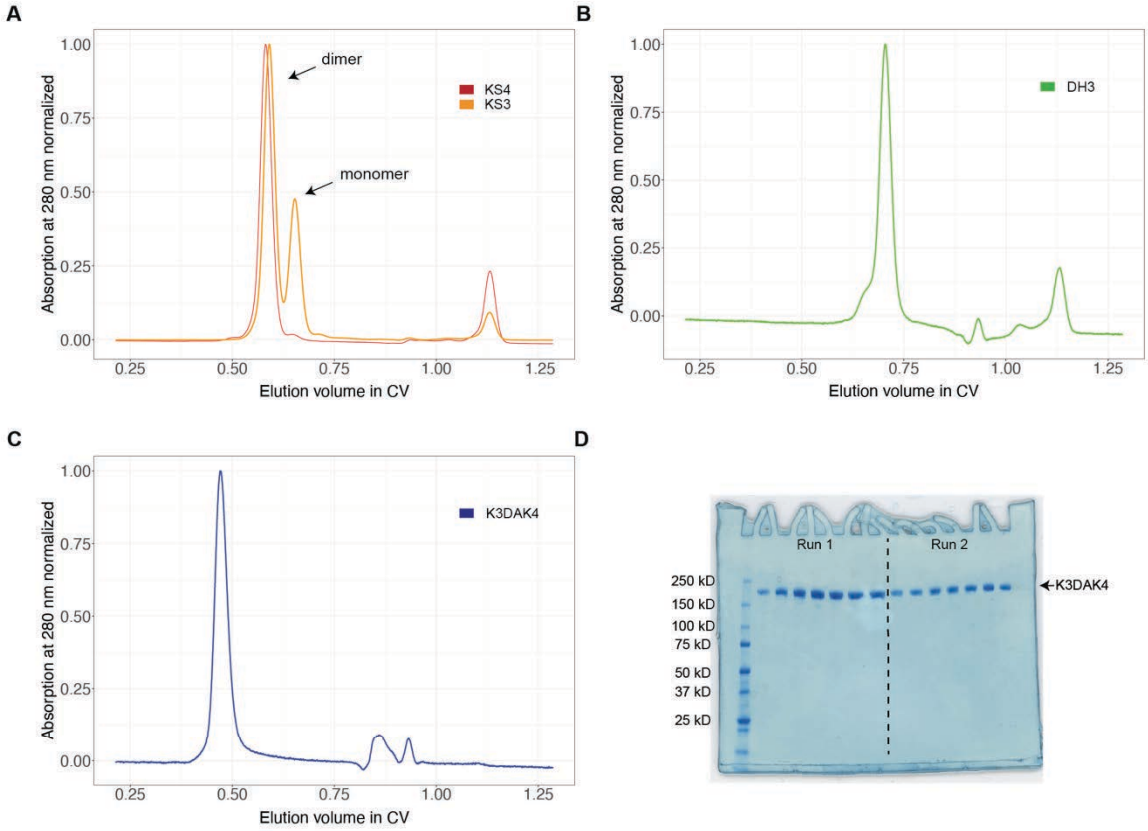

Fig. S7.

Size-exclusion chromatography of KS3, KS4, DH and the intact K3DAK4 bimodule core. Chromatography was performed on an Superdex 200 increase column at 23°C. **(A)** Overlay of KS3 and KS4 elution. KS4 elutes as a stable dimer, whereas KS3 elutes in a monomer-dimer equilibrium. **(B)** Monodisperse elution of DH3. (**C**) The K3DAK4 bimodule elutes as a monodisperse dimer. **(D)** Uncropped SDS-PAGE analysis of two individual size-exclusion chromatography runs on an Superdex 200 column of a K3DAK4 purification.

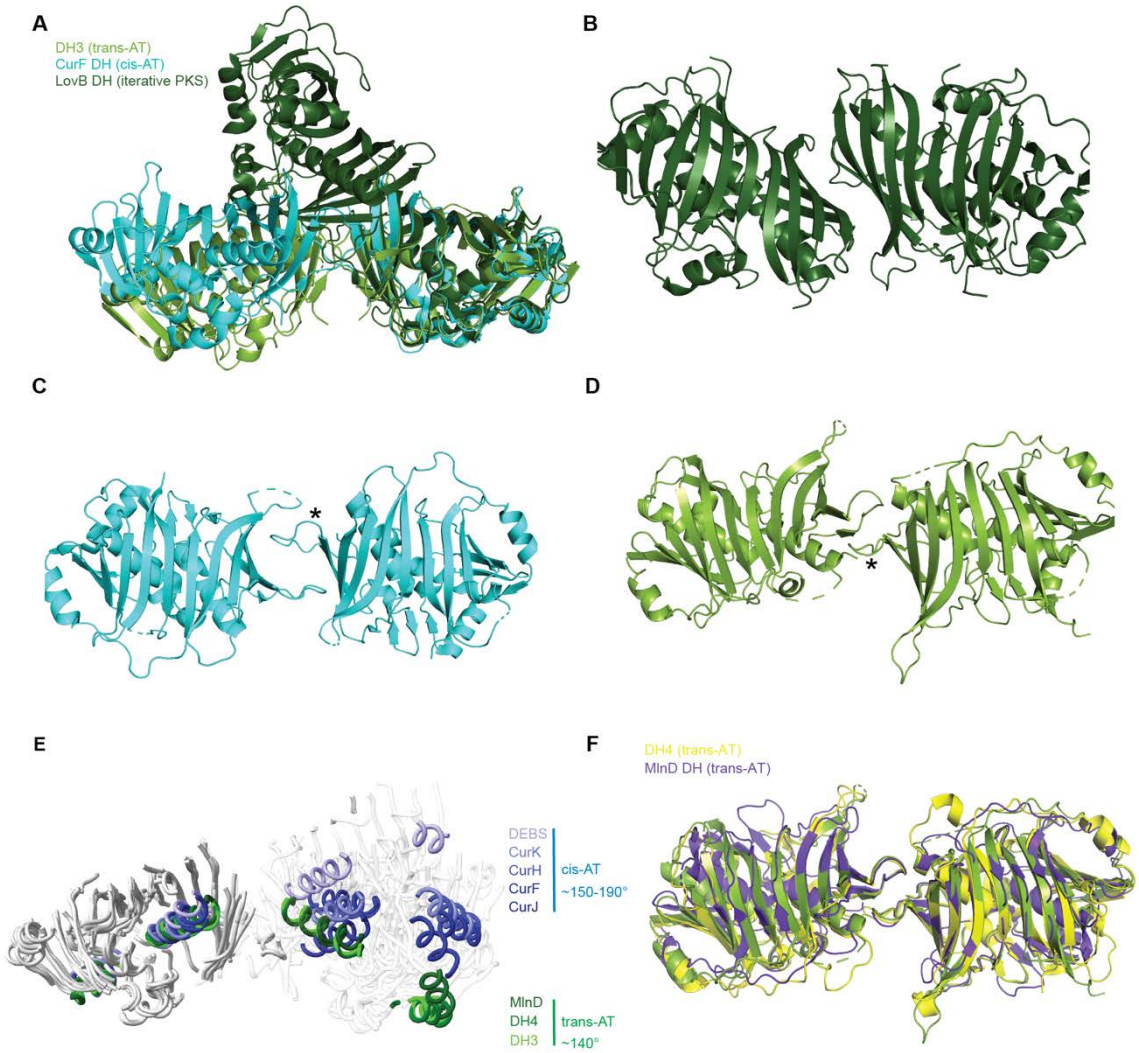

Fig. S8.

DH dimerization interfaces in different types of PKS. The *trans*-AT PKS DH3 from BGC11 (reported here) is shown in green, the DH from the *cis*-AT PKS CurF (PDB 3KG6) in cyan and the DH of the iterative PKS LovB (PDB 7CPX) in dark green. **(A)** Comparison of all three DH dimers based on superposition of the right protomer. The two *cis*-and *trans*-AT PKS DH dimers exhibit similar interdomain angles, clearly distinct from that of the iterative PKS LovB. **(B-D)** Depiction of the individual dimerization interfaces. While *cis*-and *trans*-AT PKS rely on the N-terminal loop to dimerize, the iterative LovB PKS dimerizes via interactions of the beta-strands of the double-hotdog fold. Asterisks in (C) and (D) depict the N-terminal loop of the aligned protomer in DH3 and CurF DH. *cis*-AT DHs dimerize via their N-terminal loops docking to reach β1-2 and β5-6 from one side, whereas the loops of *trans*-AT PKS DHs approach from the other side. **(E)** Shape comparison of DH dimer from *cis*- and *trans*-AT PKS. DH dimers are superimposed based on the left protomer. Superpositioning, visualization and calculation of interdomain angles is based only on ordered secondary structure elements. Two conserved helices in each protomer are shown in color to highlight differences in protomer orientation, these helices are colored in shades of blue for *cis*-AT PKS and in shades of green for *trans*-AT PKS DH domains as indicated. (F) Comparison of the dimerization mode of three trans-AT PKS DH domains, DH3 (green),

DH4 (yellow, PDB 5HQW) of BGC11 studied here, and MlnD (purple, PDB 5IL5). DH dimers are superimposed based on the left protomer only. The mode dimerization is highly similar for all three DH domains.

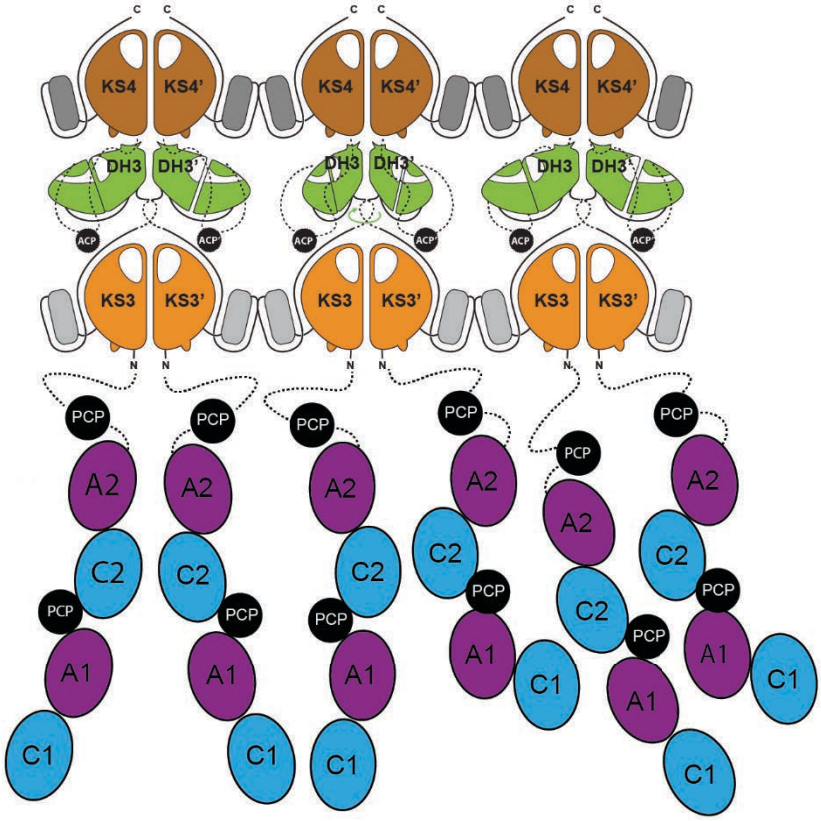

Fig. S9.

Schematic model of a hybrid NRPS-PKS assembly line. C represents the NRPS condensation domain, whereas A refers to the adenylation domain of NRPS and PCP to the peptidyl carrier protein of NRPS. While large conformational rearrangements are common in NRPS and necessary for catalytic activity, the PKS meshwork restrains conformational space.

Table S1.

X-ray crystallographic data collection and refinement statistics

	KS3 PDB 7ZM9	DH3 PDB 7ZMF	KS4 PDB 7ZMC
Data collection			
Space group	C 1 2 1	P 21 21 2	C 2 2 21
Cell dimensions			
a, *b*, *c* (Å)	126.78 92.69 99.27	72.38 197.36 39.17	73.56 191.79 293.81
α, β, γ (°)	90 92.54 90	90 90 90	90 90 90
Resolution (Å)	60.48 - 1.62 (1.66 - 1.62)[1]	48.68 - 2.21 (2.26 - 2.21)[1]	48.97 - 3.10 (3.20 - 3.10)[1]
R_{sym} or R_{merge}	4.8 (171.2)	14.68 (205.2)	28.3 (460.4)
I / σI	16.65 (1.01)	14.37 (1.33)	8.44 (0.53)
Completeness (%)	99.70 (99.40)	99.0 (98.30)	99.55 (98.11)
Redundancy	6.80 (7.00)	13.39 (13.69)	13.61 (13.66)
CC1/2	100.0 (62.5)	99.9 (66.7)	99.9 (41.9)
Refinement			
Resolution (Å)	60.48 - 1.62 (1.66 - 1.62)	48.68 - 2.21 (2.29 - 2.21)	48.97 - 3.10 (3.21 - 3.10)
No. reflections	145,063	28,951	38,108
R_{work} / R_{free}	0.15/ 0.17	0.22/ 0.25	0.26/ 0.29
No. atoms			
Protein	9,260	8,510	16,679
Ligand/ion	110	16	0
Water	518	78	0
B-factors (Å²)			
Protein	53.8	58.88	125.72
Ligand/ion	76.28	68.17	-
Water	54.87	55.85	-
R.m.s. deviations			
Bond lengths (Å)	0.012	0.005	0.004
Bond angles (°)	1.54	0.60	0.67
Clashscore	2.62	2.10	5.6

1 Dataset is derived from a single crystal.

*Values in parentheses are for highest-resolution shell.

Table S2.

Cryo-EM data collection, refinement and validation statistics

	KS3 EMDB-14795 PDB 7ZMD	KS4 EMDB-14793 PDB 7ZMA	K3DAK4 EMDB-14945 PDB 7ZSK
Data collection and processing			
Magnification	130,000	130,000	130,000
Voltage (kV)	300	300	300
Electron exposure (e–/Å^2)	67	67	67
Defocus range (μm)	1.5-2.5	1.5-2.5	0.8-1.8
Pixel size (Å)	1.058	1.058	2.116
Symmetry imposed	C2	C2	C1
Initial particle images (no.)	3,332,737	3,332,737	1,288,160
Final particle images (no.)	182,315	171,149	293,391
Map resolution (Å)	2.93	2.90	6.80
FSC threshold	0.143	0.143	0.143
Map resolution range (Å)	4.5-2.5	4.5-2.5	6.0-14.5
Refinement type	Coordinate	Coordinate	Rigid-body
Initial model used (PDB code)	Xray KS3 (7ZM9)	Xray KS4 (7ZMC)	Combined
Model resolution (Å)	3.10	3.20	10.4
FSC threshold	0.5	0.5	0.5
Map sharpening *B* factor (Å^2)	90.8	94.7	387.4
Model composition			
Non-hydrogen atoms	9,294	9,484	46,128
Protein residues	1,191	1,214	5,916
Ligands	0	0	0
B factors (Å^2)			
Protein	32.28	40.90	735.96
Ligand	-	-	-
R.m.s. deviations			
Bond lengths (Å)	0.002	0.003	
Bond angles (°)	0.522	0.894	
Validation			
MolProbity score	1.27	1.61	
Clashscore	3.75	4.52	
Poor rotamers (%)	0.00	0.00	
Ramachandran plot			
Favored (%)	97.47	94.46	
Allowed (%)	2.53	5.54	
Disallowed (%)	0.00	0.00	

Movie S1.

Overview of the bimodule core filament maps and model. Blend-over between experimental map and model-based cartoon and surface representations in different views with transition to hypothesized 2D-mesh formation, as described in the main text.

Data S1.

Analysis of clamp loop length based on experimental KS structures. Includes pdb id, type of the BGC, organism, clamp loop length and boundary residues. Structures are shown in Fig S5 A and clamp loop length distribution is visualized in Fig S5 C. For selection of clamp loop boundaries KS3 residues 920 and 936 were used as reference.

REFERENCES AND NOTES

1. D. J. Newman, G. M. Cragg, Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83**, 770–803 (2020).

2. J. Piel, Metabolites from symbiotic bacteria. *Nat. Prod. Rep.* **26**, 338–362 (2009).

3. C. Hertweck, The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed. Engl.* **48**, 4688–4716 (2009).

4. D. C. Gay, P. J. Spear, A. T. Keatinge-Clay, A double-hotdog with a new trick: Structure and mechanism of the trans-acyltransferase polyketide synthase enoyl-isomerase. *ACS Chem. Biol.* **9**, 2374–2381 (2014).

5. T. Bretschneider, J. B. Heim, D. Heine, R. Winkler, B. Busch, B. Kusebauch, T. Stehle, G. Zocher, C. Hertweck, Vinylogous chain branching catalysed by a dedicated polyketide synthase module. *Nature* **502**, 124–128 (2013).

6. J. W. Labonte, C. A. Townsend, Active site comparisons and catalytic mechanisms of the hotdog superfamily. *Chem. Rev.* **113**, 2182–2204 (2013).

7. R. W. Broadhurst, D. Nietlispach, M. P. Wheatcroft, P. F. Leadlay, K. J. Weissman, The structure of docking domains in modular polyketide synthases. *Chem. Biol.* **10**, 723–731 (2003).

8. J. Zeng, D. T. Wagner, Z. Zhang, L. Moretto, J. D. Addison, A. T. Keatinge-Clay, Portability and structure of the four-helix bundle docking domains of trans-acyltransferase modular polyketide synthases. *ACS Chem. Biol.* **11**, 2466–2474 (2016).

9. G. J. Dodge, F. P. Maloney, J. L. Smith, Protein-protein interactions in “cis-AT” polyketide synthases. *Nat. Prod. Rep.* **35**, 1082–1096 (2018).

10. E. J. Helfrich, J. Piel, Biosynthesis of polyketides by trans-AT polyketide synthases. *Nat. Prod. Rep.* **33**, 231–316 (2016).

11. K. J. Weissman, The structural biology of biosynthetic megaenzymes. *Nat. Chem. Biol.* **11**, 660–670 (2015).

12. J. Piel, Biosynthesis of polyketides by trans-AT polyketide synthases. *Nat. Prod. Rep.* **27**, 996–1047 (2010).

13. T. Nguyen, K. Ishida, H. Jenke-Kodama, E. Dittmann, C. Gurgui, T. Hochmuth, S. Taudien, M. Platzer, C. Hertweck, J. Piel, Exploiting the mosaic structure of trans-acyltransferase polyketide synthases for natural product discovery and pathway dissection. *Nat. Biotechnol.* **26**, 225–233 (2008).

14. D. C. Gay, D. T. Wagner, J. L. Meinke, C. E. Zogzas, G. R. Gay, A. T. Keatinge-Clay, The LINKS motif zippers trans-acyltransferase polyketide synthase assembly lines into a biosynthetic megacomplex. *J. Struct. Biol.* **193**, 196–205 (2016).

15. S. Dutta, J. R. Whicher, D. A. Hansen, W. A. Hale, J. A. Chemler, G. R. Congdon, A. R. Narayan, K. Håkansson, D. H. Sherman, J. L. Smith, G. Skiniotis, Structure of a modular polyketide synthase. *Nature* **510**, 512–517 (2014).

16. D. P. Cogan, K. Zhang, X. Li, S. Li, G. D. Pintilie, S. H. Roh, C. S. Craik, W. Chiu, C. Khosla, Mapping the catalytic conformations of an assembly-line polyketide synthase module. *Science* **374**, 729–734 (2021).

17. S. R. Bagde, I. I. Mathews, J. C. Fromme, C. Y. Kim, Modular polyketide synthase contains two reaction chambers that operate asynchronously. *Science* **374**, 723–729 (2021).

18. D. A. Herbst, R. P. Jakob, F. Zahringer, T. Maier, Mycocerosic acid synthase exemplifies the architecture of reducing polyketide synthases. *Nature* **531**, 533–537 (2016).

19. J. Wang, J. Liang, L. Chen, W. Zhang, L. Kong, C. Peng, C. Su, Y. Tang, Z. Deng, Z. Wang, Structural basis for the biosynthesis of lovastatin. *Nat. Commun.* **12**, 867 (2021).

20. C. Khosla, Y. Tang, A. Y. Chen, N. A. Schnarr, D. E. Cane, Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu. Rev. Biochem.* **76**, 195–221 (2007).

21. J. Staunton, P. Caffrey, J. F. Aparicio, G. A. Roberts, S. S. Bethell, P. F. Leadlay, Evidence for a double-helical structure for modular polyketide synthases. *Nat. Struct. Biol.* **3**, 188–192 (1996).

22. A. L. Edwards, T. Matsui, T. M. Weiss, C. Khosla, Architectures of whole-module and bimodular proteins from the 6-deoxyerythronolide B synthase. *J. Mol. Biol.* **426**, 2229–2245 (2014).

23. P. D. Straight, M. A. Fischbach, C. T. Walsh, D. Z. Rudner, R. Kolter, A singular enzymatic megacomplex from *Bacillus subtilis*. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 305–310 (2007).

24. M. Klaus, M. Grininger, Engineering strategies for rational polyketide synthase design. *Nat. Prod. Rep.* **35**, 1070–1081 (2018).

25. E. J. N. Helfrich, R. Ueoka, A. Dolev, M. Rust, R. A. Meoded, A. Bhushan, G. Califano, R. Costa, M. Gugger, C. Steinbeck, P. Moreno, J. Piel, Automated structure prediction of trans-acyltransferase polyketide synthase products. *Nat. Chem. Biol.* **15**, 813–821 (2019).

26. T. A. Scott, J. Piel, The hidden enzymology of bacterial natural product biosynthesis. *Nat Rev Chem* **3**, 404–425 (2019).

27. A. Nivina, K. P. Yuet, J. Hsu, C. Khosla, Evolution and diversity of assembly-line polyketide synthases. *Chem. Rev.* **119**, 12524–12547 (2019).

28. J. M. Reimer, M. Eivaskhani, I. Harb, A. Guarne, M. Weigt, T. M. Schmeing, Structures of a dimodular nonribosomal peptide synthetase reveal conformational flexibility. *Science* **366**, (2019).

29. J. M. Reimer, M. N. Aloise, P. M. Harrison, T. M. Schmeing, Synthetic cycle of the initiation module of a formylating nonribosomal peptide synthetase. *Nature* **529**, 239–242 (2016).

30. E. J. Drake, B. R. Miller, C. Shi, J. T. Tarrasch, J. A. Sundlov, C. L. Allen, G. Skiniotis, C. C. Aldrich, A. M. Gulick, Structures of two distinct conformations of holo-non-ribosomal peptide synthetases. *Nature* **529**, 235–238 (2016).

31. S. Bonhomme, A. Dessen, P. Macheboeuf, The inherent flexibility of type I non-ribosomal peptide synthetase multienzymes drives their catalytic activities. *Open Biol.* **11**, 200386 (2021).

32. J. Sambrook, D. W. Russell, Purification of nucleic acids by extraction with phenol:chloroform. *CSH Protoc.* **2006**, pdb.prot4455 (2006).

33. D. J. Fitzgerald, P. Berger, C. Schaffitzel, K. Yamada, T. J. Richmond, I. Berger, Protein complex expression by using multigene baculoviral vectors. *Nat. Methods* **3**, 1021–1032 (2006).

34. E. R. Geertsma, R. Dutzler, A versatile and efficient high-throughput cloning tool for structural biology. *Biochemistry* **50**, 3272–3278 (2011).

35. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).

36. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).

37. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).

38. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

39. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).

40. A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).

41. S. H. W. Scheres, RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).

42. G. M. Boratyn, C. Camacho, P. S. Cooper, G. Coulouris, A. Fong, N. Ma, T. L. Madden, W. T. Matten, S. D. McGinnis, Y. Merezhuk, Y. Raytselis, E. W. Sayers, T. Tao, J. Ye, I. Zaretskaya, BLAST: A more efficient report with usability improvements. *Nucleic Acids Res.* **41**, W29–W33 (2013).

43. D. A. Vander Wood, A. T. Keatinge-Clay, The modules of trans-acyltransferase assembly lines redefined with a central acyl carrier protein. *Proteins* **86**, 664–675 (2018).

44. L. Zhang, T. Hashimoto, B. Qin, J. Hashimoto, I. Kozone, T. Kawahara, M. Okada, T. Awakawa, T. Ito, Y. Asakawa, M. Ueki, S. Takahashi, H. Osada, T. Wakimoto, H. Ikeda, K. Shin-Ya, I. Abe,

Characterization of giant modular PKSs provides insight into genetic mechanism for structural diversification of aminopolyol polyketides. *Angew. Chem. Int. Ed. Engl.* **56**, 1740–1745 (2017).

45. E. Krissinel, K. Henrick, Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2256–2268 (2004).

46. E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).

47. K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S. Y. Lee, M. H. Medema, T. Weber, antiSMASH 5.0: Updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* **47**, W81–W87 (2019).

48. K. Blin, S. Shaw, S. A. Kautsar, M. H. Medema, T. Weber, The antiSMASH database version 3: Increased taxonomic coverage and new query features for modular enzymes. *Nucleic Acids Res.* **49**, D639-D643 (2021).

49. F. Madeira, M. Pearce, A. R. N. Tivey, P. Basutkar, J. Lee, O. Edbali, N. Madhusoodanan, A. Kolesnikov, R. Lopez, Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* **50**, W276–W279 (2022).