The Missing Linker: A Dimerization Motif Located within Polyketide Synthase Modules

Jianting Zheng,[†,§] Christopher D. Fage,[†,§] Borries Demeler,[‡] David W. Hoffman,[†] and Adrian T. Keatinge-Clay*,[†]

[†]Department of Chemistry & Biochemistry, The University of Texas at Austin, 1 University Station A5300, Austin, Texas 78712, United States

[‡]Department of Biochemistry, The University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, Texas 78229, United States

Ⓢ Supporting Information

ABSTRACT: The dimerization of multimodular polyketide synthases is essential for their function. Motifs that supplement the contacts made by dimeric polyketide synthase enzymes have previously been characterized outside the boundaries of modules, at the N- and C-terminal ends of polyketide synthase subunits. Here we describe a heretofore uncharacterized dimerization motif located within modules. The dimeric state of this dimerization element was elucidated through the 2.6 Å resolution crystal structure of a fragment containing a dimerization element and a ketoreductase. The solution structure of a standalone dimerization element was revealed by nuclear magnetic resonance spectroscopy to be consistent with that of the crystal structure, and its dimerization constant was measured through analytical ultracentrifugation to be ∼20 μM. The dimer buries ∼990 Å² at its interface, and its C-terminal helices rigidly connect to ketoreductase domains to constrain their locations within a module. These structural restraints permitted the construction of a common type of polyketide synthase module.
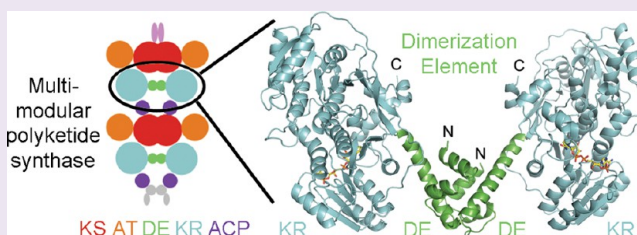
Modular polyketide synthases (PKSs) are megadalton enzymatic assembly lines that biosynthesize complex polyketides such as the antibacterial erythromycin, the antifungal amphotericin, and the insecticide spinosyn.[1−4] During the synthesis of a polyketide, each chain extension reaction is catalyzed by a separate module within the PKS. Thus, the three domains necessary for chain extension are present in all module types: an acyltransferase (AT) that selects an extender unit, a ketosynthase (KS) that catalyzes the condensation of the extender unit with a growing polyketide chain, and an acyl carrier protein (ACP) that shuttles an extender unit or an extended polyketide chain between active sites of the synthase. Within modules, processing enzymes may also be present, such as a ketoreductase (KR) that reduces the β-keto group of the newly extended polyketide chain to a β-hydroxy group, a dehydratase (DH) that removes the β-hydroxy group to form a double bond, and an enoylreductase (ER) that reduces the double bond.

Multimodular PKS assembly lines are dimeric, in contrast to the related monomeric nonribosomal peptide synthetase (NRPS) assembly lines (Figure 1).[5,6] However, the only PKS enzyme that requires a dimeric oligomerization state is KS, and the only other enzyme that makes contact across the 2-fold axis of modules is DH.[7−9] The other domains commonly within modules, AT, KR, ER, and ACP, are monomeric and loop out from the 2-fold axis of the synthase. Accessory motifs have been identified that aid KS and DH in dimerizing the polypeptides of

a PKS into its homodimeric subunits: a coiled coil is frequently located at the N-terminal end of a subunit, a four-helix bundle is often present at the C-terminal end, and a dimeric thioesterase (TE) that cleaves the elongated polyketide from the assembly line is usually positioned at the C-terminal end of the final PKS subunit.[10−12]

Here we report a dimerization motif located within PKS modules. This ∼55-residue dimerization element (DE) is immediately N-terminal to the KR domain in ∼50% of modules that contain KR as their sole processing enzyme (β-modules).[4] The 2.6 Å resolution crystal structure of a fragment from the third module of the spinosyn PKS that contains DE and KR, Spn(DE+KR)3, reveals the architecture of the DE dimer as well as a rigid connection between the last helix of DE and the KR domain. NMR analysis of a standalone DE from the first module of the spinosyn PKS is consistent with the crystal structure. Analytical ultracentrifugation revealed its dimerization constant to be 19.9 ± 1.5 μM. By positioning the C-terminal ends of a structurally characterized KS+AT didomain in proximity with the N-terminal ends of the Spn(DE+KR)3 didomain, a model of a β-module was constructed. Thus, the biophysical characterization of DE has also helped to elucidate the domain organization of a common type of PKS module.
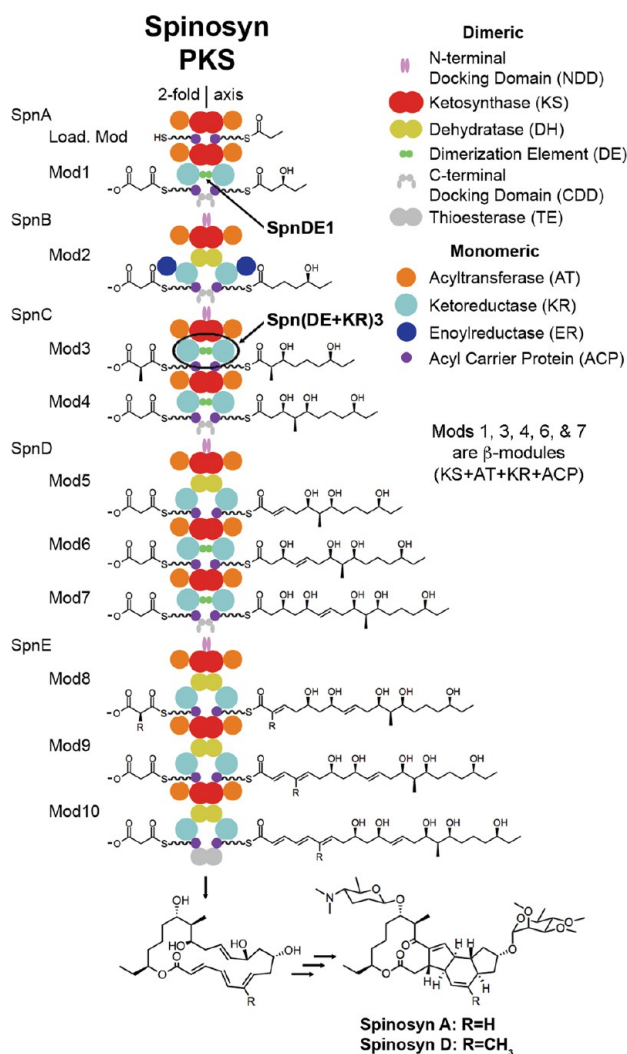
Figure 1. Subunit architecture of the spinosyn PKS. Additional dimeric interfaces aid KSs in maintaining the dimeric state of PKS subunits. These interfaces are provided by the enzymes DH and TE, as well as the dimerization motifs of NDD, CDD, and the newly identified DE. The two DE-containing fragments characterized here, SpnDE1 and Spn(DE+KR)3, are from the PKS that helps synthesize the spinosyn insecticides using malonyl and methylmalonyl building blocks.

■ RESULTS AND DISCUSSION

Crystal Structure of a DE+KR Dimer. A region containing the DE and KR domains of the third module of the spinosyn PKS (Spn(DE+KR)3; SpnC residues 913−1471, numbered here 1−559) was heterologously expressed in *E. coli* BL21-(DE3) and purified to homogeneity.[13] Crystals of Spn(DE +KR)3 diffracted to 2.6 Å resolution, and its structure was solved by molecular replacement using AmpKR2 (PDB code: 3MJS) as a search model (Figure 2).[14] The asymmetric unit contains two Spn(DE+KR)3 dimers. Well-defined electron density at the N-terminal ends of the KR domains enabled the DE domains to be built. The KR domain is similar to those already reported, particularly AmpKR2 (1.1 Å C_α rmsd), another A1-type KR. As in that structure, a conserved methionine (Met371) latches over bound NADP$^+$ and helps order the adjacent lid helix.[14−16] Searches on the Dali and CATH servers indicate that DE possesses a novel fold.[17,18]

Each DE monomer contains three helices (α1, α2, and α3; residues 12−21, 26−33, and 41−62, respectively) connected by

two loops (loop I and loop II; residues 22−25 and 34−40), nearly as predicted by the PSIPRED server from its sequence.[19,20] Within each monomer, α2 lies out of the plane formed by α1 and α3. Through the association of two monomers, a V-shaped dimer is generated with an angle of ~65° between the α3 helices (Figure 2A). A surface area of ~990 Å^2 is buried at the dimer interface.[21] The main dimeric contact is between the N-terminal end of α3 from one monomer and a pocket created by the helices of the other monomer, specifically, between residues L41, T44, L45, L48, W51 of α3 from one monomer and residues F17, W18, V21, L26, V29, L33, L35, L41, T44, L48 from the other monomer. The most significant hydrophobic contacts are made by the nearly invariant F17, W18, and W51.

Several residues from α3 mediate the connection between DE and KR (Figure 2B). Interactions with αG of the KR domain include hydrophobic contacts between E57, R58, A61, and L483, as well as a hydrogen bond between E57 and S480. At the C-terminal end of α3, W64 inserts into a hydrophobic pocket created by M476, P477, S480, I498, and D500, and R65 forms a salt bridge with E493. Rigid residues connect the end of DE to the beginning of the KR domain (α3 and β1, respectively).

The DE dimer possesses several mysterious features (Figure 2C). Its hydrophobic core appears somewhat loosely packed, perhaps due in part to interactions made by the invariant residue W18. The indole NHs of W18, and hence their dipoles, point toward one another across the DE 2-fold axis, with the hydrogens separated by only 1.6 Å. Within α3, P46 neither introduces a kink nor interacts with neighboring residues, yet it is completely conserved (Figure 2D). Both of the SpnDE3 dimers in the asymmetric unit are slightly asymmetric, with the third helix (α3) of one monomer relatively bent by ~53° (Supplementary Figure 1). The "straight" and "bent" monomers within each DE dimer show lower structural similarity with one another (0.93 and 1.0 Å C_α rmsd) than the equivalent "straight" and "bent" DE monomers from the other dimer (0.43 and 0.29 Å C_α rmsd, respectively). Usually all or none of the β-modules in a PKS contain a DE; however, some PKSs (e.g., the oleandomycin PKS) are composed of both β-modules that contain DE as well as β-modules that do not. The DE dimer is apparently unique to *cis*-AT PKSs.

Other dimerization motifs that supplement the dimeric contacts made by KS, DH, and TE have been characterized.[10,11] The N-terminal docking domain (NDD) is a ~25-residue coiled coil commonly at the N-terminal ends of subunits. The C-terminal docking domain (CDD) at the C-terminal ends of subunits is composed of three helices from each polypeptide; the first two helices dimerize to form four-helix bundles, while the last helix specifically binds to the NDD of a downstream subunit to form a second four-helix bundle. Thus, the set of nonenzymatic dimerization motifs in PKS subunits is the NDD/CDD pair and the three-helix DE described here.

NMR Analysis of a Standalone DE. The many unusual features of DE prompted us to study an isolated DE domain in solution (Figure 3). Thus, SpnDE1 (SpnA residues 1877−1944, numbered 1−68 here; ~24% identical to SpnDE3) was investigated by NMR spectroscopy (Figure 3A). Like SpnDE3, SpnDE1 is composed of three helices (α1, α2, and α3; residues 14−27, 31−37, and 44−67) and two loops (loop I and loop II; residues 28−30 and 38−43) as determined by NOE cross-peak patterns and PREDITOR (of 2037 unique NOEs observed for the SpnDE1 dimer, 548 were intra-residue and 1489 were inter-
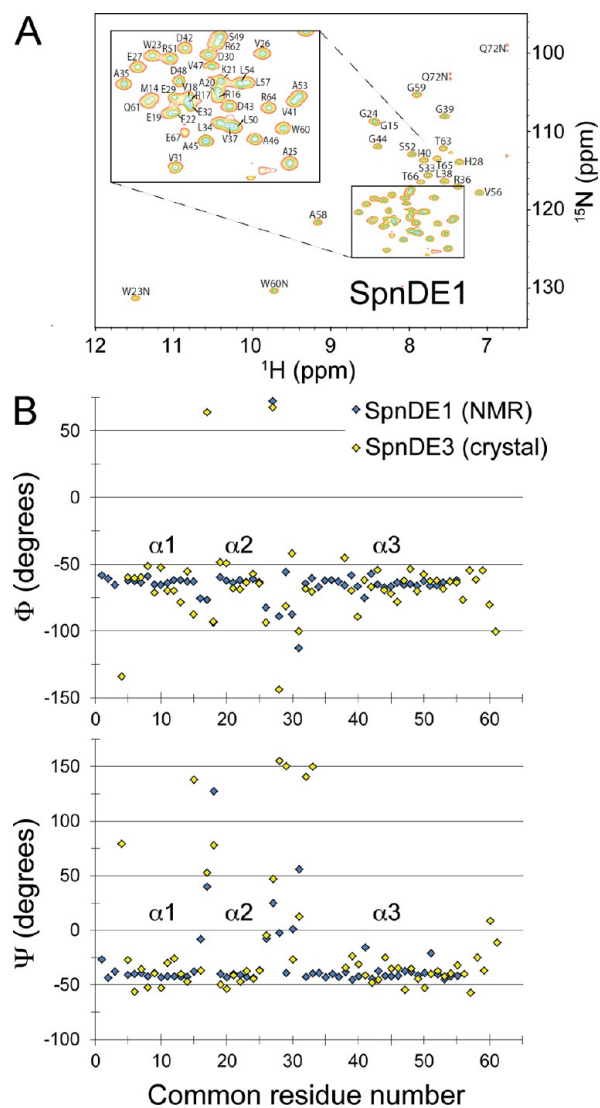
Figure 2. DE+KR structure. (A) In the crystal structure of Spn(DE+KR)3, the V-shaped DE dimer orients two KR domains bound to NADP⁺. (B) DE rigidly connects to the KR domain. No flexible loop is present between α3 of DE and β1 of KR. (C) A stereodiagram of SpnDE3 shows the nearly invariant F17, W18, W51, and other hydrophobic residues that help create the dimer interface of DE. The role of the invariant residue P46 is mysterious. (D) A sequence alignment of the ~55-residue DE from DE-containing β-modules shows the boundaries of its three helices, as observed in the Spn(DE+KR)3 crystal structure (residues labeled with * are displayed in the previous panel). Approximately half of β-modules do not contain a DE. Accession codes (GI): Amp1 and Amp2, 14794905; Amp10−Amp14, 14794893; Con3 and Con4, 74026478; Con6, 74026479; Con10 and Con11, 74026481; Ery1 and Ery2, 134097327; Ery3 and Ery4, 134097329; Ery5 and Ery6, 134097330; Pik1, 3800834; Pik3, 3800835; Pik5, 3800836; Spn1, 348173395; Spn3, 348173392; Spn4, 348173391; Spn6 and Spn7, 348173389.

residue) (Figure 3B).[22] Interestingly, the N-terminal end of α3 in SpnDE1 is six residues longer than the equivalent helix in SpnDE3.

The spectra collected for SpnDE1 are consistent with the conformation of DE observed in the crystal structure. Intermonomer NOEs indicate that the solution structure is indeed dimeric. For example, the V26 methyl protons interact with L54, L57, and A58 side-chain protons at distances of 3−6 Å (V26, L54, and L57 are highly conserved). Such NOEs could not originate from within a monomer unless substantial conformational changes occurred. The atypically high chemical shift of 11.55 ppm for the indole NH of the invariant W23 (equivalent to W18 in SpnDE3) also supports a model in which

those protons are in close proximity, with the symmetry-related tryptophan indole rings deshielding one another.

The relative stabilities of different regions of the SpnDE1 monomer were examined. Two-dimensional ¹⁵N−¹H NMR spectra collected at several pH values revealed the relative amide proton exchange rates. The amide protons of the 21 N-terminal and 10 C-terminal residues are the most susceptible to exchange, as are those of residues D42−A46 on the solvent-exposed loop II; none of these residues are highly conserved. In contrast, the amide protons of E27−D30 in loop I are protected from exchange, consistent with their being steadfastly hydrogen-bonded. Loop I is consistently four residues in length with an aspartate in the fourth position that caps α2.

Figure 3. NMR characterization of an isolated DE. (A) The peaks in the 2D ^{15}N–^{1}H NOESY spectrum of SpnDE1 represent the backbone NHs of structured residues or side-chain NHs. The first 13 and last 5 residues of the SpnDE1 construct are relatively flexible or disordered. (B) The Φ and Ψ dihedral angles for SpnDE1 (calculated by PREDITOR from chemical shifts) and SpnDE3 (obtained from monomer A of the Spn(DE+KR)3 crystal structure) display their common secondary structure. The first ordered residues of SpnDE1 and SpnDE3 are Met14 and Asp11, respectively (common residues 1 and 4, respectively, in the plot).

Contribution of DE to Dimerization. To our knowledge, no dimerization constant for a PKS dimerization motif has been determined. Therefore, we sought to measure the dimerization constant of an isolated DE. We performed sedimentation velocity analysis of SpnDE1 spanning a concentration range from 1.85 to 79.2 μM. Van Holde–Weischet analysis of the results showed a clear mass action shift of the sedimentation coefficient distribution consistent with a rapid monomer–dimer equilibrium (Figure 4).[23] To further identify the oligomerization state, the experimental data were fitted by parsimonious regularization using genetic algorithm analysis[24] followed by Monte Carlo analysis.[25] For each concentration, a major species was identified whose sedimentation coefficients were globally fitted to a binding isotherm. This approach yielded a dimerization constant of 19.9 ± 1.5 μM for SpnDE1.
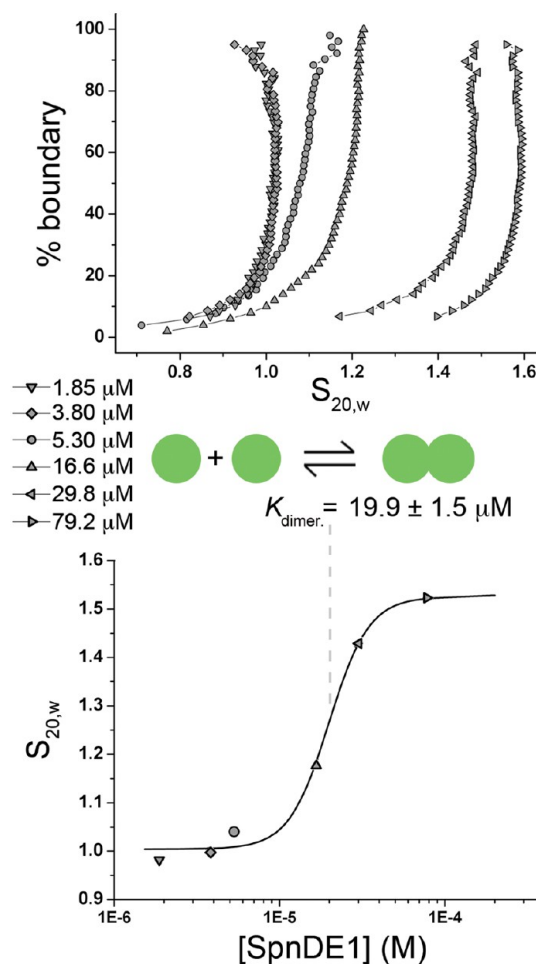
Figure 4. Analytical ultracentrifugation. From the van Holde–Weischet integral sedimentation coefficient distributions of SpnDE1 (1.85–79.2 μM) a mass action effect is clearly visible, as distributions shift to higher sedimentation coefficients with increasing concentration. From a global fit of sedimentation coefficients of the major species in each loading concentration to a monomer–dimer binding isotherm, the SpnDE1 dimerization constant was determined to be 19.9 ± 1.5 μM.

Molecular weight predictions from the genetic algorithm analysis for the lowest and highest concentrations analyzed resulted in a molecular weight of 8.61 kDa for the lowest concentration (7.8 kDa expected for the monomer) and 13.8 kDa for the highest concentration measured (15.6 kDa expected for the dimer), indicating that the fitted association reaction indeed reflects the monomer–dimer reaction.

We sought to observe how the oligomerization state of DE is affected by the addition of a KR domain. While Spn(DE+KR)3 appeared to be dimeric by gel filtration chromatography (Supplementary Figure 2), as had Amp(DE+KR)2 in previous studies,[14] we opted for sedimentation velocity analytical ultracentrifugation to more accurately determine oligomerization states. Spn(DE+KR)3 could not be used for technical reasons, and thus measurements were made with Amp(DE+KR)2. Analytical ultracentrifugation showed Amp(DE+KR)2 was not a dimer, but rather a monomer (~56 kDa as determined by genetic algorithm–Monte Carlo analysis, 56 kDa expected) when measured at 3.8 and 11.5 μM (Supplementary Figure 3). Since Spn(DE+KR)3 is dimeric within the crystal structure, the dimerization constant for a DE
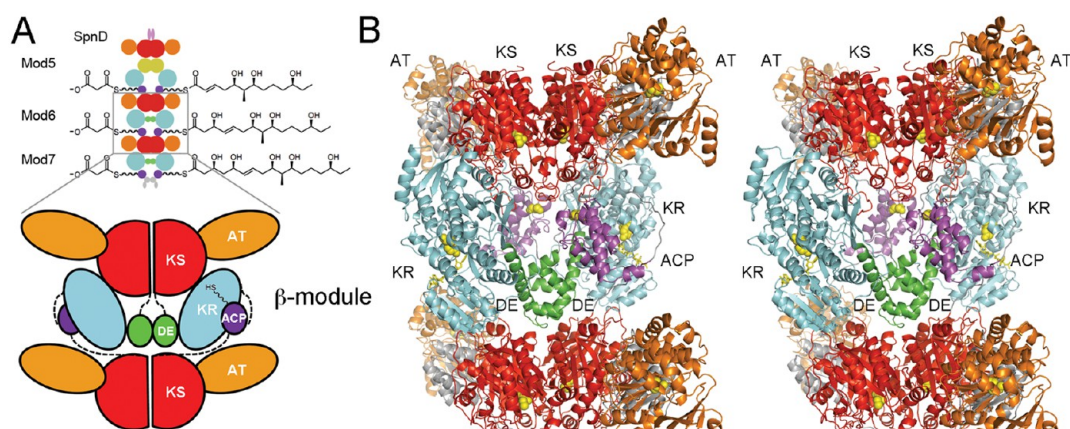
Figure 5. Model of a PKS module. (A) SpnMod6 is a DE-containing β-module (KS+AT+DE+KR+ACP). A schematic shows the architecture of this module, as well as the first part the downstream SpnMod7, in more detail. (B) A stereodiagram shows a β-module, such as SpnMod6, assembled from known structures of a KS+AT dimer, a DE+KR dimer, and two ACPs. Flexible linkers were built between these fragments as well as to the KSs of a subsequent module. The phosphopantetheinyl arm of ACP must interact with the active sites of AT, KS, KR, and the following KS, each shown with a catalytic residue displayed in spheres.

+KR fragment may be closer to crystallographic concentrations (∼10 mM). The implication is that as the number of enzymes within a PKS subunit increases, so must the number of dimerization motifs.

Modeling a β-Module. The crystal structure of Spn(DE+KR)3 enabled us to construct a model of a β-module (Figure 5). The relative positions of a KS+AT dimer, a DE+KR dimer, and two ACP domains were restrained by the average lengths of interdomain peptides (from the set of DE-containing β-modules in the MAPSI database): ∼14-residue, unconserved linkers connect the KS+AT dimer to DEs (the shortest observed was 9 residues in the third module of the geldanamycin PKS), ∼14-residue, unconserved linkers connect KRs to ACPs (the shortest observed was 9 residues in the seventh module of the spinosyn PKS), and ∼19-residue, unconserved linkers connect ACPs to KSs of the next module (the shortest observed was 16 residues in the sixth module of the spinosyn PKS) (Supplementary Figure 4).[7,26,27] The first two linkers were included in the Spn(DE+KR)3 construct that crystallized; that they did not appear in the electron density maps indicates their inherent flexibility. The low sequence conservation of the third linker also suggests its unstructured nature. Thus, flexible linkers of appropriate length were built between the structured fragments.

The model naturally raises several considerations about the assembly line architecture of a PKS: (1) Modules must be as close to one another as possible along the 2-fold axis of a PKS subunit in order to provide the ACP domains, bound by ∼19-residue linkers to the next module, enough slack to access KSs within the same module. (2) AT domains may need to adopt another conformation in order for bound extender units to be accessible to cognate ACP domains (but not to ACPs from the preceding module). (3) A physical interaction may exist between the KS and KR domains via the KR structural subdomain. A hydrophobic groove (in part created by W85 and W184 in SpnKR3) is conserved in the KR structural subdomain and faces KS in the model. An interaction between KS and KR could explain why DE is not always present in β-modules. (4) KRs may be oriented similarly in DH- and ER-containing modules (γ- and δ-modules). The DH dimer can fit between the KR domains, and ERs have sufficient space to loop out

from the KR domains in the orientation observed in the KR+ER didomain structure.[8,9,28]

Through the evolution of processive, multimodular synthases from iterative synthases that may have resembled the modern mammalian fatty acid synthase (FAS), several architectural refinements occurred.[29,30] While the relative orientations of the chain-elongation enzymes, KS and AT, did not change significantly, those of the processing enzymes diverged substantially.[7,31,32] Through the transition of a dimeric to monomeric ER, space necessary for ACPs to access their cognate enzymes became available;[28] however, without the dimeric contacts provided by ER domains, additional interfaces may have been necessary for subunits to remain dimeric. A comparison of multimodular PKSs to FASs shows the loss of the ER dimer interface, but the addition of interfaces at the N-terminal ends of DH and TE, as well as those made by CDD, NDD, and DE. Thus, enough avidity is generated through these dimerization contacts to keep the polypeptides of PKS subunits dimerized.

As the last common structured element within multimodular PKSs to be structurally elucidated, DE relates the orientations of the processing enzymes to those of the chain-elongation enzymes. In the absence of a crystal structure of a PKS module, this small three-helix element represents a very informative final piece in the quintessential puzzle that is the molecular architecture of PKS assembly lines.

■ METHODS

Cloning, Expression, and Purification. The Spn(DE+KR)3 gene fragment was amplified from genomic DNA extracted from *Saccharopolyspora spinosa* (DNeasy Blood & Tissue Kit; Qiagen) with primers 5′-ATCGTAATC*CATATG*GCTGCCTCCGACGAGGCT-CGT-3′ and 5′-TGATTCGAT*CTCGAG*TCACACCAGTTCGTGG-CGTCGTTG-3′ (restriction sites in italics; stop codon underlined), digested with *Nde*I and *Xho*I, and ligated into pET28b (Novagen). *E. coli* BL21(DE3) cells transformed with the expression plasmid were inoculated into LB media containing 50 mg/L kanamycin at 37 °C, grown to OD_{600} = 0.4, and induced with 0.5 mM IPTG. After 12 h at 15 °C, cells were harvested by centrifugation and resuspended in lysis buffer (300 mM NaCl, 10% (v/v) glycerol, 50 mM HEPES pH 7.5). Following sonication, cell debris was removed by centrifugation (30,000g for 45 min). The supernatant was poured over a column of Ni-NTA resin (Qiagen), which was then washed with 50 mL lysis

buffer containing 15 mM imidazole and eluted with 5 mL of lysis buffer containing 300 mM imidazole. Spn(DE+KR)3 was further purified using a Superdex 200 gel filtration column (GE Healthcare Life Sciences) equilibrated with 150 mM NaCl, 10% (v/v) glycerol, 10 mM HEPES pH 7.5. Spn(DE+KR)3 was concentrated to 20 mg/mL in 25 mM NaCl, 10% glycerol (v/v), 1 mM DTT, 10 mM HEPES pH 7.5.

The SpnDE1 gene fragment was amplified from *S. spinosa* genomic DNA with primers 5′-ATGCTTGCAC*CATATG*GCGCTGCGTGAA-TCTTCCGCCGGCGA-3′ and 5′-GTACGAACGG*GAATTC*TCAC-TGATCCATAATGGACTCGGT-3′. Amplicons were digested with *Nde*I and *Eco*RI and ligated into pET28b for transformation into *E. coli* BL21(DE3) cells. SpnDE1 was expressed and purified as described above; however, following elution from the Ni-NTA matrix, SpnDE1 was exchanged into thrombin cleavage buffer (20 mM Tris-HCl pH 8.4, 150 mM NaCl, 2.5 mM CaCl$_2$) and digested with 1 U of thrombin (GE Healthcare Life Sciences) per milligram of protein at 4 °C for 12 h. To remove thrombin and uncleaved protein, the digest was applied to 1 mL Ni-NTA and 0.5 mL benzamidine sepharose 6B (GE Healthcare Life Sciences) and eluted with 15 mL lysis buffer. SpnDE1 was then purified using a Superdex 75 gel filtration column (GE Healthcare Life Sciences) equilibrated with 100 mM sodium phosphate pH 7.0. SpnDE1 samples enriched in ^{15}N and/or ^{13}C were prepared as unlabeled samples were, but in M9 minimal media containing 0.6 g/L [^{15}N]ammonium chloride and/or 2 g/L [^{13}C]$_D$-glucose as the sole nitrogen and/or carbon source.

Amp(DE+KR)2 was prepared as previously described.[14]

Crystallization and Structure Determination. Crystals were grown by sitting drop vapor diffusion at 22 °C by mixing 2 μL of protein solution (20 mg/mL in 25 mM NaCl, 1 mM DTT, 10% (v/v) glycerol, 10 mM HEPES pH 7.5) with 2.5 μL of crystallization buffer (15% (w/v) poly(ethylene glycol) 3350, 0.2 M magnesium acetate, 30% (v/v) glycerol). X-ray diffraction data were collected at 100 K at Beamline 5.0.2 of the Advanced Light Source synchrotron. The collected data were processed with HKL2000.[33] The structure was solved by molecular replacement in Phaser using AmpKR2 (PDB code: 3MJS) as the search model.[34,14] The model was iteratively built in Coot and refined in Refmac (Table 1).[34,35] Local noncrystallographic symmetry (NCS) restraints, automatically generated by Refmac, were used in the refinement for the SpnKR3 domains. Atomic coordinates of the Spn(DE+KR)3 have been deposited in the Protein Data Bank (PDB code: 4IMP).

NMR Spectroscopy. NMR spectra of SpnDE1 were collected at 25 °C using a 500-MHz Varian Inova spectrometer equipped with a triple-resonance probe and a z-axis pulsed-field gradient. NMR samples consisted of 1−7 mM protein in 100 mM sodium phosphate pH 7.0 and 90% H$_2$O/10% D$_2$O. Resonances were assigned using standard three-dimensional (3D) triple-resonance methods. ^{1}H, ^{13}C, and ^{15}N nuclei of backbone atoms were assigned using 3D HNCA, HNCACB, HNCO, and HN(CO)CACB spectra;[36] side-chain resonances were assigned using 2D ^{1}H−^{1}H 2QF-COSY and TOCSY spectra as well as 3D ^{15}N−^{1}H−^{1}H HSQC-TOCSY and ^{13}C−^{1}H−^{1}H HCCH-TOCSY spectra.[37] NOE cross-peaks were identified using 2D ^{1}H−^{1}H NOESY as well as 3D ^{15}N−^{1}H−^{1}H and ^{13}C−^{1}H−^{1}H HSQC-NOESY spectra.[38] Data were processed with NMRPipe.[39] ^{1}H, ^{13}C, and ^{15}N chemical shifts were referenced using internal 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) as described,[40] where the methyl ^{1}H reference frequency of DSS was multiplied by 0.251449530 and 0.101329118 to obtain the 0-ppm frequencies for ^{13}C and ^{15}N, respectively. Helices were identified by their characteristic NOE cross-peak patterns and the chemical-shift values of C$_\alpha$, C$_\beta$, C, and H$_\alpha$ nuclei. Backbone torsion angles φ and ψ were calculated from ^{1}H, ^{13}C, and ^{15}N chemical shifts with PREDITOR.[22]

The rates at which the backbone amide protons of SpnDE1 exchange with solvent were measured using a saturation-transfer method.[41] Uniformly ^{15}N-labeled samples were equilibrated with 100 mM sodium phosphate pH 5.0, 7.0, or 8.0. At each pH, 2D ^{15}N−^{1}H spectra were obtained with and without a 2-s presaturation of the water resonance. The rate at which the amide protons exchange with solvent was estimated by comparing peak heights in spectra obtained

Table 1. Crystallization Data and Refinement Statistics

Data Collection	
space group	P2$_1$
a, b, c (Å)	56.82, 211.69, 101.74
β (deg)	95.3
resolution (Å)	50−2.60(2.64−2.60)
R$_{merge}$	0.056(0.555)
I/σ(I)	20.3(2.7)
completeness (%)	98.2(97.5)
redundancy	3.5(3.5)
Refinement	
resolution (Å)	50−2.60
no. reflections	69.297
R$_{work}$/R$_{free}$	0.204/0.245
no. atoms	
protein	15,485
NADP$^+$	192
water	37
B-factors (Å^2)	
protein	64.9
NADP$^+$	52.1
water	39.0
rms deviations	
bond lengths (Å)	0.009
bond angles (deg)	1.149
Ramachandran plot	
favored/outliers (%)	92.5/1.75
molprobity	
clashscore	19.21 (80th percentile)
overall	3.08

with and without presaturation at each pH. Nuclei in relatively well-ordered and relatively flexible regions of the protein were identified by their positive and negative peaks, respectively, in 2D ^{15}N−^{1}H heteronuclear NOESY spectra. Chemical shifts were deposited at the Biological Magnetic Resonance Bank (accession number 18947).

Gel Filtration Chromatography. A 0.1-mL protein solution of Spn(DE+KR)3 (∼100 μM) was injected onto a Superdex 200 gel filtration column equilibrated with 150 mM NaCl, 10% glycerol (v/v), 10 mM HEPES pH 7.5. The molecular weight was estimated by comparing to standards (Gel Filtration Standard; Biorad Laboratories), as previously described.[14]

Analytical Ultracentrifugation. Sedimentation velocity experiments were performed on a Beckman Optima XL-I at the Center for Analytical Ultracentrifugation of Macromolecular Assemblies (CAUMA) at the University of Texas Health Science Center at San Antonio. Calculations were performed with the UltraScan software[42] at the Texas Advanced Computing Center at the University of Texas at Austin and at the Bioinformatics Core Facility at the University of Texas Health Science Center at San Antonio.[43,44] All samples were measured in a 10 mM phosphate buffer containing 150 mM NaCl. All data were collected at 20 °C and spun at 45 krpm (Amp(KR+DE)2) or 60 krpm (SpnDE1), using titanium 2-channel centerpieces (Nanolytics; Potsdam, Germany). The partial specific volume of Amp(KR+DE)2 and SpnDE1 were determined to be 0.7329 cm^3/g and 0.7237 cm^3/g, respectively, based on protein sequence according to the method by Durchschlag, as implemented in UltraScan. All data were first analyzed by two-dimensional spectrum analysis with simultaneous removal of time-invariant noise[45,46] and then by enhanced van Holde−Weischet analysis[47] and genetic algorithm refinement[48,24] where applicable, followed by Monte Carlo analysis.[25]

Modeling a β-Module. The Ery(KS+AT)5 dimer (PDB code: 2HG4), the Spn(DE+KR)3 dimer (PDB code: 4IMP), and two EryACP2 monomers (PDB code: 2JU2) were positioned with PyMol along a common 2-fold axis.[7,26,49] The DE+KR dimer was arranged so that its N-terminal ends were as close as possible to the C-terminal

ends of the KS+AT dimer and so that its long axis was perpendicular to that of the KS+AT dimer to avoid steric clashes. Another Ery(KS+AT)5 dimer was placed along the common 2-fold axis, parallel to the first KS+AT dimer and as close as possible to the DE+KR dimer, to represent the next module in a PKS subunit and to provide an anchor for the C-terminal linker of ACP. The ACP monomers were positioned between the KS dimers. Polyalanine linkers were built using Coot in order to connect the KS+AT dimer with the DE+KR dimer, the DE+KR dimer with the ACP domains, and the ACP domains with the downstream KS+AT dimer.[35] Linker lengths were determined from sequence alignments of appropriate linkers obtained from MAPSI (Supplementary Figure 3).[50,27] Measurements were made between residues known to be structured; the linker spanning KS+AT and DE was measured between the YWL motif and 5 residues prior to the first conserved DE tryptophan; the linker spanning KR and ACP was measured between 7 residues after a highly conserved proline and helix 0;[26] the linker spanning ACP and KS was measured between the last highly conserved ACP glycine and the first highly conserved KS proline.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

Supplementary Figures 1−4. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: adriankc@utexas.edu.

Author Contributions

§These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Khosla, C., Tang, Y., Chen, A. Y., Schnarr, N. A., and Cane, D. E. (2007) Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu. Rev. Biochem. 76*, 195−221.
(2) Smith, S., and Tsai, S. C. (2007) The type I fatty acid and polyketide synthases: a tale of two megasynthases. *Nat. Prod. Rep. 24*, 1041−1072.

(3) Meier, J. L., and Burkart, M. D. (2009) The chemical biology of modular biosynthetic enzymes. *Chem. Soc. Rev. 38*, 2012−2045.
(4) Keatinge-Clay, A. T. (2012) The structures of type I polyketide synthases. *Nat. Prod. Rep. 29*, 1050−1073.
(5) Sieber, S. A., Linne, U., Hillson, N. J., Roche, E., Walsh, C. T., and Marahiel, M. A. (2002) Evidence for a monomeric structure of nonribosomal peptide synthetases. *Chem. Biol. 9*, 997−1008.
(6) Hur, G. H., Vickery, C. R., and Burkart, M. D. (2012) Explorations of catalytic domains in non-ribosomal peptide synthetase enzymology. *Nat. Prod. Rep. 29*, 1074−1098.
(7) Tang, Y., Kim, C. Y., Mathews, I. I., Cane, D. E., and Khosla, C. (2006) The 2.7-angstrom crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. *Proc. Natl. Acad. Sci. U.S.A. 103*, 11124−11129.
(8) Keatinge-Clay, A. T. (2008) Crystal structure of the erythromycin polyketide synthase dehydratase. *J. Mol. Biol. 384*, 941−953.
(9) Akey, D. L., Razelun, J. R., Tehranisa, J., Sherman, D. H., Gerwick, W. H., and Smith, J. L. (2010) Crystal structures of dehydratase domains from the curacin polyketide biosynthetic pathway. *Structure 18*, 94−105.
(10) Broadhurst, R. W., Nietlispach, D., Wheatcroft, M. P., Leadlay, P. F., and Weissman, K. J. (2003) The structure of docking domains in modular polyketide synthases. *Chem. Biol. 10*, 723−731.
(11) Buchholz, T. J., Geders, T. W., Bartley, F. E., 3rd, Reynolds, K. A., Smith, J. L., and Sherman, D. H. (2009) Structural basis for binding specificity between subclasses of modular polyketide synthase docking domains. *ACS Chem. Biol. 4*, 41−52.
(12) Tsai, S. C., Miercke, L. J., Krucinski, J., Gokhale, R., Chen, J. C., Foster, P. G., Cane, D. E., Khosla, C., and Stroud, R. M. (2001) Crystal structure of the macrocycle-forming thioesterase domain of the erythromycin polyketide synthase: versatility from a unique substrate channel. *Proc. Natl. Acad. Sci. U.S.A. 98*, 14808−14813.
(13) Waldron, C., Matsushima, P., Rosteck, P. R., Jr., Broughton, M. C., Turner, J., Madduri, K., Crawford, K. P., Merlo, D. J., and Baltz, R. H. (2001) Cloning and analysis of the spinosad biosynthetic gene cluster of Saccharopolyspora spinosa. *Chem. Biol. 8*, 487−499.
(14) Zheng, J., Taylor, C. A., Piasecki, S. K., and Keatinge-Clay, A. T. (2010) Structural and functional analysis of A-type ketoreductases from the amphotericin modular polyketide synthase. *Structure 18*, 913−922.
(15) Keatinge-Clay, A. T., and Stroud, R. M. (2006) The structure of a ketoreductase determines the organization of the beta-carbon processing enzymes of modular polyketide synthases. *Structure 14*, 737−748.
(16) Keatinge-Clay, A. T. (2007) A tylosin ketoreductase reveals how chirality is determined in polyketides. *Chem. Biol. 14*, 898−908.
(17) Greene, L. H., Lewis, T. E., Addou, S., Cuff, A., Dallman, T., Dibley, M., Redfern, O., Pearl, F., Nambudiry, R., Reid, A., Sillitoe, I., Yeats, C., Thornton, J. M., and Orengo, C. A. (2007) The CATH domain structure database: new protocols and classification levels give a more comprehensive resource for exploring evolution. *Nucleic Acids Res. 35*, D291−297.
(18) Holm, L., and Rosenstrom, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res. 38*, W545−549.
(19) Bryson, K., McGuffin, L. J., Marsden, R. L., Ward, J. J., Sodhi, J. S., and Jones, D. T. (2005) Protein structure prediction servers at University College London. *Nucleic Acids Res. 33*, W36−38.
(20) Jones, D. T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol. 292*, 195−202.
(21) Reynolds, C., Damerell, D., and Jones, S. (2009) ProtorP: a protein-protein interaction analysis server. *Bioinformatics 25*, 413−414.
(22) Berjanskii, M. V., Neal, S., and Wishart, D. S. (2006) PREDITOR: a web server for predicting protein torsion angle restrains. *Nucleic Acids Res. 34*, W63−W69.
(23) Demeler, B., Saber, H., and Hansen, J. C. (1997) Identification and interpretation of complexity in sedimentation velocity boundaries. *Biophys. J. 72*, 397−407.
(24) Brookes, E., and Demeler, B. (2007) Parsimonious regularization using genetic algorithms applied to the analysis of analytical

1269

dx.doi.org/10.1021/cb400047s | *ACS Chem. Biol.* 2013, 8, 1263−1270

ultracentrifugation experiments*GECCO Proceedings ACM 978-1-59593-697-4/07/0007.*

(25) Demeler, B., and Brookes, E. (2008) Monte Carlo analysis of sedimentation experiments. *Colloid Polym. Sci. 286*, 129−137.

(26) Alekseyev, V. Y., Liu, C. W., Cane, D. E., Puglisi, J. D., and Khosla, C. (2007) Solution structure and proposed domain domain recognition interface of an acyl carrier protein domain from a modular polyketide synthase. *Protein Sci. 16*, 2093−2107.

(27) Tae, H., Sohng, J. K., and Park, K. (2009) MapsiDB: an integrated web database for type I polyketide synthases. *Bioprocess. Biosyst. Eng. 32*, 723−727.

(28) Zheng, J., Gay, D. C., Demeler, B., White, M. A., and Keatinge-Clay, A. T. (2012) Divergent evolution of multimodular polyketide synthases revealed through the architecture of a ketoreductase-enoylreductase didomain. *Nat. Chem. Biol. 8*, 615−621.

(29) Maier, T., Leibundgut, M., and Ban, N. (2008) The crystal structure of a mammalian fatty acid synthase. *Science 321*, 1315−1322.

(30) Brignole, E. J., Smith, S., and Asturias, F. J. (2009) Conformational flexibility of metazoan fatty acid synthase enables catalysis. *Nat. Struct. Mol. Biol. 16*, 190−197.

(31) Tang, Y., Chen, A. Y., Kim, C. Y., Cane, D. E., and Khosla, C. (2007) Structural and mechanistic analysis of protein interactions in module 3 of the 6-deoxyerythronolide B synthase. *Chem. Biol. 14*, 931−943.

(32) Pappenberger, G., Benz, J., Gsell, B., Hennig, M., Ruf, A., Stihle, M., Thoma, R., and Rudolph, M. G. (2010) Structure of the human fatty acid synthase KS-MAT didomain as a framework for inhibitor design. *J. Mol. Biol. 397*, 508−519.

(33) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, in *Methods in Enzymology* (Carter, C.W., Jr., and Sweet, R.M., Eds.) Vol. 276, Macromolecular Crystallography, Part A, pp 307−326, Academic Press, New York.

(34) Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 59*, 1131−1137.

(35) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 2126−2132.

(36) Muhandiram, D. R., and Kay, L. E. (1994) Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. *J. Magn. Reson., Ser. B 103*, 203−216.

(37) Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R., and Formankay, J. D. (1993) A gradient-enhanced HCCH-TOCSY experiment for recording side-chain ^{1}H and ^{13}C correlations in H_2O samples of proteins. *J. Magn. Reson., Ser. B 101*, 333−337.

(38) Pascal, S. M., Muhandiram, D. R., Yamazaki, T., Formankay, J. D., and Kay, L. E. (1994) Simultaneous acquisition of ^{15}N- and ^{13}C-edited NOE spectra of proteins dissolved in H_2O. *J. Magn. Reson., Ser. B 103*, 197−201.

(39) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277−293.

(40) Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) ^{1}H, ^{13}C, and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR 6*, 135−140.

(41) Lillemoen, J., Cameron, C. S., and Hoffman, D. W. (1997) The stability and dynamics of ribosomal protein L9: investigations of a molecular strut by amide proton exchange and circular dichroism. *J. Mol. Biol. 268*, 482−493.

(42) Demeler, B. (2012) UltraScan III, version 1473: An integrated framework for the analysis of sedimentation velocity experiments. http://www.ultrascan.uthscsa.edu/.

(43) Brookes, E., and Demeler, B. (2008) Parallel computational techniques for the analysis of sedimentation velocity experiments in UltraScan. *Colloid Polym. Sci. 286*, 138−148.

(44) Demeler, B. (2010) Methods for the design and analysis of sedimentation velocity and sedimentation equilibrium experiments with proteins. *Curr. Protoc. Protein Sci. 60*, 7.13.1−7.13.24.

(45) Brookes, E., Boppana, R. V., and Demeler, B. (2006) Computing large sparse multivariate optimization problems with an application in biophysics, *Supercomputing '06 ACM 0-7695-2700-0/06.*

(46) Brookes, E., Cao, W., and Demeler, B. (2010) A two-dimensional spectrum analysis for sedimentation velocity experiments of mixtures with heterogeneity in molecular weight and shape. *Eur. Biophys. J. 39*, 423−435.

(47) Demeler, B., and van Holde, K. E. (2004) Sedimentation velocity analysis of highly heterogeneous systems. *Anal. Biochem. 335*, 279−288.

(48) Brookes, E., and Demeler, B. (2006) Genetic algorithm optimization for obtaining accurate molecular weight distributions from sedimentation velocity experiments. *Prog. Colloid Polym. Sci. 131*, 78−82.

(49) DeLano, W. L. *The PyMOL Molecular Graphics System*, DeLano Scientific LLC, San Carlos, CA, USA; http://www.pymol.org.

(50) Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics 23*, 2947−2948.