Solution Structure and Conformational Flexibility of a Polyketide Synthase Module

Maja Klaus,† Emanuele Rossini,† Andreas Linden, Karthik S. Paithankar, Matthias Zeug, Zoya Ignatova, Henning Urlaub, Chaitan Khosla,* Jürgen Köfinger, Gerhard Hummer,* and Martin Grininger*

ABSTRACT: Polyketide synthases (PKSs) are versatile C–C bond-forming enzymes that are broadly distributed in bacteria and fungi. The polyketide compound family includes many clinically useful drugs such as the antibiotic erythromycin, the antineoplastic epothilone, and the cholesterol-lowering lovastatin. Harnessing PKSs for custom compound synthesis remains an open challenge, largely because of the lack of knowledge about key structural properties. Particularly, the domains—well characterized on their own—are poorly understood in their arrangement, conformational dynamics, and interplay in the intricate quaternary structure of modular PKSs. Here, we characterize module 2 from the 6-deoxyerythronolide B synthase by small-angle X-ray scattering and cross-linking mass spectrometry with coarse-grained structural modeling. The results of this hybrid approach shed light on the solution structure of a cis-AT type PKS module as well as its inherent conformational dynamics. Supported by a directed evolution approach, we also find that acyl carrier protein (ACP)-mediated substrate shuttling appears to be steered by a nonspecific electrostatic interaction network.

KEYWORDS: natural compound synthesis, assembly lines, multidomain proteins, hybrid approach, directed evolution

INTRODUCTION

Modular polyketide synthases (PKSs) are a class of megasynthases that produce compounds of high pharmaceutical importance.1 In modular PKSs, modules are linked covalently or noncovalently,2−6 and are hypothesized to self-organize into linear scaffolds that feature product synthesis in a vectorial manner (Figure 1A).5,6 In cis-AT PKSs, the catalytic core of each PKS module comprises the domains needed for the condensation of small acyl precursors, the ketosynthase (KS), the acyl transferase (AT), and the acyl carrier protein (ACP). In addition, the domains ketoreductase (KR), dehydratase (DH), and enoylreductase (ER) can be present to modify the condensation product at the β-position. In most cases, a thioesterase (TE) domain at the C-terminus of the modular PKSs is responsible for product release.7,8

Over the last years, structural evidence has been obtained for two distinct structural arrangements of PKS modules that differ most substantially in the relative arrangement of the condensation domains KS and AT. Small-angle X-ray scattering (SAXS) data obtained on the 6-deoxyerythronolide B synthase module 3 (DEBS M3) revealed an extended conformation of the KS-AT dimer9 in support of previous X-ray crystallographic data collected on the didomain subunit.10−12 In contrast, a cryo-electron microscopy (cryoEM) study on the related pikromycin synthase module 5 (PIKS M5), reporting the first structure of a complete PKS module received at moderate resolution, suggested an arch-shaped KS-AT dimer conformation (Figure 1B).13,14 Whereas the DEBS M3 structural model agrees with the structure of the evolutionarily related mammalian fatty acid synthase (FAS),15 with the cryoEM structure of the lovastatin synthase iterative module LovB,16 and with 3D-models generated from partial X-ray structures (mycocerosic acid synthase (MAS)-like PKS17 and spinosyn synthase module 2 (SPNS M2)18), the structural model derived from the EM reconstruction of PKS M5 differs substantially from the others. The two structural models lead to substantial divergence in the molecular mechanism, notably concerning both the intramodular elongation and intermodular translocation steps.19,20 The model derived from PIKS M5 structural data suggested loading of the KS active site by the upstream ACP (translocation step) via a side entrance, whereas, for elongation, the extender substrate reaches the binding site via a newly observed bottom entrance. Understanding the structural and mechanistic key features of PKSs is essential to inform their engineering for producing novel compounds with potentially new bioactivities.11,22

Received: February 2, 2021
Published: October 18, 2021

Supporting Information

10.1021/jacsau.1c00043
JACS Au 2021, 1, 2162−2171
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.
To address the question of the structural and conformational properties of a PKS module, we have investigated DEBS module 2 (DEBS M2) by integrating SAXS and cross-linking mass spectrometry (XL-MS) with coarse-grained (CG) structural modeling and simulation. We considered the combination of noninvasive solution structure methods with robust structure calculation by Bayesian inference as particularly suited for addressing the question of the overall structural appearance of a PKS module. The hybrid approach enables the detailed study of the ensemble of conformations of a PKS module in solution, providing insight beyond the currently available structural information. DEBS is the ideal modular PKS for this approach, because it is a prototypical PKS and it has served as a model system before to examine the mechanism and engineering potential as well as the structure of the enzymatic domains or larger assemblies, e.g., KR1, DH4, KS3-AT3, KS5-AT5, ACP2 (numbers indicate module), docking domains (DDs) between DEBS2 and DEBS3, and TE. Further, DEBS is phylogenetically situated in the same clade with other actinobacterial modular PKSs, so that insight gained likely applies to PKS modules in general.

We applied our approach to two DEBS M2 constructs (Figure 1C). Construct A contains the native M2 (KS-AT-KR-ACP-DD) with an N-terminally fused maltose binding protein (MBP). It represents a stand-alone PKS module with a domain organization comparable to the previously studied DEBS M3 and PIKS M5 (essentially just differing in the N-terminal MBP fusion). In construct B, M2 is decorated with the natural upstream domains KR1 and ACP1 (“1” for M1), constituting the processing part of DEBS M1, and a C-terminal TE domain (Figure 1C). The two constructs were chosen to provide complementary information on domain-domain interfaces in PKS modules (Table S1).

Although our data show that extended (“DEBS M3 model”) and arch-shaped (“PIKS M5 model”) conformations of module DEBS M2 are possible in solution, they likewise reveal the propensity of DEBS M2 to adopt the extended conformation in solution. In this conformation, domains KR and ACP are conformationally dynamic processing part of a PKS module in accordance with previous data. In a focus on the domain-domain interplay in modular PKSs, we complemented our hybrid approach with a directed evolution experiment based on phage display. Enhanced CG simulations, XL-MS analysis, and phage display suggest that interactions of ACP with catalytic domains are steered by networks of weak electrostatic interactions.

Our findings shed light on the structure and dynamics of a PKS module, and offer new insight into the interplay between domains in PKSs. Collectively, these results provide guidance in the engineering of PKSs for the custom synthesis of novel polyketides.

## EXPERIMENTAL SECTION

### Plasmids, Bacterial Cell Culture, and Protein Purification

Plasmids harboring genes encoding individual PKS modules were generated via In-Fusion Cloning (Takara) and restriction and ligation-based techniques. Proteins were expressed and purified using similar protocols. For holo-proteins (where the ACP domain is post-translationally modified with a phosphopantetheine arm), E. coli BL21 cells were cotransformed with a plasmid encoding for the phosphopantetheine transferase Sfp from B. subtilis (pAR357). Proteins either contained a C-terminal His6-tag or an N-terminal MBP-tag for purification. Twin-Strep-tagged proteins (KS3-AT3 for...
was set to placed in a cubic box of 50 nm size. The solvent dielectric constant conformations produced at highest temperature for further analysis. For more information, see the Supporting Information.

Enzymatic Assays and Triketide Analysis
PKS enzymatic assay of KR1-ACP1-M2-TE was performed according to published procedures. The resulting product was extracted twice with 450 µL dried in vacuo and submitted for LC-MS analysis. For more information, see the Supporting Information.

SAXS Analysis and CG Structural Modeling
SEC-SAXS analysis was performed on the Bio-SAXS beamline BM29 at the European Synchrotron Radiation Facility (ESRF). We computed the SAXS intensities for each of the individual CG conformations using the package FoXS. We used the Bayesian Inference of Ensembles (BioEn) method to determine the statistical weight of $A_{\text{org}}$, $A_{\text{coh}}$, and $B_{\text{ext}}$ ensembles of conformations underlying the measured SAXS intensities. BioEn computations were performed considering the SAXS intensities in the range of $0.02 < q < 0.3$ Å$^{-1}$. From the scattering curves, the radius of gyration $R_g$ was calculated using the Guinier approximation. We reduced the sampled ensembles to the conformations that together fulfill more than 99% of the statistical cumulative weight for evaluation of the measured cross-links. Systematic deviations between the measured and simulated scattering curves at low $q$ range suggested moderate aggregation of the systems in solution. To quantify and account for the fraction of aggregation, we combined each simulated curve with the intensities of a globular model at a statistical weight that matched the fraction of aggregation, we combined each simulated scattering curves at low $q$ range and simulated scattering curves at low $q$ range.

SAXS Analysis and XL-MS Data Collection
We mapped the measured cross-links of construct $A$ over the 1000 $A_{\text{org}}$ and $A_{\text{coh}}$ conformations ranked by weight that fulfill 99.9% of the BioEn scattering curves. The same analysis was performed to map the measured XL-MS data of construct $B$ over the 5000 conformations of model $B_{\text{ext}}$ that fulfill 99% of the BioEn scattering curve. For all selected conformations, we calculated the $C_r$ pair-distances across the reported cross-linked residues. For more information, see the Supporting Information.

Phage Display Methodology
Phage display experiments were carried out according to the protocol described by Tonikian et al. A library of ACP1(2) mutants fused to the N-terminus of the minor coat protein P3 of the M13 bacteriophage was generated as described before. A primer for randomization of five amino acids in the chain translation epitope of ACP1 was designed using NNK at the target positions (P-MK162:5' CTG GCC TCG CTG CCC GCG NNK GAG CGC CNK NKC GCC CTG TTC TNC CTC GTG CGC NNK NGC GCC GGC GTC CTC 3' (N: A/C/G/T K: G/T)). Panning was performed in a StreptactinXT-coated microtiter plate with wells coated with (3)KS3-AT3-Strep. The blocked phage library was incubated for 1 h on the coated and blocked wells prior to washing. Wash steps were increased each round over the course of the panning protocol (10×, 12×, 14×, 16×, 18×). Elution of bound phage was done with 100 µL of 100 mM hydrochloric acid for 20 min. The elution was propagated and the titer was determined. For propagation, 2 mL of log phase E. coli SS320 (OD600 0.8) was added to 200 µL of eluted phage and infection was allowed for 30 min at 37 °C, 200 rpm, whereupon 20 µL of M13K07 helper phage (1.0 × 10$^{12}$ PFU/mL) was added. Phages were harvested and pellets containing the phage were resuspended in 1 mL of blocking buffer. To test the eluted phage for increased and specific binding, a specificity ELISA was performed. To further confirm the relative binding intensity of the newly enriched mutants, we used a titration ELISA using purified phage. The advantage of this assay is that defined phage amounts can be used and the signal can be normalized across different phage preparations.

RESULTS

**SAXS Analysis and XL-MS Data Collection**
Constructs $A$ and $B$ were purified as both apo- and phosphopantetheinylated holo-proteins using established protocols, yet codon harmonization of the M2 gene was used to increase expression yields for both constructs (Figure S1, Table S2). We note that the body of data has been collected on the apoprotein (CG structural modeling, SAXS, XL-MS, and phage display) unless otherwise stated. The proteins purified to homogeneity were subjected to tandem size exclusion chromatography small-angle X-ray scattering (SEC-SAXS) and eluted as dimeric species (Figure S2A). Scattering data from the region of stable $R_g$ values were analyzed, assuming that higher oligomer and aggregates are essentially eliminated.
Coarse-Grained Modeling of Two PKS Module Conformations

In integrating SAXS and XL-MS into a complex computational model building approach, we decided to employ structural models of the extended and the arch-shaped KS-AT dimer (KS-AT2) conformation for CG modeling of construct \( A \) (\( A_{\text{ext}} \) and \( A_{\text{arch}} \). Figure 2A, B, Figure S5, and SI Material and Methods). Model \( A_{\text{ext}} \) was based on available X-ray crystallographic data and exhibits an extended KS-AT2 conformation with rigid post-AT linkers folding back to the KS domain. Model \( A_{\text{arch}} \) was informed by the PIKS M5 cryoEM density and carries an arch-shaped KS-AT2, with flexible post-AT loops (Figure S6). We generated an ensemble of structural models in CG simulations and performed Bayesian ensemble refinement. An entropic penalty ensured that the weights of the individual conformations were minimally adjusted to match the measured SAXS intensities (Figure S7). We note that for CG structural modeling, SEC-SAXS frames over the entire dimeric elution peak were averaged for constructs \( A \) and \( B \), respectively (Figure S2A). The Bayesian inference highlighted a small degree of protein aggregation in solution, also visible for both constructs in the SEC profiles (Figure S2A), for which we accounted in silico by aggregate correction (SI Note 2).

Computed \( R_g \) values from the BioEn ensembles of \( A_{\text{ext}} \) and \( A_{\text{arch}} \) were equal to 67.6 and 64.0 Å based on the Guinier approximation. The in silico intensities computed for \( A_{\text{ext}} \) are in excellent agreement with the previous data collected on DEBS M3-TE. The in silico intensities computed for \( A_{\text{arch}} \) are essentially similar, molecular weight \( M_w \) estimates by the volume of correlation \( V_p \) well reflect calculated \( M_w; \) i.e., 408 (exp.) vs 403 kDa (calc.) for construct \( A \), and 484 vs 488 kDa for construct \( B \) (SI Note 1, Figure S3B and Tables S3 and S4). In addition to SAXS analysis, XL-MS was employed to gain distance information for specific residues in solution. We used the heterobifunctional cross-linker succinimidyl 4,4′-azipentanoate to map interactions of lysine residues with other residues within and across domain boundaries (Cα-Cα distance ~20 Å (ref19) (Tables S5–S7)). Focusing on construct \( A \), we observed a number of XLs between the flexible ACP domain and catalytic domains (from ACP\textsubscript{K1774} and ACP\textsubscript{K2754}), yet also across interfaces of less flexible domains and units, e.g., KS-AT and KS-KR (from KS\textsubscript{K273}), AT-LD and AT-KR (from AT\textsubscript{K1071}), and the post-AT linker (from K1273) and domains LD, AT, and KR (Table S6). Overall constructs \( A \) and \( B \) showed a strong overlap in mapped interactions (Table S6 and S7). The biological relevance of the M2-TE region of construct \( B \) was established by measuring triketide lactone product formation using a synthetic substrate (Figure S4).

XL-MS Distance Information within PKS Modules

In a next step, we mapped the measured XLs of construct \( A \) over the extended and arch-shaped configurations to locate candidate regions of domain–domain interactions in M2 (SI Note 5). Overall, the \( A_{\text{ext}} \) model is in very good agreement with the measured cross-linking data (Tables S6 and S7, SI Note 6). High numbers of cross-linked peptide spectral matches (CSMs), supporting a specific interaction, are observed across the post-AT linker and KR, particularly between K1273 (post-AT linker) and E1474, Y1475, and A1476 (KR, Figure 3A). These residues emerge as a hot-spot of interaction that constrain an otherwise undirected random motion of KR around the post-AT linker and KS, supporting the high flexibility of M2 in solution (Figure 3B and Movie S1). The attachment of the post-AT linker to the KS-AT2 core and the treatment of the KS-AT2 core as rigid body generally prevented interactions between KR and the peripheral regions of AT. As such, one observed cross-link in this region (XL7, AT\textsubscript{K1071}–KR\textsubscript{Q1623}), which cannot be explained by conformational variability of the \( A_{\text{ext}} \) model, indicates that dissociation of the post-AT linker from the KS-AT2 core, providing a higher degree of flexibility to KR, or conformational variability within the KS-AT2 core may be possible at low frequency.
Alternatively, XL7 may result from formation of higher oligomers, as observed in SEC (Figure S2A), in which the peripheral stretch of the AT could interact with the mobile KR of another module (Figure S13, Table S6). Another set of XLs was observed between MBP and KS-AT2, in which the MBP domain interacts ubiquitously, yet at low frequencies with the KS-AT dimer (Figure S14, Table S6). This can be rationalized by the high mobility of the MBP domain, which is connected by long linkers to the condensing domain (Table S9). In contrast, Aext satisfies only a small set of the XLs measured mainly across KR and ACP (Table S6). Overall, XL-MS data thus support an extended conformation of M2 in solution.

Positional Variability of ACP

The attachment of ACP via a 14-amino acid linker to the mobile KR implies a high conformational variability. By computational modeling and XL-MS, we isolated candidate regions of domain–domain interaction of ACP with KS-AT2 and KR (Figure 3C–E and Figure S15). We observed high frequencies of interactions across K1775 and K1784 of ACP and D538, G539, and Y540 of KS and across K1775 of ACP and V1295, S1303, and E1304 of KR. The candidate regions are populated by charged residues: D1172, D1773, R1778, E1785 and D1789 of ACP; D535, R533, Y536, R612, R711, and D714 of KS; and D1296, E1299, R1302, E1449, D1450, R1716, R1718, D1722, and R1723 of KR (Table S6). These data indicate that domain–domain interactions are based on networks of electrostatic interactions, similarly as found before in type I and type II FASs.20–22 In our simulations, ACP extensively samples the entrance of the KS binding pocket (Figure S15), and we isolated a candidate orientation with ACP51790 pointing toward the active site and the catalytic C571 of KS (Figure 3D). Note that, because of its small size, the ACP domain is represented in computational models with low weight, thereby impeding a more detailed description of interfaces. However, XL-MS confirms modeling data and highlights interaction hot spots of ACP distributed over KS-AT2 and KR that are steered by networks of electrostatic interactions (Table S6).

Structural Properties of Construct B

Analogously to construct A, the SAXS intensities measured for construct B in solution are explained by a structural model (Bext) that carries the extended KS-AT2 conformation (Figure S16, SI Note 7). In Bext, KR1 (KR of M1) is highly mobile and occupies a large conformational space, whereas ACP1 displays low mobility and samples specific regions of KS-AT2 (Figure S17, SI Note 8). In addition, the lowered flexibility of KR (of M2) in Bext confirms the structural constraints imposed by the TE (Figure S18, SI Note 8). We note that XL-MS was performed with phosphopantetheinylated construct B. Overall fewer CSMs were observed compared to (apo-)construct A (Table S7), which was also found in a single experiment with phosphopantetheinylated construct A (yielding fewer CSMs compared to (apo-)construct A; Table S6). This observation indicates that the phosphopantetheine moiety constrains ACP docking during substrate shuttling, as suggested previously in structural13,53,54 and functional studies55 on type I PKSs and the related FASs (SI Note 9). In construct B, no CSMs were observed between ACP1 and KS, which may be due to the suppression of this interaction by the competitive ACP1-KRI interactions or by the high dynamics of ACP1 due to the unconstrained KR1 (no N-terminal KS1-AT1 dimer in construct B). In conclusion, Bext is in good agreement with the measured CSMs (Table S5, SI Note 10). Interestingly, the interaction pattern of Aext is also captured in Bext, highlighting the transferability of the XL-MS data to a different construct of DEBS (Table S6, SI Note 11).
transacylation, which we 300 enriched mutants were selected for ELISA screening, from only one that has been described so far in modular PKSs in locating the growing acyl chain, because this interface is the interacting with KS of the downstream module for trans-
translocation interface between ACP and KS, i.e., ACP phage display. We decided to work with the intermodular additional experiment based on directed evolution through understanding of catalytic domain recognition by ACP in an

from selection for the translocation interface. For the two most active mutations, Lib1-Mut3 and Lib1-Mut5, new positions at H1 were chosen for further randomization. Both libraries were treated with the same protocol and analyzed with the same methods as the original library, except for an additional ELISA titration experiment that was conducted to further evaluate binding of ACP1H1 mutants compared to wild-type. The new libraries did neither produce a consensus sequence nor result in high affinity or high turnover bimodular PKSs (Figure S20, Table S8, and SI Note 13). Collectively, these data argue against the occurrence of a specific ACP:KS translocation interface, and, thus, agree with the positional variability of ACP in docking at a catalytic domain suggested from cross-linking and modeling data. We note that several studies propose that the spatial distribution of ACP is sensitive to the acyl-phosphopantetheine moiety,13,53,54 and future studies with phosphopantetheinylated ACP and the native substrate will be needed to further understand the specificity in substrate shuttling as well as the impact of acyl-

Characterization of an Interface via Phage Display

Having observed that the ACP-covering cross-links do not converge on specific interfaces on either catalytic domain of the KS-AT core (Tables S6 and S7), we sought to further our understanding of catalytic domain recognition by ACP in an additional experiment based on directed evolution through phage display. We decided to work with the intermodular translocation interface between ACP and KS, i.e., ACP interacting with KS of the downstream module for translocating the growing acyl chain, because this interface is the only one that has been described so far in modular PKSs in more detail.55 Specifically, we set up a phage display approach with the aim of increasing the affinity of the non-native interface between ACP of module 1 (ACP1) and KS of module 3 (KS3). The sequence identity between KS of M2 (KS2) and KS3 is 58% on the protein level (EMBOSS pairwise assembly, www.ebi.ac.uk).56 For phage display, ACP was expressed on phage, and its helix 1 (H1), previously determined to play a crucial role during intermodular chain translocation,58 was mutagenized (Figure S19). We note that phage display was performed with apo-ACP phage library due to the difficulty in performing and monitoring phosphopante-

The structure of PKSs is challenging. Structural characterization of flexible proteins is possible by the synergistic application of SAXS, XL-MS, and integrative modeling that enables the quantitative assessment and modeling of structural heterogeneity and conformational variability. In applying the combined approach, we were able to substantiate the relevance of the two strikingly different
arch-shaped and extended conformations that have both been proposed as prevalent structures of a PKS module. Overall, our experimental data and modeling results on DEBS M2 show that both conformations are possible in principle; however, the extended structure appears to be dominant in solution.

The extended structure of modular PKSs in solution is itself dynamic, and we present direct insight into the conformational variability of a PKS module. Whereas the KS-AT dimeric unit is relatively rigid, as reported by previous structural studies on type I PKSs, and treated as rigid core in this study, computational modeling constrained by SAXS data suggests a high conformational variability of KR. Combined with XL-MS data highlighting the interaction of KR with the post-AT linker, a picture of KR moving undirected and randomly just constrained by the post-AT linker as an anchoring point emerges. A PKS module can also feature dimerizing units in the processing part; i.e., the dimeric domain DH, and a dimerization element present in about half of the modules with only KR as modifying domain. Both units will constrain conformational dynamics, and KR will no longer be a separately moving domain (Movie S1), but part of a larger assembly with decreased mobility. In conclusion, this study argues that in PKS modules, a compact KS-AT dimeric unit carries a processing part with conformational dynamics depending on its module organization.

The substrate-shuttling ACP domain is of key importance for any PKS module. In our study, performed on the nonphosphopantetheinylated domain, we observe a lack of high-frequency XLS from ACP to KS, AT and KR, which indicates that ACP does not form highly specific interfaces with those domains. To further our understanding of the specificity of ACP-domain interactions, we employed a directed evolution approach for evolving the ACP:KS translocation interface toward higher affinities. Directed evolution approaches are widely used to narrow down specific interactions as a result of an increase in affinity. Our results show that despite intensive screening only few higher affinity hits emerged from ELISA analysis, among which no consensus sequence could be found. We interpret this observation as arguing against the existence of specific ACP:KS translocation interfaces, in line with our solution structure data. Overall, our data suggest that substrate shuttling by ACP of type I PKS is facilitated by weak interactions, similarly as described earlier for type I and type II FAS. We conclude that ACP moves freely, and rather randomly collides with domains of the type I assembly, where electrostatic networks influence the approach of ACP to the individual domains. As such our findings essentially support a previously suggested model in which ACP is unrestrained, but restricted in its trajectory by the other domains of the multienzyme assembly.

We note that the body of data for this study was collected with nonphosphopantetheinylated proteins, mainly due to technical reasons. For example, domains were represented as rigid bodies in CG structural modeling, and the phosphopantetheine moiety was not represented in the model. Since ACPs of type I systems do not accommodate phosphopantetheine and acyl-chains, leading just to local changes around the active serine in apo-, holo-, and acylated-ACP, but leaving the ACP fold overall unchanged, our data are representative of the domain–domain portion of the interactions. However, several structural studies revealed that the spatial distribution of ACP is sensitive to the acyl moiety in type I PKSs and the related FASs. Further, recent enzyme kinetic studies on ATs from type I PKSs and FAs reported the variation in $K_M$ with the identity of the acyl unit, suggesting the impact of phosphopantetheine moiety on ACP docking. Accordingly, studies are required on phosphopantetheinylated and acylated ACPs to understand substrate shuttling. In particular, the impact of the phosphopantetheine and the identity of the acyl-moiety on this process is just poorly characterized for type I PKSs and FAs until to date.

Besides giving insight into the structural and conformational properties of PKSs, our study provides guidelines for PKS engineering. Taking into account recently published data, the extended conformation is likely self-sufficient for synthetic progress in PKS assembly lines. We therefore suggest to work with the extended model in PKS engineering approaches, as well as to base active site remodeling on the biomolecular processes connected to this structural arrangement. For example, our data support the KS functional mode to be “conventional”, i.e., using the generic KS binding tunnel identical to type I FASs and iterative PKSs, and approaches on active site design can be informed by data available also on those systems. Our findings can further explain why engineering approaches involving exchanges of the KR domain or the entire processing part have been successful in the past, as well as encourage consideration of such approaches in PKS engineering projects. XL-MS data indicate just few cross-links between KR and other domains (to the AT domain, Tables S6 and S7), so that KR exchanges may be well accessible in modules with only KR as modifying domain. We assume that also exchanges of the entire processing part (in larger PKS modules) are amenable, because condensing and processing parts appear separated in structural and conformational properties; as also implied by several successful engineering approaches (e.g., refs 72 and 73). Data for ACP should be interpreted in a similar light. As our direct insight in ACP-mediated domain–domain interactions argues against well-defined docking events (between apo-ACP and catalytic domains), productive substrate shuttling appears possible in chimeric PKS with non-native ACP as long as conflicting electrostatic interactions are avoided.

**ASSOCIATED CONTENT**

**Supporting Information**

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacsau.1c00043.

Description of plasmids and chemicals, enzymatic assays, and compound analysis; protocols for data collection and analysis in XL-MS, CG structural modeling, SAXS, and phage display; this includes Supporting Notes 1–13, Figures S1–S20, and Tables S1–S10 for details and additional data on all the methods and material used in this study (PDF)

Movie S1, prevalent conformations of construct A in solution (MP4)

**AUTHOR INFORMATION**

**Corresponding Authors**

Martin Grininger – Institute of Organic Chemistry and Chemical Biology, Buchmann Institute for Molecular Life Sciences, Goethe University Frankfurt, Frankfurt am Main 60438, Germany; orcid.org/0000-0002-7269-0667; Email: grininger@chemie.uni-frankfurt.de
Gerhard Hummer — Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Frankfurt am Main 60438, Germany; Institute of Biophysics, Goethe University Frankfurt, Frankfurt am Main 60438, Germany; orcid.org/0000-0001-7768-746X; Email: gerhard.hummer@biophys.mpg.de

Chaitan Khosla — Department of Chemistry, Stanford ChEM-H, Department of Chemical Engineering, Stanford University, Stanford, California 94305, United States; orcid.org/0000-0001-6529-495X; Email: khosla@stanford.edu

Authors

Maja Klaus — Institute of Organic Chemistry and Chemical Biology, Buchmann Institute for Molecular Life Sciences, Goethe University Frankfurt, Frankfurt am Main 60438, Germany

Emanuele Rossini — Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Frankfurt am Main 60438, Germany

Andreas Linden — Max Planck Institute for Biophysical Chemistry, Goettingen 37077, Germany; Institute for Clinical Chemistry, University Medical Center Göttingen, Göttingen 37075, Germany

Karthsik S. Pathankar — Institute of Organic Chemistry and Chemical Biology, Buchmann Institute for Molecular Life Sciences, Goethe University Frankfurt, Frankfurt am Main 60438, Germany

Matthias Zeug — Institute of Organic Chemistry and Chemical Biology, Buchmann Institute for Molecular Life Sciences, Goethe University Frankfurt, Frankfurt am Main 60438, Germany; orcid.org/0000-0002-7326-319X

Zoya Ignatova — Institute for Biochemistry and Molecular Biology, University of Hamburg, Hamburg 22607, Germany

Henning Urlaub — Max Planck Institute for Biophysical Chemistry, Goettingen 37077, Germany; Institute for Clinical Chemistry, University Medical Center Göttingen, Göttingen 37075, Germany

Jürgen Köfinger — Department of Theoretical Biophysics, Max Planck Institute of Biophysics, Frankfurt am Main 60438, Germany; orcid.org/0000-0001-8367-1077

Complete contact information is available at: https://pubs.acs.org/10.1021/jacsau.1c00043

Author Contributions

M.K. and E.R. contributed equally to this work. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding

Support was received from the Volkswagen Foundation (Lichtenberg Professorship to M.G.; grants 85701 and 91809), the LOEWE program (Landes-Offensive zur Entwicklung wissenschaftlich-ökonomischer Exzellenz) of the state of Hesse conducted within the framework of the MegaSyn Research Cluster (to M.G. and G.H.), the Max Planck Society (to G.H.), the Deutsche Forschungsgemeinschaft (to H.U., SFB886), and the NIH (to C.K., grant number R01 GM087934).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Mirko Joppe for providing the diketide substrate, Alicia D’Souza for assistance in cloning pADD01, and Lynn Buyachuihan for help with the protein purification. We are grateful for Florian Bühr’s and Harald Schwabe’s support in generating the harmonized gene sequences, Gregor Witte for initial support in analyzing the SAXS data, and Andreas Ernst and Mateusz Putyrski for support in setting up and running the phage display experiments.

ABBREVIATIONS

PKS, polyketide synthase; DEBS, deoxyerythronolide B synthase; PIKS, pikromycin synthase; KS, ketosynthase; AT, acyl transferase; ACP, acyl carrier protein; KR, ketoreductase; DH, dehydratase; ER, enoylreductase; TE, thioesterase

REFERENCES

(1) Staunton, J.; Weissman, K. J. Polyketide biosynthesis: a millennium review. Nat. Prod. Rep. 2001, 18, 380–410.
(2) Broadhurst, R. W.; Nietlispach, D.; Wheatcroft, M. P.; Leadlay, P. F.; Weissman, K. J. The structure of docking domains in modular polyketide synthases. Chem. Biol. 2003, 10, 723–731.
(3) Buchholz, T. J.; Geders, T. W.; Bartley, F. E., 3rd.; Reynolds, K. A.; Smith, J. L.; Sherman, D. H. Structural basis for binding specificity between subclasses of modular polyketide synthase docking domains. ACS Chem. Biol. 2009, 4, 41–52.
(4) Rissler, F.; Collin, S.; Dos Santos-Morais, R.; Gruez, A.; Chagot, B.; Weissman, K. J. Towards improved understanding of intersubunit interactions in modular polyketide biosynthesis: Docking in the enacyloxin IIa polyketide synthase. J. Struct. Biol. 2020, 212, 107581.
(5) Cortes, J.; Haydock, S. F.; Roberts, G. A.; Bevitt, D. J.; Leadlay, P. F. An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of Saccharopolyspora erythraea. Nature 1990, 348, 176–178.
(6) McDaniel, R.; Ebert-Kholsa, S.; Hopwood, D. A.; Khosla, C. Engineered biosynthesis of novel polyketides. Science 1993, 262, 1546–1550.
(7) Tsai, S. C.; Miercke, L. J.; Krucinski, J.; Gokhale, R.; Chen, J. C.; Foster, P. G.; Cane, D. E.; Khosla, C.; Stroud, R. M. 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. 2001, 98, 14808–14813.
(8) Giraldes, J. W.; Akey, D. L.; Kittendorf, J. D.; Sherman, D. H.; Smith, J. L.; Feick, R. A. Structural and mechanistic insights into polyketide macroalactonization from polyketide-based affinity labels. Nat. Chem. Biol. 2006, 2, 531–536.
(9) Edwards, A. L.; Matsui, T.; Weiss, T. M.; Khosla, C. Architectures of whole-module and bimodular proteins from the 6-deoxyerythronolide B synthase. J. Mol. Biol. 2014, 426, 2229–2245.
(10) Tang, Y.; Chen, A. Y.; Kim, C. Y.; Cane, D. E.; Khosla, C. Structural and mechanistic analysis of protein interactions in module 3 of the 6-deoxyerythronolide B synthase. Chem. Biol. 2007, 14, 931–943.
(11) Tang, Y.; Kim, C.-Y.; Mathews, I. I.; Cane, D. E.; Khosla, C. The 2.7 Å crystal structure of a 194-kDa homodimeric fragment of the 6-deoxyerythronolide B synthase. Proc. Natl. Acad. Sci. U. S. A. 2006, 103, 11124–11129.
(12) Whicher, J. R.; Smaga, S. S.; Hansen, D. A.; Brown, W. C.; Gerwick, W. H.; Sherman, D. H.; Smith, J. L. Cyanobacterial polyketide synthase docking domains: a tool for engineering natural product biosynthesis. Chem. Biol. 2013, 20, 1340–1351.
(13) Whicher, J. R.; Dutta, S.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Dosey, A. M.; Narayan, A. R.; Hākansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G. Structural rearrangements of a polyketide synthase module during its catalytic cycle. Nature 2014, 510, 560–564.
(14) Dutta, S.; Whicher, J. R.; Hansen, D. A.; Hale, W. A.; Chemler, J. A.; Congdon, G. R.; Narayan, A. R.; Håkansson, K.; Sherman, D. H.; Smith, J. L.; Skiniotis, G. Structure of a modular polyketide synthase. *Nature* **2014**, *510*, S12–S17.

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

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

(17) Herbst, D. A.; Townsend, C. A.; Maier, T. The architectures of iterative type I PKS and FAS. *Nat. Prod. Rep.* **2018**, *35*, 1046–1069.

(18) Zheng, J.; Gay, D. C.; Demeler, B.; White, M. A.; Keatinge-Bruegger, J.; Caldara-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, S.; McSweeney, S. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **2013**, *20*, 660–664.

(19) Schneidman-Duhovny, D.; Hammel, M.; Tainer, J. A.; Sali, A. Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys. J.* **2013**, *105*, 962–974.

(20) Schneidman-Duhovny, D.; Hammel, M.; Tainer, J. A.; Sali, A. FoXS, FoXS-Dock and MultiFoXS: Single-state and multi-state structural modeling of their proteins and complexes based on SAXS profiles. *Nucleic Acids Res.* **2016**, *44*, W424–429.

(21) Rozyczki, B.; Kim, Y. C.; Hummer, G. SAXS ensemble refinement of ESCRT-III CHMP3 conformational transitions. *Structure* **2011**, *19*, 109–116.

(22) Köfinger, J.; Stelzl, L. S.; Reuter, K.; Allande, C.; Reichel, K.; Hummer, G. Efficient Ensemble Refinement by Reweighting. *J. Chem. Theory Comput.* **2019**, *15*, 3390–3401.

(23) Humer, G.; Köfinger, J. Bayesian ensemble refinement by replica simulations and reweighting. *J. Chem. Phys.* **2015**, *143*, 243150.

(24) Svergun, D. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **1992**, *25*, 495–503.

(25) Kim, Y. C.; Hummer, G. Coarse-grained models for simulations of multipeptide complexes: application to ubiquitin binding. *J. Mol. Biol.* **2008**, *375*, 1416–1433.

(26) Yang, B.; Wu, Y. J.; Zhu, M.; Fan, S. B.; Lin, J.; Zhang, K.; Li, S.; Chi, H.; Li, Y. X.; Chen, H. F.; Luo, S. K.; Ding, Y. H.; Wang, L. H.; Hao, Z.; Xiu, L. Y.; Chen, S.; Ye, K.; He, S. M.; Dong, M. Q. Identification of cross-linked peptides from complex samples. Nat. Methods **2012**, *9*, 904–906.

(27) Tonkian, R.; Zhang, Y.; Boone, C.; Sidhu, S. S. Identifying specificity profiles for peptide recognition modules from phase-displayed peptide libraries. *Nat. Protoc.* **2007**, *2*, 1368–1386.

(28) Lowry, B.; Robbins, T.; Cane, D. E.; Khosla, C. Solution structure and proposed domain domain recognition interface of an acyl carrier protein domain from a modular polyketide synthase. *Protein Sci.* **2007**, *16*, 2093–2107.

(29) Nivina, A.; Yuet, K. P.; Hsu, J.; Khosla, C. Evolution and Diversity of Assembly-Line Polyketide Synthases. *ACS Chem. Biol.* **2020**, *15*, 2422–2432.

(30) Ahringer, J.; Jang, S.; Zhang, Z.; Wang, T.; Zheng, J.; Gay, D. C.; Demeler, B.; White, M. A.; Keatinge-Bruegger, J.; Caldara-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, S.; McSweeney, S. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **2013**, *20*, 660–664.

(31) Ahringer, J.; Jang, S.; Zhang, Z.; Wang, T.; Zheng, J.; Gay, D. C.; Demeler, B.; White, M. A.; Keatinge-Bruegger, J.; Caldara-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, S.; McSweeney, S. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **2013**, *20*, 660–664.

(32) Ahringer, J.; Jang, S.; Zhang, Z.; Wang, T.; Zheng, J.; Gay, D. C.; Demeler, B.; White, M. A.; Keatinge-Bruegger, J.; Caldara-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, S.; McSweeney, S. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **2013**, *20*, 660–664.

(33) Ahringer, J.; Jang, S.; Zhang, Z.; Wang, T.; Zheng, J.; Gay, D. C.; Demeler, B.; White, M. A.; Keatinge-Bruegger, J.; Caldara-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, S.; McSweeney, S. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **2013**, *20*, 660–664.

(34) Ahringer, J.; Jang, S.; Zhang, Z.; Wang, T.; Zheng, J.; Gay, D. C.; Demeler, B.; White, M. A.; Keatinge-Bruegger, J.; Caldara-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, S.; McSweeney, S. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **2013**, *20*, 660–664.

(35) Ahringer, J.; Jang, S.; Zhang, Z.; Wang, T.; Zheng, J.; Gay, D. C.; Demeler, B.; White, M. A.; Keatinge-Bruegger, J.; Caldara-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, S.; McSweeney, S. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **2013**, *20*, 660–664.

(36) Schmidt, T.; Kessler, A. The Streptag system for one-step purification and high-affinity detection or capturing of proteins. *Nat. Protoc.* **2007**, *2*, 1528–1535.

(37) Vernot, P.; Round, A.; Barrett, R.; De Maria Antonilos, A.; Gobbo, A.; Gordon, E.; Huet, J.; Kieffer, J.; Lentiini, M.; Mattenet, M.; Morawe, C.; Mueller-Dieckmann, C.; Ohlsson, S.; Schmid, W.; Surr, J.; Theveneau, P.; Zerrad, L.; McSweeney, S. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **2013**, *20*, 660–664.

(38) Ahringer, J.; Jang, S.; Zhang, Z.; Wang, T.; Zheng, J.; Gay, D. C.; Demeler, B.; White, M. A.; Keatinge-Bruegger, J.; Caldara-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, S.; McSweeney, S. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **2013**, *20*, 660–664.

(39) Ahringer, J.; Jang, S.; Zhang, Z.; Wang, T.; Zheng, J.; Gay, D. C.; Demeler, B.; White, M. A.; Keatinge-Bruegger, J.; Caldara-Festin, G.; Finzel, K.; Jackson, D. R.; Ishikawa, S.; McSweeney, S. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J. Synchrotron Radiat.* **2013**, *20*, 660–664.
(54) Lou, J. W.; Iyer, K. R.; Hasan, S. M. N.; Cowen, L. E.; Mazhab-Jafari, M. T. Electron cryomicroscopy observation of acyl carrier protein translocation in type I fungal fatty acid synthase. *Sci. Rep.* 2019, 9, 12987.

(55) Rittner, A.; Paithankar, K. S.; Huu, K. V.; Grininger, M. Characterization of the Polyspecific Transferase of Murine Type I Fatty Acid Synthase (FAS) and Implications for Polyketide Synthase (PKS) Engineering. *ACS Chem. Biol.* 2018, 13, 723−732.

(56) Geyer, K.; Sundaram, S.; Susnik, P.; Koert, U.; Erb, T. J. Understanding Substrate Selectivity of Phospholactomycin Polyketide Synthase by Using Reconstituted in Vitro Systems. *ChemBioChem* 2020, 21, 2080−2085.

(57) Stegemann, F.; Grininger, M. Transacylation Kinetics in Fatty Acid and Polyketide Synthases and its Sensitivity to Point Mutations**.** *ChemCatChem* 2021, 13, 2771−2782.

(58) Kapur, S.; Chen, A. Y.; Cane, D. E.; Khosla, C. Molecular recognition between ketosynthase and acyl carrier protein domains of the 6-deoxyerythronolide B synthase. *Proc. Natl. Acad. Sci. U. S. A.* 2010, 107, 22066−22071.

(59) Needleman, S. B.; Wunsch, C. D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 1970, 48, 443−453.

(60) Wong, F. T.; Jin, X.; Mathews, I. I.; Cane, D. E.; Khosla, C. Structure and mechanism of the trans-acylating transferase from the disorazole synthase. *Biochemistry* 2011, 50, 6539−6548.

(61) Wong, F. T.; Chen, A. Y.; Cane, D. E.; Khosla, C. Protein−protein recognition between acyltransferases and acyl carrier proteins in multimodular polyketide synthases. *Biochemistry* 2010, 49, 95−102.

(62) Miyanaga, A.; Iwasawa, S.; Shinohara, Y.; Kudo, F.; Eguchi, T. Structure-based analysis of the molecular interactions between acyltransferase and acyl carrier protein in vicenistatin biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 2016, 113, 1802−1807.

(63) Benning, F. M. C.; Sakaiyama, Y.; Mazur, A.; Bukhari, H. S. T.; Lim, R. Y. H.; Maier, T. High-Speed Atomic Force Microscopy Visualization of the Dynamics of the Multienzyme Fatty Acid Synthase. *ACS Nano* 2013, 8, 1964−1971.

(64) Zhong, J.; Piasecki, S. K.; Keatinge-Clay, A. T. Structural studies of an A2-type modular polyketide synthase ketoreductase reveal features controlling alpha-substituent stereochemistry. *ACS Chem. Biol.* 2015, 54, 2205−2213.

(65) Tran, L.; Broadhurst, R. W.; Tosin, M.; Cavalli, A.; Weissman, K. J. Insights into Protein-Protein and Enzyme-Substrate Interactions in Modular Polyketide Syntheses. *Chem. Biol.* 2010, 17, 705−716.

(66) Perez, D. R.; Leibundgut, M.; Wider, G. Interactions of the Acyl Chain with the Saccharomyces cerevisiae Acyl Carrier Protein. *Biochemistry* 2015, 54, 2205−2213.

(67) Poskon, E.; Arthur, C.; Evans, S.; Williams, C.; Crosby, J.; Simpson, T.; Crump, M. A mammalian type I fatty acid synthase acyl carrier protein domain does not sequester acyl chains. *J. Biol. Chem.* 2008, 283, 518−528.

(68) Cogan, D. P.; Li, X.; Sevillano, N.; Mathews, I. I.; Matsui, T.; Craik, C. S.; Khosla, C. Antibody Probes of Module 1 of the 6-Deoxyerythronolide B Synthase Reveal an Extended Conformation During Ketoreduction. *J. Am. Chem. Soc.* 2020, 142, 14933−14939.

(69) Bedford, D. J.; Jacobsen, J. R.; Luo, G.; Cane, D. E.; Khosla, C. A functional chimeric modular polyketide synthase generated via domain replacement. *Chem. Biol.* 1996, 3, 827−831.

(70) Kellenberger, L.; Galloway, I. S.; Sauter, G.; Bohn, G.; Hanefeld, U.; Cortes, J.; Staunton, J.; Leadlay, P. F. A polylinker approach to reductive loop swaps in modular polyketide synthases. *ChemBioChem* 2008, 9, 2740−2749.

(71) Annal, T.; Paris, C.; Leadlay, P. F.; Jacob, C.; Weissman, K. J. Evaluating Ketoreductase Exchanges as a Means of Rationally Altering Polyketide Stereochemistry. *ChemBioChem* 2015, 16, 1357−1364.

(72) Hagen, A.; Pou, S.; Rond, T.; Fortman, J. L.; Katz, L.; Petzold, C. J.; Keasling, J. D. Engineering a Polyketide Synthase for In Vitro Production of Adipic Acid. *ACS Synth. Biol.* 2016, 5, 21−27.

(73) Eng, C. H.; Yuzawa, S.; Wang, G.; Baidoo, E. E.; Katz, L.; Keasling, J. D. Alteration of Polyketide Stereochemistry from anti to syn by a Ketoreductase Domain Exchange in a Type I Modular Polyketide Synthase Subunit. *Biochemistry* 2016, 55, 1677−1680.