Title

Structural and functional analysis of two di-domain aromatase/cyclases from type II polyketide synthases

Journal

Proceedings of the National Academy of Sciences of the United States of America, 112(50)

ISSN

0027-8424

Authors

Caldara-Festin, Grace
Jackson, David R
Barajas, Jesus F
et al.

Publication Date

2015-12-15

DOI

10.1073/pnas.1512976112

Structural and functional analysis of two di-domain aromatase/cyclases from type II polyketide synthases

Grace Caldara-Festin[a,b,c,1], David R. Jackson[a,b,c,1,2], Jesus F. Barajas[a,b,c], Timothy R. Valentic[a,b,c], Avinash B. Patel[a,b,c], Stephanie Aguilar[a,b,c], MyChi Nguyen[a,b,c], Michael Vo[a,b,c], Avinash Khanna[a,b,c], Eita Sasaki[d,e], Hung-wen Liu[d,e], and Shiou-Chuan Tsai[a,b,c,2]

[a]Department of Molecular Biology and Biochemistry, University of California, Irvine, CA 92697; [b]Department of Chemistry, University of California, Irvine, CA 92697; [c]Department of Pharmaceutical Sciences, University of California, Irvine, CA 92697; [d]Division of Medicinal Chemistry, College of Pharmacy, University of Texas, Austin, TX 78712; and [e]Department of Chemistry, University of Texas, Austin, TX 78712

Edited by Janet L. Smith, University of Michigan, Ann Arbor, MI, and accepted by the Editorial Board November 4, 2015 (received for review July 2, 2015)

Aromatic polyketides make up a large class of natural products with diverse bioactivity. During biosynthesis, linear poly-β-ketone intermediates are regiospecifically cyclized, yielding molecules with defined cyclization patterns that are crucial for polyketide bioactivity. The aromatase/cyclases (ARO/CYCs) are responsible for regiospecific cyclization of bacterial polyketides. The two most common cyclization patterns are C7–C12 and C9–C14 cyclizations. We have previously characterized three monodomain ARO/CYCs: ZhuI, TcmN, and WhiE. The last remaining uncharacterized class of ARO/CYCs is the di-domain ARO/CYCs, which catalyze C7–C12 cyclization and/or aromatization. Di-domain ARO/CYCs can further be separated into two subclasses: "nonreducing" ARO/CYCs, which act on nonreduced poly-β-ketones, and "reducing" ARO/CYCs, which act on cyclized C9 reduced poly-β-ketones. For years, the functional role of each domain in cyclization and aromatization for di-domain ARO/CYCs has remained a mystery. Here we present what is to our knowledge the first structural and functional analysis, along with an in-depth comparison, of the nonreducing (StfQ) and reducing (BexL) di-domain ARO/CYCs. This work completes the structural and functional characterization of mono- and di-domain ARO/CYCs in bacterial type II polyketide synthases and lays the groundwork for engineered biosynthesis of new bioactive polyketides.

polyketide biosynthesis | structural biology | aromatase/cyclase

The biosynthesis of type II aromatic polyketide natural products has been extensively investigated because of the versatile pharmacological properties of these compounds (1–7). The type II polyketide synthase (PKS) is composed of dissociated enzymes that are used iteratively and are responsible for the elongation, cyclization, and modification of the polyketide chain (Fig. 1) (3, 4, 8, 9). The regiospecific cyclization of an acyl carrier protein (ACP)-linked linear poly-β-ketone intermediate is a key transformation catalyzed by type II PKSs. However, the enzymatic mechanism of cyclization remains poorly understood (10–14). Without such knowledge, the polyketide cyclization pattern cannot be predicted; a full understanding of this process at the molecular level is essential for future biosynthetic engineering efforts.

In 2008, we reported the crystal structure of the first aromatase/cyclase (ARO/CYC) (TcmN ARO/CYC), which is a single-domain protein (15). On the basis of the structural analysis and mutagenesis results, we proposed that monodomain ARO/CYCs contain an active site and are capable of catalyzing polyketide cyclization and aromatization. Since then, we have performed structural and biochemical studies of two other monodomain ARO/CYCs: WhiE and ZhuI (Fig. 1) (16, 17). These studies provided strong evidence supporting our hypothesis that ARO/CYC is the site of polyketide cyclization. However, many type II PKSs contain di-domain ARO/CYCs that have two seemingly identical domains (18–23). Why these enzymes require two domains (as opposed to just one) and how they conduct the cyclization/aromatization are not understood.

The di-domain ARO/CYCs are found in both nonreducing and reducing PKSs (18, 24). In nonreducing systems, the di-domain ARO/CYCs regiospecifically cyclize a polyketide between C7 and C12, followed by aromatization (18). In reducing systems, a ketoreductase (KR) first regiospecifically cyclizes the linear poly-β-ketone from C12 to C7, followed by a highly specific C9-carbonyl reduction (7, 25). A di-domain ARO/CYC then catalyzes the dehydration of the C9 hydroxyl, followed by first-ring aromatization (Fig. 1) (14). Therefore, in a nonreducing system, the growing poly-β-ketone intermediate is transported directly from the ketosynthase (KS) to the ARO/CYC. In contrast, in a reducing system, the intermediate needs to be transported from KS to KR and then to ARO/CYC. Before this study, there was no knowledge of whether there is any difference between the di-domain ARO/CYCs in reducing versus nonreducing PKSs, nor any information on the role that ARO/CYCs may play in determining the product specificity of reducing and nonreducing PKSs. The key issue that has hampered the investigation of this group of enzymes is the difficulty in protein crystallization.

To critically compare the reducing and nonreducing di-domain ARO/CYCs, we have chosen StfQ as a model nonreducing di-domain ARO/CYC and BexL as a model reducing di-domain ARO/CYC (Fig. 1). StfQ is a di-domain ARO/CYC from

Significance

Polyketides are a class of diverse natural products with well-documented bioactivity and medicinal importance. Enzymes known as aromatase/cyclases (ARO/CYCs) catalyze regiospecific cyclization and aromatization during type II polyketide biosynthesis. Understanding how ARO/CYCs catalyze cyclization and aromatization is critical for developing strategies for engineering biosynthetic pathways. This is the first study, to our knowledge, to use X-ray crystallography, bioinformatic and structural analysis, and in vitro functional assays to critically compare a reducing di-domain ARO/CYC (BexL) and a nonreducing di-domain ARO/CYC (StfQ). Together, these results fill in a missing link in the structural enzymology of polyketide biosynthesis and will have a direct effect on future biosynthetic engineering efforts and bioinformatic analysis of type II PKS gene clusters.

Author contributions: H.-w.L. and S.-C.T. designed research; G.C.-F., D.R.J., T.R.V., A.B.P., S.A., M.N., M.V., A.K., and E.S. performed research; G.C.-F., D.R.J., J.F.B., H.-w.L., and S.-C.T. analyzed data; and G.C.-F., D.R.J., H.-w.L., and S.-C.T. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. J.L.S. is a guest editor invited by the Editorial Board.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 4XRT and 4XRW).

[1]G.C.-F. and D.R.J. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: drjackso@uci.edu or sctsai@uci.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1512976112/-/DCSupplemental.

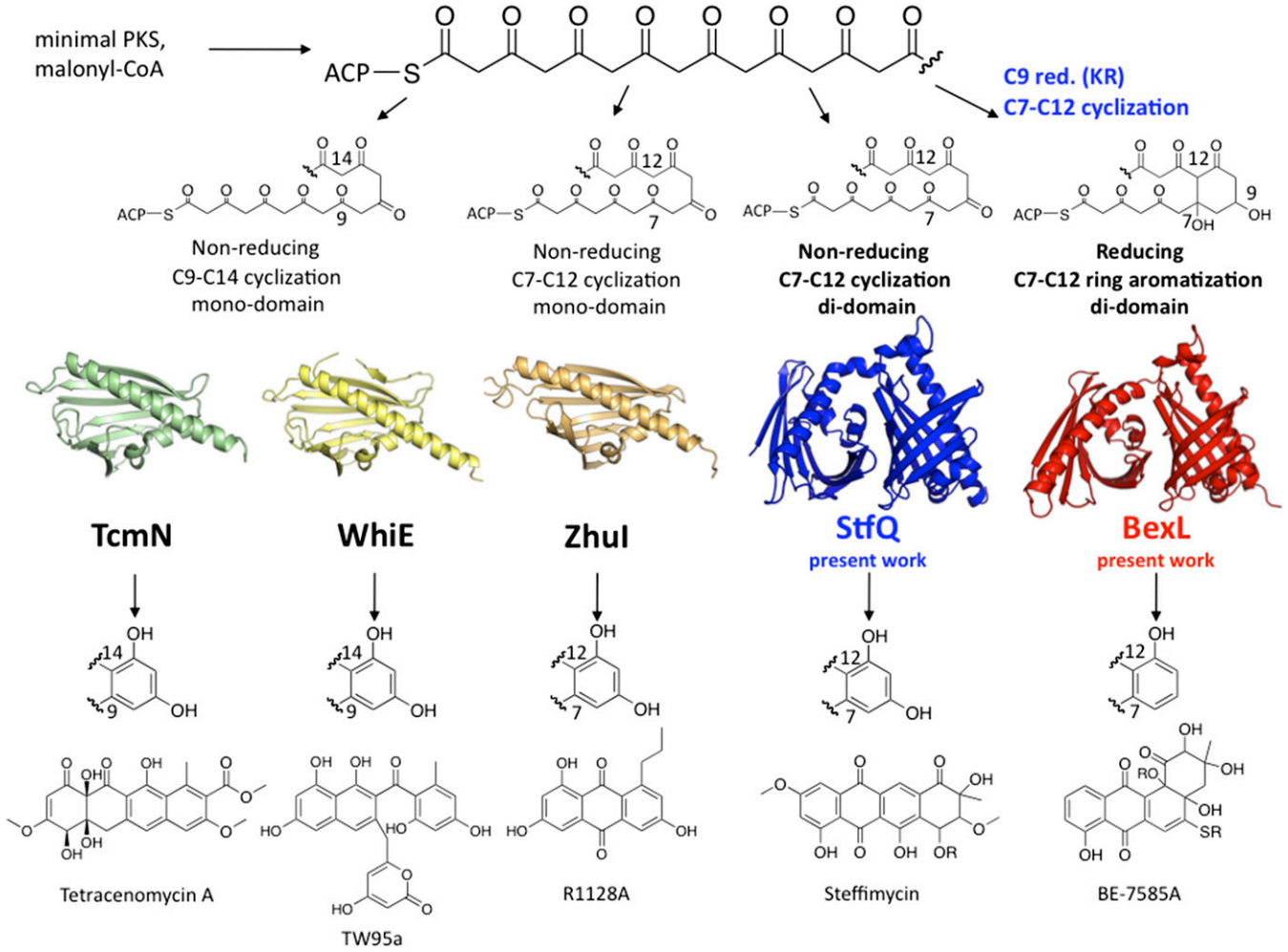

Fig. 1. Schematic diagram of ARO/CYC activity and cyclization specificity in representative type II PKSs. The monodomain ARO/CYCs TcmN (PDB ID 2RER) and WhiE (PDB ID 3TVR) act on unreduced polyketide intermediates to generate C9–C14 cyclized and aromatized products. The monodomain ARO/CYC ZhuI (PDB ID 3TFZ) and di-domain ARO/CYC StfQ act on unreduced polyketide intermediates to generate C7–C12 cyclized and aromatized products. The di-domain ARO/CYC BexL acts on C9 reduced, C7–C12 cyclized intermediates and catalyzes the aromatization of the C7–12 cyclized ring by dehydration of the C9 hydroxyl group.

Streptomyces steffisburgensis and is part of the nonreducing PKS catalyzing the biosynthesis of steffimycin, which shows promising antitumor activities (26–28). StfQ is the enzyme responsible for the C7–C12 first-ring cyclization and aromatization of the elongated poly-β-ketone substrate. In contrast, BexL is a di-domain ARO/CYC in the reducing PKS responsible for the biosynthesis of the anticancer agent BE-7585A. The linear poly-β-ketone precursor of BE-7585A is first cyclized, then reduced at the C9 position by KR, and then aromatized by BexL (Fig. 1) (24). Both StfQ and BexL use 20-carbon poly-β-ketone substrates; therefore, the structural enzymology of these two ARO/CYCs offers a great opportunity for side-by-side comparison with insight into the catalytic mechanisms of di-domain ARO/CYCs. Here we present the first structural and biochemical characterization, to our knowledge, of two di-domain ARO/CYCs, StfQ and BexL, using X-ray crystallography, structure-based mutagenesis, in silico docking, and in vitro functional assays.

Results and Discussion

Overall Structures. The monomers of BexL and StfQ both contain 309 amino acids that consist of two separate domains: the N-terminal (N-term) and C-terminal (C-term) domains (Fig. 2 and Fig. S1). In both enzymes, the N- and C-term domains adopt a helix-grip fold, which is found in other type II PKS ARO/CYCs (15–17). The BexL N-term consists of 146 amino acids, followed by a four-amino acid linker; the C-term has 160 amino acids. The StfQ N-term contains 144 amino acids, followed by a six-amino acid linker; the C-term has 157 amino acids. The overall sequence identity between BexL and StfQ is 32.0%. The identity of the BexL N-term (BexL residues 1–146) and StfQ N-term (residues 1–146) is 25.0%, whereas the C-term domains (BexL residues 147–309, StfQ residues 147–309) are 39.8% identical (Fig. S2). A detailed description of the secondary structural elements is given in the *SI Results and Discussion*. The individual N- and C-term domains of BexL and StfQ are structurally very similar to previously characterized monodomain ARO/CYCs, with the exception of the N- to C-term domain linkers (15–17). In StfQ, Nα3 turns into a six-amino acid α-helix linker to the C terminus. BexL contains a short unstructured linker sequence between the N-term α-helix and the start of the C-term domain.

The two domains of both proteins lay back-to-back, such that the β-sheets interact through hydrophobic interactions. This involves Nβ5–Nβ7 of the N-term and Cβ4–Cβ6 of the C-term domains. In BexL, both the N- and C-term pocket entrances are on the same face of the protein and border the axis that sits along the di-domain interface. Both the N- and C-term pocket entrances of

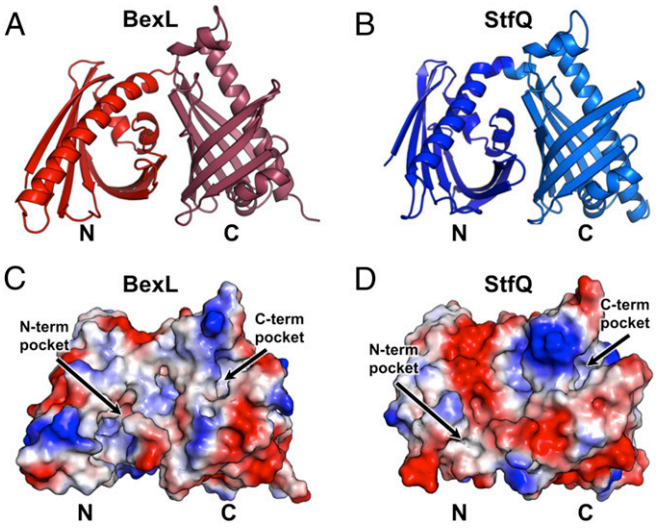

Fig. 2. A comparison of the tertiary structures and electrostatic surfaces of BexL versus StfQ. (*A*) Overall structure of BexL shows two distinct domains: N-term (*Left*) and C-term (*Right*). Both domains have a helix-grip fold with a long central α-helix surrounded by a series of β-strands. (*B*) Overall structure of StfQ similarly shows two distinct domains: N-term (*Left*) and C-term (*Right*). (*C*) Electrostatic surface potential map of BexL shows multiple positive patches (blue) that may bind ACP. (*D*) Electrostatic surface potential map of an StfQ monomer showing complementary positive (blue) and negative (red) patches, which form stabilizing interactions between the two StfQ monomers.

BexL display a wide entrance and are considerably positively charged, which may facilitate ACP docking for substrate delivery (Fig. 2). In contrast, the StfQ N- and C-term pocket entrances are pointed away from each other and open toward opposite faces of the StfQ dimer (Fig. 2). The StfQ N-term pocket entrance is smaller and has less positive character compared with the StfQ C-term pocket entrance. The size difference of the N- and C-term pocket entrances may have a direct effect on StfQ–ACP interactions, which would favor the ACP to deliver the growing polyketide into the C-term (as opposed to the N-term) for cyclization and aromatization.

StfQ and BexL N-Term Domain Comparison. The N-term of BexL has a much larger pocket than that of StfQ (Fig. 3). The rmsd of N-term (146 residues) between BexL and StfQ is 1.0 Å. The N-term of BexL has a pocket length of 22.2 Å, whereas StfQ has a much shorter pocket length of 15.7 Å (Fig. 3 *A* and *C*). This length difference is reflected in the pocket volume and area: the StfQ N-term pocket only has a volume of 523.7 $Å^3$ and an area of 500.5 $Å^2$, whereas the BexL N-term has a much larger volume of 1,722.4 $Å^3$ and an area of 885.7 $Å^2$. When comparing the BexL and StfQ N-terms, the central α-helix (Nα3) has the largest shift, with a 3.2 Å deviation (Fig. S1*C*). The Nα3 shift is a major element that accounts for the great difference in pocket volumes. Further, the loop connecting Nα3 to Nβ2 is longer in StfQ, resulting in a shorter central α-helix. In StfQ, the Nβ2 is puckered inward compared with BexL, which also pushes Nβ3 toward the pocket and repositions the Nβ2/Nβ3 loop that effectively closes the N-term pocket of StfQ.

The N-term domains of BexL and StfQ only have a 25.0% sequence identity. We analyzed the pocket in detail by dissecting the pocket from the back, to the middle, to the front. For both BexL and StfQ, there are several hydrophilic and bulky hydrophobic residues at the back of the N-term pocket, with notable differences (Fig. S2*A*). For example, R66 of BexL overlays perfectly with the previously proposed catalytic arginine of the monodomain ARO/CYCs TcmN, WhiE, and ZhuI (Fig. S3) (15–17). However, the StfQ N-term has a bulky W72 that perfectly aligns with R66 of BexL, resulting in the reduction of the pocket size of the StfQ N-term pocket. Other differences are F31 of BexL versus H36 of StfQ, F77 of BexL versus Y83 of StfQ, and H109 of BexL versus Y115 of StfQ. The middle of the pocket for both enzymes has several polar residues. The only identical residue is W62 in BexL and W68 in StfQ. The front of the N-term pocket for StfQ and BexL is flanked by a series of loops consisting of polar and nonpolar residues. It is thus clear that the most significant difference between the two ARO/CYCs is within the N-term domain. Our results reveal that StfQ has a relatively small N-term pocket, whereas BexL has an N-term pocket with a size comparable to that of the interior pockets of monodomain ARO/CYCs.

StfQ and BexL C-Term Domain Comparison. Unlike the N-term, the C-terms of StfQ and BexL are highly similar, with an rmsd of 0.78 Å and 39.8% sequence identity (Fig. S2). As a consequence, the C-term volume and area are highly similar, with an area of 796.7 $Å^2$ and a volume of 1,391.6 $Å^3$ for StfQ, and an area of 813.0 $Å^2$ with a volume of 1,339.8 $Å^3$ for BexL (Fig. 3 *B* and *D*). The sequence alignments also show that the C-term of BexL and StfQ are much more similar than their respective N-term domains (Fig. S2). The major catalytic residues identified in the monodomain ARO/CYC ZhuI (R66, H109, and D146) have similar counterparts in the C-term of StfQ (R218, H259, and N295), as well as both the N-term (R66, H109, and E135) and C-term (R218, H260, and T297) domains of BexL. When we overlaid the structures of ZhuI, StfQ (C-term), and BexL (N-term), the key residues matched perfectly (Fig. S3*A*). Again, we analyzed the pocket in detail by dissecting the pocket from the back, to the middle, to the front (Fig. S2*A*). At the bottom of the pocket, there are only a few residues that are different; for example, F200 in StfQ versus L200 in BexL, and N295 in StfQ versus T297 in BexL. Similarly, in the middle of the pocket, only one residue is different: L237 in StfQ versus M237 in BexL. The disparity is more apparent at the entrance of the pocket, which is defined by loop regions and may dictate specific interactions with their respective ACPs. The fact that the C-terms of both ARO/CYCs are highly similar to the C7–C12 monodomain ZhuI ARO/CYC strongly suggests that for StfQ, the catalytic event may occur in the C-term, but not N-term. In contrast, both the N- and C-terms of BexL may be functional.

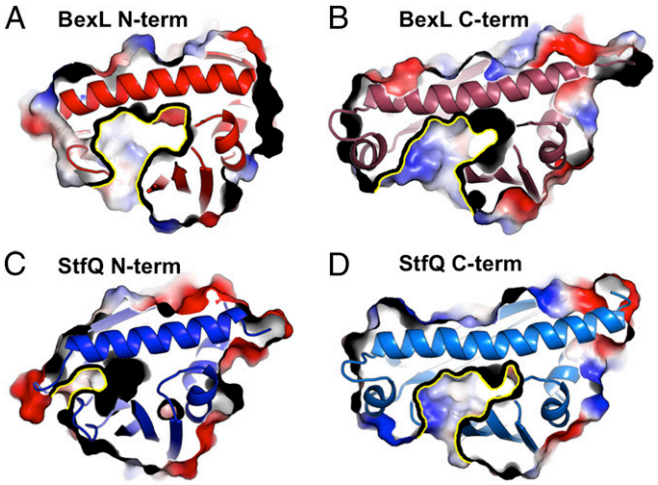

Fig. 3. Cutaway view of BexL and StfQ N- and C-term domains for a critical comparison of pocket sizes. (*A*) BexL N-term pocket. (*B*) BexL C-term pocket. (*C*) StfQ N-term pocket. (*D*) StfQ C-term pocket.

Sequence Comparisons of Reducing and Nonreducing ARO/CYCs. To further explore the similarities and differences between the reducing and nonreducing ARO/CYCs, a systematic search of reducing and nonreducing di-domain ARO/CYCs was conducted (Fig. S4). The N-terms of BexL and StfQ share a much lower sequence identity than that of the C-terms. When extending this comparison with the di-domain ARO/CYCs ActVII (reducing) and MtmQ (nonreducing), we observe a similar trend. For example, comparing ActVII and BexL, the N-term and C-term share 44% and 41% identity, respectively. In contrast, for the di-domain nonreducing ARO/CYCs, the N-term and C-term of StfQ and MtmQ share a 36% and 74% identity, respectively. Evidently, in the reducing ARO/CYCs, both the N-term and C-term domains are moderately conserved between related enzymes. However, the nonreducing ARO/CYCs display high protein sequence conservation of the C-term and low conservation of the N-term (Fig. S4).

StfQ N-Term Domain Mutagenesis. To determine residues that are important for catalysis in each domain of StfQ, we conducted extensive mutagenesis of both the N-term and the C-term of StfQ and evaluated the activity of the resultant StfQ mutants, using a linear poly-β-ketone substrate generated by a fungal PKS (PKS4) (29). If StfQ is active, we would observe C7–C12 cyclized product. If a mutation abolishes activity, we would observe only the minimal PKS shunt products (Fig. 4*C* and Fig. S5). Residues for mutagenesis were selected on the basis of a structural alignment with the ZhuI monodomain ARO/CYC, which has the same cyclization specificity (C7–C12) as StfQ. The StfQ N-term contains a tryptophan (W72) at the position of the conserved bottom-pocket arginine of known ARO/CYCs. Mutating W72 of StfQ resulted in insoluble inclusion bodies, suggesting W72 may be important for maintaining the overall fold of StfQ. Other putative catalytic N-term residues (H36, Y85, Q95, and Y115) were also mutated, but all these N-term mutants retain full enzyme activity. These structure-guided mutagenesis results suggest that the N-term domain of StfQ is not crucial for catalysis. However, some residues (such as W72) may be important for structural integrity.

StfQ C-Term Domain Mutagenesis. The StfQ C-term contains many hydrophilic residues, and its structure aligns well with both ZhuI and the BexL C-term. Ten site-specific mutants of StfQ C-term pocket residues were generated and assayed for activity (Fig. 4*D* and Fig. S5). Mutation of the key residues R218, Y228, W244, and H259, as predicted by the alignment with ZhuI, all led to >70% reduction in activity compared with wild-type (WT) StfQ. In particular, the R218A mutant exhibited a 90% reduction in activity. Similar to the ZhuI mutation result, the R218 to lysine or glutamine mutation significantly deactivated StfQ (16). Therefore, as previously noted for the monodomain ZhuI, the highly conserved R218 at the bottom of the StfQ C-term pocket plays a critical role in determining enzyme activity. Surprisingly, R218K mutant is significantly less active, suggesting the guanidinium moiety, not just a positively charged residue, is necessary for cyclization. Taken together, the side chain of R218 may serve as the active site acid or hydrogen bond donor, similar to the role played by its counterpart in the mechanism previously proposed for ZhuI. Because the substrate of StfQ is an unreduced, 20-carbon poly-β-ketone, it may interact with multiple hydrogen bond donors and acceptors on binding to the C-term pocket, to be correctly oriented for cyclization in the pocket, allowing R218-promoted C7–C12 aldol cyclization to occur. Additional C-term mutants including Y228F, Q230A, H240A, W244F, H259A, N291A, S292A, and N295A were also prepared. All eight mutants exhibited diminished activities (Fig. 4*D* and Fig. S5). Thus, contrary to the noncatalytic N-term, mutations of the C-term of StfQ significantly affect the cyclization activity in a manner that is consistent with the cyclization mechanism proposed for the monodomain C7–C12 ZhuI ARO/CYC (16).

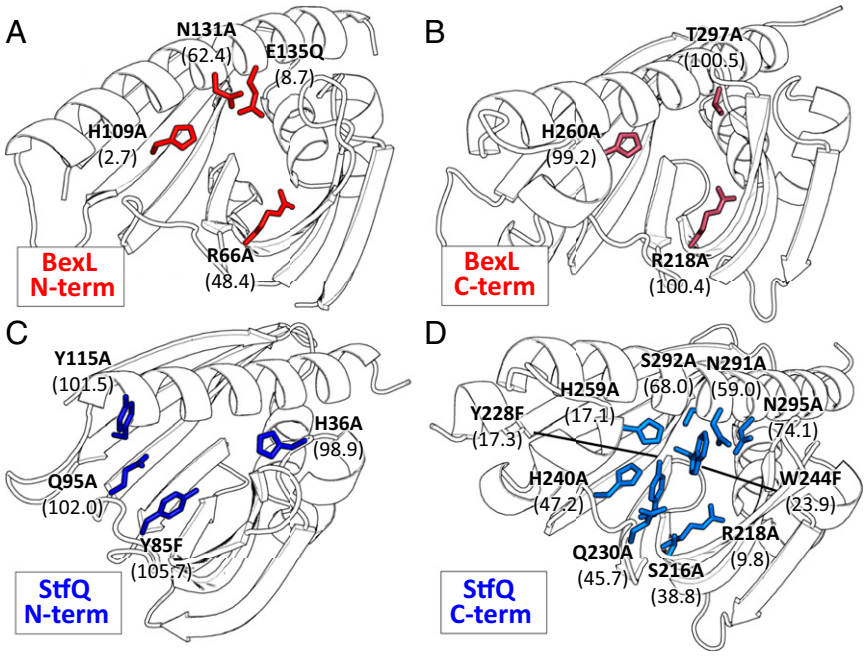

Fig. 4. In vitro assay and mutagenesis results. Each residue is labeled with the corresponding mutation and the percentage activity compared with WT. (*A*) BexL N-term pocket mutagenesis reveals that R66, H109, N131, and E135 are important for activity. Mutations of H109 and E135 are especially detrimental. (*B*) BexL C-term pocket mutagenesis reveals that C-term pocket residues do not contribute to BexL activity. (*C*) StfQ N-term pocket mutagenesis reveals that N-term pocket residues do not contribute to StfQ activity. (*D*) StfQ C-term pocket mutagenesis reveals that S216, R218, Y228, Q230, H240, W244, H259, N291, S292, and N295 are all important for StfQ activity. Mutations of R218, H259, and Y228 are especially detrimental.

BexL N-Term Domain Mutagenesis. On the basis of a structural alignment of ZhuI and the C-term domain of StfQ, we conducted extensive mutagenesis of both the N- and C-term of BexL and measured the activity of the mutants, using a coupled actinorhodin minimal PKS assay (30). Our results showed that the N-term R66, H109, and E135 are critical for activity (Fig. 4*A* and Fig. S6). However, R66A still retained 48.4% activity, whereas 66.3% and 82.2% activity were observed with a lysine and glutamine mutation, respectively. E135A was also constructed, but it resulted in inclusion bodies during protein expression. Fortunately, E135D and E135Q were generated as soluble proteins. E135Q had little activity (8.7%), but E135D had an activity of 56.4%. Hence, an acidic residue at this position seems to be important for the aromatization activity. This matches well with ZhuI, which has an aspartic acid instead of glutamic acid at this position and was found to be important for activity (16). Because N131A also retains 62.4% activity, N131 is unlikely directly involved in the aromatization catalyzed by BexL. Overall, contrary to StfQ, mutations of several key residues in the BexL N-term pocket do affect the enzyme activity.

BexL C-Term Domain Mutagenesis. In sharp contrast to the BexL N-term mutagenesis results, R218A, H260A, and T297A retain full activity comparable to that of WT BexL (Fig. 4*B* and Fig. S6). These results indicate that the C-term of BexL is not involved in catalysis. Although StfQ and BexL both have two domains with similar folds, the C-term pocket of StfQ is catalytically important, but not the N-term, whereas the opposite is true for BexL. This marked difference reflects the fact that one aromatizes a reduced substrate, whereas the other cyclizes an unreduced substrate.

Docking of a 20-Carbon Unreduced Poly-β-ketone with StfQ. A 20-carbon pPant-linked unreduced poly-β-ketone was docked in the N-term and C-term pockets of StfQ to investigate whether the internal pockets could accommodate the proposed substrate. The substrate docking experiment reveals potential interactions between the substrate and proposed active site residues in the C-term. Corroborating our mutagenesis results, the 20-carbon intermediate could not be docked to the StfQ N-term because of the small pocket dimensions observed in the crystal structure. In contrast, the StfQ C-term can accommodate a 20-carbon pPant-linked intermediate, and docking solutions revealed a myriad of solutions in different conformations. Of these solutions, a subset positions the C7 and C12 carbons within 5 Å of each other. At the same time, the C9 carbonyl group is within hydrogen bonding distance of R218 (Fig. S7*B*). Although many active site residues that are important for StfQ activity in vitro were identified in the mutagenesis experiments, it is difficult to assign specific roles of these residues such as general acid or base for the promotion of the C7–C12 aldol cyclization. Residues important for activity, such as R218, H259, and N295, are located away from each other in the crystal structure. When docked with the substrate, their specific roles can either be catalytic or promote a proper orientation of the poly-β-ketone for the C7–C12 aldol cyclization.

Alternatively, the StfQ C-term may adopt a more closed conformation on ACP/substrate binding, which may position these active site residues differently than in the crystal structure. This may account for the variation of their activities resulting from mutagenesis. Another possibility is that the putative active site residues simply act as a "mold" to help draw in and position the 20-carbon intermediate, which then spontaneously cyclizes between C7 and C12, with water serving as the general acid/base. This is an attractive hypothesis, because multiple docking solutions anchor the C9 carbonyl group near R218. This kinks the poly-β-ketone and forces C7 and C12 to be near each other, allowing spontaneous aldol cyclization. Therefore, it is reasonable to propose that StfQ (and other nonreducing ARO/CYCs) may simply fold the poly-β-ketone, using the arginine at the bottom of its pocket (R218) as an anchoring point to facilitate the proper cyclization pattern. Additional pocket residues may play assisting roles to accelerate the kinetics of the ARO/CYC-mediated aldol cyclization by providing protons, interacting with waters, and facilitating the binding and transfer of the poly-β-ketone intermediate in and out of the ARO/CYC pocket.

Docking of a 20-Carbon C9-Reduced C7–C12 Cyclized Poly-β-Ketone with BexL. To investigate whether the internal pockets of BexL can accommodate the proposed substrate, we docked a 20-carbon pPant-linked, C9-reduced, C7–C12 cyclized polyketide intermediate in the N-term and C-term pockets of BexL. It was found that the 20-carbon intermediate could readily be docked in both the BexL N-term and C-term because of the large pocket dimensions observed in the crystal structure. Of these results, a subset of solutions positioned the C9 hydroxyl group in hydrogen bonding distance to R66, which was shown by mutagenesis to be important for activity (Fig. S7*A*). The existence of multiple docking solutions suggest that the active site residues such as H109, E135, and R66 may have multiple roles, ranging from general acid or general base for promoting aromatization of C7–C12 ring (such as R66), or they may simply orient the incoming substrate for the aromatization to spontaneously occur (such as H109 and E135). In addition, the BexL N-term may adopt a more closed conformation on substrate-ACP binding. The BexL C-term domain was not analyzed in detail because site-directed mutagenesis suggested that the C-term is not important for catalysis. In summary, the active sites of BexL and StfQ are quite similar, and both bind C7–C12 cyclized intermediates. The active site arginine of both enzymes may be important for anchoring the respective substrate in such a way that promotes either cyclization/aromatization (for StfQ) or dehydration/aromatization (for BexL).

Functional Roles of the BexL N- and C-Term Domain Pockets. Both BexL and StfQ have two domains, and each domain has one interior pocket. In BexL, the functional role of each pocket was probed by site-directed mutagenesis. The BexL N-term pocket contains essential catalytic residues, but not the C-term pocket. Although the C-term pocket contains the putative active site residues (R218, H260, and T297) based on homology to other ARO/CYCs, mutations of these residues to alanine had no effect on enzyme activity. These results suggest that the N-term domain alone of BexL is responsible for dehydration/aromatization, and the C-term domain may be important for overall enzyme stability or protein–protein interactions with the ACP or other minimal PKS enzymes. In addition, we were unable to express the BexL N-term domain as a soluble protein, which further supports the hypothesis that the C-term domain may aid in overall protein stability.

Functional Roles of the StfQ N- and C-Term Domain Pockets. The StfQ N-term pocket contains many hydrophobic residues whose mutations resulted in very low expression and insoluble aggregates. This suggests that residues in the StfQ N-term are important for overall protein stability. Docking studies indicate that the StfQ N-term pocket is not large enough to accept a linear 20-carbon poly-β-ketone. However, a different conformation where the pocket volume increases may allow the binding of a cyclized intermediate. If this is the case, the hydrophobic nature of the StfQ N-term pocket may promote aromatization of a cyclized substrate. Docking studies show that the StfQ C-term domain can accept the full-length 20-carbon poly-β-ketone and allows the chain to fold appropriately for a C7–C12 cyclization. Mutational studies also strongly support that R218, H259, N295, and other hydrophilic residues are important for the cyclization of the poly-β-ketone. Hence, in contrast to BexL, residues in the StfQ N-term pocket appear to be structurally important, whereas

the C-term domain is responsible for catalyzing the first-ring C7–C12 cyclization (Fig. 5). We were unable to express the C-term domain as a soluble protein, which further supports the above hypothesis that the N-term domain aids in the structural integrity.

Functional Considerations of BexL and StfQ in Polyketide Biosynthesis. StfQ and BexL represent di-domain ARO/CYCs from two different type II polyketide synthases. Each of these PKSs (Bex and Stf) is responsible for the synthesis of aromatic polyketides with the same C7–C12 first-ring cyclization pattern. However, there is a key difference between these two PKSs: the presence (Bex) or absence (Stf) of a NADPH-dependent KR. On the basis of previous results (7, 25), we hypothesize that for the reducing PKS (Bex), the KR first cyclizes the elongated polyketide chain between C7 and C12 and then reduces the C9 carbonyl to a hydroxyl group. Upon aromatization by BexL, the hydroxyl group is eliminated as water. In contrast, for the nonreducing PKS (Stf), which lacks a KR, StfQ receives the unreduced, linear poly-β-ketone directly after elongation and proceeds with cyclization specifically between C7 and C12, followed by aromatization to form the first ring (Fig. 5). Our results clearly demonstrate that the two domains in di-domain ARO/CYCs do not have the same functions, and this is true for both reducing and nonreducing systems.

Biological Significance

Biosynthetic control at the enzyme level has been demonstrated to be an attractive means for the construction of new pharmaceuticals and bioactive molecules. However, the use of the biosynthetic machinery to make new natural products through control of functional group formation and cyclization patterns requires detailed knowledge of how the individual enzyme works in the biosynthetic pathway.

The reported work on the structures and functions of the StfQ and BexL di-domain ARO/CYCs has greatly advanced our understanding of the role of the di-domain within their respective PKSs. BexL is the first model of a di-domain ARO/CYC present in reducing PKSs that produce polyketides such as doxorubicin (anticancer), actinorhodin (antibiotic), and griseusin (anticancer) (19, 31, 32). Study of StfQ has given us the first insight into di-domains involved in the nonreducing PKSs that biosynthesize polyketides such as steffimycin (anticancer) and mithramycin (anticancer) (27, 33). StfQ mutagenesis has revealed that the conserved residues in the C-term that are important for the ZhuI mechanism are also important for StfQ. This suggests that ZhuI may be a bridge between the monodomain and the di-domain ARO/CYCs. In the reducing system, we demonstrated that the N-term domain of BexL, instead of the C-term in the nonreducing system, acts as a dehydratase/aromatase catalyzing the key reaction in the biosynthesis of aromatic polyketides.

Overall, this is the first molecular view, to our knowledge, of any di-domain ARO/CYC from either a reducing or nonreducing type II PKS system. Examples of both types are presented, displaying significant differences between these two classes of ARO/CYCs. The combination of this information with what we already know about monodomain ARO/CYCs will facilitate future engineering of PKSs to generate novel polyketide metabolites with alternative cyclization patterns.

Materials and Methods

Cloning, Expression, and Purification of BexL. A pET-28b(+) (Novagen)-derived DNA plasmid encoding N-term His-tagged BexL (BexL/pET28b) was prepared from cosmid C006 previously constructed from the genomic DNA of *Amycolatopsis orientalis* subsp. *Vinearia* (24). The BexL/pET28b plasmid was transformed into *Escherichia coli* BL21(DE3)-competent cells and plated on LB-agar plates containing kanamycin (50 μg/mL). The plates were incubated overnight at 37 °C. Positive transformants were transferred to a 5-mL starter culture of LB broth containing kanamycin (50 μg/mL) and grown overnight at 37 °C with shaking. The cultures were used to inoculate 1 L of LB with kanamycin (50 μg/mL) and grown at 37 °C until the A_{600} reached 0.4–0.6. The cells were then cooled to 18 °C, and 0.1 mM isopropyl β-thiogalactopyranoside (IPTG) was added to induce protein expression. After 12–18 h of incubation at 18 °C, the cells were harvested by centrifugation at 5,500 × *g* for 15 min. The cell pellets were flash-frozen in liquid nitrogen and stored at −80 °C. The frozen cell pellets were thawed on ice and resuspended in lysis buffer [50 mM Tris at pH 8.0, 300 mM NaCl, 15% (vol/vol) glycerol, and 10 mM imidazole]. The cell suspension was lysed using sonication (8 × 30 s cycles), and cellular debris was removed by centrifugation at 21,000 × *g* for 45 min. The lysate was incubated with 5 mL Ni-IMAC resin (BioRad) at 4 °C for 1 h. The resin was poured into a fritted column, and the flow through fraction was collected. The resin was washed with 100 mL lysis buffer and then eluted with lysis buffer plus increasing amounts of imidazole (20–500 mM). The elutions were analyzed using SDS/PAGE, and elutions containing the protein of interest were combined and concentrated by centrifugal filtration. The theoretical molecular weight of BexL is 36,148.4 Da, and was confirmed using SDS/PAGE.

Cloning, Expression, and Purification of StfQ. The expression and purification of StfQ is similar to the above procedure. The plasmid pYR31 encoding N-term His-tagged StfQ was provided by Yi Tang at the University of California, Los Angeles. pYR31 was transformed into *E. coli* BL21(DE3) and verified by gene sequencing. StfQ was expressed and purified using a similar procedure as described earlier. After Ni-IMAC purification, StfQ was buffer exchanged into the crystallization buffer (20 mM Tris at pH 8.0), using a PD-10 column (GE Healthcare), and the resulting pure native protein was concentrated to 13 mg/mL.

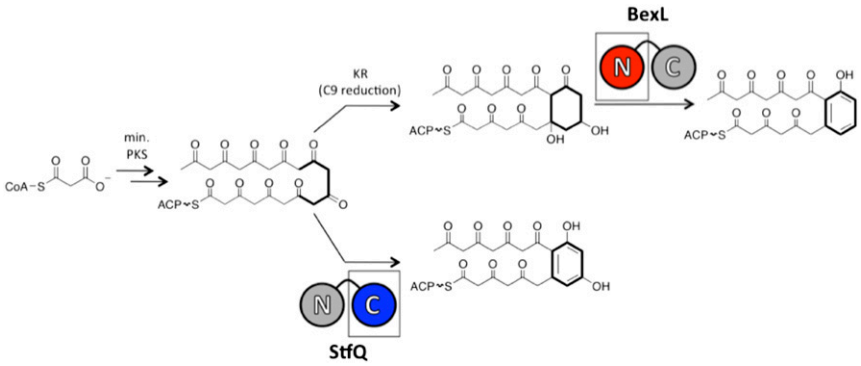

Fig. 5. The proposed mechanism for reducing versus nonreducing di-domain ARO/CYCs. During type II PKS biosynthesis, the ACP, KS-CLF, and MAT use malonyl-CoA to generate an unreduced polyketide intermediate. In nonreducing PKSs, KR first cyclizes the unreduced polyketide with C7–C12 cyclization specificity followed by reduction of the C9 carbonyl to a hydroxyl. BexL acts on the C7–C12 cyclized, C9-reduced intermediate to catalyze aromatization of the C7–C12 cyclized ring. The N-term domain of BexL is responsible for catalysis. In contrast, StfQ acts directly on the unreduced polyketide intermediate to yield a C7–C12 cyclized and aromatized product. The C-term domain of StfQ is responsible for catalysis.

The selenomethionine-substituted StfQ was produced in *E. coli* strain BL21 (DE3) in 2 × 1 LB broth containing kanamycin at 37 °C until A_{600} reached ~ 0.7. The cells were harvested (3,000 × *g* for 20 min) and resuspended three times with a total of 100 mL of M9 minimal medium containing the following amino acids: lysine hydrochloride, phenylalanine, threonine (5 mg each), isoleucine, leucine, valine (2.5 mg each), and L-selenomethionine (3 mg) (Sigma). The third resuspension was added to 2 × 1 L M9 minimal media plus kanamycin and the following amino acids: lysine hydrochloride, phenylalanine, threonine (50 mg each), isoleucine, leucine, valine (25 mg each); L-selenomethionine (30 mg) plus 1 mM IPTG to induce protein expression. Cells were grown for an additional 18 h at 18 °C. The cells were then harvested by centrifugation at 5,500 × *g*. The selenomethionine-substituted StfQ protein purification procedure was completed as described earlier for WT StfQ.

Crystallization and Data Collection. After Ni-IMAC purification, BexL was buffer exchanged into the crystallization buffer (20 mM Tris at pH 8.0), using a PD-10 column (GE Healthcare). BexL was crystallized using the sitting drop vapor diffusion method at room temperature from crystal seed stocks generated using a Seed Bead (Hampton). One microliter of 4 mg/mL protein solution and 1 μL well solution [0.1 M Mes (2-[*N*-morpholino]ethanesulfonic acid) at pH 7.0, 15% PEG 20000] were mixed and allowed to equilibrate over 500 μL well solution. Three-dimensional trapezoidal crystals formed overnight. Crystals were washed in well solution and then flash frozen in liquid nitrogen. X-ray diffraction data using monochromatic X-rays (0.96110 Å) were collected for native BexL crystals to 1.79 Å at Stanford Synchrotron Radiation Laboratory on beamline 12–2. Data were processed using HKL2000 (34).

Native and selenomethionine-substituted StfQ were crystallized by the hanging drop vapor diffusion method at room temperature. Two microliters of 8–10 mg/mL protein were mixed with 2 μL well solution [0.2 M sodium formate, 15% (vol/vol) PEG 3350] and equilibrated over 500 μL well solution. The crystals were cryopreserved in 20% glycerol and then flash-frozen in liquid nitrogen. X-ray diffraction data using monochromatic X-rays (0.96110 Å) were collected for native StfQ crystals to 2.2 Å at Stanford Synchrotron Radiation Laboratory on beamline 12–2. The multiwavelength anomalous diffraction data of selenomethionine-substituted StfQ crystals were collected to 2.70 Å at Advanced Light Source on beamline 8.2.1. Data were processed using HKL2000. The crystallographic statistics are listed in Table S1.

StfQ Phasing, Model Building, and Refinement. StfQ was crystallized in the space group $P2_12_12_1$, with two StfQ molecules forming a dimer in the asymmetric unit. The StfQ structure was solved using heavy atom phasing by multiwavelength anomalous diffraction to 2.7 Å by SHELX (35). A preliminary model was built using ARP/WARP, which was used for iterative rounds of model building (COOT) and refinement (PHENIX Refine) (36–38). The resulting multiwavelength anomalous diffraction structure was used to solve the 2.2-Å native StfQ structure by molecular replacement (Phaser) (39). This was followed by refinement (PHENIX Refine) and model building (COOT) to yield a structure with a R_{work} of 18% and R_{free} value of 21%. The following residues could not be confidently placed in the StfQ model because of missing electron density: M1-D7 (chain A), M1-D7, and D119-G122 (chain B). The StfQ structure has been deposited to the protein data bank (PDB ID 4XRT). The crystallographic statistics are listed in Table S1.

BexL Phasing, Model Building, and Refinement. BexL was crystallized in the space group $P4_12_12$, and Matthews coefficient analysis suggested one BexL molecule in the asymmetric unit (Vm. of 2.39 $Å^3$/Da and solvent content 48.56%). Initial phases for BexL were determined using molecular replacement (Phaser) with StfQ as a search model looking for one BexL molecule in the asymmetric unit (39). A preliminary model was built (PHENIX AutoBuild), and then this model was used for iterative rounds of model building and refinement (PHENIX Refine) (37, 40). This cycle was continued until R_{free} of the model was significantly reduced. At this point, Mes and waters were added to the model using COOT and PHENIX Refine, respectively. The model was further refined to achieve an R_{work} of 18% and an R_{free} of 22%. Data collection and refinement statistics can be found in Table S1. The following residues could not be confidently placed in the BexL model because of missing electron density: E44, G45, G251, and G252. The BexL structure has been deposited to the protein data bank (PDB ID 4XRW).

Site-Directed Mutagenesis. BexL and StfQ mutants were generated by site-directed mutagenesis using PCR, with mutagenic primers using BexL/pET28b and pYR31 as templates, respectively. Mutagenic primers are listed in Table S2.

Actinorhodin KS/CLF Expression and Purification. Actinorhodin KS (ActKS)/chain-length factor (CLF) was expressed in *Streptomyces coelicolor* CH999 containing the pRJC006 expression vector, as previously described, but with slight modifications (30). Fresh spores were used to inoculate 50 mL Super YEME containing 50 μg/mL kanamycin. The 50-mL cultures were grown for 3 d at 30 °C with shaking at 250 rpm. The 50-mL cultures were transferred to 2.5-L baffled flasks containing 500 mL Super YEME and 50 μg/mL kanamycin. The cultures were grown for an additional 2 d under the same conditions. Protein expression was induced with 10 μg/mL thiostrepton. Cells were harvested 24 h after induction by centrifugation (20 min at 5,500 × *g*), and the cell pellets were stored at −80 °C. Cell pellets from 2 L of cell culture were resuspended in 200 mL of lysis buffer [100 mM KP_i, 15% (vol/vol) glycerol, 300 mM NaCl at pH 7.5, with 2 EDTA-free protease inhibitor mixture (Roche) tablets]. The cell suspension was lysed by sonication on ice (10 × 1 min cycles). Cell debris was pelleted by centrifugation (1 h at 21,000 × *g*), and DNA was precipitated using streptomycin sulfate (2% final concentration), followed by centrifugation (45 min at 21,000 × *g*). The supernatant was filtered using a 0.45-μm filter, and 30–50% $(NH_4)_2SO_4$ was used to precipitate ActKS/CLF overnight. Precipitated protein was pelleted by centrifugation (30 min at 21,000 × *g*) and redissolved in Ni-binding buffer [50 mM KP_i, 10% (vol/vol) glycerol, 500 mM NaCl, 5 mM imidazole at pH 7.5]. After resuspension, the solution was filtered, using a 0.45-μm filter, and then bound to 2 mL Ni-IMAC resin (Bio-Rad) pre-equilibrated with Ni-binding buffer and stirred at 4 °C for 1 h. ActKS/CLF was eluted using Ni-binding buffer containing increasing amounts of imidazole. Fractions (2.5 mL each) containing between 375 and 750 mM imidazole were individually buffer exchanged into storage buffer [100 mM KP_i, 10% (vol/vol) glycerol at pH 7.5], using a PD-10 column (GE Healthcare), flash frozen in liquid nitrogen, and stored at −80 °C.

***Holo*-ActACP, ActKR, and MAT Expression and Purification.** The *S. coelicolor* MAT was purified from *E. coli*, as previously described (41). A pET-28a-derived ActACP expression plasmid was transformed into *E. coli* BAP1 cells that contain a chromosomally encoded copy of the phosphopantetheinyl transferase (Sfp) that ensures the production of *holo*-ACP. Expression and purification was conducted as described above for BexL. Purified ActACP was dialyzed into storage buffer [50 mM Tris, 100 mM NaCl, 10% (vol/vol) glycerol, 2 mM DTT at pH 8.0] and concentrated to 7 mg/mL. A pET28c-derived ActKR expression plasmid was transformed into *E. coli* BL21(DE3) cells and expressed and purified as described earlier for BexL, then dialyzed into storage buffer and concentrated to 8 mg/mL.

BexL in Vitro Assays. WT and mutant BexL activity was measured using an in vitro assay based on the actinorhodin minimal PKS (30). Next, 250-μL reactions containing 5 μM ActKS/CLF, 50 μM ActACP, 30 μM ActKR, BexL WT (or a BexL mutant), 2 μM MAT, 5 mM NADPH (Sigma) (when ActKR is present), and 5 mM malonyl-CoA (Sigma) were carried out in 100 mM KP_i at pH 7.5 at room temperature for 12 h. The reaction mixtures were extracted with 300 μL 94:5:1 of ethyl acetate:methanol:acetic acid, and the organic layer was dried using Speed Vac Concentrator. The residual oil was resuspended in DMSO (Sigma) and subject to reversed-phase HPLC analysis, using a Synergi Hydro-RP column (Phenomenex). The HPLC gradient was from 5 to 50% MeCN in a H_2O/0.1% formic acid mixture over the course of 15 min, followed by 50–95% MeCN in a H_2O/0.1% formic acid mixture over the course of 5 min. SEK34 and mutactin were identified using ESI-MS, UV-Vis absorbance, and relative retention times (42, 43).

PKS4 KS-AT and PKS4 ACP Expression and Purification. A PKS4 KS-AT expression plasmid was provided by Yi Tang, University of California, Los Angeles, transformed into *E. coli* BL21(DE3), and expressed and purified as described earlier for BexL. Purified PKS4 KS-AT was dialyzed into storage buffer [50 mM Tris at pH 7.5, 10% (vol/vol) glycerol, 100 mM NaCl, and 2 mM DTT] and then concentrated to 8 mg/mL The PKS4 ACP expression plasmid was transformed into *E. coli* BAP1 cells that contain a chromosomally encoded copy of the phosphopantetheinyl transferase (Sfp) that ensures the production of holo-ACP. Expression and purification was conducted as described earlier for BexL. Purified PKS4 ACP was dialyzed into storage buffer and concentrated to 7 mg/mL.

StfQ in Vitro Assays. WT and mutant StfQ activity was measured using an in vitro assay based on the fungal PKS4 minimal PKS (29). Next, 250-μL reactions containing 10 μM PKS4-KSAT, 50 μM PKS4 ACP, 30 μM, StfQ WT (or an StfQ mutant), 2 μM MAT, and 5 mM malonyl-CoA (Sigma) were carried out in 100 mM KP_i at pH 7.5 at room temperature for 12 h. Reactions were extracted and analyzed using the same procedure described earlier for the BexL assay. SEK4, nonaSEK4, PK8, and napthapyrone were identified using ESI-MS, UV-Vis absorbance, and relative retention times (29).

In Silico Docking Studies. The ligands were drawn in ChemDraw and then converted to PDB files using the NCI SMILES converter server (cactus.nci.nih.gov/translate/). The ligands were energy minimized and converted to .mol2 files using Chimera (44). The program GOLD was used for docking putative phosphopantetheine-linked intermediates into the C-term domain of StfQ and N-term domain of BexL (45). Docking runs for BexL and StfQ were performed by defining the ligand binding site within 20 Å of the active site arginine (R66 of BexL, or R218 of StfQ), using default settings for all parameters, with 100 docking trials per docking run.

Circular Dichroism of BexL and StfQ Mutants. Samples were collected on a JASCO J-810 Circular Dichroism Spectropolarimeter. All samples were prepared by diluting protein into 25 mM Tris at pH 7.5 to a final concentration of 0.1–0.2 mg/mL (Fig. S8).

ACKNOWLEDGMENTS. This work was supported by National Institute of General Medical Sciences Grant R01GM076330 (to S.-C.T.), NIH Grant GM040541 (to H.-w.L.), and Welch Foundation Grant F-1511 (to H.-w.L.).

1. Marti T, Hu Z, Pohl NL, Shah AN, Khosla C (2000) Cloning, nucleotide sequence, and heterologous expression of the biosynthetic gene cluster for R1128, a non-steroidal estrogen receptor antagonist. Insights into an unusual priming mechanism. *J Biol Chem* 275(43):33443–33448.
2. Rix U, Fischer C, Remsing LL, Rohr J (2002) Modification of post-PKS tailoring steps through combinatorial biosynthesis. *Nat Prod Rep* 19(5):542–580.
3. Shen B (2003) Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol* 7(2):285–295.
4. Hertweck C, Luzhetskyy A, Rebets Y, Bechthold A (2007) Type II polyketide synthases: Gaining a deeper insight into enzymatic teamwork. *Nat Prod Rep* 24(1):162–190.
5. Das A, Khosla C (2009) Biosynthesis of aromatic polyketides in bacteria. *Acc Chem Res* 42(5):631–639.
6. Zhou H, Li Y, Tang Y (2010) Cyclization of aromatic polyketides from bacteria and fungi. *Nat Prod Rep* 27(6):839–868.
7. Javidpour P, Korman TP, Shakya G, Tsai SC (2011) Structural and biochemical analyses of regio- and stereospecificities observed in a type II polyketide ketoreductase. *Biochemistry* 50(21):4638–4649.
8. Bartel PL, et al. (1990) Biosynthesis of anthraquinones by interspecies cloning of actinorhodin biosynthesis genes in streptomycetes: Clarification of actinorhodin gene functions. *J Bacteriol* 172(9):4816–4826.
9. Hopwood DA, Sherman DH (1990) Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu Rev Genet* 24:37–66.
10. McDaniel R, Ebert-Khosla S, Fu H, Hopwood DA, Khosla C (1994) Engineered biosynthesis of novel polyketides: Influence of a downstream enzyme on the catalytic specificity of a minimal aromatic polyketide synthase. *Proc Natl Acad Sci USA* 91(24):11542–11546.
11. Alvarez MA, Fu H, Khosla C, Hopwood DA, Bailey JE (1996) Engineered biosynthesis of novel polyketides: Properties of the whiE aromatase/cyclase. *Nat Biotechnol* 14(3):335–338.
12. Shen B, Hutchinson CR (1996) Deciphering the mechanism for the assembly of aromatic polyketides by a bacterial polyketide synthase. *Proc Natl Acad Sci USA* 93(13):6600–6604.
13. Zawada RJ, Khosla C (1997) Domain analysis of the molecular recognition features of aromatic polyketide synthase subunits. *J Biol Chem* 272(26):16184–16188.
14. Mcdaniel R, Ebertkhosla S, Hopwood DA, Khosla C (1994) Engineered Biosynthesis of Novel Polyketides - Act(Vii) and Act(Iv) Genes Encode Aromatase and Cyclase Enzymes, Respectively. *J Am Chem Soc* 116(24):10855–10859.
15. Ames BD, et al. (2008) Crystal structure and functional analysis of tetracenomycin ARO/CYC: Implications for cyclization specificity of aromatic polyketides. *Proc Natl Acad Sci USA* 105(14):5349–5354.
16. Ames BD, et al. (2011) Structural and biochemical characterization of ZhuI aromatase/cyclase from the R1128 polyketide pathway. *Biochemistry* 50(39):8392–8406.
17. Lee MY, Ames BD, Tsai SC (2012) Insight into the molecular basis of aromatic polyketide cyclization: Crystal structure and in vitro characterization of WhiE-ORFVI. *Biochemistry* 51(14):3079–3091.
18. Gullón S, et al. (2006) Isolation, characterization, and heterologous expression of the biosynthesis gene cluster for the antitumor anthracycline steffimycin. *Appl Environ Microbiol* 72(6):4172–4183.
19. Rajgarhia VB, Strohl WR (1997) Minimal Streptomyces sp. strain C5 daunorubicin polyketide biosynthesis genes required for aklanonic acid biosynthesis. *J Bacteriol* 179(8):2690–2696.
20. Menéndez N, et al. (2004) Biosynthesis of the antitumor chromomycin A3 in Streptomyces griseus: Analysis of the gene cluster and rational design of novel chromomycin analogs. *Chem Biol* 11(1):21–32.
21. Daum M, et al. (2009) Organisation of the biosynthetic gene cluster and tailoring enzymes in the biosynthesis of the tetracyclic quinone glycoside antibiotic polyketomycin. *ChemBioChem* 10(6):1073–1083.
22. Trefzer A, et al. (2002) Biosynthetic gene cluster of simocyclinone, a natural multihybrid antibiotic. *Antimicrob Agents Chemother* 46(5):1174–1182.
23. Xu Z, Jakobi K, Welzel K, Hertweck C (2005) Biosynthesis of the antitumor agent chartreusin involves the oxidative rearrangement of an anthracyclic polyketide. *Chem Biol* 12(5):579–588.
24. Sasaki E, Ogasawara Y, Liu HW (2010) A biosynthetic pathway for BE-7585A, a 2-thiosugar-containing angucycline-type natural product. *J Am Chem Soc* 132(21):7405–7417.
25. Javidpour P, et al. (2013) The determinants of activity and specificity in actinorhodin type II polyketide ketoreductase. *Chem Biol* 20(10):1225–1234.
26. Sriram M, Liaw YC, Gao YG, Wang AH (1991) Molecular structure of antitumor drug steffimycin and modelling of its binding to DNA. *J Biomol Struct Dyn* 9(2):251–269.
27. Brodasky TF, Mizsak S, Hoffstetter JR (1985) Steffimycin C, a new member of the steffimycin anthracyclines. Isolation and structural characterization. *J Antibiot (Tokyo)* 38(7):849–855.
28. Olano C, et al. (2008) Glycosylated derivatives of steffimycin: Insights into the role of the sugar moieties for the biological activity. *ChemBioChem* 9(4):624–633.
29. Zhang W, Li Y, Tang Y (2008) Engineered biosynthesis of bacterial aromatic polyketides in Escherichia coli. *Proc Natl Acad Sci USA* 105(52):20683–20688.
30. Matharu AL, Cox RJ, Crosby J, Byrom KJ, Simpson TJ (1998) MCAT is not required for in vitro polyketide synthesis in a minimal actinorhodin polyketide synthase from Streptomyces coelicolor. *Chem Biol* 5(12):699–711.
31. Rudd BA, Hopwood DA (1979) Genetics of actinorhodin biosynthesis by Streptomyces coelicolor A3(2). *J Gen Microbiol* 114(1):35–43.
32. Yu TW, Bibb MJ, Revill WP, Hopwood DA (1994) Cloning, sequencing, and analysis of the griseusin polyketide synthase gene cluster from Streptomyces griseus. *J Bacteriol* 176(9):2627–2634.
33. Lombó F, Braña AF, Méndez C, Salas JA (1999) The mithramycin gene cluster of Streptomyces argillaceus contains a positive regulatory gene and two repeated DNA sequences that are located at both ends of the cluster. *J Bacteriol* 181(2):642–647.
34. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
35. Schneider TR, Sheldrick GM (2002) Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr* 58(Pt 10 Pt 2):1772–1779.
36. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
37. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
38. Langer G, Cohen SX, Lamzin VS, Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3(7):1171–1179.
39. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
40. Terwilliger TC, et al. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* 64(Pt 1):61–69.
41. Kumar P, Koppisch AT, Cane DE, Khosla C (2003) Enhancing the modularity of the modular polyketide synthases: Transacylation in modular polyketide synthases catalyzed by malonyl-CoA:ACP transacylase. *J Am Chem Soc* 125(47):14307–14312.
42. Ma SM, et al. (2008) Redirecting the cyclization steps of fungal polyketide synthase. *J Am Chem Soc* 130(1):38–39.
43. Zawada RJ, Khosla C (1999) Heterologous expression, purification, reconstitution and kinetic analysis of an extended type II polyketide synthase. *Chem Biol* 6(9):607–615.
44. Pettersen EF, et al. (2004) UCSF Chimera–a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
45. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD (2003) Improved protein-ligand docking using GOLD. *Proteins* 52(4):609–623.