ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

Enzyme-Directed Mutasynthesis: A Combined Experimental and Theoretical Approach to Substrate Recognition of a Polyketide Synthase

Uschi Sundermann,[†,‡] Kenny Bravo-Rodriguez,[§] Stephan Klopries,[†] Susanna Kushnir,[†] Hansel Gomez,[§,⊥] Elsa Sanchez-Garcia,*[,§] and Frank Schulz*[,†,‡]

[†]Fakultät für Chemie, Chemische Biologie, Technische Universität Dortmund, Otto-Hahn-Str. 6, 44221 Dortmund, Germany

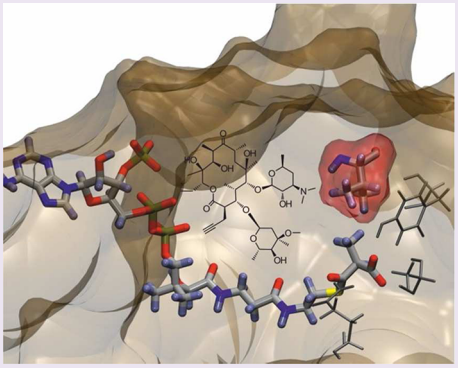
[‡]Max-Planck-Institut für molekulare Physiologie, Abteilung für Chemische Biologie, Otto-Hahn-Str. 11, 44227 Dortmund, Germany

[§]Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr, Germany

[⊥]Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès (Bellaterra), Spain

Supporting Information

ABSTRACT: Acyltransferase domains control the extender unit recognition in Polyketide Synthases (PKS) and thereby the side-chain diversity of the resulting natural products. The enzyme engineering strategy presented here allows the alteration of the acyltransferase substrate profile to enable an engineered biosynthesis of natural product derivatives through the incorporation of a synthetic malonic acid thioester. Experimental sequence−function correlations combined with computational modeling revealed the origins of substrate recognition in these PKS domains and enabled a targeted mutagenesis. We show how a single point mutation was able to direct the incorporation of a malonic acid building block with a non-native functional group into erythromycin. This approach, introduced here as enzyme-directed mutasynthesis, opens a new field of possibilities beyond the state of the art for the combination of organic chemistry and biosynthesis toward natural product analogues.



Natural products and their derivatives are frequently used as tools in modern biochemistry and are of high importance in medicine for the treatment of many different diseases. Of special relevance are polyketides, which form a large family of natural products and are mainly isolated from bacteria, fungi, and plants. They show diverse bioactivities and highly complex structures and have sparked the interest of biochemists and chemists alike.[1,2] Albeit being structurally diverse, polyketides share a common biosynthetic logic. In bacteria, giant Type I Polyketide Synthases (PKS) play a pivotal role in their biogenesis and share similarities with mammalian Fatty Acid Synthases (FAS).[3,4] In contrast to FAS, however, PKS direct the formation of a diverse array of functional groups and stereocenters.[2,5,6]

PKS catalyze a cascade of Claisen condensations between enzyme-bound acyl thioesters and malonic acid thioesters that serve as extender units of a nascent polyketide chain.[7] In optional steps further domains catalyze the stepwise reduction of the primary β-keto thioester product to the corresponding alcohol, olefin, or methylene group. During this reaction cascade, the growing polyketide chain is sequentially passed on from one catalytic enzyme domain to the next.[6,8] Structural diversity is brought about by variations in the redox pattern and stereochemistry as well as in the choice of biosynthetic building blocks.[2,9] Considering the complexity of polyketides found in various natural sources, redox variations are rather restrictive as only a small number of different functional groups can be introduced. A considerably larger diversity, ranging from plain hydrogen over long alkyl chains to primary amines in the side chain of the nascent polyketide, is achieved through the incorporation of different natural malonyl extender units.[2,10]

As a crucial means for the engineered biosynthesis of polyketide derivatives, generalizable strategies to influence building block diversity in PKS are highly sought after.[11−15] Current understanding of PKS enzymology suggests that the complex process of extender unit building block selection and incorporation is mainly controlled by acyltransferase (AT) domains, which catalyze the malonylation of acyl carrier proteins (ACP).[6] Building block specificity within PKS can be altered by transplanting AT domains between different PKS.[16−18] However, in many cases this domain swapping results in nonfunctional hybrid enzymes. In particular for AT domains, swapping has been shown to result in impaired protein folding.[19−21] A potentially more reliable alternative is the direct engineering of innate AT domains to alter their substrate scope by the replacement of selected amino acids instead of whole catalytic domains.[22,23] Nevertheless, the

Received: September 23, 2012
Accepted: November 18, 2012
Published: November 26, 2012

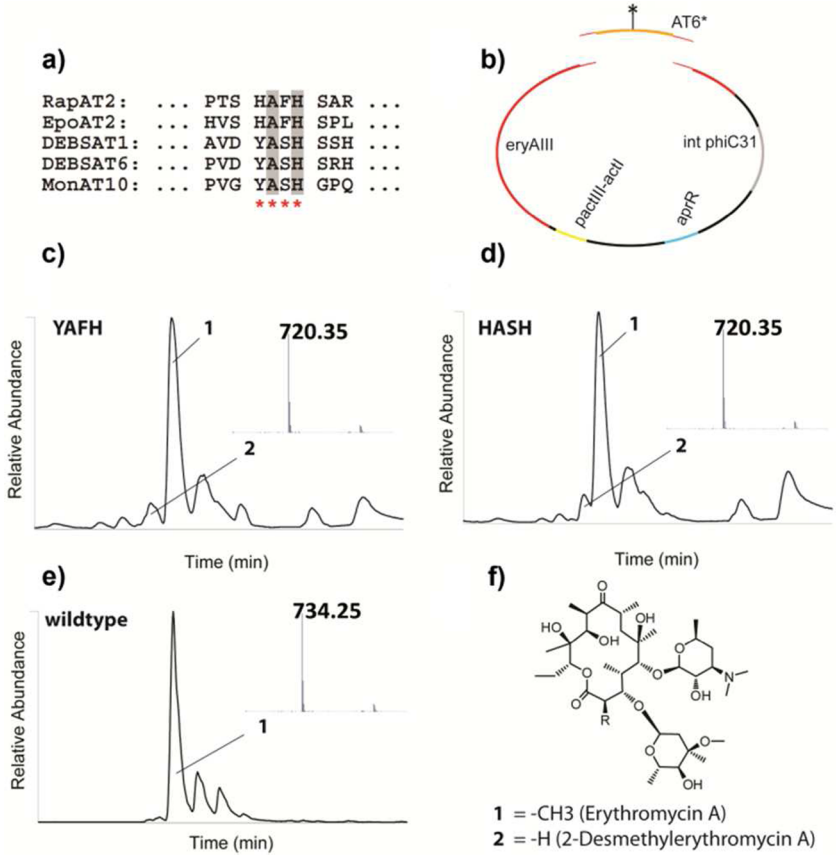


Figure 1. Experimental sequence–function correlations. (a) Protein sequence alignments guide identification of residues that correlate with substrate specificity of AT domains. Rap, Rapamycin PKS; Epo, Epothilon PKS; Mon, Monensin PKS. (b) Sequence-and-ligation independent cloning is used to construct a library of 266 different expression plasmids for DEBS3 carrying defined mutations in AT6. An overview of the complete biosynthetic pathway toward erythromycin A can be found in Supplementary Figure 1. Vector characteristics: *eryAIII*, gene encoding DEBS3; pactIII-actI, promoter sequence from the Actinorhodin biosynthetic gene cluster; *aprR*, Apramycin resistance cassette; int phiC31, integrase from the bacteriophage φC31. (c) LC–ESI-MS of a crude fermentation extract, showing the production of erythromycin A (**1**) and the desmethyl derivative (**2**) by a *S. erythraea* AT6-variant carrying the previously described[28] YAFH-hybrid motif. The chromatograms show the base peak from 2 to 10 min; the most significant signal in the relative abundance is erythromycin A, and the desmethyl derivative elutes 0.2 min ahead of the wild-type product. (d) Analysis of a variant carrying the analogous mutation in the YASH-fingerprint motif toward HASH. (e) LC–ESI-MS analysis of the wild-type erythromycin-producer for comparison, showing no formation of **2**. (f) Structures of erythromycin A and its biosynthetic derivative 2-desmethylerythromyin A.

complexity of PKS and the lack of understanding of the recognition of malonyl thioester building blocks by these enzymes render the directed mutagenesis of AT domains a formidable task, thereby leaving the search for a reliable and versatile strategy for extender unit engineering in PKS open.[24]

Herein, we present a model for substrate recognition and turnover by a generic AT domain based on a combination of systematic enzyme mutagenesis and molecular modeling. Analysis of the sequence–function relationship provides a dynamic picture of the interaction between enzyme and substrate. On the basis of these findings, we present an enzyme-directed mutasynthesis in which enzyme mutagenesis is used to govern the incorporation of a fully synthetic and not biosynthetically known building block into a polyketide. By allowing rational enzyme manipulation, this methodology opens a whole new range of so far unexplored possibilities in the enzyme engineering field.

RESULTS AND DISCUSSION

Establishing Sequence–Function Relationships. A number of different sequence regions in AT domains are discussed as potential determinants of their substrate scope.[17,22,25−29] A transferable strategy for the mutagenesis-induced acceptance of entirely artificial building blocks with non-natural side chains could dispense with the need to transplant AT domains. Instead, it preferably relies on a mechanism-based mutation of few residues rather than the swapping of long stretches of amino acids. Since accumulated evolutionary differences would have blurred the results on the comparatively rarely found AT domains with uncommon substrate profiles (Supplementary Figure 2), a sequence–function relationship study was initiated here on the widely distributed methylmalonyl-CoA (MMCoA)-specific domains. We anticipated to identify broadly conserved critical residues in their sequences, which would potentially enable a transferability of the results without the need for domain swapping.[13,20,30−34]

An archetypical example for MMCoA-specific AT domains is found in 6-deoxyerythronolide B Synthase (DEBS),[35,36] the key multienzyme complex in the biosynthesis of the antibiotic erythromycin from the Gram-positive bacterium *Saccharopolyspora erythraea*. In this PKS, all AT domains exclusively accept MMCoA as substrate[10,21] and share significant sequence homology among each other. As model for our investigations the final acyltransferase domain in the assembly line, AT6, was

dx.doi.org/10.1021/cb300505w | *ACS Chem. Biol.* 2013, 8, 443–450

chosen. AT6 shares high sequence homology with the previously crystallized AT5[29] from the adjacent module 5 and AT3 from module 3[27] and is located in the last extension module of the PKS. This might reduce the probability of a potentially stalled biosynthesis upon incorporation of a non-native extender unit when being passed on to a downstream module, thereby delivering false negative results on the AT domain specificity. However, the terminal thioesterase domain might still discriminate a non-native substrate. Protein sequence analysis indicated that a relatively small number of amino acid residues might determine the substrate specificity of AT domains. Previous analysis of the myxobacterial Epothilon biosynthetic gene cluster revealed two AT domains that differ in only eight residues but possess different substrate specificities.[37] We identified in total 14 residues in DEBS AT6 that were homologous in either the sequence or the modeled structures to the critical positions in Epothilon PKS AT2 and AT3 (Supplementary Figure 2).

To establish sequence−function relationships for the selected residues, we decided to explore the sequence space using saturation mutagenesis. This technique is often applied in the directed evolution of biocatalytically used enzymes[38−40] but has so far not been successfully transferred to multienzyme complexes such as PKS. Henceforth, we implemented a saturation mutagenesis methodology adapted to the large size and high GC-content of DEBS, which enabled the systematic exploration of the relevance of a given position in DEBS AT6. Importantly, the experiments were all carried out in a *S. erythraea* strain carrying the complete assembly line of erythromycin A. This significantly complicated experimentation in comparison to the use of simplified PKS systems in heterologous expression hosts such as *E. coli*.[41] The use of simplified systems, however, has been shown in the past to potentially yield non-transferable results and was thus not considered for our experimentation.[42]

To obtain clear data on AT sequence−function correlation with reasonable screening effort, we decided to prepare each of the resulting 266 enzyme variants in the saturation libraries individually with defined mutagenic oligonucleotides. The residual erythromycin production level was taken as indicator of general functionality of the enzyme variant; for the initial detection of an increased substrate promiscuity of AT6 variants, the incorporation of the endogenous malonyl-CoA (MCoA) in the last extension module of DEBS was used, leading to 2-desmethylerythromyin. Detection and relative quantification of the two metabolites were accomplished by comparative LC−ESI-MS and MS/MS (Figure 1).

The 266 individual expression plasmids were constructed from a bacteriophage-integrase based vector backbone, enabling comparatively rapid and reliable genetic manipulation of *S. erythraea*. The PCR-based mutagenesis was accomplished employing the recently developed SLIC-MIX protocol,[43] and the prepared plasmids were individually introduced into a *S. erythraea* ΔAT6hygR variant carrying a nonfunctional copy of DEBS3 (Figure 1 and Supporting Information sections I.3 and I.4 for details on the hygR-cassette integration in place of the native *eryAIII* gene). The identity of each mutated plasmid was confirmed by DNA sequencing. Copies of *S. erythraea* NRRL B-24071 wild type and a simultaneously prepared reconstituted wild type based on *S. erythraea* ΔAT6hygR were cultivated in parallel, serving as standards for the relative quantification of erythromycin fermentation levels. To assess the statistical significance of observed productivity changes, four clones of each variant were searched for outliers, and subsequently the averaged erythromycin productivity value was compared to 60 independent clones of the reconstituted wild type. This determined the extent of the influence of the genetic manipulation itself onto the productivity. The relative quantification allowed for the identification of individual mutations which exert a significant effect on the enzyme activity.

LC−ESI-MS screening of the 266 AT6 variants revealed fewer than 20% variants with a complete loss of activity (Figure 2), indicating that the targeted residues have on the average limited impact on the structural integrity of the enzyme.

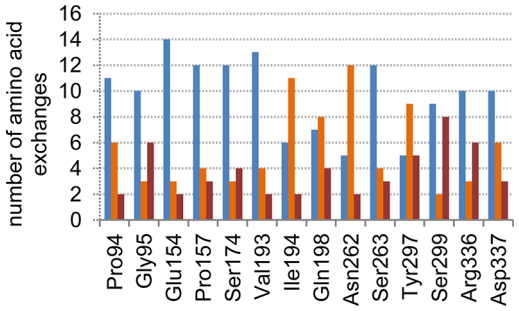


Figure 2. Distribution of residual activity of the erythromycin assembly line upon saturation mutagenesis at the 14 selected positions. The fraction of enzyme variants with no significant changes in activity is often high, correlating with the location of the residues further away from the active site. Highly sensitive regions were chosen for theoretical investigation as these residues apparently had a strong impact on catalysis or potentially protein folding. Blue: number of mutations with no significant impact on productivity levels (10−50% of wt activity). Orange: mutations with negative impact on activity (<10% of wt activity). Red: loss-of-function mutations in which over several independent clones no residual erythromycin production was detected.

A frequently made observation in enzyme engineering experiments is that a range of different mutations lead to similar results, especially when saturation mutagenesis is applied. However, in the case of AT6, out of all tested enzyme variants, many of them carrying mutations around the active site, only three resulted in the formation of 2-desmethylerythromyin (**2**) (Figure 1c−e). Two of these variants carried a single mutation in the YASH fingerprint motif that is known to correlate with MMCoA specificity in AT domains.[22,26] The two variants possess the sequences HASH and YAFH in place of the fingerprint motif, both of them hybrids of YASH and the analogous HAFH motif, which indicates MCoA specificity. Additionally, the mutation Q198L resulted in very small quantities of **2** as a fermentation byproduct.[22]

The influence of the mutations is highly specific; the very similar YAYH in place of the fingerprint motif, for example, rendered the enzyme almost completely inactive, the low residual activity exclusively delivering the wild-type product. Also at other positions, subtle alterations in the amino acid residue sometimes led to drastic changes in overall activity. For example, the exchange of Gln198 to the basic residue arginine (Q198R) resulted in no significant alteration of fermentation activity. However, mutation to another basic residue, lysine (Q198K), nearly completely abolished the activity.

Role of Specific Residues for Enzymatic Activity. To initiate a systematic interpretation of the complex experimental findings, we constructed a homology model of the DEBS AT6

dx.doi.org/10.1021/cb300505w | *ACS Chem. Biol.* 2013, 8, 443−450

domain with its substrate MMCoA bound to the active site. So far, no crystal structures for AT domains with bound substrate are described, leaving the binding mode of MMCoA unclear. The crystallized acyltransferase domains of modules 3 and 5 of DEBS[27,29] share a high sequence identity with the query sequence (>40% of sequence identity) and therefore were considered as suitable templates to deliver a reliable structural model. The resulting DEBS AT6 structure was used as starting point for docking (2*S*)-methylmalonyl-CoA into the proposed active site (variants with docked (2*R*)-methylmalonyl-CoA were also theoretically and experimentally investigated and will be published elsewhere). The obtained structure was subjected to a 30 ns molecular dynamics simulation (MD) (Figure 3).

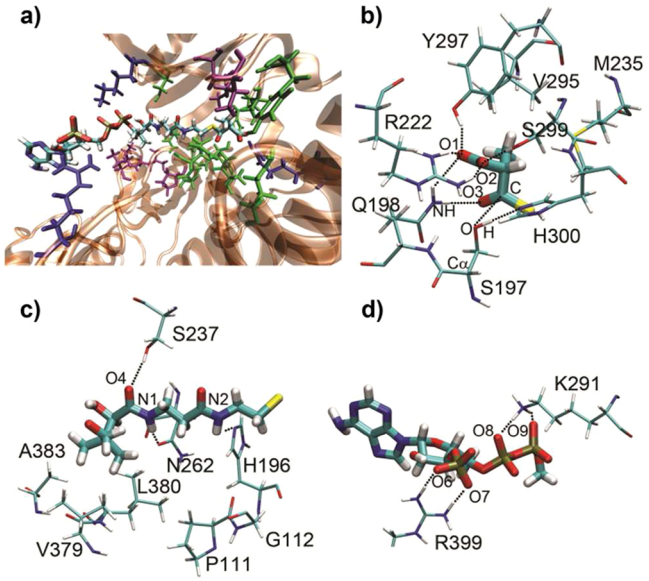


Figure 3. Molecular interactions between MMCoA and the binding pocket of AT6. (a) Residues interacting with MMCoA highlighted by polarity: green, polar; magenta, nonpolar; blue, basic. (b) Interactions in the thioester region. (c) Interactions in the isopropyl region. (d) Interactions in the phosphate region.

After that, randomly selected snaphots were optimized at the Quantum Mechanics/Molecular Mechanics (QM/MM) level. The same methodology (MD simulations and QM/MM optimizations) was also applied to the mutated systems. Since MD simulations allow sampling a larger conformational space, we focus our discussion on the MD results, which are supported by the good agreement with the QM/MM calculations (see Supplementary Table 7 and computational section for details).

Analysis of the molecular interactions that were time-conserved during the molecular dynamics simulations of the wild-type AT6-MMCoA system allowed us to assess the quality of the model. Importantly, our model indeed showed the previously proposed role of the conserved His300 in activating the catalytic Ser197 for nucleophilic attack on the substrate thioester. Accordingly, the hydroxyl hydrogen atom of Ser197 interacts with the unprotonated nitrogen atom of His300 ($H_{197}\cdots N_{300}$ distance is 2.13 ± 0.36 Å), and the distance between the nucleophilic oxygen atom of Ser197 (O) and the electrophilic carbon atom of MMCoA (C) (3.69 ± 0.34 Å, Figure 3b) suggests the presence of a prereactive complex.

We also found that in the wild type system, the residues Arg222, Tyr297, and Ser299 are important for substrate positioning. Thus, Arg222 interacts frontally and not symmetrically with both the carboxyl oxygen atoms of MMCoA ($NH1_{222}\cdots O1_{MMCoA}$ and $NH2_{222}\cdots O2_{MMCoA}$ distances of 2.50 ± 0.31 and 1.87 ± 0.21 Å) (Figure 3b). At the same time, Tyr297 and Ser299 also interact from above the plane defined by the carboxylate group with the O1 and O2 oxygen atoms (Figure 3b) ($OH_{297}\cdots O1_{MMCoA}$ distance of 1.72 ± 0.14 Å and $OH_{299}\cdots O2_{MMCoA}$ distance of 1.76 ± 0.14 Å).

Importantly, residue Gln198 interacts with the thioester group ($NH_{198}\cdots O3_{MMCoA}$ distance of 1.98 ± 0.23 Å) and is located 3.36 ± 0.34 Å away from the O1 atom, which also contributes to restrict the rotation freedom of the carboxyl group of MMCoA. Thus, Gln198 plays two roles: it promotes the activation of the thioester group and helps to keep MMCoA in a conformation suitable for the nucleophilic attack of Ser197.

It is noteworthy that Asn262, Gln112, His196, and Ser237 create a polar environment around the amide groups of MMCoA while the nonpolar isopropyl moiety of MMCoA is located in a nonpolar region formed by Pro111, Val379, Leu380, and Ala383 (Figure 3c). Although the phosphate region of MMCoA is exposed to the solvent, its conformational flexibility is restricted by the interactions of Lys291 and Arg399 with the phosphate groups (Figure 3d).

Not only the wild type but also selected mutated systems were modeled. An interesting starting point was the comparison of the functionally neutral mutation Q198R with the chemically similar but deleterious mutation Q198K. Our MD simulations indicate that the difference in length and flexibility of the lateral chain of the amino acids glutamine, arginine, and lysine is crucial for the explanation of this experimental observation. In the case of Q198R, the rigid lateral chain of arginine and its side by side interaction with MMCoA keep the malonyl-thioester in a position susceptible to be attacked by Ser197. In the case of Q198K, the greater flexibility of the lateral chain of lysine allows MMCoA to be positioned on top of Lys198 at a MMCoA−Ser197 distance that prevents the nucleophilic attack of Ser197 ($C\alpha_{197}\cdots C_{MMCoA}$ distance 6.12 ± 0.34 Å, see Supplementary Table 7 for more details). Another factor contributing to the forfeiture of biological activity is that the interaction between Tyr297 and the carboxylate group of MMCoA appears to be lost in the Q198K variant.

Tyr297 is one of the residues in the YASH fingerprint sequence. Mutation Y297H leads to a variant that retains the native catalytic activity on top of delivering 2-desmethylerythromyin (**2**) (Figure 1). The mutation appears to result in MMCoA being positioned over the plane defined by the guanidinium group of the lateral chain of Arg222. It is worth remarking that in all other systems studied here MMCoA is placed under this plane. The over the plane positioning of MMCoA is an exceptional behavior related to the partial loss of polarity in the Y297H binding pocket region close to the thioester region of MMCoA. The $O_{197}\cdots C_{MMCoA}$ distance of 4.34 ± 0.44 Å and the $C\alpha_{197}\cdots C_{MMCoA}$ distance of 6.48 ± 0.36 Å also reflect the different positioning of MMCoA in Y297H with respect to the wild type (Figure 3b, Supplementary Figure 8 and Supplementary Table 7). Since here the $C\alpha_{197}\cdots C_{MMCoA}$ distance is even larger than in Q198K the question of why this mutant still shows biological activity (unlike Q198K) arises. This can be understood in terms of the different susceptibilities of both systems to be involved in molecular interactions that restrain the mobility of MMCoA after activation of Ser297. Although in Y297H the smaller size of histidine with respect to

dx.doi.org/10.1021/cb300505w | *ACS Chem. Biol.* 2013, 8, 443−450

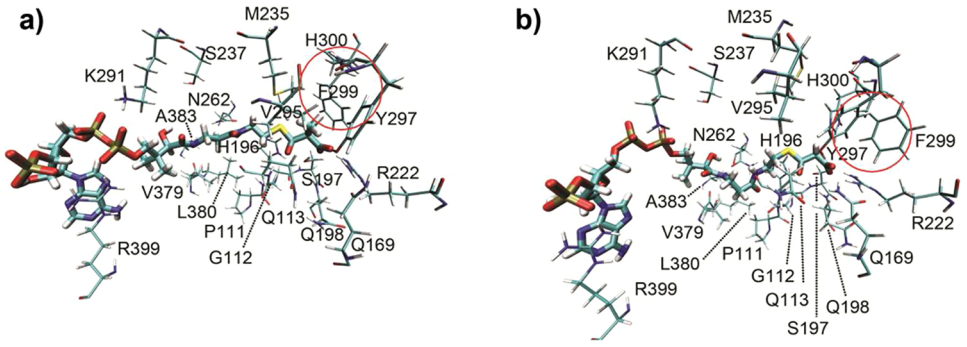


Figure 4. Incorporation of non-native malonates into polyketide biosynthesis. (a) S299F variant with bound MMCoA. (b) The same variant with bound MCoA. Relevant conformational changes are highlighted.

tyrosine results in larger MMCoA–Ser197 distances, the absence of secondary strong interactions with a charged amino acid (unlike in Q198K) results in MMCoA being more flexible and therefore able to approach Ser197 for the nucleophilic attack once the activation has taken place.

Ser299 is also part of the YASH fingerprint motif and was subjected to saturation mutagenesis. The selected variants S299V, S299F, and S299Y were modeled as variants with significantly different experimentally determined properties. S299V experimentally leads to the wild-type product, S299F to the production of 2-desmethylerythromyin, and the closely related S299Y to a loss-of-function variant. In the MD simulations we found a correlation between the loss of activity and the size of the amino acid side chain. The larger size of the phenylalanine and tyrosine side chains with respect to serine results in the lateral chain of the catalytic His300 being pushed away from the binding pocket with the consequent weakening of the His300–Ser197 interaction (Supplementary Table 7).

Thus, in these enzyme variants the activation of the catalytic Ser197 probably takes place in a less efficient fashion through the carboxylate group of MMCoA itself, which is a weaker activating group than His300. The bigger size of the amino acid in position 299 pushes down MMCoA and Gln198 can no longer interact with the thioester group of the substrate as it does in the wild type. Thereby, all three mutations appear to have a non-favorable effect on the interactions that activate both the substrate thioester and the catalytic Ser197 (Figure 3 and Supplementary Table 7). The S299Y mutation, however, causes a complete disruption of the interaction pattern between the binding pocket and MMCoA. Almost all key interactions are lost so that MMCoA is free to move inside the binding pocket, which becomes wider and unstructured (Supplementary Figure 9), potentially leading to a largely misfolded enzyme.

A particularly interesting case is the mutation S299F, which was found to lower the substrate specificity of the protein in a similar fashion as Y297H (Figure 1d). The MD simulations of S299F show that the benzyl group of the phenylalanine is able to rotate (Figure 4). This rotation partially reduces the space available to accommodate MMCoA inside the binding pocket, thereby resulting in erythromycin and 2-desmethylerythromyin as fermentation products. MD simulations with the non-native MCoA substrate inside the S299F variant show that in the absence of the methyl group, the phenyl ring is placed outside the binding pocket, allowing for a better positioning of MCoA (Figure 4b and Supplementary Table 7).

Enzyme-Directed Mutasynthesis. A major objective of this study was to devise a strategy for the engineered incorporation of non-natural, ideally fully synthetic, malonates into polyketide biosynthesis as an extension to the hardly predictable classical mutasynthesis approach.[24] Accordingly, the whole AT6 variant library was recultivated in the presence of 2-propargylmalonyl-SNAC (**3**) (Figure 5), a building block that would introduce an easily modifiable functional group into the natural product, provided that one of the AT6 variants could accept it as substrate. However, none of the test fermentations showed the formation of the sought-after 2-propargylerythromycin (**4**) to a significant extent. Further analysis of the MD and QM/MM structures of AT6-MMCoA pointed out a residue that did not appear in the sequence alignment studies to be critical for substrate selection. In the modeled structures, however, the largely conserved Val295 (Figure 6, see

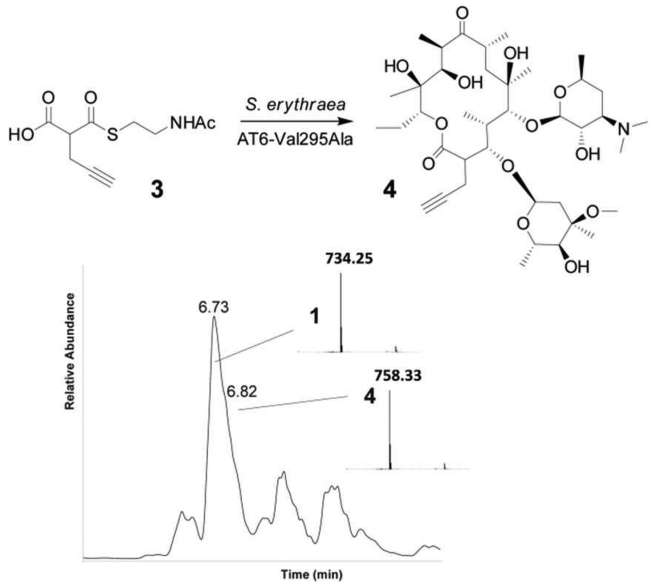


Figure 5. Scheme for the formation of compound **4** by feeding the synthetic malonate **3** to a mutated strain of *S. erythraea* and TIC of an LC–ESI-MS analysis of the crude 2-propargylerythromycin (**4**), as obtained in the enzyme-directed mutasynthesis approach upon feeding the *S. erythraea* mutant DEBS AT6-V295A with compound **3** in the growth medium. The chromatogram reveals the co-fermentation of wild-type product **1** in lower quantities than in the reconstituted wild type with a slight retention time difference, indicating that the substrate scope is partially shifted. Interestingly, the post-PKS processing of the primary product 6-deoxyerythronolide B *via* a series of glycosylations and hydroxylations proceeds in the same fashion for the polyketide derivative.

dx.doi.org/10.1021/cb300505w | *ACS Chem. Biol.* 2013, 8, 443–450

Supplementary Figure 2 for alignment) seemed to constrain the size of the malonate side chain, preventing more spacious polyketide building blocks from incorporation.

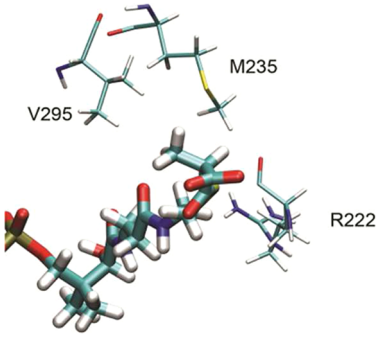


Figure 6. Position of Val295 and the opposite Met235, acting together in constraining the choice of the substituent at position 2 of the enzyme-bound malonic acid thioester.

We thus devised mutations at position 295 with a lower sterical hindrance toward position 2 of the enzyme-bound malonate. Screening a mini-library of three variants in fermentations in the presence of 2-propargylmalonyl-SNAC *via* LC–ESI-MS revealed that variant V295L gives only wild-type product and variant V295G is non-functional. V295A variant, however, was found to incorporate the artificial building block to yield 2-propargylerythromycin A as revealed by HRMS and MS/MS-analysis (Figure 5 and Supplementary Table 6, section II). The new fermentation product is a mixture of the wild-type erythromycin and its propargylated derivative and is not yet produced in synthetically useful quantities, but this experiment clearly points out that a single well targeted point mutation in a PKS multienzyme complex of ~1,800 kDa can trigger the incorporation of an entirely non-natural building block, indicating that an enzyme-guided mutasynthesis can indeed be feasible.

Mutagenesis of PKS domains for this purpose, however, has to rely not only on sequence comparison for the identification of mutagenesis sites but should take the experimental and theoretical analysis of enzyme–substrate interaction into account. Importantly, no accumulation of biosynthetic intermediates from post-PKS processing was found. This indicates a sufficient substrate flexibility of the glycosyltransferases and monooxygenases that elaborate the primary DEBS product polyketide 6-deoxyerthronolide D toward erythromycin A. With this new strategy at hand, the portfolio of functional groups in polyketides can be significantly expanded, providing opportunities for the targeted semisynthesis of further derivatives.

Conclusion. Engineering of Polyketide Synthase has yielded hundreds of new polyketide derivatives in recent years. However, the lack of transferable strategies for the predictable engineering of these giant enzymes provides formidable challenges to modern enzymology. Especially, the swapping of domains or modules between individual PKS has proven to be conceptually and experimentally highly demanding and often unpredictable.

Recently, the systematic use of point mutations instead of domain swappings to diversify a complex polyketide has been shown to provide access to a library of new bioactive compounds with a high degree of predictability.[44] Of particular importance in PKS engineering are ways to alter the building block specificity in the polyketide assembly line. By reporting the first systematic exploration of sequence–function correlations in the context of a whole Polyketide Synthase, we show that saturation mutagenesis can be used for analyzing and manipulating catalytic PKS domains in the framework of a medium-throughput fermentation screening system. Analysis of the screening results depends on careful comparison with reference cultures and enables the generation of clear data on activity changes in these large enzymes. Importantly, it was shown that the minimally invasive approach to PKS mutagenesis *via* site-directed mutagenesis does not typically lead to deleterious effects on the overall protein, thereby enabling not only a deeper understanding of PKS enzymology but also more targeted engineering of the enzymatic machinery.

The experimental data, in full agreement with theoretical findings on a model of methylmalonyl-CoA placed in an acyltransferase domain, support the previous hypothesis on the catalytic mechanism in PKS AT domains and provide a clear picture of the molecular interactions between substrate and enzyme that lead to product formation in these complex enzymes. The model was used to predict the effect of a point mutation in the DEBS AT6 domain toward the incorporation of a non-natural building block into the hallmark antibiotic erythromycin.

Thus, we can conclude that in addition to reporting here the sequence–function- correlation of a Polyketide Synthase Acyltransferase domain, theoretical investigations based on a large experimental data set enabled the design of Polyketide Synthase point mutations that led to the incorporation of an entirely non-natural building block into a reduced polyketide. This extents the current state of the art, in which typically naturally occurring extender units are transferred from one biosynthetic pathway to another through either domain swapping or site-directed mutagenesis and will potentially pave the way toward an extensive elaboration of polyketides through a combination of organic synthesis and engineered fermentation.

METHODS

PKS Mutagenesis. For the purpose of mutagenesis, a φC31-bacteriophage-based integrative vector[45] for the expression of mutated DEBS3 in the native erythromycin producer *S. erythraea* was employed. *EryAIII*-AT6 mutants were constructed *via* overlap-extension PCR and introduced into the vector using sequence-and-ligation-independent cloning (SLIC-MIX).[43] Identity of all plasmids was confirmed by DNA sequencing. Detailed protocols can be found in the Supporting Information (pages 4–14).

Fermentation Experiments and Analytics. Plasmids encoding for mutated DEBS3-AT6 were introduced *via* conjugative plasmid transfer from *E. coli* ET12567/pUZ8002 into *S. erythraea* ΔAT6hygR carrying a nonfunctional copy of *eryAIII*. Four independent clones from each resulting mutant were cultivated in 24-well plates employing the Duetz fermentation system,[46] using *S. erythraea* NRRL-B-24071 and *S. erythraea* ΔAT6hygR carrying a nonmutated DEBS3-expression construct as controls in parallel fermentations. For feeding studies cultures were supplemented with 10 mM freshly synthesized 2-propargylmalonyl-SNAC. Fermentation results were analyzed by LC–ESI-MS from ethyl acetate extracts of the fermentation broth. Precise protocols and example chromatograms can be found in the Supporting Information (pages 15–22).

Computer Simulations. The homology modeling of the acyltransferase domain of the sixth module of 6-deoxyerythronolide synthase (DEBS-AT6) was performed using the I-Tasser server.[47] The resulting DEBS-AT6 model was used as starting point for docking the (2*S*)-methylmalonyl-CoA (MMCoA) substrate into AT6 using the

dx.doi.org/10.1021/cb300505w | *ACS Chem. Biol.* 2013, 8, 443–450

Schrödinger Suite.[48] The best resulting model was selected on the basis of the predicted free binding energy and the conservation of key interactions that support the specificity of the enzyme.[27] Next, the models for the wild type protein and selected mutants were solvated with a box of water of 80 × 105 × 85 Å dimensions, and the system's charge was neutralized using the VMD 1.9 software.[49] After that, molecular dynamics simulations (MD) were performed with the NAMD 2.7 program using the CHARMM22 force field and the TIP3P water model.[50−52] Quantum Mechanics/Molecular Mechanics (QM/MM) optimizations of three snapshots from the MD simulations were performed with the program ChemShell v3.4[53] at the BP86-D2/SVP//CHARMM level of theory.[54−56] More computational details are reported in the Supporting Information (pages 23 and 24).

■ ASSOCIATED CONTENT

Supporting Information

Detailed methods, analytical data, and further modeling details. This material is available free of charge *via* the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: esanchez@mpi-muelheim.mpg.de; frank3.schulz@tu-dortmund.de.

Notes

The authors declare no competing financial interests.

■ ACKNOWLEDGMENTS

F.S. acknowledges generous financial support from the Beilstein Institut zur Förderung der chemischen Wissenschaften. E.S.-G. and F.S. both acknowledge Liebig stipends, and U.S. and K.B.-R. both acknowledge predoctoral stipends from the Fonds der Chemischen Industrie. U.S. is a fellow of the IMPRS of Chemical Biology. E.S.-G acknowledges financial support by the Cluster of Excellence RESOLV (EXC 1069) funded by the Deutsche Forschungsgemeinschaft. P. Janning and C. Sevenich are thanked for help with MS analysis. Roquette GmbH is thanked for the donation of Glucidex IT29. E.S.-G. thanks W. Thiel for his support and useful suggestions. F.S. thanks H. Waldmann for unconditional support and encouragement.

■ REFERENCES

(1) Walsh, C. T. (2004) Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science 303*, 1805−1810.
(2) Hertweck, C. (2009) The biosynthetic logic of polyketide diversity. *Angew. Chem., Int. Ed. 48*, 4688−4716.
(3) Fischbach, M. A., and Walsh, C. T. (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev. 106*, 3468−3496.
(4) Hopwood, D. A., and Sherman, D. H. (1990) Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu. Rev. Genet. 24*, 37−66.
(5) Van Lanen, S. G., and Shen, B. (2008) Advances in polyketide synthase structure and function. *Curr. Opin. Drug Discovery Dev. 11*, 186−195.
(6) Weissman, K. J., and Leadlay, P. F. (2005) Combinatorial biosynthesis of reduced polyketides. *Nat. Rev. Microbiol. 3*, 925−936.
(7) Staunton, J., and Weissman, K. J. (2001) Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep. 18*, 380−416.
(8) Sundermann, U., Kushnir, S., and Schulz, F. (2011) Naturstoff−Lego. *Nachr. Chem. 59*, 29−35.
(9) Kwan, D. H., Sun, Y., Schulz, F., Hong, H., Popovic, B., Sim-Stark, J. C. C., Haydock, S. F., and Leadlay, P. F. (2008) Prediction and manipulation of the stereochemistry of enoylreduction in modular polyketide synthases. *Chem. Biol. 15*, 1231−1240.
(10) Khosla, C., Gokhale, R. S., Jacobsen, J. R., and Cane, D. E. (1999) Tolerance and specificity of polyketide synthases. *Annu. Rev. Biochem. 68*, 219−253.
(11) Wilson, M. C., and Moore, B. S. (2012) Beyond ethylmalonyl-CoA: the functional role of crotonyl-CoA carboxylase/reductase homologs in expanding polyketide diversity. *Nat. Prod. Rep. 29*, 72−86.
(12) Mo, S., Kim, D. H., Lee, J. H., Park, J. W., Basnet, D. B., Ban, Y. H., Yoo, Y. J., Chen, S.-w., Park, S. R., Choi, E. A., Kim, E., Jin, Y.-Y., Lee, S.-K., Park, J. Y., Liu, Y., Lee, M. O., Lee, K. S., Kim, S. J., Kim, D., Park, B. C., Lee, S.-g., Kwon, H. J., Suh, J.-W., Moore, B. S., Lim, S.-K., and Yoon, Y. J. (2010) Biosynthesis of the allylmalonyl-CoA extender unit for the FK506 polyketide synthase proceeds through a dedicated polyketide synthase and facilitates the mutasynthesis of analogues. *J. Am. Chem. Soc. 133*, 976−985.
(13) Quade, N., Huo, L., Rachid, S., Heinz, D. W., and Müller, R. (2012) Unusual carbon fixation gives rise to diverse polyketide extender units. *Nat. Chem. Biol. 8*, 117−124.
(14) Koryakina, I., and Williams, G. J. (2011) Mutant malonyl-CoA synthetases with altered specificity for polyketide synthase extender unit generation. *ChemBioChem 12*, 2289−2293.
(15) Hughes, Amanda J., and Keatinge-Clay, A. (2011) Enzymatic extender unit generation for in vitro polyketide synthase reactions: structural and functional showcasing of Streptomyces coelicolor MatB. *ChemBiol 18*, 165−176.
(16) Stassi, D. L., Kakavas, S. J., Reynolds, K. A., Gunawardana, G., Swanson, S., Zeidner, D., Jackson, M., Liu, H., Buko, A., and Katz, L. (1998) Ethyl-substituted erythromycin derivatives produced by directed metabolic engineering. *Proc. Natl. Acad. Sci. U.S.A. 95*, 7305−7309.
(17) Lau, J., Fu, H., Cane, D. E., and Khosla, C. (1999) Dissecting the role of acyltransferase domains of modular polyketide synthases in the choice and stereochemical fate of extender units. *Biochemistry 38*, 1643−1651.
(18) Oliynyk, M., Brown, M. J. B., Cortés, J., Staunton, J., and Leadlay, P. F. (1996) A hybrid modular polyketide synthase obtained by domain swapping. *ChemBiol 3*, 833−839.
(19) Pickens, L. B., Tang, Y., and Chooi, Y.-H. (2011) Metabolic engineering for the production of natural products. *Annu. Rev. Chem. Biomol. Eng. 2*, 211−236.
(20) Ridley, C. P., Lee, H. Y., and Khosla, C. (2008) Evolution of polyketide synthases in bacteria. *Proc. Natl. Acad. Sci. U.S.A. 105*, 4595−4600.
(21) Hans, M., Hornung, A., Dziarnowski, A., Cane, D. E., and Khosla, C. (2003) Mechanistic analysis of acyl transferase domain exchange in polyketide synthase modules. *J. Am. Chem. Soc. 125*, 5366−5374.
(22) Reeves, C. D., Murli, S., Ashley, G. W., Piagentini, M., Hutchinson, C. R., and McDaniel, R. (2001) Alteration of the substrate specificity of a modular polyketide synthase acyltransferase domain through site-specific mutations. *Biochemistry 40*, 15464−15470.
(23) Condurso, H. L., and Bruner, S. D. (2012) Structure guided approaches toward exploiting and manipulating nonribosomal peptide and polyketide biosynthetic pathways. *Curr. Opin. Chem. Biol. 16*, 162−169.
(24) Kirschning, A., and Hahn, F. (2012) Merging chemical synthesis and biosynthesis: a new chapter in the total synthesis of natural products and natural product libraries. *Angew. Chem., Int. Ed. 51*, 4012−4022.
(25) Haydock, S. F., Aparicio, J. F., Molnár, I., Schwecke, T., Khaw, L. E., König, A., Marsden, A. F. A., Galloway, I. S., Staunton, J., and Leadlay, P. F. (1995) Divergent sequence motifs correlated with the substrate specificity of (methyl)malonyl-CoA:acyl carrier protein transacylase domains in modular polyketide synthases. *FEBS Lett. 374*, 246−248.
(26) Del Vecchio, F., Petkovic, H., Kendrew, S. G., Low, L., Wilkinson, B., Lill, R., Cortes, J., Rudd, B. A., Staunton, J., and Leadlay, P. F. (2003) Active-site residue, domain and module swaps in modular polyketide synthases. *J. Ind. Microbiol. Biotechnol. 30*, 489−494.

(27) Tang, Y., Chen, A. Y., Kim, C.-Y., Cane, D. E., and Khosla, C. (2007) Structural and mechanistic analysis of protein interactions in module 3 of the 6-deoxyerythronolide B synthase. *Chem. Biol. 14*, 931−943.
(28) Petkovic, H., Sandmann, A., Challis, I. R., Hecht, H. J., Silakowski, B., Low, L., Beeston, N., Kuscer, E., Garcia-Bernardo, J., Leadlay, P. F., Kendrew, S. G., Wilkinson, B., and Müller, R. (2008) Substrate specificity of the acyl transferase domains of EpoC from the epothilone polyketide synthase. *Org. Biomol. Chem. 6*, 500−506.
(29) Tang, Y., Kim, C.-Y., Mathews, I. I., Cane, D. E., and Khosla, C. (2006) 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. 103*, 11124−11129.
(30) Gulder, T. A. M., and Moore, B. S. (2010) Salinosporamide natural products: potent 20S proteasome inhibitors as promising cancer chemotherapeutics. *Angew. Chem., Int. Ed. 49*, 9346−9367.
(31) Liu, Y., Hazzard, C., Eustáquio, A. S., Reynolds, K. A., and Moore, B. S. (2009) Biosynthesis of salinosporamides from α,β-unsaturated fatty acids: implications for extending polyketide synthase diversity. *J. Am. Chem. Soc. 131*, 10376−10377.
(32) Chan, Y. A., Podevels, A. M., Kevany, B. M., and Thomas, M. G. (2009) Biosynthesis of polyketide synthase extender units. *Nat. Prod. Rep. 26*, 90−114.
(33) Bergeret, F., Gavalda, S., Chalut, C., Malaga, W., Quémard, A., Pedelacq, J.-D., Daffé, M., Guilhot, C., Mourey, L., and Bon, C. (2012) Biochemical and structural study of the atypical acyltransferase domain from the mycobacterial polyketide synthase Pks13. *J. Biol. Chem. 287*, 33675−33690.
(34) Laureti, L., Song, L., Huang, S., Corre, C., Leblond, P., Challis, G. L., and Aigle, B. (2011) Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in Streptomyces ambofaciens. *Proc. Natl. Acad. Sci. U.S.A. 108*, 6258−6263.
(35) Cortes, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J., and Leadlay, P. F. (1990) An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of Saccharopolyspora erythraea. *Nature 348*, 176−178.
(36) Donadio, S., Staver, M., McAlpine, J., Swanson, S., and Katz, L. (1991) Modular organization of genes required for complex polyketide biosynthesis. *Science 252*, 675−679.
(37) Khosla, C., Tang, Y., Chen, A. Y., Schnarr, N. A., and Cane, D. E. (2007) Structure and mechanism of the 6-deoxyerythronolide B synthase. *Annu. Rev. Biochem. 76*, 195−221.
(38) Romero, P. A., and Arnold, F. H. (2009) Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol. 10*, 866−876.
(39) Reetz, M. T., and Carballeira, J. D. (2007) Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat. Protoc. 2*, 891−903.
(40) Kille, S., Zilly, F. E., Acevedo, J. P., and Reetz, M. T. (2011) Regio and stereoselectivity of P450-catalysed hydroxylation of steroids controlled by laboratory evolution. *Nat. Chem. 3*, 738−743.
(41) Kao, C. M., Katz, L., and Khosla, C. (1994) Engineered biosynthesis of a complete macrolactone in a heterologous host. *Science 265*, 509−512.
(42) Kwan, D. H., Tosin, M., Schlager, N., Schulz, F., and Leadlay, P. F. (2011) Insights into the stereospecificity of ketoreduction in a modular polyketide synthase. *Org. Biomol. Chem. 9*, 2053−2056.
(43) Kushnir, S., Sundermann, U., and Schulz, F. (2012) A technique for modifying biosynthetic pathways for natural products in bacteria. *Asian J. Biotechnol.* Submitted for publication.
(44) Kushnir, S., Sundermann, U., Yahiaoui, S., Brockmeyer, A., Janning, P., and Schulz, F. (2012) Minimally invasive mutagenesis gives rise to a biosynthetic polyketide library. *Angew. Chem., Int. Ed. 51*, 10664−10669.
(45) Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000) *Practical Streptomyces Genetics*, John Innes Foundation, U.K.
(46) Duetz, W. A., Ruedi, L., Hermann, R., O'Connor, K., Buchs, J., and Witholt, B. (2000) Methods for intense aeration, growth, storage, and replication of bacterial strains in microtiter plates. *Appl. Environ. Microbiol. 66*, 2641−2646.
(47) Roy, A., Kucukural, A., and Zhang, Y. (2010) I−TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc. 5*, 725−738.
(48) *Schrödinger Suite 2009 Protein Preparation Wizard*; Epik version 2.0, Schrödinger, LLC, New York, NY, 2009; Impact version 5.5, Schrödinger, LLC, New York, NY, 2009; Prime version 2.1, Schrödinger, LLC, New York, NY, 2009.
(49) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics 14*, 33−38.
(50) Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kalé, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD. *J. Comput. Chem. 26*, 1781−1802.
(51) Mackerell, A. D., Feig, M., and Brooks, C. L. (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas−phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem. 25*, 1400−1415.
(52) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys. 79*, 926−935.
(53) Chemshell, a Computational Chemistry Shell (www.chemshell.org).
(54) Becke, A. D. (1992) Density functional thermochemistry. II. The effect of the Perdew-Wang generalized-gradient correlation correction. *J. Chem. Phys. 97*, 9173−9177.
(55) Grimme, S. (2006) Semiempirical GGA-type density functional constructed with a long−range dispersion correction. *J. Comput. Chem. 27*, 1787−1799.
(56) Ahlrichs, R., Bär, M., Häser, M., Horn, H., and Kölmel, C. (1989) Electronic structure calculations on workstation computers: the program system TURBOMOLE. *Chem. Phys. Lett. 162*, 165−169.