The Loading and Initial Elongation Modules of Rifamycin Synthetase Collaborate To Produce Mixed Aryl Ketide Products†

Suzanne J. Admiraal,‡ Chaitan Khosla,§ and Christopher T. Walsh*,‡

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, and Departments of Chemical Engineering, Chemistry, and Biochemistry, Stanford University, Stanford, California 94305-5025

Received January 11, 2002; Revised Manuscript Received March 4, 2002

ABSTRACT: Rifamycin synthetase assembles the chemical backbone that members of the rifamycin family of antibiotics have in common. The synthetase contains a mixed biosynthetic interface between its loading module, which uses a nonribosomal peptide synthetase mechanism, and its initial elongation module, which uses a polyketide synthase mechanism. Biochemical studies of the loading and initial elongation modules of rifamycin synthetase reveal that this bimodular protein (LM-M1) catalyzes the formation of the phenyl ketide 3-hydroxy-2-methyl-3-phenylpropionate via a series of reactions that require benzoate, Mg·ATP, methylmalonyl-CoA, and NADPH. The overall rate of phenyl ketide production appears to be determined by the covalent loading of benzoate onto LM-M1, rather than by subsequent steps such as intermodular transfer of benzoate or condensation of benzoate and methylmalonate. Substituted benzoates that have previously been shown to be substrates for the loading module alone can also be incorporated into the corresponding aryl ketides by LM-M1, suggesting that the bimodular protein has a broad substrate tolerance. Discrimination between the substituted benzoates appears to reside in the benzoate loading reaction, and preincubation of LM-M1 with substituted benzoates and Mg·ATP allows faster downstream reactions to be unmasked. LM-M1 may be a useful biochemical system for exploring interactions between nonribosomal peptide synthetase and polyketide synthase modules.

The rifamycin synthetase of *Amycolatopsis mediterranei* directs the formation of proansamycin X, a precursor of the rifamycin class of antibiotics (*4*−*7*). These antibiotics are widely used to combat tuberculosis, and a recent crystal structure revealed that rifampicin, a member of the rifamycin family, exerts its cytotoxic effect by binding deep within the main DNA/RNA channel of the β subunit of bacterial RNA polymerase and blocking the path of the elongating RNA transcript (*8*). The synthetase is a mixed nonribosomal peptide synthetase (NRPS)[1]−polyketide synthase (PKS), consisting of an NRPS-like loading module and 10 PKS elongation modules (Figure 1). The loading module has been shown to activate 3-amino-5-hydroxybenzoate (AHB) as AHB-AMP and link it to a thiolation domain, a mechanism that is analogous to the activation and covalent attachment of amino acids in canonical NRPS modules (*9*). The benzoate is then transferred across the mixed NRPS−PKS interface to PKS module 1, where it reacts with a methylmalonyl moiety to form an enzyme-linked aryl ketide[2] product. Downstream PKS modules extend the ketide backbone of the molecule, which is finally cyclized to yield proansamycin X (Figure 1).

Other therapeutically important mixed products of NRPS−PKS systems include the antitumor agents bleomycin (*10*) and epothilone (*11*) and the immunosuppressants rapamycin (*12*) and FK506 (*13*). An understanding of the mechanisms of the biosynthetic machineries that produce mixed molecules may enable these machineries to be manipulated, advancing drug discovery. We have chosen the loading and initial elongation modules of rifamycin synthetase (LM-M1) as a defined biochemical system in which to investigate properties of an NRPS−PKS interface.

A wealth of information from previous studies of NRPS and PKS biosynthetic systems suggests mechanisms of

† This work was supported by National Institutes of Health Grants GM20011 (C.T.W.) and AI38947 (C.K.). S.J.A. is a Fellow of the Damon Runyon Cancer Research Foundation (DRG-1573).

* To whom correspondence should be addressed. Phone: 617-432-1715. FAX: 617-432-0438. E-mail: christopher_walsh@hms.harvard.edu.

‡ Harvard Medical School.

§ Stanford University.

[1] Abbreviations: NRPS, nonribosomal peptide synthetase; PKS, polyketide synthase; AHB, 3-amino-5-hydroxybenzoate; LM, NRPS-like loading module of rifamycin synthetase; M1, initial PKS elongation module of rifamycin synthetase; A, adenylation domain; T, thiolation domain; KS, ketosynthase domain; mmCoA, methylmalonyl-CoA; AT, acyltransferase domain; KR, ketoreductase domain; DH, dehydratase domain; phenyl ketide or PK, 3-hydroxy-2-methyl-3-phenylpropionate; Rif, rifamycin; DEBS, 6-deoxyerythronolide B synthase; TE, thioesterase domain; IPTG, isopropyl-β-D-thiogalactopyranoside; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; B, benzoate; 3-HB, 3-hydroxybenzoate; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; TLC, thin-layer chromatography; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; TCA, trichloroacetic acid; HPLC, high-performance liquid chromatography; MS, mass spectrometry; TCEP, tris(2-carboxyethyl)phosphine hydrochloride; E-B, enzyme-linked benzoate; E-3-HB, enzyme-linked 3-hydroxybenzoate; XB, substituted benzoate; E-XB, enzyme-linked substituted benzoate; k_{rel}, rate constant relative to analogous reaction with benzoate; E-PK, enzyme-linked phenyl ketide; V_{rel}, velocity of aryl ketide formation relative to analogous velocity of phenyl ketide formation.

[2] We refer to the product of the first two modules (LM-M1) of rifamycin synthetase as an aryl ketide rather than a diketide because the electrophile in the C−C bond formation reaction is an aryl moiety. For the specific case in which the electrophile is an unsubstituted benzoyl moiety, we refer to the product as a phenyl ketide.

10.1021/bi0200312 CCC: $22.00
Published on Web 03/29/2002

substrate loading (Figure 2A) and product formation (Figure 2B) for LM-M1 by analogy. For covalent loading of LM (Figure 2A, left), AHB is expected to be activated as AHB−AMP by the A domain, and the thiol of the phosphopantetheine cofactor of the T_1 domain would attack AHB−AMP to form a covalent arylthioester enzyme intermediate. Transfer of the AHB moiety from the T_1 domain of LM to the active site cysteine of the KS domain of M1 may occur before or after covalent loading of M1 with methylmalonate. For covalent loading of M1 (Figure 2A, right), the methylmalonyl moiety of mmCoA is expected to be transferred to the active site serine of the AT domain, and the thiol of the phosphopantetheine cofactor of the T_2 domain would subsequently attack the AT intermediate to form a second covalent acylthioester enzyme intermediate. Formation of an aryl ketide product by LM-M1 would presumably occur by KS-catalyzed decarboxylative condensation of enzyme-linked AHB and enzyme-linked methylmalonate (Figure 2B). The KR domain of M1 is expected to reduce the aryl ketide while it is covalently attached to the T_2 domain, and the soluble aryl ketide product would result from hydrolysis of this covalent intermediate.

Using the mechanistic model described above and shown in Figure 2 as a guide, we investigated the ability of LM-M1 to catalyze formation of a mixed product, analyzed its ability to use alternative benzoates as substrates, and performed preliminary kinetic analyses.

EXPERIMENTAL PROCEDURES

Materials. [7-^{14}C]-Benzoic acid (57 mCi/mmol) and [7-^{14}C]-3-hydroxybenzoic acid (55 mCi/mmol) were obtained from American Radiolabeled Chemicals. All other substituted benzoic acids, including benzoic-d_5 acid, were obtained from Aldrich. DL-2-[*methyl*-^{14}C]-Methylmalonyl-CoA (60 mCi/mmol) was from New England Nuclear. ATP was supplied by Sigma Chemical Co. Authentic 3-hydroxy-2-methyl-3-phenylpropionic acid (phenyl ketide) was obtained from Maybridge. AHB was synthesized according to a previously published protocol (*14*). Restriction enzymes were from New England Biolabs.

Manipulation of DNA and Strains. DNA manipulations were performed in *E. coli* XL1 Blue (Stratagene) using standard culture conditions (*15*). Polymerase chain reactions were carried out using Pfu polymerase (Stratagene) as recommended by the manufacturer.

Construction of an Expression Vector for LM-M1. Introduction of an *Nde*I restriction site at the start codon of the *rifA* gene, corresponding to the N-terminal end of the LM, has been described previously (*9*). The natural sequence 5′-CGCGCC-3′ at nt 10402−10407 (GenBank accession no. AF040570), corresponding to the C-terminal end of M1, was chosen on the basis of an alignment of DEBS and Rif ACP sequences (*16*) for replacement with the *Spe*I recognition sequence 5′-ACTAGT-3′. The *Nde*I−*Spe*I fragment encoding Rif LM-M1 was then fused to the *Spe*I−*Eco*RI fragment encoding the DEBS TE via replacement of the *Nde*I−*Spe*I fragment encoding DEBS M3 in pST132 (*17*). The *Nde*I−*Eco*RI fragment encoding Rif LM-M1-DEBS TE was cloned into pET21c (Novagen) to produce pSA19, an expression vector for Rif LM-M1-DEBS TE with hexahistidine appended to the C-terminus.

Expression and Purification of LM-M1. Plasmid pSA19 was introduced via transformation into *E. coli* BL21 (Stratagene) for expression of *apo* LM-M1. One liter cultures of BL21/pSA19 were grown at 37 °C in 2 L flasks containing LB medium supplemented with 100 μg/mL carbenicillin. Expression of LM-M1 was induced with 100 μM IPTG at an optical density at 600 nm of 0.7. After induction, incubation was continued for 24 h at 20 °C. The cells were then harvested by centrifugation at 2500*g* and resuspended in disruption buffer [200 mM sodium phosphate (pH 7.2), 200 mM sodium chloride, 2.5 mM DTT, 2.5 mM EDTA, 1.5 mM benzamidine, pepstatin (2 mg/L), leupeptin (2 mg/L), and 30% v/v glycerol].

All purification procedures were performed at 4 °C. The resuspended cells were disrupted by two passages through a French press at 13 000 psi, and the lysate was collected by centrifugation at 40000*g*. Nucleic acids were precipitated with polyethylenimine (0.15%) and removed via centrifugation. The supernatant was made 40% (w/v) saturated with ammonium sulfate and precipitated overnight. After centrifugation, the pellet containing protein was redissolved in 50 mM Tris-HCl (pH 8), 300 mM sodium chloride, 10 mM imidazole, and 10% v/v glycerol. This solution was loaded onto a previously equilibrated nickel-nitrilotriacetic acid (Ni-NTA) column (10 mL, Qiagen). The column was washed with 10 mM imidazole in 50 mM Tris-HCl (pH 8), 300 mM sodium chloride, and 10% v/v glycerol, and LM-M1 was eluted with 100 mM imidazole in the same solution. Pooled fractions containing LM-M1 were loaded onto a previously equilibrated HiTrap Q anion-exchange column (5 mL, Amersham Pharmacia Biotech). LM-M1 was eluted by a NaCl gradient (150−500 mM) in 100 mM sodium phosphate (pH 7.2), 2.5 mM DTT, 2 mM EDTA, and 20% v/v glycerol. Pooled fractions containing LM-M1 were buffer-exchanged into 100 mM sodium phosphate (pH 7.2), 2.5 mM DTT, 2 mM EDTA, and 20% v/v glycerol by gel filtration (PD-10, Amersham Pharmacia Biotech) and concentrated with a Centriprep-50 concentrator (Amicon). The purified protein was flash-frozen in liquid nitrogen and stored at −80 °C. Protein concentration was determined using the following calculated extinction coefficient at 280 nm: 258 460 M^{-1} cm^{-1} (*18*). A typical 1 L culture produced about 5 mg of purified protein.

For expression of *holo* LM-M1, plasmid pSA19 was transformed into BL21 containing the plasmid pRSG56 (*16*), which carries a kanamycin resistance gene and the *sfp* gene. The *sfp* gene expresses Sfp, a nonspecific phosphopantetheinyl transferase from *B. subtilis* that converts the *apo* protein into the *holo* protein (*19*, *20*). One liter cultures of this recombinant *E. coli* strain were grown at 37 °C in 2 L flasks containing LB medium supplemented with 100 μg/mL carbenicillin and 50 μg/mL kanamycin. The expression and purification steps for *holo* LM-M1 were performed as described above for *apo* LM-M1.

Covalent Loading of LM-M1 with Substrates. For qualitatively assessing the incorporation of B, 3-HB, or mmCoA into LM-M1, reactions contained 0.5 μM *apo* or *holo* LM-M1, 50 mM sodium phosphate (pH 7.2), 1 mM DTT, 1 mM EDTA, 15 mM $MgCl_2$, 10% glycerol, and 50 μM [7-^{14}C]-B, [7-^{14}C]-3-HB, or DL-2-[*methyl*-^{14}C]-mmCoA. In reactions where ATP was included, 5 mM was present. After incubation at 30 °C for 30 min, reactions were quenched with

SDS−PAGE sample buffer and electrophoresed on a 4−15% polyacrylamide gel (Bio-Rad). The gel was briefly stained with Coomassie blue, destained, dried, and autoradiographed.

Detection of the Putative Phenyl Ketide Product of LM-M1 by TLC. Complete reactions shown in Figure 5 contained 5 μM *apo* or *holo* LM-M1, 50 mM Tris (pH 7.4), 15 mM $MgCl_2$, 5 mM ATP, 5 mM NADPH, 50 μM B, and 50 μM mmCoA. Individual components of the complete reaction were omitted from individual reactions as indicated in the Figure 5 legend. Reactions were incubated at 30 °C for 120 min, and 20 μL aliquots were quenched in 180 μL of ice-cold 10% TCA. Precipitates were pelleted by centrifugation and washed with 190 μL of 10% TCA. After recentrifugation, the precipitates were washed with 190 μL of ice-cold 100% ethanol. The pellets were allowed to air-dry prior to addition of 20 μL of 200 mM KOH. Samples were vortexed and then incubated at 65 °C for 20 min. After brief centrifugation, hydrolyzed samples were spotted onto silica gel TLC plates (250 μm layer, Whatman). Plates were developed using 60:30:10 toluene/ether/acetic acid and subsequently exposed to phosphorimager screens overnight. Screens were scanned by a Fuji BAS1000 phosphorimager.

Confirming the Identity of Phenyl Ketide by HPLC-MS. Reactions contained 50 μM *holo* LM-M1, 50 mM Tris (pH 7.4), 15 mM $MgCl_2$, 5 mM ATP, 5 mM NADPH, 1 mM mmCoA, 10% glycerol, and 100 μM B or pentadeuterated B. Reactions were incubated at 30 °C, and 5 M KOH was added to a final concentration of 350 mM upon completion of the enzymatic reaction. The mixtures were incubated at 65 °C for 20 min to release covalently bound product from the enzyme. Formic acid was added to a final concentration of 2.5% to acidify each sample, and samples were centrifuged to pellet the resulting protein precipitate. The supernatant was injected onto an HPLC equipped with a C18 reverse-phase column (VYDAC, 250 × 4.6 mm) with the UV-detector monitoring at 254 nm. A linear gradient between buffer A (0.01% formic acid, 0.1% trifluoroacetic acid, 0.02% triethylamine) and buffer B (4:1 acetonitrile/buffer A) from 10% to 100% B was run over 25 min with a flow rate of 1 mL/min. Peaks that possessed the same 16.5 min retention time as the 3-hydroxy-2-methyl-3-phenylpropionate standard were collected, lyophilized, and subjected to MALDI-TOF MS analysis.

Kinetics of Arylation. Typical reactions contained 5 μM *holo* LM-M1, 100 mM Tris (pH 7.4), 15 mM $MgCl_2$, 5 mM ATP, 5 mM TCEP, 0.5−10 μCi/mL [7-^{14}C]-B or [7-^{14}C]-3-HB, and varying concentrations of unlabeled B or HB. Unlabeled B and 3-HB stocks were adjusted to the reaction pH prior to addition. Addition of 5 mM NADPH to reactions had no effect. Reactions were incubated at 30 °C, and at desired time points, 20 μL aliquots were quenched in 0.75 mL of ice-cold 5% TCA and 200 μg of bovine serum albumin (Sigma) was added to this mixture to aid precipitation of the protein. The precipitate was pelleted by centrifugation, washed with 0.5 mL of 5% TCA, and solubilized in 0.5 mL of a 100 mM phosphate (pH 8), 2% SDS solution. This solution was combined with 6 mL of liquid scintillation fluid (Ultima Gold, Packard), and the incorporated [^{14}C]-label, corresponding to E-B or E-3-HB covalent intermediate, was quantified by liquid scintillation counting. Reaction rates were linearly dependent on enzyme concentration. Data analysis was performed with Kaleidagraph (Synergy Software), and exponential fits to the data typically gave $R \geq 0.95$.

To measure the relative rate constants for arylation of LM-M1, reactions were performed as described above but in the presence of 20 μM−5 mM of a series of substituted benzoates. Substituted benzoate stocks were adjusted to the reaction pH prior to addition. The relative rate constant (k_{rel}) for reaction of a given substituted benzoate with respect to B was determined from the concentrations of B and substituted benzoate in the original reaction ([B], [XB]) and the amount of product present as E-B and E-XB, according to the equation in Scheme 1 (*21*). The amount of E-XB product in each reaction at a given time point was determined by subtracting the amount of radiolabeled E-B in the presence of the competing substituted benzoate from that obtained at the same time point in an identical reaction lacking competitor. The ratio of E-B to E-XB was constant, within error, throughout a given time course, indicating that no secondary reactions involving the reaction products were occurring. For each substituted benzoate, the same k_{rel} value, within error, was obtained for reactions performed at different substituted benzoate concentrations. The reactions were repeated for selected substituted benzoates using radiolabeled 3-HB instead of B, and the same k_{rel} values (with respect to B), within error, were obtained. Each k_{rel} value in Table 1 represents an average of at least four separate determinations.

Kinetics of Aryl Ketide Formation. Typical reactions to measure overall aryl ketide formation contained 5 μM *holo* LM-M1, 100 mM Tris (pH 7.4), 15 mM $MgCl_2$, 5 mM ATP, 5 mM NADPH, 5 mM TCEP, 0.5−10 μCi/mL [7-^{14}C]-B or [*methyl*-^{14}C]-mmCoA, and varying concentrations of unlabeled B, substituted benzoate, or mmCoA. Reactions were incubated at 30 °C, and at desired time points 4 μL aliquots were quenched in 1 μL of 0.4 M KOH and incubated at 65 °C for 20 min. The hydrolyzed samples were centrifuged briefly and then spotted onto TLC plates. Alternatively, at desired time points 20 μL aliquots were quenched in 180 μL of ice-cold 10% TCA, precipitates were pelleted by centrifugation and washed with 190 μL of 10% TCA, and after recentrifugation the precipitates were washed with 190 μL of ice-cold 100% ethanol. The pellets were allowed to air-dry prior to addition of 20 μL of 200 mM KOH. Samples were vortexed and incubated at 65 °C for 20 min, and after brief centrifugation, hydrolyzed samples were spotted onto TLC plates. Most separations between substrates and aryl ketide products were effected using the silica gel TLC protocol described above for phenyl ketide. PEI-cellulose TLC plates (Aldrich) developed in 4:1:1 butanol/acetic acid/water were used to resolve the aryl ketide products of the reactions containing AHB or 3,5-dihydroxybenzoate, and silica gel TLC plates (Whatman) developed in 60:30:10 chloroform/methanol/water were used to resolve the aryl ketide product of the reactions containing 3-aminobenzoate. A phosphorimager (Fuji BAS1000) was used to quantitate product formation, and data analysis was performed with Kaleidagraph (Synergy Software). Exponential fits to the data typically gave $R \geq 0.95$.

To measure the post-arylation formation of aryl ketide, reactions were preincubated at 30 °C in the absence of mmCoA. At desired time points following addition of mmCoA, 4 or 20 μL aliquots were quenched, hydrolyzed, and analyzed as described above. The observed rate constant

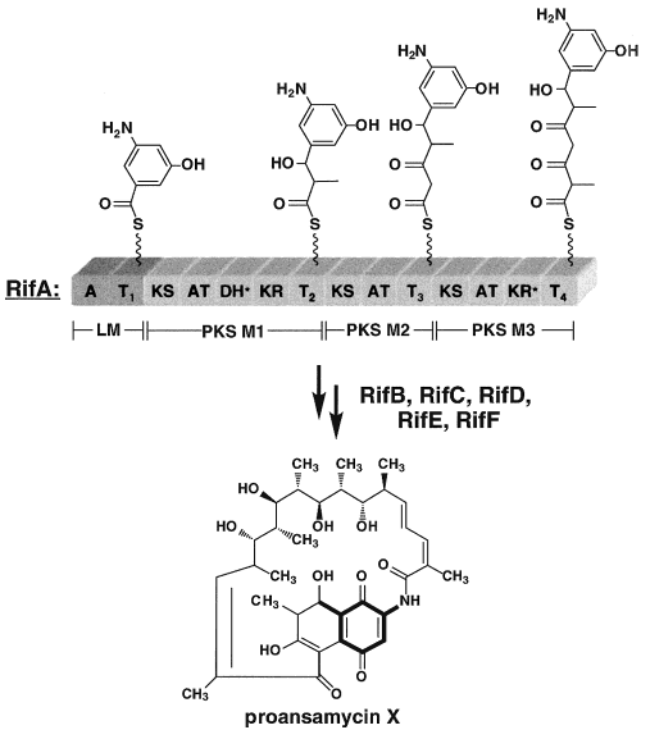


FIGURE 1: Chain elongation on the RifA subunit of rifamycin synthetase. The NRPS-like LM of RifA activates and loads AHB, and the three PKS modules extend the chain using mmCoA (M1, M3) or malonyl-CoA (M2) as substrates. The seven remaining PKS modules of rifamycin synthetase are distributed on the RifB, RifC, RifD, and RifE subunits, and RiF is believed to catalyze cyclization via intramolecular amide formation to yield proansamycin X, the precursor of the rifamycin class of antibiotics. The location of the mC_7N unit derived from AHB is shown in boldface in the proansamycin X structure. The active sites denote adenylation (A), thiolation (T), acyltransferase (AT), ketosynthase (KS), β-keto-reductase (KR), or dehydratase (DH) domains. Starred domains are believed to be inactive due to active site amino acid deletions (*4*−*6*).

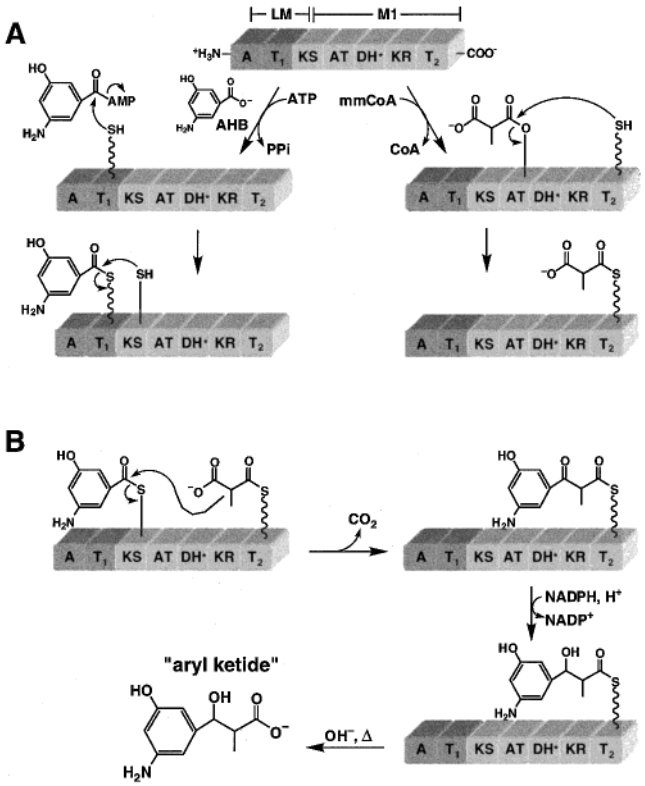


FIGURE 2: Putative mechanisms for substrate loading and product formation by the bimodular LM-M1 protein derived from rifamycin synthetase. The two domains of the NRPS-like LM are shaded more darkly than the five domains of PKS M1, and enzymatic thiol and hydroxyl nucleophiles are depicted schematically. The starred DH domain is inactive in rifamycin synthetase due to active site amino acid deletions (*4*−*6*). (A) Covalent loading of LM with AHB is shown on the left, and covalent loading of M1 with methylmalonate is shown on the right. (B) Formation of an aryl ketide product by LM-M1 would presumably occur by KS-catalyzed decarboxylative condensation of enzyme-linked AHB and enzyme-linked methylmalonate followed by KR-catalyzed reduction. Treatment with base at elevated temperatures should hydrolyze the thioester bond to yield soluble aryl ketide.

obtained for aryl ketide formation in reactions preincubated with a given benzoate was independent of the length of preincubation, the concentration of the benzoate species in the preincubation reaction, LM-M1 concentration, and final mmCoA concentration (>20 μM), although burst amplitudes varied as expected.

RESULTS

Construction and Purification of LM-M1. The loading and initial elongation modules of rifamycin synthetase are located at the N-terminus of the 500 kDa RifA polypeptide (Figure 1). To investigate the bimodular LM-M1 system biochemically, these two modules were removed from the RifA protein context. The sequence encoding the 260 kDa LM-M1 protein was subcloned into an expression vector, using an *Nde*I restriction site engineered at the transcriptional start site of RifA and a *Spe*I restriction site introduced at the C-terminal end of the consensus T domain of module 1 (T_2; Figure 2). Thiolation domains require covalent attachment of the 4′-phosphopantetheine moiety of CoA to a conserved serine for activity (*22*). The Sfp phosphopantetheinyl transferase from *B. subtilis*, which is capable of converting the *apo* forms of many heterologous recombinant proteins into the *holo* forms, was therefore coexpressed with the LM-M1 protein in the *holo* enzyme preparation (*19*, *20*). The *apo* and *holo*

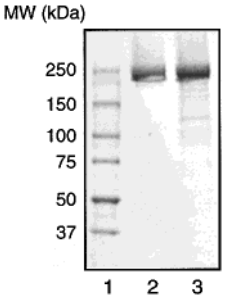


FIGURE 3: Purified recombinant *apo* and *holo* LM-M1 (260 kDa, encoded by plasmid pSA19) overproduced in *E. coli*. Protein samples were resolved by SDS−PAGE (4−15%, Bio-Rad) and Coomassie-stained. Lanes 1−3 correspond to the molecular mass markers, *holo* LM-M1, and *apo* LM-M1, respectively.

forms of LM-M1 were produced in *E. coli* as C-terminal hexahistidine-tagged fusion proteins and were purified by nickel affinity chromatography and anion exchange chromatography to >90% homogeneity (Figure 3).

Arylation and Acylation of LM-M1. The mechanism of the loading module of LM-M1 has been shown to involve activation of AHB as the aryl-AMP by the A domain, followed by formation of a covalent arylthioester enzyme

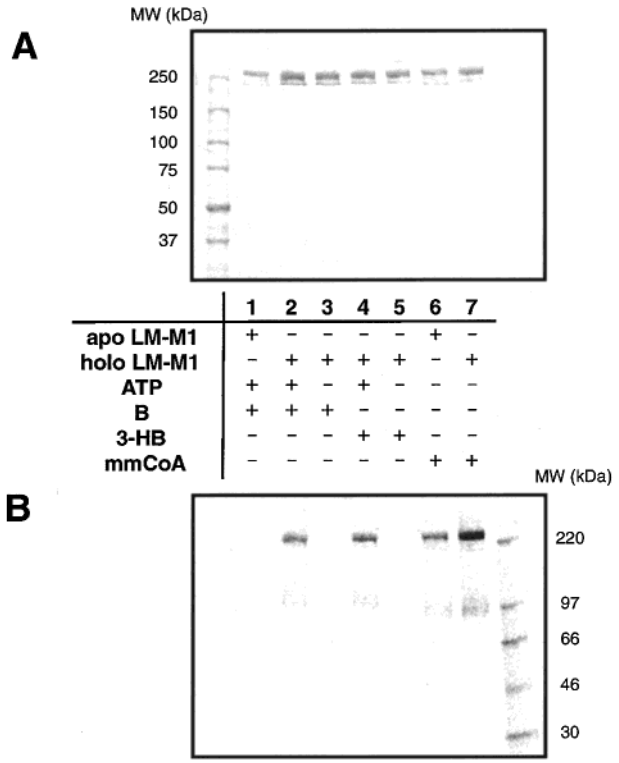


FIGURE 4: Covalent loading of LM-M1 with B, 3-HB, and mmCoA. (A) Coomassie-stained gel (4−15% gradient) of the reaction mixtures. (B) Autoradiograph of this gel, including a [^{14}C]-methylated protein ladder (Amersham) in the far right lane. The presence or absence of *apo* or *holo* LM-M1, ATP, and [^{14}C]-B, [^{14}C]-3-HB, or [^{14}C]-mmCoA is indicated in the chart.

intermediate from attack of the aryl-AMP by the thiol nucleophile of the phosphopantetheine cofactor of the T_1 domain (*9*; Figure 2A, left). To investigate the activity of the larger LM-M1 protein, we sought to covalently load it with benzoate substrate. Although AHB is not available in radiolabeled form, previous work had shown that the loading module can also be loaded by benzoate (B) or 3-hydroxybenzoate (3-HB) (*9*). Reactions containing [^{14}C]-B or [^{14}C]-3-HB and *apo* or *holo* LM-M1 were incubated in the presence or absence of Mg·ATP and subsequently analyzed by SDS−PAGE autoradiography (Figure 4). Lacking the phosphopantetheine cofactor of the T_1 domain, *apo* LM-M1 could not be covalently loaded with B (lane 1). However, *holo* LM-M1 is covalently loaded with either B or 3-HB in reactions that require Mg·ATP (lanes 2−5).

Based on previous studies of PKS modules, there are two expected sites of covalent acylation of M1 by methylmalonate, supplied to the protein as mmCoA. The active site serine of the AT domain is expected to attack mmCoA to yield acylated enzyme (*23*), and subsequent nucleophilic attack of the thioester bond of this intermediate by the phosphopantetheine thiol of the T_2 domain would produce a second acylated enzyme (Figure 2A, right). Reactions containing [^{14}C]-mmCoA and *apo* or *holo* LM-M1 were analyzed by SDS−PAGE autoradiography (Figure 4). *Apo* LM-M1 could be loaded by mmCoA despite the absence of the phosphopantetheine arm of the T_2 domain, presumably because acylation of the AT domain serine was unaffected (lane 6). However, the extent of acylation increased when the reaction contained *holo* LM-M1, consistent with acylation of both AT and T_2 domains of this species (lane 7).

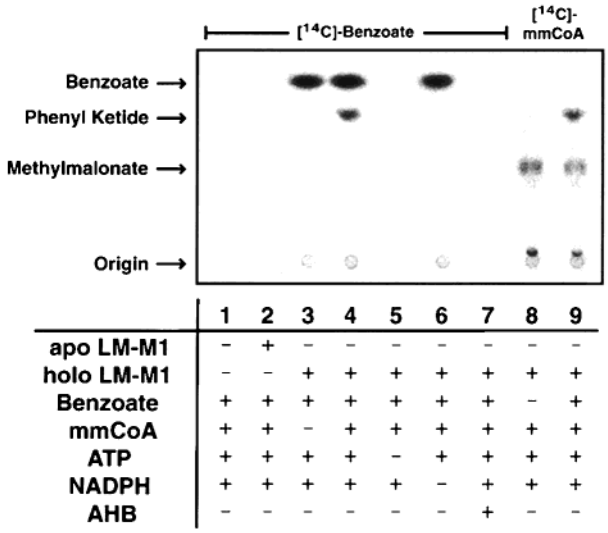


FIGURE 5: LM-M1 produces phenyl ketide in the presence of B, mmCoA, Mg·ATP, and NADPH. [^{14}C]-Phenyl ketide product was resolved from [^{14}C]-B or [^{14}C]-methylmalonate by TLC following TCA precipitation and hydrolysis of reactions. The presence or absence of *apo* or *holo* LM-M1, B, mmCoA, ATP, NADPH, or AHB in the original reactions is indicated in the chart. Lanes 1−7 show the hydrolysis products of reactions containing [^{14}C]-B in addition to the components indicated in the chart, whereas lanes 8 and 9 show the hydrolysis products of reactions that contained [^{14}C]-mmCoA in addition to the other components.

Formation of Phenyl Ketide Product by LM-M1. After confirming that LM-M1 was active for self-arylation and self-acylation using the experiments just described (Figure 4), the ability of the protein to *condense* B and methylmalonate to form the corresponding mixed phenyl ketide product was investigated. Following incubation of LM-M1 with different combinations of Mg·ATP, NADPH, B, and mmCoA, the enzyme was precipitated with TCA and subjected to hydrolysis to release any enzyme-linked products (Figure 5).[3] In the absence of enzyme, the reaction components are not precipitated by TCA, and no substrate or product is visible in the TLC lane (lane 1). When *apo* LM-M1 is incubated with all of the substrates using [^{14}C]-B as the reporter, the lane is also empty (lane 2), presumably because the loading module cannot be arylated with B in the absence of its phosphopantetheine arm. Only B is released from the enzyme upon incubation of *holo* LM-M1 with all substrates except mmCoA (lane 3). When all substrates including mmCoA are present in the reaction, a new band that comigrates with authentic phenyl ketide, 3-hydroxy-2-methyl-3-phenylpropionate (data not shown), is visible in the TLC lane (lane 4). Activation of B as the benzoyl-AMP cannot occur in the absence of ATP, so neither B nor phenyl ketide is observed in such reactions (lane 5). Phenyl ketide is absent when

[3] The LM-M1 protein described herein has the TE domain of DEBS appended to its C-terminus (see Experimental Procedures). This TE has previously been shown to catalyze hydrolysis of acylthioester bonds when placed in unnatural contexts (*1*−*3*). Therefore, we hoped that the TE would enable LM-M1 to undergo multiple turnovers by catalyzing hydrolysis of the acylthioester bond linking the aryl ketide product to the T_2 domain of LM-M1. Neither reduced nor unreduced aryl ketide accumulated appreciably in reactions that contained LM-M1 and all substrates, suggesting that the TE did not catalyze hydrolysis as had been hoped. Aryl ketide was, however, released when reactions were treated with hydroxide at elevated temperatures, so this method was used to analyze the products of single turnover reactions.

NADPH is omitted (lane 6), presumably because the reduced product does not form in its absence. Preliminary results suggest that the unreduced phenyl ketide product may form when NADPH is absent, but this putative product is not visible in lane 6 because it comigrates with B (data not shown). B and phenyl ketide are not released upon hydrolysis of reactions that contain the natural substrate AHB (lane 7), presumably because unlabeled AHB outcompetes radiolabeled B for arylation of the loading module. Radiolabeled mmCoA can be used instead of radiolabeled B to monitor reactions. Only methylmalonate is released from the enzyme upon incubation of *holo* LM-M1 with all substrates except B (lane 8), but a product with the same mobility as 3-hydroxy-2-methyl-3-phenylpropionate is visible when B is present (lane 9; compare to lane 4).

The results shown in Figure 5 are consistent with formation of phenyl ketide, as the product incorporates both B and methylmalonate, and *holo* LM-M1, Mg·ATP, and NADPH are required. The identity of this product was verified by mass spectrometry as follows. Reactions containing the required components *holo* LM-M1, Mg·ATP, NADPH, and mmCoA were incubated with either B or pentadeuterated B and subsequently hydrolyzed. The products of hydrolysis were resolved by HPLC, the peak that eluted with the same retention time as the 3-hydroxy-2-methyl-3-phenylpropionate standard was collected for each reaction, and these products were analyzed by MALDI-TOF MS. For the B reaction, masses of 203 and 219, corresponding to the expected masses of the [M + Na] and [M + K] ions of phenyl ketide, were observed (data not shown). The product of the pentadeuterated B reaction gave the expected shift of 5 mass units, for masses of 208 and 224 (data not shown). Taken together, the results in Figure 5 and the mass spectrometry results provide strong evidence for the formation of enzyme-linked phenyl ketide by LM-M1. Preliminary analysis of the kinetics of the process is described below.

Substrate Tolerance of Arylation of LM-M1. B and 3-HB are substrates for *holo* LM-M1, as shown qualitatively in Figure 4. To quantitatively assess these benzoates as substrates for aryl-AMP formation followed by arylation of the thiol of the phosphopantetheine cofactor of the T_1 domain, we utilized a protein precipitation assay. Aliquots from reactions containing *holo* LM-M1, 0.5−5 μCi/mL [7-^{14}C]-B or [7-^{14}C]-3-HB, and varying concentrations of unlabeled B or 3-HB were quenched with TCA, and the amount of radiolabeled protein in each washed protein pellet was determined by liquid scintillation counting. Initial velocities of E-B or E-3-HB formation as a function of B or 3-HB concentrations were obtained using this method and used to generate the curves shown in Figure 6. Best fits of the data to a saturation model give lower limits for k_{cat} of 0.11 min^{-1} and K_M of 1100 μM for B, and lower limits for k_{cat} of 0.5 min^{-1} and K_M of 1100 μM for 3-HB. These kinetic parameters are lower limits because concentrations of B and 3-HB that saturate LM-M1 were not reached. However, k_{cat}/K_M values of 100 M^{-1} min^{-1} for B and 460 M^{-1} min^{-1} for 3-HB can be accurately determined from the linear portion of each curve. Unexpectedly, these values are ~8-fold (B) and ~20-fold (3-HB) less than the analogous k_{cat}/K_M values for arylation of isolated LM (*9*). It remains to be determined how the M1 domains that are not present in LM alone apparently decrease the rate of covalent arylation of the LM module when embedded in LM-M1.

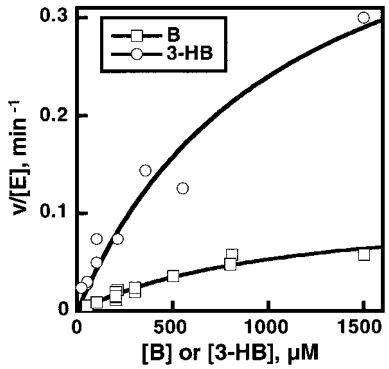
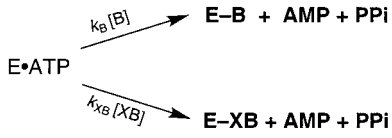


FIGURE 6: Dependence of the rate of covalent loading of *holo* LM-M1 on the concentration of the substrates B (○) or 3-HB (□). Although LM-M1 could not be saturated with either substrate, the lines are best fits of the data to a simple saturation model and give lower limits for k_{cat} of 0.11 min^{-1} and K_M of 1100 μM for B, and lower limits for k_{cat} of 0.5 min^{-1} and K_M of 1100 μM for 3-HB. Despite the inability to saturate LM-M1, valid k_{cat}/K_M values of 100 M^{-1} min^{-1} for B and 460 M^{-1} min^{-1} for 3-HB can be determined from the linear portion of each curve.

Scheme 1[a]



$$k_{rel} = \frac{k_{XB}}{k_B} = \frac{[E\text{–}XB][B]}{[XB][E\text{–}B]}$$

[a] E = Rif LM-M1, B = benzoate, XB = substituted benzoate.

Based on previous in vivo feeding experiments, 3-HB and 3,5-dihydroxybenzoate, in addition to the natural substrate AHB, are known to be processed by the loading and three initial elongation modules (*24*). In vitro results obtained for LM alone revealed 11 additional substituted benzoates to be substrates for this module (*9*). Prior to determining whether the LM-M1 protein could process these benzoates into the corresponding aryl ketide products, we investigated whether they were accepted as substrates for covalent arylation of LM-M1.

Addition of a substituted benzoate to a reaction mixture containing radiolabeled B and *holo* LM-M1 allowed partitioning between reaction with the substituted benzoate and reaction with B to be followed. The relative rate constant (k_{rel}) for reaction of a given substituted benzoate with respect to reaction of B was determined from the concentrations of B and substituted benzoate in the original reaction ([B], [XB]) and the amount of product present as E-B and E-XB, according to the equation in Scheme 1 (*21*). The ratio of E-B to E-XB was constant throughout a given time course, indicating that no secondary reactions involving the reaction products were occurring. The constant ratios also support the view that the substituted benzoates are true substrates and not high-affinity competitive inhibitors, as E-B would continue to accumulate in the presence of a competitive inhibitor, resulting in a ratio of E-B to apparent E-XB that increases as a function of time. For each substituted benzoate, the same k_{rel} value, within error, was obtained for reactions

Table 1: Relative Rate Constants for Covalent Arylation of LM-M1 and LM by Substituted Benzoates[a]

substrate	LM-M1 k_{rel}[b]	LM k_{rel}[c]	k_{rel} ratio[d]
3-amino-5-hydroxybenzoate	66 ± 20	120 ± 10	0.6
3,5-diaminobenzoate	10 ± 2	16 ± 1	0.6
3-hydroxybenzoate	6 ± 3	12 ± 2	0.5
3,5-dihydroxybenzoate	3.8 ± 0.9	3.1 ± 0.5	1.2
3,5-dichlorobenzoate	3.5 ± 0.6	4.0 ± 0.5	0.9
3-aminobenzoate	2.5 ± 1.1	6.6 ± 0.6	0.4
3,5-dibromobenzoate	1.8 ± 0.4	4.1 ± 0.5	0.4
3-chlorobenzoate	1.1 ± 0.4	2.1 ± 0.2	0.5
3-bromobenzoate	1.1 ± 0.5	1.9 ± 0.2	0.6
2-aminobenzoate	1.1 ± 0.3	0.62 ± 0.08	1.8
benzoate	(1)	(1)	(1)
3-methoxybenzoate	0.30 ± 0.06	0.43 ± 0.06	0.7
3-fluorobenzoate	0.22 ± 0.08	0.42 ± 0.11	0.5
3,5-difluorobenzoate	0.12 ± 0.03	0.13 ± 0.02	0.9

[a] Rate constant for arylation, relative to arylation by benzoate (Scheme 1). These k_{rel} values, obtained from substrate competition experiments, represent the k_{cat}/K_M ratio for a given substituted benzoate and B (*21*). [b] 30 °C, 100 mM Tris, pH 7.4, 5 mM TCEP, 15 mM $MgCl_2$, 5 mM ATP. [c] 30 °C, 50 mM sodium phosphate, pH 7.2, 1 mM DTT, 1 mM EDTA, 15 mM $MgCl_2$, 5 mM ATP, 10% glycerol. From ref *9*. [d] Ratio is k_{rel} for LM-M1 divided by k_{rel} for LM.

performed at different substituted benzoate concentrations. The reactions were repeated for selected substituted benzoates using radiolabeled 3-HB instead of B, and the same k_{rel} values (with respect to B), within error, were obtained. Each k_{rel} value for LM-M1 in Table 1 represents an average of at least four separate determinations.

The k_{rel} values for LM-M1 in Table 1 represent the k_{cat}/K_M ratio for a given substituted benzoate and B, and as such provide a measure of the specificity of LM-M1 for each substrate (*21*). The validity of this approach is demonstrated by comparing the k_{rel} value of 6 ± 3 obtained for 3-HB with the k_{cat}/K_M ratio of 4.6 obtained from direct measurement of k_{cat}/K_M for 3-HB and B (Figure 6). The k_{rel} values for LM-M1 are compared to previously determined k_{rel} values for LM alone in Table 1. The overall trends in k_{rel} correlate for LM-M1 and LM. For example, the natural substrate AHB is preferred by both LM-M1 and LM, whereas 3,5-difluorobenzoate is the poorest substrate for both proteins. It is possible that the small differences in k_{rel}, reflected by deviations from one in the ratios of the two sets of values, are due to experimental error or to the slightly different reaction conditions used for the two sets of measurements (see Table 1).

Substrate Tolerance of Aryl Ketide Formation by LM-M1. The substituted benzoates that were shown to arylate LM-M1 above (Table 1) were next surveyed for their ability to be processed into the corresponding aryl ketide products by LM-M1 (Figure 7). Aryl ketide product is not formed in the absence of any benzoate, so only methylmalonate is visible upon TCA precipitation and hydrolysis of the control reaction (lane 9). Lanes 1−8 show products of reactions that contained the substituted benzoates indicated in the figure legend; new aryl ketides are observed in each case. Aryl ketide products of reactions that contained 2-aminobenzoate, 3-aminobenzoate, AHB, 3,5-dihydroxybenzoate, and 3-HB could also be detected, although different TLC conditions than those depicted in Figure 7 were required to resolve these products from methylmalonate. No product of reactions containing 3,5-diaminobenzoate has been detected, although

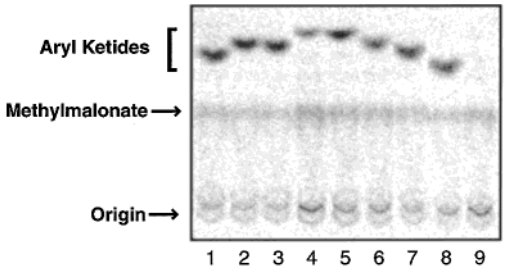


FIGURE 7: LM-M1 converts a series of alternative benzoate substrates to the corresponding aryl ketides. [^{14}C]-Aryl ketides were resolved from [^{14}C]-methylmalonate by TLC following TCA precipitation and hydrolysis of reactions containing LM-M1, [^{14}C]-mmCoA, and the following substrates: B (lane 1); 3-bromobenzoate (lane 2); 3-chlorobenzoate (lane 3); 3,5-dibromobenzoate (lane 4); 3,5-dichlorobenzoate (lane 5); 3,5-difluorobenzoate (lane 6); 3-fluorobenzoate (lane 7); 3-methoxybenzoate (lane 8); no benzoate (lane 9). Aryl ketide products for reactions of LM-M1 and mmCoA with 2-aminobenzoate, 3-aminobenzoate, AHB, 3,5-dihydroxybenzoate, and 3-HB were observed but are not shown.

3,5-diaminobenzoate has been shown to covalently load both isolated LM (*9*) and LM-M1 (Table 1). It is possible that methylmalonate and the 3,5-diaminobenzoyl moiety are not converted into the corresponding aryl ketide by M1; we consider it more likely that TLC conditions that resolve the product have yet to be been identified. Thirteen structurally distinct benzoates appear to be converted to the corresponding aryl ketides by LM-M1, suggesting that LM-M1 has a broad substrate tolerance.

Preliminary Kinetic Analysis of Phenyl Ketide Formation by LM-M1. These studies reveal that LM-M1 catalyzes the formation of a variety of aryl ketides. We turned to kinetic methods to compare the different benzoates as substrates for aryl ketide formation, and to determine what steps limit the overall process. It is important to note that multiple reactions are involved in the overall formation of aryl ketides, including covalent loading of substrates onto LM-M1, interdomain transfer of these substrates, the condensation reaction itself, and the reduction reaction catalyzed by the KR domain (Figure 2). Therefore, interpretation of the kinetic results should take this complexity into account.

The LM-M1-catalyzed conversion of B and mmCoA to phenyl ketide was investigated in the most detail, since this reaction is experimentally convenient. Both B and mmCoA are available in radiolabeled form, allowing the reaction to be monitored using either substrate (Figure 5), and the radiolabeled B has a high specific activity which facilitates experiments that require substrate concentrations to be varied over a wide range. Additionally, no degradation products of 3-hydroxy-2-methyl-3-phenylpropionate were detected by HPLC after incubation of the authentic standard in 0.5 M KOH for up to 60 min at 65 °C, indicating that this phenyl ketide product of the enzymatic reaction is stable under conditions that are even more harsh than the conditions used to hydrolytically release it from LM-M1 (0.2 M KOH, 65 °C, 20 min). Therefore, measuring the amount of phenyl ketide in reaction time points following hydrolysis of the covalent linkage between it and the enzyme is a quantitative assay for the enzymatic formation of phenyl ketide.

The accumulation of enzyme-linked B (E-B) and enzyme-linked phenyl ketide (E-PK) products in a representative reaction containing LM-M1, radiolabeled B, mmCoA, and

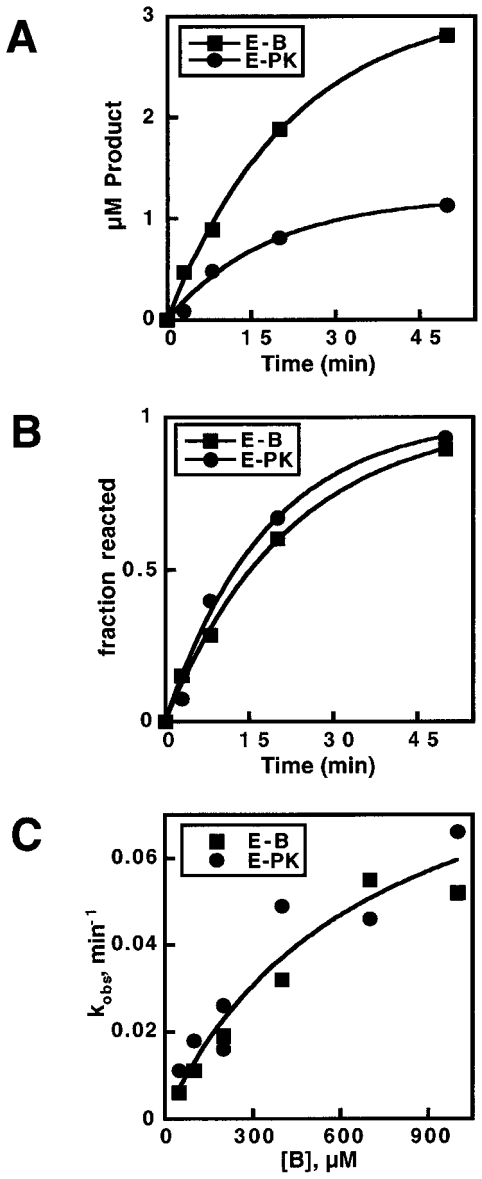


FIGURE 8: Rate of overall phenyl ketide formation is limited by the rate of covalent arylation of LM-M1 by B. (A) Time course of E-B (■) and E-PK (●) formation in reactions that contained 5 μM LM-M1, 1 mM B, and 5 mM mmCoA, in addition to other reaction components. The lines are best exponential fits to each data set and give k_{obs} of 0.05 min^{-1} and an endpoint of 3.2 μM for E-B and k_{obs} of 0.06 min^{-1} and an endpoint of 1.2 μM for E-PK. (B) Time course in (A) replotted using the fraction of E-B or E-PK formed (calculated from the reaction endpoints determined above) to illustrate that k_{obs} values for E-B and E-PK formation are identical, within error. (C) Dependence of E-B and E-PK formation on the concentration of the substrate B. Although LM-M1 could not be saturated with B, the line is a best fit of the combined data to a simple saturation model and gives lower limits for k_{cat} of 0.10 min^{-1} and K_M of 700 μM. The k_{cat}/K_M value of 140 M^{-1} min^{-1} that is determined from the linear portion of the curve is in good agreement with the value of 100 M^{-1} min^{-1} for E-B formation measured using TCA precipitation followed by scintillation counting (Figure 6).

the remaining reaction components is shown in Figure 8A. For this representative reaction and for other reactions in which the concentrations of B and mmCoA were varied widely, two observations were made. First, the endpoint for accumulation of E-B typically exceeded the endpoint for accumulation of E-PK by at least 2-fold. For example, for the reaction shown in Figure 8A, which contained 5 μM total enzyme, E-B formation leveled off at approximately 3.2 μM and E-PK formation leveled off at approximately 1.2 μM. Second, despite the different endpoints for the two products, E-B and E-PK accumulated simultaneously, with the same apparent rate constant (k_{obs}). This is illustrated in Figure 8B, where the fraction of total E-B and E-PK from Figure 8A is plotted as a function of time.

The different endpoints that were consistently observed for the formation of E-B and E-PK were unexpected, as the simplest model for aryl ketide formation predicts that each active LM-M1 molecule would have a molecule of B covalently bound to its T_1 domain and a molecule of phenyl ketide covalently bound to its T_2 domain, giving stoichiometric amounts of B and phenyl ketide upon hydrolysis. Several specific models for the nonstoichiometric product amounts that are observed are encompassed by the general model that M1 is only competent for condensation of B and methylmalonate in a fraction of the LM-M1 proteins. If the AT or T_2 domains were inactive in a fraction of the LM-M1 proteins, these proteins could not be covalently acylated with methylmalonate and would therefore not produce phenyl ketide. This model is unlikely, as LM-M1 is able to covalently load 2 equiv of radiolabeled mmCoA (data not shown), suggesting that both the AT and T_2 domains are active for acylation. The T_2 domain could nevertheless become inactivated if decarboxylation of its methylmalonyl moiety ever occurred without concomitant condensation, because the resultant dead-end propionate adduct could not participate in phenyl ketide formation. However, propionate was not detected upon hydrolysis of LM-M1 that had been incubated with mmCoA (data not shown), suggesting that this uncoupling model does not account for the endpoint results. The possibility that an internal enzyme equilibrium governs the ratio of E-B and E-PK is unlikely, because the release of CO_2 upon E-PK formation is expected to render this reaction effectively irreversible. Alternative models invoke the KS domain, the catalyst of the condensation reaction. If only a fraction of the LM-M1 proteins possessed functional KS domains, only this fraction could produce phenyl ketide, accounting for the observation that E-B exceeds E-PK. Nonstoichiometric amounts of E-B and E-PK could also result if LM-M1 is stably loaded with two molecules of B, one covalently linked to the T_1 domain and another covalently linked to the active site cysteine of the KS domain. In the simplest incarnation of this model, twice as much E-B as E-PK would be present at the endpoint. Cross-talk between protein subunits could also contribute to the complexity of the data, if models containing higher order protein species are invoked.

Despite the different endpoints for E-B and E-PK, the two product-forming reactions can be directly compared by considering the fraction of LM-M1 that does form each product and determining the corresponding k_{obs} values (Figure 8B). When k_{obs} values obtained in this way are plotted as a function of B concentration, the data for E-B and E-PK coincide (Figure 8C), suggesting that the same reaction step is rate-limiting for product formation in each case. The k_{cat}/K_M value of 140 M^{-1} min^{-1} that is determined from the linear portion of the combined data is in good agreement with the k_{cat}/K_M value of 100 M^{-1} min^{-1} for E-B formation that was measured using TCA precipitation followed by scintillation counting (Figure 6). E-B formation is on the pathway to

E-PK formation, and the kinetics of E-B formation and overall E-PK formation coincide, so the formation of E-B appears to be rate-limiting for the formation of E-PK. In other words, compared to formation of the E-B species, all of the subsequent reactions that are required for E-PK formation are fast.

If E-PK formation is indeed limited by E-B formation, then preincubation of LM-M1 with B followed by addition of mmCoA to the reaction should result in a burst of E-PK formation. Figure 9A shows the formation of E-B and E-PK in a control reaction in which mmCoA was present from the start, whereas the formation of E-PK upon addition of mmCoA following either 20 or 50 min of preincubation with B is shown in Figure 9B. For both preincubation durations, a burst of E-PK formation is observed upon mmCoA addition, and in both cases, the burst amplitude approximately matches the E-PK amount that is present in the control reaction which lacked a preincubation period (Figure 9B). The bursts of E-PK formation are replotted in Figure 9C as a function of time after addition of mmCoA to each reaction; although the burst amplitudes vary as expected based on the duration of preincubation (and, therefore, the amount of E-B available for immediate conversion to E-PK), the apparent rate constants of 1.1 and 1.5 min^{-1} for the bursts are experimentally indistinguishable. The observation that a burst of E-PK formation occurs upon addition of mmCoA to a preformed E-B complex strengthens the conclusion that all of the downstream reactions that lead to E-PK formation are fast relative to formation of the E-B species.

Preliminary Kinetic Analysis of Aryl Ketide Formation by LM-M1. The LM-M1-catalyzed conversions of substituted benzoate substrates and mmCoA into aryl ketides were investigated and compared to the phenyl ketide forming reaction. A concentration of 50 μM B is subsaturating for LM-M1, based on the concentration dependence of the phenyl ketide forming reaction (Figure 8C). This concentration of each benzoate substrate was therefore selected to measure the velocity of formation of the corresponding aryl ketide, with the expectation that this concentration would also be subsaturating for most of the other benzoates. If this proved to be the case, then the velocities obtained under these conditions would be proportional to k_{cat}/K_M, allowing the benzoates to be directly compared as substrates of LM-M1. The results of these experiments are shown in Table 2, which lists both the initial velocity of aryl ketide formation for each benzoate under these defined conditions, as well as the relative velocity (V_{rel}) for formation of each aryl ketide with respect to the reaction of B to form phenyl ketide.

A comparison of the k_{rel} values for LM-M1 arylation in Table 1 with the V_{rel} values for aryl ketide formation in Table 2 reveals the same overall trends. For example, 3,5-difluorobenzoate is the poorest substrate with respect to both arylation and aryl ketide formation. Furthermore, there is a close correspondence between the values of k_{rel} and V_{rel} for all substrates except AHB (see below). This suggests that, as for the reaction of B and mmCoA to form phenyl ketide, the overall formation of each aryl ketide from a benzoate and mmCoA is limited by arylation of LM-M1.

The one outlier in this comparison between the relative rates of arylation and aryl ketide formation is the physiological substrate AHB. Although the rate constant for arylation of LM-M1 with AHB is 66-fold larger than for arylation of LM-M1 with B (Table 1), the velocity of 3-amino-5-hydroxyphenyl ketide formation is only 9.9-fold faster than the velocity of phenyl ketide formation for reactions containing 50 μM AHB or B, respectively (Table 2). We believe that this discrepancy is due to the conditions chosen for comparing aryl ketide formation, not due to an actual change in the rate-limiting step for the AHB reaction. We suspected that AHB, the natural substrate of rifamycin

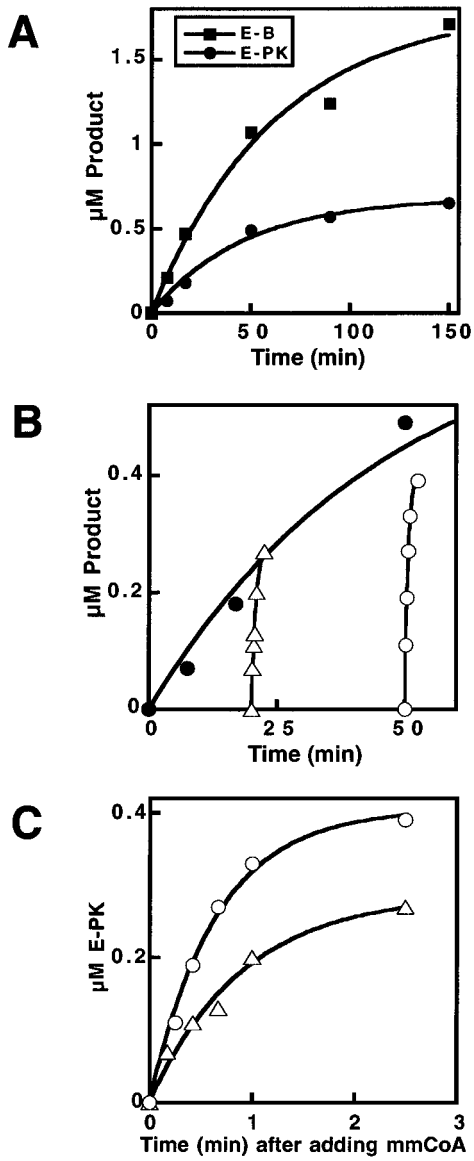


FIGURE 9: Direct comparison of overall and pre-arylation velocities of phenyl ketide formation. (A) Time course of E-B (■) and E-PK (●) formation in control reactions that contained 5 μM LM-M1, 200 μM B, and 50 μM mmCoA, in addition to other reaction components. The lines are best exponential fits to each data set and give k_{obs} of 0.02 min^{-1} and an endpoint of 2 μM for E-B and k_{obs} of 0.02 min^{-1} and an endpoint of 0.7 μM for E-PK. (B) 50 μM mmCoA was added to comparable reactions containing 5 μM LM-M1 and 200 μM B at times 20 min (△) and 50 min (○), and the fast accumulation of E-PK was monitored in each case. The data for E-PK (●) formation from the control reaction shown in (A) are replotted here. (C) Data from reactions in (B) that were preincubated with B for 20 min (△) and 50 min (○) are displayed as a function of time after addition of mmCoA. The lines are best exponential fits to each data set and give k_{obs} of 1.1 min^{-1} and a burst amplitude of 0.3 μM, and k_{obs} of 1.5 min^{-1} and a burst amplitude of 0.4 μM, for the reactions that were preincubated with B for 20 and 50 min, respectively.

Table 2: Velocity of Aryl Ketide Formation by LM-M1

substrate	overall velocity[a] (μM/min)	V_{rel}[b]
3-amino-5-hydroxybenzoate	0.072	9.9
3-hydroxybenzoate	0.026	3.6
3,5-dihydroxybenzoate	0.017	2.3
3,5-dichlorobenzoate	0.023	3.2
3-aminobenzoate	0.012	1.6
3,5-dibromobenzoate	0.018	2.5
3-chlorobenzoate	0.010	1.4
3-bromobenzoate	0.0099	1.4
2-aminobenzoate	0.0027	0.37
benzoate	0.0073	(1)
3-methoxybenzoate	0.0046	0.63
3-fluorobenzoate	0.0046	0.63
3,5-difluorobenzoate	0.00074	0.10

[a] 30 °C, 100 mM Tris, pH 7.4, 5 mM TCEP, 15 mM $MgCl_2$, 5 mM ATP, 5 mM NADPH, 5 μM LM-M1, 50 μM mmCoA, 50 μM substituted benzoate substrate. The velocities obtained under these conditions are expected to be proportional to k_{cat}/K_M for all substrates except AHB (see Results), by analogy to the B reaction in which 50 μM substrate is subsaturating. [b] Ratio of the overall velocity of aryl ketide formation and the overall velocity of phenyl ketide formation in the analogous reaction containing B.

synthetase, would saturate LM-M1 at a lower concentration than does B, so we tested the possibility that the 50 μM benzoate conditions selected for comparison of aryl ketide formation were saturating for AHB. Detection limitations prevented measurement of K_M for AHB, but a limit of $K_M = 10$ μM could be obtained (data not shown). Therefore, the velocity reported for aryl ketide formation in the presence of 50 μM AHB is not proportional to k_{cat}/K_M, as are the velocities reported for B and presumably for the other benzoates (Table 2). This means that V_{rel} for AHB in Table 2 is a lower limit for the V_{rel} that would be obtained if AHB were subsaturating, leaving open the possibility that arylation of LM-M1 limits aryl ketide formation in the AHB reaction, as for reactions of the other benzoate substrates.

If aryl ketide formation is limited by arylation of LM-M1 for all of the benzoate substrates, then preincubation of LM-M1 with any of the benzoates followed by addition of mmCoA to the reaction should result in a burst of aryl ketide formation, as depicted in Figure 9 for the B reaction. Bursts of aryl ketide formation were observed following addition of mmCoA to LM-M1 that had been preincubated with any of the substrates in Table 2. Each burst could be described by an apparent rate constant between 1 and 3 min^{-1} (data not shown). These apparent rate constants are at the practical limit of measurements that can be made using manual techniques (half-lives 10−30 s), so it is not clear whether individual k_{obs} values truly differ or whether the variability is due to experimental error. The fast accumulation of aryl ketides following pre-arylation suggests that for all of the substrates in Table 2, aryl ketide formation is limited by formation of arylated LM-M1.

DISCUSSION

Reconstitution of LM-M1. Our results demonstrate that the excised loading and initial elongation modules of rifamycin synthetase collaborate to form aryl ketide products. Successful reconstitution of the two modules of LM-M1 requires a benzoate, mmCoA, Mg·ATP, and NADPH. These experiments verify the functioning of four catalytic and two carrier domains and the inactivity of a seventh domain (Figure 2). The A domain catalyzes the reaction between a benzoate and ATP to produce an activated aryl-AMP species. The thiol of the phosphopantetheine cofactor of the T_1 domain attacks the aryl-AMP to form an arylated enzyme intermediate. The AT domain catalyzes the transfer of the methylmalonyl moiety of mmCoA to the thiol of the phosphopantetheine cofactor of the T_2 domain via an acylserylenzyme intermediate. The KS domain catalyzes the transfer of the aryl moiety from T_1 to its active site cysteine, and then effects decarboxylative condensation between the KS-linked aryl moiety and T_2-linked methylmalonate. Finally, the KR domain uses NADPH to reduce the T_2-linked aryl ketide, and this reduced product is released from LM-M1 upon hydrolysis. The absence of further reduction in the final product is consistent with an inactive DH domain, as expected due to active site amino acid deletions in this domain (*4*−*6*).

Substrate Tolerance of LM-M1. Although the natural substrate of the loading module of rifamycin synthetase is AHB, it also accepts alternative benzoates. LM-M1 can become covalently arylated with 14 different benzoate substrates, recapitulating the substrate tolerance of the isolated LM (*9*; Table 1). Moreover, the benzoates are subsequently transferred to the KS domain of M1 and processed by this module, as evidenced by the formation of aryl ketide products in reactions containing these benzoates (Figure 7, Table 2). It remains to be determined whether downstream PKS modules will tolerate the unnatural benzoate component of these aryl ketides and extend them further. Aryl triketide shunt products were detected upon feeding of 3-HB or 3,5-dihydroxybenzoate to the rifamycin producer *Amycolatopsis mediterranei* (*24*), suggesting that at least the RifA protein, which contains the LM and the first three PKS modules of rifamycin synthetase, accommodates these alternative benzoates.

Gatekeeper Function for the Loading Module of LM-M1. The kinetics of arylation and overall aryl ketide formation are the same for a given benzoate, and pre-arylation of LM-M1 with a benzoate substrate prior to addition of mmCoA results in a burst of aryl ketide formation, due to the unmasking of faster reactions that are downstream of arylation. These results suggest that for the benzoates we have studied, arylation of LM-M1 is rate-limiting for the formation of aryl ketide from a benzoate and mmCoA, and intermodular aryl transfer, carbon−carbon bond formation, and β-ketone reduction are fast by comparison. Our assay for LM-M1 arylation does not differentiate between initial formation of aryl-AMP and subsequent transfer of the aryl moiety to the T_1 domain, so further experimentation will be required to pinpoint the reaction step or steps responsible for the rate-limiting effect.

These results are consistent with a gatekeeper function for the loading module of rifamycin synthetase. The loading module exhibits a 10−1000-fold preference for AHB, its biological substrate, over all other substrates tested. This discrimination between benzoates ensures that AHB is loaded when comparable concentrations of AHB and competing benzoates are available. However, once the loading module is covalently arylated with a benzoate substrate, all of the benzoates surveyed are converted into aryl ketides by M1 with very similar apparent rate constants of 1−3 min^{-1} (see Results). These observations do not rule out the possibility that there is discrimination between the benzoates at reaction

steps other than loading, but at the level of resolution of our initial experiments, none is detected. Discrimination at other reaction steps could be masked by a slow step. For example, even a large difference between the rates of intermodular transfer of two benzoates from the T_1 domain to the KS domain would be masked if the rates of condensation were slower and equivalent for both substrates. The gatekeeping ability of the loading module is likely to be the result of natural selection for binding and reaction of AHB, rather than an evolved discrimination against other benzoates, because most of the benzoates that have been identified as substrates are not commonly found in rifamycin-producing organisms.

Our preliminary kinetic analyses set the stage for dissection of the pathway to aryl ketide formation into its individual reactions. Once a mechanistic framework that includes all of the reactions that convert free B and mmCoA into phenyl ketide has been established, analogous experiments could be performed using alternative benzoate substrates to reveal additional steps, if any, at which discrimination occurs and the magnitude of such discrimination.

Single or Multiple Turnover in Modular Assembly Lines. Upon completion of passage through multimodular NRPS, PKS, or mixed NRPS−PKS assembly lines, natural product chains must be released from their final thioester tether to regenerate the system for another round of biosynthesis. The necessary chain termination reaction is typically catalyzed by hydrolytic or cyclizing thioesterases, aldehyde-forming reductases, or amide synthases (*25*).

To facilitate biochemical analysis, portions of assembly line machineries are often studied in isolation, so the natural mechanism of chain termination is lost. In some cases, engineering a TE domain onto the C-terminal end of the protein of interest has restored chain termination ability (*1*−*3*). In other cases, multiple turnover of partial assembly lines occurs because the thioester linkage between the product and the protein is cleaved by a nucleophilic component of the product, an exogenous nucleophile, or water during the reaction time course (*26, 27*). In the absence of multiple turnover, the covalent thioester attachment between a protein and its acyl product can be hydrolyzed by base so that the detached product can be analyzed (*28*−*30*). For certain types of analyses in specific NRPS systems, the use of aminoacyl-*S*-(*N*-acetyl)cysteamine substrates in place of aminoacyl-*S*-enzyme substrates obviates the need for a chain termination mechanism, because the final products are soluble (*31, 32*).

Intact rifamycin synthetase has an amide synthase (RifF) that releases and cyclizes the rifamycin backbone, but the truncated LM-M1 protein lacks an intrinsic chain terminator. Although we attempted to engineer a multiple turnover system by appending the TE domain of 6-deoxyerythronolide B synthase to the C-terminal end of LM-M1, TE-catalyzed hydrolysis of the enzyme-linked aryl ketide product was not observed.[3] No promising candidates were identified when several nucleophiles were surveyed for their ability to release a nucleophile−product adduct. Thus, the activity of LM-M1 was analyzed using single turnover reactions, requiring hydrolytic release of covalently bound products and intermediates.

There are several constraints associated with hydrolytic release of protein-bound products. The major drawback to this method is product yield, as at most 1 equiv of product is formed per equivalent of protein. This can be a serious limitation if reasonable quantities of protein are difficult to obtain. For example, 2.5 mg of the LM-M1 protein (260 kDa) was required to obtain ~300 ng of phenyl ketide (180 Da), an amount sufficient for product analysis by mass spectrometry. Furthermore, certain types of experiments that require low enzyme concentrations cannot be performed effectively, because detection of the correspondingly low product amounts (≤stoichiometric) is prohibitive. For example, for a substrate with a low K_M, enzyme concentration would need to be even lower to measure k_{cat}/K_M accurately (*33*), and the amount of product attainable for detection would be limited by this enzyme concentration. This situation was encountered for AHB, the biological substrate of LM-M1, preventing us from measuring a V_{rel} value for 3-amino-5-hydroxyphenyl ketide formation that could be directly compared to V_{rel} values for formation of other aryl ketides (see Results). A final problem associated with hydrolytic release is that acyl products of interest may be degraded when subjected to base at an elevated temperature, preventing or complicating product analysis (*29*).

Despite the shortcomings of hydrolytic release of acyl-enzyme intermediates as an experimental method, we were able to use this approach in conjunction with analytical techniques to identify the product of reactions containing LM-M1, benzoate, mmCoA, Mg·ATP, and NADPH as 3-hydroxy-2-methyl-3-phenylpropionate (phenyl ketide). Hydrolytic release of radiolabeled phenyl ketide and other aryl ketides followed by analytical TLC and radiography further enabled us to obtain preliminary kinetic data for the LM-M1 system. Hydrolytic release and other single turnover methods have the advantage of providing insights into the stoichiometries of total enzyme, enzyme intermediates, and enzyme-bound product that would be missed by multiple turnover methods. We were able to discern that the amount of product formed in LM-M1 reactions was consistently substoichiometric with respect to the amounts of total enzyme and arylated enzyme intermediate. This result would have been invisible in a multiple turnover system, preventing potential complexities such as dysfunctional domains and internal equilibria from being discovered. Although natural product assembly lines are typically depicted as entities that are fully occupied at each T domain with covalently attached intermediates, such stoichiometry has not been generally validated, and little is known about how the occupancy of one module of a natural product assembly line affects the occupancy of other modules. Hydrolytic release of intermediates and products from modular assembly lines may allow the distribution of covalent species, dictated by the relative rates of chain initiation, elongation, and termination, to be investigated in detail.

Engineering Potential of LM-M1. As a model NRPS−PKS system, LM-M1 could serve as a template for engineering novel NRPS−PKS enzymes. The sequence encoding M1 could be replaced by genes encoding other representative PKS modules, using known homologies between modules to choose appropriate fusion sites. Likewise, LM could be systematically replaced with other NRPS modules. Comparison of the kinetic parameters of the parent LM-M1 system with the parameters of the variant systems is expected to reveal information about the steps at which interactions between NRPS and PKS modules occur, and therefore the

individual reactions and domains at which efforts to engineer novel, functional NRPS−PKS interfaces should be directed.

ACKNOWLEDGMENT

We thank Daniel Hunziker for synthesis of AHB and Dr. Lusong Luo, Dr. Patrick O'Brien, and Dr. Sarah O'Connor for comments on the manuscript.

REFERENCES

1. Cortes, J., Wiesmann, K. E. H., Roberts, G. A., Brown, M. J. B., Staunton, J., and Leadlay, P. F. (1995) *Science 268*, 1487−1489.
2. Kao, C. M., Luo, G., Katz, L., Cane, D. E., and Khosla, C. (1995) *J. Am. Chem. Soc. 117*, 9105−9106.
3. Kao, C. M., Luo, G., Katz, L., Cane, D. E., and Khosla, C. (1996) *J. Am. Chem. Soc. 118*, 9184−9185.
4. Schupp, T., Toupet, C., Engel, N., and Goff, S. (1998) *FEMS Microbiol. Lett. 159*, 201−207.
5. August, P. R., Tang, L., Yoon, Y. J., Ning, S., Müller, R., Yu, T.-W., Taylor, M., Hoffmann, D., Kim, C.-G., Zhang, X., Hutchinson, C. R., and Floss, H. G. (1998) *Chem. Biol. 5*, 69−79.
6. Tang, L., Yoon, Y. J., Choi, C.-Y., and Hutchinson, C. R. (1998) *Gene 216*, 255−265.
7. Floss, H. G., and Yu, T.-W. (1999) *Curr. Opin. Chem. Biol. 3*, 592−597.
8. Campbell, E. A., Korzheva, N., Mustaev, A., Murakami, K., Nair, S., Goldfarb, A., and Darst, S. A. (2001) *Cell 104*, 901−912.
9. Admiraal, S. J., Walsh, C. T., and Khosla, C. (2001) *Biochemistry 40*, 6116−6123.
10. Du, L., Sánchez, C., Chen, M., Edwards, D. J., and Shen, B. (2000) *Chem. Biol. 7*, 623−642.
11. Julien, B., Shah, S., Ziermann, R., Goldman, R., Katz, L., and Khosla, C. (2000) *Gene 249*, 153−160.
12. Schwecke, T., Aparicio, J. R., Molnár, I., König, A., Khaw, L. E., Haydock, S. F., Oliykyk, M., Caffrey, P., Cortés, J., Lester, J. B., Böhm, G. A., Staunton, J., and Leadlay, P. F. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 7839−7843.
13. Motamedi, H., and Shafiee, A. (1998) *Eur. J. Biochem. 256*, 528−534.
14. Ghisalba, O., and Nüesch, J. (1981) *J. Antibiot. 34*, 64−71.
15. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.
16. Gokhale, R. S., Tsuji, S. Y., Cane, D. E., and Khosla, C. (1999) *Science 284*, 482−485.
17. Tsuji, S. Y., Cane, D. E., and Khosla, C. (2001) *Biochemistry 40*, 2326−2331.
18. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
19. Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., Reid, R., Khosla, C., and Walsh, C. T. (1996) *Chem. Biol. 3*, 923−936.
20. Quadri, L. E. N., Weinreb, P. H., Lei, M., Nakano, M. M., Zuber, P., and Walsh, C. T. (1998) *Biochemistry 37*, 1585−1595.
21. Fersht, A. R. (1998) in *Structure and Mechanism in Protein Science*, pp 116−117, W. H. Freeman, New York.
22. Walsh, C. T., Gehring, A. M., Weinreb, P. H., Quadri, L. E. N., and Flugel, R. S. (1997) *Curr. Opin. Chem. Biol. 1*, 309−315.
23. Khosla, C. (1997) *Chem. Rev. 97*, 2577−2590.
24. Hunziker, D., Yu, T.-W., Hutchinson, C. R., Floss, H. G., and Khosla, C. (1998) *J. Am. Chem. Soc. 120*, 1092−1093.
25. Keating, T. A., Ehmann, D. E., Kohli, R. M., Marshall, C. G., Trauger, J. W., and Walsh, C. T. (2001) *ChemBioChem. 2*, 99−107.
26. Stachelhaus, T., Mootz, H. D., Bergendahl, V., and Marahiel, M. A. (1998) *J. Biol. Chem. 273*, 22773−22781.
27. Gehring, A. M., Mori, I., Perry, R. D., and Walsh, C. T. (1998) *Biochemistry 37*, 11637−11650.
28. Stachelhaus, T., and Walsh, C. T. (2000) *Biochemistry 39*, 5775−5787.
29. Chen, H., and Walsh, C. T. (2001) *Chem. Biol. 8*, 301−312.
30. Chen, H., O'Connor, S., Cane, D. E., and Walsh, C. T. (2001) *Chem. Biol. 8*, 899−912.
31. Ehmann, D. E., Trauger, J. W., Stachelhaus, T., and Walsh, C. T. (2000) *Chem. Biol. 7*, 765−772.
32. Marshall, C. G., Burkart, M. D., Keating, T. A., and Walsh, C. T. (2001) *Biochemistry 40*, 1066−10663.
33. Fersht, A. R. (1998) in *Structure and Mechanism in Protein Science*, pp 103−106, W. H. Freeman, New York.

BI0200312