

A New Family of Type III Polyketide Synthases in *Mycobacterium tuberculosis**

Received for publication, June 24, 2003, and in revised form, July 25, 2003
Published, JBC Papers in Press, August 26, 2003, DOI 10.1074/jbc.M306714200

Priti Saxena, Gitanjali Yadav‡, Debasisa Mohanty, and Rajesh S. Gokhale§

From the National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110 067, India

The *Mycobacterium tuberculosis* genome has revealed a remarkable array of polyketide synthases (PKSs); however, no polyketide product has been isolated thus far. Most of the PKS genes have been implicated in the biosynthesis of complex lipids. We report here the characterization of two novel type III PKSs from *M. tuberculosis* that are involved in the biosynthesis of long-chain α -pyrones. Measurement of steady-state kinetic parameters demonstrated that the catalytic efficiency of PKS18 protein was severalfold higher for long-chain acyl-coenzyme A substrates as compared with the small-chain precursors. The specificity of PKS18 and PKS11 proteins toward long-chain aliphatic acyl-coenzyme A (C_{12} to C_{20}) substrates is unprecedented in the chalcone synthase (CHS) family of condensing enzymes. Based on comparative modeling studies, we propose that these proteins might have evolved by fusing the catalytic machinery of CHS and β -ketoacyl synthases, the two evolutionarily related members with conserved thiolase fold. The mechanistic and structural importance of several active site residues, as predicted by our structural model, was investigated by performing site-directed mutagenesis. The functional identification of diverse catalytic activity in mycobacterial type III PKSs provide a fascinating example of metabolite divergence in CHS-like proteins.

Mycobacteria are classified in the phylogeny of the *Actinomycetes*, along with the *Streptomyces* bacteria. Interestingly, these two actinomycete genera have received immense attention due to their contrasting effects on human society. Whereas *Streptomyces* have provided a rich source of antibiotics and other therapeutic products for human diseases, *Mycobacterium tuberculosis* and *Mycobacterium leprae* have been two of humankind's greatest scourges. Although the original phylogenetic classifications of *Mycobacterium* and *Streptomyces* were based primarily on morphology, the genome sequences of these organisms have further established their common lineage (1, 2). These genome sequences have revealed several families of genes that are common between them, which include an unusually large number of gene clusters that are homologous to polyketide synthases (PKSs).¹ Although a polyketide product

has not yet been isolated from *M. tuberculosis*, recent studies have implicated some of its PKS genes in the biosynthesis of complex lipids (3). In this report, we have characterized two novel polyketide synthases from *M. tuberculosis*, which are involved in the biosynthesis of unusual long-chain α -pyrones.

Polyketides are a diverse class of secondary metabolites that possess a broad range of biological activities (4). Despite structural diversity of these natural products, PKSs synthesize polyketides by a common chemical strategy. Initial priming by a starter molecule is followed by repetitive decarboxylative condensation of coenzyme A (CoA) analogues of simple carboxylic acids. PKS elongates the polyketide chain either by repetitively using a single active site to perform multiple condensation reactions or by using a modular assembly line mechanism (5–7). Based on protein architecture, PKSs have been classified in three distinct families. Type I PKSs are giant assemblies of multifunctional polypeptides and include both iterative and modular mechanism of biosynthesis. Type II PKSs are a large multienzyme complex of discrete proteins and resemble type II fatty acid synthases found in bacteria and plants. The type III PKS has been recently discovered in bacteria and belongs to the plant chalcone synthases (CHSs) superfamily of condensing enzymes (8). These are homodimeric proteins and are structurally and mechanistically quite distinct from type I and type II PKSs (9). Type III PKSs use free CoA substrates without the involvement of a 4'-phosphopantetheine group.

CHSs are ubiquitously present in higher plants and catalyze biosynthesis of starting materials for many flavonoids (10). Several new additions to the CHS superfamily, such as acridone synthase, 2-pyrone synthase (2PS), and benzalacetone synthase, deviate from the chalcone biosynthetic model by either utilizing non-phenylpropanoid starter units or varying the number of condensation reactions (11). Some of the plant enzymes also show different cyclization patterns, emphasizing the growing diversity of the CHS superfamily. The three-dimensional crystal structures of CHS from *Medicago sativa* (alfalfa) (12) and 2PS from *Gerbera hybrida* (daisy) (13) have led to a more comprehensive understanding to structurally rationalize the mechanistic differences between enzymes within the thiolase family. Presently, the thiolase fold enzymes can be grouped into three families (11, 14, 15). β -Ketoacyl acyl carrier protein synthase (KAS) I and KAS II proteins comprise one group, and KAS III and type III PKS form another group. The third group consists of several thiolases. These enzymes have a conserved $\alpha\beta\alpha\beta$ fold architecture that forms one side of the active site cavity. The positioning of the catalytic cysteine with respect to the dimer interface is also conserved in this family of enzymes. The differences primarily exist in the structural loops that form the other face of the cavity. The identity

* This work was supported by grants from ICMR, India (to R. S. G.) and by DBT, India (to the National Institute of Immunology). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ A Senior Research Fellow of CSIR, India.

§ A Wellcome Trust International Senior Research Fellow in India. To whom correspondence should be addressed. Tel.: 91-11-26717130-461; Fax: 91-11-26162125; E-mail: rsg@nii.res.in.

¹ The abbreviations used are: PKS, polyketide synthase; CoA, coenzyme A; CHS, chalcone synthase; 2PS, 2-pyrone synthase; THN,

1,3,6,8-tetrahydroxynaphthalene; THNS, 1,3,6,8-tetrahydroxynaphthalene synthase; ESI, electrospray ionization; MS, mass spectrometry; HPLC, high pressure liquid chromatography.

and the number of the key catalytic amino acid residues might account for the observed differences in the catalytic machinery. The crystal structures also provide a framework for developing structure-based approaches for characterizing other CHS homologues. Over the past several years, CHS homologues have also been reported in bacteria. The type III PKS involved in the biosynthesis of 1,3,6,8-tetrahydroxynaphthalene (THN) has now been characterized from *S. griseus* (16), *S. erythraea* (17), *S. antibioticus* (18) and *S. coelicolor* (19). All of these THN synthases (THNSs) use malonyl-CoA as starter and extender units. The chain elongation proceeds through successive decarboxylative condensations of C-2 units to form pentaketide, which is then cyclized into THN. The THNS from *S. griseus* (also known as RppA) has been shown to possess broad substrate specificity to yield a wide variety of products (20). In several strains that produce vancomycin group of natural products, a CHS-like protein performs the biosynthesis of 3,5-dihydroxyphenylglycine by using several molecules of malonyl-CoA as substrate (21–23). The precursor to the broad spectrum antimicrobial agent 2,4-diacetylphloroglucinol is produced by a CHS homologue (*phlD*) in many strains of fluorescent *Pseudomonas* sp. (24). The PhlD protein uses acetoacetyl-CoA and two molecules of malonyl-CoA to form monoacetyl phloroglucinol, which is then converted to 2,4-diacetylphloroglucinol.

In this study, we have characterized two type III PKS proteins, PKS18 and PKS11, from *M. tuberculosis* that are involved in the biosynthesis of long-chain α -pyrones. Kinetic analysis indicated that the catalytic efficiency of PKS18 protein to use long-chain acyl-CoA substrates is severalfold higher than that for small-chain substrates. This remarkable specificity of mycobacterial proteins is unprecedented in the family of CHS-related proteins. Molecular modeling studies suggested that these proteins might have evolved by combining the active sites of CHS and KAS proteins, the two mechanistically related members of the thiolase fold enzymes. The catalytic and structural importance of several active site residues, as predicted by our structural model, was investigated by performing site-directed mutagenesis. These studies thus reveal a new class of function in this protein superfamily.

EXPERIMENTAL PROCEDURES

Materials—BAC genomic library of *M. tuberculosis* was obtained from Prof. Stewart Cole of Pasteur Research Institute (1). [2-¹⁴C]Malonyl-CoA (58.40 mCi/mmol) was procured from PerkinElmer Life Sciences, and nonradioactive acyl-CoA starter substrates were purchased from Sigma. Plant-specific starter substrates were kindly donated by Joseph P. Noel's group. [2-¹⁴C]Lauroyl-CoA (57 mCi/mmol) was synthesized using an acyl-CoA synthetase protein.²

Cloning, Expression, and Purification—Type III PKS genes, *pks10*, *pks11*, and *pks18*, were amplified by PCR from the BAC genomic library of *M. tuberculosis* using gene-specific primers and cloned into pBlue-script SK⁺ (Stratagene). The authenticity of clones was confirmed by automated nucleotide sequencing. These genes were subcloned into the pET21c vector system (Novagen) and expressed as C-terminal hexahistidine-tagged proteins in the BL21(DE3) strain of *Escherichia coli*. Optimum soluble expression was achieved by growing cultures at low temperature. Whereas induction at 22 °C with 0.5 mM isopropyl-1-thio- β -D-galactopyranoside led to expression of PKS18 in soluble form, PKS11 protein (~20%) was obtained in the supernatant by growing uninduced cultures at 18 °C for 24 h. Under the experimental conditions tested, PKS10 persistently expressed as inclusion bodies and could not be rescued in the soluble form. Proteins were partially purified using Ni²⁺-nitrilotriacetic acid-agarose (Qiagen) and further resolved on a Resource Q anion exchange column (Amersham Biosciences).

Mutagenesis—PKS18 mutants A148T, A148M, A148F, L348S, and K318A were generated using the QuikChange site-directed mutagenesis kit (Stratagene). Mutagenesis reaction was carried out in accordance with the manufacturer's protocol. Mutant clones were screened by re-

striction endonuclease digestion and confirmed by automated DNA sequencing. All of the mutant proteins were expressed and purified analogous to the wild type PKS18 protein. A148T, A148M, and A148F mutants expressed as inclusion bodies and could not be expressed in soluble form.

Enzyme Assay and Product Characterization Using Mass Spectrometry—The standard reaction conditions involved 100 μ M starter molecule and 50 μ M malonyl-CoA (inclusive of 9.12 μ M [2-¹⁴C]malonyl-CoA (58.40 mCi/mmol)). Reactions were carried out with 45 μ g of protein at 30 °C for 10–60 min and quenched with 5% acetic acid. Products were extracted with 2 \times 300 μ l of ethyl acetate and dried under vacuum. Radiolabeled products were resolved on silica gel 60 F₂₅₄ TLC plates (Merck) in ethyl acetate/hexanes/acetic acid (63:27:5, v/v/v). Reaction products were resolved on reverse phase columns (Phenomenex). A linear gradient of 30% CH₃CN in H₂O (each containing 2% acetic acid) to 60% CH₃CN in H₂O over 30 min was used for separating hexanoyl-CoA-primed reaction products. Products of lauroyl-CoA-primed reaction were separated using a solvent system of 20 mM ammonium acetate, pH 5.4 (solvent A), and CH₃CN/CH₃OH (42.5:7.5, v/v) (solvent B). Polyketide products were characterized using nanospray electrospray ionization (ESI)-mass spectrometry (MS) (API QSTAR Pulsar i MS/MS; Applied Biosystems).

Determination of Kinetic Parameters—A standard reaction volume of 75 μ l contained 100 μ M malonyl-CoA, inclusive of 9.12 μ M [2-¹⁴C]malonyl-CoA (58.40 mCi/mmol), 45 μ g of protein, 50 mM Tris, pH 7.5, 1 mM EDTA, 1 mM dithiothreitol, 10% glycerol and varied (0.5–200 μ M) starter molecule concentrations. Reactions were preincubated at 30 °C for 2 min, initiated by the addition of substrates, and then continued for 3 min. Products were extracted with 2 \times 300 μ l of ethyl acetate after quenching the reaction with 5% acetic acid. The organic layer was vacuum-dried and run on silica gel 60 F₂₅₄ TLC plates (Merck) as described above. Resolved radiolabeled products were quantified using a Fuji FLA-5000 phosphor imager.

Substrate Binding Assays—30 μ g of protein was incubated with [2-¹⁴C]malonyl-CoA (58.40 mCi/mmol) or [2-¹⁴C]lauroyl CoA (57 mCi/mmol) for 15 min at 30 °C. Reactions were quenched by the addition of SDS gel loading dye and resolved on SDS-PAGE. Radioactivity bound to the proteins was detected using a Fuji FLA-5000 phosphor imager.

Malonyl-CoA Decarboxylase Assay—A reaction volume of 20 μ l containing 24 μ g of protein and 34.2 μ M [2-¹⁴C]malonyl-CoA (58.40 mCi/mmol) was incubated at 30 °C for 30 min. Reactions were quenched with 6 M HCl and applied directly on silica gel 60 F₂₅₄ TLC plates (Merck). A solvent system of isopropyl alcohol/H₂O/25% aqueous ammonia (80:5:15, v/v/v) was used to resolve malonyl-CoA and acetyl-CoA. The *R_f* of these compounds was determined using authentic standards.

Structural Modeling Studies—The sequences of various CHS-like proteins were downloaded from GenBankTM (25) and structures of various enzymes of thiolase fold family were obtained from the Protein Data Bank (26). All pairwise alignments were carried out using BLAST (27) at NCBI and ClustalW program (28) from the EBI server was used for generating multiple alignments and dendrograms. The active site cavities in various structures were identified by using the program ACSITE. The 3D-PSSM server was used for threading calculations. Structural superpositions were carried out using CE server (29). Structural models were generated using the rotamer library approach of the SCWRL program and the INSIGHT II package (version 2.3.5; Biosym Technologies, San Diego) was used for energy minimization and depiction of structural models.

RESULTS

Homology-based Structure/Function Analysis of Type III PKS Proteins—The genome sequencing projects of various microbes have resulted in identification of several genes similar to type III PKSs. A number of CHS homologues are present in *S. coelicolor*, *Bacillus subtilis*, *Bacillus halodurans*, *Deinococcus radiodurans*, *Cytophaga hutchinsonii*, and *M. tuberculosis* (8). Although none of the type III PKS identified through genome sequencing efforts has been characterized, these results suggest that type III PKS is prevalent in bacteria as well. Our analysis of mycobacterial genome sequences revealed three type III PKS genes in *M. tuberculosis*. Two of these genes, *pks10* and *pks11*, are annotated putative chalcone synthases, whereas *pks18* (which shows similar sequence similarity) is surprisingly listed as a conserved hypothetical protein in the revised *M. tuberculosis* genome annotations (30). The *pks10*

² P. Arora and R. S. Gokhale, unpublished results.

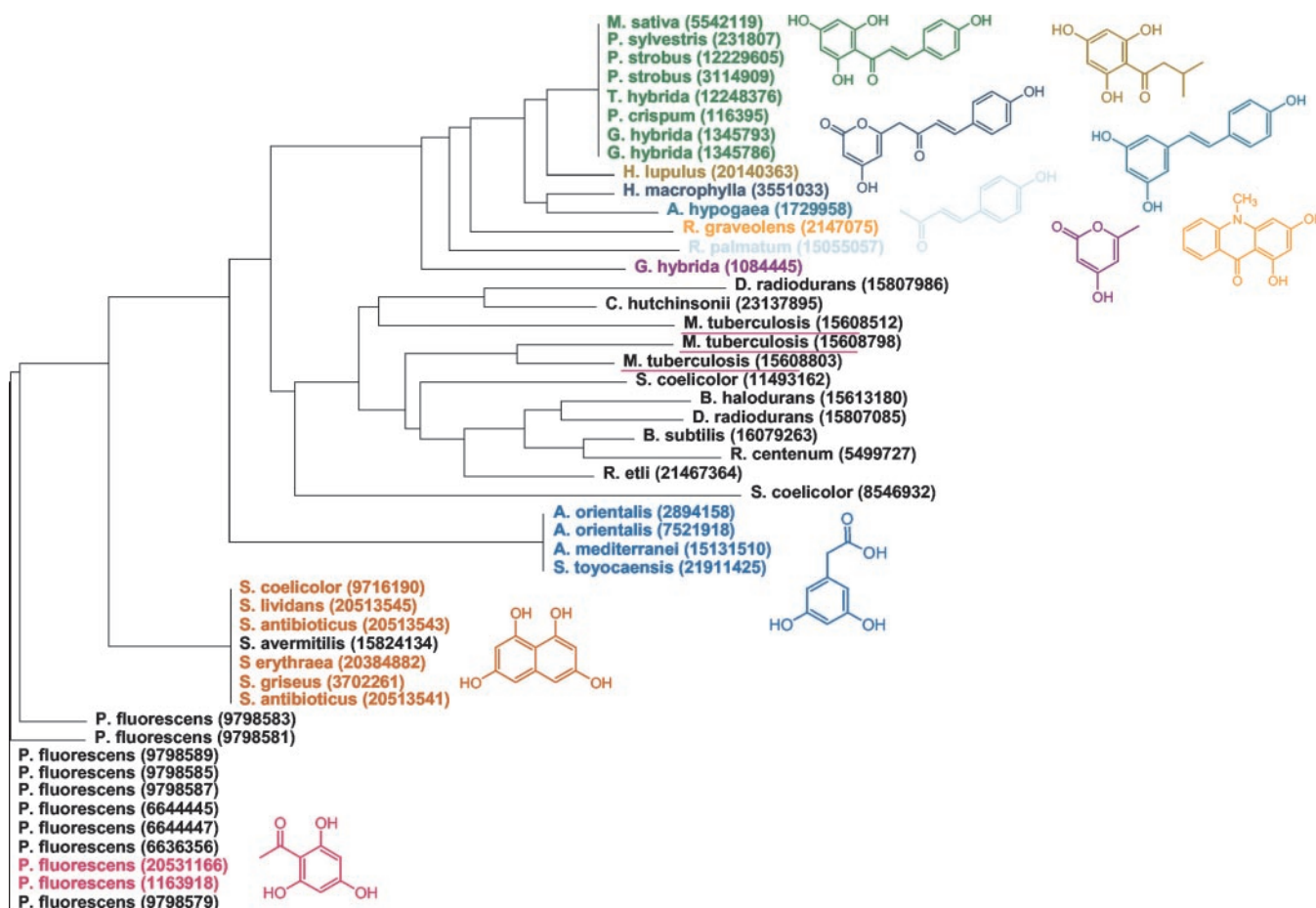


FIG. 1. **Phylogenetic analysis of plant and bacterial type III PKSs.** The dendrogram was generated for all bacterial and a few representative plant sequences using 32 residues corresponding to the alfalfa chalcone synthase (Protein Data Bank code 1CGZ) active site amino acids. Each protein and its polyketide product are drawn in the same color. The uncharacterized proteins are shown in black (gi numbers are shown in parenthesis). Mycobacterial type III PKSs have been underlined in pink.

and *pks11* genes are clustered in an unusual organization with other PKS genes that encompass 18 kb of genomic DNA. The *pks18*, on the other hand, is not flanked by any PKS-related genes. All three genes are completely conserved in *M. bovis* BCG, suggesting their functional relevance in mycobacteria.

Comparative analysis of protein sequences of plants and microbial type III PKSs have previously led to the clustering of these proteins based on species. The dendrogram based on complete protein sequences had two major branches corresponding to plants and microbial PKSs; further categorization within the plant family was not feasible (8, 17). Since protein sequence diversity between plant and bacterial type III PKSs correlates with functional differences between the two systems, it is possible that the active site residues in the catalytic pocket can facilitate more reliable clustering. Similar active site sequence analyses of acyl transferase domains from modular PKSs (31) and adenylation domain of nonribosomal peptide synthetases (32) have indicated much better correlation with their functions. Fig. 1 shows a dendrogram of all microbial and some of the plant type III PKSs, constructed by using 32 active site residues that have been identified in the three-dimensional structure of CHS. The product biosynthesized by the corresponding PKS is drawn next to its branch (proteins and their products are drawn in the same color). It can be noticed that all the plant polyketide synthases cluster together. The eight CHS proteins that have been selected randomly have identical sets of active site residues. Interestingly, the plant PKSs associated with divergent products appear at different branches in the

dendrogram, which was not the case when analysis was done with the entire protein sequence. The three distinct types of bacterial type III PKSs that have been experimentally characterized also cluster in three distinct branches. Here again we observe that active site residues are completely conserved among their family members. This trend of complete conservation of active site residues may be a manifestation of biased data set, since most of the sequences are available from the same genus. There are presently seven protein sequences in the THN synthase family, and six of them have been experimentally characterized. It can be predicted with reasonable confidence that the seventh sequence would also be involved in THN biosynthesis, which was also recently suggested by Austin and Noel (11). This protein was probably wrongly annotated as PhlD homologue during *S. avermitilis* genome sequencing (33). Several PhlD homologues from *Pseudomonas fluorescens* also possess an identical set of active site residues, indicating that they are involved in the biosynthesis of the same metabolite. The other bacterial type III proteins of unknown function are present in a separate branch in the dendrogram. Although several important catalytic and structural residues are conserved, these CHS-like enzymes from bacterial genomes have significant variation at several amino acid positions. Our sequence comparison analysis therefore could not predict putative function of mycobacterial type III PKSs. In order to characterize the function of these mycobacterial type III PKSs, we have used two distinct approaches. In the genetic-based approach, the *pks10-pks11* cluster of mycobacteria has been

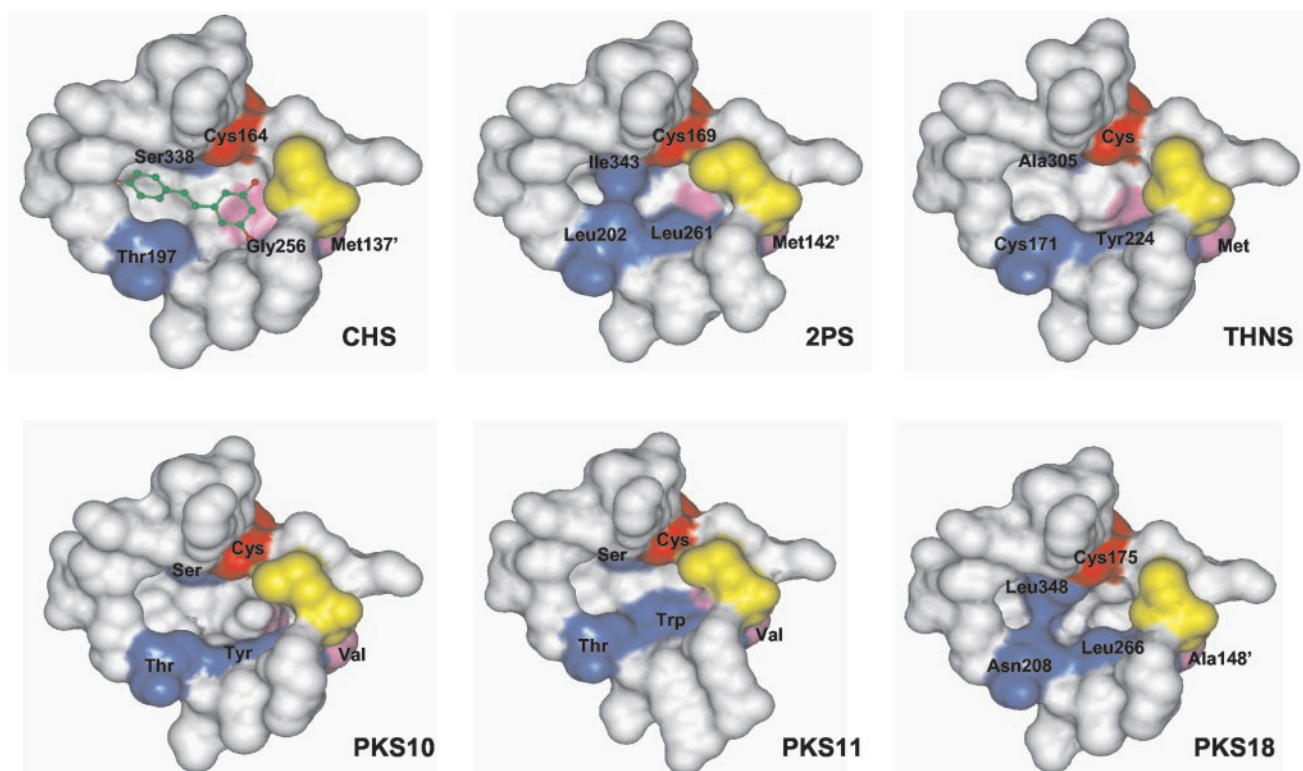


FIG. 2. Comparison of the active site cavities of mycobacterial PKS10, PKS11, and PKS18 proteins with the experimentally determined CHS initiation/elongation cavity complexed with resveratrol (1CGZ). Some of the crucial active site residues along with the catalytic Cys (in red) are shown here. Chimeric models of PKS10, PKS11, and PKS18 were generated by amino acid replacements using the SCWRL amino acid rotamer library. The 2PS and THNS chimeras were constructed for comparison. 2PS cavity resembles the crystallographically elucidated structure validating the homology-based active site predictions. Since active site cavity is buried, two residues were removed for clarity.

knocked out by using double recombination.³ The metabolic profiles for the wild type and mutant *Mycobacteria* can reveal the function of PKS genes. The results of this study would be discussed elsewhere.

In the structure-based approach, we have probed the differences in the initiation/elongation cavities of all three mycobacterial CHS homologues. The chimeric models of PKS10, PKS11, and PKS18 were constructed by replacing variable amino acids in the catalytic pocket of alfalfa CHS (Protein Data Bank code 1CGZ) by analogous residues in these proteins. For example, the PKS18 chimeric model was constructed by replacing 11 of 32 residues that constituted the initiation/elongation cavity of CHS. The amino acid replacements were carried out by the rotamer library approach using the SCWRL program (34). This comparative modeling approach is based on the premise that the overall fold as well as the active site geometry of protein would be similar. In one such successful attempt, CHS was transformed into 2PS by introducing bulkier side chains at three crucial positions (13). The reduction in cavity volume matched with the change in usage of the starter substrates, number of extensions, and final cyclization. However, similar homology models with stilbene synthases have failed to identify the structural basis for alternative cyclizations (11). The comparison of CHS and stilbene synthase three-dimensional structures has attributed these cyclization differences to subtle active site changes, apparently involving electronic rather than steric interactions. Although this example illustrates the pitfalls of homology modeling, it should be noted that the homology modeling predictions are more reliable at the gross level than for subtle rearrangements. Fig. 2 compares the active site cavities of chimeric models generated for mycobacterial type III

PKSs with that of the experimentally determined structure of CHS. Chimeric models were also generated for 2PS and THNS for comparisons. The presence of three amino acids, Leu²⁰², Leu²⁶¹, and Ile³⁴³, in 2PS considerably decreases the cavity volume due to increased steric bulk of these residues, as observed in its three-dimensional crystal structure (13). This provides support to the validity of homology-based active site predictions. In the chimeric model of THNS, these three positions correspond to Cys¹⁷¹, Tyr²²⁴, and Ala³⁰⁵. The changes of Tyr and Cys are compensatory in terms of changes of the cavity volume for THNS. In PKS10 and PKS11, Gly²⁵⁶ of CHS (corresponding to Leu²⁶¹ of 2PS) is changed to Tyr and Trp, respectively. Although there is a reduction in the volume of the cavity compared with CHS, substrate preference of these proteins is not apparent from these homology-modeled chimeric structures. In contrast, the PKS18 homology model shows remarkable constriction in the active site cavity. The steric bulk from three active site residues, Asn²⁰⁸, Leu²⁶⁶, and Leu³⁴⁸, significantly affects the active site volume. We therefore predicted that PKS18 protein might resemble 2PS and may catalyze biosynthesis of pyrones. In order to improve our predictability for PKS10 and PKS11, various CHS-related plant products were docked in the initiation/elongation cavity. Most of these docked structures resulted in steric clashes with the protein side chains and backbone, suggesting that *p*-coumaroyl-CoA-like starter units may not be a substrate for these proteins.

Characterization of Mycobacterial Type III PKSs—Three mycobacterial type III PKSs showed 40–45% similarity with both plant and bacterial CHS-like sequences. These genes were cloned from the H37Rv BAC genomic library and were then expressed in *E. coli* using the T7 expression system. Although overexpression for all three proteins, PKS10, PKS11, and PKS18, was observed, most of the protein was present in in-

³ Yogyata and R. S. Gokhale, unpublished results.

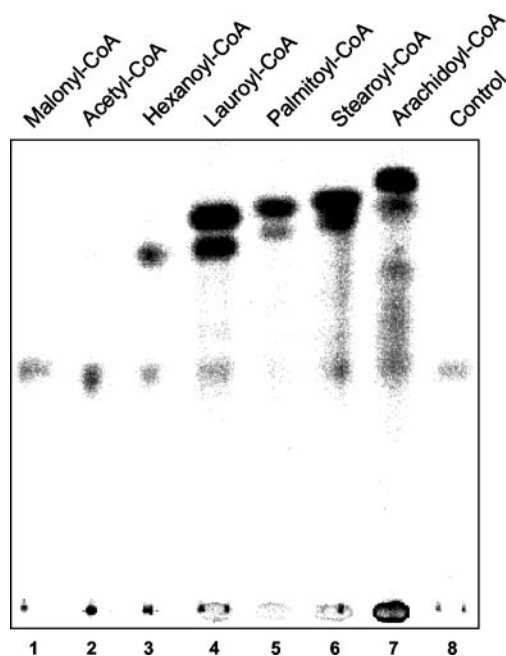


FIG. 3. TLC based analysis of radiolabeled products of PKS18. Reactions were carried out for 10 min using $[2-^{14}\text{C}]$ malonyl-CoA as an extender unit along with different starter substrates. The starter substrates used in the assay are indicated above each lane.

clusion bodies. The PKS11 protein was purified from inclusion bodies and was used to raise polyclonal antibodies. The expression of PKS18 protein was modulated to soluble form by inducing cultures with a low isopropyl-1-thio- β -D-galactopyranoside concentration at 22 °C. PKS11 protein could be purified in the soluble form by growing cultures without induction at 18 °C for 24 h. The antiserum generated from PKS11 protein specifically recognized the recombinant hexahistidine-tagged 40-kDa protein, and no other cross-reacting bands were observed. PKS11 antiserum was also able to weakly recognize PKS18 protein, which may be a result of homology between these two proteins. Despite our best efforts, PKS10 protein could not be expressed in soluble form. PKS18 and PKS11 proteins were purified by using Ni^{2+} -nitrilotriacetic acid affinity chromatography and were then purified to homogeneity using anion exchange chromatography. The identities of proteins were confirmed by N-terminal protein sequencing and mass spectrometry.

Although our modeling studies predicted that mycobacterial proteins would not be able to use plant-specific CoA-esters, we investigated the function of these proteins with several different CoA-esters due to their sequence similarity with plant PKSs. Several plant-specific CoA-esters were incubated with PKS18 and PKS11 proteins along with radiolabeled malonyl-CoA. Analysis of the reaction mixture by TLC showed one faint radioactive band for PKS18 reactions with an R_F value different from that of the expected plant products. Further examinations with longer time incubations clearly showed that PKS18 protein could synthesize this product from malonyl-CoA itself. Since 2PS uses acetyl-CoA as a starter substrate, we incubated PKS18 protein with acetyl-CoA and $[2-^{14}\text{C}]$ malonyl-CoA. Under the assay condition used in this study, there was no significant change in the product formation. Since PKS11 protein did not show any appreciable activity under these conditions, we proceeded to identify the PKS18 reaction product.

Remarkable Substrate Specificity for PKS18 Protein—Type III PKSs such as CHS and THNS have been shown to have rather promiscuous starter unit specificity (20, 35). However, in most cases the incorporation of suboptimal moiety has often adversely affected the catalytic activity, both in terms of for-

mation of end products and in the amount of products. Since PKS18 protein did not exhibit any activity with plant-specific acyl-CoA substrates, we used several medium- and long-chain fatty acid CoA-thioesters as starter substrates. Fig. 3 shows radio-TLC of a 10-min assay performed using several starter units and $[2-^{14}\text{C}]$ malonyl-CoA as extender substrate with PKS18 protein. It can be noticed that several significant bands appear on the TLC plate, particularly for the long-chain acyl-CoA esters. There is a systematic increase in the R_F value of the products with increasing length of the fatty acyl chain, clearly indicating that PKS18 protein is able to use medium- and long-chain aliphatic acyl-CoAs as substrates. Whereas incubation of PKS18 protein with hexanoyl-CoA shows one major product, incubation with long-chain acyl-CoAs (C_{12} to C_{20}) produces two significant products. Under the assay conditions used in this study, no activity was detected for small-chain, aromatic, and branched-chain acyl-CoA analogues. These studies clearly suggested that PKS18 protein can extend long-chain acyl-CoA substrates by using malonyl-CoA as an extender molecule to synthesize polyketide products.

Identification of Novel Reaction Products—Polyketide products biosynthesized from purified enzymes were separated on reverse phase HPLC and characterized by nanospray ESI MS. Since hexanoyl-CoA-primed reaction showed a reasonable amount of product on radio-TLC assays, we proceeded to characterize this product. The HPLC profile showed a single major peak *a* at 280 nm (Fig. 4A). Incubation of PKS18 with $[2-^{14}\text{C}]$ malonyl-CoA along with hexanoyl-CoA as starter gave one major radioactive product that co-migrated with the peak *a*. This product was then subjected to nanospray MS analysis, and a molecular ion peak (M-H^-) at m/z 181.08 was observed. Based on our understanding of polyketide biosynthesis, this molecular mass is possible for a cyclized triketide synthesized from one molecule of hexanoyl-CoA and two molecules of malonyl-CoA. In order to characterize products from lauroyl-CoA, a ternary solvent system was developed, where the two main products could be separated as shown in Fig. 4B. Nanospray mass spectrometric analysis showed molecular ion peaks (M-H^-) at m/z 265.17 for *c* and m/z 307.20 for *b*, which could be expected from triketide and tetraketide products, respectively, of lauroyl-CoA reactions with extensions from malonyl-CoA.

Tetraketide products of type III PKSs are known to cyclize in three distinct ways, where different polyketide products are formed from aldol condensation, Claisen condensation, and lactonization of the reaction intermediate (10, 11). The structural elucidation of PKS18 reaction products was completed by using tandem mass spectrometry. Fig. 4C shows the fragmentation pattern for the molecular ion (M-H^-) m/z 307.20 (obtained for *b*), which yielded a fragment of m/z 125.03 corresponding to $(\text{C}_6\text{H}_5\text{O}_3)^-$ and m/z 263.21 corresponding to $(\text{M-CO}_2\text{-H})^-$ ion. This fragmentation pattern has been previously reported for the tetraketide α -pyrones biosynthesized by RppA from *S. griseus* (20). The molecular ion peak (M-H^-) at m/z 265.17 for *c* yielded a $(\text{M-CO}_2\text{-H})^-$ peak at m/z 221.18, which indicated the presence of an α -pyrone ring (Fig. 4D). Products biosynthesized by PKS18 protein with other acyl chains were similarly characterized by nanospray MS and determined to be triketide and tetraketide α -pyrones.

The starter unit specificity of PKS18 was investigated by performing steady state kinetic analysis using C_2 - C_{20} acyl-CoA substrates. PKS18 protein obeyed saturation kinetics in response to the increasing concentration of acyl-CoA substrates, which is summarized in Table I. It can be observed that the k_{cat}/K_m values for the biosynthesis of triketide pyrones of long-chain acyl-CoA substrates are severalfold higher than that observed for small-chain substrates. Specificity of PKS18 pro-

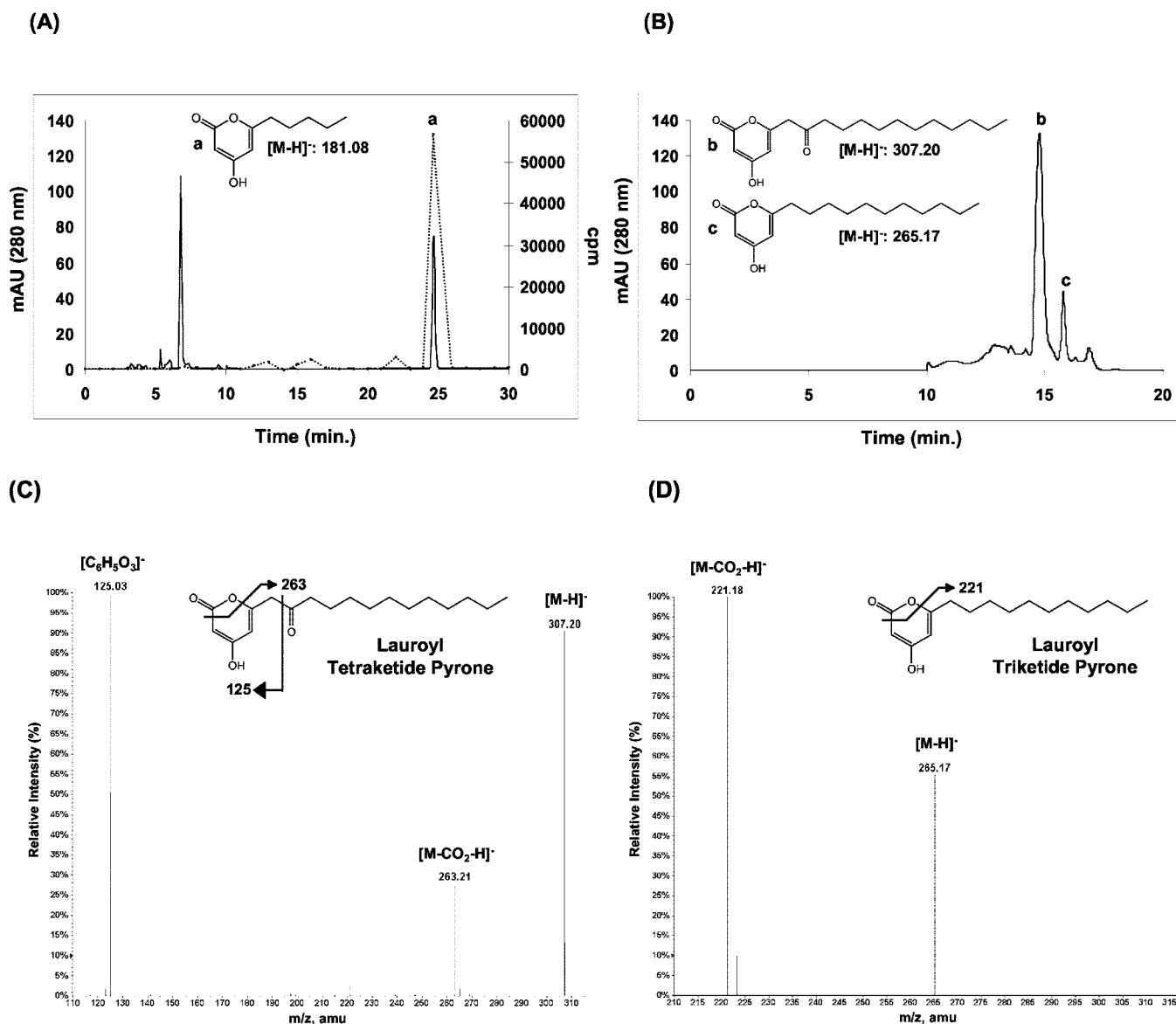


FIG. 4. **Identification of PKS18 products.** HPLC profile was monitored at 280 nm. A shows the superimposition of radioactive measurements (dotted line) on the 280-nm measurement (continuous line) for hexanoyl-CoA-primed reaction (major product peak is marked a). B shows the HPLC chromatogram for products formed from lauroyl-CoA-primed reaction (marked b and c) separated using a ternary gradient system. The molecular mass obtained for each peak by nanospray mass spectrometric analysis is indicated below the structures for the expected products. C and D illustrate the fragmentation pattern observed by tandem mass spectrometry for lauroyl-CoA-primed products tetraketide pyrone (M-H)⁻ 307.20 (C) and triketide pyrone (M-H)⁻ 265.17 (D). Fragment ions are indicated above the corresponding peaks. The structure of the molecule with the expected fragmentation profile is drawn beside its MS/MS spectra.

TABLE I
Steady state kinetic parameters for
PKS18 protein with different starter units

Results are means ($n = 3$) with S.E. values less than 15%.

	k_{cat}	K_m	k_{cat}/K_m
	min^{-1}	μM	$\text{s}^{-1} \text{M}^{-1}$
Acetyl-CoA	0.045	49.2	15.2
Malonyl-CoA	0.017	58.4	4.8
Hexanoyl-CoA	0.33	19.2	286.4
Lauroyl-CoA	13.94	4.2	55317.4
Palmitoyl-CoA	15.83	5.4	48858.2
Arachidoyl-CoA	15.47	6.1	42267.7

teins for long-chain acyl-CoA starters, as reflected in the K_m values, is similar for C_{12} to C_{20} acyl chains. Aliphatic CoA substrates longer than C_{20} were not investigated in this study. This unusual specificity toward long-chain aliphatic fatty acids is unparalleled in the CHS family of condensing enzymes.

Previous studies have shown octanoyl-CoA to be the largest starter unit extended by CHS and THNS (20, 35).

Based on this unusual ability of CHS-like proteins to use long-chain substrates, reactions were carried out with PKS11 protein using medium- and long-chain acyl-CoA units. Fig. 5 shows the radio-TLC of reactions of PKS11 and PKS18 proteins. As can be visualized, various different products were obtained when the proteins were incubated for 1 h with these acyl-CoA starters and [2- ^{14}C]malonyl-CoA. Like PKS18, PKS11 readily accepted these starter units and formed α -pyrone products, as confirmed by nano-ESI-MS (data not shown).

Structural Analysis of Long-chain Specificity for Mycobacterial Proteins—The analysis of the structural models of PKS10, PKS11, and PKS18 proteins clearly indicated that the putative CHS active site cannot accommodate pyrone rings with long aliphatic side groups, suggesting the possibility of an altered binding pocket in these type III PKS proteins from *M. tuberculosis*. It is interesting to note that some members belonging

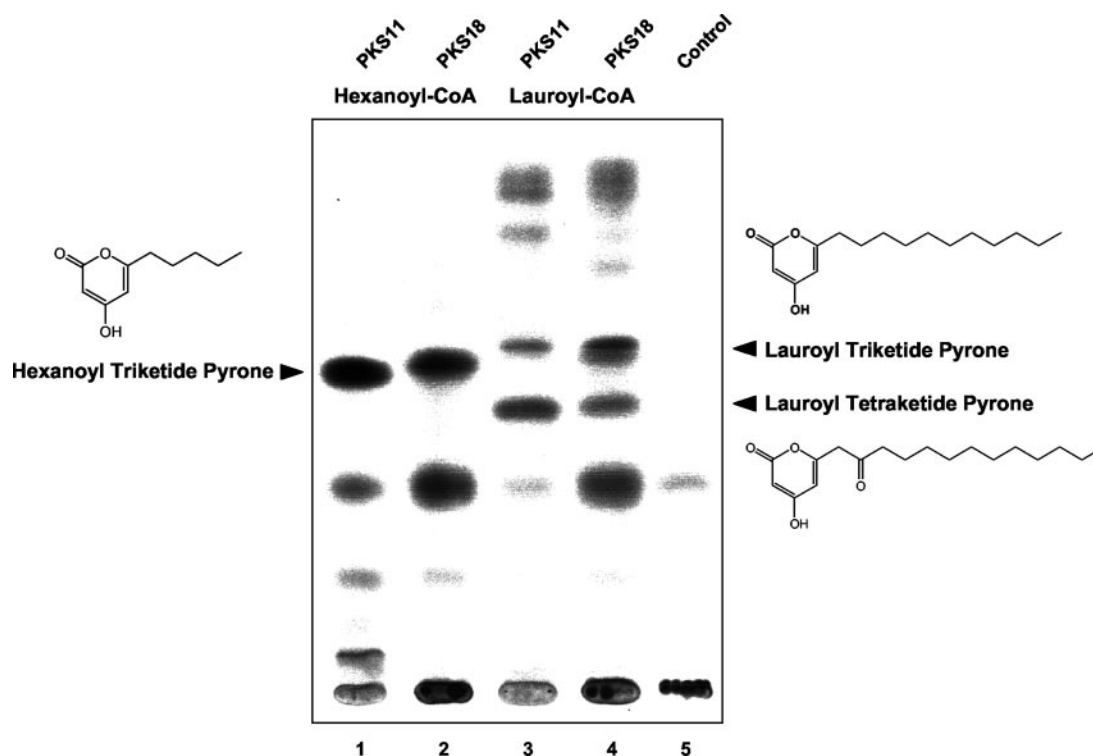


FIG. 5. **Comparative TLC of radiolabeled products of PKS11 and PKS18 proteins.** Lanes 1 and 2, hexanoyl-CoA-primed reactions; lanes 3 and 4, lauroyl-CoA-primed reactions. $[2-^{14}\text{C}]$ malonyl-CoA was used as an extender unit in all of the reactions. Chemical structures of the major products are indicated *beside* the radioactive bands.

to the thiolase fold family possess the capability to use long aliphatic chains as substrates (15, 36, 37). The analysis of the three-dimensional crystal structure of KAS III enzyme (1HZP) from *M. tuberculosis* (referred to as mtFabH) has revealed the mechanism by which it can catalyze the decarboxylative condensation of bulky myristoyl-CoA with malonyl-acyl carrier protein. Apart from the overall thiolase $\alpha\beta\alpha\beta$ fold, KAS III and CHS-like proteins also share the Cys-His-Asn catalytic triad. The structural superposition of 1HZP on 1CGZ, as shown in Fig. 6A, gives a root mean square deviation of 2.9 Å at a Z score of 6.8. The $\alpha\beta\alpha\beta$ core of the two structures superposes quite well, and the differences are restricted to the loop regions. The catalytic triad also superposes with a root mean square deviation of 0.54 Å. It is interesting to note from Fig. 6A that the two active site cavities (*red dots* for CHS and *green dots* for KAS III) are oriented differently with respect to the CoA-binding tunnel. Whereas the KAS III orients the starter units to the right and toward the dimer interface, CHS orients them toward the left. Unlike CHS, mtFabH enzyme can elongate long-chain aliphatic chains and has been shown to be specific for C_8 to C_{20} aliphatic substrates (36). The structural basis of the substrate specificity for six priming ketosynthases was investigated recently (36, 38). These studies indicated that the selectivity and primer unit size are primarily controlled by the side chains of three residues from the loops present close to the dimer interface. In mtFabH, all of the side chains of the cavity-limiting residues (Asn⁹⁰, Thr⁹⁶, and Glu²⁰⁰) are either short or point outwards, resulting in a long tunnel. A recent review article has discussed the structural and mechanistic implications of various members of the thiolase family of enzymes (11).

Since mycobacterial type III PKSs biosynthesize products that have characteristics of both CHS and KAS III proteins, we set out to test the hypothesis that these enzymes would have evolved by combining the two catalytic centers of CHS and KAS III. PKS18 displays a homology of 42% with alfalfa CHS over the entire protein sequence. As described earlier, a compara-

tive model of PKS18 based on 1CGZ structure was constructed, which demonstrated a constricted CHS initiation/elongation cavity. In the CHS-related structural family of condensing enzymes, several proteins such as mtFabH use long-chain substrates. The attempt to model long-chain aliphatic substrate in the putative active site of PKS18 was tested by using 1HZP as template. Sequence similarity of PKS18 to mtFabH is restricted to a small region consisting of amino acids 111–239. The C-terminal domain does not show any homology with PKS18 in sequence comparisons. Since threading methods are capable of detecting structural similarity at low sequence identities, PKS18 was threaded using the 3D-PSSM server (39). The threading results indicated the plant CHS structure (1CGZ) to be the highest scoring fold, although the structure of mtFabH (1HZP) was also suggested to be a possible template. Upon detailed inspection of the alignments given by the threading servers, it was found that 3D-PSSM gives a lower score for 1HZP due to poor alignment of the 90 N-terminal residues. Deletion of 90 N terminus residues improved the threading score (3D-PSSM E-value from 1.17 to 0.037) and thus provided an optimal alignment of PKS18 sequence on the structural fold of mtFabH. However, plant CHS structure was still the highest scoring fold. We decided to use 1HZP as a template for homology modeling, since this protein is known to bind long-chain aliphatic substrates. Using the program ACSITE (40), an elongated cavity was identified in mtFabH along the dimer interface, indicating that elongated substrate can be accommodated in the active site of the 1HZP fold. Since the mtFabH coordinates complexed with substrate are not available in the Protein Data Bank, lauric acid was modeled in the active site of PKS18 using the information from the crystal structure of lauric acid-KAS I complex (1EK4) (37). Although KAS I proteins have no detectable sequence homology to the CHS family of condensing proteins, crystal structure analysis has revealed a conserved thiolase fold (41). Superposition of the substrate-bound 1EK4 on 1HZP indicated that the transformed substrate fits in the

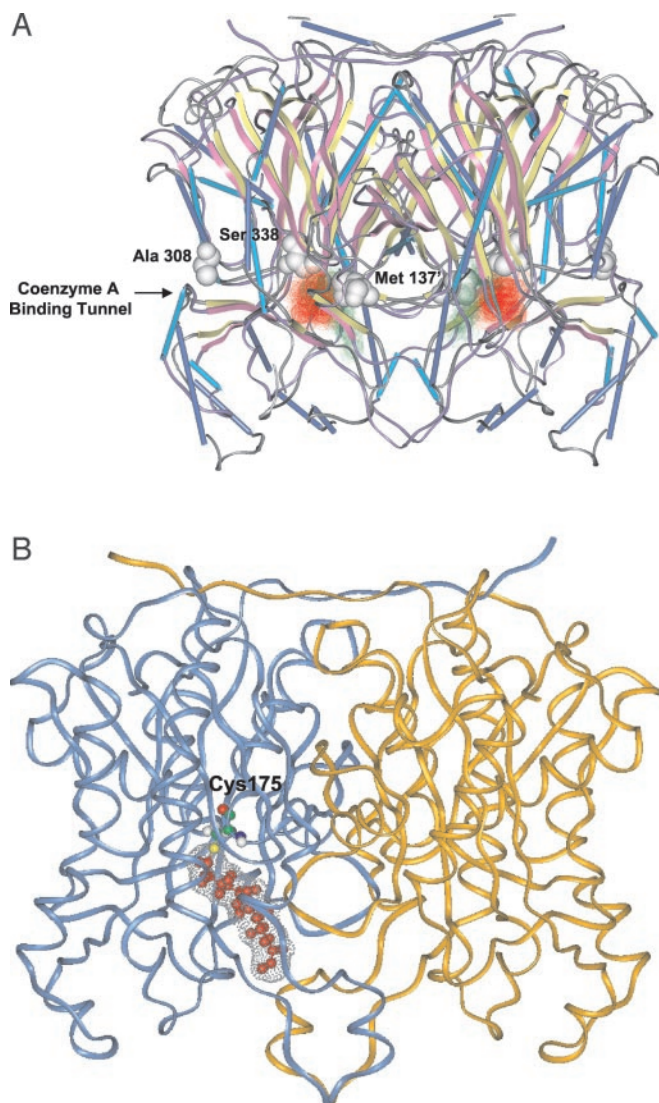


FIG. 6. A, the structural superposition of mtFabH (1HZIP) on alfalfa chalcone synthase (1CGZ). The two active site cavities predicted by ACSITE are marked on the structure (red dots for CHS and green dots for mtFabH). It can be observed that the two active site cavities have overlapping regions. The three residues that are mutated in the present study are shown on the CHS backbone. B, chimeric model of PKS18 (based on 1HZIP) in complex with lauric acid (shown in red). The amino acids lining the acyl binding cavity are replaced with PKS18 residues. Ligand is acylated at the catalytic cysteine residue (shown in a ball and stick model).

cavity of 1HZIP. These transformed coordinates were then used as starting orientation in the active site of 1HZIP fold. The amino acids of 1HZIP that were in contact with the lauric acid were replaced with equivalent residues of PKS18 based on threading alignment. The carboxyl carbon of the lauric acid was covalently attached to the thiol of active site cysteine, and the complex was energy-minimized using the CVFF force field (42). The energy-minimized structural model of PKS18 complexed with covalently tethered lauric acid is shown in Fig. 6B. The residues in contact with the substrate form a hydrophobic acyl-binding channel, and no steric clashes were observed. These observations offer a structural model for understanding the unusual substrate specificity of PKS18 protein.

Probing the Active Site of PKS18 Protein by Site-directed Mutagenesis—Structural analysis has provided a testable model to probe into the functional relevance of catalytic and structural residues of mycobacterial PKSs. Our modeling stud-

ies in conjunction with the analysis of the substrate specificity for priming ketosynthases suggested that amino acid residue Ala¹⁴⁸ from PKS18 may be a crucial residue. CHS contains a Met¹³⁷ residue at the equivalent position, which is contributed from the second monomer and lines the active site cavity. It can be observed from Fig. 6A that the Met¹³⁷ amino acid restricts the extension of the CHS cavity (red dots) further into the acyl-chain binding pocket (green dots). The Ala¹⁴⁸ residue in PKS18 was mutated to bulkier side chains of threonine, methionine, and phenylalanine to restrict the extension of the cavity. Similar strategy was previously successfully employed, where the KAS II specificity was modulated to favorably accept shorter acyl-chain length substrates (43). Mutation of the Ala¹⁴⁸ residue of PKS18 protein resulted in expression of these three proteins in inclusion bodies. Despite our best efforts, these mutant proteins could not be recovered in the soluble form. Although we could not directly probe the functional significance of this residue, it can be inferred that Ala¹⁴⁸ may be a structurally important residue.

Another residue predicted to be involved in the formation of active site cavity, Leu³⁴⁸ of PKS18, was investigated by performing mutagenesis studies. Previous mutagenesis studies of the equivalent residue in CHS (Ser³³⁸) (Fig. 6A) have implicated its role in starter unit selectivity and determination of the chain length of the growing polyketide (13). Mutation of the analogous Ala³⁰⁵ residue to a bulkier isoleucine residue in THNS protein abolished the synthesis of tetraketide pyrones, and only triketide pyrones were observed (18). In PKS18, Leu³⁴⁸ was mutated to a smaller serine residue. Biochemical characterization carried out with [2-¹⁴C]malonyl-CoA and the various starter substrates resulted in a relative increase in the synthesis of the tetraketide products. Fig. 7 shows the nano-spray tandem mass spectrometric analysis of the tetraketide pyrone of the hexanoyl-CoA-primed reaction of L348S mutant. The wild type PKS18 protein does not synthesize this product. An increase in the tetraketide product for the L348S mutant suggested that the Leu³⁴⁸ residue might be located in the elongation/cyclization cavity of PKS18 protein, consistent with our homology model.

CHS condensing enzymes contain three conserved basic amino acid residues that are located at the outer rim of the CoA-binding tunnel for electrostatic interactions with CoA phosphates (12). Bacterial type III PKSs contain an extra conserved basic amino acid residue (Lys³¹⁸ in PKS18) in proximity to the predicted CoA-binding tunnel (11). Plant CHSs have a conserved alanine residue at this position, as shown in Fig. 6A. The equivalent residue is arginine in several KAS III enzymes, where the guanidino side chain facilitates the binding of acyl carrier protein (46). It has been speculated that some bacterial type III PKSs may also have specificity for ACP-thioester substrates (11). In an attempt to elucidate the role of this conserved residue, Lys³¹⁸ of PKS18 was mutated to alanine. The K318A mutant was unable to synthesize any polyketide products in our standard PKS18 assay. Biophysical analysis of the mutant protein showed no change in the overall secondary and tertiary structure as compared with wild type PKS18 protein. The inactivity of the mutant protein was probed by carrying out binding assays with radiolabeled starters and extender units. Fig. 8 shows the autoradiogram of the SDS-PAGE gel containing wild type PKS18 protein and K318A mutant protein incubated independently with [2-¹⁴C]lauroyl-CoA and [2-¹⁴C]malonyl-CoA. It can be observed that the wild type PKS18 protein can be radiolabeled with both malonyl-CoA and lauroyl-CoA (Fig. 8A). In contrast, the K318A mutant retains the ability to accept the starter lauroyl-CoA but could not bind to the extender malonyl-CoA (Fig. 8B). Further mechanistic investiga-

tion suggested that this mutant protein was severely impaired in malonyl-CoA decarboxylation (Fig. 8C).

DISCUSSION

The identification of diverse catalytic activity in mycobacterial type III PKSs provides an intriguing example of metabolic divergence in CHS-like proteins. Interestingly, the plant prototype CHS protein has not yet been characterized in bacteria. In this article, we discuss the characterization of two novel type III polyketide synthases from *M. tuberculosis*. PKS11 and PKS18 enzymes have remarkable specificity to use long-chain aliphatic-CoA analogues (C_{12} to C_{20}) as starter units to biosynthesize α -pyrones. The reaction catalyzed by PKS18 protein is summarized in Fig. 9. The mechanism of PKS18 protein to synthesize pyrones was examined to explore whether the pyrone formation was a result of hydrolysis followed by nonenzymatic cyclization or if this was produced by intramolecular lactonization. The pyrone products of PKS18 were extractable with organic solvent even without acidification of the reaction

mixture, suggesting that the acidic conditions were not required for pyrone formation. Although acidification does promote spontaneous lactonization, it could also stabilize and promote the extraction into organic solvents of an enzymatically formed pyrone. Moreover, pyrone products with long aliphatic chains would enhance extractability of these products in organic solvents and can complicate the interpretation of results. The nanospray ESI-MS analysis did not detect any hemiacetal or carboxylic acid compounds, indicating that the release of final product proceeded by intramolecular lactonization. A similar mechanism of ring closure has been previously proposed for RppA protein (20).

Pyrene synthases have been previously characterized in plants (47), although such proteins have not been identified from the microbial world. Several plant and bacterial type III PKSs are also known to synthesize lactones as derailment products upon incorporation of a suboptimal starter moiety (20, 35). However, this often results in relatively less activity. Whereas the broad starter molecule specificity of PKS18 and PKS11 enzymes obscures the assessment of *in vivo* biological function, the general preference for the type of starter units (preference of aliphatic or aromatic, smaller or larger, straight-chain or branched) provides clues about the likely physiological substrate. This fact combined with the *in vivo* availability of CoA-thioesters should determine the biological function. The notable preference of mycobacterial type III PKSs for long-chain acyl-CoAs makes it tempting to speculate that these are the preferred substrates *in vivo*, although such pyrone-containing long-chain derived metabolites have not been isolated from *Mycobacterium*. It should be noted that *M. tuberculosis* contains a rich pool of a remarkable array of long-chain and very long-chain fatty acids (48). Moreover, the *M. tuberculosis* genome also includes 36 copies of *fadD* homologues that are typically involved in the biosynthesis of acyl-CoA substrates (1).

Comparative modeling studies suggest that the unusual capability to accept long-chain acyl-CoA substrates might have evolved by combining the catalytic machinery of CHS and KAS proteins. Apart from a constricted CHS active site initiation/elongation cavity, an additional acyl group binding pocket (typically observed in KAS-related proteins) has been identified along the dimer interface in these mycobacterial proteins. Since homology modeling studies at low sequence identity cannot give unambiguous results, three-dimensional crystal struc-

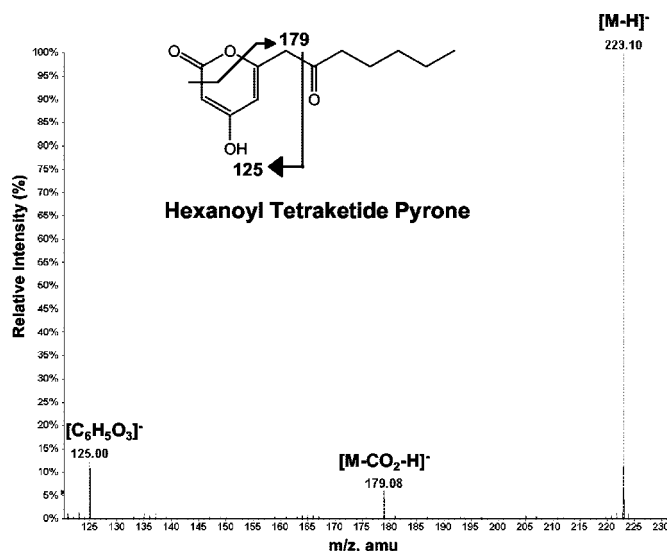


FIG. 7. Tandem mass spectrometric analysis of hexanoyl tetraketide pyrone synthesized by L348S mutant. The molecular ion peak ($M-H$)⁻ at m/z 223 fragments into m/z 125 corresponding to $(C_6H_5O_3)^-$ and m/z 179 corresponding to $(M-CO_2-H)^-$ ion.

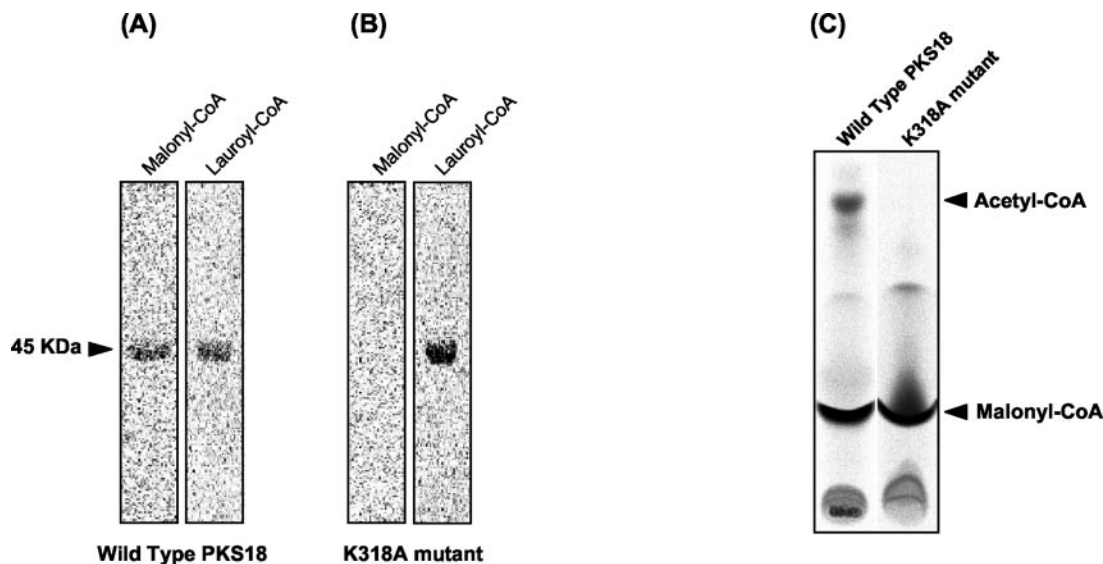


FIG. 8. Biochemical analysis of PKS18 mutant. The autoradiographs of substrate binding assays are shown for wild type PKS18 (A) and K318A mutant (B). The substrate used for labeling is indicated above the lane. C shows the TLC-based analysis of malonyl-CoA decarboxylation assay carried out for wild type and mutant protein (K318A).

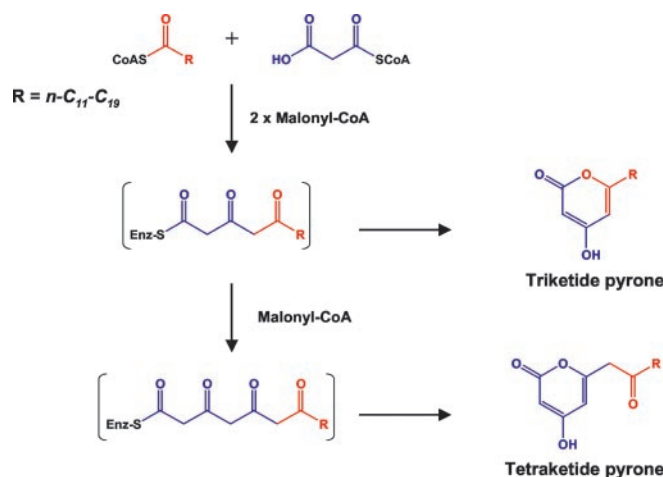


FIG. 9. Reactions catalyzed by mycobacterial type III PKSs. The starter molecule and its corresponding component in the polyketide metabolite are shown in red. The extender unit is indicated in blue.

ture analysis of PKS18 protein should provide a clear picture. Based on our structural models, we have investigated the importance of three catalytically crucial amino acid residues of PKS18 protein. The mutation of Ala¹⁴⁸ to threonine, methionine, and phenylalanine resulted in mutant proteins that expressed persistently in inclusion bodies. Previous structural analysis of priming ketosynthases (36, 38) and our modeling studies had suggested that this residue may be important in dictating long-chain substrate specificity. It is possible that the mutagenesis to a bulkier residue has destabilized the hydrophobic cavity, resulting in unstable protein. The thermodynamic and structural consequences of mutations in the hydrophobic cavity are often difficult to rationalize (44, 45). The CHS elongation/cyclization cavity was investigated by mutating Leu³⁴⁸ to a smaller Ser residue. The characteristics of this L348S mutant protein showed increased biosynthesis of triketide pyrone product. This catalytic profile corroborated with other CHS-like proteins, where the presence of a smaller amino acid residue at the analogous position enhanced the number of iterations (13, 18). The K318A mutant showed rather interesting characteristics, where the mutant protein could load the monocarboxylic starter units but lost the capability of using dicarboxylic malonyl-CoA as a substrate. The K318A mutant protein was incapable of decarboxylating malonyl-CoA. Similar studies have previously been reported for a mechanistically parallel lysine residue from *E. coli* KAS I, where the K328A mutant protein could not decarboxylate malonyl-CoA (49). We predict that this lysine residue may be required very early in the catalysis, probably acting as a docking station for dicarboxylic acid extender units. Although malonyl-CoA is an extender in almost all plant and microbial type III PKSs, it is possible that there are variations in the binding of extender substrate that might influence the catalytic steps of reaction intermediates. The presence of an additional positively charged residue adjacent to the outer rim of the CoA binding tunnel in 2PS structure has indicated the possibility of variations in CoA binding that are permitted within the type III PKS superfamily (11, 13).

In conclusion, we have characterized two novel type III PKS proteins that catalyze the formation of α -pyrones using long-chain acyl-CoA substrates. Comparison of the structural and mechanistic features of thiolase fold enzymes have suggested that type III PKS proteins might have emerged by the gain of function from structurally homologous KAS III proteins. The functional equivalence of mycobacterial type III PKSs to both

CHS and KAS III proteins presents fascinating insights into the complexity of the evolutionary process.

Acknowledgment—We thank Dr. Sandip K. Basu for helpful discussions.

REFERENCES

- Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., III, Teklaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Barrell, B. G. *et al.* (1998) *Nature* **393**, 537–544
- Bentley, S. D., Chater, K. F., Cerdeno-Tarraga, A. M., Challis, G. L., Thomson, N. R., James, K. D., Harris, D. E., Quail, M. A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C. W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C. H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabinowitsch, E., Rajandream, M. A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B. G., Parkhill, J., and Hopwood, D. A. (2002) *Nature* **417**, 141–147
- Kolattukudy, P. E., Fernandes, N. D., Azad, A. K., Fitzmaurice, A. M., and Sirakova, T. D. (1997) *Mol. Microbiol.* **24**, 263–270
- O'Hagan, D. (1993) *Nat. Prod. Rep.* **10**, 593–624
- Hopwood, D. A. (1997) *Chem. Rev.* **97**, 2465–2498
- Khosla, C., Gokhale, R. S., Jacobsen, J. R., and Cane, D. E. (1999) *Annu. Rev. Biochem.* **68**, 219–253
- Katz, L., and Donadio, S. (1993) *Annu. Rev. Microbiol.* **47**, 875–912
- Moore, B. S., and Hopke, J. N. (2001) *ChemBioChem* **2**, 35–38
- Gokhale, R. S., and Tuteja, D. (2001) in *Biotechnology* (Rehm, H.-J., ed) Vol. 10, 2nd Ed., pp. 341–372, Wiley-VCH, Weinheim, Germany
- Schroder, J. (2000) in *Evolution of Metabolic Pathways* (Romeo, J. T., Ibrahim, R. K., Varin, L., and De Luca, V., ed) Vol. 34, pp. 55–89, Pergamon Press, Amsterdam
- Austin, M. B., and Noel, J. P. (2003) *Nat. Prod. Rep.* **20**, 79–110
- Ferrer, J. L., Jez, J. M., Bowman, M. E., Dixon, R. A., and Noel, J. P. (1999) *Nat. Struct. Biol.* **6**, 775–784
- Jez, J. M., Austin, M. B., Ferrer, J., Bowman, M. E., Schroder, J., and Noel, J. P. (2000) *Chem. Biol.* **7**, 919–930
- Heath, R. J., and Rock, C. O. (2002) *Nat. Prod. Rep.* **19**, 581–596
- Moche, M., Dehesh, K., Edwards, P., and Lindqvist, Y. (2001) *J. Mol. Biol.* **305**, 491–503
- Funa, N., Ohnishi, Y., Fujii, I., Shibuya, M., Ebizuka, Y., and Horinouchi, S. (1999) *Nature* **400**, 897–899
- Cortes, J., Velasco, J., Foster, G., Blackaby, A. P., Rudd, B. A., and Wilkinson, B. (2002) *Mol. Microbiol.* **44**, 1213–1224
- Funa, N., Ohnishi, Y., Ebizuka, Y., and Horinouchi, S. (2002) *Biochem. J.* **367**, 781–789
- Moore, B. S., Hertweck, C., Hopke, J. N., Izumikawa, M., Kalaitzis, J. A., Nilsen, G., O'Hare, T., Piel, J., Shipley, P. R., Xiang, L., Austin, M. B., and Noel, J. P. (2002) *J. Nat. Prod.* **65**, 1956–1962
- Funa, N., Ohnishi, Y., Ebizuka, Y., and Horinouchi, S. (2002) *J. Biol. Chem.* **277**, 4628–4635
- Pfeifer, V., Nicholson, G. J., Ries, J., Recktenwald, J., Schefer, A. B., Shawky, R. M., Schroder, J., Wohlleben, W., and Pelzer, S. (2001) *J. Biol. Chem.* **276**, 38370–38377
- Chen, H., Tseng, C. C., Hubbard, B. K., and Walsh, C. T. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 14901–14906
- Li, T. L., Choroba, O. W., Hong, H., Williams, D. H., and Spencer, J. B. (2001) *Chem. Commun. (Camb.)* **20**, 2156–2157
- Bangera, M. G., and Thomasow, L. S. (1999) *J. Bacteriol.* **181**, 3155–3163
- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Rapp, B. A., and Wheeler, D. L. (2002) *Nucleic Acids Res.* **30**, 17–20
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res.* **28**, 235–242
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22**, 4673–4680
- Shindyalov, I. N., and Bourne, P. E. (1998) *Protein Eng.* **11**, 739–747
- Camus, J. C., Pryor, M. J., Medigue, C., and Cole, S. T. (2002) *Microbiology* **148**, 2967–2973
- Yadav, G., Gokhale, R. S., and Mohanty, D. (2003) *J. Mol. Biol.* **328**, 335–363
- Challis, G. L., Ravel, J., and Townsend, C. A. (2000) *Chem. Biol.* **7**, 211–224
- Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., Kikuchi, H., Shiba, T., Sakaki, Y., and Hattori, M. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98**, 12215–12220
- Bower, M. J., Cohen, F. E., and Dunbrack, R. L., Jr. (1997) *J. Mol. Biol.* **267**, 1268–1282
- Jez, J. M., Bowman, M. E., and Noel, J. P. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99**, 5319–5324
- Scarsdale, J. N., Kazanina, G., He, X., Reynolds, K. A., and Wright, H. T. (2001) *J. Biol. Chem.* **276**, 20516–20522
- Olsen, J. G., Kadziola, A., von Wettstein-Knowles, P., Siggaard-Andersen, M., and Larsen, S. (2001) *Structure* **9**, 233–243
- Pan, H., Tsai, S., Meadows, E. S., Miercke, L. J., Keatinge-Clay, A. T., O'Connell, J., Khosla, C., and Stroud, R. M. (2002) *Structure* **10**, 1559–1568
- Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. (2000) *J. Mol. Biol.* **299**, 499–520
- Choudhury, D. (2001) *Functional Implications of Macromolecular Recognition:*

- Assembly of Adhesive Pili and Enzyme Substrate Interactions*. Ph.D. thesis, Swedish University of Agricultural Sciences, Uppsala
41. Olsen, J. G., Kadziola, A., von Wettstein-Knowles, P., Siggaard-Andersen, M., Lindquist, Y., and Larsen, S. (1999) *FEBS Lett.* **460**, 46–52
42. Dauber-Osguthorpe, P., Roberts, V. A., Osguthorpe, D. J., Wolff, J., Genest, M., and Hagler, A. T. (1988) *Proteins* **4**, 31–47
43. Val, D., Banu, G., Seshadri, K., Lindqvist, Y., and Dehesh, K. (2000) *Struct. Fold Des.* **8**, 565–566
44. Buckle, A. M., Cramer, P., and Fersht, A. R. (1996) *Biochemistry* **35**, 4298–4305
45. Ratnaparkhi, G. S., and Varadarajan, R. (2000) *Biochemistry* **39**, 12365–12374
46. Zhang, Y. M., Rao, M. S., Heath, R. J., Price, A. C., Olson, A. J., Rock, C. O., and White, S. W. (2001) *J. Biol. Chem.* **276**, 8231–8238
47. Eckermann, S., Schröder, G., Schmidt, J., Strack, D., Edrada, R. A., Helariutta, Y., Elomaa, P., Kotilainen, M., Kilpelainen, I., Proksch, P., Teeri, T. H., and Schroder, J. (1998) *Nature* **396**, 387–390
48. Russell, D. G., Mwandumba, H. C., and Rhoades, E. E. (2002) *J. Cell Biol.* **158**, 421–426
49. McGuire, K. A., Siggaard-Andersen, M., Bangera, M. G., Olsen, J. G., and von Wettstein-Knowles, P. (2001) *Biochemistry* **40**, 9836–9845