THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 36, Issue of September 3, pp. 37956–37963, 2004
Printed in U.S.A.
© 2004 by The American Society for Biochemistry and Molecular Biology, Inc.

Structural and Functional Analysis of Tetracenomycin F2 Cyclase from *Streptomyces glaucescens*

A TYPE II POLYKETIDE CYCLASE*

Received for publication, June 2, 2004, and in revised form, June 28, 2004
Published, JBC Papers in Press, June 30, 2004, DOI 10.1074/jbc.M406144200

Thomas B. Thompson‡§, Kinya Katayama¶, Kenji Watanabe‖, C. Richard Hutchinson‖‡‡, and Ivan Rayment‡§§**

From the ‡Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706, the ¶Department of Applied Bioscience, Graduate School of Agriculture, Hokkaido University, Sapporo 060-8589, Japan, and the ‖School of Pharmacy, University of Wisconsin, Madison, Wisconsin 53705

Tetracenomycin F2 cyclase (*tcmI* gene product), catalyzes an aromatic rearrangement in the biosynthetic pathway for tetracenomycin C in *Streptomyces glaucescens*. The x-ray structure of this small enzyme has been determined to 1.9-Å resolution together with an analysis of site-directed mutants of potential catalytic residues. The protein exhibits a dimeric $\beta\alpha\beta$ ferredoxin-like fold that utilizes strand swapping between subunits in its assembly. The fold is dominated by four strands of antiparallel sheet and a layer of α-helices, which creates a cavity that is proposed to be the active site. This type of secondary structural arrangement has been previously observed in polyketide monooxygenases and suggests an evolutionary relationship between enzymes that catalyze adjacent steps in these biosynthetic pathways. Mutational analysis of all of the obvious catalytic bases within the active site suggests that the enzyme functions to steer the chemical outcome of the cyclization rather than providing a specific catalytic group. Together, the structure and functional analysis provide insight into the structural framework necessary to perform the complex rearrangements catalyzed by this class of polyketide cyclases.

Polyketides represent a diverse group of natural products produced by plants, fungi, and bacteria that have well established therapeutic and antibiotic properties (1). The biosynthesis of polyketides resembles fatty acid synthesis, in which a series of small CoA precursors, such as acetate, propionate, or butyrate, are combined by the polyketide synthase enzyme(s) (2, 3). Polyketide synthases are organized into two classes: type I polyketide synthases, which are giant multifunctional, multidomain proteins that contain multiple active sites on a single enzyme, and type II polyketide synthases, which are discrete multienzyme systems made up of largely monofunctional enzymes. Type I polyketide synthases contain motifs that resemble the acyl carrier protein (ACP),[1] β-ketoacyl:ACP synthase, and an ACP acyltransferase from the domains of animal fatty acid biosynthesis. In addition, ketoreductase and dehydratase activities play an important role in modifying the polyketide intermediates.

Type II polyketide synthases mostly generate aromatic ring-containing compounds from the combination of acetyl and malonyl-CoA (4, 5). Examination of gene clusters in *Streptomyces* and other bacterial species indicates that the bacterial set of biosynthetic enzymes includes a β-ketoacyl:ACP synthase, ACP, and ketoreductase to generate poly-β-ketone intermediates, which are subsequently condensed by a unique set of enzymes called cyclases or aromatases. These latter enzymes catalyze an intermolecular aldol condensation of the poly-β-ketone intermediate. Since the poly-β-ketone intermediates are highly unstable, much of the work done thus far has been on the terminal cyclases. Enzymes such as JadI from the jadomycin pathway and DnrD from the doxorubicin pathway catalyze the final aldol condensation reaction to form a fused aromatic ring.

Tetracenomycin C is an antibiotic/antitumor aromatic polyketide produced by *Streptomyces glaucescens* (5). It has been shown that the enzymes TcmJKLM can synthesize a linear decaketide that can spontaneously condense into TcmF2 (Fig. 1) (6, 7); however, aberrant side products are generated. The addition of TcmN to the other polyketide synthase components enhances the production of TcmF2. TcmF2 is converted to TcmF1 by the enzyme TcmI (Fig. 1) to form the final fused aromatic ring (8). Several additional steps catalyze the conversion of TcmF1 to the final product tetracenomycin C.

The polyketide cyclase enzyme TcmI has been characterized, and initial enzymatic studies have been performed (8). It is a small enzyme with a molecular weight of 12,728, where gel filtration suggested that the enzyme functions as a trimer. The reaction has been shown to yield two products when the pH of the reaction is varied. At pH >8.0 the carboxylated form of TcmF1 is generated, whereas the decarboxylated form is most prevalent at a pH of 6.5 or lower. Additionally, there is a strong

* This research was supported by National Institutes of Health Grant AR-35186 (to I. R.). Use of the Argonne National Laboratory Structural Biology Center beamlines at the Advanced Photon Source was supported by the United States Department of Energy, Office of Energy Research, under Contract W-31-109-ENG-38. 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.

The atomic coordinates and structure factors (code 1TUW) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§ Supported by the NIH Biophysics Training Grant GM08293. Present address: Dept. of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston IL 60208.

** Present address: Dept. of Applied Bioscience Hokkaido University Sapporo 060-8589, Japan.

‡‡ To whom correspondence may be addressed: School of Pharmacy, University of Wisconsin, Madison, WI 53705. E-mail: hutchinson@kosan.com.

§§ To whom correspondence may be addressed: Dept. of Biochemistry, 433 Babcock Dr., Madison, WI 53706. Tel.: 608-262-0437; Fax: 608-262-1319; E-mail: Ivan_Rayment@biochem.wisc.edu.

[1] The abbreviations used are: ACP, acyl carrier protein; DTT, dithiothreitol; HPLC, high pressure liquid chromatography.

This paper is available on line at http://www.jbc.org

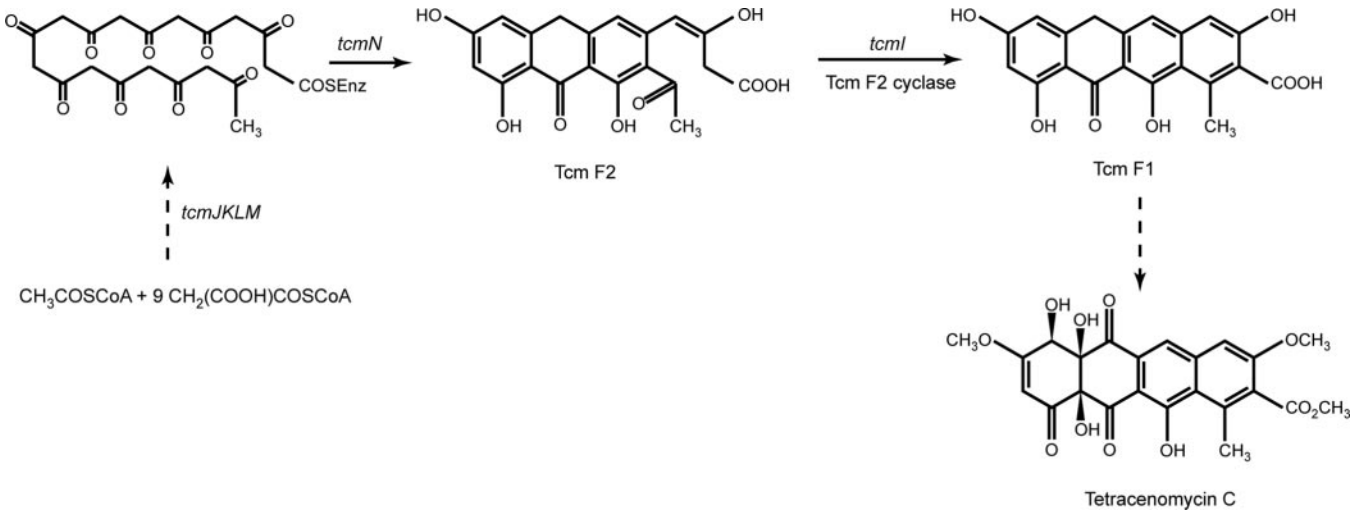

FIG. 1. **Biosynthetic step catalyzed by tetracenomycin F2 cyclase (TcmI) in *S. glaucescens*.**

drop off of product formation at pH 6.5. This suggests that a catalytic residue, such as a histidine, may play a role in catalysis.

In order to establish a molecular framework for understanding the reaction mechanism and specificity of this enzyme, a structural study coupled with site-directed mutagenesis was initiated. We report here the structure of tetracenomycin F2 cyclase (TcmI) from *S. glaucescens* to 1.9-Å resolution. This shows that the enzyme exhibits a fold that is similar to that found in other polyketide monooxygenases and provides insight into structural framework necessary to perform the complex rearrangements catalyzed by this class of polyketide cyclases. Although a structure with bound substrate could not be obtained, the close relationship of the structures of TcmI and the actva-orf6 monooxygenase (9) allowed a prediction of the active site of TcmI and enabled examination of the role of four amino acids in enzyme catalysis.

EXPERIMENTAL PROCEDURES

Materials—Oligonucleotides were custom-synthesized by University of Wisconsin Biotechnology Center. Restriction endonucleases and T4 DNA ligase were purchased from Invitrogen, and Pwo DNA polymerase, purchased from Roche Applied Science, was used for polymerase chain reaction. *Escherichia coli* strain BL21(DE3), and plasmids pET-15b, pET-28b(+), and pET-30a(+) were purchased from Novagen.

Protein Cloning and Site-directed Mutagenesis—The open reading frame of the *tcmI* gene was inserted into the NcoI-BamHI sites of the pET-15b vectors with the PCR primers 5′-AGC GTG CCA TGG CTT ACC GTG CTC TG-3′ and 5′-GTT AGC AGC CGG ATC CCG GAG GTC GC-3′ to yield plasmid pET15bTcmI.

Site-directed mutant proteins were made directly in plasmids pET15bTcmI, pET28bTcmI, and pET30aTcmI according to the QuikChange protocol (Stratagene). Two double mutants, D27N/R40G and R40G/H51A, were made utilizing the R40G plasmid as a template. The mutagenic primers in this study were 5′-GCC TTC GCC GAG **GC**C GAC ACG ACC-3′, 5′-GGT CGT GTC G**GC** CTC GGC GAA GGC-3′ (H26A complementary primer), 5′-GCC TTC GCC GAG CA**G** GAC ACG ACC-3′, 5′-GGT CGT GTC **C**TG CTC GGC GAA GGC-3′ (H26Q complementary primer), 5′-TTC GCC GAG CAC AAC ACG **A**CC GAA CTG-3′ (D27N), 5′-CAG TTC GG**T** CGT GTT GTG CTC GGC GAA-3′ (D27N complementary primer), 5′-GTG CGA CGG **G**GG GTC CTG TTC-3′ (R40G), 5′-GAA CAG GAC CC**C** CCG TCG CAC-3′ (R40G complementary primer), 5′-GTG CGA CGG **AA**G GTC CTG TTC CGC-3′ (R40K), 5′-GCG GAA CAG GAC C**TT** CCG TCG CAC-3′ (R40K complementary primer), 5′-GAT CTG TAC ATG CA**A** CTC ATC GAG GCC-3′ (H51Q), 5′-GGC CTC GAT GAG **T**TG CAT GTA CAG ATC-3′ (H51Q complementary primer), 5′-GAT CTG TAC ATG **GC**C CTC ATC GAG GCC-3′ (H51A), and 5′-GGC CTC GAT GAG G**GC** CAT GTA CAG ATC-3′ (H51A complementary primer) (the bases in boldface type indicate the mutated sites). The mutations were confirmed by sequencing.

Cloning and Expression of the Wild-type Protein—For protein expression of the wild type protein, the plasmids were transformed into BL21 (DE3). In general, a single colony was used to inoculate a 3-ml liquid culture of LB medium containing ampicillin (100 μg/ml). After incubation at 37 °C and 200 rpm for 16 h, 0.6 ml of culture was transferred into a 2-liter baffled Erlenmeyer flask containing 600 ml of medium as described above. The culture was grown under the same conditions until the A_{600} was ~1.0. Thereafter, 3 ml of this culture was used to inoculate a New Brunswick Scientific MPPF vessel (60-liter working volume). The pH and the dissolved oxygen were maintained at 7.0 and 20% of air saturation, respectively. Ampicillin was not added to the culture medium. The culture was grown at 37 °C until the A_{600} was ~1.0, at which point the culture was induced by a batch addition of isopropyl thio-β-D-galactoside (Amresco) (0.5 mM), and the culture growth was allowed to continue for 4 h. A yield of ~5 g of wet cell paste was obtained per liter of culture medium.

Selenomethionine-labeled TcmI was obtained by controlled fermentation (10). *E. coli* B834 (DE3) was transformed with pKI1. A single colony from a LB agar plate containing 100 μg/ml ampicillin was used to inoculate a 3-ml liquid culture of the same medium. When this culture reached A_{600} 0.2, 0.2-ml aliquots were used to inoculate two 2-liter baffled Erlenmeyer flasks containing 500 ml of a minimal medium (10). The 500-ml shaken flask cultures were grown to $A_{600} \approx 0.8$ and were then used to inoculate a fermentor containing 10 liters of the same medium prepared without ampicillin. During the fermentation, the pH and the dissolved O_2 were maintained at 7.0 and 20% of air saturation, respectively. The rate of O_2 utilization was then closely monitored to detect when methionine was depleted, and when the culture reached $A_{600} \approx 5$, the rate of O_2 utilization dropped, indicating the onset of methionine depletion. At this time, 2.0 g of (±)-selenomethionine (ACROS) was added, and 2 min later the culture was induced by the addition of 4 g/liter of D-(+)-lactose. A yield of ~8 g of wet cell paste was obtained per liter of culture medium. 60–70% of the recombinant TcmI and selenomethionine-labeled TcmI were in the insoluble form. This material could not be solubilized with detergent or refolded from urea or guanidine HCl.

Overproduction of TcmI Mutant Proteins—The resultant plasmids were introduced via transformation into *E. coli* strain BL21(DE3). Five ml of LB medium containing 100 μg/ml sodium ampicillin or 50 mg/ml kanamycin was inoculated with a single colony of *E. coli* strain BL21(DE3) transformed and incubated overnight at 37 °C. One liter of the same medium with 100 μg/ml sodium ampicillin or 50 μg/ml kanamycin was seeded with the densely grown preculture and shaken at 150 rpm at 27 °C for 12 h. Then 0.5 mM isopropyl-β-D-thiogalactopyranoside was added to induce expressions. After induction, incubation was contained for 12 h at 27 °C. The cells were harvested by centrifugation at 2,500 × *g*, and the cell pellet was stored at −80 °C.

Purification of Wild-type TcmI and Selenomethionine-labeled TcmI—Cell pellets were resuspended in 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 1 mM DTT, 0.1 mM phenylmethanesulfonyl fluoride and sonicated with a Sonifer cell disruptor, model 185 (Branson). DNase I (Sigma) and $MgCl_2$ were added to the lysate to final concentrations of 1 μg/ml and 10 mM, respectively, and the mixture was left to incubate on ice for 30 min. Clear lysate (13,680 × *g*, 20 min, 4 °C) was dialyzed against 10 mM Tris-HCl, pH 8.0, 1 mM DTT and loaded on to Q-Sepharose FF (Amer-

TABLE I
Kinetic data for the TcmI active site mutants

Variant	K_m	V_{max}	Relative activity	k_{cat}	K_{cat}/K_m
	μM	$\mu mol\ min^{-1}\ mg^{-1}$		s^{-1}	$M^{-1}\ s^{-1}$
Wild type	155 ± 12.7	0.680 ± 0.041	100	0.142 ± 0.0086	916 ± 75
H26Q	203 ± 26.4	0.627 ± 0.082	91	0.131 ± 0.0017	645 ± 84
H26A	732 ± 58.6	0.0920 ± 0.0074	15	0.0192 ± 0.0015	26.2 ± 2.1
D27N	(1.05 ± 0.0861) × 10^3	0.0948 ± 0.0118	14	0.0198 ± 0.0025	18.8 ± 2.4
R40G	(1.20 ± 0.0972) × 10^3	0.0703 ± 0.0072	10	0.0146 ± 0.0015	12.1 ± 1.2
R40K	816 ± 67.7	0.108 ± 0.012	16	0.0250 ± 0.0028	30.6 ± 3.4
H51Q	191 ± 28.7	0.651 ± 0.064	96	0.136 ± 0.013	712. ± 10.
H51A	892 ± 79.4	0.105 ± 0.013	16	0.0219 ± 0.0026	24.5 ± 2.9

sham Biosciences) column and eluted with 10 column volumes of a linear NaCl gradient (0–1.0 M). Fractions were analyzed by cyclase activity and SDS-PAGE. Fractions contained TcmI were brought to 1.2 M $(NH_4)_2SO_4$ by the addition of solid ammonium sulfate, and pH was adjusted to 8.0 by the addition of NH_4OH. The resultant solution was loaded onto a phenyl-Sepharose CL6B (Amersham Biosciences) column and eluted with 10 column volumes of a linear $(NH_4)_2SO_4$ gradient (1.2 to 0 M). The fractions containing TcmI were concentrated to ~20 mg/ml and buffer-exchanged with 10 mM Tris-HCl, pH 8.0, 1 mM DTT utilizing Centriprep YM-3 centrifugal units (Millipore). The protein was further purified on a MonoQ HR10/10 column (Amersham Biosciences) and eluted with five column volumes of a linear NaCl gradient (0–0.6 M). Minor contaminants were removed thereafter by gel filtration on a Sephacryl S-200 (Amersham Biosciences) column and eluted with 10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM DTT. Purified TcmI was then concentrated and buffer-exchanged with 5 mM Tris-HCl, pH 7.5, 1 mM DTT using Centriprep YM-3. The resultant protein was drop-frozen into liquid nitrogen and stored at −80 °C.

The selenomethionine-labeled TcmI was purified in a similar manner to that of the native protein except that all buffers were freshly degassed by aspiration and that dithiothreitol was added to all purification buffers to a final concentration of 10 mM.

Purification of TcmI Variants—The purification of the mutant proteins was performed in a similar manner to that of the wild type protein except that the thawed cells were gently resuspended in 25 ml of 50 mM sodium phosphate, pH 7.2, containing 1 mM EDTA, 1 mM DTT, 0.1 mM phenylmethanesulfonyl fluoride, 10% (w/v) glycerol. The filtrate was loaded onto a Q-Sepharose Fast Flow column (160 ml; Amersham Biosciences) previously equilibrated in degassed buffer A (50 mM Tris-HCl, pH 7.5, 1 mM DTT) at 2 ml/min. The column was washed extensively with buffer A and then eluted with an NaCl gradient (0–1 M) in buffer A. The active fraction was dialyzed overnight against buffer B (50 mM Tris-HCl, pH 7.5, 1 mM DTT, 1.7 M NH_4SO_2), filtered, and absorbed on to a HiPrep 16/10 Butyl column (Amersham Biosciences) previously equilibrated in degassed buffer B at 1 ml/min. The column was washed with buffer B and then eluted with an NH_4SO_2 gradient (1.7 to 0 M) in buffer B. The protein was further purified by anion exchange chromatography, concentrated, and stored at −80 °C as before. A typical 1-liter culture produced about 0.2 mg of purified enzyme.

Kinetic Assays—Kinetic assays were carried out at 30 °C in 0.1 M Tris-HCl, pH 7.5, by monitoring the consumption of TcmF2 and the formation of TcmF1 by HPLC. For determination of Michaelis-Menten parameters k_{cat} (per active site) and K_m, the series of substrate concentrations was as follows: 0.062, 0.081, 0.12, 0.24, and 5.0 mM or 0.31, 0.38, 0.50, 0.71, 1.3, and 5.0 mM. The assay solution was preincubated at 30 °C for 10 min. The reaction then was initiated by the addition of TcmF2 and terminated after a 5- or 30-min incubation at 30 °C by the addition of solid NaH_2PO_4 (100 mg) to saturation and extraction with ethyl acetate (200 μl). The aliquot was directly injected into the HPLC system. The HPLC analytical conditions were as described previously (8). The results of the kinetic assays are given in Table I.

Crystallization and Data Collection—Crystals of tetracenomycin F2 cyclase were grown by microbatch from a solution composed of 2.2 M $(NH_4)_2SO_4$, 3 mg/ml protein, 50 mM HEPES, pH 7.5, at 4 °C (11). Large crystals were obtained by macroseeding with small crystals (0.06 × 0.08 mm) washed in macroseeding 1.4 M $(NH_4)_2SO_4$, 50 mM HEPES, pH 7.5, to slightly dissolve the crystal surfaces. After 3–4 weeks, the crystals attained a size of 0.3 × 0.3 × 0.6 mm. Crystals of the selenomethionine-labeled protein were obtained in a similar manner.

The crystals belong to the space group $P6_522$ and have unit cell dimensions of $a = b = 46.6$ Å, $c = 188.3$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$ with one molecule per asymmetric unit. Prior to x-ray data collection, crystals were transferred stepwise into a solution containing 2.0 M $(NH_4)_2SO_4$, 20% ethylene glycol, 50 mM HEPES at pH 7.5.

Multiple anomalous dispersion selenomethionine-labeled data were collected at the Advanced Photon Source on the 19-ID beam-line of the Structural Biology Center at Argonne National Laboratory on a single crystal at three wavelengths (Table I). Data were recorded with a 3-s exposure time and an oscillation angle of 1° at a crystal to detector distance of 235 mm. The data were processed with DENZO and subsequently scaled with SCALEPACK (12). Data collection statistics are shown in Table II.

For the final structure, data were collected from a native crystal with a Bruker HiStar multiwire area detector at −4 °C to 1.9-Å resolution. Data were recorded at a crystal to detector distance of 22 cm with CuK_α radiation generated by a Rigaku RU200 x-ray generator operated at 50 kV and 90 mA and equipped with Göbel focusing mirrors. Data were collected with an oscillation angle of 0.15° for 45–60 s, integrated with XDS (13) and internally scaled with XSCALIBRE. Data collection statistics for the native data are presented in Table II.

Structural Determination and Refinement—The structure of tetracenomycin F2 cyclase was solved to 2.4-Å resolution by multiwavelength anomalous dispersions methods. The atomic positions of the seleniums were determined with the program SOLVE (14) and refined with the program CNS (15). The correct enantiomer for the space group was chosen by visual inspection of the resultant electron density maps. A model for the protein was built with TURBO-FRODO (16) and refined with the maximum likelihood refinement of CNS and the least squares refinement of TNT. The model derived from the Se-Met data (R_{work} of 23.0% and R_{free} of 28.4%) was transferred to the native data to complete the refinement. Iterative cycles of least squares refinement and manual model building with the programs TNT and Turbo (16, 17) reduced the R_{work} to 19.9% for all measured x-ray data from 30.0- to 1.9-Å resolution. The R_{free} was 24.5% for 5% of the data that were excluded from the refinement. Least squares refinement statistics are presented in Table III. Analysis of the coordinates with the program PROCHECK (18) revealed that 91.8% of the residues lie in the most favored regions of the Ramachandran plot, whereas the remaining 8.1% of the residues lie in additionally allowed areas. A section of representative electron density is shown in Fig. 2.

RESULTS AND DISCUSSION

The final model for the tetracenomycin F2 cyclase contains 106 of 109 amino acids. The electron density starts at Ala^2 and extends continuously to Pro^{107}, where the two remaining residues in the sequence could not be modeled. The protein assembles as a tight dimer, where each subunit exhibits a fold that belongs to the ferredoxin fold superfamily that is characterized by βαβ repeats in the SCOP data base (19). This disagrees with the earlier prediction, based on gel filtration, that the enzyme might assemble as a trimer (8); however, this is not surprising because of the nonspherical shape of the dimer discussed below. The globular domain consists of two layers of secondary structural elements: an antiparallel β-sheet in one layer and three helices in the second (Fig. 3). The sheet exhibits the strand order *1*, *3*, *2*, (*4′*), where the first and second strands are connected by the first helix, and the second and third are connected by a β-turn. The second and third helix follow the third β-strand and contribute to the second layer of secondary structural elements. Thereafter, the final 10 residues form an extended section of the polypeptide chain, starting at Ala^{98}, that domain-swaps with the dimer-related subunit. This segment forms a fourth β-strand that increases the breadth of the antiparallel β-sheet. As such, the fold for tetracenomycin F2 cyclase belongs to the ferredoxin

TABLE II
Data collection statistics for the native and Se-Met MAD data

Parameters	Values			
	1.5418-Å wavelength Native	1.0194-Å wavelength Remote λ_1 Se-Met	0.9793-Å wavelength Peak λ_2 Se-Met	0.9792-Å wavelength Edge λ_2 Se-Met
Resolution (Å)	30–1.9	30.0–2.4	30.0–2.4	30.0–2.4
Reflections	33,040	69,565	62,155	62,293
Unique reflections	9599	8012	8488	8457
Average redundancy	8.8	8.7	7.3	7.4
Completeness (%)[a]	92.0 (84.3)	92.4 (52.6)	98.3 (98.7)	98.0 (99.1)
$\langle I\rangle/\langle\sigma\rangle$[a]	23.05 (5.36)	32.8 (19.2)	44.7 (24.0)	45.5 (26.8)
R_{merge} (%)[a,b]	2.7 (8.8)	4.9 (7.2)	5.0 (7.4)	4.9 (6.8)

[a] Value in parenthesis represents data from the resolution shell of 2.42 to 2.34 Å.
[b] $R_{merge} = (\Sigma|I_{hkl} - \langle I\rangle|)/(\Sigma I_{hkl})$, where the average intensity $\langle I\rangle$ is taken over all symmetry equivalent measurements, and I_{hkl} is the measured intensity for any given reflection.

TABLE III
Least squares refinement statistics

Parameters	Values
Resolution (Å)	30 to 1.9
Reflections	9599
Final R_{work} (%)[a]	19.9
Final R_{free} (%)[b]	24.5
Protein atoms	23.05 (5.36)
Solvent molecules	2.65 (8.8)
Other molecules, ions	1 sulfate ion
Average B values (Å^2)	
Main chain atoms	32.1
All protein atoms	38.7
Solvent molecules	49.1
Sulfate ion	63.4
Weighted root mean square deviation from ideality	
Bond lengths (Å)	0.017
Bond angles (degrees)	2.61
Planarity (trigonal) (Å)	0.007
Planarity (others) (Å)	0.015
Torsional angles (degrees)	19.3
Disordered side chains	Leu[33], Arg[38], Glu[54], Asp[57], Gln[65], Arg[67], Gln[73], and Lys[94]

[a] $R_{work} = (\Sigma\ ||F_o| - k|F_c||)/(\Sigma|F_o|)$.
[b] R_{free} was calculated with 5% of the data. The data utilized in the R_{free} calculation were removed prior to least squares refinement

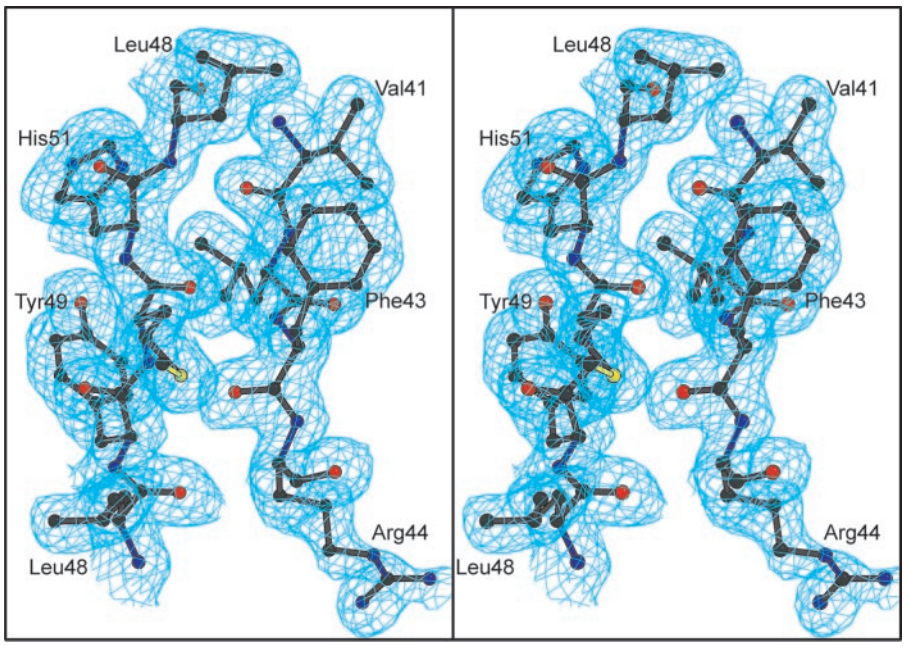

FIG. 2. **Stereo view of representative electron density from the TcmI.** The electron density, contoured at 1 σ, belongs to the second and third β-strands and was calculated with coefficients of the form $2F_o - F_c$ map. The figure was prepared with the program Bobscript (29).

superfamily that is characterized by $\beta\alpha\beta$ repeats. Overall, the β-sheet exhibits a pronounced curvature.

Two subunits of TcmI assemble to form an elongated dimer with a compact with and extensive interface between subunits. The interface is created by the juxtaposition of the dimer related β-sheets, which creates a distorted antiparallel β-barrel that accommodates the pronounced curvature of the β-sheets (Fig. 4). A substantial fraction of the surface area of each subunit, 1951 Å^2, is buried as determined with the program GETAREA (20). The interface between the monomers is highly hydrophobic as indicated by the burial of 1459 Å^2 of nonpolar surface area per subunit.

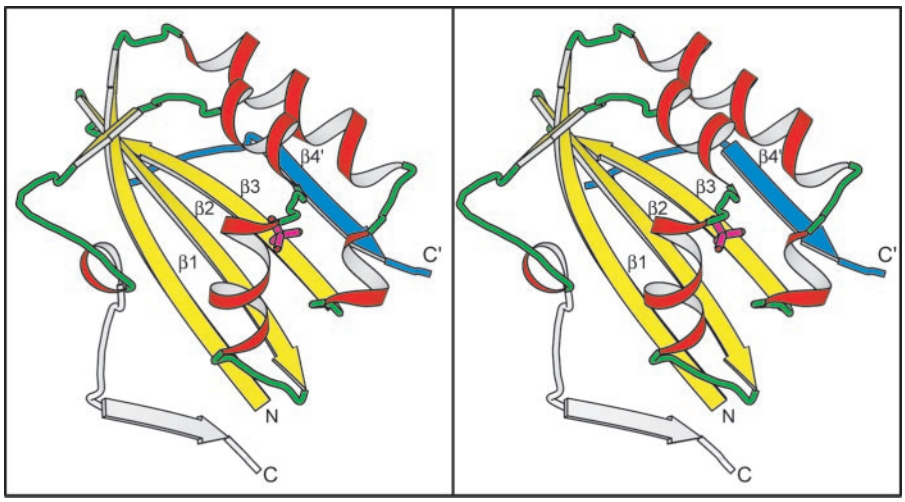

FIG. 3. **Stereo ribbon representation of the monomer of TcmI that shows the strand order in this ferredoxin-like fold.** The final section of the polypeptide is depicted in *white*, since this "domain-swaps" with the dimer-related subunit. The equivalent segment from the 2-fold related molecule is depicted in *blue*. The figure was prepared with the program Bobscript (29).

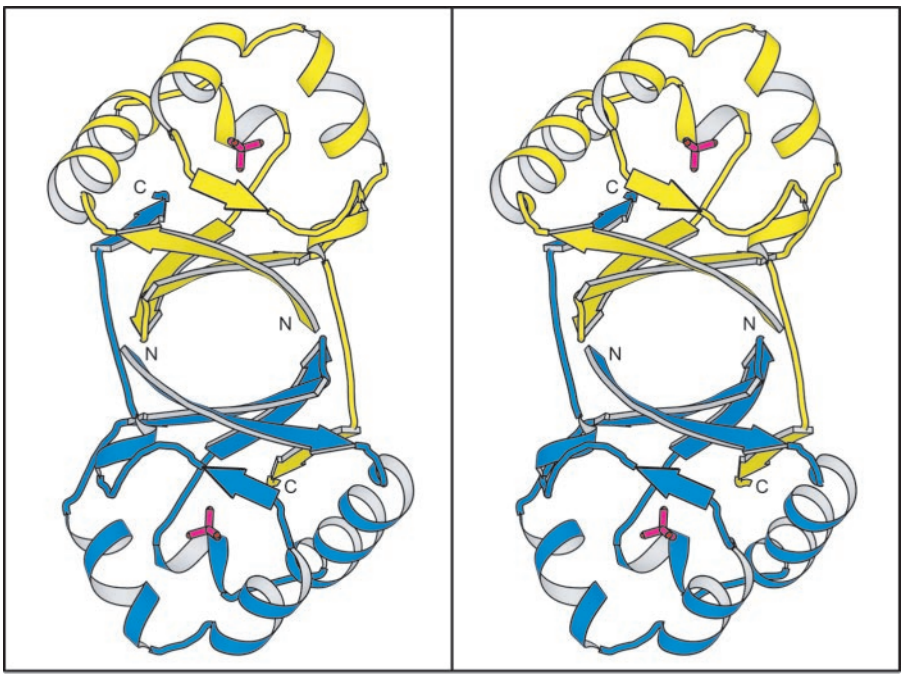

FIG. 4. **Stereo ribbon representation of the quaternary structure of TcmI.** The molecule assembles as a dimer where the 2-fold axis is coincident with an axis of symmetry in the crystal lattice. The figure was prepared with the program Molscript (30).

A search with the DALI server reveals a strong topological similarity between TcmI and actva-orf6 monooxygenase from *Streptomyces coelicolor* (Protein Data Bank accession number 1LQ9, Z-score 7.0) (9, 21) with the next closest similarity to a subdomain of cholesterol oxidase from *Brevibacterium sterolicum* (Protein Data Bank accession number 1IL9, Z-score 4.5) (22). The relationship to cholesterol oxidase is purely topological, where the fold observed in TcmI occurs as a component of a domain that is dominated by an eight-stranded antiparallel β-sheet. This will not be discussed further. In contrast to cholesterol oxidase, there does appear to be a structural and functional relationship with the monooxygenase, since this catalyzes the oxidation of 6-deoxydihydrokalafungin to dihydrokalafungin in the biosynthetic pathway for actinorhodin. Actinorhodin is also a polyketide antibiotic with a similar molecular framework to that seen in tetracenomycin C. In chemical terms, this latter enzyme catalyzes an earlier reaction to that performed by the *tcmI* gene product in the tetracenomycin C biosynthetic pathway.

The fold for the actva-orf6 monooxygenase contains five antiparallel β-strands with strand order *4*, *1*, *3*, *2*, *5′*, where the first three strands have the same topology as that now observed in tetracenomycin F2 cyclase. Furthermore, both enzymes exhibit the strand-swapping that serves to increase the width of the sheet adjacent to the second strand. The major difference between these folds is the inclusion of an additional antiparallel β-strand adjacent to the first strand in the fold for actva-orf6 monooxygenase. Even with this disparity, the root mean square difference between 85 structurally equivalent α-carbon atoms is only 2.2 Å (Fig. 5*a*).

Comparison of the two structures reveals that although tetracenomycin F2 cyclase lacks the additional strand seen in actva-orf6 monooxygenase, there is an extended section of random coil that follows the end of the second α-helix and precedes the swapped strand. This plays a role in closing off the large cavity that lies between the α-helices and β-sheets that is presumed to constitute the active site.

Because of the differences in the number of strands, the relative orientation of the β-sheets relative to the dimer 2-fold is different in the two enzymes. In tetracenomycin F2 cyclase, the presence of four antiparallel β-strands that form a curved sheet allows the two subunits to assemble so that the strands are inclined at an angle of ~30° to the dimer axis to form a cylindrical barrel (Fig. 5). This orientation of sheets is analogous to that seen in parallel $(\beta/\alpha)_8$ barrels.

The inclusion of an extra strand in the actva-orf6 monooxy-

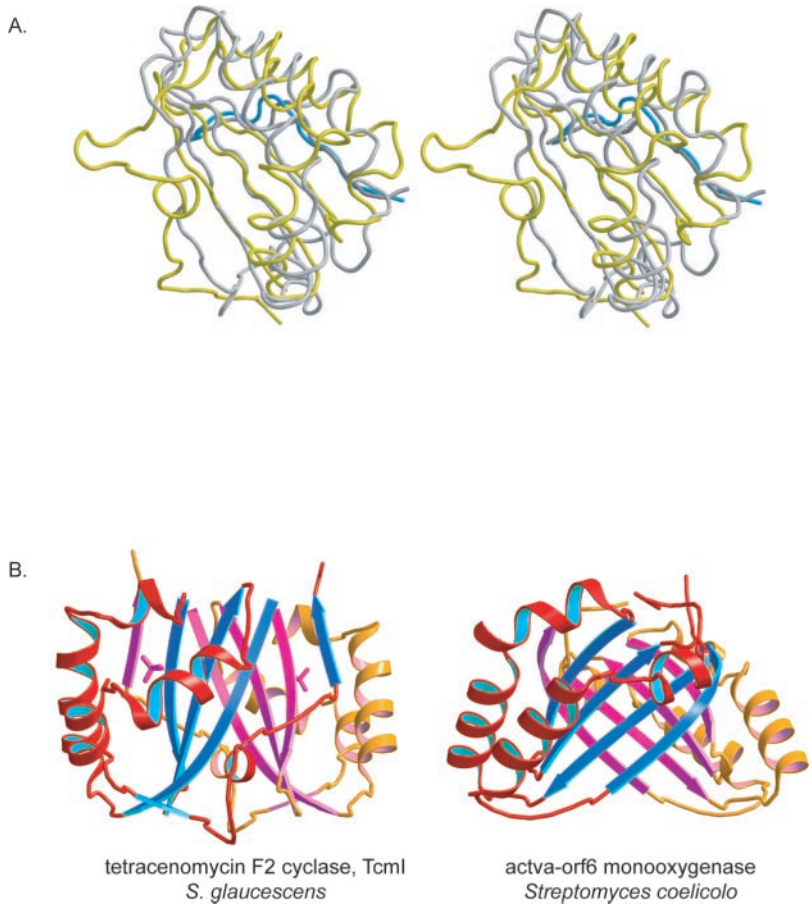

FIG. 5. **Comparison of the structures for TcmI and actva-orf6 monooxygenase from *Streptomyces coelicolor*.** *A*, a stereo comparison of TcmI and the monooxygenase depicted in *yellow* and *gray*, respectively. The domain-swapped β-strand from the dimer-related subunit is depicted in *blue* and *dark gray* for TcmI and actva-orf6 monooxygenase, respectively. *B*, a side-by-side comparison of the quaternary structures for TcmI and actva-orf6 monooxygenase from *S. coelicolor*. The figure was prepared with Molscript and Raster3d (31).

genase would create a hollow barrel if the strands were arranged at an angle of ~30° to the dimer axis. As a consequence, the sheets on opposing dimers are inclined approximately perpendicular to each other, which allows for a closely packed interface (Fig. 5). This orientation of strands is common in proteins that contain two layers of β-sheet as observed in the lipocalin family of proteins (23), although the latter normally utilize the space between the layers as a ligand binding site. The argument against 10-stranded parallel β-barrels was proposed many years ago to explain why compact barrels typically commonly contain eight strands of sheet (24, 25).

Active Site Pocket—The exact manner in which TcmI binds its substrates is unknown; however, the presence of a large cavity at the same general position as that seen in the actva-orf6 monooxygenase is highly suggestive that this is the active site (Fig. 6). Considerable effort was devoted to trying to obtain crystals of TcmI bound to substrate or products, but none of these gave crystals with meaningful electron density within the cavity. The reason for this was most likely the very high salt concentration needed to obtain crystals. In particular, the sulfate ion that is observed at the back of the cavity would be expected to inhibit substrate binding. Even in the absence of this information, the nature of the cavity provides insight into how the substrate might bind.

The cavity lies between the β-sheet and the layer formed by the three α-helices and is dominated by hydrophobic residues, which include Met^{7}, Leu^{7}, Phe^{23}, Pro^{32}, Ile^{35}, Ile^{53}, Leu^{63}, Phe^{72}, Val^{75}, Val^{79}, Leu^{83}, Trp^{90}, and Leu^{93}. These residues are ideally situated for binding a polyaromatic species. The cavity is empty except for numerous water molecules and a sulfate ion that lies at the back of the pocket, where the latter ion is coordinated by Arg^{40} and His^{26}.

Earlier biochemical characterization of TcmI showed a pronounced pH dependence in the cyclization reaction (8) with a pH optimum of 8 or higher, which suggests the involvement of a group with a high pK_a. Furthermore, when the pH is less than 6.5, the reaction product changes to 9-deoxy-TcmF1, a known shunt metabolite of the tetracenomycin biosynthetic pathway (26). Examination of the putative active site reveals that His^{51} is in close proximity to Arg^{40}. Furthermore, there is a salt bridge between Asp^{27} and Arg^{40}, which suggests that this acidic residue may be important in maintaining the orientation of the two former side chains. The only other potential catalytic residue is His^{26}. None of these residues, except for Arg^{40} are expected to have an anomalously high pK_a. To examine the role of these residues in the catalytic reaction, mutant proteins were prepared for His^{26}, Asp^{27}, Arg^{40}, and His^{51}.

Site-directed Mutagenesis of Active Site Residues—TcmI mutants with single or double amino acid changes in the four residues predicted to play key roles in catalysis (Fig. 6) were made by site-directed mutagenesis and purified from *E. coli* BL21(DE3) as described. The relative activity for mutations at His^{26}, Asp^{27}, Arg^{40}, and His^{51} was determined in enzymatic assays by monitoring the consumption of TcmF2 and the formation of TcmI at 30 °C in Tris buffer, pH 7.5, as described following the methods of Shen and Hutchinson (8). The results (Table I) show that the V_{max} was reduced to ~15% of the value of the wild type enzyme in the H26A, D27N, R40K, R40G, and H51A mutants, whereas the H26A and H51Q mutants retained 91 and 96% of that activity, respectively. The D27N/R40G and R40G/H51A double mutants had ≤1.1% of the activity of the wild type enzyme (data not shown). Although the K_m increased about 10-fold for the D27N and R40G mutants, the three other mutants displayed only 1.2- to ~5-fold increases. The catalytic efficiency (k_{cat}/K_m) underwent an ~30–76-fold decrease for all mutants with the exception of the H26A and H51Q mutants.

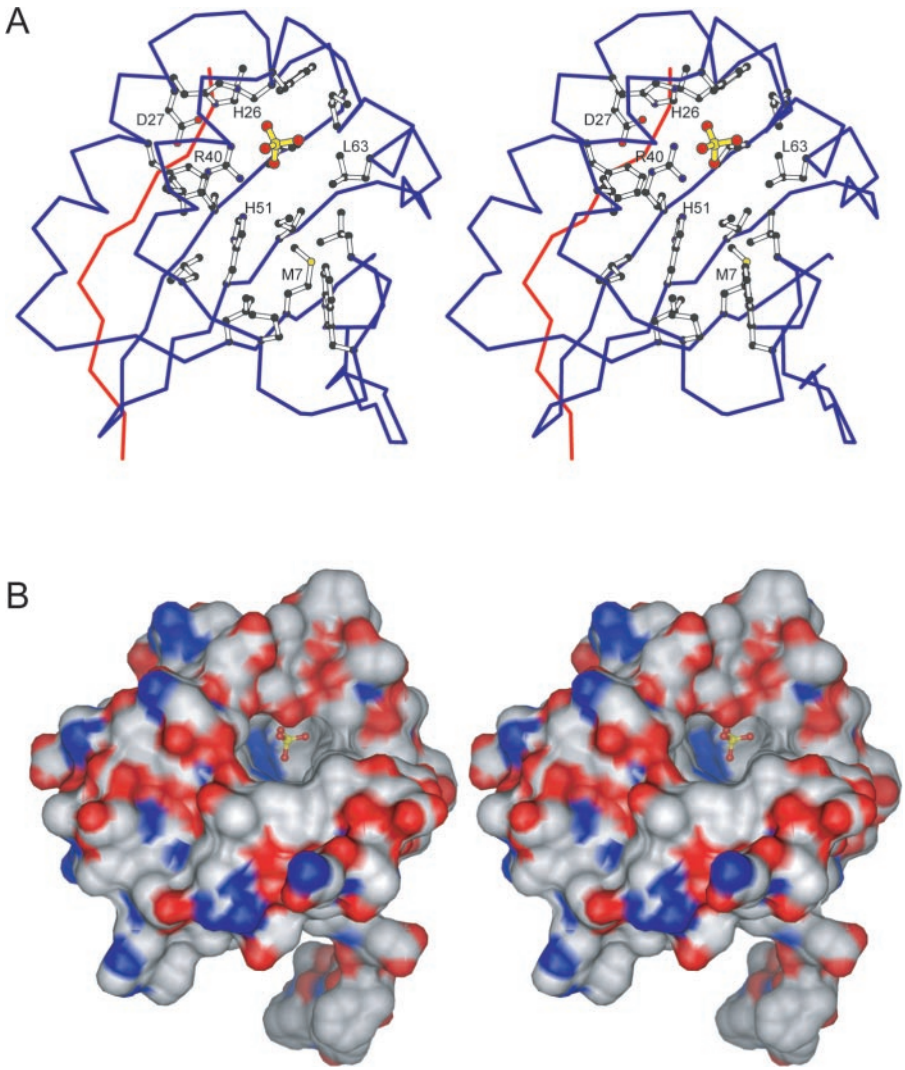

FIG. 6. **Stereo views of the active site cavity.** *A*, an α-carbon trace of TcmI decorated with the hydrophobic residues and major polar residues that form the active site cavity together with the sulfate ion, which is coordinated by Arg^{40} and His^{51}. *B*, a space-filling representation of the opening to the cavity. The figure was prepared with the programs Bobscript and Chimera (29, 32, 33).

Consequently, each of the four amino acids tested appear to be important to the function of TcmI, and thus their involvement supports the supposition that they are part of the active site of the enzyme. None of the mutations cause a decrease in activity that would be consistent with a role as a catalytic acid or base (a reduction of at least 10^5 would be expected).

It is particularly interesting that the effect of a mutation on catalysis depends on the nature of the amino acid substitution at His^{26} and His^{51}. Clearly, if either of these were to function as an obligatory catalytic base, almost any change should result in a large decrease in activity. Instead, replacement with a glutamine preserves much of the activity, whereas change to an alanine reduces the catalytic efficiency in both cases by ~35-fold. This implies that neither of these side chains is the catalytic base. Rather, it suggests that these side chains are involved in some form of hydrogen bonding or polar interactions with either the substrate or water molecules in the active site, since the conversion of a histidine to glutamine is expected to retain some of the hydrogen bonding characteristics of histidine as opposed to alanine.

At the time of the structural determination, it was assumed that an amino acid with an altered pK_a would be the catalytic base. Based on site-directed mutagenesis of the obvious candidates in the putative active site, this is now called into question, since there are no other residues in the cavity with a pK_a in an appropriate range. Given the relatively slow rate of reaction for this enzyme, it is possible that the active site serves to provide specificity or direction for a water-facilitated reaction. This is consistent with the variation in activity depending on the nature of the substitution. Clearly, more study of the reaction catalyzed by TcmI is needed before the interplay between the protein and its substrate becomes clear.

Conclusions—The structure of TcmI provides insight into the structure and potential mechanisms of type II polyketide cyclases. Such enzymes can serve as a mechanistic model for the type II polyketide cyclases that catalyze a series of intramolecular aldol reactions upon enzyme-bound poly-β-ketone intermediates (27, 28). Importantly, this structure shows strong topological similarity to the actva-orf6 monoxygenase in the actinorhodin biosynthetic pathway, which suggests that this fold is ideally suited to serve as a framework for rearrangements and chemical modification of polyaromatic substrates. It is expected that this motif will be found elsewhere in the biosynthetic pathways for polyketide antibiotics.

Acknowledgments—We thank Drs. P. Johnson and J. M. Studts (Department of Biochemistry, University of Wisconsin) for help with cell growth.

REFERENCES

1. O'Hagan, D. (1991) *The Polyketide Metabolites*, Ellis Horwood, New York
2. Staunton, J., and Weissman, K. J. (2001) *Nat. Prod. Rep.* **18,** 380–416
3. Cane, D. E., Walsh, C. T., and Khosla, C. (1998) *Science* **282,** 63–68
4. Hopwood, D. A. (1997) *Chem. Rev.* **97,** 2465–2498
5. Hutchinson, C. R. (1997) *Chem. Rev.* **97,** 2525–2536
6. Shen, B., and Hutchinson, C. R. (1993) *Science* **262,** 1535–1540
7. Bao, W., Wendt-Pienkowski, E., and Hutchinson, C. R. (1998) *Biochemistry* **37,**

8132–8138
8. Shen, B., and Hutchinson, C. R. (1993) *Biochemistry* **32,** 11149–11154
9. Sciara, G., Kendrew, S. G., Miele, A. E., Marsh, N. G., Federici, L., Malatesta, F., Schimperna, G., Savino, C., and Vallone, B. (2003) *EMBO J.* **22,** 205–215
10. Studts, J. M., and Fox, B. G. (1999) *Protein Expr. Purif.* **16,** 109–119
11. Rayment, I. (2002) *Structure* (*Camb.*) **10,** 147–151
12. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
13. Kabsch, W. (1988) *J. Appl. Crystallogr.* **21,** 916–924
14. Terwilliger, T. C. (1997) *Methods Enzymol.* **276,** 530–537
15. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr.* **D54**, 905–921
16. Roussel, A., and Cambillau, C. (1991) *Silicon Graphics Geometry Partners Directory*, Vol. 86, Silicon Graphics, Mountain View, CA
17. Tronrud, D. E. (1997) *Methods Enzymol.* **277,** 306–319
18. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Applied Crystallogr.* **26,** 283–291
19. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) *J. Mol. Biol.* **247,** 536–540
20. Fraczkiewicz, R., and Braun, W. (1998) *J. Comput. Chem.* **19,** 319–333
21. Holm, L., and Sander, C. (1999) *Nucleic Acids Res.* **27,** 244–247
22. Coulombe, R., Yue, K. Q., Ghisla, S., and Vrielink, A. (2001) *J. Biol. Chem.* **276,** 30435–30441
23. Flower, D. R., North, A. C., and Sansom, C. E. (2000) *Biochim. Biophys. Acta* **1482,** 9–24
24. Murzin, A. G., Lesk, A. M., and Chothia, C. (1994) *J. Mol. Biol.* **236,** 1382–1400
25. Murzin, A. G., Lesk, A. M., and Chothia, C. (1994) *J. Mol. Biol.* **236,** 1369–1381
26. Shen, B., Nakayama, H., and Hutchinson, C. R. (1993) *J. Nat. Prod.* **56,** 1288–1293
27. Dreier, J., and Khosla, C. (2000) *Biochemistry* **39,** 2088–2095
28. 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* (*Camb.*) **10,** 1559–1568
29. Esnouf, R. M. (1999) *Acta Crystallogr. Sec. D* **55**, 938–940
30. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
31. Merritt, E. A., and Bacon, D. J. (1997) *Macromol. Crystallogr.* **277,** 505–524
32. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) *J. Comput. Chem.* **25,** 1605–1612
33. Sanner, M. F., Olson, A. J., and Spehner, J. C. (1996) *Biopolymers* **38,** 305–320