THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 277, No. 27, Issue of July 5, pp. 24000–24005, 2002

Printed in U.S.A.

The Cytochrome P450 Complement (CYPome) of *Streptomyces coelicolor* A3(2)*

Received for publication, November 20, 2001, and in revised form, March 22, 2002
Published, JBC Papers in Press, April 9, 2002, DOI 10.1074/jbc.M111109200

David C. Lamb‡, Tove Skaug‡, Hong-Lin Song‡, Colin J. Jackson‡, Larissa M. Podust§, Michael R. Waterman§, Douglas B. Kell‡, Diane E. Kelly‡, and Steven L. Kelly‡¶

From the ‡Wolfson Laboratory of P450 Biodiversity, Institute of Biological Sciences, University of Wales Aberystwyth, Aberystwyth, Wales SY23 3DA, United Kingdom and the §Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, Tennessee 37232-0146

In the present study we describe the complete cytochrome P450 complement, the "CYPome," of *Streptomyces coelicolor* A3(2). Eighteen cytochromes P450 (CYP) are described, in contrast to the absence of CYPs in *Escherichia coli*, and the twenty observed in *Mycobacterium tuberculosis*. Here we confirm protein identity as cytochromes P450 by heterologous expression in *E. coli* and measurement of reduced carbon monoxide difference spectra. We also report on their arrangement in the linear chromosome and relatedness to other CYPs in the superfamily. The future development of manipulation of antibiotic pathways and the use of streptomycetes in bioremediation and biotransformations will involve many of the new CYP forms identified here.

Cytochrome P450 (CYP)[1]-dependent monooxygenases are a superfamily of heme-containing enzymes that are involved in a wide array of NADPH-/NADH- and O_2-dependent reactions (1, 2). Extensive studies in animals have established their role in metabolism of xenobiotic drugs and toxic chemicals as well as endogenous compounds such as sterols, fatty acids, and prostaglandins (3–5). Recently, the completion of fully sequenced genomes of prokaryotes and eukaryotes has revealed the extraordinary biodiversity of this superfamily. To date, as judged from sequence motifs alone, there are ~57 CYP genes in humans, ~90 in *Drosophila melanogaster*, ~80 in *Caenorhabditis elegans*, and ~275 in *Arabidopsis thaliana* (6). Additionally, CYP genes are also found in many microorganisms: three in the bakers' yeast *Saccharomyces cerevisiae*, ~150 in the white rot fungus *Phanerochaete chrysosporium*, and 20 in *Mycobacterium tuberculosis*, but *Escherichia coli* carries none (7). CYP proteins are extremely diverse in amino acid sequence with CYPs having less than 40% amino acid identity being placed in different families (8). Common features, such as the cysteine ligand to the prosthetic heme and its flanking conserved residues, allow most CYP sequences to be recognized (9). For catalytic activity, CYPs must be associated with their electron donor partner proteins, either NADPH cytochrome P450 reductase for microsomal eukaryotic CYPs or ferredoxin/ferredoxin reductase complex for prokaryotic and mitochondrial CYPs (for review see Ref. 10).

Streptomycetes produce a vast array of antibiotics applied in human and veterinary medicine and agriculture, as well as anti-parasitic agents, herbicides, and pharmacologically active metabolites (*e.g.* immunosuppressants). Streptomycetes also catalyze numerous transformations of xenobiotics of industrial and environmental importance (11, 12). These oxidative transformations have been observed with alkaloids (13), coumarins (14), retinoids (15), and other complex xenobiotics (16). Industrial application has been exploited in the synthesis of pravastatin utilizing a streptomycete biotransformation step (17) as well as in 16-hydroxylation of steroids (18) and in the preparation of drug metabolites for toxicological evaluation (19).

Streptomyces coelicolor A3(2) is the most studied member of the genus in molecular genetic investigations and has become a paradigm for the actinomycetes. In 1997, the Sanger Centre initiated a program to sequence the 8-Mb genome of *S. coelicolor* A3(2) (www.sanger.ac.uk/Projects/S_coelicolor/). The project has been carried out at the Sanger Centre at Hinxton, Cambridge, UK, was funded by the Biotechnology and Biological Science Research Council and the Beowulf Genomics initiative of the Wellcome Trust, and was completed in July 2001. Eighteen CYP sequences from *S. coelicolor* A3(2) have been revealed from the genome-sequencing project, but their endogenous roles and contribution to *S. coelicolor* A3(2) biology remain unknown. The present paper describes the biochemical verification of the complete cytochrome P450 complement (the so-called CYPome) of *S. coelicolor* A3(2) and discusses the potential function. Results are described in consideration of the evolution of the cytochrome P450 superfamily as a whole, particularly among the actinomycetes.

MATERIALS AND METHODS

DNA Manipulations—The open reading frames (ORFs) of the putative cytochrome P450 enzymes present in the *S. coelicolor* A3(2) genome were revealed through analysis of the presence of CYP heme-binding motifs in the genomic information as it was deposited. Eighteen CYP ORFs were identified, and the corresponding DNA sequences were isolated by PCR and cloned into the *E. coli* expression vector pET17b (Novagen) using the *Nde*I and *Hin*dIII sites. Forward and reverse primers to isolate each CYP are described in Table I. In each PCR product an *Nde*I site encoding ATG was incorporated defining the 5′ of the open reading frame. Also the stop triplet was removed and triplets encoding four histidine were inserted for subsequent protein purification studies. These were followed by a new stop signal, and beyond the reading frame a *Hin*dIII site was included to facilitate cloning into pET17b. In four cases the predicted initiating codon was not ATG, CYP155A1, CYP157A1, CYP157B1, and CYP157C1, and in these cases the initiating triplet of the gene construct for expression was replaced by ATG within an *Nde*I site. The expression vector for each individual CYP was transformed separately into competent HMS174 (Novagen) cells of *E. coli* prior to heterologous expression and purification according to Bellamine *et al.* (20).

Bioinformatic Analysis—Data were collected at various addresses

* This work was partly supported by Biotechnology and Biological Sciences Research Council and Medical Research Council. 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.

¶ To whom correspondence should be addressed. Tel.: 0044-1970-621515; Fax: 0044-1970-622350; E-mail: steven.kelly@aber.ac.uk.

[1] The abbreviations used are: CYP, cytochrome P450; ORF, open reading frame; Ni^{2+}-NTA, nickel-nitrilotriacetic acid.

TABLE I

Consensus sequences in the CYPome of Streptomyces coelicolor

Residues conserved in the I-helix (Thr), the K-helix (Glu and Arg), and the heme-binding site (Gly and Cys) are highlighted in bold.

CYP name[1]	Gene name[2]	I-helix	K-helix	Heme binding motif
CYP51	SC7E4.20	^{274}PGSE**T**^{278}	^{330}**E**AM**R**^{333}	^{403}F**G**AG**R**K**R**K**C**P**S**^{412}
CYP102B1	SCF43.12	^{301}AGHE**T**^{305}	^{358}**E**SL**R**^{361}	^{433}F**G**T**G**ARAC**I**G^{442}
CYP105D5	3SCF60.06c	^{251}AGHE**T**^{255}	^{290}**E**LM**R**^{293}	^{354}F**G**F**G**IHQC**L**G^{363}
CYP105N1	SC4C2.21	^{249}AGRE**T**^{253}	^{288}**E**LL**R**^{291}	^{353}F**G**Y**G**VHQC**V**G^{362}
CYP107P1	SCH10.14c	^{253}AGHEA^{257}	^{292}**E**LM**R**^{295}	^{356}F**S**AG**I**HYC**I**G^{365}
CYP107T1	SCH63.17	^{231}AGHE**T**^{235}	^{270}**E**SL**R**^{273}	^{335}F**G**H**G**PHHC**L**G^{344}
CYP107U1	SCE41.08c	^{250}AGFE**T**^{254}	^{294}**E**LL**R**^{297}	^{359}Y**G**H**G**IHYC**L**G^{368}
CYP154A1	SCE6.21	^{242}AGYE**T**^{246}	^{281}**E**TL**R**^{284}	^{347}F**G**H**G**VHFC**L**G^{356}
CYP154C1	SC6D11.13c	^{242}AGHE**T**^{246}	^{281}**E**TL**R**^{284}	^{348}F**G**H**G**PHVC**P**G^{357}
CYP155A1	SC6D11.40c	^{250}AGMV**T**^{254}	^{290}**E**LL**R**^{293}	^{361}F**G**D**G**PHRC**P**G^{370}
CYP156A1	SCE6.20	^{241}AGIE**P**^{245}	^{272}STV**R**^{275}	^{344}W**S**T**G**PHTC**P**A^{353}
CYP156B1	SCIF3.12	^{248}AGAD**P**^{252}	^{269}**E**YA**R**^{272}	^{344}W**S**A**G**PHAC**P**S^{353}
CYP157A1	SC6D11.14c	^{240}AGHQ**P**^{244}	^{279}**E**VL**W**^{282}	^{342}F**G**H**G**EHRC**P**F^{351}
CYP157B1	SCF55.08c	^{255}AQQ**P**T^{259}	^{293}**E**VL**W**^{296}	^{355}F**S**N**G**EHRC**P**Y^{364}
CYP157C1	SCI41.09c	^{259}AAYEA^{263}	^{297}**E**QSL**W**^{301}	^{361}F**G**G**G**PHEC**P**G^{370}
CYP158A1	SC8F11.24c	^{245}GGEAV^{249}	^{283}**E**LL**R**^{286}	^{349}Y**G**N**G**HHFC**T**G^{358}
CYP158A2	2SCG58.07	^{242}GGEAV^{246}	^{280}**E**LL**R**^{283}	^{346}F**G**F**G**PHYC**P**G^{355}
CYP159A1	SCF55.07	^{234}AGGE**T**^{238}	^{273}**E**TL**R**^{276}	^{347}FAL**G**RHFC**V**G^{356}

[1] CYP names as annotated at website drnelson.utmem.edu/CytochromeP450.html.

[2] Gene name as annotated at website www.sanger.ac.uk/Projects/S_coelicolor/.

indicated in the table legends. CYPs were detected in homology searches (BLAST and Pfam) using the conserved heme-binding domain signature and were identified also at ScoDB using a similar approach. Other features conserved in CYPs were identified by amino acid alignment and included visual assessment with reference to the known features of the superfamily. Phylogenetic trees were calculated using ClustalX and TreeView. The names of genes are given in their *S. coelicolor* A3(2) order in the cosmids sequenced and available at ScoDB (jic-bioinfo.bbsrc.ac.uk/streptomyces/). CYP family and subfamily assignments were made by Dr. David Nelson according to the P450 nomenclature (8, 9). In this, >40% identity places a CYP in the same family (unless homologues with lower identity are demonstrated to be functionally equivalent) and >55% places it in the same subfamily (*e.g.* CYP154A1, CYP154C1, etc).

General Methods—Sodium dithionite-reduced carbon monoxide difference spectra for quantification of cytochrome P450 contents were measured and calculated according to the method described by Omura and Sato (21). Similar spectra obtained using ferredoxin and ferredoxin reductase proteins to reduce CYP used the electron donor system method described previously for CYP51 from *M. tuberculosis* (20). For this system proteins were first purified using Ni^{2+}-NTA chromatography. Protein quantification was performed by using the Bicinchoninic acid assay. Unless otherwise stated all chemicals were supplied by Sigma.

RESULTS AND DISCUSSION

The cytochrome P450 superfamily has members present in all kingdoms of life, although they can be highly diverse in activity and function. Their role in secondary metabolism is well established as is likely reflected in the ~275 CYPs observed in *A. thaliana.* The genome of *S. coelicolor* A3(2) is of applied interest as the model for the category of microbial organisms producing the highest number of medicinal natural products in use today. The primary amino acid sequence of CYPs is highly variable, but secondary and tertiary structures are generally conserved as revealed by the known structures.

In *S. coelicolor* A3(2) 18 CYPs were detected through continued analysis of the progressing genome-sequencing project, which was recently completed. Detection of probable CYPs within a genome is facilitated by the presence of consensus sequences within the primary structure that are known to be present in all CYPs. Primary sequence motifs are important to the secondary and tertiary structure with three amino acids conserved throughout the superfamily. These are E*XX*R in the K-helix and the cysteine residue, which forms a fifth axial ligand to the heme iron. Two other very highly conserved residues in CYPs are two glycine residues, four amino acids upstream of the conserved cysteine, and a threonine residue in the I-helix, which is proposed to be involved in oxygen activation. The sequences near these highly conserved amino acids are sometimes conserved within CYP gene families, but variability can also be accommodated.

TABLE II

The oligonucleotide sequences (forward and reverse) used to isolate each CYP gene, the expression yields, and the spectral maxima found by reduced carbon monoxide difference assay of cytosolic fractions of E. coli

CYP name[1]	Oligonucleotide sequences (forward-top; reverse-below)	Yield (nmol/L)	Spectral maximum (nm)
CYP51	5′ CGCCATATGACCGTCGAGTCCGTCAAC 3′	<10	445-450
	5′ GCTAAGCTTTCAGTGATGGTGATGCCGCGCCCACGGGCCTGAC 3′		
CYP102B1	5′ CGCCATATGGCCCAGACAGCGAGGGAA 3′	<10	448
	5′ GCTAAGCTTTCAGTGATGGTGATGGTCACCCGCCCGGTGCAC 3′		
CYP105D5	5′ CGCCATATGACGGACACCGACACGACG 3′	~6500	448
	5′ GCTAAGCTTCTAGTGATGGTGATGCCAGGCCACGGGGAGTTC 3′		
CYP105N1	5′ CGCCATATGACACCCCCCGAATCCCCG 3′	<10	448
	5′ GCTAAGCTTTCAGTGATGGTGATGCCAGGTCACCATCAGTTC 3′		
CYP107P1	5′ CGCCATATGACGGCTGCAACCGATGGTCCC 3′	~200	450
	5′ GCTAAGCTTTCAGTGATGGTGATGGAGTTCCACGCGTAGCTC 3′		
CYP107T1	5′ CGCCATATGGGGAGTGCTCCCCGTTTAATT 3′	~50	447
	5′ GCTAAGCTTTCAGTGATGGTGATGGCCGAGCCGCACCGGCAG 3′		
CYP107U1	5′ CGCCATATGACCGGCAGCTCTTCCGCCC 3′	<50	448
	5′ GCTAAGCTTTCAGTGATGGTGATGCTTTTGCGTCGGCGACGG 3′		
CYP154A1	5′ CGCCATATGGCGACCCAGCAGCCCGCC 3′	~6500	448
	5′ GCTAAGCTTTCAGTGATGGTGATGGCCGGCGTGCAGCAGGAC 3′		
CYP154C1	5′ CGCCATATGACGACCGGCACCGAAGAAGCC 3′	~6500	448
	5′ GCTAAGCTTTCAGTGATGGTGATGGGCCAGCCTGACCGGCAG 3′		
CYP155A1	5′ CGCCATATGGCTGATCCTGGTGACCACGTC 3′	~250	448
	5′ GCTAAGCTTTCAGTGATGGTGATGGCCGCGGCCGGCCCGGGG 3′		
CYP156A1	5′ CGCCATATGACGTTGCCCTCCACCGAGACC 3′	~50	450
	5′ GCTAAGCTTTTAGTGATGGTGATGGTGTGCGGCGGATTGCGC 3′		
CYP156B1	5′ CGCCATATGGACGCCACCACCCCCGCG 3′	~50	450
	5′ GCTAAGCTTTCAGTGATGGTGATGCCTCGTGAGCCAGTTGAG 3′		
CYP157A1	5′ CGCCATATGAGCACCGACGCCCACGACG 3′	~50	448
	5′ GCTAAGCTTTCAGTGATGGTGATGTGCTGACGTGCCTCCCAG 3′		
CYP157B1	5′ CGCCATATGACCGACATCGACCCGTCACC 3′	<10	447
	5′ GCTAAGCTTTCAGTGATGGTGATGGGCCACGACGGGTGTGAA 3′		
CYP157C1	5′ CGCCATATGACGCCTGAAAGCCACTCCCC 3′	<10	448
	5′ GCTAAGCTTTCAGTGATGGTGATGGTAGCCGCGCCACCAGAG 3′		
CYP158A1	5′ CGCCATATGACGCAGGAGACCACCACG 3′	~350	450
	5′ GCTAAGCTTCTAGTGATGGTGATGCCAGGTGCAGGGCAG 3′		
CYP158A2	5′ CGCCATATGACTGAAGAAACGATTTCC 3′	~500	448
	5′ GCTAAGCTTTCAGTGATGGTGATGCCACGTCACCGGCAG 3′		
CYP159A1	5′ CGCCATATGTCCACCGCGCAGCAGGTCC 3′	~250	450
	5′ GCTAAGCTTTCAGTGATGGTGATGGGCCGCGGGTGTGAA 3′		

[1] CYP names as annotated at website drnelson.utmem.edu/CytochromeP450.html.

Table I shows the motifs selected as most closely corresponding to these conserved CYP expectations; the genes were detected using the consensus heme-binding domain and known variants to do a protein motif search in the *S. coelicolor* A3(2) chromosomal sequence data base (www.sanger.ac.uk/Projects/S_coelicolor/). Some of the CYP sequences contain variant sequences, particularly in the E*XX*R domain in the K-helix, and further investigation of the basis and relevance of these alterations is underway experimentally. All of these CYPs were subject to heterologous expression, and this confirmed their identity as cytochromes P450. All CYPs were soluble, locating to the cytosolic fraction on expression in *E. coli* with reduced carbon monoxide difference spectra of soluble protein giving maxima located around 448 nm (Table II). Despite the high G and C content of *S. coelicolor* A3(2) genome expression levels were often extremely high, but five proved difficult to express, often with low and intermittent yield. CYP51 and CYP107U1 sometimes gave variable spectral maxima toward 440 nm in different preparations. Four genes had valine triplets at the 5′ of the putative open reading frame, and these were replaced with ATG in the engineered genes used for expression in *E. coli*. The identification of heterologous expression of these CYPs supports the correct identification of the initiation sites. The importance of rare *E. coli* codons in the *S. coelicolor* A3(2)

TABLE III
Putative ferredoxin reductases and ferredoxins in Streptomyces coelicolor

Gene name[1]	Amino acids, no.	Predicted function	Match in databases[2]				
			Species	Gene name	Protein identifier	% identity	Amino acids overlap
SCF15.02	454	Ferredoxin-NADP-reductase	*Deinococcus radiodurans*	DR0496	Q9RX19	44.3	449
			Mycobacterium tuberculosis	Rv3106	O05783	40.9	460
SC4B10.18c	420	Ferredoxin reductase	*Mycobacterium tuberculosis*	Rv1869c	P95146	51.2	410
			Nocardioides sp.	PhdD	Q9R6Z4	40.0	420
SC7A8.08c	421	Probable ferredoxin-reductase	*Mycobacterium tuberculosis*	Rv1869c	P95146	50.7	414
			Streptomyces coelicolor	SC4B10.18c	Q9FC69	50.8	415
SC4B10.11	129	Ferredoxin, fdxA	*Streptomyces coelicolor*	SC9E12.20	gi|9967627	66.7	105
			Streptomyces griseus	7Fe	P13279	64.4	104
SC9E12.20	106	Ferredoxin, fdxA1	*Streptomyces griseus*	7Fe	P13279	98.1	104
			Streptomyces erythraeus	fdxA	P24496	78.3	106
SC4C2.11	73	Ferredoxin	*Streptomyces griseolus*	FD-2	P18325	55.7	61
			Streptomyces natalensis	PimF	Q9EW93	55.7	61
3SCF60.05c	66	Ferredoxin	*Streptomyces lividans*	ORF3	gi|3293538	98.5	66
			Streptomyces griseus	SoyB	P26910	61.9	63
SCH18.04c	70	Ferredoxin	*Streptomyces noursei*	NysM	Q9L4W9	48.4	62
			Streptomyces coelicolor	3SCF60.05c	Q9EWS5	45.2	62
SCI41.32	101	Ferredoxin	*Bacillus halodurans*	BH1605	Q9KCG6	34.8	69
			Bacillus stearothermophilus	FER	P00212	33.3	69

[1] As annotated at www.sanger.ac.uk/Projects/S_coelicolor/.
[2] Database search at FASTA website fasta.bioch.virginia.edu/.

CYPs may contribute to low level expression, but these codons were also observed in highly expressing CYPs such as CYP105D5. Examples of the reduced carbon monoxide difference spectra of highly expressed and lowly expressed CYP forms are presented in Fig. 1.

Table III gives the potential redox partners for CYPs detectable in the genome. Interestingly, a ferredoxin (3SCF60.05c) is probably associated in an operon with *CYP105D5* whose most well known homologue is *CYP105D1* from *Streptomyces griseus* (69.5% amino acid identity) (22) and has even stronger identity with *S. lividans CYP105D4* (>90% amino acid identity). Three ferredoxin reductase homologues were identified in the genome and were evaluated for their role in support of the CYP proteins. The production and purification of these putative redox partners and the ferredoxin product of 3SCF60.05c was undertaken using recombinant methods as for CYPs. Reduction of CYP was confirmed in the case of CYP105D5 and CYP105N1 as shown by the generation of carbon monoxide-reduced difference spectra measured after 10 min in the presence of cofactor and compared with that obtained using sodium dithionite reduction. Using SC4B10.18c protein and NADH, we observed 90 and 70% of the reduction observed using dithionite for CYP105D5 and CYP105N1, respectively. Using SCF15.02 protein and NADPH, 80% reduction of CYP105D5 and 16% reduction of CYP105N1 was observed while for SCF15.02 protein and NADH only <10% reduction of CYP105D5 and 20% reduction of CYP105N1 occurred. This showed that SC7A8.08c was the least effective ferredoxin reductase partner. In the complete evaluation of redox partners for *S. coelicolor* CYPs the knowledge of endogenous substrate is a prerequisite for identification of reduction partners and kinetics for this matrix of combinations. Natural electron transfer from reductase in the CYP catalytic cycle occurs to substrate-bound enzyme.

A phylogenetic tree was constructed of the *S. coelicolor* A3(2) CYPome in conjunction with the CYPome of *M. tuberculosis*, another closely related actinomycete (Fig. 2). The *M. tuberculosis* genome contains 20 CYPs in ~3800 predicted ORFs in contrast to the 18 CYPs in ~7500 predicted ORFs of *S. coelicolor* A3(2). One clan of five CYPs consists entirely of *S. coelicolor* A3(2) proteins; other groups are mixed or predominantly from one genome or the other. No striking level of conserved homology across the CYPome was observed, which probably reflects the divergence of these bacterial species into their different niches. The major classes of the actinomycetes have been estimated to have originated in the last 500 million years (23). Possibly the environmental soil organism, *S. coelicolor* A3(2), has need of detoxifying, and secondary metabolism resulting in the evolution of a distinct CYPome from *M. tuberculosis* where the CYPome may also be closely involved in synthesis of its complex lipid wall.

Table IV shows information on each CYP concerning the number of amino acids and homology to other CYPs. Some show moderate to high level homologies to CYPs with known activities, others less so. One of the low-level homologies is a CYP51-like protein (SC 7E4.20). CYP51 is the only CYP found in bacteria, plants, fungi, protists, and animals. In bacteria it was first shown in *M. tuberculosis* (20). Phylogenetically, it is located in a cluster of CYPs associated with likely lipid metabolism roles including sterol 14-demethylase of *M. tuberculosis* (Rv0764c) and SC F43.12 (*CYP102B1*). CYP102B1 is the longest bacterial CYP encountered so far at 527 amino acids.

CYP105D5 and surrounding genes had the closest homology to the *S. lividans* AUD4 gene cluster (24). A *S. griseus* protein associated with the metabolism of a vast array of xenobiotics is also encoded by a homologue with high identity (16, 25). As in both species the *CYP105D5* is adjacent to a ferredoxin but not a ferredoxin reductase. This CYP is particularly interesting from the viewpoint of a detoxifying role in the soil environment where the bacteria may be exposed to toxic secondary metabolites from live organisms and decaying organic material. Furthermore, CYP105 proteins are important in biotransformations and bioremediation, including their use in manipulating plant herbicide tolerance (16, 26). *CYP107U1*, *CYP107P1*, and *CYP107T1* are three CYP107 members in the genome. The first has homology to CYP113 as well as the latter to cytochrome P450 EryF, a known enzyme of antibiotic production (27). These proteins may also participate in secondary metabolism.

CYP155A1 has low-level homology to other CYPs. *CYP156A1* and *CYP154A1* are located adjacent on the same strand with only a 12-bp separation. This points to a related function with

polycistronic translation. Homology to CYPs with known function, such as in tylosin biosynthesis (28), points to an important secondary metabolism function. *CYP157A1* and *CYP154C1* also lie adjacent on the chromosome and may reflect a gene duplication event and a putative operon involved in secondary metabolism. Within the operon is a sensor kinase, an ATP binding protein, as well as two unknown ORFs prior to CYP. *CYP156A1*, *CYP157B1*, and *CYP157C1* also appear in operons with the same upstream gene order.[2] Many of the CYPs have closest identity with each other.

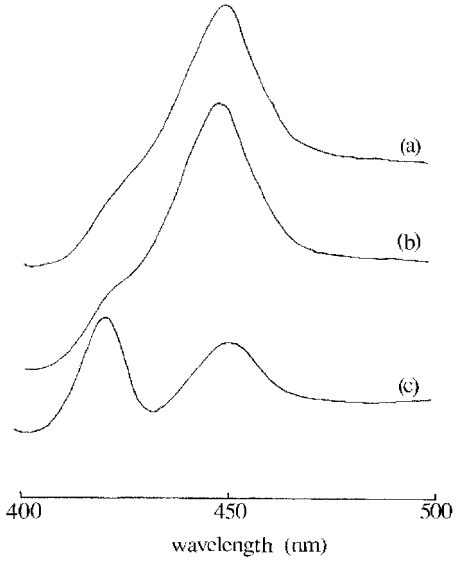

FIG. 1. **Spectroscopic comparison of CYPs expressing at different levels.** Reduced carbon monoxide spectra of different CYP forms expressing at high, intermediate, and low levels respectively for CYP154C1 (*a*), CYP158A1 (*b*), and CYP105N1 (*c*).

CYP159A1 and *CYP157B1* are also adjacent genes but on opposite strands. Their homologies indicate different functions for each. *CYP157C1* has homology to known CYP proteins that are orphans with little clue as to function. CYP158A2 has 81.4% identity with a CYP from *Streptomyces avermitilis* (29), and also the genetic locus is similar on the *S. avermitilis* chromosome. That includes the adjacent polyketide synthase suggesting a role in secondary metabolism. CYP158A1 also has 58.7% identity with this *S. avermitilis* CYP. *CYP105N1* is again similar to other CYPs involved in xenobiotic metabolism but is located in a likely operon of secondary metabolism including probable non-ribosomal peptide synthases, a probable thioesterase, and probable ABC transporter.

CONCLUSIONS

The identification of 18 novel CYPs in the genome of *S. coelicolor* A3(2) represents the first systematic production of a CYPome of an organism for functional genomic/metabolomic characterization. There are ~2000 CYP genes known with

[2] D. Hopwood, personal communication.

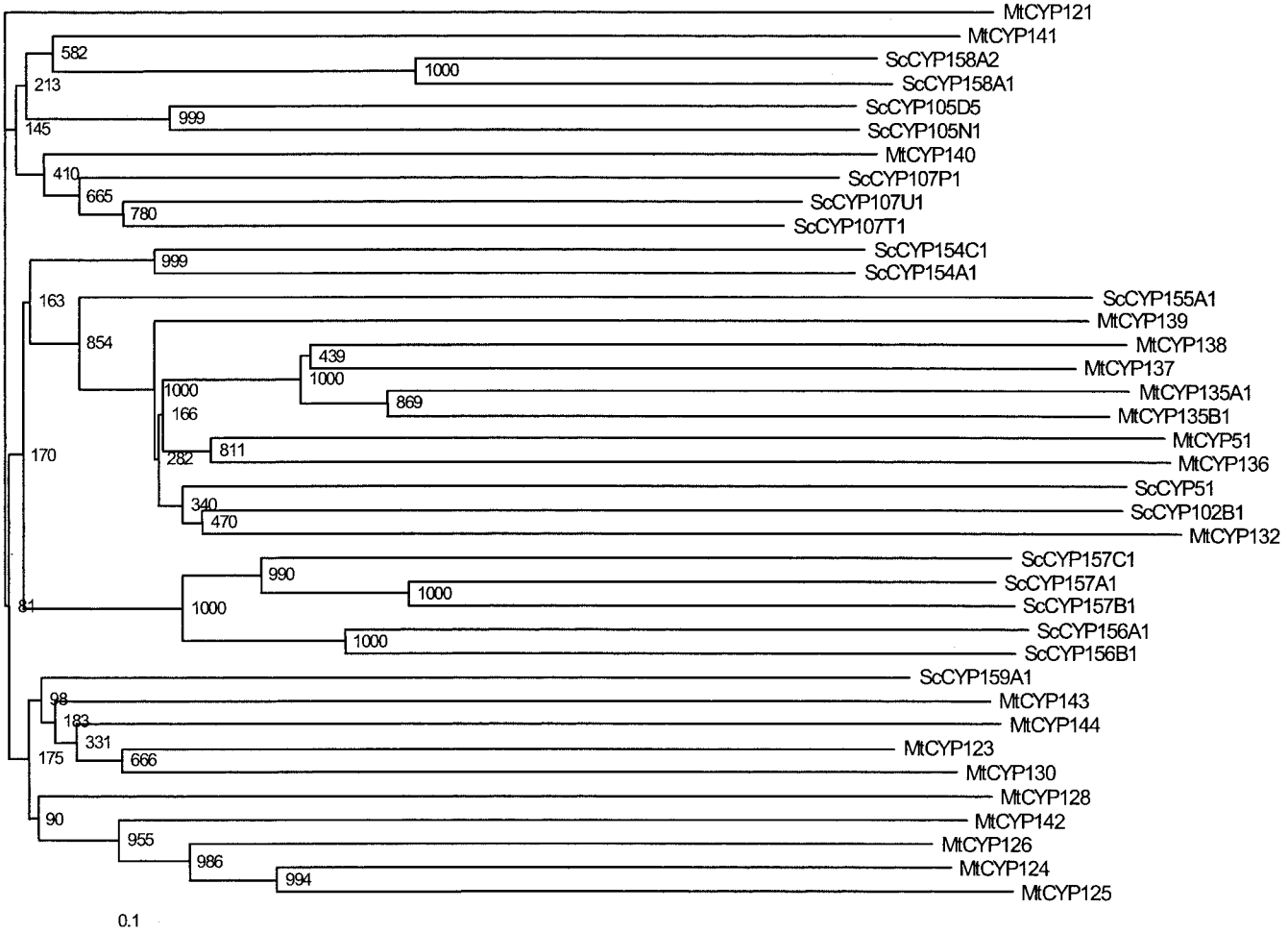

FIG. 2. **Relationship of the CYPomes of *S. coelicolor* A3(2) with *M. tuberculosis*.** Phylogenetic tree of the CYPome of *S. coelicolor* A3(2) (18 ScCYPs) and *M. tuberculosis* (20 MtCYPs) shown together and produced using ClustalX and TreeView. *Numbers* indicate the bootstrap probability values for the branch topology shown.

TABLE IV
The CYPs of Streptomyces coelicolor with their closest homologues

CYP name[1]	Gene name[2]	Amino acid, no.	Match in the databases[3]				
			Species	CYP name	Protein identifier	% identity	Amino acid overlap
CYP51	SC7E4.20	461	*Streptomyces avermitilis*	CYP171A1	Q9S0R5	32.4	429
			Mycobacterium tuberculosis	CYP51	P77901	26.8	421
CYP102B1	SC F43.12	527	*Bacillus megaterium*	CYP102A1	P14779	42.8	488
			Bacillus subtilis	-	gi|16077792	42.2	502
CYP105D5	3SCF60.06c	412	*Streptomyces lividans*	CYP105D4	O85697	97.5	405
			Streptomyces griseus	CYP105D1	P26911	67.5	412
CYP105N1	SC4C2.21	411	*Streptomyces griseolus*	CYP105A1	P18326	45.5	407
			Streptomyces lividans	CYP105D4	O85697	44.3	411
CYP107P1	SCH10.14c	411	*Streptomyces coelicolor*	CYP157C1	Q9RJ75	40.3	426
			Streptomyces erythraeus	CYP107B1	P33271	39.2	406
CYP107T1	SCH63.17	394	*Streptomyces venezuelae*	CYP107L1	O87605	51.5	394
			Streptomyces erythraeus	CYP107B1	P33271	46.9	373
CYP107U1	SCE41.08c	433	*Streptomyces coelicolor*	CYP107P1	Q9X8Q3	40.4	426
			Actinomadura hibisca	CYP107M1	O32460	40.0	434
CYP154A1	SCE6.21	408	*Streptomyces fradiae*	CYP154B1	Q9XCC6	50.9	409
			Streptomyces coelicolor	CYP154C1	Q9L142	42.1	411
CYP154C1	SC6D11.13c	407	*Streptomyces coelicolor*	CYP154A1	Q9KZR7	42.1	411
			Streptomyces erythraeus	CYP107B1	P33271	38.4	409
CYP155A1	SC6D11.40c	423	*Brevibacterium linens*	-	Q9KK77	38.8	408
			Streptomyces tendae	CYP162A1	Q9L465	27.9	394
CYP156A1	SCE6.20	410	*Streptomyces coelicolor*	CYP156B1	gi|13872774	43.2	410
			Streptomyces griseus	ORF5	Q9AJP0	31.3	403
CYP156B1	SCIF3.12	447	*Streptomyces coelicolor*	CYP156A1	Q9KZR8	43.2	410
			Streptomyces griseus	ORF5	Q9AJP0	34.7	452
CYP157A1	SC6D11.14c	413	*Streptomyces coelicolor*	CYP157B1	Q9RJQ6	51.7	404
			Streptomyces griseus	ORF5	Q9AJP0	41.4	406
CYP157B1	SCF55.08c	420	*Streptomyces coelicolor*	CYP157A1	Q9L141	51.7	404
			Streptomyces griseus	ORF5	Q9AJP0	41.3	421
CYP157C1	SCI41.09c	498	*Streptomyces griseus*	ORF5	Q9AJP0	50.8	519
			Streptomyces coelicolor	CYP157A1	Q9L141	39.8	400
CYP158A1	SC8F11.24c	407	*Streptomyces coelicolor*	CYP158A2	Q9FCA6	61.2	407
			Streptomyces avermitilis	-	gi|15824133	58.7	407
CYP158A2	2SCG58.07	404	*Streptomyces avermitilis*	-	gi|15824133	81.4	404
			Streptomyces coelicolor	CYP158A1	Q9KZF5	61.2	407
CYP159A1	SCF55.07	407	*Bacillus subtilis*	CYPX	O34926	41.3	407
			Mycobacterium tuberculosis	CYP123	P77902	33.3	403

[1] CYP names as annotated at website drnelson.utmem.edu/CytochromeP450.html.
[2] Gene names as annotated at the Sanger Centre website www.sanger.ac.uk/Projects/S_coelicolor/.
[3] Data base search at FASTA website fasta.bioch.virginia.edu/.

others appearing from almost every known genome sequence. The vast majority of these are of unknown function, and assigning function is the challenge for this and many other projects. However, the information on the *S. coelicolor* A3(2) CYPome has allowed progress to structural biology studies and proteomics, and analyzing mutant metabolomes represents the best way forward toward identifying function.

Acknowledgments—We are grateful for comments and advice from Prof. David Hopwood and Dr. David Nelson.

REFERENCES

1. Guengerich, F. P. (1991) *J. Biol. Chem.* **266,** 10019–10022
2. Porter, T. D., and Coon, M. J. (1991) *J. Biol. Chem.* **266,** 13469–13472
3. Ortiz de Montellano, P. R. (ed) (1995) *Cytochrome P450*, 2nd Ed., Plenum Publishing Corp., New York
4. Waterman, M. R., and Johnson, E. F. (eds) (1991) *Methods in Enzymology*, Vol. 206, Academic Press, San Diego, CA
5. Schenkman, J. B., and Kupfer, D. (eds) (1982) in *Hepatic Cytochrome P450 Monooxygenase System*, Pergamon Press, New York
6. Nelson, D. R. (1999) *Arch. Biochem. Biophys.* **369,** 1–10
7. Kelly S. L., Lamb D. C., Cannieux, M., Greetham D., Jackson C. J., Marczylo T., Ugochukwu C., and Kelly D. E. (2001) *Biochem. Soc. Trans.* **29,** 122–128
8. Nelson, D. R., Koymans, L., Kamataki, T., Stegman, J. J., Feyereisen, R., Waxman, D. J., Waterman, M. R., Gotoh, O., Coon, M. J., Estabrook, R. W., Gunsalus, I. C., and Nebert, D. W. (1996) *Pharmacogenetics* **6,** 1–42
9. Nelson, D. R., Kamataki, T., Waxman, D. J., Guengerich, F. P., Estabrook, R. W., Feyereisen, R., Gonzalez, F. J., Coon, M. J., Gunsalus, I. C., Gotoh, O., Okuda, K., and Nebert, D. W. (1993) *DNA Cell Biol.* **12,** 1–51
10. Takemori, S., Yamazaki, T., and Ikushiro, S. (1993) In *Cytochrome P450*, (Omura, T., Ishimura, Y., and Fujii-Kuriyama, Y., eds) 2nd Ed., VCH Publishers, Inc., New York
11. Hopwood, D. A. (1999) *Microbiology* **145,** 2183–2202
12. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F., and Hopwood, D. A. (2000) *Practical Streptomyces Genetics* The John Innes Foundation, Norwich, UK
13. Sariaslani, F. S., Eckenrode, F. M., Beale, J. M., and Rosazza, J. P. (1984) *J. Med. Chem.* **27,** 749–754
14. Sariaslani, F. S., and Rosazza, J. P. (1983) *Appl. Environ. Microbiol.* **46,** 468–474
15. Sariaslani, F. S., and Rosazza, J. P. (1983) *Appl. Environ. Microbiol.* **49,** 451–452
16. Taylor, M., Lamb, D. C., Cannell, R., Dawson, M., and Kelly, S. L. (1999) *Biochem. Biophys. Res. Commun.* **263,** 838–842
17. Serizawa, N., and Matsuoka, T. (1991) *Biochim. Biophys. Acta* **1084,** 35–40
18. Berrie, J. R., Williams, R. A. D., and Smith, K. E. (1999) *J. Steroid Biochem.* **71,** 153–165
19. Cannell R. J. P., Knaggs A. R., Dawson M. J., Manchee G. R., Eddershaw P. J.,

Waterhouse I., Sutherland D. R., Bowers G. D., and Sidebottom P. J. (1995) *Drug. Metab. Disp.* **23,** 724–729
20. Bellamine, A., Mangla, A. T., Nes, W. D., and Waterman, M. R. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 8937–8942
21. Omura, T., and Sato, R. (1964) *J. Biol. Chem.* **239,** 2370–2378
22. Trower, M. K., Lenstra, R., Omer, C., Bucholz, S. E., and Sariaslani, F. S. (1992) *Mol. Microbiol.* **6,** 2125–2134
23. Embley, T. M., and Stackebrandt, E. (1994) *Ann. Rev. Microbiol.* **48,** 257–289
24. Schmid, E., Buchler, C., and Altenbuchner, J. (1999) *Microbiol.* **145,** 3331–3341
25. Trower, M. K., Sariaslani, F. S., and Kitson F. G. (1988) *Biochem. Biophys. Res. Commun.* **157,** 1417–1422
26. O'Keefe, D. P., Tepperman, J. M., Dean, C., Leto, K. J., Erbes, D. L., and O'Dell, J. T. (1994) *Plant Physiol.* **105,** 473–482
27. Cuppvickery, J. R., and Poulos, T. L. (1995) *Nat. Struct. Biol.* **2,** 144–153
28. Merson-Davies, L. A., and Cundliffe, E. (1994) *Mol. Microbiol.* **13,** 349–355
29. 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