Available online at www.sciencedirect.com
SCIENCE DIRECT®
Archives of Biochemistry and Biophysics 425 (2004) 233–241
ABB
www.elsevier.com/locate/yabbi

Genome analyses of *Streptomyces peucetius* ATCC 27952 for the identification and comparison of cytochrome P450 complement with other *Streptomyces*

Niranjan Parajuli, Devi B. Basnet, Hei Chan Lee, Jae Kyung Sohng, and Kwangkyoung Liou*

Institute of Biomolecule Reconstruction, Sun Moon University, Asan, 336-708, Republic of Korea

Received 5 February 2004, and in revised form 4 March 2004

Abstract

We have determined the genome sequence of 8.7 Mb chromosome of *Streptomyces peucetius* ATCC 27952, which produces clinically important anthracycline chemotherapeutic agents of the polyketide class of antibiotics, daunorubicin and doxorubicin. The cytochrome P450 (CYP) superfamily is represented by 19 sequences in the *S. peucetius*. Among those, 15 code for functional genes, whereas the remaining four are pseudo genes. CYPs from *S. peucetius* are phylogenetically close to those of *Streptomyces amermitilis*. Four CYPs are associated with modular PKS of avermectin and two with doxorubicin biosynthetic gene cluster. CYP252A1 is the new family found in *S. peucetius*, which shares 38% identity to CYP51 from *Streptomyces coelicolor* A3 (2). Nine CYPs from *S. peucetius* are found in the cluster containing various regulatory genes including *rar* operon, conserved in *S. coelicolor* A3 (2) and *Streptomyces griseus*. Although two ferredoxins and four ferredoxin reductases have been identified so far, only one ferredoxin reductase was found in the cluster of *CYP147F1* in *S. peucetius.* To date, 174 CYPs have been described from 45 *Streptomyces* species in all searchable databases. However, only 18 CYPs are clustered with ferredoxin. The comparative study of cytochrome P450s, ferredoxins, and ferredoxin reductases should be useful for the future development and manipulation of antibiotic biosynthetic pathways.

Keywords: Cytochrome P450; Ferredoxin; Ferredoxin reductase; Secondary metabolites; *Streptomyces*

Cytochrome P450[1] enzymes belonging to the superfamily of genus exist in all eukaryotic organisms, animals, plants, fungi, and microorganisms, and have evolved a function in the initial oxidation of xenobiotics. About 40% (174) of all known bacterial CYPs are found in 45 different species of the industrially important genus *Streptomyces*. In contrast to membrane bound CYPs found in eukaryotes, most of the CYPs from *Streptomyces* are soluble, but only a few of them have been characterized. However, no CYP has been reported in the genomes of several bacterial species examined, including *Escherichia coli* [1]. CYP does not seem to be essential for the basic metabolism of most prokaryotes, but some of them are involved in catabolism of hydrocarbons, terpenes and others, and provide the sole carbon and energy sources for bacteria [2]. They are also involved in the oxidative, peroxidative, and reductive metabolisms of various endogenous compounds such as steroids, bile acids, fatty acids, prostaglandins, leukotrienes, biogenic amines, and other secondary metabolites.

The main function of CYP is the monooxygenation of various substrates. This requires molecular oxygen and the supply of reducing equivalents from NADPH or NADH. However, most of the bacterial CYPs receive electrons from NADH [2]. They are able to introduce atomic oxygen into allylic positions, double bonds, or even into non-activated C–H bonds. They encode a superfamily of heme-thiolate containing enzymes often

* Corresponding author. Fax: +82-41544-2919.
E-mail address: liou@sunmoon.ac.kr (K. Liou).

[1] *Abbreviations used:* CYP, cytochrome P450; ORF, open reading frame; fdx, ferredoxin; fdr, ferredoxin reductase.

0003-9861/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.abb.2004.03.011

located in macrolide antibiotic biosynthetic gene clusters, where they catalyze stereo- and region-specific oxidation of precursors, leading to structural diversity within these molecules [3]. CYP107A1 (EryF) is responsible for hydroxylation of 6-deoxyerythronolide B to erythronolide B in the biosynthesis of erythromycin from *Saccharopolyspora erythraea* [4]. Macrolactones, YC-17 and narbomycin, are hydroxylated by CYP107L1 (PikC) involved in pikromycin biosynthesis from *Streptomyces venezuelae* [5]. Tylosin biosynthetic gene cluster includes *CYP113B1* and *CYP105L1* in *Streptomyces fradiae* [6,7]. CYP107D1, encoded *oleP*, could be responsible for the epoxidation of C-8 in the oleandomycin lactone ring [8]. Two cytochromes, CYP161A3 (AmphL) and CYP105H4 (AmphN), probably carry out post-polyketide modifications during biosynthesis of amphotericin from *Streptomyces nodosus* [9]. CYP161A3 also complements the CYP161A1 (NysL) involved in nystatin biosynthesis from *Streptomyces noursei*. C-6 and C-8a of avermectin algycone biosynthesis from *Streptomyces avermitilis* are hydroxylated by CYP171A1 (AveE) to form a furan ring [10]. CYP122A2 (RapJ) and CYP107G1 (RapN) are possibly involved in hydroxylation of the macrolide at C-9 and C-27 or C-26 and C-32 of rapamycin in *Streptomyces hygroscopicus* [11]. CYP129A2 and CYP131A1 are components of the anthracycline class of antibiotic, doxorubicin biosynthetic gene clusters from *Streptomyces peucetius* [12,13]. Other members of the same subfamily (CYP129A1 and CYP131A12) participate in daunorubicin biosynthesis from *Streptomyces* sp. strain C5 [14]. Glycopeptide antibiotic, complestatin from *Streptomyces lavendulae* NRRL 2564 also includes the two CYPs: CYP165B5 and CYP165E1 [15]. In *Streptomyces griseolus*, herbicide metabolism requires CYP105A1 and CYP105B1 [16]. *CYP211A1* is clustered in enediyne antitumor antibiotic C-1027 from *Streptomyces globisporus* [17]. Ansamycin type antibiotic, geldanamycin biosynthetic gene cluster from *S. hygroscopicus*, involves *CYP105U1* [18]. CYP162A1 (NikQ) from *Streptomyces tendae* [19] and CYP105K2 (SanL) and SanQ from *Streptomyces ansochromogenes* are associated with biosynthesis of nucleoside antibiotic, nikkomycin [20]. Two cytochromes of the same subfamily, CYP163A1 (NovI) and CYP163A2 (CloI), are involved in coumarin formation in aminocoumarin antibiotics such as novobiocin from *Streptomyces spheroids* and clorobiocin from *Streptomyces roseochromogenes* [21]. Another aminocoumarin antibiotic, simocyclinone biosynthetic gene clusters from *Streptomyces antibioticus* Tü 6040, was cloned with a heterologous probe from a gene encoding CYP163A1 [22]. The existence of CYP105M1 (Orf10) in biosynthesis of the β-lactum antibiotic "clavulanic acid" from *Streptomyces clavuligerus* has also been reported [23]. Therefore, one may conclude that CYPs are clustered with biosynthesis of various kinds of antibiotics except aminoglycosides and deoxysugar biosynthesis in *Streptomyces* (Table 1).

Although substantial knowledge has been obtained about various CYPs from plants, animals, and microorganisms over the past four decades, a number of major questions remain regarding the structure, function, and mechanism of these enzymes. Further investigations on CYPs in aminoglycosides are essential for the development of new hybrid antibiotics. An approach to produce deoxynitro sugars would facilitate the discovery of new antibiotics, and it is well known that this could be achieved only a novel CYP. Therefore, a complete understanding of the CYP superfamily of *Streptomyces* is essential for the development of novel secondary metabolites. The fully sequenced genome of *Streptomyces coelicolor* A3 (2) revealed 7825 ORFs with 18 CYPs genes in the linear 8.7 Mb chromosome [24]. Another fully sequenced genome of *S. avermitilis* contained approximately 7600 ORFs with 33 CYPs within the 9 Mb chromosome [3]. Database search reveals that *Streptomyces* sp. and *Streptomyces tubercidicus* have 16 and 12 CYPs, respectively.

We have sequenced the whole genome (i.e., 8.7 Mb chromosome of *S. peucetius* ATCC 27952) and it will be disclosed in the near future. Reported here are the 19 putative CYPs including DoxA and DnrQ, which are previously known. The classification of CYPs and identification of ferredoxins and ferredoxin reductases in the *S. peucetius* genome are also discussed.

Experimental procedures

Total genome sequencing

Streptomyces peucetius ATCC 27952 was cultured in TSB medium at 28 °C for 2 days for isolation of genomic DNA [25]. pSuperCosI was used to construct a genomic library by digestion of genomic DNA with *Sau*3AI. The ligated DNA was packaged in vitro using Gigapack III XL packaging extract (Stratagene). The genomic library of *S. peucetius* was screened, and several corresponding cosmids were isolated by internal sequences of *dnrF* (5′-AGG TTT GAG GTG GCC TTG ACG-3′ and 5′-TCC GCG TCA GTT CGC CGG AGG-3′) and *dpsY* (5′-GGA CTG CCG GTG TGC TGT GGT-3′ and 5′-CCG GAA CGT TCA TTC GTC GAC-3′). The total genome of *S. percetius* has been sequenced using various cosmids, and the shotgun approach [26] generated 2–4 kb fragments of the genome. The all non-redundant fragments were then assembled by PHRED [27] and PHRAP (http://www.phrap.org). An additional genomic library was also prepared to assist the whole genomic sequencing.

Table 1
Cytochrome P450s involved in secondary metabolites production in *Streptomyces*

Species	CYP name[a]	Gene name[b]	Secondary metabolites	Applications
S. lavendulae	CYP107N1	*orf3*	MitomycinC	AntiHIV
	CYP160A1	*mmcN*		
	CYP105F1	*orf4*		
	CYP165B5[c]	*com02*		
	CYP165E1	*com01*	Complestatin	
S. antibioticus	CYP107D1	*oleP*	Oleandomycin	Antibacterial
	CYP235A1	*oleP1*		
	CYP163A3	*simI*	Simocyclinone	Antifungal
S. avermitilis	CYP105P1	*pteC*	Filipin	Antifungal
	CYP105D6[c]	*pteD*		
	CYP171A1	*aveE*	Avermectin	Antiparasitic
S. fradiae	CYP113B1	*orf1*	Tylosin	Antibacterial
	CYP105L1[c]	*tylH1*		
	CYP154B1	*orf16*		
S. hygroscopicus	CYP122A2	*rapJ*	Rapamycin	Antibacterial
	CYP107G1[c]	*rapN*		
	CYP105U1	*gdmP*	Geldanamycin	
S. carzinostaticus	CYP154J1		Neocarzinostatin	Antitumor
	CYP208A2			
S. griseolus	CYP105A1	*suaC*	7-ethoxycoumarin	Antimicrobial
	CYP105B1[c]	*subC*		
S. nanchangensis	CYP124B2	*nanP*	Nanchangmycin	Anticoccidial
	CYP171A2	*meiE*	Meilingmycin	
S. natalensis	CYP161A2	*pimD*	Pimaricin	Antifungal
	CYP105H3[c]	*pimG*		
S. nodosus	CYP161A3	*amphL*	Amphotericin	Antifungal
	CYP105H4[c]	*amphN*		
S. noursei	CYP161A1[c]	*nysL*	Nystatin	Antifungal
	CYP105H1[c]	*nysN*		
S. tendae	CYP162A1	*nikQ*	Nikkomycin	Insecticidal
	CYP105K1[c]	*nikF*		
S. peucetius	CYP131A1	*dnrQ*	Doxorubicin	Antitumor
	CYP129A2	*doxA*		
S. sp. strain C5	CYP131A2	*dauQ*	Daunorubicin	Antitumor
	CYP129A1	*doxA*		
S. ansochromogenes	CYP105K2	*SanL*	Nikkomycin	Insecticidal
		SanQ		
S. carbophilus	CYP105A3		Pravastatin	AntiHIV
S. clavuligerus	CYP105M1[c]	*orf10*	Clavulinic acid	Antibacterial
S. caelestis	CYP113B2	*orf6*	Niddamycin	Antibacterial
S. cinnamonensis	CYP124B1	*monD*	Monensin	Anticoccidial
S. globisporus	CYP211A1	*orf29*	Antibiotic C-1027	Antitumor
S. narbonensis	CYP107L7P	*nbmL*	Narbomycin	Antifungal
S. spheroides	CYP163A1	*novI*	Novobiocin	Antibacterial
S. thermotolerans	CYP107C1	*orfA*	Carbomycin	Antibacterial
S. roseochromogenes	CYP163A2	*cloI*	Clorobiocin	Antibacterial
S. venezuelae	CYP107L1	*pikC*	Pikromycin	Antibacterial

[a] CYP names as annotated at website: ''http://drnelson.utmem.edu/CytochromeP450.html.''
[b] Gene names as described in NCBI database.
[c] CYPs clustered with ferredoxin.

Identification and analyses of CYPs from database

Genome databases of *S. peucetius* (unpublished) were formulated in our laboratory for the first time (cf., "http://203.230.12.60"). Other data were collected at various addresses mentioned in the table legends. CYP genes were initially searched based on the heme-binding domain signature, GXXXCXG. The ORFs of all CYP genes obtained using Glimmer 2.0 were annotated by BLAST, followed by manual inspection [28]. The ORFs containing such motifs were further screened for the presence of a highly conserved threonine in the putative I-helix, and the conserved EXXR motif present in the K-helix [29]. The genes including all three motifs were then used as queries for the local BLAST searches of the GenBank, using recent non-redundant protein database from NCBI to identify their closest homologes in other organisms. Each gene was translated to amino acid by GeneDoc program [30]. Dr. David Nelson had carried out the assignment of all CYPs according to the P450 nomenclature [31]. CYPs showing more than 40% identity were placed in the same family, and CYPs having more than 55% were categorized in the same subfamily. New family was assigned for the CYP showing less than 40% similarity to other organisms.

Multiple alignments, phylogenetic analyses, and comparison

The deduced amino acid sequences of the putative *S. peucetius* CYPs were aligned with the CYPs from *S. coelicolor* A3 (2) and *S. amermitilis* by ClustalX [32]. Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 2.1 [33]. Other CYPs involved in secondary metabolites production were searched at "http://drnelson.utmem.edu/CytochromeP450.html" and NCBI from 44 *Streptomyces* species. The accession number and protein identifier in database identified their positions as mentioned in Table 1.

DNA sequence accession numbers

The DNA sequences of all 19 putative CYPs, ferredoxin, and ferredoxin reductases have been deposited in the EMBL GenBank international nucleotide sequence database under accession numbers shown in Tables 3 and 4.

Results and discussion

S. peucetius CYP sequence alignments

Streptomyces peucetius ATCC 27952 is the only organism that produces clinically important anthracycline chemotherapeutic agents of the polyketide class of antibiotics, daunorubicin and doxorubicin. The 8.7 Mb sequence of *S. peucetius* ATCC 27952 revealed 19 CYPs, which cover approximately 0.2% of all coding sequences. Among those, the 5 CYPs; CYP107P3, CYP107U3, CYP107N3, CYP105F2, and CYP105P2 show more than 74% identity to cytochrome P450s of other organisms (Table 2). Threonine in I-helix (oxygen activation motif), glutamic acid and arginine in an EXXR motif of K-helix, and glycine and cysteine in the GXXXCXG signature of heme-binding domain are conserved in 15 CYPs (Table 3). CYP107N3 of *S. peucetius* has the highest homology 91% to CYP107N1 (Orf3) involved in the biosynthesis of mitomycin C from *S. lavendulae*.

The phylogenetic tree comprising CYPs of *S. avermitilis*, *S. coelicolor* A3 (2), and *S. peucetius* indicated that CYPs from *S. peucetius* are closer to those from *S. avermitilis* than *S. coelicolor* A3 (2) (Fig. 1). New CYP family found in *S. peucetius* is remotely related to CYP171A1 from *S. avermitilis*. CYPs belonging to the four subfamilies CYP107P, CYP107U, CYP154A, and CYP157C are present in *S. coelicolor* A3 (2), *S. avermitilis*, and *S. peucetius*. Five CYP subfamilies CYP102B, CYP105D, CYP154C, CYP157A, and CYP158A are also common in *S. coelicolor* A3 (2) and *S. avermitilis*. Similarly, CYPs of three subfamilies CYP105P, CYP107L, and CYP125A, are conserved only in *S. avermitilis* and *S. peucetius*.

Relationships to CYPs from other organisms

Our sequence analysis reveals that the four CYPs CYP147F1, CYP107AH1, CYP107AJ1, and CYP107L9 from *S. peucetius* have significant identity to CYP107L1 (PikC) found in macrolide antibiotics, pikromycin, from *S. venezuelae*. CYP251A1 is the new family found in *S. peucetius*, which shares 38% identity to CYP51 from *S. coelicolor* A3 (2) and eukaryotes [34]. Two other CYPs (CYP252A1 and CYP253A1) from *S. peucetius* also have 32% identity to CYP51. No CYP homologous to CYP51 is found in *S. avermitilis*. The remaining 18 CYPs fit into existing CYP families. Sequence analysis of *S. peucetius* revealed that majority of CYPs belong to CYP107 family.

The CYP164B1 from *S. peucetius* shows 39% identity to CYP164A1 from *Mycobacteriom leprae*, and hypothetical protein and PKS genes follow it. However, two cytochromes are clustered in *M. leprae*. CYP166B1 shows 44% identity to CYP105C1, clustered with cholesterol oxidation from *Streptomyces* species. CYP166B1 is clustered with modular PKS of avermectin in *S. peucetius*.

ATP-binding proteins and FAD-dependent monooxygenase proteins are found to be associated with CYP253A1. Four CYPs CYP107AJ1, CYP107AH1, CYP107U3, and CYP107L9 from *S. peucetius* exhibit

Table 2
Putative cytochrome P450s in *S. peucetius* with their closest homologues

CYP name[a]	Size[b]	Match in the databases[c]				
		Species	CYP name[a]	Protein identifier	Identity (%)	Overlap[d]
CYP147F1	412	*Myxococcus xanthus*	CYP147A1	AAD26638.1	53	317
		Streptomyces venezuelae	CYP107L1	AAC64105.1	39	402
CYP157C4	476	*Streptomyces griseus*	CYP157C3	BAB40762.1	54	470
		Streptomyces coelicolor	CYP157C1	NP_625901.1	51	465
CYP154A3	325	*Streptomyces coelicolor*	CYP154A1	NP_627112.1	72	320
		Streptomyces fradiae	CYP154B1	AAD40802.1	50	324
CYP107AH1	560	*Saccharopolyspora erythraea*	CYP107B1	AAA26483.1	45	396
		Streptomyces venezuelae	CYP107L1	AAC64105.1	44	410
CYP105P2	399	*Streptomyces avermitilis*	CYP105P1	BAB69309.1	89	399
		Streptomyces lavendulae	CYP105F1	AAD28448.1	40	403
CYP107U3	338	*Streptomyces coelicolor*	CYP107U1	NP_627317.1	88	323
		Saccharopolyspora erythraea	CYP107B1	AAA26483.1	43	315
CYP166B1	460	*Streptomyces* sp.	CYP105C1	AAA26718.1	44	382
		Amycolatopsis mediterranei	CYP166A1	AAC01709.1	43	381
CYP105F2	409	*Streptomyces lavendulae*	CYP105F1	AAD28448.1	83	406
		Streptomyces griseolus	CYP105B1	AAA26825.1	43	399
CYP107N3	414	*Streptomyces lavendulae*	CYP107N1	AAD28449.1	91	409
		Streptomyces maritimus	CYP107R1	AAF81737.1	52	391
CYP252A1	504	*Bacillus halodurans*	CYP197A1	NP_241445.1	36	457
		Streptomyces coelicolor	CYP51	NP_629370.1	32	456
CYP164B1	347	*Mycobacterium leprae*	CYP164A1	NP_302391.1	39	320
		Bacillus subtilis	CYP107H1	BAC03244.1	36	316
CYP107AJ1	400	*Saccharopolyspora erythraea*	CYP107B1	AAA26483.1	50	333
		Streptomyces venezuelae	CYP107L1	AAC64105.1	45	317
CYP253A1	429	*Bacillus halodurans*	CYP197A1	NP_241445.1	33	457
		Streptomyces coelicolor	CYP51	NP_629370.1	32	441
CYP125A4	327	*Mycobacterium tuberculosis*	CYP125A1	AAK48007.1	41	297
		Streptomyces avermitilis	CYP107W1	BAB69197.1	29	229
CYP107L9	449	*Streptomyces venezuelae*	CYP107L1	AAC64105.1	57	304
		Saccharopolyspora erythraea	CYP107B1	AAA26483.1	48	314
CYP107P3	373	*Streptomyces coelicolor*	CYP107P1	NP_627830.1	77	213
		Sinorhizobium meliloti	CYP202A1	NP_385358.1	45	211
CYP251A1	448	*Streptomyces coelicolor*	CYP51	NP_629370.1	38	420
		Streptomyces avermitilis	CYP171A1	BAA84477.1	36	452
CYP131A1	438	*Streptomyces peucetius*	CYP131A1	AAD15266.1	100	438
		Streptomyces antibioticus	CYP235A1	CAA05640.1	30	416
CYP129A2	415	*Streptomyces peucetius*	CYP129A2	AAD04715.1	100	415
		Streptomyces sp. strain C5	CYP129A1	AAB08049.1	85	223

[a] CYP names as annotated at website: "http://drnelson.utmem.edu/CytochromeP450.html."
[b] Number of amino acids.
[c] Database search at NCBI.
[d] Number of amino acid overlap, which exceeds the protein size, is due to the introduction of gaps during BLAST comparison.

similarity to CYP107B1 (EryF) involved in erythromycin biosynthetic pathway from *S. erythaea*. CYP107L1 from *S. venezuelae* has 45% identity to CYP107AJ1, and is clustered with ATP/GTP-binding protein. CYP107L1 also shows 44% similarity to CYP107AH1, and it is associated with putative dehydrogenase and expression regulator. Eukaryotic-type protein is located in the vicinity of CYP107U3, which shows 88% identity to

Table 3
Classification of putative cytochrome P450s in *S. peucetius*

CYP name[a]	Gene name[b]	I-helix	K-helix	Heme-binding motif	Accession number
CYP147F1	*cyp0549*	^{247}AGHET^{251}	^{286}ELLR^{289}	^{351}LGSGIHSCFG^{360}	AJ605536
CYP157C4	*cyp0618*	^{201}GSET^{205}	^{250}EHLR^{253}	^{360}FSSGPHECPG^{369}	AJ605537
CYP154A3	*cyp0673*	^{156}GYET^{260}	^{195}ETLR^{198}	^{262}FGHGVHFCLG^{271}	AJ605538
CYP107AH1	*cyp0749*	^{242}GHKT^{246}	^{281}ELVR^{284}	^{349}FGHGSHYCLG^{358}	AJ605539
CYP105P2	*cyp7863*	^{237}AHDT^{241}	^{276}ELLR^{279}	^{341}FGFGAHQCIG^{350}	AJ605540
CYP107U3	*cyp0819*	^{154}GFET^{158}	^{198}ELLR^{201}	^{263}YGHGIHYCLG^{272}	AJ605541
CYP166B1	*cyp0879*	^{243}GLET^{247}	^{282}ELLR^{285}	^{347}FGSGPHQCMG^{356}	AJ605542
CYP105F2	*cyp8812*	^{241}GHET^{245}	^{280}ELLR^{283}	^{345}FGHGIHLCLG^{354}	AJ605543
CYP107N3	*cyp0881*	^{251}GYET^{255}	^{290}ELLR^{293}	^{356}FGAGPHHCLG^{365}	AJ605544
CYP252A1	*cyp0812*	^{288}GYET^{292}	^{344}EVLR^{347}	^{417}YAGGPRNCVG^{426}	AJ605545
CYP164B1	*cyp0831*	^{165}GHET^{169}	^{204}EVLR^{207}	^{268}FAAGPHFCLG^{277}	AJ605546
CYP253A1	*cyp0913*	^{285}GHET^{289}	^{341}ESMR^{344}	^{414}FGAGPRKCIG^{423}	AJ605548
CYP251A1	*cyp0759*	^{260}GAET^{264}	^{316}EVIR^{319}	^{388}FGAGNRKCIG^{397}	AJ605552
CYP107P3	*cyp0887*	^{90}HEAT^{95}	^{129}ELMR^{132}	^{193}FGAGIHYCLG^{202}	AJ605551
CYP125A4	*cyp0266*	^{110}DRPT^{114}	^{144}ETGR^{147}	C-terminal missing	AJ605549
CYP129A2	*cyp0772*	^{277}GHDT^{281}	^{313}EALR^{316}	^{378}FGDGPHYCIG^{387}	AF403708
CYP107AJ1	*cyp0908*	^{149}PLSVT^{252}	^{205}ERRR^{208}	Unidentified	AJ605547
CYP107L9	*cyp0854*	^{230}GHET^{234}	^{269}EMLR^{272}	Unidentified	AJ605550
CYP131A1	*cyp0880*	^{275}GAET^{279}	^{314}ETLR^{317}	Unidentified	L47164

[a] CYP names as annotated at website: "http://drnelson.utmem.edu/CytochromeP450.html."
[b] Cytochrome P450 deposited in EMBL.

CYP107U1 from *S. coelicolor* A3 (2). CYP107L9 and CYP105L1 share 57% similarity within the same subfamily.

With the exception of *doxA* and *dnrQ*, other four CYPs, *CYP105P2*, *105F2*, *166B1*, and *107N3*, are clustered with putative modular PKS of avermectin biosynthesis in *S. peucetius*. Oxidation of deoxyavermectin aglycon is catalyzed by CYP171A1 (AveE) in *S. avermitilis* (10). CYP251A1 in *S. peucetius* represents AveE homologue protein (36%), but no such gene is found in the *S. coelicolor* A3 (2) genome [35].

Horizontally transferred CYPs

CYP105P2 has 89% similarity to CYP105P1 (PteC), and it is involved in the biosynthesis of antifungal polyene, filipin, and clustered with modular PKS and dehydrogenase in *S. avermitilis*. Interestingly, the same PKS and dehydrogenase homologue are arranged with CYP105P2 in *S. peucetius.* Despite being from different families, CYP105F2 and CYP107N3 are clustered together adjacent to LinA homolog and putative dehydrogenase in *S. peucetius.* Similar organizations are found in the mitomycin C biosynthetic gene cluster from *S. lavendulae.* CYP157C4 is clustered with *rar* operon of *S. peucetius*, which is also conserved in *S. coelicolor* A3 (2) and *S. griseus* [36]. This behavior probably resulted due to the horizontal gene transfer between those species.

CYPs clustered with ferredoxin and ferredoxin reductase

CYP147A1 from *Myxococcus xanthus* shares 53% identity with CYP147F1, and is the only CYP clustered with ferredoxin reductase in *S. peucetius*. One CYP from *S. coelicolor* A3 (2) and five CYPs from *S. avermitilis* are found to cluster with ferredoxin. No ferredoxin reductase is directly clustered with CYPs in *S. avermitilis* and *S. coelicolor* A3 (2); however, *CYP147B1* and *105Q1* from *S. avermitilis* are associated with ferredoxin, followed by ferredoxin reductase. 11 CYPs are clustered with ferredoxin in other *Streptomyces* (Table 1). One of the most studied CYPs in prokaryotes is CYP105D1 from *S. griseus* [37]. CYP105D4 from *S. lividans* and CYP105D5 from *S. coelicolor* A3 (2) share high homology to CYP105D1, and each of them is clustered with ferredoxin.

It has been suggested that ferredoxin and ferredoxin reductase enhance the expression and activity of each cytochrome [16,39], but the number of ferredoxin genes is limited to 9 and 6 in *S. avermitilis* and *S. coelicolor* A3 (2), respectively. Similarly, 6 and 3 ferredoxin reductases are found in *S. avermitilis* and *S. coelicolor* A3 (2), respectively. On the other hand, *S. peucetius* has 2 ferredoxins and 4 ferredoxin reductases, which will presumably support the activity of the 19 CYPs (Table 4).

CYPs clustered with regulatory elements

CYP251A1 from *S. peucetius* is found in a cluster containing putative RoK-family transcriptional regulatory protein. Another cytochrome of the same species, CYP252A1 is clustered with ATP/GTP-binding protein, threonine 3-dehydrogenase and LysR-family transcriptional regulator. The CYP147F1 is the only cytochrome P450 clustered with ferredoxin reductase, followed by

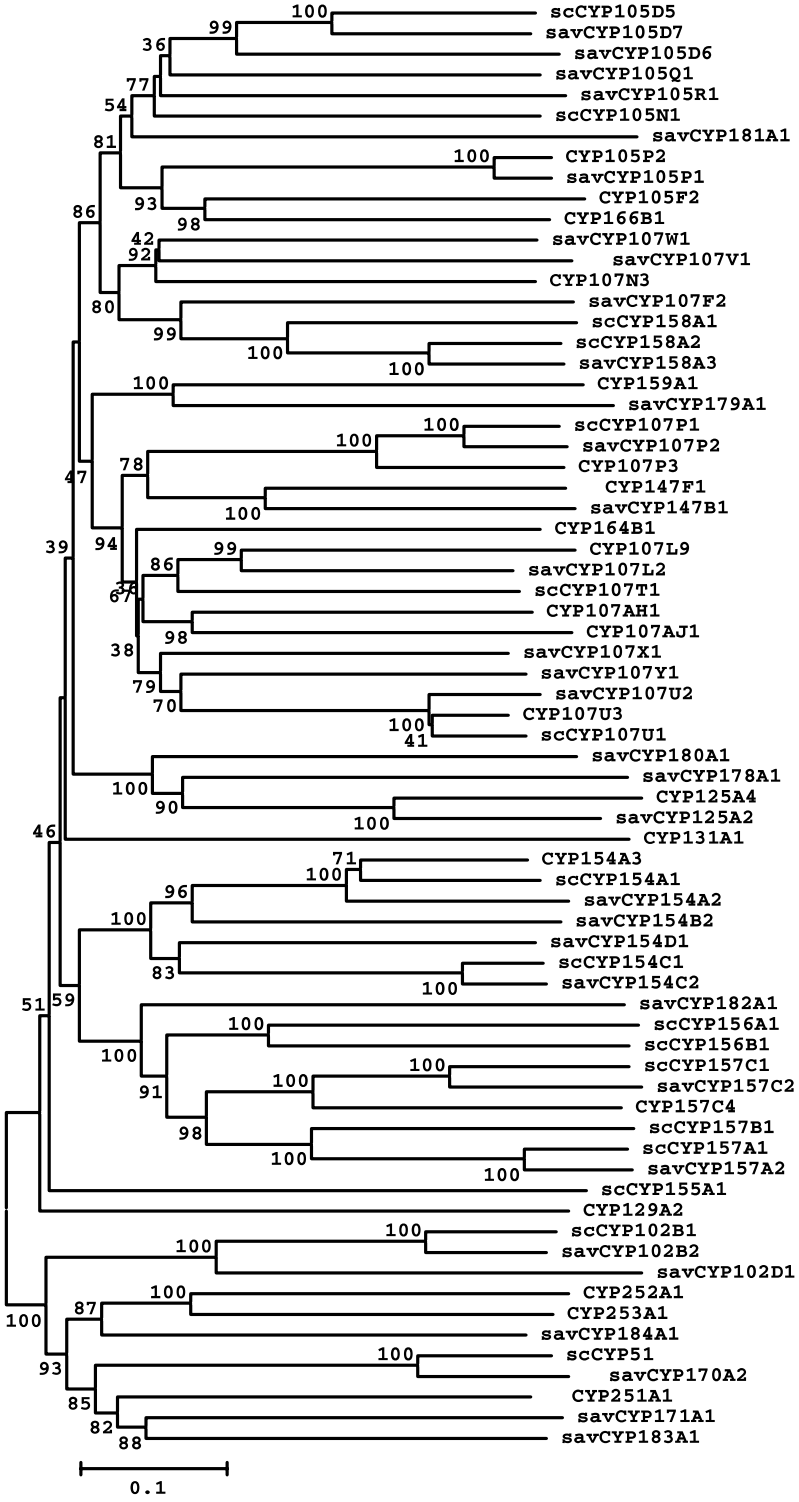

Fig. 1. NJ tree of CYPs from *S. peucetius*, *S. coelicolor* A3 (2) (prefix sc), and *S. avermitilis* (prefix sav). Horizontal branch lengths proportional to the estimated number of nucleotide substitutions, and bootstrap probabilities (as percentages), as determined for 1000 resamplings, are given *above* or *beside* the internal branches. The bar in the lower left corner indicates 0.1 amino acid substitutions per amino acid for the branch length.

ATP/GTP-binding protein and putative transcriptional regulator in *S. peucetius*. CYP157C4 from *S. peucetius* shows 54% identity to RarE of *rar* operon from *S. griseus*. In *S. coelicolor* A3 (2) *cvn9* showed a remarkable similarity to *rar* operon of *S. griseus*. *CYP154A3* lies in a gene cluster containing RNA polymerase sigma factor and MerR-family transcriptional regulatory protein in *S. peucetius*. CYP129A2 and 131A1 are associated with

Table 4
Putative ferredoxins and ferredoxin reductases in *S. peucetius* with their closest homologs

Gene name[a]	Accession number	Size[b]	Match in the databases[c]			
			Species	Gene name	Identity %	Predicted function
fdr549	AJ605536	422	*Myxococcus xanthus*	ORF1	57	CYP reductase
			Mycobacterium tuberculosis	MT0716	35	Ferredoxin reductase
fdr903	AJ628767	260	*Deinococcus radiodurans*	DR0496	48	Ferredoxin reductase
			Streptomyces coelicolor	SCF15.02	41	
fdr743	AJ628764	566	*Deinococcus radiodurans*	DRA0013	42	Ferredoxin reductase
			Trichodesmium sp.	nirA	36	
fdr793	AJ628765	450	*Streptomyces coelicolor*	SC4B10.18c	62	Ferredoxin reductase
			Streptomyces coelicolor	SC7AB.08c	55	
fdx808	AJ628766	352	*Mycobacterium tuberculosis*	fdxB	42	Ferredoxin
			Escherichia coli	D90778.1	41	
fdx734	AJ628763	244	*Amycolatopsis mediterranei*	ORF15B	65	Transketolase
			Rhodococcus fascians	fdx	55	Ferredoxin

[a] Ferredoxins and ferredoxin reductases deposited in EMBL.
[b] Number of amino acids.
[c] Database search at NCBI.

the pathway-specific regulatory gene *dnrI*. Similarly, CYP107AH1 is clustered with putative dehydrogenase and expression regulator in *S. peucetius*. CYP107L9 is found to cluster with the putative transcriptional regulator in *S. peucetius*. Surprisingly, a cytochrome of same subfamily, CYP105L1 (PikC) from *S. venezuelae*, is also located with the cluster containing ATP-binding protein and transcriptional regulator; therefore, a relationship between regulatory genes and CYPs could be expected in metabolism.

CYPs involved in the production of secondary metabolites

About 174 (40%) bacterial CYPs are reported from 45 different *Streptomyces* species in database, but only 50 CYPs are predicted to be involved in the production of secondary metabolites. The remaining CYPs are presumably involved in catabolism of carbon compounds. Although, none of the CYPs, involved in the biosynthesis of aminoglycoside antibiotics, have been assigned in database, many of them have been found to be associated with a wide range of other antibiotics (Table 1). CYP105D1 from *S. griseus* is one of the most studied bacterial cytochrome, which is involved in the biotransformation and degradation of a diverse array of structurally complex xenobiotics. Other CYPs of the same family: 105D2 from *S. griseus*, 105D3 from *Streptomyces sclerotialus*, 105D4 from *S. lividans*, 105D5 from *S. coelicolor* A3 (2), and 105D6, 105D7 from *S. avermitilis*, are suitable candidates for further investigations, since each might complement CYP105D1. The CYP154A1 from *S. coelicolor* A3 (2) shares 72% identity to CYP154A3 and lies with ATP/GTP-binding protein. CYP154A1 and CYP154C1 share 42% similarity in *S. coelicolor* A3 (2), but later catalyzes hydroxylation in 12- and 14-membered ring macrolactones in vitro [38]. CYP105D6 and 105D7 from the same subfamily are clustered in *S. avermitilis*, but they are absent in *S. coelicolor* A3 (2).

Significance of CYPs involved in streptomyces

Pathway modification of mitomycin C, oleandomycin, amphotericin, avermectin, and rapamycin can be carried out by a CYP belonging to CYP107L subfamily, since CYP107L1 (PikC) of this family shows flexibility over many macrolites. CYPs of two families, CYP154J and CYP211A, found in the biosynthetic pathway of enediyne antibiotics neocarzinostatin and C-1027, might be candidates for the screening of biosynthetic pathways of other enediyne antibiotics such as kedarcidin, esperamicin, and calicheamicin. Similarly, CYPs belonging to subfamilies CYP129A and CYP131A are also important for the development of hybrid anthracycline antibiotics. CYP161A and CYP105H subfamilies possibly can complement those involved in amphotericin and nystatin antibiotics, respectively.

Conclusions

S. peucetius ATCC 27952 includes 19 cytochrome P450s, 2 ferredoxins, and 4 ferredoxin reductases. Only one cytochrome, CYP147F1 from *S. peucetius*, is clustered with ferredoxin reductase. CYP251A1 is a new CYP found in *S. peucetius*, which shares sequence similarity with CYP51 from *S. coelicolor* A3 (2) and CYP171A1 from *S. avermitilis*. Six CYPs are associated

with secondary metabolite production in *S. peucetius*. Nine CYPs from *S. peucetius* are clustered with regulatory genes, and four CYPs are arranged with either ATP/GTP-binding proteins or dehydrogenases. We also observe such events in other *Streptomyces* species. The involvement of regulatory genes, ATP/GTP-binding proteins, and dehydrogenases with cytochrome P450 is a subject of further study. Distributions of CYPs in various species of *Streptomyces* reflect their participation either with secondary metabolite genes, ferredoxin, ferredoxin reductase or regulatory operon.

Acknowledgments

We are grateful to Dr. David Nelson for the annotation of CYPs. This project was supported by IMT2000 program, Ministry of Science & Technology, Republic of Korea.

References

[1] D.R. Nelson, Arch. Biochem. Biophys. 369 (1999) 1–10.
[2] T. Omura, Biochem. Biophys. Res. Commun. 266 (1999) 690–698.
[3] D.C. Lamb, H. Ikeda, D.R. Nelson, J. Ishikawa, T. Skaug, C. Jackson, S. Omura, M.R. Waterman, S.L. Kelly, Biochem. Biophys. Res. Commun. 307 (2003) 610–619.
[4] J.M. Weber, J.O. Leung, S.J. Swanson, K.B. Idler, J.B. McAlpine, Science 252 (1991) 114–117.
[5] Y. Xue, D. Wilson, L. Zhao, H. Liu, D.H. Sherman, Chem. Biol. 5 (1998) 661–667.
[6] L.A.M. Davies, E. Cundliffe, Mol. Microbiol. 13 (1994) 349–355.
[7] R. Fouces, E. Mellado, B. Diez, J.L. Barredo, Microbiology 145 (1999) 855–868.
[8] A.M. Rodriguez, C. Olano, C. Mendez, C.R. Hutchinson, J.A. Salas, FEMS Microbiol. Lett. 127 (1995) 117–120.
[9] P. Caffrey, S. Lynch, E. Flood, S. Finnan, M. Oliynyk, Chem. Biol. 8 (2001) 713–723.
[10] H. Ikeda, T. Nonomiya, M. Usami, T. Ohta, S. Omura, Proc. Natl. Acad. Sci. USA 96 (1999) 9509–9514.
[11] Molnar, J.F. Aparicio, S.F. Haydock, L.E. Khaw, T. Schwecke, A. Konig, J. Staunton, P.F. Leadlay, Gene 169 (1996) 1–7.
[12] N. Lomovskaya, S.L. Otten, Y.D. Katayama, L. Fonstein, X.C. Liu, T. Takatsu, A.I. Solari, S. Filippini, F. Torti, A.L. Colombo, C.R. Hutchinson, J. Bacteriol. 181 (1999) 305–318.
[13] S.L. Otten, X. Liu, J. Ferguson, C.R. Hutchinson, J. Bacteriol. 177 (1995) 6688–6692.
[14] M.L. Dickens, N.D. Priestley, W.R. Strohl, J. Bacteriol. 179 (1997) 2641–2650.
[15] H.T. Chiu, B.K. Hubbard, A.N. Shah, J. Eide, R.A. Fredenburg, C.T. Walsh, C. Khosla, Proc. Natl. Acad. Sci. USA 98 (2001) 8548–8553.
[16] H.A. Hussain, J.M. Ward, Appl. Environ. Microbiol. 69 (2003) 373–382.
[17] W. Liu, S.D. Christenson, S. Standage, B. Shen, Science 297 (2002) 1170–1173.
[18] Rascher, Z. Hu, N. Viswanathan, A. Schirmer, R. Reid, W.C. Nierman, M. Lewis, C.R. Hutchinson, FEMS Microbiol. Lett. 218 (2003) 223–230.
[19] C. Huawei, K.H. Brian, E.C. Sarah, T.W. Christopher, Chem. Biol. 9 (2002) 103–112.
[20] H. Zeng, H. Tan, J. Li, Curr. Microbiol. 45 (2002) 175–179.
[21] H. Chen, C.T. Walsh, Chem. Biol. 8 (2001) 301–312.
[22] U. Galm, J. Schimana, H.P. Fiedler, J. Schmidt, S.M. Li, L. Heide, Arch. Microbiol. 178 (2002) 102–114.
[23] R. Li, N. Khaleeli, C.A. Townsend, J. Bacteriol. 182 (2000) 4087–4095.
[24] D.C. Lamb, T. Skaug, H.L. Song, C.J. Jackson, L.M. Podust, M.R. Waterman, D.B. Kell, D.E. Kelly, S.L. Kelly, J. Biol. Chem. 277 (2002) 24000–24005.
[25] J. Sambrook, D.W. Russell, Molecular Cloning, Cold Spring harbor Laboratory Press, New York, 2001.
[26] R. Himmelreich, H. Hilbert, H. Plagens, E. Pirkl, B.C. Li, R.R. Herrmann, Nucleic Acids Res. 24 (1996) 4420–4449.
[27] B. Ewing, P. Green, Genome Res. 8 (1998) 186–194.
[28] A.L. Delcher, D. Harmon, S. Kasif, O. White, S.L. Salzberg, Nucleic Acids Res. 27 (1999) 4636–4641.
[29] D.R. Nelson, L. Koymans, T. Kamataki, J.J. Stegman, R. Feyereisen, D.J. Waxman, M.R. Waterman, O. Gotoh, M.J. Coon, R.W. Estabrook, I.C. Gunsalus, D.W. Nebert, Pharmacogenetics 6 (1996) 1–42.
[30] K.B. Nicholas, H.B. Nicholas Jr., GeneDoc: a tool for editing and annotating multiple sequence alignments, Editor and Shading Utility Version 2.6.002, 1997.
[31] D.R. Nelson, Methods Mol. Biol. 107 (1998) 15–24.
[32] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, Nucleic Acids Res. 24 (1997) 4876–4882.
[33] S. Kumar, K. Tamura, I.B. Jakobsen, M. Nei, Bioinformatics 17 (2001) 1244–1245.
[34] Y. Aoyama, T. Horiuchi, O. Gotoh, M. Noshiro, Y. Yoshida, J. Biochem. Tokyo 124 (1998) 694–696.
[35] S.D. Bentley, K.F. Chater, A.M. Cerdeno-Tarraga, G.L. Challis, N.R. Thomson, K.D. James, D.E. Harris, M.A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C.W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C.H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M.A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B.G. Barrell, J. Parkhill, D.A. Hopwood, Nature 417 (2002) 141–147.
[36] M. Komatsu, Y. Kuwahara, A. Hiroishi, K. Hosono, T. Beppu, K. Ueda, Gene 306 (2003) 79–89.
[37] H. Hussain, M.J. Ward, Enzyme Microb. Technol. 32 (2003) 790–800.
[38] L.M. Podust, Y. Kim, M. Arase, B.A. Neely, B.J. Beck, H. Bach, D.H. Sherman, D.C. Lamb, S.L. Kelly, M.R. Waterman, J. Biol. Chem. 278 (2003) 12214–12221.
[39] L. Lei, M.R. Waterman, A.J. Fulco, S.L. Kelly, D.C. Lamb, Proc. Natl. Acad. Sci. USA 101 (2004) 494–499.