CHAPTER III

RESULTS

Screening of Polyketide Cytochrome P-450 Hydroxylase Gene

Dot blot hybridization of 100 strains *Streptomyces* genomic DNA samples using a *picK* gene probe showed a variety of signal strength. When intensity of hybridization signals of the samples was at least one fold higher than that of *E.coli*, it was interpreted to a positive result. The positive and negative results are indicated in parenthesis and shown in Figure 9 and 10. *S. rimosus* (-), *S. hygroscopicus* (+), SMC 1 (-), SMC 2 (-), SMC 3 (-), SMC 4 (-), SMC 5 (-), SMC 6 (-), SMC 7 (-), SMC 8 (+), SMC 9 (+), SMC 10 (-), SMC 12 (-), SMC 13 (-), SMC 14 (-), SMC 15 (-), SMC 16 (+), SMC 17 (+), SMC 18 (+), SMC 19 (+), SMC 20 (-), SMC 21(-), SMC 24 (+), SMC 25 (-), SMC 26 (-), SMC 27 (-), SMC 28 (-), SMC 29 (+), SMC 30 (+), SMC 31 (+), SMC 32 (-), SMC 33 (-), SMC 35 (-), SMC 36 (-), SMC 37 (-), SMC 38 (+), SMC 39 (-), SMC 40 (+), SMC 41(-), SMC 42 (-), SMC 43 (-), SMC 44 (+), *S. narbonensis* (-), SMC 46 (-), SMC 47(-), SMC 49 (-), SMC 50 (+), SMC 51 (-), SMC 52 (+), SMC 54 (-), SMC 55 (-), SMC 56 (+), SMC 57 (-), SMC 58 (-), SMC 59 (+), SMC 60 (+), SMC 61 (-), SMC 62 (-), SMC 63 (-), SMC 64 (-), SMC 65 (-), SMC 66 (-), SMC 67 (-), SMC 68 (-), SMC 69 (-), SMC 70 (-), SMC 71 (-), SMC 72 (-), SMC 73 (-), SMC 74 (+), SMC 75 (-), SMC 76 (-), SMC 77 (-), SMC 79 (-), SMC 80 (-), SMC 81(-), SMC 82 (-), *S. coerulescens* (+), SMC 83 (-), *S. roseoviolaceus* (+), SMC 84(-), SMC 85 (-), SMC 86 (+), SMC 87 (-), SMC 88 (+), SMC 90 (-), SMC 91 (+), SMC 92 (-), SMC 93 (-), SMC 94 (-), SMC 95 (-), SMC 96 (-), *S. flavoviridis* (+), SMC 22 (+), SMC 23 (+), SMC 11 (+), SMC 89 (+), *S. lividans* (+), SMC 78 (+) and SMC 48 (+).

	1	2	3	4	5	6	7	8	9
1	picK gene 1000 ng	picK gene probe 500ng	picK gene probe 50 ng	E.coli 2000 ng	E.coli 500 ng	E.coli 100 ng	S.rimosus	S.hygro scopicus	SMC 1
2	SMC 2	SMC 3	SMC 4	SMC 5	SMC 6	SMC 7	SMC 8	SMC 9	SMC 10
3	SMC 12	SMC 13	SMC 14	SMC 15	SMC 16	SMC 17	SMC 18	SMC 19	SMC 20
4	SMC 21	SMC 24	SMC 25	SMC 26	SMC 27	SMC 28	SMC 29	SMC 30	SMC 31
5	SMC 32	SMC 33	SMC 35	SMC 36	SMC 37	SMC 38	SMC 39	SMC 40	SMC 41
6	SMC 42	SMC 43	SMC 44	S.narbo nensis	SMC 46	SMC 47	SMC 49	SMC 50	SMC 51
7	SMC 52	SMC 54	SMC 55	SMC 56	SMC 57	SMC 58	SMC 59	SMC 60	SMC 61
8	SMC 62	SMC 63	SMC 64	SMC 65	SMC 66	SMC 67	SMC 68	SMC 69	SMC 70
9	SMC 71	SMC 72	SMC 73	SMC 74	SMC 75	SMC 76	SMC 77	SMC 79	SMC 80
10	SMC 81	SMC 82	S.coeru lescens	SMC 83	S.roseo violaceus	SMC 84	SMC 85	SMC 86	SMC 87
11	SMC 88	SMC 90	SMC 91	SMC 92	SMC 93	SMC 94	SMC 95	SMC 96	
12							S.venezu-elae 50 ng	S.venezu-elae 500 ng	S.venezu-elae 1000 ng
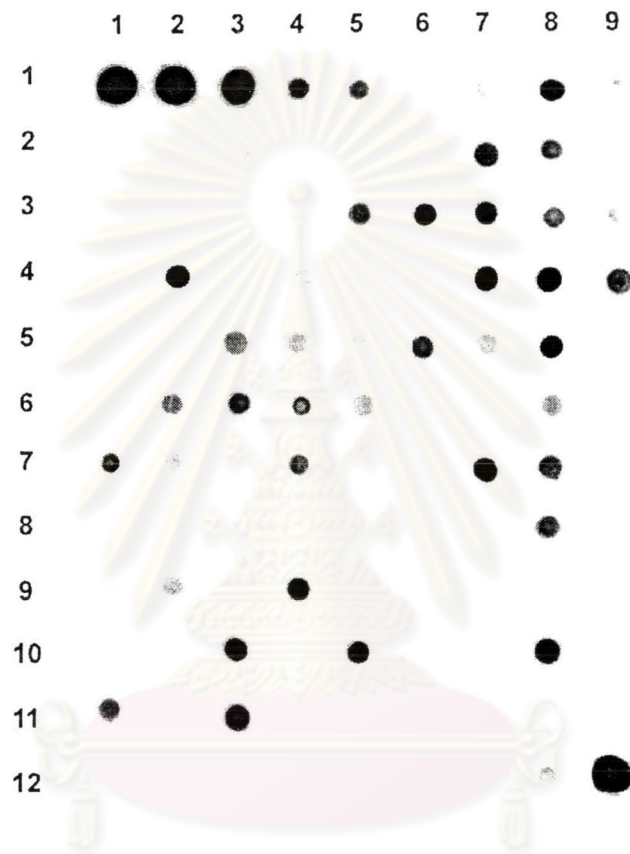
Figure 9. Dot blot hybridization of 92 samples of genomic DNA isolated from *Streptomyces* spp.

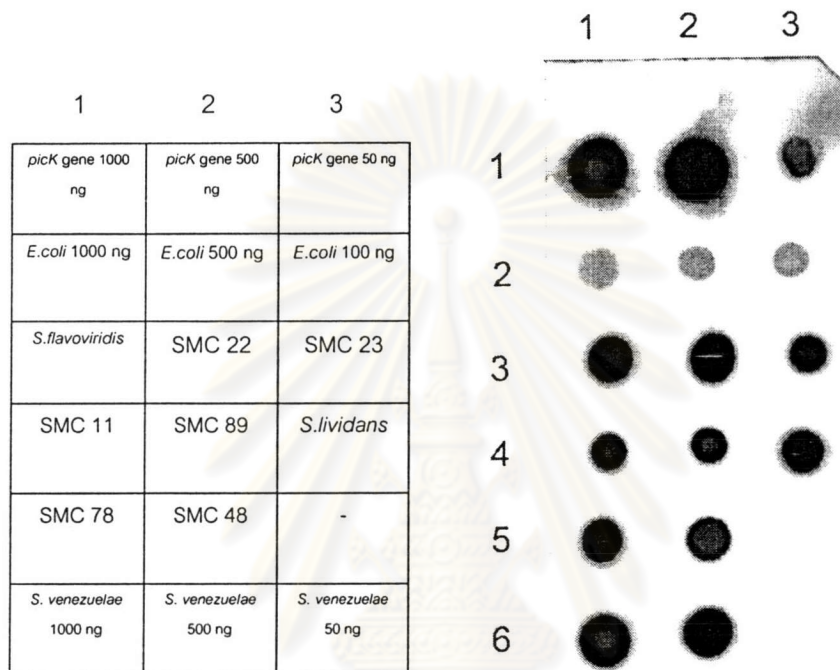|	1	2	3
1	picK gene 1000 ng	picK gene 500 ng	picK gene 50 ng
2	E.coli 1000 ng	E.coli 500 ng	E.coli 100 ng
3	S.flavoviridis	SMC 22	SMC 23
4	SMC 11	SMC 89	S.lividans
5	SMC 78	SMC 48	-
6	S. venezuelae 1000 ng	S. venezuelae 500 ng	S. venezuelae 50 ng

Figure 10. Dot blot hybridization of 8 samples of genomic DNA from *Streptomyces* spp.

Isolation of Putative Polyketide Cytochrome P-450 Gene From *Streptomyces* Genomic DNA

Twenty positively samples (intensity of hybridization signals is 3-4 fold) were selected to identify the polyketide P-450 gene using the PCR. The annealing temperature (Ta) for samples SMC 29, SMC 40, SMC 86, SMC 89, SMC 91, *S. coerulescens* and *S. roseoviolaceus* was 62 ^{o}C; *S. hygroscopicus* and SMC 78 was 65 ^{o}C; SMC 30, SMC 48, SMC 59, *S. lividans* and *S. flavoviridis* was 67 ^{o}C. Acheivement of PCR as visualized by agarose gel electrophoresis was obtained from fourteen samples (*S.hygroscopicus*, SMC 30, *S.flavoviridis*, SMC 29, SMC 40, SMC 48, SMC 59, *S.lividans*, SMC 78, SMC 89, *S.coerulescens*, *S.roseoviolaceus*, SMC 86 and SMC 91) and the products were about 800 bp in size (Figure 11 and 12).

Identification of Polyketide Cytochrome P-450 Hydroxylase Gene

The PCR products of SMC 48, SMC 59 and SMC 78 were cloned in pCR®4-TOPO vector (4000 bp) and transformed into *E.coli* TOP10. Recombinant clones were screened and identified by miniprep plasmid isolation and plasmid digestion with *EcoR*I. The right recombinant clones showed two fragments of approximately 4000 bp and 800 bp. Screening of recombinant clones was shown in Figure 13, 14 and 15. DNA sequences of putative polyketide P-450 genes of *Streptomyces* strains SMC 48, SMC 59 and SMC 78 were shown in Figure 20, 27 and 34 , respectively.

Figure 11.   Agarose gel electrophoresis of PCR products of SMC 12, SMC 24,

S. hygroscopicus, SMC 30, SMC 59, SMC 78, SMC 23, *S. lividans*, SMC 23, SMC 48, SMC 40

and SMC 11.

Lane M : Gene ruler™ 1 kb Ladder, lane 1 : negative control, lane 2 : positive control, lane 3 :

SMC 12, lane 4 : SMC 24, lane 5 : *S. hygroscopicus*, lane 6 : SMC 30, lane 7 : SMC 59, lane 8 :

SMC 78, lane 9 : SMC 22, lane 10 : *S. lividans*, lane 11 : SMC 23, lane 12 : SMC 48, lane 13 :

SMC 40, and lane 14 : SMC 11. The arrow indicates the desired band.

Figure 12.   Agarose gel electrophoresis of PCR products of SMC 74, *S. flavoviridis*, SMC 29,

SMC 89, *S. coerulescens*, *S. roseoviolaceus*, SMC 86 and SMC 91.

Lane M : Gene ruler™ 1 kb Ladder, lane 1 : negative control, lane 2 : SMC 74, lane 3 : *S.*

flavoviridis, lane 4 : SMC 29, lane 5 : SMC 89, lane 6 : *S. coerulescens*, lane 7 : *S.*

roseoviolaceus, lane 8 : SMC 86 and lane 9 : SMC 91, and lane 10 : positive control. The arrow

indicates the desired band.

Figure 13. Agarose gel electrophoresis of plasmid isolated from recombinant clone SMC 48. Lane M : Gene ruler™ 1 kb Ladder, lane 1 : uncut SMC 48 clone1, lane 2 : EcoRI digested SMC 48 clone1, lane 3 : uncut SMC 48 clone2 and lane 4 : EcoRI digested SMC 48 clone2. The arrow indicates the insert fragment.

Figure 14. Agarose gel electrophoresis of plasmid isolated from recombinant clone SMC 59.

Lane M : Gene ruler™ 1 kb Ladder, lane 1 : uncut SMC 59 clone 1, lane 2 : EcoRI digested

SMC 59 clone 1, lane 3 : uncut SMC 59 clone 2, lane 4 : EcoRI digested SMC 59 clone 2, lane

5 : uncut SMC 59 clone 3, lane 6 : EcoRI digested SMC 59 clone 3, lane 7 : uncut SMC 59

clone 4 and lane 8 : EcoRI digested SMC 59 clone 4. The arrow indicates the insert fragment.
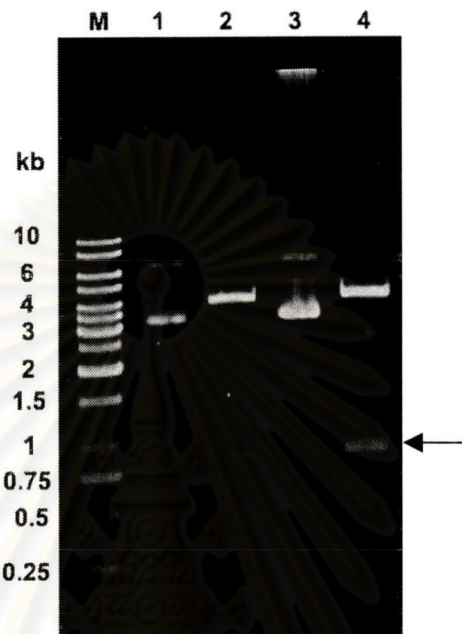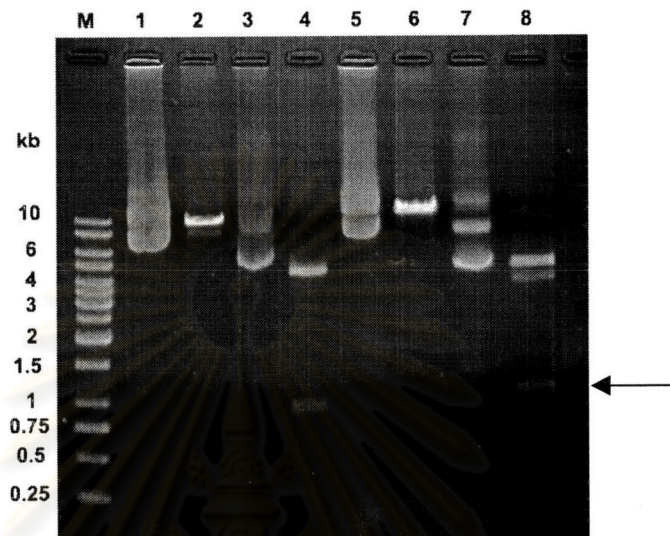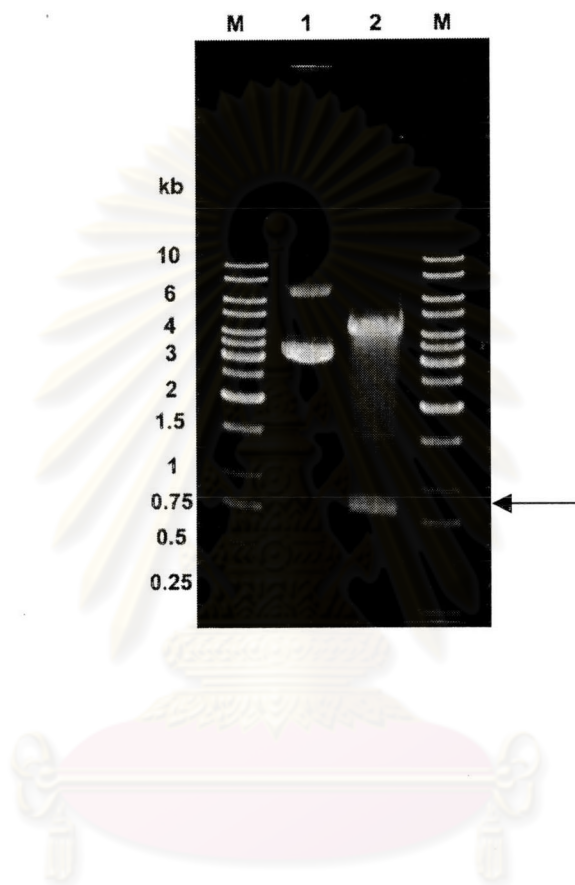
Figure 15. Agarose gel electrophoresis of plasmid isolated from recombinant clone SMC 78.

Lane M : Gene ruler™ 1 kb Ladder, lane 1 is uncut SMC 78, lane 2 is *Eco*RI digested SMC 78,

lane M : Gene ruler™ 1 kb Ladder. The arrow indicates the insert fragment.

Sequence Analysis of Putative Cytochrome P-450 Genes

To predict their biological function of the isolated genes, the nucleotide sequences and their deduced amino acid sequences of SMC 48, SMC 59 and SMC 78 were used as a query for the BLAST search. Results from BLASTN, TBLASTX and BLASTP of our sequences showed the sequence similarity to those polyketide P-450 of *Streptomyces* species as shown in Figure 16-18, 23-25, 30-32. The DNA sequence of SMC 48 showed high homology to cyp7, a putative cytochrome P-450 of *Streptomyces avermitilis* (GenBank No. AP004027) (45% identity) as shown in Figure 21. Similarly, these sequences of SMC 59 and SMC 78 showed high similarity to the sequences of putative cytochrome P-450 of *Streptomyces coelicolor* A3(2) (GenBank No. AL939114) (41% identity) as shown in Figure 26 and 35.

Analysis of amino acid sequences of putative P-450 from SMC 48 , SMC 59 and SMC 78 revealed the significant motifs of the cytochrome P-450 (oxygen-binding and heme-binding region). The oxygen-binding domain of translated amino acid sequences SMC 48 were TGHET, and SMC 59 was AGYET, and SMC 78 was AGFET. The heme-binding domain of translated amino acid sequences SMC 48 were FGYGVHLCIG, and that of SMC 59 and SMC 78 was FGIGVHHCIG. The putative macrolide   P-450 were illustrated in Figure 21, 28 and 35 respectively.

Secondary structures of the translated amino acid sequences of SMC 48, SMC 59 and SMC 78 were predicted and aligned to those of the highest homologous EryF P-450 sequence as shown in Figure 22, 29 and 36 , respectively. The topology of putative P-450s of unknown samples was similar to the topology of EryF from *Saccharopolyspora erythraea* (GenBank No. M54983).

To demonstrate the relationship of our unknown DNA sequences to other P-450s, phylogenetic trees were constructed using a maximum parsimony algorithm as shown in Figure 37 and 38. The trees implied that nucleotide and amino acid sequences of SMC 48 was closely related to

P-450s of *Streptomyces griseus* (GenBank No. X63601), of *Streptomyces carbophilus* (responsible for production of pravastatin) (GenBank No. D30815), of *Streptomyces griseolus* (responsible in herbicide monooxygenase systems) (GenBank No. M32238) and of *Streptomyces diastaticus* (responsible for macrolides rimocidin and CE-108 biosynthesis) (GenBank No. AY442225). Phylogenetic relationship of the sequence of SMC 59 and SMC 78 exhibited that they were claded in the group of P-450 involved in tylosin-biosynthetic regulatory gene cluster isolated from *Streptomyces fradiae* (GenBank No. AF145049).

Figure 16. Result from Basic Local alignment Search for nucleotide similarity to nucleotide sequence of SMC 48

Sequences producing significant alignments:

		Score (bits)	E Value
gi|296051102|dbj|AP005027.1|	Streptomyces avermitilis genomic DNA, complete genome, section 7/30	90	7e-15
gi|153480|gb|M32239.1|STMSUBCB	S.griseolus cytochrome P-450-SU2 (subC) and ferredoxin-2 (subB) genes, complete cds	62	2e-06
gi|561881|gb|L37200.1|STMOLEP	Streptomyces antibioticus cytochrome P-450 (oleP) gene, complete cds	60	7e-06
gi|24413728|emb|AL939106.1|SCO939106	Streptomyces coelicolor A3(2) complete genome; segment 3/29	54	4e-04
gi|3293538|gb|AF072709.1|AF072709	Streptomyces lividans amplifiable element AUD4; putative transcriptional regulator, putative ferredoxin, putative cytochrome P450 oxidoreductase, and putative oxidoreductase genes, complete cds; and unknown genes	54	4e-04
gi|3243066|gb|AF071148.1|AF071148	Amycolata autotrophica cytochrome P450 hydroxylase gene, partial cds	54	4e-04
gi|3243062|gb|AF071146.1|AF071146	Micromonospora inyoensis cytochrome P450 hydroxylase gene, partial cds	54	4e-04
gi|3243068|gb|AF071149.1|AF071149	Streptomyces sclerotialus cytochrome P450 hydroxylase gene, partial cds	52	0.002
gi|11993518|gb|AF306830.1|AF306830	Streptomyces aureofaciens Ren71 (ren71) gene, complete cds; and unknown genes	52	0.002
gi|28192593|gb|AY179507.1|	Streptomyces hygroscopicus strain NRRL 3602 geldanamycin biosynthesis gene cluster, partial sequence	50	0.006

Figure 17. Result from Basic Local alignment Search for translated amino acid similarity to translated amino acid sequence of SMC 48

Sequences producing significant alignments:

		Score (bits)	E Value
gi|296051102|dbj|AP005027.1|	Streptomyces avermitilis genomic DNA, complete genome, section 7/30	204	4e-86 3
gi|153480|gb|M32239.1|STMSUBCB	S.griseolus cytochrome P-450-SU2 (subC) and ferredoxin-2 (subB)genes, complete cds	139	2e-76 3
gi|107232160|dbj|D30815.1|STMCYP450	Streptomyces carbophilus gene for cytochrome P-450sca-2, complete cds	126	5e-64 4
gi|153477|gb|M32238.1|STMSUACB	S.griseolus cytochrome P-450-SU1 (suaC) and ferredoxin-1 (suaB)genes, complete cds	122	9e-62 5
gi|296109620|dbj|AP005050.1|	Streptomyces avermitilis genomic DNA, complete genome, section 30/30	104	8e-61 4
gi|8050835|gb|AF263912.1|AF263912	Streptomyces noursei ATCC 11455 nystatin biosynthetic gene cluster, complete sequence	130	1e-60 3
gi|24413728|emb|AL939106.1|SCO939106	Streptomyces coelicolor A3(2) complete genome; segment 3/29	104	2e-59 4
gi|32935381|gb|AF072709.1|AF072709	Streptomyces lividans amplifiable element AUD4: putative transcriptional regulator, putative ferredoxin, putative cytochrome P450 oxidoreductase, and putative oxidoreductase genes, complete cds; and unknown genes	104	2e-59 4
gi|24413926|emb|AL939132.1|SCO939132	Streptomyces coelicolor A3(2) complete genome; segment 29/29	148	4e-58 4
gi|5420037|emb|Y18574.1|STE18574	Streptomyces tendae strain Tue901, nikA, nikB, nikC, nikD, nikE, nikF and nikG genes, partial	108	1e-57 3

Figure 18. Result from Basic Local alignment Search for protein similarity to translated amino acid sequence of SMC 48

Sequences producing significant alignments:

		Score (bits)	E Value
gi|2982815|ref|NP_822787.1|	putative cytochrome P450 [Streptomyces avermitilis MA-4680]	270	2e-71
gi|117302|sp|P18327|CPXF_STRGO	Cytochrome P450-SU2 (P450-CVB1) (CYP105B1)	248	8e-65
gi|2983401|ref|NP_828645.1|	putative cytochrome P450 [Streptomyces avermitilis MA-4680]	215	6e-55
gi|23396528|sp|Q59831|CPS2_STRCC	Cytochrome P450 105A3 (Cytochrome P450 sca-2) (CYT P-450sca-2)	212	5e-54
gi|117301|sp|P18326|CPXE_STRGO	Cytochrome P450-SU1 (P450-CVA1) (CYP105A1)	212	5e-54
gi|28192614|gb|AAO06929.1|	GdmP [Streptomyces hygroscopicus]	208	6e-53
gi|8050845|gb|AAF71771.1|	NysN [Streptomyces noursei]	207	2e-52
gi|21039505|gb|AAM33670.1|	putative cytochrome P450 GrhO3 [Streptomyces sp. JP95]	204	2e-51
gi|3293542|gb|AAC25766.1|	putative cytochrome P450 oxidoreductase [Streptomyces lividans]	202	3e-51
gi|21219297|ref|NP_625076.1|	putative cytochrome P450 [Streptomyces coelicolor A3(2)]	202	3e-51

Figure 19.   Pairwise alignment of nucleotide sequence of SMC 48 and P-450 from *S.avermitilis* (AP005027)

```
SMC 48      ------------------------------------------------------------
AP005027    ATGGCCGATGCCCTGGCGGGCCGCGCGCCCGACGCCACCCCGCCGGTCGCCGCGTACCCG


SMC 48      ------------------------------------------------------------
AP005027    ATGCCCCGGGCGGTGAGCTGCCCGCTCGCCCCGCCGCCCGCCCTGCAACCGTTGCGCGAC


SMC 48      ------------------------------------------------------------
AP005027    GAGCAGCCGATCACCAAGGTGCGGATCTGGAACGGCAGCACGCCCTGGCTCATCACCCGC


SMC 48      ------------------------------------------------------------
AP005027    CACGCCGATCAGCGCGCCCTGCTCACCGACCCCCGCGTCAGCAACGACGACCGCGATCCC


SMC 48      ------------------------------------------------------------
AP005027    GGCTTCCCCTACGTGAACGCGCACCGCGCGCGGAGATCGCGCACGCGACCCCGAGGCTCATC


SMC 48      -----------------GG-------CCGGCTGCGCAGGCTGGT-GACGGCGCCCTTCGC
AP005027    ACCAACACCGACGCACCGGAACACACCCGGCTGCGCCGGACGGTCAACG-CACCTTTCCT
                             **    ********** **   ***   *** * ** ***


SMC 48      GATCAAGAAGGTGGAGGCGCTTCGGCC-GTCGGTCCAGAAGATCGTCGATGACCTGATCG
AP005027    CATCAAGCGGATCGAGGCCATGCGGCCCGCCG-TCCAGCGGATCGTCGACGGTCTGATCG
            ****** * * ***** * ***** * ** ***** ********* *  *******


SMC 48      ACGCGATGCTCGCCGGGCCGAAGCCGGTCGATCTGGTTGAGGCGTTCGCGCTGCCGGTCC
AP005027    ACGACATGCTGGACGGACCGAACCCGGCCGATCTCCTCACCGCGCTCGCGCTGCCCGTGC
            ***   ***** * *** ***** **** **** * * ***  *** ********** ** *


SMC 48      CCAGCCTGGTCATCTGCGAGCTGCTCGGCGTGCCCTACGAGGACCACGACCTCTTCCAGC
AP005027    CGTCCCTGGTCATCGCCCAGTTGCTCGGTGTGCCGTACGCGGACCACGAGTTCTTCCAGC
            *  ********** *  ** *  ** ******* ***** * **** **** ********* ** *********


SMC 48      GCAACACCAGGGTCGCCGTCCGCCGCACCTCGA--CAGCGGAGGAGGTGGTCGCGGCCAT
AP005027    GCAACAGCAACCTCGT-GCTCGACA-ACTCCGTTCCGGCCGAGGAGGCGCGGGCGGCGA-
            ****** **   *** *  ** **  ** *  * ** ******* *     ***** *


SMC 48      GCAGA-CGCTCACCGACTACCTGGACGACCTCCTGACCGTCAAGCGCGCCCGGCCCGGGG
AP005027    GCGGGGCCCTCGCCGCCTACCTCGACACCCTGCTCGCGGAGAAGACCGCC-GCCCCCGAC
            ** **  * *** *** ***** *** *  *   * *  ** ***  **** * *** *


SMC 48      ACG-ACCTGCTCTCCGTACTGGCCA--CGGAGCGGGTGGCCAAGGGCGAGATGACGCAGC
AP005027    GCGGACGTGCTGTCCGAGATGGGCGGCCGGATCAAG--GCC---GGCGAGATGACGCACC
            ** ** **** ****    *** *     **** *  *   ***   ************* *


SMC 48      GCCAGGCCGCCGAGAACGGGTGT-GCTGCTGCTGGCCACGGGCCACGAGACCACCGCGAA
AP005027    GCGAGGCTGTCCACA-TGGGCGTCGCCA-TGCTCATCGCCGGACACGAGACCACCGCGAC
            ** **** * * *   *** ** **    **** *  * * ** **************** *


SMC 48      CATGATC-GCGCTCGGCACCCTCGCCCTGCTGCGGA-ACCCGGACCAGCTCGCTCTGCCC
AP005027    GATGATCAGC-CTCGGCACCCTCGCCCTGCT-CGAACACCCGGAGCAGCTGGCCGTGCTG
            ****** ** ******************* **  * ******* ***** ***** **  ***
```

```
SMC 48      CGCGACACCGAGGACCCCAAGGTGGTCGCCGGAGCCGTGGAGGAACTGCTGCGCTACCTG
AP005027    CGGGACGCCGAGGACCCGAAGGTGATCGCGGCCGCGGTGGAGGAACTGCTGCGCTATCTG
            ** *** ********** ****** ****  *  ** ****************** ***

SMC 48      AACATCACCCACTCCGGGCG-CCGCCGCGTGGCGCT--GGCGGACATCGAGGTCGGCGGC
AP005027    ACCATCGTCCACTCCGG-CATCCGCCGGGT--CGCCAAGGAGGACATCGAGATCGGCGAC
            *  **** ********* *  ****** **    ***  ** ********** ****** *

SMC 48      GAGGTCATCCGCGCCGGCGACGGCGTCATCTTCG-CCAACGAGATCGCCAACCGGGACCC
AP005027    CGGGTCATCGGCGCCGGGGACGGCCTTCTGTTCGATCTGC-ACGCCGCGAACTGGGACTC
            ******* ******* ****** *  * ****  *  *    *** *** ***** *

SMC 48      GGAAGCCTTCTCCGACCCCGACCGGCTCGACATCACCCGTGACGCCCGCCGCCACGTCGC
AP005027    GGAGGCCTTTCCCGAGGCCGACCGGCTGGACCTGAGCCGTCCCGCCCGCCACCACCAGGC
            *** *****  ****  ** ********* *** * * **** ******** ****   **

SMC 48      CTTCGGCTACGGCGTGCACC----TCTG---CA-TCGGCG--------------------
AP005027    ATTCGGCTACGGGCCCCACCAGTGTCTGGGACAGTCGCTGGCCCGGCTGGAGCTCCAGGT
             ********** ****  ****     ****  ** ** *  *

SMC 48      ------------------------------------------------------------
AP005027    CGTCTACGGCACGCTCTACCGCCGCTTCCCCACACTGCGTCTGGCCGCCCCGATCGACCG

SMC 48      ------------------------------------------------------------
AP005027    GCTCGCGTTCCACCACACCGGCACCACCTACGGCGTCCGCTGTCTGCCCGTCACCTGGTG

SMC 48      -
AP005027    A
```

Figure 20. DNA sequence of putative macrolide P-450 from SMC 48 and its translated amino acid sequence

```
         10        20        30        40        50        60
GGCCGGCTGCGCAGGCTGGTGACGGCGCCCTTCGCGATCAAGAAGGTGGAGGCGCTTCGG
 G  R  L  R  R  L  V  T  A  P  F  A  I  K  K  V  E  A  L  R

         70        80        90       100       110       120
CCGTCGGTCCAGAAGATCGTCGATGACCTGATCGACGCGATGCTCGCCGGGCCGAAGCCG
 P  S  V  Q  K  I  V  D  D  L  I  D  A  M  L  A  G  P  K  P

        130       140       150       160       170       180
GTCGATCTGGTTGAGGCGTTCGCGCTGCCGGTCCCCAGCCTGGTCATCTGCGAGCTGCTC
 V  D  L  V  E  A  F  A  L  P  V  P  S  L  V  I  C  E  L  L

        190       200       210       220       230       240
GGCGTGCCCTACGAGGACCACGACCTCTTCCAGCGCAACACCAGGGTCGCCGTCCGCCGC
 G  V  P  Y  E  D  H  D  L  F  Q  R  N  T  R  V  A  V  R  R

        250       260       270       280       290       300
ACCTCGACAGCGGAGGAGGTGGTCGCGGCCATGCAGACGCTCACCGACTACCTGGACGAC
 T  S  T  A  E  E  V  V  A  A  M  Q  T  L  T  D  Y  L  D  D

        310       320       330       340       350       360
CTCCTGACCGTCAAGCGCGCCCGGCCCGGGGACGACCTGCTCTCCGTACTGGCCACGGAG
 L  L  T  V  K  R  A  R  P  G  D  D  L  L  S  V  L  A  T  E

        370       380       390       400       410       420
CGGGTGGCCAAGGGCGAGATGACGCAGCGCCAGGCCGCCGAGACGGGTGTGCTGCTGCTG
 R  V  A  K  G  E  M  T  Q  R  Q  A  A  E  T  G  V  L  L  L

        430       440       450       460       470       480
GCCACGGGCCACGAGACCACCGCGAACATGATCGCGCTCGGCACCCTCGCCCTGCTGCGG
 A  T  G  H  E  T  T  A  N  M  I  A  L  G  T  L  A  L  L  R

        490       500       510       520       530       540
AACCCGGACCAGCTCGCTCTGCCCCGCGACACCGAGGACCCCAAGGTGGTCGCCGGAGCC
 N  P  D  Q  L  A  L  P  R  D  T  E  D  P  K  V  V  A  G  A

        550       560       570       580       590       600
GTGGAGGAACTGCTGCGCTACCTGAACATCACCCACTCCGGGCGCCGCCGCGTGGCGCTG
 V  E  E  L  L  R  Y  L  N  I  T  H  S  G  R  R  R  V  A  L

        610       620       630       640       650       660
GCGGACATCGAGGTCGGCGGCGAGGTCATCCGCGCCGGCGACGGCGTCATCTTCGCCAAC
 A  D  I  E  V  G  G  E  V  I  R  A  G  D  G  V  I  F  A  N

        670       680       690       700       710       720
GAGATCGCCAACCGGGACCCGGAAGCCTTCTCCGACCCCGACCGGCTCGACATCACCCGT
 E  I  A  N  R  D  P  E  A  F  S  D  P  D  R  L  D  I  T  R

        730       740       750       760       770
GACGCCCGCCGCCACGTCGCCTTCGGCTACGGCGTGCACCTCTGCATCGGCG
 D  A  R  R  H  V  A  F  G  Y  G  V  H  L  C  I  G
```

Figure 20. Pairwise alignment between translated amino acid sequences of SMC 48 and P-450 from S. avermitilis (AP005027). Block indicated the oxygen-binding and heme-binding motif of P-450. Consensus of amino acids is represented by asterisk (*)

```
SMC 48      ------------------------------------------------------------
AP005027    MADALAGRAPDATPPVAAYPMPRAVSCPLAPPPALQPLRDEQPITKVRIWNGSTPWLITR


SMC 48      ---------------------------------------------------GRLRRLVTAPFA
AP005027    HADQRALLTDPRVSNDDRDPGFPYVNAHRAEIAHATPRLITNTDAPEHTRLRRTVNAPFL
                                                               ****  * ***

SMC 48      IKKVEALRPSVQKIVDDLIDAMLAGPKPVDLVEAFALPVPSLVICELLGVPYEDHDLFQR
AP005027    IKRIEAMRPAVQRIVDGLIDDMLDGPNPADLLTALALPVPSLVIAQLLGVPYADHEFFQR
            **:.**:**:**:.***.*** ** **:*.**: *:*********.:****** **:.***

SMC 48      NTRVAVRRTSTAEEVVAAMQTLTDYLDDLLTVKRARPGDDLLSVLATERVAKGEMTQRQA
AP005027    NSNLVLDNSVPAEEARAASGALAAYLDTLLAEKTAAPDADVLSEMG-GRIKAGEMTHREA
            *:.:.:  .: .***. **  :*: *** **: * * *. *:**  :.  *: ****:*:*

                oxygen-binding domain
SMC 48      AETGVLLLATGHETTANMIALGTLALLRNPDQLALPRDTEDPKVVAGAVEELLRYLNITH
AP005027    VHMGVAMLIAGHETTATMISLGTLALLEHPEQLAVLRDAEDPKVIAAAVEELLRYLTIVH
            .. **  :* :.****.**.**:*******..:*:*** **:*****:* .*********.*.*

                                                        Heme-binding domain
SMC 48      SGRRRVALADIEVGGEVIRAGDGVIFANEIANRDPEAFSDPDRLDITRDARRHVAFGYGV
AP005027    SGIRRVAKEDIEIGDRVIGAGDGLLFDLHAANWDSEAFPEADRLDLSRPARHHQAFGYGP
            ** ****  ***:*..** ****::*  . ** *.***.::.****::* **:* *****

SMC 48      HLCIG------------------------------------------------------
AP005027    HQCLGQSLARLELQVVYGTLYRRFPTLRLAAPIDRLAFHHTGTTYGVRCLPVTW
            *:*
```

Figure 22.   Multiple sequence alignment of predicted secondary structures of putative P-450 of

SMC 48, P-450 from *S. avermitilis* (AP005027) and P-450eryF from *S. erythraea* (M54983)

```
                                                          A helix      1 sheet
SMC 48          ------------------------------------------ --------------- --------
AP005027        ------HHH--------------------------------- --------------- -EEEE---
M54983          ------------------------------------------E--HH--HH------ ---EE---
M54983_known    ------------------------------------------HHHHHHHHHHHH-HH EEEEEEE-E

                              B helix                          B'helix
SMC 48          -------     ---------------------------------- -------     ----
AP005027        ---EEEE     ------------------------E---HH HHHH----- --EE
M54983          --HEEE-     --HHHHHHHHH------------EEE------ ------E-     -----
M54983_known    EEEEEEE     HHHHHHHHHHH--------------------- HHHHHHHH     ----

                    C helix          D helix          2 sheet
SMC 48          -----  ------EE---- HHHH-H--------EHHHHHHHHH -------HHH -----
AP005027        -----  ------------ HHHH-H----HHHHEH---H---- ------HHHHH -----
M54983          -----  ------HH---H EHHHH-------HHHHHHHHH-- ---EEEE HH -----
M54983_known    -----  HHHHHHHHHHHH HHHHHHHHHHHHHHHHHHHHHHH E--EEEE --HHHHH

                 E helix          F helix          G helix
SMC 48          ---EEEE---  ------- HHHHHHHEEE ---  --HHHHHHHHHHHHHHHHHH-------
AP005027        --HHHHEH--  ------- HHHH----E- ---  --HHHHHHHHHHHHHHHHHH-----
M54983          --HEEEEHH-  ------- ------EEE ---  --HHH---H-HHHHHHHHH-------
M54983_known    HHHHHHHHHH  ------- HHHHHHHHHH ---  HHHHHHHHHHHHHHHHHHHHHHHH-H

                  H helix 3 sheet          I helix                  J helix
SMC 48          --HHEHHHHH H ---- HHHHH-HHHHHEEE------HHHHHHHHHHH--- ------
AP005027        --HHHHH--- E ---- HHHHHHHHHHHHEEEH-----HEEEEH-HHHHHH-  -HHHHH
M54983          --HHHHHHE- ----- ---HHHHHHHEHHH-HHH-EEEEE--EEEE--- ----HE-
M54983_known    --HHHHHHHHH EEEEE HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH HHHHHH

                  K helix     4 sheet    5 sheet    6 sheet
SMC 48          ----- -EEH-HHHHHHHHHE -- ----EE- HH---HEEE ---EEHH -------
AP005027        ----- HHHHHHHHHHHHHHHEEH --HHHH- -HHHH--EEE ----HHHHH -------
M54983          ----- -----HHHHHHH-- ---HHHHHHHH---E- --EEEEH- -------
M54983_known    H--HH- HHHHHHHHHHHH EEEE-EEE---EEEEEEEE- EEEEEEEE -------

                                                L helix      7 sheet
SMC 48          ------------HHHHHHH----EEEE --------------------- --- -------
AP005027        ---HH--------HH--------- HHHHHHHEEH--E---- --- HHHHH
M54983          --------------------EE- -----HHHHHHHHHHH--- --- ----H
M54983_known    --------------------- HHHHHHHHHHHHHHHHHHHH EEE -----

                         8 sheet
SMC 48          -------------- ---- ----------
AP005027        HE------E-EE- ---- ----------
M54983          HHHHHH-E----- ---- ----------
M54983_known    ------------- EEEEE ----------
```

Figure 23. Result from Basic Local alignment Search for nucleotide similarity to nucleotide sequence of SMC 59

Sequences producing significant alignments:	Score (bits)	E Value
gi|2247711|gb|AF293355.2| Streptomyces collinus DSM2012 rubrinomycin gene cluster, complete sequence	60	7e-06
gi|24413773|emb|AL939114.1|SCO939114 Streptomyces coelicolor A3(2) complete genome; segment 11/29	58	3e-05
gi|48941|emb|X60379.1|SEERYABIO S.erythraea erythromycin A biosynthesis gene cluster encoding hydrolase; thioesterase; methylase	50	0.006
gi|152696|gb|M54983.1|SERERYEGH Saccharopolyspora erythraea EryF (eryF) gene, complete cds; EryG (eryG) gene, partial cds; and unknown gene	50	0.006
gi|26004851|dbj|AP005026.1| Streptomyces avermitilis genomic DNA, complete genome, section 6/30	48	0.025
gi|24426505|emb|AL939115.1|SCO939115 Streptomyces coelicolor A3(2) complete genome; segment 12/29	46	0.10
gi|15029316|gb|AF386507.1|AF386507 Streptomyces lavendulae complestatin biosynthetic gene cluster,complete sequence	44	0.40
gi|24429533|emb|AL939121.1|SCO939121 Streptomyces coelicolor A3(2) complete genome; segment 18/29	42	1.6
gi|3746883|gb|AF087022.1|AF087022 Streptomyces venezuelae cytochrome P450 monooxygenase (picK) gene, complete cds	42	1.6
gi|29609004|dbj|AP005046.1| Streptomyces avermitilis genomic DNA, complete genome, section 26/30	42	1.6

Figure 24. Result from Basic Local alignment Search for translated amino acid similarity to translated amino acid sequence of SMC 59

Sequences producing significant alignments:

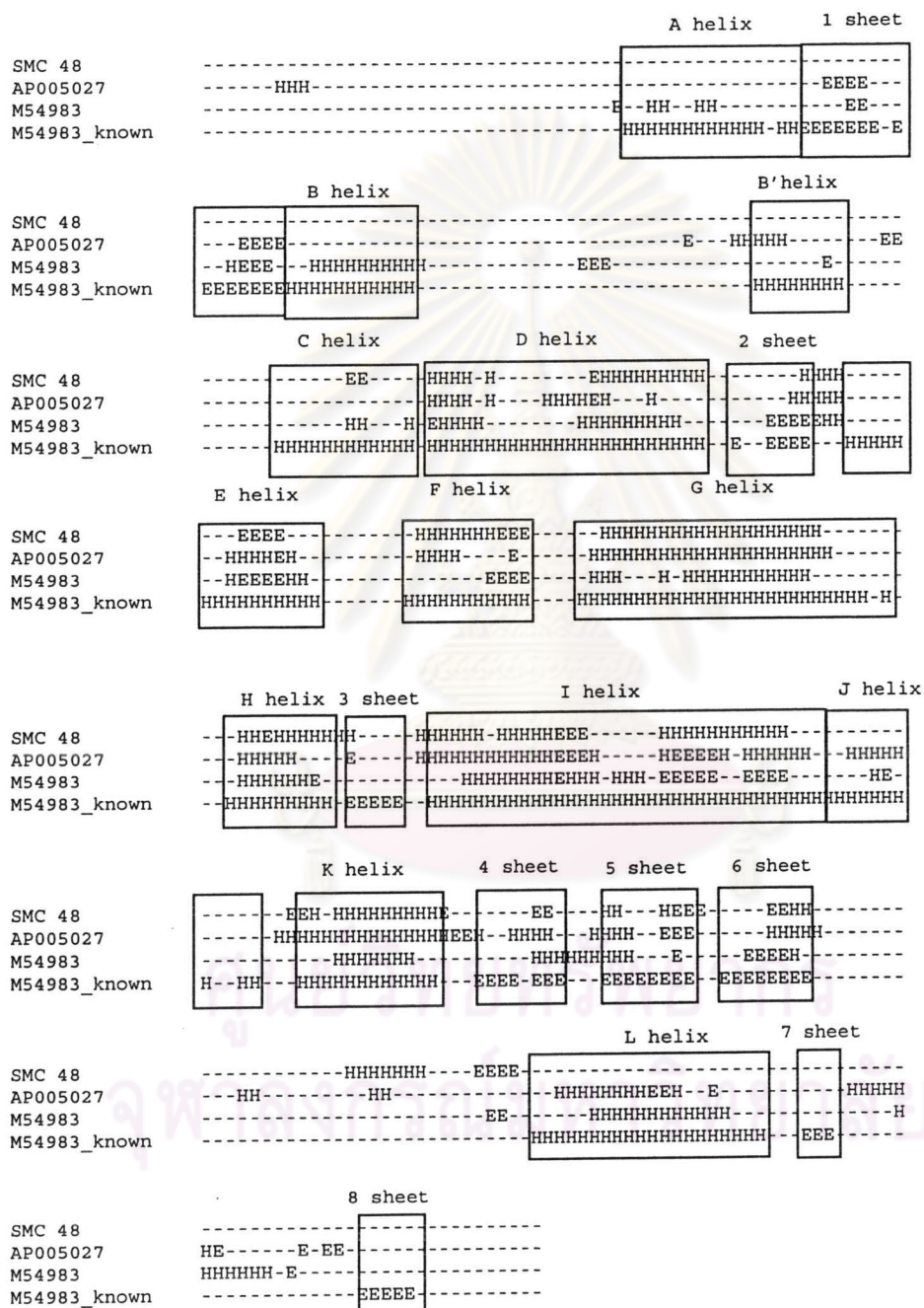|	Score (bits)	E Value
gii24413773lemblAL939114.1ISCO939114 *Streptomyces coelicolor A3(2) complete genome; segment 11/29*	138	6e-53
gii519317lgblAF145049.1IAF145049 *Streptomyces fradiae tylosin-biosynthetic regulatory gene cluster,complete sequence*	143	4e-51
gii30698345ldbjlAB088224.1I *Streptomyces rochei plasmid pSLA2-L DNA, complete sequence*	152	4e-49
gii29604851ldbjlAP005026.1I *Streptomyces avermitilis genomic DNA, complete genome, section 6/30*	158	9e-44
gii38008039lgblAF079139.1IAF079139 *Streptomyces venezuelae pikCD operon, complete sequence*	104	2e-43
gii3746883lgblAF087022.1IAF087022 *Streptomyces venezuelae cytochrome P450 monooxygenase (pick) gene, complete cds*	104	2e-43
gii29605609ldbjlAP005029.1I *Streptomyces avermitilis genomic DNA, complete genome, section 9/30*	95	4e-40
gii29603636ldbjlAP005021.1I *Streptomyces avermitilis genomic DNA, complete genome, section 1/30*	119	2e-38
gii152696lgblM54983.1ISERERYFGH *Saccharopolyspora erythraea EryF (eryF) gene, complete cds;* EryG(eryG) gene, partial cds; and unknown gene	113	2e-38
gii29607295ldbjlAP005036.1I *Streptomyces avermitilis genomic DNA, complete genome, section 16/30*	138	6e-38

Figure 25. Result from Basic Local alignment Search for protein similarity to translated amino acid sequence of SMC 59

Sequences producing significant alignments:		Score (bits)	E Value
gi|11249736|pir||T44587	cytochrome P450 homolog [imported] - *Streptomyces fradiae*	202	3e-51
gi|30795056|ref|NP_851506.1|	P450-like hydroxylase [*Streptomyces roche*]	201	1e-50
gi|21221333|ref|NP_627112.1|	putative cytochrome P450. [*Streptomyces coelicolor A3(2)*]	196	3e-49
gi|29830247|ref|NP_824881.1|	putative cytochrome P450 [*Streptomyces avermitilis*]	186	3e-46
gi|29827849|ref|NP_822483.1|	putative cytochrome P450 [*Streptomyces avermitilis MA-4680*]	182	7e-45
gi|28192484|gb|AAM77997.1|	cytochrome P-450 [*Streptomyces carzinostaticus subsp. neocarzinostaticus*]	179	4e-44
gi|23018492|ref|ZP_00058208.1|	Cytochrome P450 [*Thermobifida fusca*]	177	1e-43
gi|21225687|ref|NP_631466.1|	putative cytochrome P450-family protein. [*Streptomyces coelicolor A3(2)*]	174	1e-42
gi|28373301|pdb|1GWI|A	Chain A, The 1.92 A Structure Of *Streptomyces Coelicolor A3(2)* Cyp154c1: A New Monooxygenase That Functionalizes	174	1e-42
gi|23019921|ref|ZP_00059630.1|	COG2124: Cytochrome P450 [*Thermobifida fusca*]	171	1e-41

Figure 26. Pairwise alignment of nucleotide sequence of SMC 59 and P-450 from *S. coelicolor* A3(2) (AL939114)

```
SMC 59      ------------------------------------------------------------
AL939114    ATGGCGACCCAGCAGCCCGCCCTCGTCCTCGACCCCACCGGCGCCGACCACCACACCGAG


SMC 59      ------------------------------------------------------------
AL939114    CACCGCACCCTGCGGGAGGGCGGCCCGGCCACCTGGGTGGACGTCCTCGGGGTGCAGGCC


SMC 59      ------------------------------------------------------------
AL939114    TGGTCGGTCAGCGACCCCGTCCTCCTCAAGCAGCTGCTCACCAGCTCCGACGTCTCCAAG


SMC 59      ------------------------------------------------------------
AL939114    GACGCCCGGGCGCACTGGCCCGCCTTCGGGGAGGTCGTCGGCACCTGGCCGCTGGCCCTG


SMC 59      -----------------------------------------------GGCCGGCTGCGCAAG
AL939114    TGGGTGGCGGTGGAGAACATGTTCACCGCGTACGGGCCCAATCACCGCAAGCTGCGCCGG
                                                            **  ******* *

SMC 59      CTGGTGGCTCCCAGCTTCACGCACCGCCGCACCGAGGCGATGCGGCCCCGCGTGGAGGCC
AL939114    CTGGTGGCGCCCGCCTTCAGCGCCCGTCGCGTCGACGCGATGCGGCCGGCCGTCGAGGCG
            ******** ***  *****   *** ***  *** ********** ***  *** *****

SMC 59      ATCACGGCGGAGCTGCTGGACGCGCTCGACGCGAG---CGGCGACGGGGTCGTGGACATC
AL939114    ATGGTGACCGGTCTCGTCGACCGGCTCGCCGAGCTTCCCGCCGGTGAGCCGGTGGACCTG
            **    *  *   **   * ***  ***** **  *     **  **   *   ****** *

SMC 59      AAGGAGGGGTTCGCCCATCCGCTTCCGATGCGGATGATCTGCGAACTGTTCGGCGTGCCG
AL939114    CGGCAGGAGCTGGCCTACCCGCTGCCCATCGCGGTGATCGGTCACCTCATGGGCGTGCCT
              * *** *  * *** ***** ** **  * ***** *    * **  * ** ********

SMC 59      GACGAGTTGCGCGAGGCCACCGGCCGGATGGTCGCGGCCATCATGGACACCACCG-ACAC
AL939114    CAGGACCGGCGCGACGGCTTCCGCGCGCCCTCGTGGACGGCGTCTTCGACACCACCCTGGAC
            * **  ******* * *  * ***     * **  * * ** * ********* **

SMC 59      GAGCCCGGAGCACGCGGCGTCCGTCCAGCAGCAGATCGGCACCGTGCTGCCCGCGCTGAT
AL939114    CAGGCCGAGGCCCA-GGCCAACACCGCGCGCCTGTACGAGGTCCT--CGACCA-GCTCAT
            ** * ***  ** *    **  **  *  * ** **    * ** *  *** ***  **

SMC 59      CACCCACAAGGCCGCGCACCCCGGTGACGACATGACCACCGAGCTGATCCGGGTCCGGGA
AL939114    CGCGGCCAAGCGCGCCACCCCGGGCGACGACATGACCTCGCTGCTCATAGCCGCGCGGGA
            * *  **** *** ***  *** **  *********** *  *** **     *  *****

SMC 59      CGA------GGACGGCGACCGGCTCAGCGACGAGGAGCTGCTGTACACGCTGCTGCTGGT
AL939114    CGACGAGGGGGACGGCGACCGGCTCTCCCCCGAGGAGCTGCGCGACACCCTGCTGCTGAT
            ***       ************** *  * * ********** ****  ********* *

SMC 59      CATCGGGGCGGGGTACGAGACCACGGTCAACCTCATCGGCAACGCGGCGGTCGCCCTGCT
AL939114    GATCAGCGCCGGGTACGAGACCACCGTCAACGTCATCGACCAGGCCGTGCACACCCTGCT
            *** * ** ************** ******  ***** * *  *  * ** * ********

SMC 59      CCGCCGTCCCGAGCAGCTGGCGGCGGTCCGGTCCGGGGAGATCGGCTGGGACGCCGTCGT
AL939114    GACCCGCCCCGACCAGCTCGCGCTGGTCCGCAAGGGCGAGGTCACCTGGGCGGACGTGGT
            *** ***** ***** ***  *  ****** **  * *  * **  * **** * ** **
```

```
SMC 59     GGACGAGACGCTGCGCGCGCACCCGTCGATCGCCTCGCTGCCGCTCCGGTTCGCCGTCAC
AL939114   GGAGGAGACGCTGCGCCACGAACCGGCGGTCAAGCACCTGCCGCTGCGGTACGCGGTCAC
           *** *********** * *** ** **      ******** **** *** *****

SMC 59     CGACCTCACCGTCGGCGACGTCAGGGTCCCGGCCGGCGACG---CCATCATCACGACGTA
AL939114   CGACATCGCCCTGCCGGACGGGCGGACCATCGCCCGCGGGGAGCCGATCCTCGCCTCGTA
           **** ** ** *    ****   **   *   *** ***   *  * *** ** *  ****

SMC 59     CGCCGCCGCGGGGCTCGACCCCGACCACTACGGGCCGGACGCCGACGTCCTCGACGCCAC
AL939114   CGCCGCCGCCAACCGCCATCCGGACTGGCACGA---GGACGCCGACACCTTCGACGCGAC
           *********   * * ** ***   ***    ********** * ******* **

SMC 59     CCGGGCCTCAGACGACCATCTCGCGTTCGGGATCGGCGTCCACCACTGCATCGGCG----
AL939114   CCGCACCGTCAAGGAGCACCTGGCCTTCGGCCACGGCGTCCACTTCTGCCTGGGCGCGCC
           ***  **   * ** ** ** ** *****  ********** **** * ****

SMC 59     ------------------------------------------------------------
AL939114   GCTGGCCCGCATGGAGGTCACCCTCGCGCTGGAGAGTCTCTTCGGCCGCTTCCCGGACCT

SMC 59     ------------------------------------------------------------
AL939114   CCGCCTCGCCGATCCGGCCGAGGAGCTGCCGCCCGTGCCCTCCCTGATCAGCAACGGCCA

SMC 59     ------------------------------------
AL939114   CCAGCGGCTCCCGGTCCTGCTGCACGCCGGCTGA
```

Figure 27.   DNA sequence of putative macrolide P-450 from SMC 59 and its translated amino

acid sequence

```
             10        20        30        40        50        60
GGCCGGCTGCGCAAGCTGGTGGCTCCCAGCTTCACGCACCGCCGCACCGAGGCGATGCGG
 G   R   L   R   K   L   V   A   P   S   F   T   H   R   R   T   E   A   M   R

             70        80        90       100       110       120
CCCCGCGTGGAGGCCATCACGGCGGAGCTGCTGGACGCGCTCGACGCGAGCGGCGACGGG
 P   R   V   E   A   I   T   A   E   L   L   D   A   L   D   A   S   G   D   G

            130       140       150       160       170       180
GTCGTGGACATCAAGGAGGGGTTCGCCCATCCGCTTCCGATGCGGATGATCTGCGAACTG
 V   V   D   I   K   E   G   F   A   H   P   L   P   M   R   M   I   C   E   L

            190       200       210       220       230       240
TTCGGCGTGCCGGACGAGTTGCGCGAGGCCACCGGCCGGATGGTCGCGGCCATCATGGAC
 F   G   V   P   D   E   L   R   E   A   T   G   R   M   V   A   A   I   M   D

            250       260       270       280       290       300
ACCACCGACACGAGCCCGGAGCACGCGGCGTCCGTCCAGCAGCAGATCGGCACCGTGCTG
 T   T   D   T   S   P   E   H   A   A   S   V   Q   Q   Q   I   G   T   V   L

            310       320       330       340       350       360
CCCGCGCTGATCACCCACAAGGCCGCGCACCCCGGTGACGACATGACCACCGAGCTGATC
 P   A   L   I   T   H   K   A   A   H   P   G   D   D   M   T   T   E   L   I

            370       380       390       400       410       420
CGGGTCCGGGACGAGGACGGCGACCGGCTCAGCGACGAGGAGCTGCTGTACACGCTGCTG
 R   V   R   D   E   D   G   D   R   L   S   D   E   E   L   L   Y   T   L   L

            430       440       450       460       470       480
CTGGTCATCGGGGCGGGGTACGAGACCACGGTCAACCTCATCGGCAACGCGGCGGTCGCC
 L   V   I   G   A   G   Y   E   T   T   V   N   L   I   G   N   A   A   V   A

            490       500       510       520       530       540
CTGCTCCGCCGTCCCGAGCAGCTGGCGGCGGTCCGGTCCGGGGAGATCGGCTGGGACGCC
 L   L   R   R   P   E   Q   L   A   A   V   R   S   G   E   I   G   W   D   A

            550       560       570       580       590       600
GTCGTGGACGAGACGCTGCGCGCGCACCCGTCGATCGCCTCGCTGCCGCTCCGGTTCGCC
 V   V   D   E   T   L   R   A   H   P   S   I   A   S   L   P   L   R   F   A

            610       620       630       640       650       660
GTCACCGACCTCACCGTCGGCGACGTCAGGGTCCCGGCCGGCGACGCCATCATCACGACG
 V   T   D   L   T   V   G   D   V   R   V   P   A   G   D   A   I   I   T   T

            670       680       690       700       710       720
TACGCCGCCGCGGGGCTCGACCCCGACCACTACGGGCCGGACGCCGACGTCCTCGACGCC
 Y   A   A   A   G   L   D   P   D   H   Y   G   P   D   A   D   V   L   D   A

            730       740       750       760       770       780
ACCCGGGCCTCAGACGACCATCTCGCGTTCGGGATCGGCGTCCACCACTGCATCGGC
 T   R   A   S   D   D   H   L   A   F   G   I   G   V   H   H   C   I   G
```

Figure 28. Pairwise alignment between translated amino acid sequences of SMC 59 and P-450 from *S. coelicolor* A3(2) (AL939114). Block indicated the oxygen-binding and heme-binding motif of P-450. Consensus of amino acids is represented by asterisk (*)

```
SMC 59      ------------------------------------------------------------
AL939114    MATQQPALVLDPTGADHHTEHRTLREGGPATWVDVLGVQAWSVSDPVLLKQLLTSSDVSK


SMC 59      ------------------------------G----RLRKLVAPSFTHRRTEAMRPRVEA
AL939114    DARAHWPAFGEVVGTWPLALWVAVENMFTAYGPNHRKLRRLVAPAFSARRVDAMRPAVEA
                                            *   :**:****:*: **.:**** ***

SMC 59      ITAELLDALDASGDGV-VDIKEGFAHPLPMRMICELFGVPDELREATGRMVAAIMDTTDT
AL939114    MVTGLVDRLAELPAGEPVDLRQELAYPLPIAVIGHLMGVPQDRRDGFRALVDGVFDTTLD
            :.: *:* *      * **::: :*:***: :* .*:***:: *:.   :* .::***

SMC 59      SPEHAASVQQQIGTVLPALITHKAAHPGDDMTTELIRVRDE--DGDRLSDEELLYTLLLV
AL939114    QAEAQANTARLY-EVLDQLIAAKRATPGDDMTSLLIAARDDEGDGDRLSPEELRDTLLLM
            ..*  *.. :  ** **: * * ******: ** .**: ****** *** ****:

                       Oxygen-binding domain
SMC 59      IGAGYETTVNLIGNAAVALLRRPEQLAAVRSGEIGWDAVVDETLRAHPSIASLPLRFAVT
AL939114    ISAGYETTVNVIDQAVHTLLTRPDQLALVRKGEVTWADVVEETLRHEPAVKHLPLRYAVT
            *.******:**:*.:*. :** **:*** **.** *  **:**** .*:: ****:***

                                                Heme-binding domain
SMC 59      DLTVGDVR-VPAGDAIITTYAAAGLDPDHYGPDADVLDATRASDDHLAFGIGVHHCIG--
AL939114    DIALPDGRTIARGEPILASYAAANRHPDWH-EDADTFDATRTVKEHLAFGHGVHFCLGAP
            *::: * * :. *:.*::****. .** :  ***.:****: .:***** ***.*:*

SMC 59      ------------------------------------------------------------
AL939114    LARMEVTLALESLFGRFPDLRLADPAEELPPVPSLISNGHQRLPVLLHAG
```

Figure 29.   Multiple sequence alignment of predicted secondary structures of putative P-450

of SMC 59, P-450 from *S. coelicolor* A3(2) (AL939114) and P-450eryF from *S. erythraea*

(M54983)

```
SMC 59          ------------------------------------------------------------
AL939114        ------------------------------------------------------------
M54983          ------------------------------------------------------------
M54983_known    ------------------------------------------------------------

                                  A helix            1 sheet          B helix
SMC 59          -----------------------------------------------------------
AL939114        --------------E------------------------EEEHH--EE--------HHHHH
M54983          -------------E--HH--HH------------------EE----H---EEE--HHHHH
M54983_known    -------------HHHHHHHHHH-HHH-EEEEEEEEEEEE---EEE-HHHHHHH

                                            B'helix          C helix
SMC 59          --------------------------------------------------------
AL939114        H-------H----------HH-------E----HH----EHHHHHHH-------HHH
M54983          HHHHHH---------------------EEE---------------E--------HHH
M54983_known    HHHH------------------------------HHHHHHHH----HHHHH

                          D helix                    2 sheet        E helix
SMC 59          -EE-------------------HHHHHHHHHHH-----E-EEHHH-------HHHH
AL939114        HHHH-----------------HHHHHHHHHH-----------H-H------EEEE
M54983          -HH---H-EHHHH-------HHHHHHHHH----------E-EEEHHH------HEEEEH
M54983_known    HHHHHHH-HHHHHHHHHHHHHHHHHHHHHHHH-----EE-EEE---HHHHHHHHHHHH

                          F helix                  G helix               H helix
SMC 59          --------HHHHHHHHEEE--------H-HHHE---E--HHHH-H---------HHHH
AL939114        ---------EEE--------HHHHHH-HHHHHHHHH---------------HHHH
M54983          H-------------EEEE----HHH---H-HHHHHHHHHHHH----------HHHHH
M54983_known    HH------HHHHHHHHHHH---HHHHHHHHHHHHHHHHHHH-HHHHHH--HHHHHH
                LGVDEKYRGEFGRWSSEILVMDPERAEQRGQAAREVVNFILDLVE-RRRTEPGDDLLSAL
                *.                   .                        : :    :      *:    :

                  3 sheet               I helix              J helix
SMC 59          EEE-------HHHHHHHHEEE------EEEHHHHHHHHHH-HHHH-H----
AL939114        EE-------HHHHHHEEE------EEHHHHHHH---------HHHH---H-
M54983          HE------HHHHHHHHEHHH-HHH-EEEEE--EEEE-------HE----
M54983_known    HHH-EEEEE-HHHHHHHHHHHHHHHHHH-HHH-EEEEE--EEEE-HHHHHHHHHHHH---

                  K helix       4 sheet   5 sheet   6 sheet
SMC 59          -E-HHHHHH-------HEE-----------EEEHH-----------
AL939114        -HHHHHHHH---------E----------HHHHHH----------
M54983          ---HHHHHH-------------HHHHHHHH----E---EEEEH-----------
M54983_known    HHHHHHHHHHHH--EEEE-EEE--EEEE-EEEE-EEEEEEE------------

                                              L helix    7 sheet
SMC 59          -HHH---------HHE----E---------------------------
AL939114        ----HHHHH----HHHH----EEE------HHHHHHHHHHH-------------
M54983          -----------------------EE----------HHHHHHHHHHH-----------H
M54983_known    ----------------------------HHHHHHHHHHHHHHHHHHHHHH--EEE-----

                              8 sheet
SMC 59          ------------------------------------
AL939114        --------------HHHH-------------
M54983          HHHHHH-E-------------------------
M54983_known    --------------EEEE-------------
```

Figure 30. Result from Basic Local alignment Search for nucleotide similarity to nucleotide sequence of SMC 78

Sequences producing significant alignments:		Score (bits)	E Value
gi|24413773|emb|AL939114.1|SCO939114	Streptomyces coelicolor A3(2) complete genome; segment 11/29	68	3e-08
gi|45124719|emb|AJ605542.1|	Streptomyces peucetius cyp0879 gene for putative cytochrome P450,strain ATCC 27952	54	4e-04
gi|45124717|emb|AJ605541.1|	Streptomyces peucetius cyp0819 gene for putative cytochrome P450,strain ATCC 27952	52	0.002
gi|48941|emb|X60379.1|SEERYABIO	S.erythraea erythromycin A biosynthesis gene cluster encoding hydrolase; thioesterase; methylase	48	0.026
gi|152696|gb|M54983.1|SERERYFGH	Saccharopolyspora erythraea EryF (eryF) gene, complete cds; EryG(eryG) gene, partial cds; and unknown gene	48	0.026
gi|6933878|gb|AF071516.1|AF071516	Streptomyces noursei putative P450 hydroxylase gene, partial cds	48	0.026
gi|29603636|dbj|AP005021.1|	Streptomyces avermitilis genomic DNA, complete genome, section 1/30	48	0.026
gi|45124711|emb|AJ605538.1|	Streptomyces peucetius cyp0673 gene for putative cytochrome P450,strain ATCC 27952	46	0.10
gi|24413781|emb|AL939116.1|SCO939116	Streptomyces coelicolor A3(2) complete genome; segment 13/29	44	0.41
gi|29603894|dbj|AP005022.1|	Streptomyces avermitilis genomic DNA, complete genome, section 2/30	44	0.41

Figure 31. Result from Basic Local alignment Search for translated amino acid similarity to translated amino acid sequence of SMC 78

Sequences producing significant alignments:	Score (bits)	E Value
gii30698345ldbjiAB088224.1 Streptomyces rochei plasmid pSLA2-L DNA, complete sequence	159	5e-48 3
gii5199317lgbiAF145049.1IAF145049 Streptomyces fradiae tylosin-biosynthetic regulatory gene cluster,complete sequence	138	3e-44 3
gii45124711lembiAJ605538.1 Streptomyces peucetius cyp0673 gene for putative cytochrome P450,strain ATCC 27952	84	2e-42 4
gii29604851ldbjiAP005026.1 Streptomyces avermitilis genomic DNA, complete genome, section 6/30	159	3e-41 2
gii24413773lembiAL939114.1ISCO939114 Streptomyces coelicolor A3(2) complete genome; segment 11/29	92	5e-41 4
gii15269961lgbiM54983.1ISERERYFGH Saccharopolyspora erythraea EryF (eryF) gene, complete genome; EryG(eryG) gene, partial cds; and unknown gene	113	1e-40 4
gii45124735lembiAJ605550.1 Streptomyces peucetius cyp0854 gene for putative cytochrome P450,strain ATCC 27952	101	4e-40 5
gii152682lgbiM83110.1ISERCP450A Saccharopolyspora erythraea ORF 1 gene, partial cds; cytochrome P-450 gene, complete cds; ORF 2 gene, partial cds	82	2e-39 5
gii29603636ldbjiAP005021.1 Streptomyces avermitilis genomic DNA, complete genome, section 1/30	76	5e-39 5
gii24418961lembiAL939131.1ISCO939131 Streptomyces coelicolor A3(2) complete genome; segment 28/29	119	1e-36 4

Figure 32. Result from Basic Local alignment Search for protein similarity to translated amino acid sequence of SMC 78

Sequences producing significant alignments:		Score (bits)	E Value
gi|30698345|dbj|AB088224.1|	Streptomyces rochei plasmid pSLA2-L DNA, complete sequence	159	5e-48 3
gi|5199317|gb|AF145049.1|AF145049	Streptomyces fradiae tylosin-biosynthetic regulatory gene cluster, complete sequence	138	3e-44 3
gi|4512471|emb|AJ605538.1|	Streptomyces peucetius cyp0673 gene for putative cytochrome P450, strain ATCC 27952	84	2e-42 4
gi|29604851|dbj|AP005026.1|	Streptomyces avermitilis genomic DNA, complete genome, section 6/30	159	3e-41 2
gi|24413773|emb|AL939114.1|SCO939114	Streptomyces coelicolor A3(2) complete genome; segment 11/29	92	5e-41 4
gi|15269696|gb|M54983.1|SERERYFGH	Saccharopolyspora erythraea EryF (eryF) gene, complete cds; EryG (eryG) gene, partial cds; and unknown gene	113	1e-40 4
gi|45124735|emb|AJ605550.1|	Streptomyces peucetius cyp0854 gene for putative cytochrome P450, strain ATCC 27952	101	4e-40 5
gi|152682|gb|M83110.1|SERCP450A	Saccharopolyspora erythraea ORF 1 gene, complete cds; ORF 2 gene, partial cds; cytochrome P-450 gene, complete cds	82	2e-39 5
gi|29603636|dbj|AP005021.1|	Streptomyces avermitilis genomic DNA, complete genome, section 1/30	76	5e-39 5
gi|24418961|emb|AL939131.1|SCO939131	Streptomyces coelicolor A3(2) complete genome; segment 28/29	119	1e-36 4

Figure 33. Pairwise alignment of nucleotide sequence of SMC 78 and P-450 from *S. rochei* (AB088224)

```
SMC 78      --------------------------------------------------AGCCGGCTGCGCAAGCT
AB088224    TCATGCCGCGGCTCCGGGCACCAGCACGACCGGCAGTGCGACATGCCCGCTGGAGATGAA
                                                         *** **** * *

SMC 78      GG--TGGCGCCC----AGTTTCACGCACCGCCGGACCGAGGCGATGCGG---CCCCGGGT
AB088224    GGACTCGCGCACCGCGAGTTCCGCCTCCGGCACCGCCGCCGCCATCCGGGGATAGCGCGC
            **  * **** *   **** * *   * **   *** ** ** ***      ** *

SMC 78      GGAGA-CGATCACGG------CGGAGCTGCTGGCTGCGCTGGAGGCCG-CGGAGGGCGAG
AB088224    GAACAGCGACTCCAGGGCGATCGAGACCTCCAGCCGGGCCAGCGGCCGGCCGACGCAGTG
            * * ***  * *    **   * *  ** *** * ***** * **   * *

SMC 78      GTCGTGGACGTCAAGGAGGGGCTCGCCCACCCGCTGCCGATGCGGATGATCTGCGAACTC
AB088224    GTGGACGCCGTAACCGAAGGGGACGTGGTCCCGGCGGCTCGGCCGGGCCACGTCGAACTG
            ** *  * *** *  ** *** **   **** * *   ** *    *  * ******

SMC 78      ATGGGTGTGCCGGACGCGTTGCGGGAGGACACCGGCCGCCTCATC----GCGGCCATCAT
AB088224    GTCGGCGTCCGCGCCGTGCCGGTGCGGGCAGCGGGCTGCCGCGGCCAGGGAGAGCAGGAT
             * ** ** * ** *  * **   * **  * *** *** * *     * *  **  **

SMC 78      GGACAC---CTCCGA---CCCGAGCCCGGAGCACGCGGCGTCCGTG-CAGCGG--CAGAT
AB088224    GGGATCGCCCTTCGGGATCGTGACCCCGCCGACCTCGATGTCCTCGACCGCGAAACGGAA
            **  *   ** **    *  ** ****  *   ** * ****  * * ***    * **

SMC 78      CGGGACGGTGCTGCC--CGCGCTGATCGCCCACCGGGCCGCGCA---------CCCCGGG
AB088224    CGGGACATTGGCGACGACGGGTTGATGGCGCAGCGTTTCGTCGATGACGCTGTCCCAGGT
            ******  **  * *  ** * **** ** ** **   **  *         *** **

SMC 78      GACGACATCAC----GACCG-----AGCTCATCCGGGTCCGCGACGAGGACGGCGACCGG
AB088224    GGCCTCGCCGCCGAGGACCAGCGCCAGTTGGGCGGGGTGGGTGATCAGCCCGTGCACGGC
            * * * *   ****    ****   ** *  * **** * **  **  **    ** *

SMC 78      CTGAGCGACGAGGA----GCTGCTCTACACGCTGCTGCTGGTGATCGGCGCCGGGTT---
AB088224    CTGGTCGATGAGATTGACGGTGGTCTCCGTACCCGCGCCGATGAGCAGCAGCAGGGTGTC
            *** *** ***      * ** *** *       *  ** * *** * **  * ** *

SMC 78      --CGAGACCACCGTGAATCACATCGGCAACGCGGTGGTCGCCCTGCTCCGCCACCCCGAG
AB088224    GACGAGCTCCTTCTCGCTCAGCCCCTCTCCCCCGTCCTCGTCCCGCGCCGCTATC--AGT
              ****  *    * ***   *  *   ***  *** *   ** * * *** * *

SMC 78      CAGCTGGCGGCCGTGCGGTCCGGGGAGATCGGCTGGGACGCCGTC------GTCGACGAG
AB088224    TCGCTGGTGAGGTCGGTGCCGGGGGGCGCTGCGCCGGGCGGCCACCAGTTCCGTCAGCAGT
            ***** *  *  * **   ** ***** ** *** ** *** ***    *

SMC 78      ACGCTGCGCGCCCACCCGTCGAT-CGCGTCGCCGCCGCTGCGGTTCGCCGTCACCGACCT
AB088224    GCGTACAGCTCGTACTGGTTGGCACGGGCCTCCTCGGGCTCGGCT-GTCGTCTTGAACAG
            **     ** * **  * *    * ** ** *    * *** ** *     ** **

SMC 78      CGCCGTCGGCGAGGTGACG-ATCCCGGCCGGCGACGCCATCACCACGACGTACGCCGCGG
AB088224    CAC-GTCGACCATGCGCTGGAGGTCGGCGTGGAGGTCCTCCGGCAGCCCCAGCGTCTCGC
            * * **** * * *   * **   **** *  **   * *  *  *  ** * **

SMC 78      CGGGGCTGGATCCCGAGCACTACGGGACGGAC--GCCGACCGCTTCGACG-CCACGC---
AB088224    CCATGACGAGCACCGACAGGGGGCGGGCGAACTCGTTGCGCACGTCGACGGCCGCACCGG
            *   *   **** **    ** ** ** *   * *  * ******* ** * *
```

```
SMC 78      GGGGCGCCGACGACCATCTGGCCTTCGGCATCGGCGTGCACCACTGCATCGGCG------
AB088224    GGGGCACCGCGGCCAGGCCGTCCAGCAGCCGCGCGGTGATCCGCTCCACCCGCGGCCGCA
            ***** ***   *  *    * *  **   *  **   **   ***  ** ** **  * ***

SMC 78      ------------------------------------------------------------
AB088224    GGTCGGCGGTGCGGCGCGCGGTGAACGCGCCGGCCACCAGGCGGCGCAGCCGGGCGTGTT

SMC 78      ------------------------------------------------------------
AB088224    CCTCGCCGTACGCGGTCAGCATGTTGCGCACCGACACCCACATGGACATCTGCCAGGACT

SMC 78      ------------------------------------------------------------
AB088224    CGTCCACCTCCCCGCCGGCCCAGGCCGGCCAGTGCCGGTAGGCGTCCTTCGAGGCGCGCG

SMC 78      ------------------------------------------------------------
AB088224    AGTCGGTCAGCAGCTGCTTGATCAGGTCGAGGTCCGTCACCGCCCACGCCGAGACACCCC

SMC 78      ------------------------------------------------------------
AB088224    CCGGCAGCACCACCCGTACGACCGGGCCGCGCTCCCGCAGCCGTGCCGCCTCCCCCGGCA

SMC 78      --------------------------------------------------
AB088224    GGTCACGGCCGGCCGAATCGATGACGTAGGGCGCTTCCTGACGAAGCAT
```

Figure 34. DNA sequence of putative macrolide P-450 from SMC 78 and its translated amino

acid sequence

```
         10        20        30        40        50        60
AGCCGGCTGCGCAAGCTGGTGGCGCCCAGTTTCACGCACCGCCGGACCGAGGCGATGCGG
 S  R  L  R  K  L  V  A  P  S  F  T  H  R  R  T  E  A  M  R

         70        80        90       100       110       120
CCCCGGGTGGAGACGATCACGGCGGAGCTGCTGGCTGCGCTGGAGGCCGCGGAGGGCGAG
 P  R  V  E  T  I  T  A  E  L  L  A  A  L  E  A  A  E  G  E

        130       140       150       160       170       180
GTCGTGGACGTCAAGGAGGGGCTCGCCCACCCGCTGCCGATGCGGATGATCTGCGAACTC
 V  V  D  V  K  E  G  L  A  H  P  L  P  M  R  M  I  C  E  L

        190       200       210       220       230       240
ATGGGTGTGCCGGACGCGTTGCGGGAGGACACCGGCCGCCTCATCGCGGCCATCATGGAC
 M  G  V  P  D  A  L  R  E  D  T  G  R  L  I  A  A  I  M  D

        250       260       270       280       290       300
ACCTCCGACCCGAGCCCGGAGCACGCGGCGTCCGTGCAGCGGCAGATCGGGACGGTGCTG
 T  S  D  P  S  P  E  H  A  A  S  V  Q  R  Q  I  G  T  V  L

        310       320       330       340       350       360
CCCGCGCTGATCGCCCACCGGGCCGCGCACCCCGGGGACGACATCACGACCGAGCTCATC
 P  A  L  I  A  H  R  A  A  H  P  G  D  D  I  T  T  E  L  I

        370       380       390       400       410       420
CGGGTCCGCGACGAGGACGGCGACCGGCTGAGCGACGAGGAGCTGCTCTACACGCTGCTG
 R  V  R  D  E  D  G  D  R  L  S  D  E  E  L  L  Y  T  L  L

        430       440       450       460       470       480
CTGGTGATCGGCGCCGGGTTCGAGACCACCGTGAATCACATCGGCAACGCGGTGGTCGCC
 L  V  I  G  A  G  F  E  T  T  V  N  H  I  G  N  A  V  V  A

        490       500       510       520       530       540
CTGCTCCGCCACCCCGAGCAGCTGGCGGCCGTGCGGTCCGGGGAGATCGGCTGGGACGCC
 L  L  R  H  P  E  Q  L  A  A  V  R  S  G  E  I  G  W  D  A

        550       560       570       580       590       600
GTCGTCGACGAGACGCTGCGCGCCCACCCGTCGATCGCGTCGCCGCCGCTGCGGTTCGCC
 V  V  D  E  T  L  R  A  H  P  S  I  A  S  P  P  L  R  F  A

        610       620       630       640       650       660
GTCACCGACCTCGCCGTCGGCGAGGTGACGATCCCGGCCGGCGACGCCATCACCACGACG
 V  T  D  L  A  V  G  E  V  T  I  P  A  G  D  A  I  T  T  T

        670       680       690       700       710       720
TACGCCGCGGCGGGGCTGGATCCCGAGCACTACGGGACGGACGCCGACCGCTTCGACGCC
 Y  A  A  A  G  L  D  P  E  H  Y  G  T  D  A  D  R  F  D  A

        730       740       750       760       770       780
ACGCGGGGCGCCGACGACCATCTGGCCTTCGGCATCGGCGTGCACCACTGCATCGGCG
 T  R  G  A  D  D  H  L  A  F  G  I  G  V  H  H  C  I  G
```

Figure 35. Pairwise alignment between translated amino acid sequences of SMC 78 and P-450 from *S. coelicolor* A3(2) (AL939114). Block indicated the oxygen-binding and heme-binding motif of P-450. Consensus of amino acids is represented by asterisk (*)

```
SMC 78      ------------------------------------------------------------
AL939114    MATQQPALVLDPTGADHHTEHRTLREGGPATWVDVLGVQAWSVSDPVLLKQLLTSSDVSK


SMC 78      ------------------------------------SRLRKLVAPSFTHRRTEAMRPRVET
AL939114    DARAHWPAFGEVVGTWPLALWVAVENMFTAYGPNHRKLRRLVAPAFSARRVDAMRPAVEA
                                                :**:****:*: **.:**** **:

SMC 78      ITAELLAAL-EAAEGEVVDVKEGLAHPLPMRMICELMGVPDALREDTGRLIAAIMDTSDP
AL939114    MVTGLVDRLAELPAGEPVDLRQELAYPLPIAVIGHLMGVPQDRRDGFRALVDGVFDTTLD
            :.: *: * * . ** **::: **:***: :* .*****: *:.  *: .::**:

SMC 78      SPEHAASVQRQIGTVLPALIAHRAAHPGDDITTELIRVRDE--DGDRLSDEELLYTLLLV
AL939114    QAEAQANTARLY-EVLDQLIAAKRATPGDDMTSLLIAARDDEGDGDRLSPEELRDTLLLM
            ..* *.. *   ** ***  : * ****:*: ** .**: ****** ***  ****:

            Oxygen-binding domain
SMC 78      IGAGFET TVNHIGNAVVALLRHPEQLAAVRSGEIGWDAVVDETLRAHPSIASPPLRFAVT
AL939114    ISAGYET TVNVIDQAVHTLLTRPDQLALVRKGEVTWADVVEETLRHEPAVKHLPLRYAVT
            * **:**  *** *.:** :** :*:*** **.**: *  **:**** .*:: ***:* **

                                                   Heme-binding domain
SMC 78      DLAVGE-VTIPAGDAITTTYAAAGLDPEHYGTDADRFDATRGADDHLAFGIGVHHCIG--
AL939114    DIALPDGRTIARGEPILASYAAANRHPDWH-EDADTFDATRTVKEHLAFGHGVHFCLGAP
            *:*: :  **. *:.* ::****. .*: : *** *****  ..:***** ***.*:*

SMC 78      ------------------------------------------------------------
AL939114    LARMEVTLALESLFGRFPDLRLADPAEELPPVPSLISNGHQRLPVLLHAG
```

Figure 36. Multiple sequence alignment of predicted secondary structures of putative P-450 of SMC 78, P-450 from *S. coelicolor* (AL939114) and P-450eryF from *S. erythraea* (M54983)

```
SMC 78       ------------------------------------------------------------
AL939114     ------------------------------------------------------------
M54983       ------------------------------------------------------------
M54983_known ------------------------------------------------------------

                              A helix          1 sheet        B helix
SMC 78       -------------------|------------|--EEEHH---EE----|-HHHHHH--
AL939114     -----------E------|-E--HH--HH-----|----EE-----HEEE-|-HHHHHHHH
M54983       ----------------E-|HHHHHHHHH-HHH|EEEEEEEE-EEEEEEEE|-HHHHHHHH
M54983_known -----------------|HHHHHHHHH-HHH|EEEEEEEE-EEEEEEEE|-HHHHHHHH

                                    B'helix        C helix
SMC 78       -------------------------------------|-------EE----|----
AL939114     -----H--HH---------E-----|HEHHH--H|HHH-----|--HHHHHH---|----
M54983       HH------------EEE-------|------E|------|------HH---H|EHHH
M54983_known H--------------|HHHHHHHH|------|HHHHHHHHHHHHH|HHHH

                D helix              2 sheet    E helix
SMC 78       ----------HHHHHHHHH|-HHHH-H-HHHHHH|--------HHHH----|-----H----
AL939114     ------HHHHHHHHHHHHH|-H----|----H-H|-------EEEE-----|-------
M54983       H--------HHHHHHHHH-|-------|-EEEEEHH|------HEEEEEHH-|-------
M54983_known HHHHHHHHHHHHHHHHHHH|-------|EEEEE|-HHHHHHHHHHHHHHHH|------HHH

                F helix           G helix             H helix   3 sheet
SMC 78       HHHEEE--|---|-----HHHHH-----HHHHHHH---|-|---HHHEEE|-|-----|--
AL939114     EEE-----|---|HHHHHHHHHHHHHHHHHHHHHH|-|---HHHHEE-|-|-----|--
M54983       ----EEEE|---|-HHH---H-HHHHHHHHHHH------|-|--HHHHHH-|-|EEEEE|---
M54983_known HHHHHHHH|---|HHHHHHHHHHHHHHHHHHHHHHHHH|-|-HHHHHHHHH|-|EEEEE|---

                              I helix         J helix     K helix
SMC 78       |--HHHHHHHHEEE-----EEEHH--HHHHHHHH-|H-H-HH|-------|-E-HHHHHHH--
AL939114     |-----HHHHHEEE------EEEHHHHHHH-----|--HHHH|----H|-HHHHHHHH--
M54983       |---HHHHHHHHEHHH-HHH-EEEEE--EEEE---|----HE-|------|----HHHHHH--
M54983_known |HHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH|HHHHHHH|HH---|HHHHHHHHHHHH

                4 sheet    5 sheet    6 sheet
SMC 78       ---|-----EE|-|----EE-|-|-EHHH-|-----------------------
AL939114     ---|-----E-|-|--HHHHHH|-|-HHHHHH-|-------HHHHH-------
M54983       ---|----HHHHHHHHH----E-|-|-EEEEH-|----------------
M54983_known ---|EEEE-EEE|---|EEE-EEEEE|-|-EEEEEEEE-|----------------

                              L helix      7 sheet
SMC 78       ---HHE----E---|-----------------|-|---|-----------------
AL939114     ---HHHH----EEE|--------HHHHHHHHH---|-|---|-----HHHHHHH-E----
M54983       ----------EE-|--------HHHHHHHHH---|-|---|----HHHHHHH-E----
M54983_known -------------|HH--HHHHHHHHHHHHHHHHH|-|EEE|----------------

                8 sheet
SMC 78       -|---|------------
AL939114     -|HHHH|------------
M54983       -|---|------------
M54983_known -|EEEE|------------
```
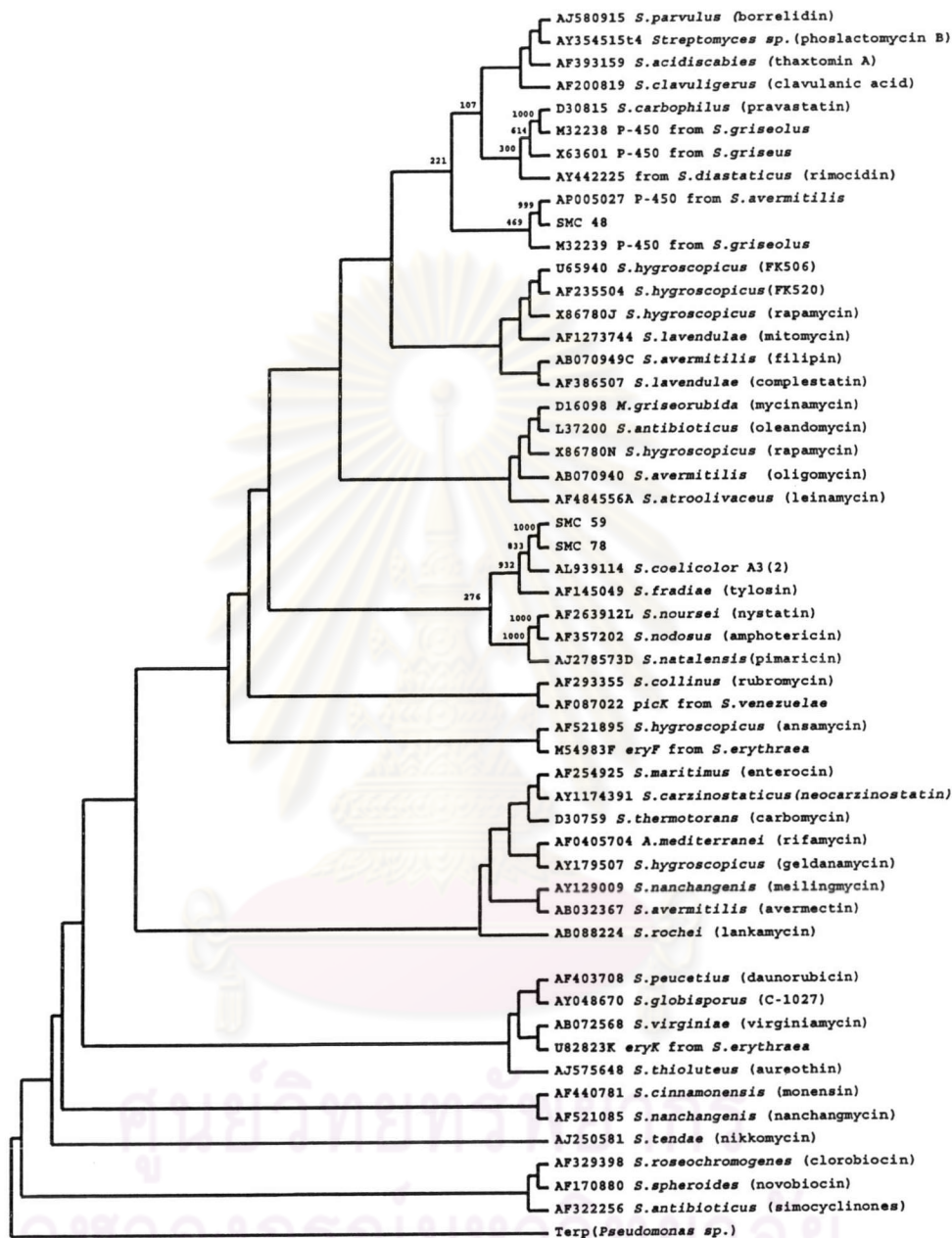
Figure 37.   The parsimony tree of nucleotide sequences of SMC 48, SMC 59 and SMC 78
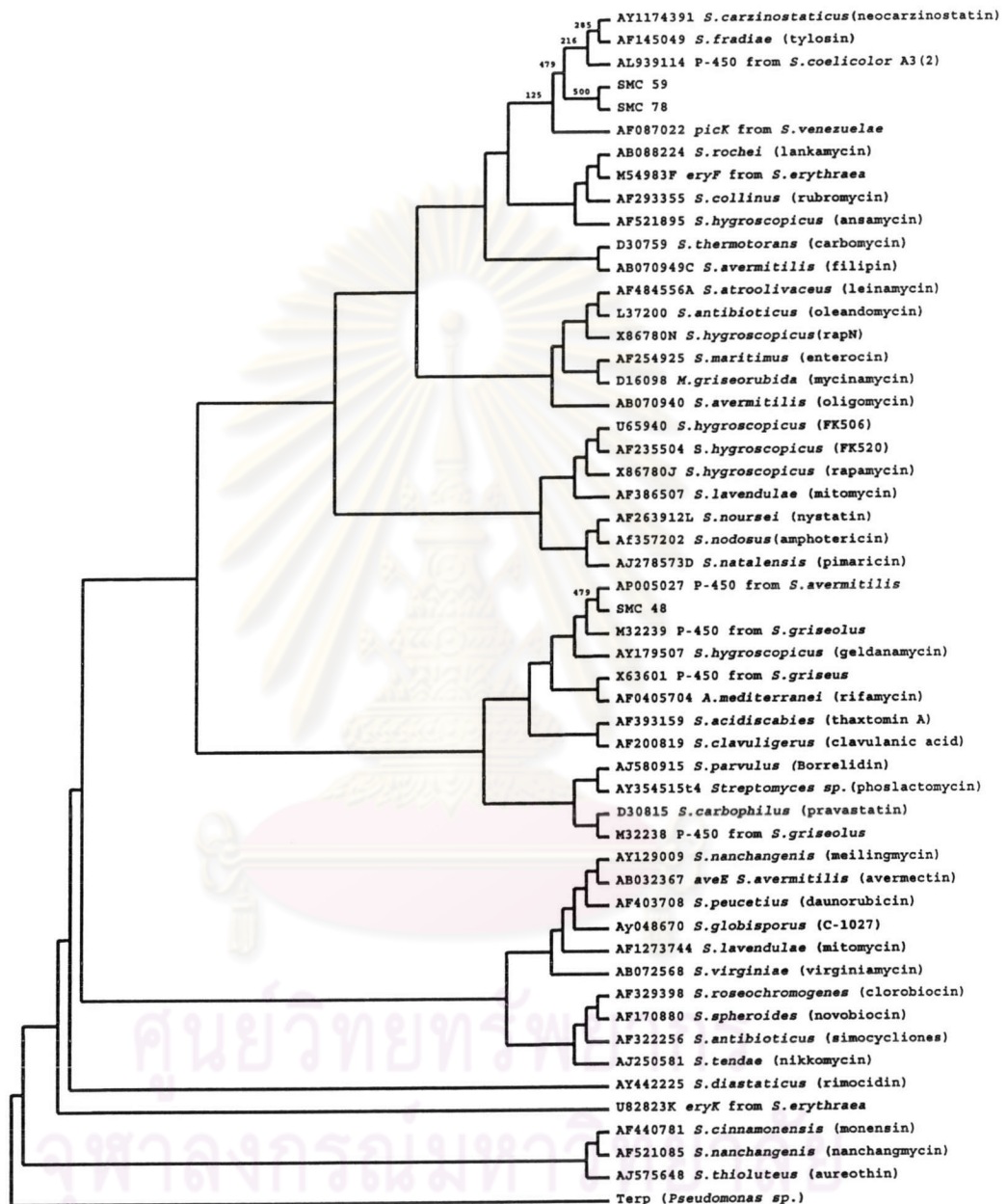(bootstrap 1000)

Figure 38. The parsimony tree of amino acid sequences of SMC 48, SMC 59 and SMC 78