JOURNAL OF BACTERIOLOGY, Dec. 1993, p. 8043–8048
0021-9193/93/8043-06$02.00/0

Vol. 175, No. 24

A Hydroxylase-Like Gene Product Contributes to Synthesis of a Polyketide Spore Pigment in *Streptomyces halstedii*

GLORIA BLANCO,[1] ANA PEREDA,[1]† PAUL BRIAN,[2]‡ CARMEN MÉNDEZ,[1] KEITH F. CHATER,[2] AND JOSÉ A. SALAS[1]*

Departamento de Biología Funcional, Universidad de Oviedo, 33006 Oviedo, Spain,[1] and John Innes Institute, Norwich NR4 7UH, United Kingdom[2]

Received 30 July 1993/Accepted 14 October 1993

A gene, *schC*, adjacent to the *sch* gene cluster encoding the biosynthesis of a polyketide spore pigment in *Streptomyces halstedii* was sequenced. Its deduced product resembled flavin adenine nucleotide-containing hydroxylases involved in the biosynthesis of polycyclic aromatic polyketide antibiotics and in catabolic pathways of aromatic compounds. When *schC* was disrupted, the normally green spores of *S. halstedii* became lilac. An *schC*-like gene was located in an equivalent position next to a large gene cluster (*whiE*) known to determine spore pigment in *Streptomyces coelicolor* A3(2).

Spores of *Streptomyces* spp. are often pigmented. Although these spore pigments are chemically uncharacterized, the sequencing of their genetic determinants in two species leaves little doubt that in these cases they are polyketides. Thus, the *whiE* cluster (4) from *Streptomyces coelicolor* and the *sch* cluster (1, 2) from *Streptomyces halstedii* both contain a core set of homologs of four genes characteristic of type II synthases for polycyclic aromatic polyketides in streptomycetes (11). Immediately upstream of the characterized set of six consecutive *sch* open reading frames (ORFs), the DNA sequence revealed part of a divergent ORF, whose deduced amino acid sequence resembled sequences of certain aromatic hydroxylases (1). Here we present the complete sequence of this ORF (*schC*) and show that its disruption affects spore color. An *schC* homolog is also demonstrated in *S. coelicolor* A3 (2).

Isolation and sequencing of the *schC* region from *S. halstedii*. A genomic library of *S. halstedii* NRRL 2381 chromosomal DNA in cosmid *cos*4 (10) was probed with the previously sequenced (1) *sch* 5.2-kb *Bam*HI (sites 7 to 15 in Fig. 3A) fragment by in situ colony hybridization. From one of these clones, all of which contained the 5.2-kb *Bam*HI fragment, we isolated an overlapping 4.0-kb *Pvu*II fragment (sites 1 to 7 in Fig. 3A) and sequenced 2,893 bp of it (Fig. 1). The sequence was analyzed for potential coding regions, using the CODON-PREFERENCE program (6) and looking for the characteristic codon usage and third-position bias of *Streptomyces* genes (28). This analysis showed that the whole of *schC* had been sequenced, together with the 5′ end of a further ORF (ORFD) beyond it. *schC* would encode a polypeptide of 555 amino acids with an estimated M_r of 59,551. The truncated part of the adjacent ORFD would encode 213 amino acids from the N terminus of a polypeptide (data base searching revealed no obvious homologs of this polypeptide). The intergenic region (170 bp) between *schC* and ORFD has a very high GC content and contains several inverted repeat sequences potentially capable of forming stem-loop-like secondary structures if transcribed as RNA.

A sequence similar to *schC* is located in an equivalent position next to the *whiE* spore pigment biosynthetic gene cluster of *S. coelicolor* A3(2). The spore pigment biosynthetic gene clusters of *S. halstedii* (*sch*) and *S. coelicolor* A3(2) (*whiE*) have extensive homology (1). The previously determined *whiE* sequence did not extend far enough to reveal the presence or absence of an *schC* homolog, though the DNA cloned in that earlier work in the primary clone pIJ2156 included about 900 bp of DNA to the left of the sequenced region (4). Therefore, we have now sequenced most of this region. This analysis revealed part of an ORF (*whiE* ORFVIII) that would be transcribed divergently from the major *whiE* cluster and whose derived gene product would be 82% identical to that of *schC*, with two insertions (of six amino acids and one amino acid).

Comparison of the deduced *schC* and *whiE* ORFVIII gene products with protein sequences in data bases. When compared with sequences in protein data bases, the *schC* and *whiE* ORFVIII gene products showed significant similarity to two groups of hydroxylase enzymes (Fig. 2): those catalyzing the introduction of hydroxyl groups into polyketide antibiotics such as tetracenomycin C (5), oxytetracycline (13), and daunorubicin (8) and those involved in catabolism of aromatic compounds in different organisms (14, 17, 19, 20, 25). There was no similarity to other classes of hydroxylases such as the P-450 cytochromes involved in hydroxylating the erythromycin macrolactone ring (9, 23, 24) (the O_2 binding site and the heme ligand pocket characteristic of these monooxygenases are absent [16, 18, 21]) or to any of the *actVA* gene products involved in hydroxylations of the ring structures in actinorhodin biosynthesis (3). Instead, and like hydroxylases that modify the oxytetracycline and tetracenomycin precursors, there are well-conserved amino acid sequences around two motifs which are common to a number of flavin adenine dinucleotide (FAD)- and NADPH-dependent enzymes (Fig. 2) (7, 26, 27). The first one is the so-called βαβ fold, which is involved in binding of the ADP moiety of FAD (27) and is located near the N terminus of the enzyme (amino acids 10 to 54 in *S. halstedii* and amino acids 16 to 60 in *S. coelicolor*). In the second motif, which may be involved in FAD binding, the aspartic residue (amino acid 313 in *schC*) could be particularly

* Corresponding author. Electronic mail address: JASF@DWARF1.QUIMICA.UNIOVI.ES.
† Present address: Abbott Laboratories, Corporate Molecular Biology, North Chicago, IL 60064.
‡ Present address: Department of Microbiology, Michigan State University, East Lansing, MI 48824.

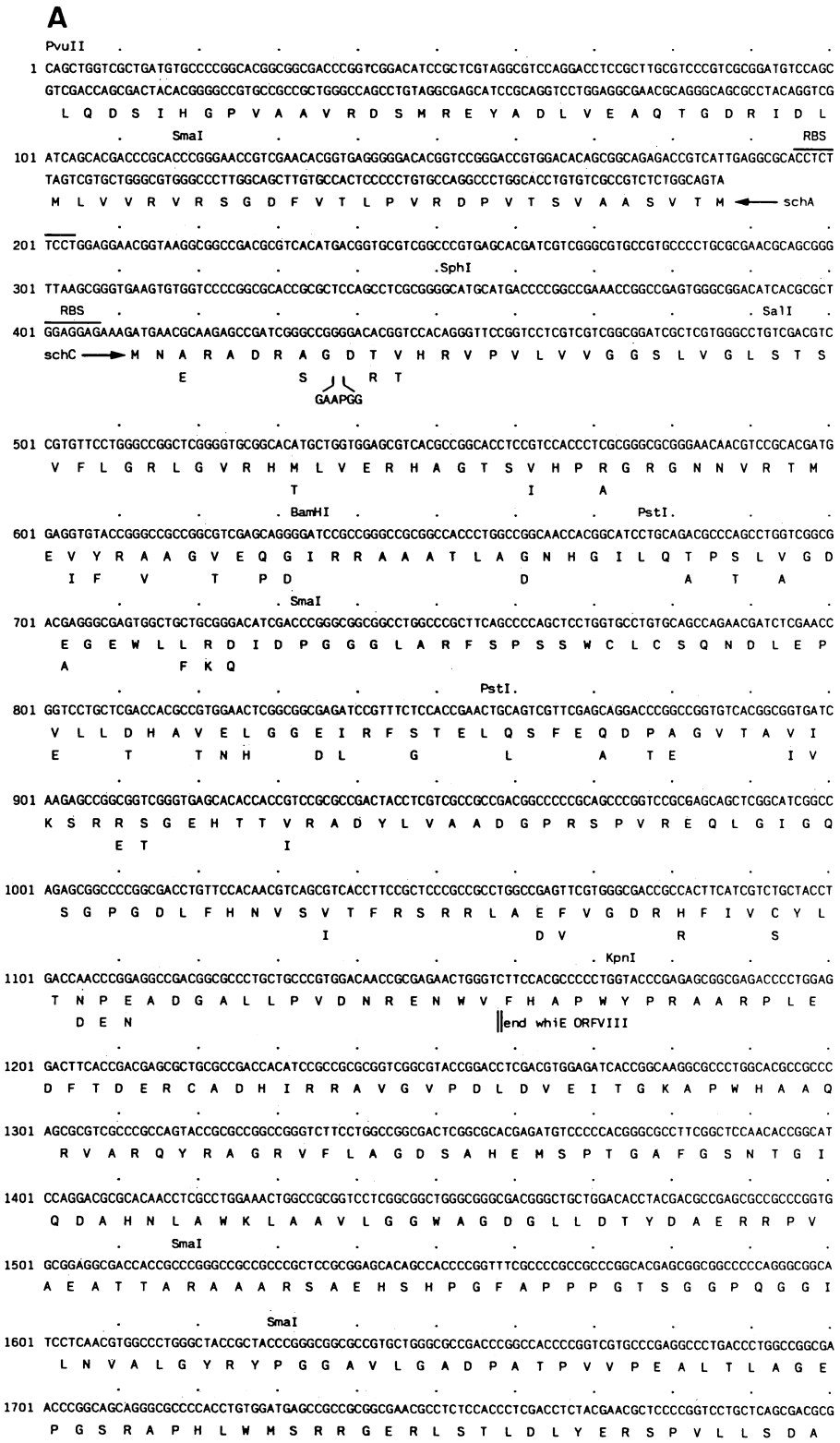

FIG. 1. Nucleotide sequence of a 2.8-kb *Pvu*II-*Pst*I fragment from the *sch* cluster (A) and of *whiE* ORFVIII from the *whiE* cluster (B). The nucleotide sequences of the fragments are shown with the deduced amino acid sequences of the different ORFs in the single-letter code. Potential ribosomal binding sites are indicated by RBS. Convergent arrows indicate stem-loop-like structures. Amino acid differences between the *schC* and *whiE* ORFVIII gene products are indicated under the relevant residues for the *schC* gene product. The *whiE* ORFVIII sequence determined does not extend up to the *Sph*I site used for subcloning (only a few bases were undetermined). The last 107 nucleotides (residues 816 to 922) of the *whiE* ORFVIII sequence were determined on one DNA strand only.

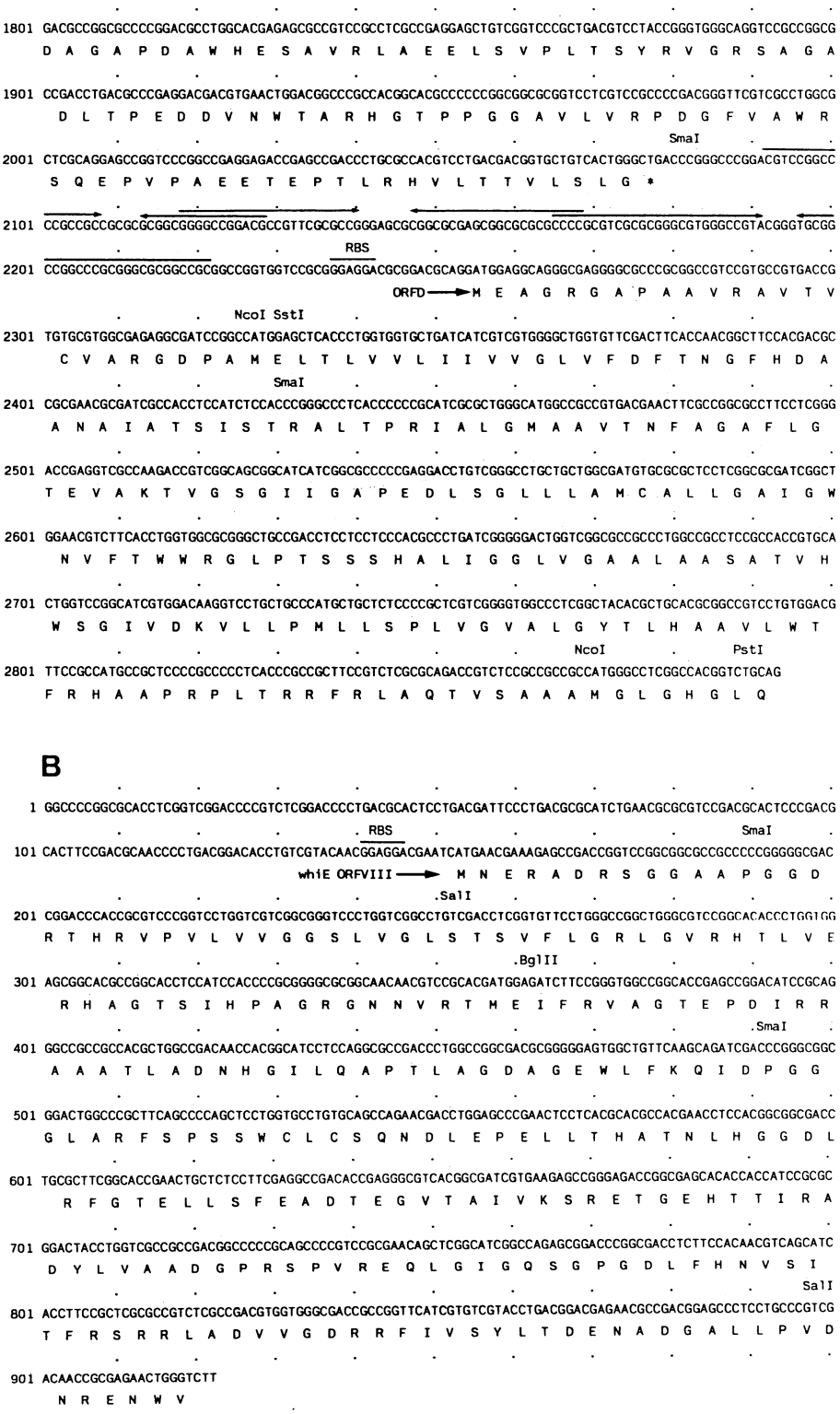

FIG. 1—*Continued.*

important in binding the ribityl chain of the flavin moiety of FAD (7, 22). These considerations strongly suggest that *schC* might function as an aromatic hydroxylase and, since the *sch* cluster resembles gene sets making polycyclic aromatic compounds, that *schC* is involved in spore pigment biosynthesis.

Insertional inactivation of *schC*. To determine whether *schC* was involved in spore pigment biosynthesis, we carried out insertional inactivation. A 540-bp *Kpn*I-*Bam*HI fragment (sites 4 to 7 in Fig. 3A) internal to *schC* was subcloned into the thermosensitive vector pGM160 (15), resulting in pUO618

```
                ******** ***** **      * ***  * **   *   **
PHEA   31  TNVVETEVLTVGSGPAGSSAAMFLS.....TQGISNIMITKYRWTANTPR   75
TFDB    5  ...IETDVLVVGTGPAGASAGALLA.....RYGVRTMLINKYNWTAPTPR   46
DNRF    3  LTKPDVDVLVVGGGLGGLSTALFLA.....RRGARVLLVERHAGTSVHPR   47
SCHC   10  DTVHRVPVLVVGGSLVGLSTSVFLG.....RLGVRHMLVERHAGTSVHPR   54
WHIE   16  DRTHRVPVLVVGGSLVGLSTSVFLG.....RLGVRHTLVERHAGTSIHPA   60
TCMG   15  LSTEEVPVLIVGGGLTGLSAALFLS.....QHGVSCRLVEKHRGTTVLTR   59
OTCC    1  ...MRYDVVIAGAGPTGLMLACELR.....LAGARTLVLERLAERVDFSK   39
PCLO   10  GQSADAAVLIVGGGTTGLIAANELL.....RRGVSCRMIDRLPVAHQTSK   54
PHYA    4  YSESYCDVLIVGAGPAGLMAARVLSEYVRQKPDLKVRIIDKRSTKVYNGQ   53
HBH     1  ...MKTKVAIIGAGPSGLLLGQLL.....HKAGIDNVILERQTPDYVLGR   42

        * ********** * * * ** * ******** ** **** ** * *   *****  **
PHEA 320 QKGRVCCAGDAIHKHPPSHGLGSNTSIQDSYNLCWKLACVLKGQAGPELLETYSTERA 377
TFDB 302 QQGRVFCAGDAVHRHPPTNGLGSNTSIQDSFNLAWKIAMVLNGTADESLLDTYTIERA 359
SCHC 304 RAGRVFLAGDSAHEMSPTGAFGSNTGIQDAHNLAWKLAAVLGGWAGDGLLDTYDAERR 361
TCMG 327 RSGRVFLAGDAAHVHPPAGAFGANGGIQDAHNLAWKLAAVLKGTAGDALLDTYEQERL 384
OTCC 279 RDGRVLLAGDACHIHLPAGGQGLNLGFQDAVNLGWKLGATIAGTAPPELLDTYEAERR 336
PCLO 289 RKGNVFLAGDAAHCHSPSGGSGMNVGMQDAFNLGWKIAMVERGEAKPDLLDTYHTERT 346
PHYA 349 KDERVFIAGDACHTHSPKAGQGMNTSMMDTYNLGWKLGLVLTGRAKRDILKTYEEERH 406
HBH  277 QHGRLFLAGDAAHIVPPTGAKGLNLAASDVSTLYRLLLKAYRE.GRGELLERYS.... 329
```

FIG. 2. Comparison of the deduced amino acid sequence of the *schC* gene with sequences of amino acids encoded by hydroxylase genes. The alignment shows the similarity among the *schC* product and different hydroxylases around the two motifs present in many FAD- and NADPH-dependent enzymes (7, 26, 27). The asterisks indicate amino acids which are present in at least half of the hydroxylases compared. The hydroxylase genes compared and the percentages of similar/identical amino acids against the *schC* product were as follows: PHEA, phenol monooxygenase from *Pseudomonas* sp. strain EST1001 (17), 53.61/32.70; TFDB, 2,4-dichlorophenol hydroxylase from *Alcaligenes eutrophus* (20), 54.25/32.22; DNRF, hydroxylase involved in daunorubicin biosynthesis by *S. peucetius* (8), 72.15/53.16 (only for the 79 amino acids available from the data base); SCHC, *S. halstedii* hydroxylase (this work); WHIE, *whiE* ORFVIII from *S. coelicolor* (this work); TCMG, tetracenomycin hydroxylase from *S. glaucescens* (5), 62.59/43.51; OTCC, anhydrooxytetracycline monooxygenase from *S. rimosus* (13), 52.59/30.76; PCLO, pentachlorophenol 4-monooxygenase from *Flavobacterium* sp. (19), 50.79/26.87; PHYA, phenol hydroxylase from *Trichosporon cutaneum* (14), 49.90/24.76; HBH, 4-hydroxybenzoate hydroxylase from *Pseudomonas fluorescens* (25), 46.56/24.33.

(Fig. 3B). Integration of pUO618 into the *S. halstedii* chromosome was carried out as described previously (2). As a consequence of this integration, some mutant lilac colonies appeared, and phase-contrast microscopy indicated that they formed morphologically normal spores. The lilac phenotype

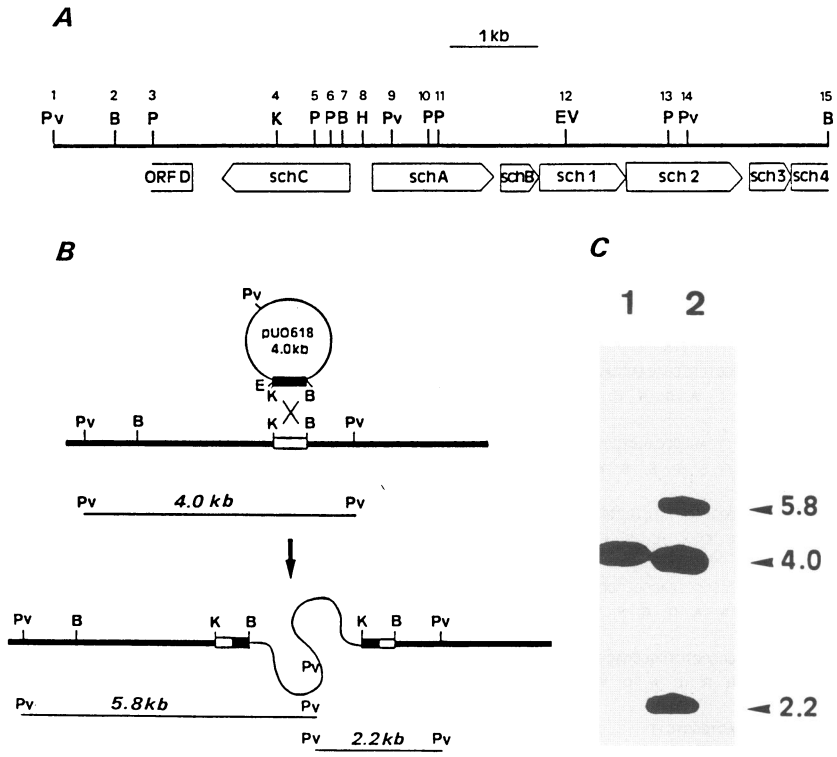

FIG. 3. (A) Restriction map of the *sch* region from *S. halstedii*. The restriction map of the previously cloned and sequenced 5.2-kb *Bam*HI fragment (1) (sites 7 to 15) and the adjacent DNA sequenced in this study (sites 3 to 7) are shown together with the different ORFs. *schA* and *schB*, unknown function; *schC*, hydroxylase (this work); ORFD, unknown function (this work); *sch1* and *sch2*, β-ketoacyl synthase; *sch3*, acyl carrier protein; *sch4*, cyclase. Abbreviations: B, *Bam*HI; EV, *Eco*RV; H, *Sph*I; K, *Kpn*I; Pv, *Pvu*II; P, *Pst*I. (B) Scheme representing the integration of pUO618 into the homologous region of the *S. halstedii* chromosomal DNA indicating the fragments generated after *Pvu*II digestion. Abbreviations: B, *Bam*HI; E, *Eco*RI; K, *Kpn*I; Pv, *Pvu*II. (C) Analysis of chromosomal DNA by Southern hybridization. *Pvu*II-digested chromosomal DNA of the wild-type *S. halstedii* strain (lane 1) and of mutant strain G8 (lane 2) were analyzed by Southern hybridization, using as a probe the 540-bp *Kpn*I-*Bam*HI fragment (sites 4 to 7 in Fig. 3A) cloned into M13mp19 (pUO6018). Sizes are indicated in kilobases.

was stable only at the nonpermissive temperature and in the presence of thiostrepton. Southern hybridization (Fig. 3C) confirmed the expected structure of this region of the chromosome in the lilac mutant (strain G8). The existence of a 4-kb *Pvu*II band, corresponding both to the size of the linearized pUO618 and to the wild-type chromosomal DNA band, can probably be explained by a dynamic equilibrium between free and integrated plasmid or because a tandem integration of pUO618 into the chromosome had occurred. The changed pigmentation resulting from *schC* disruption could be due either to inactivation of *schC* itself or to polar effects on the expression of sequences downstream of *schC*. In either case, there is no doubt that the *schC* transcription unit (and, by implication, the *whiE* ORFVIII transcription unit) is implicated in normal pigmentation. Since the characterized *schC* and *whiE* ORFVIII homologs all hydroxylate aromatic rings, and in view of the overall relatedness of the *schC* and *whiE* clusters to gene clusters encoding biosynthesis of polycyclic aromatic compounds, it seems very likely that the spore pigments of *S. halstedii* and *S. coelicolor* are themselves based on hydroxylated polycyclic aromatic structures.

The discovery of *whiE* ORFVIII extends the minimum length of the *whiE* cluster to 7.5 kb. *whiE* ORFVIII is absent from, or incomplete in, previously cloned segments of *whiE*, i.e., the brown, soluble pigment-encoding plasmid pARC1 (12) and the *whiE*-complementing pIJ2156 (4). In the case of pARC1, which causes pigment production only when expressed from an inserted promoter, it seems likely that synthesis of the brown pigment does not involve *whiE* ORFVIII, so this pigment is probably significantly different from the wild-type spore-associated pigment. Strains carrying pIJ2156 produce spores that are unusually dark and greenish, as if they are overproducing pigment (4), suggesting that the levels of hydroxylase expressed from the chromosomally located *whiE* ORFVIII are by no means limiting for synthesis of this dark green pigment.

Nucleotide sequence accession numbers. The sequences shown in Fig. 1 have been deposited in GenBank under accession numbers L05390 (*sch*) and X74213 (*whiE*).

We thank W. Wohlleben for providing plasmid pGM160 and S. J. Lucania, Squibb Institute, Princeton, N.J., for the kind gift of thiostrepton. We are also grateful to I. Hunter and C. R. Hutchinson for kindly providing unpublished sequence information.

Analysis of the sequence was made on a MicroVAX 3400 computer at the Scientific Computer Center of the University of Oviedo. This research was supported in part by grant BAP-0358-E from the European Community (to J.A.S.) and a John Innes Foundation studentship (to P.B.).

REFERENCES

1. **Blanco, G., P. Brian, A. Pereda, C. Méndez, J. A. Salas, and K. F. Chater.** 1993. Hybridization and DNA sequence analyses suggest an early evolutionary divergence of related biosynthetic gene sets encoding polyketide antibiotics and spore pigments in *Streptomyces* spp. Gene **130:**107–116.
2. **Blanco, G., A. Pereda, C. Méndez, and J. A. Salas.** 1992. Cloning and disruption of a DNA fragment of *Streptomyces halstedii* involved in the biosynthesis of a spore pigment. Gene **112:**59–65.
3. **Caballero, J. L., E. Martinez, F. Malpartida, and D. A. Hopwood.** 1991. Organisation and functions of the *actVA* region of the actinorhodin biosynthetic gene cluster of *Streptomyces coelicolor*. Mol. Gen. Genet. **230:**401–412.
4. **Davis, N. K., and K. F. Chater.** 1990. Spore colour in *Streptomyces coelicolor* A3(2) involves the developmentally regulated synthesis of a compound biosynthetically related to polyketide antibiotics. Mol. Microbiol. **4:**1679–1691.
5. **Decker, H., H. Motamedi, and C. R. Hutchinson.** 1992. Nucleotide sequences and heterologous expression of *tcmG* and *tcmP*, biosynthetic genes for tetracenomycin C in *Streptomyces glaucescens*. J. Bacteriol. **175:**3876–3886.
6. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
7. **Eggink, G., H. Engel, G. Vriend, P. Terpstra, and B. Witholt.** 1990. Rubredoxin reductase of *Pseudomonas oleovorans*: structural relationship to other flavoprotein oxidoreductases based on one NAD and two FAD fingerprints. J. Mol. Biol. **212:**135–142.
8. **Guilfoile, P. G., and C. R. Hutchinson.** 1991. A bacterial analog of the *mdr* gene of mammalian tumor cells is present in *Streptomyces peucetius*, the producer of daunorubicin and doxorubicin. Proc. Natl. Acad. Sci. USA **88:**8553–8557.
9. **Haydock, S. F., J. A. Dowson, N. Dhillon, G. A. Roberts, J. Cortés, and P. F. Leadlay.** 1991. Cloning and sequence analysis of the genes involved in erythromycin biosynthesis in *Saccharopolyspora erythraea*. Sequence similarities between *eryG* and a family of S-adenosylmethionine-dependent methyltransferases. Mol. Gen. Genet. **230:**120–128.
10. **Hohn, B., and J. Collins.** 1980. A small cosmid for efficient cloning of large DNA fragments. Gene **11:**291–298.
11. **Hopwood, D. A., and D. H. Sherman.** 1990. Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. Annu. Rev. Genet. **24:**37–66.
12. **Horinouchi, S., and T. Beppu.** 1985. Construction and application of a promoter-probe plasmid that allows chromogenic identification in *Streptomyces lividans*. J. Bacteriol. **162:**406–412.
13. **Hunter, I.** Personal communication.
14. **Kalin, M., H. Y. Neujahr, R. N. Weissmahr, T. Sejlitz, R. Jöhl, A. Fiechter, and J. Reiser.** 1992. Phenol hydroxylase from *Trichosporon cutaneum*: gene cloning, sequence analysis, and functional expression in *Escherichia coli*. J. Bacteriol. **174:**7112–7120.
15. **Muth, G., B. Nußbaumer, W. Wohlleben, and A. Puhler.** 1989. A vector system with temperature-sensitive replication for gene disruption and mutational cloning in streptomycetes. Mol. Gen. Genet. **219:**341–348.
16. **Nebert, D. W., and F. J. Gonzales.** 1987. P450 genes: structure, evolution, and regulation. Annu. Rev. Biochem. **56:**945–993.
17. **Nurk, A., L. Kasak, and M. Kivisaar.** 1991. Sequence of the gene (*pheA*) encoding phenol monooxygenase from *Pseudomonas* sp. EST1001: expression in *Escherichia coli* and *Pseudomonas putida*. Gene **102:**13–18.
18. **Omer, C. A., R. Lenstra, P. J. Litle, C. Dean, J. M. Tepperman, K. J. Leto, J. A. Romesser, and D. P. O'Keefe.** 1990. Genes for two herbicide-inducible cytochromes P-450 from *Streptomyces griseolus*. J. Bacteriol. **172:**3335–3345.
19. **Orsen, C. S., C. C. Lange, L. Xun, T. C. Zahrt, and B. J. Schneider.** Unpublished results.
20. **Perkins, E. J., M. P. Gordon, O. Caceres, and P. F. Lurquin.** 1990. Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. J. Bacteriol. **172:**2351–2359.
21. **Poulos, T. L., B. C. Finzel, and A. J. Howard.** 1987. High-resolution crystal structure of cytochrome $P450_{cam}$. J. Mol. Biol. **195:**687–700.
22. **Russel, M., and P. Model.** 1988. Sequence of thioredoxin reductase from *Escherichia coli*: relationship to other flavoprotein disulfide oxidoreductases. J. Biol. Chem. **263:**9015–9019.
23. **Stassi, D., S. Donadio, M. J. Staver, and L. Katz.** 1993. Identification of a *Saccharopolyspora erythraea* gene required for the final hydroxylation step in erythromycin biosynthesis. J. Bacteriol. **175:**182–189.
24. **Weber, J. M., J. O. Leung, S. J. Swanson, K. B. Idler, and J. B. McAlpine.** 1991. An erythromycin derivative produced by targeted gene disruption in *Saccharopolyspora erythraea*. Science **252:**114–116.

25. **Weijer, W. J., J. Hofsteenge, J. M. Vereijken, P. A. Jekel, and J. J. Beintema.** 1982. Primary structure of *p*-hydroxybenzoate hydroxylase from *Pseudomonas fluorescens*. Biochim. Biophys. Acta **704:** 385–388.
26. **Wierenga, R. K., P. Terpstra, and W. G. J. Hol.** 1985. Interaction of pyrophosphate moieties with α-helices in dinucleotide binding proteins. J. Mol. Biol. **24:**1346–1357.
27. **Wierenga, R. K., P. Terpstra, and W. G. J. Hol.** 1986. Prediction of the occurrence of the ADP-binding β-α-β fold in proteins using an amino acid fingerprint. J. Mol. Biol. **187:**101–107.
28. **Wright, F., and M. J. Bibb.** 1992. Codon usage in the G+C-rich *Streptomyces* genome. Gene **113:**55–65.