Mol Gen Genet (1992) 232:351–358

MGG

The *Streptomyces coelicolor whiB* gene encodes a small transcription factor-like protein dispensable for growth but essential for sporulation

N.K. Davis* and K.F. Chater

John Innes Institute, John Innes Centre for Plant Science Research, Colney Lane, Norwich NR4 7UH, UK

Received October 4, 1991 / Accepted November 20, 1991

Summary. A non-sporulating mutant *(whiB218)* of *Streptomyces coelicolor* A3(2), which proved to contain a deletion of more than 5 kb of DNA including *whiB*, was complemented by a small cloned DNA fragment, deduced from DNA sequencing* to encode a protein of only 87 amino acids. This protein would bear some similarities to transcription factors, including an acidic, somewhat amphipathic α-helical region predicted near its N-terminus, and a basic α-helical region predicted at its C-terminus. A point mutation *(whiB70)* giving a phenotype indistinguishable from that of the *whiB218* deletion mutant would cause a leucine to proline change at the start of the latter region. The *whiB* homologue from the closely related species *S. lividans* differed at only one base from its *S. coelicolor* counterpart, and would specify an identical polypeptide.

Key words: Developmental mutants – Sporulation – *Streptomyces coelicolor* A3(2) – *Streptomyces lividans* 66 – Transcription factor

Introduction

In the differentiating mycelial prokaryote *Streptomyces coelicolor* A3(2), nine mutant classes *(whiA,B,C,D,E, F,G,H* and *I)* were recognised among a collection of 50 mutants that show defects in the conversion of white multinucleoidal aseptate aerial hyphae into chains of mature grey uninucleoidal spores, while retaining apparently normal vegetative growth and antibiotic production (Hopwood et al. 1970; Chater 1972; Chater and Merrick 1976). Six of these genes (*whiA,B,C,G,H* and *I*) are required for the initial subdivision of aerial hyphae into spore-sized compartments. Five of the mutant classes were multiply represented in the collection, each class having its own typical aerial hyphal morphology: *whiG*, straight; *whiH*, loosely coiled; *whiA*, *whiB*, tightly coiled; *whiI*, tightly coiled and somewhat fragmented (Chater 1972; McVittie 1974). These phenotypes (except for *whiI*) appear to coincide with simple blocks at particular morphological stages (Wildermuth and Hopwood 1970), a view supported by the morphology of representative double mutants: a *whiG* mutation was epistatic to *whiH,A,B* and *I* mutations; *whiH* to *whiA,B* and *I;* and *whiA* and *B* to *whiI* (Chater 1975).

The availability of cloning systems for *S. coelicolor* (Hopwood et al. 1985) makes it possible to isolate DNA that complements these *whi* mutants (Méndez and Chater 1987; Davis and Chater 1990). Analysis of such cloned DNA has so far indicated that the crucial decision to switch from continued extension of aerial hyphae to spore formation is controlled by the level of a specific RNA polymerase sigma factor specified by the *whiG* gene (Chater et al. 1989); and that grey spore pigment formation depends on the (presumably developmentally regulated) expression of a cluster of genes (the *whiE* cluster) related to those involved in polyketide antibiotic formation (Davis and Chater 1990; Hopwood and Sherman 1990). The *whiG* and *whiE* gene products may thus simplistically be considered to control the first and the last steps in sporulation in aerial hyphae.

As part of a fuller analysis of the regulatory network that determines the development of aerial hyphae into spore chains, we describe here the cloning and sequencing of *whiB* from *S. coelicolor* and from the closely related strain *S. lividans* 66, and the characterization of the only two known *whiB* mutant alleles, *whiB70* and *whiB218* (Chater 1972). In mutants carrying either allele, the aerial hyphae as viewed by phase contrast microscopy (Chater 1972; 1975) consist of a straight stem surmounted by a long, tightly coiled, non-fragmenting tip region. Thin sections of a *whiB70* mutant confirmed the absence of the specialised septa involved in normal sporulation (McVittie 1974). This phenotype, coupled with the results of the morphological epistasis tests sum-

* *Present address:* Glaxo Group Research, Genetics Department, Greenford Road, Greenford, Middlesex UB6 OHE, UK

Offprint requests to: K.F. Chater

* The accession number in the EMBL Data Library is X62287

marised above, led to the proposal that the *whiB* gene product might be directly involved as a structural or catalytic element in the development of sporulation septa (Chater 1975). The results described here, on the other hand, tend to suggest a regulatory role.

Materials and methods

Strains and media. *Streptomyces* strains used were *S. coelicolor* A3(2) and its derivatives C70 (*whiB70;* Chater 1972), C218 (*whiB218;* Chater 1972), J1501 (*hisA1, uraA1, strA1,* Pgl$^-$ SCP1$^-$, SCP2$^-$, Chater et al. 1982), M124 (*proA1, argA1, cysD18,* SCP1$^-$, SCP2$^-$; Hopwood et al. 1985), M145 (SCP1$^-$ SCP2$^-$; Hopwood et al. 1985) and J171 (*cysD18 strA1 whiB218,* SCP1$^+$ or SCP1^{NF}, SCP2$^+$; Chater 1972), as well as *S. lividans* 66. J171 derivatives lacking SCP2 (J171.B1, J171.B2) were identified by colony hybridisation after regeneration of J171 protoplasts. Minimal medium (MM), protoplast regeneration medium (R2YE), liquid yeast extract-malt extract medium (YEME) and supplements were as in Hopwood et al. (1985). Conditions for selection and maintenance of plasmids and for morphological studies were as in Davis and Chater (1990). *Escherichia coli* TG2 was maintained as in Maniatis et al. (1982) and transformed as in Hanahan (1983).

Plasmids and phages and their manipulation. The *Streptomyces* low copy-number, transmissible SCP2-based vector pIJ698 (Kieser and Melton 1988) was used to clone *whiB*. The *Streptomyces* Tn*4560*-delivery plasmid pUC1169 (Chung 1987) was kindly provided by Dr S.-T. Chung, The Upjohn Company, Kalamazoo, Mich. The high copy-number *Streptomyces* vector pIJ486 (Ward et al. 1986) was used in a subcloning experiment. The *E. coli* vectors were pUC19 (Yannisch-Perron et al. 1985) and M13mp10 (Messing 1983). Techniques for DNA isolation and manipulation were as in Hopwood et al. (1985) and Maniatis et al. (1982), except that pIJ698 derivatives were prepared as described for pIJ922 derivatives in Davis and Chater (1990).

Colony hybridisation. Colonies of strain J171 obtained after protoplast regeneration were patched on plates of R2YE, and then Whatman 541 filters were laid on top. After incubation for 48 h the filters were prepared for hybridisation essentially as described by Maas (1983) but with a lysozyme pretreatment step as for nitrocellulose filters (Hopwood et al. 1985). Incubation of the plates was continued and the desired patches (in this case, those failing to hybridise to an SCP2-specific ^{32}P-labelled probe) were picked.

Isolation of whiB clones. In separate experiments, DNA isolated from *S. coelicolor* A3(2) was digested with *Bcl*I, *Bgl*II or *Xho*I, and ligated to pIJ698 cleaved with *Bgl*II (in the case of the *Xho*I experiment, the single-stranded ends of the vector and insert molecules were partially filled to generate complementary 2 bp ends; Zabarovsky and Allikmets 1986). The ligation mixtures were used to transform protoplasts of the *whiB* mutants J171.B1 or J171.B2, with selection for thiostrepton resistance (Hopwood et al. 1985). Transformed colonies were replicated onto MM plus mannitol, lacking thiostrepton, to screen for grey colonies. In the *Bcl*I experiment, several clones were obtained, of which two (pIJ2157 and pIJ2158) were used in further studies.

Transposon mutagenesis. Mutagenesis with transposon Tn*4560*, which confers viomycin resistance (Chung 1987), was essentially as in Davis and Chater (1990). The non-transmissible Tn*4560* donor plasmid (pUC1169) was used to transform J1501 carrying pIJ2157 or pIJ2158, and M124 was used as the recipient in mobilisation of Tn*4560* from the pUC1169-containing transformant colonies during plate matings. Although pUC1169 is not transmissible by itself, plasmid preparations from pooled M124 exconjugants showed limited co-transfer (probably in about 1 mating in 50), so the exconjugant pool was restreaked on suitably supplemented R2YE lacking viomycin (to allow further loss of the unstable pUC1169) but containing hygromycin (200 μg/ml) (to maintain selection for the pIJ698 derivatives), before restreaking individual colonies to medium containing viomycin. The resulting cultures, now lacking pUC1169, but containing Tn*4560* inserted into pIJ2157 or pIJ2158, were replicated to J171.B1 on fully supplemented, antibiotic-free R2YE to allow plate mating. J171.B1 exconjugants that had received plasmids containing Tn*4560* were finally selected by replication of the plate matings to MM supplemented with mannitol, cystine and viomycin, to allow scoring of the white/grey colony phenotypes. Physical analysis of transposon inserts was done by Southern blotting of total DNA preparations after digestion with *Hin*dIII plus *Pvu*II (and in some cases with *Eco*RV plus *Xho*I), using the nick-translated entire 5.7 kb cloned *whiB* region as probe.

Nucleotide sequence determination and analysis. For the sequencing of *whiB*, a 3.5 kb fragment capable of complementing a *whiB* mutant was isolated from an *E. coli* subclone (the fragment also contained a small amount of pUC19 vector DNA). It was sequenced by the shotgun sequencing method of Bankier and Barrell (1983) as in Davis and Chater (1990). The sequence of both strands was determined, each base in each strand being sequenced on average 3 times. Primary compilation and analysis of sequence data used the program DBUTIL (Staden 1980). Open reading frames were determined by the bias in base composition at first and third codon positions in high GC DNA using the program FRAME (Bibb et al. 1984). Protein structure prediction used the programs PEPPLOT and PLOTSTRUCTURE from the Sequence Analysis software package (version 7) supplied by the University of Wisconsin Genetics Computer Group (Devereux et al. 1984). In order to find proteins homologous to WhiB, the protein sequence database OWL6, comprising 19231 proteins, was searched by PROSRCH, which implements an exhaustive inexact string-matching algorithm, on an ICL 64 × 64 Distributed Array Processor, using the 100 PAM similarity table of

Collins and Coulson (1987). (This seach was kindly done by Dr. A. Lyall.) In addition, the program TFASTA (based on the algorithm of Pearson and Lipmann 1988 and supplied by the University of Wisconsin Genetics Computer Group) was used to compare the WhiB amino acid sequence with the NBRF database (release 26) and with all possible translations of the EMBL database (release 25).

For the sequencing of the *whiB70* allele and of the *S. lividans whiB* homologue, the *whiB* open reading frame (ORF) was amplified from total DNA preparations using the polymerase chain reaction (PCR; Sambrook et al. 1989). The oligonucleotides used as primers corresponded to residues 294–310 and 582–598 in the sequence shown in Fig. 2 and carried non-homologous 5′ tails with restriction enzyme recognition sites. The amplified fragments were phosphorylated with T4 polynucleotide kinase and ligated with M13 mp10 DNA that had been cut with *Sma*I and dephosphorylated with calf intestine alkaline phosphatase. After transfection, plaques with a LacZ$^-$ phenotype were used as a source of template for DNA sequencing.

Results

Cloning and localisation of whiB

Among ca. 10 000 colonies obtained after transformation of the *whiB218* mutant J171.B1 with the *Bcl*I library, 6 were scored as grey, and 5 of these were clearly sporulating when examined microscopically. No grey colonies were obtained with the libraries generated with *Bgl*II (ca. 9500 colonies) or *Xho*I (ca. 4000 colonies) even though each library should have been almost completely representative. This was consistent with later information about the distribution of sites for *Bgl*II and *Xho*I near or in the *whiB* gene.

Plasmids isolated from the five positive clones all contained the same 5.7 kb insert (Fig. 1). One orientation was represented in pIJ2157, the other in pIJ2158. Since both orientations gave *whiB* complementation, and since the cloning site in the vector pIJ698 is flanked by transcription terminators (Kieser and Melton 1988), the cloned fragment presumably contained the *whiB* promoter. Both plasmids were subjected to mutagenesis with Tn*4560*, using the Tn*4560* mobilisation procedure previously employed for the analysis of cloned *whiE* DNA (Davis and Chater 1990). Of 127 potential transposition events studied, nearly 30% gave white colonies when the relevant plasmids were transferred into the *whiB* mutant J171.B1. Many of these also gave very poor sporulation in the *whiB*$^+$ strain M124 for unknown reasons, and were not studied further. Southern blot analysis of 79 total DNA samples from grey and from white colonies showed only 4 cases where Tn*4560* was inserted in the 5.7 kb insert without accompanying rearrangements (Fig. 1). Of these four, one, with Tn*4560* located near the left end, was from a white colony. The other three, with Tn*4560* further to the right, were from grey colonies. Thus, *whiB* was apparently contained within the leftmost 2.2 kb of the cloned DNA. (In addition to deletions and other rearrangements associated with many of the other transposition events, in several cases there appeared to be integration of the transposon delivery plasmid into pIJ2157 and pIJ2158.)

The ability of smaller subclones of the 5.7 kb cloned fragment to complement the *whiB* mutant was also tested (Fig. 1). The leftmost 3.5 kb (up to *Sno*I site 9) gave complementation, but the rightmost 4.2 kb (from the *Sph*I site 6) did not. The plasmids used in these tests (based on pIJ698) were structurally unstable when introduced into strain J171.B1. This was probably due to pUC19 sequences present as a result of the particular subcloning strategies used.

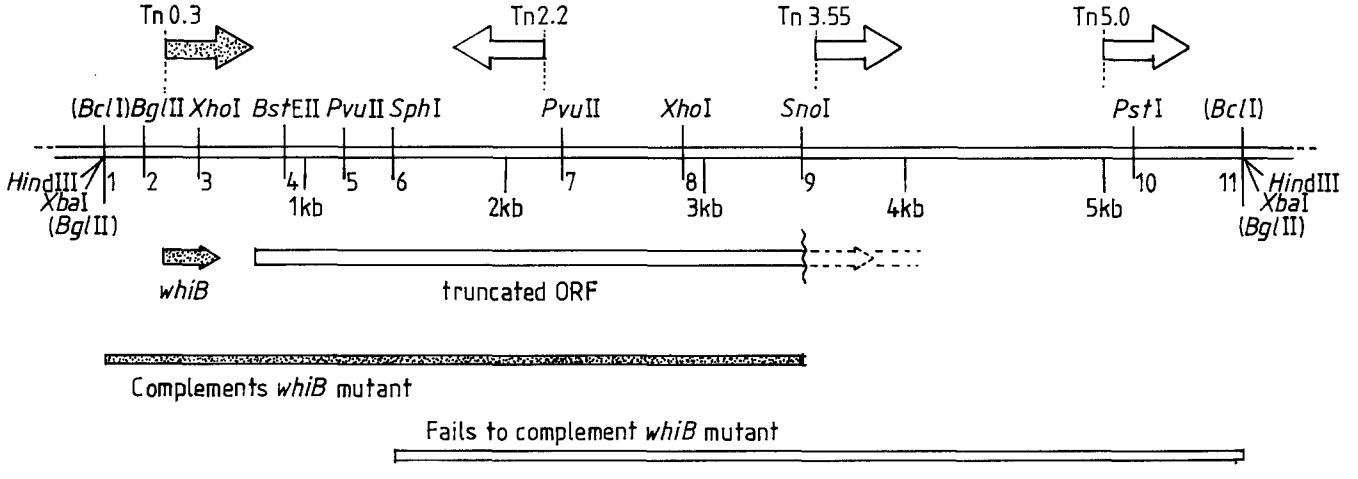

Fig. 1. Restriction map of the cloned *whiB* region of *Streptomyces coelicolor* A(3)2. The *arrows* below the map indicate the position and direction of open reading frames (ORFs) deduced from sequencing the interval between restriction sites 1 and 9. The *whiB* ORF is *stippled*. The *arrows* above the map show the positions and orientations (in relation to the map of Chung 1987) of Tn*4560* insertions; the insertion marked by a *stippled arrow* inactivated *whiB*-complementing activity. The *bars* at the bottom of the figure indicate subcloned fragments: the ability to complement *whiB* mutants is shown by *stippling*. The whole map is shown flanked by restriction sites present in the vector pIJ698

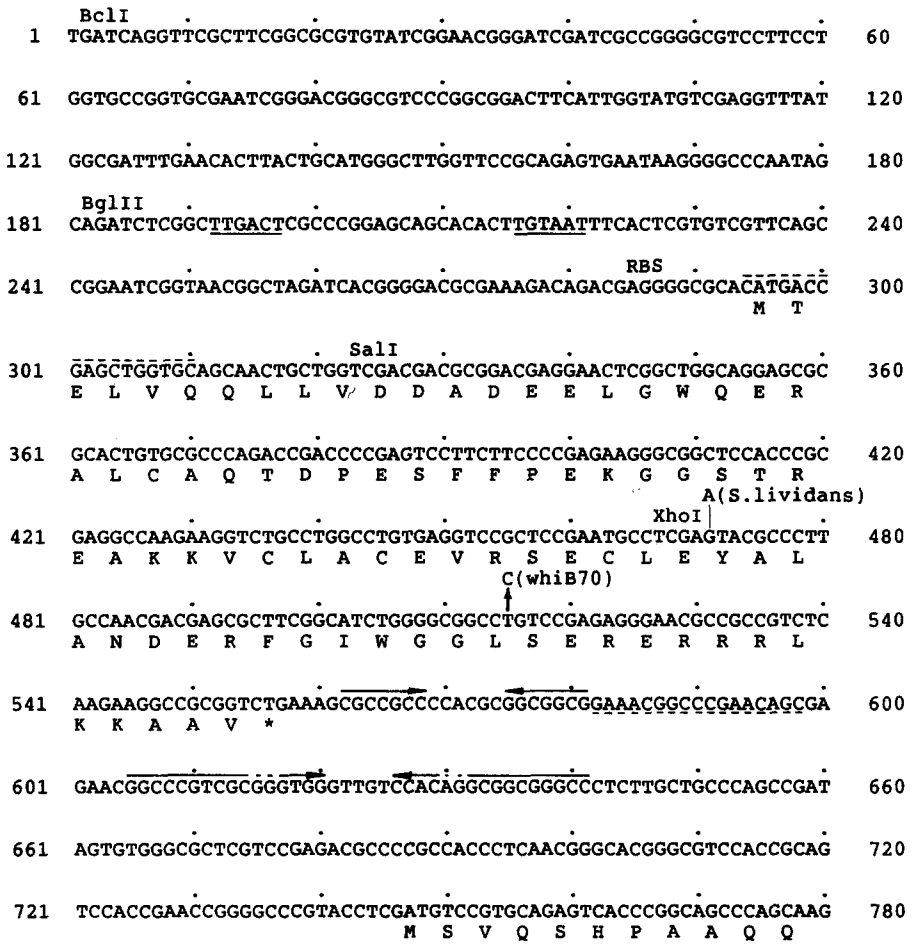

Fig. 2. Nucleotide sequence of *whiB* and flanking DNA. Selected restriction sites are indicated. *Underlined* sequences at nucleotides 192–7 and 216–221 resemble the −35 and −10 regions of vegetative promoters (Hopwood et al. 1986). RBS indicates a potential ribosome binding site. The deduced *whiB* gene product is given in the single-letter code for amino acids. *Dashed lines* indicate the positions of primers used for amplification by the polymerase chain reaction of the homologous region of the C70 and *S. lividans* genomes. The positions of a G or A difference between the *S. coelicolor* and *S. lividans* sequences, and of a T to C difference between the wild-type and *whiB70* sequences, are indicated. Inverted repeat sequences that might be involved in transcription termination are shown by *horizontal arrows*. A further 2708 bp of sequence was determined downstream of that shown here. It is deposited, together with the sequence shown, in the EMBL Data Library, accession no X62287

Sequence of whiB

The leftmost 3488 bp of the cloned DNA (up to *Sno*I site 9) was sequenced. The high overall GC content of the sequence (73.6%, typical for *Streptomyces* DNA) made it straightforward to use the FRAME analysis to detect meaningful open reading frames (Bibb et al. 1984). One complete and one partial ORF were found, both reading left to right on the restriction map. The righthand, incomplete, ORF encodes 914 amino acids up to the *Sno*I site (data not shown, but deposited in the EMBL database; accession no X62287). The lefthand, complete, ORF (Fig. 2) contains the position corresponding to the Tn*4560* insertion that inactivated *whiB* complementation. This ORF, containing 67.8% GC, was assumed to represent *whiB*. A site for *Xho*I within it provided a sufficient explanation for the failure to clone *whiB* using this enzyme. The 294 residues sequenced upstream of the *whiB* ORF did not exhibit coding character, and had a markedly lower GC content (58.8%). This region contained appropriately spaced hexanucleotide sequences resembling "vegetative" consensus promoter −35 and −10 regions (Hopwood et al. 1986) between nucleotides 192 and 221. Cleavage at a *Bgl*II site just upstream of the "-35-like" sequence, at nucleotides 182–187, may conceivably separate the ORF from *cis*-acting sequences essential for expression, and thereby account for failure to clone *whiB* using *Bgl*II. Alternatively, the finding from Southern blot analysis (not shown) that *whiB* is located on a *Bgl*II fragment of ≥20 kb suggests that the fragment could have been too big for efficient cloning. Inverted repeats, potentially specifying transcription terminators, were located downstream of the *whiB* ORF (Fig. 2).

The deduced *whiB* product, WhiB, contains only 87 amino acid residues (overall M_r 9878), of which an unusually high proportion (35%) are charged. Overall, there is a net excess of acidic residues. The amino-terminus is strongly negatively charged, and the carboxy-terminus is strongly positively charged. Both these regions contain predicted α-helices (Fig. 3). The central region of the protein contains four cysteine residues (not arranged as in zinc finger motifs), three of which are in a third predicted α-helical region. Even very sensitive database searches failed to detect other known or deduced proteins similar to WhiB. The possible properties of WhiB are further discussed below.

The whiB *gene locus comprises a single ORF dispensable for vegetative and aerial growth, but necessary for sporulation*

Southern blotting experiments, using the 5.7 kb *Bcl*I fragment as probe, revealed a single hybridising *Bcl*I-

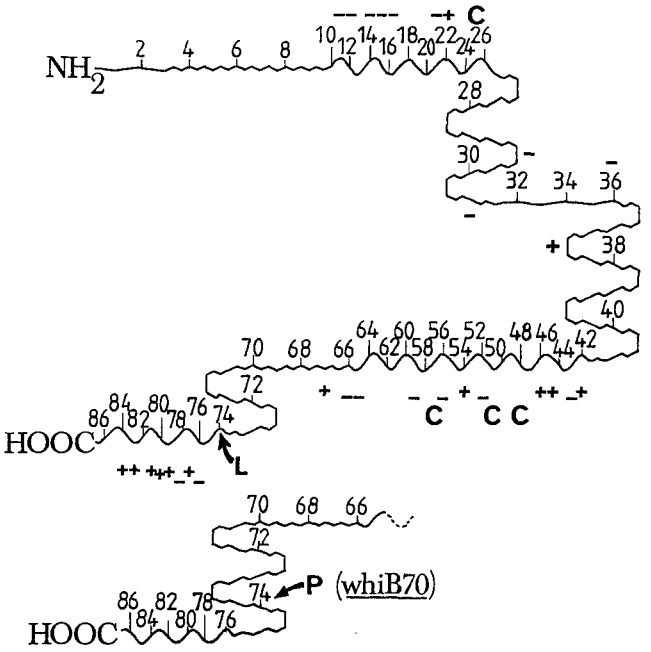

Fig. 3. Annotated structure predictions for the *whiB* gene product. The program used was PLOTSTRUCTURE (Devereux et al. 1984). The positions of charged residues (+, −) and cysteine residues (C) are noted. Below, the predicted carboxy-terminal region of the *whiB70* gene product. The positions of a leucine residue (L) and its proline (P) replacement in *whiB70* are also shown. Alpha-helices are shown with a sine wave (e.g. residues 10–26), beta-sheets with a sharp saw-tooth wave (e.g. residues 4–8), turns with 180 degree turns (e.g. residues 72–73) and coils with a dull saw-tooth wave (e.g. residues 32–36)

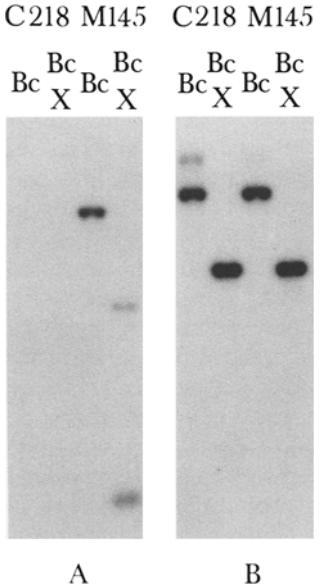

Fig. 4A, B. The *whiB218* allele is a deletion. DNA from C218 or M145 digested with *Bcl*I (Bc) or *Bcl*I plus *Xho*I (Bc X) was transferred to a nylon membrane (Hi-Bond, Amersham) and probed with (**A**) ^{32}P[ATP]-labelled *whiB*-specific probe (a 467 bp *Apa*I fragment, nucleotides 175–641 in Fig. 2; Random Priming Kit, Amersham), then stripped of probe as recommended by the manufacturer and re-probed (**B**) with *whiG*-specific probe (C. Granozzi, personal communication) to demonstrate the presence of the expected DNA in another region of the genome. Hybridization was at 65° C in 2 × SSC, and the final wash was at 65° C in 0.2 × SSC. The autoradiograph was provided by J. Soliveri

generated band of 5.7 kb in DNA of *S. coelicolor* M145 (wild type) and strain C70 (carrying the *whiB70* mutation). The cloned DNA was therefore not overtly rearranged, and *whiB70* is not a major deletion or insertion mutation. Chromosomal DNA from the *whiB218* mutant did not hybridise at all with the 5.7 kb *whiB*-containing probe, or with a smaller, *whiB*-specific probe (Fig. 4A), but it hybridised normally with a probe to a different sporulation-specific gene (*whiG;* Fig. 4B). Thus, *whiB218* is a deletion. This proves that *whiB* is not needed for vegetative or aerial growth, though it is essential for sporulation. To verify that only the putative *whiB* ORF, and not further coding sequences downstream of it, was necessary to restore sporulation to *whiB218,* pIJ2160, a high copy-number *Streptomyces* plasmid carrying little more than the small ORF (*Bcl*I site 1 to *Bst*EII site 4 in Fig. 1), was constructed using the vector pIJ486. Plasmid pIJ2160 restored sporulation to the *whiB218* mutant. This observation, coupled with the sequence information showing that the small ORF is relatively well separated from neighbouring ORFs, suggests that *whiB* consists of a single gene, and that the adjacent, partially sequenced ORF plays no essential role in normal growth and sporulation. No conspicuous morphological abnormality was observed when pIJ2160 was introduced into the wild type (in contrast to the hypersporulation induced by multiple copies of another sporulation regulatory gene, *whiG;* Chater et al. 1989).

Sequencing of a whiB *mutant allele and of the* S. lividans whiB *gene following PCR amplification*

Oligonucleotides were designed from the *S. coelicolor whiB* sequence to allow PCR-mediated amplification of a 304 bp segment carrying the *whiB* ORF from total DNA samples. In this way, the *whiB70* mutant allele was amplified. It was then inserted into M13 mpl0 and five separate clones were sequenced, revealing a single base-change (a TA to CG transition at residue 515) in all five, which would cause a Leu to Pro change at position 74 of the deduced gene product (Figs. 2, 3). (One of the clones contained an additional base change presumed to have occurred during PCR amplification.) Since a *whiB70* mutant is phenotypically the same as a *whiB218* deletion mutant, the change in the product of the *whiB70* allele evidently causes a dramatic loss of function.

In view of the generally high (ca. 98%) similarity of *S. lividans* DNA to that of *S. coelicolor* (Biró and Chater 1987; Bolotin and Biró 1989; Malpartida and Hopwood 1986; Henderson et al. 1989; Westpheling and Brawner 1989), it appeared likely that the primers used to amplify the *S. coelicolor whiB70* allele would also allow amplification of the *S. lividans whiB* homologue. This proved to be the case, and led to the cloning and sequencing of *whiB* from *S. lividans.* Within the amplified and cloned segment, only a single base difference from the *S. coelicolor whiB* sequence was discerned (residue 471; Fig. 2). This difference did not cause a change in the deduced gene product, but it did cause a *Xho*I restriction fragment length polymorphism between *S. lividans* and *S. coeli-*

color, since the difference falls within a sequence recognised in *S. coelicolor* DNA by *Xho*I. This polymorphism had already been observed in Southern blots of total *S. lividans* DNA. Just as for the *whiB70* clones, one in five of the *S. lividans* clones sequenced after PCR amplification and cloning contained a single base difference from the others, which was attributed to an error during amplification.

Discussion

This analysis has revealed that *whiB* consists of a single small gene, preceded by at least 130 bp (and probably more than 293 bp) of apparently non-coding DNA. The next gene downstream of *whiB* is large and dispensable for growth and differentiation, at least under normal laboratory conditions, and is separated from it by 189 bp of non-coding DNA. The *whiB* gene itself can be completely deleted without causing any obvious deficiencies in vegetative growth, aerial mycelium formation, or production of several antibiotics; but it is essential for an early stage in sporulation at aerial hyphal tips.

The phenotype of *whiB* mutants, and the results of epistasis studies, led Chater (1975) to propose that *whiB* was involved in sporulation septum formation or its regulation. The high proportion of charged and polar residues in WhiB does not suggest a close association with membranes, tending to rule out a structural role in the septum, and suggesting that it is a cytoplasmic protein. Although comparison to other gene products of known or unknown function revealed no obvious homologies, some features of WhiB are reminiscent of certain prokaryotic and eukaryotic transcription factors. The PLOTSTRUCTURE program predicts that WhiB contains three α-helical regions: a highly acidic region near the N-terminus, a cysteine-containing central region, and a highly basic region near the C-terminus. Acidic regions in transcriptional activators are often involved in protein:protein contacts with the transcription apparatus. Examples include repressors of lambdoid phages (Ptashne 1986), and eukaryotic proteins GCN4 (Hope and Struhl 1986), GAL4 (Ma and Ptashne 1987) and HAP1 (Pfeifer et al. 1989). Often, these acidic regions are α-helical (Giniger and Ptashne 1987). The presence of suitably spaced hydrophobic residues in α-helices may allow interactions within or between proteins, and these may be especially important in some cases for allowing strong DNA-binding, as in the coiled coil type of leucine zipper involved in Fos-Jun interaction (O'Shea et al. 1989) and the "Myc-like leucine repeats" in which hydrophobic (preferably leucine) residues occur at 7 amino acid intervals instead of intervals of 3–4 as in the Fos–Jun paradigm (Hu et al. 1990; Lüscher and Eisenman 1990). The acidic α-helix of WhiB has a somewhat Myc-like character.

The C-terminal 14 residue basic α-helix predicted for WhiB follows a glycine pair that would permit free turning. In the *whiB70* mutation, WhiB function is lost as a result of a Leu to Pro change next to this glycine pair: a proline residue might well alter the relationship of the putative C-terminal α-helix to features upstream of this hinge point (Fig. 3). Such a C-terminal basic helix might possibly be involved in sequence-specific DNA binding. Although it lacks important features of the classical helix-turn-helix motif widespread among prokaryotic DNA-binding proteins (Harrison and Aggarwal 1990), it nevertheless somewhat resembles the similarly located sequence-specific DNA-binding domain of Fis, a small *E. coli* protein containing many charged residues (Johnson et al. 1988) and capable of transcriptional activation (Nilsson et al. 1990; Ross et al. 1990): X-ray crystallography of Fis-DNA complexes has shown that the basic region aligned with WhiB in Fig. 5 is involved in contacts with DNA (Kostrewa et al. 1991). It is also interesting to note that the potentially α-helical region 2.4 of sigma factors, believed to contact the −10 region of promoters, is not part of a classical helix-turn-helix motif, and ends in a region that is often rich in basic residues even though sigma factors are acidic proteins overall (Helmann and Chamberlin 1988). In the more classical helix-turn-helix motif located at the C-terminal region 4.2 of sigma factors, the recognition helix that appears to contact the −35 region of promoters is also followed by a highly basic region (Helmann and Chamberlin 1988). Several kinds of transcriptional activators of eukaryotes appear to interact with specific DNA sequences through α-helices with basic C-terminal regions. Examples include Fos-Jun (Nakabeppu and Nathans 1989) and homeobox proteins (Hanes and Brent 1991).

The central region (residues 25–65) of the predicted WhiB molecule (Fig. 3) is marked by two prolines (residues 30 and 35) and a glycine pair (residues 43 and 44), potentially allowing major turns in the structure, and four cysteines (residues 25, 48, 51 and 57). There is no obvious similarity to typical zinc finger motifs of some eukaryotic transcriptional activators (Cys-X-X-Cys....Cys-X-X-Cys; Berg 1988), but the cysteine residues might be involved in metal ion coordination in some different configuration, perhaps leading either to intramolecular interactions as in the oestrogen receptor (Schwabe et al. 1990)

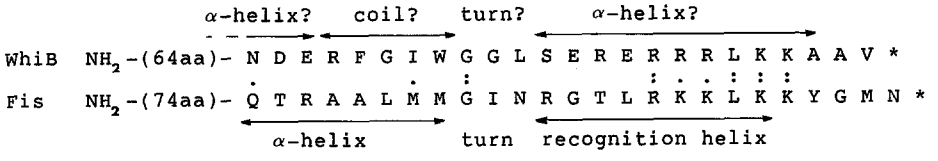

Fig. 5. Possible similarity between the C-terminal regions of the deduced *whiB* gene product and the Fis protein, a transcriptional activator from *Escherichia coli*. The Fis sequence is from Johnson et al. (1988) and the structural information about Fis is from Kostrewa et al. (1991). *Question marks* indicate computer-predicted structures (see Fig. 3)

or to intersubunit interactions as in the Tat protein of HIV (Frankel et al. 1988). The cysteines 48, 51 and 57 are in a region predicted to be α-helical, and would all be on the same face of the α-helix.

Taken together these considerations encourage the working hypothesis that WhiB may be directly involved both in DNA binding and in interactions with other proteins, most probably leading to transcriptional activation. Recent experiments by P. Brian (personal communication) have shown that promoters of the spore pigment gene cluster *whiE* are not transcribed in a *whiB* mutant, strengthening the view that WhiB may play a regulatory role. Perhaps this role is similar to that of the similarly sized (93 amino acids) "switch protein" encoded by the *spo*IIID gene in *Bacillus subtilis*. This protein contains potential amphipathic helices which could be involved in its interactions with RNA polymerase and DNA: interactions which direct some aspects of mother cell gene expression during sporulation (Kroos et al. 1989; Kunkel et al. 1989). Another small, highly charged, regulatory protein with multiple attributes of transcriptional activators has also been described in *B. subtilis*: this is the Sin protein, which is involved in both positive and negative regulation of various stationary phase phenotypes (Gaur et al. 1991). The paradigm for positive and negative control by small proteins, the 66 amino acids Cro protein of phage λ, is less highly charged (Ptashne 1986).

Acknowledgements. We thank Andrew Lyall of the University of Edinburgh for carrying out the PROSRCH analysis and Juan Soliveri for providing Fig. 4. We also thank Mervyn Bibb, Mark Buttner, David Hopwood, Tobias Kieser and Erik Vijgenboom for comments on the manuscript. The work was supported by grants from the Agricultural and Food Research Council and the John Innes Foundation.

References

Bankier AT, Barrell BG (1983) Shotgun DNA sequencing. In: Flavell RA (ed) Techniques in the Life Sciences. Elsevier Press, Ireland, pp 1–34

Berg JM (1988) Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. Proc Natl Acad Sci USA 85:99–102

Bibb MJ, Findlay PR, Johnson MW (1984) The relationship between base composition and codon usage in bacterial genes and its use in the simple and reliable identification of protein-coding sequences. Gene 30:157–166

Biró S, Chater KF (1987) Cloning of *Streptomyces lividans* and *Streptomyces griseus* genes for glycerol dissimilation. Gene 56:79–86

Bolotin A, Biró S (1989) Nucleotide sequence of the putative regulatory gene and major promoter region of the *Streptomyces griseus* glycerol operon. Gene 87:151–152

Chater KF (1972) A morphological and genetic mapping study of white colony mutants of *Streptomyces coelicolor*. J Gen Microbiol 72:9–28

Chater KF (1975) Construction and phenotypes of double sporulation deficient mutants in *Streptomyces coelicolor* A(3)2. J Gen Microbiol 87:312–325

Chater KF, Merrick MJ (1976) Approaches to the study of differentiation in *Streptomycescoelicolor* A3(2). In: Macdonald KD (ed) Second International Symposium on the Genetics of Industrial Microorganisms. Academic Press, London, pp 583–593

Chater KF, Bruton CJ, King AA, Suarez JE (1982) The expression of *Streptomyces* and *Escherichia coli* drug resistance determinants cloned into the *Streptomyces* phage øC31. Gene 19:21–32

Chater KF, Bruton CJ, Plaskitt KA, Buttner MJ, Méndez C, Helmann J (1989) The developmental fate of *Streptomyces coelicolor* hyphae depends crucially on a gene product homologous with the motility sigma factor of *Bacillus subtilis*. Cell 59:133–143

Chung ST (1987) Tn*4556* a 6.8 kb transposable element of *Streptomyces fradiae*. J Bacteriol 169:4436–4441

Collins JF, Coulson AFW (1987) Molecular sequence comparison and alignment. In: Bishop M, Rawlings C (eds) Nucleic Acid and Protein Sequence Analysis: a Practical Approach. IRL Press, Oxford, pp 323–358

Davis NK, Chater KF (1990) Spore colour in *Streptomyces coelicolor* A(3)2 involves the developmentally regulated synthesis of a compound biosynthetically related to polyketide antibiotics. Mol Microbiol 4:1679–1692

Devereux T, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12:387–395

Frankel AD, Bredt DS, Pabo CO (1988) Tat protein from human immunodeficiency virus forms a metal-linked dimer. Science 240:70–73

Gaur NK, Oppenheim J, Smith I (1991) The *Bacillus subtilis sin* gene, a regulator of alternate developmental processes, codes for a DNA-binding protein. J Bacteriol 173:678–686

Giniger E, Ptashne M (1987) Transcription in yeast activated by a putative amphipathic α helix linked to a DNA binding unit. Nature 330:670–672

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol 166:557–580

Hanes SD, Brent R (1991) A genetic model for interaction of the homeodomain recognition helix with DNA. Science 251:426–430

Harrison SC, Aggarwal AK (1990) DNA recognition by proteins with the helix-turn-helix motif. Annu Rev Biochem 59:933–969

Helmann JD, Chamberlin MJ (1988) Structure and function of bacterial sigma factors. Annu Rev Biochem 42:839–872

Henderson DJ, Lydiate DJ, Hopwood DA (1989) Structural and functional analysis of the minicircle, a transposable element of *Streptomyces coelicolor* A3(2). Mol Microbiol 3:1307–1318

Hope IA, Struhl K (1986) Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell 46:885–894

Hopwood DA, Sherman DH (1990) Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. Annu Rev Genet 24:37–66

Hopwood DA, Wildermuth H, Palmer HM (1970) Mutants of *Streptomyces coelicolor* defective in sporulation. J Gen Microbiol 61:397–408

Hopwood DA, Bibb MJ, Chater KF, Kieser T, Bruton CJ, Kieser HM, Lydiate DJ, Smith CP, Ward JM, Schrempf H (1985) Genetic Manipulation of *Streptomyces* – A Laboratory Manual. The John Innes Foundation, Norwich

Hopwood DA, Bibb MJ, Chater KF, Janssen GR, Malpartida F, Smith CP (1986) Regulation of gene expression in antibiotic-producing *Streptomyces*. In: Booth I, Higgins C (eds) Regulation of Gene Expression – 25 Years on. 39th Symposium of Society for General Microbiology. Cambridge University Press, Cambridge, pp 251–276

Hu Y-F, Lüscher B, Admon A, Mermod N, Tjian R (1990) Transcription factor AP-4 contains multiple dimerization domains that regulate dimer specificity. Genes Dev 4:1741–1752

Johnson RC, Ball CA, Pfeffer D, Simon MI (1988) Isolation of the gene encoding the Hin recombinational enhancer binding protein. Proc Natl Acad Sci USA 85:3484–3488

Kieser T, Melton R (1988) Plasmid pIJ699, a multi-copy positive-selection vector for *Streptomyces*. Gene 65:83–91

Kostrewa D, Granzin J, Koch C, Choe H-W, Raghunathan S, Wolf W, Labahn J, Kahmann R, Saenger W (1991) Three-dimensional structure of the *E. coli* DNA-binding protein FIS. Nature 349:178–180

Kroos L, Kunkel B, Losick R (1989) Switch protein alters specificity of RNA polymerase containing a compartment-specific sigma factor. Science 243:526–529

Kunkel B, Kroos L, Poth H, Youngman P, Losick R (1989) Temporal and spatial control of the mother-cell regulatory gene *spoIIID* of *Bacillus subtilis*. Genes Dev 3:1735–1744

Lüscher B, Eisenman RN (1990) New light on Myc and Myb. Part I. Myc. Genes Dev 4:2025–2035

Ma J, Ptashne M (1987) Deletion analysis of GAL4 defines two transcriptional activating segments. Cell 48:847–853

Maas R (1983) An improved colony hybridization method with significantly increased sensitivity for detection of single genes. Plasmid 10:296–298

Malpartida F, Hopwood DA (1986) Physical and genetic characterisation of the gene cluster for the antibiotic actinorhodin in *Streptomyces coelicolor* A(3)2. Mol Gen Genet 205:66–73

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

McVittie A (1974) Ultrastructural studies on sporulation in wild-type and white colony mutants of *Streptomyces coelicolor*. J Gen Microbiol 81:291–302

Méndez C, Chater KF (1987) Cloning of *whiG*, a gene critical for sporulation of *Streptomyces coelicolor* A(3)2. J Bacteriol 169:5715–5720

Messing J (1983) New M13 vectors for cloning. Methods Enzymol 101: 20–78

Nakabeppu Y, Nathans D (1989) The basic region of Fos mediates specific DNA binding. EMBO J 8:3833–3841

Nilsson L, Vanet A, Vijgenboom E, Bosch L (1990) The role of FIS in *trans* activation of stable RNA operons of *E. coli*. EMBO J 9:727–734

O'Shea EK, Rutkowski R, Kim PS (1989) Evidence that the leucine zipper is a coiled coil. Science 243:538–542

Pearson WR, Lipmann DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85:2444–2448

Pfeifer K, Kim K-S, Kogan S, Guarente L (1989) Functional dissection and sequence of yeast HAP1 activator. Cell 56:291–301

Ptashne M (1986) A Genetic Switch. Cell Press and Blackwell Scientific Publications, San Francisco

Ross W, Thompson JF, Newland JT, Gourse RL (1990) *E. coli* Fis protein activates ribosomal RNA transcription *in vitro* and *in vivo*. EMBO J 9:3733–3742

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: a Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Schwabe JWR, Neuhaus D, Rhodes D (1990) Solution structure of the DNA-binding domain of the oestrogen receptor. Nature 348:458–461

Staden R (1980) A new computer method for the storage and manipulation of DNA gel reading data. Nucleic Acids Res 8:3673–3694

Ward JM, Janssen GR, Kieser T, Bibb MJ, Buttner MJ, Bibb MJ (1986) Construction and characterisation of a series of multi-copy promoter-probe plasmid vectors for *Streptomyces* using the aminoglycoside phosphotransferase gene from Tn*5* as indicator. Mol Gen Genet 203:468–478

Wildermuth H, Hopwood DA (1970) Septation during sporulation in *Streptomyces coelicolor*. J Gen Microbiol 60:57–59

Westpheling J, Brawner MW (1989) Two transcribing activities are involved in expression of the *Streptomyces* galactose operon. J Bacteriol 171:1355–1361

Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 vectors and host strains: nucleotide sequences of the M13 mp18 and pUC19 vectors. Gene 33:103–119

Zabarovsky ER, Allikmets RK (1986) An improved technique for the efficient construction of gene libraries by partial filling-in of cohesive ends. Gene 42:119–123

Communicated by H. Hennecke