Characterization of the gene for factor C, an extracellular signal protein involved in morphological differentiation of *Streptomyces griseus*

Zsuzsa Birkó,¹ Andrea Sümegi,¹ Andrea Vinnai,¹ Gilles van Wezel,³ Ferenc Szeszák,¹ Sándor Vitális,¹ Pál T. Szabó,² Zoltán Kele,² Tamás Janáky² and Sándor Biró¹

Author for correspondence: Sándor Biró. Tel: +36 52 416 531. Fax: +36 52 416 531.
e-mail: sbiro@jaguar.dote.hu

¹ Department of Human Genetics, University Medical School of Debrecen, H-4012 Debrecen Nagyerdei körút 98, Hungary

² Department of Medicinal Chemistry, Albert Szent-Györgyi Medical University, H-6720 Szeged Dóm tér 8, Hungary

³ Department of Biochemistry, Leiden University, Gorlaeus Laboratory, PO Box 9502, 2300 RA Leiden, The Netherlands

The gene encoding factor C (*facC*), an extracellular signal protein involved in cellular differentiation, was cloned from *Streptomyces griseus* 45H, and the complete nucleotide sequence was determined. The deduced amino acid sequence was confirmed by HPLC/electrospray ionization-mass spectrometry analysis. The full-length protein consists of 324 amino acids and has a predicted molecular mass of 34523 Da. The mature extracellular 286 amino acid protein (31038 Da) is probably produced by cleaving off a 38 amino acid secretion signal sequence. Southern hybridization detected *facC* in several other *Streptomyces* strains, but database searches failed to identify a protein with significant homology to factor C. Expression of *facC* from a low-copy-number vector in *S. griseus* 52-1 resulted in a phenotypic effect similar to that given by exogenously added factor C protein.

Keywords: extracellular signalling, cell differentiation, sporulation, autoregulator, *Streptomyces*

INTRODUCTION

Streptomyces are Gram-positive soil bacteria. Their study is particularly attractive because their mycelial form of growth develops into spores, and the concomitant production of secondary metabolites includes many compounds of great medical and/or industrial importance. Extracellular regulatory molecules – also called autoregulators (Khokhlov, 1991) – play a key role in controlling cellular differentiation and secondary metabolism in streptomycetes (Horinouchi & Beppu, 1992).

Studies on morphological differentiation and its intimate connection to secondary metabolism in *Streptomyces coelicolor*, the prototype and most extensively studied *Streptomyces* strain, have recently been reviewed by

Chater (1998). In this strain, continuously increasing collections of developmental mutants and the corresponding cloned genes are available.

In *Streptomyces griseus* the loss of aerial mycelium formation, streptomycin production and ability to produce the low-molecular-mass γ -butyrolactone A-factor are frequently associated. Aerial mycelium formation and antibiotic synthesis can be restored by the addition of A-factor at nanomolar concentrations (Khokhlov, 1991). Studies of the genetics of A-factor biosynthesis have shown that A-factor-dependent initiation of sporulation involves phosphorylation of two regulatory proteins by cognate membrane-bound protein kinases (Horinouchi, 1996) and have also led to an understanding of the regulatory cascades involved in antibiotic biosynthesis (Horinouchi & Beppu, 1992).

Factor C was detected and later isolated as a protein from the culture fluid of *S. griseus* 45H (Szabó *et al.*, 1962), a strain that readily sporulates in liquid medium (Szabó *et al.*, 1961). Factor C induced the formation of precondia in liquid cultures of the susceptible *S. griseus*

This paper is dedicated to the memory of Professor Gábor Szabó.

Abbreviations: DIG, digoxigenin; ESI, electrospray ionization.

The GenBank accession number for the sequence reported in this paper is AF103943.

strain 52-1 (Vitális & Szabó, 1969), at concentrations as low as 0.5–1 ng ml⁻¹ (Szeszák *et al.*, 1991). This strain is otherwise blocked in submerged sporulation (Szabó *et al.*, 1961). Factor C also stimulated sporulation on solid medium (Szabó *et al.*, 1995). The pattern of proteins present in *S. griseus* 52-1 mycelium showed a specific change after the administration of factor C (Vitális *et al.*, 1988). The majority or all of the factor-C-like antigen was released into the fermentation liquid of *Streptomyces* cultures (Szeszák *et al.*, 1990, 1991).

Factor C has been purified to electrophoretic homogeneity by phosphocellulose and DNA-agarose chromatography. Its molecular mass was estimated to be about 34 kDa (Biró *et al.*, 1980), and its isoelectric point is around 9.9 (Szeszák *et al.*, 1991). By using polyclonal and monoclonal antibodies raised against factor C in conjunction with ELISA and immunoblotting techniques, the factor-C-like antigen was detected in 23 *Streptomyces* strains, in *Bacillus subtilis*, *Escherichia coli* and an archaeon, as well as in mammalian cells. On this evidence, factor C was assumed to be evolutionarily conserved.

Preliminary experiments aimed at examining the possible mechanism of action of factor C detected an effect on transcription (Szeszák & Szabó, 1973, Szabó *et al.*, 1984), on the release of potassium (Szeszák *et al.*, 1989) and on NAD glycohydrolase activity in washed mycelia (Szabó *et al.*, 1988). Factor C was also shown to bind to single-stranded and double-stranded heterologous DNA (Szabó *et al.*, 1984).

In the present paper the cloning, sequence analysis and expression of the cloned factor C gene, *facC*, are described.

METHODS

Strains and growth conditions. *E. coli* strains JM109 (Messing *et al.*, 1981) and XL-1 Blue (Stratagene) were hosts for pUC19 (Yanisch-Perron *et al.*, 1985) and pBluescript II KS+ (Stratagene), respectively, and were grown for 16 h at 37 °C in Luria-Bertani (LB) medium supplemented with 100 µg ampicillin ml⁻¹. Submerged cultures of *Streptomyces* strains were grown in filtered soybean (Szabó *et al.*, 1961) or tryptone-soy broth (Hopwood *et al.*, 1985) medium for 48 or 72 h on a rotary shaker at 28 °C. *S. griseus* 52-1 transformants were grown on R2YE medium (Hopwood *et al.*, 1985) supplemented with 50 µg thiostrepton ml⁻¹.

N-terminal amino acid sequencing. Factor C protein was blotted electrophoretically to PVDF membrane and sequenced by the Edman degradation method.

Molecular cloning and DNA manipulations. Purification of *Streptomyces* chromosomal DNA and transformation of *S. griseus* 52-1 protoplasts were carried out by standard methods (Hopwood *et al.*, 1985). Plasmid DNA was isolated from *E. coli* strains using the techniques described by Sambrook *et al.* (1989), except if required for DNA sequencing, when Wizard miniprep columns (Promega) were used. Restriction enzymes were used according to the instructions of the suppliers. To construct the mini gene-library of *S. griseus* 45H, chromosomal DNA was digested with *Sac*II. The DNA fragments

were separated electrophoretically on a 1% agarose gel in TAE buffer and transferred to a Hybond-N membrane (Amersham). The membrane was hybridized with the 5'-digoxigenin (DIG)-labelled 39-mer oligonucleotide (TIB MOLBIOL Syntheselabor; Table 1) in 5 × SSC, 2% (w/v) blocking reagent, 0.1% (w/v) N-lauroylsarcosine, 0.02% (w/v) SDS, and 50% (v/v) formamide at 42 °C, overnight. At highest stringency the membrane was washed in 0.5 × SSC, 0.1% (w/v) SDS at 55 °C for 30 min. Hybridizing bands were detected by immunostaining using a DIG nucleic acid detection kit (Boehringer Mannheim). On the basis of the hybridization signal, the *Sac*II fragments in the 2.8 kb region from a parallel gel were purified using a DNA recovery purification kit (Hybaid). The purified DNA fragments were cloned in *Sac*II-digested pBluescript II KS+ (Stratagene) and the recombinant vector was used to transform *E. coli* XL-1 Blue. Ampicillin-resistant white colonies were screened by colony hybridization with the 39-mer DIG-labelled probe as described above.

PCR amplification of the 5' end of the gene. The 5' end of the coding region of the factor C gene was amplified from *Sac*I-digested chromosomal DNA of *S. griseus* 45H: 100 ng DNA, 2.5 U *Taq* DNA polymerase (Amersham), 100 pmol degenerate synthetic oligonucleotide primers (Table 1), and 30 µM dNTP were reacted in an amplification protocol consisting of 1 cycle at 94 °C (5 min), 50 °C (2 min), 72 °C (2 min); 8 cycles at 94 °C (1 min), 50 °C (2 min), 72 °C (1 min); 50 cycles at 92 °C (1 min), 50 °C (1 min), 72 °C (1 min); 1 cycle at 94 °C (1 min) 50 °C (2 min), 72 °C (5 min). The amplified 76 bp DNA fragment was cloned into the *Sma*I site of pUC19 to form p8B and sequenced to verify amplification of the desired DNA fragment.

DNA sequencing. Plasmids were sequenced either by the dideoxy chain-termination method (Sanger *et al.*, 1977) using the T7 Sequencing Kit (Pharmacia Biotech) and [α -³⁵S]dCTP α S as the labelled nucleotide, or with an ABI 373 automated sequencer (Applied Biosystems) using a dye terminator cycle sequencing kit with Amplitaq DNA polymerase FS (Applied Biosystems). In all sequencing reactions, double-stranded plasmids were used with universal or specific oligonucleotides (Table 1) as primers. Specific oligonucleotide primers were synthesized by Pharmacia Biotech. All sequence information reported was obtained from the independent analysis of both strands.

Computer analysis of nucleotide and amino acid sequence. Sequence data were analysed with the PC/GENE nucleic acid and protein sequence analysis software system (IntelliGenetics). For coding region analysis, FramePlot Version 2.2.1 (based on Bibb *et al.*, 1984, developed by Ishikawa Jun) was also used. Amino acid sequence homology searches made use of the BLAST network service and non-redundant protein sequence databases (Altschul *et al.*, 1997).

Molecular mass and sequence determination of tryptic fragments. Gel bands containing about 13 µg protein from SDS-PAGE were cut into 1 mm² pieces and soaked in 400 µl 100 mM NH₄HCO₃ (pH 8) with stirring for 10 min. The supernatant was discarded and the gel slices were dehydrated in 800 µl acetonitrile. The protein within the gel slices was reduced with 300 µl 10 mM dithiothreitol at 56 °C for 1 h and repeatedly washed with NH₄HCO₃ and acetonitrile. For alkylation, 300 µl 55 mM iodoacetamide was added at room temperature for 45 min. The washing procedure was repeated three times and the gel slices were dried under vacuum. For tryptic digestion (18 h at 37 °C) the ratio of factor C to trypsin (w/w) was 10:1. The peptide fragments were then washed

Table 1. Oligonucleotides used in PCR, screening and sequencing

Oligonucleotide	Nucleotide sequence	Nucleotide position
Oligonucleotides used in PCR amplification		
1. Forward primer	5'-GC(GC)GT(GC)CC(GC)GC(GC)AC(GC)AAG-3'	1121-1138
2. Reverse primer	5'-CGTCGTG(CG)AGCTT(GC)GCGTG-3'	1196-1178
Oligonucleotide used as a probe in screening		
1.	5'-CGGTTTCAGCCTGACCGAGCCGTCCCACGACCTGTTCCGG-3'	1139-1177
Oligonucleotides used in sequencing		
Forward primers		
1.	5'-GGCCGACCTGTGCATCAC-3'	1283-1299
3.	5'-GCACCTCCTCGGCACTCG-3'	1485-1502
5.	5'-GAGCCCGAGGGCATGGGC-3'	1859-1876
7.	5'-AACGAGGCTCCCGGCGGG-3'	307-324
8.	5'-GCTGGGCGCGGGCTCGGC-3'	633-650
9.	5'-CGCACCACCTGTGCTCCC-3'	933-950
14.	5'-GGGAGGTCCGCCCGGCG-3'	2029-2046
Reverse primers		
2.	5'-GCCGGAACAGGTCGTGGG-3'	1178-1160
4.	5'-CCGAGCCGCAGCAGCCCG-3'	1059-1042
6.	5'-TCCTCCTGTGCGCCCCC-3'	736-719
10.	5'-GGCCACGAACAGCCGCCG-3'	1252-1235
11.	5'-GATGGCGCAGGTCATCTC-3'	1549-1532
12.	5'-CAGCGTGGAACCGGCGCG-3'	1849-1832
13.	5'-ATCGGCTTCCTCATGCTG-3'	2128-2111

with 400 µl NH₄HCO₃ solution and extracted with a 1:1 (v/v) acetonitrile/water mixture containing 5% (v/v) formic acid, concentrated in a Speedvac Evaporator, and dissolved for HPLC-MS analysis in 1% (v/v) acetic acid. The peptide fragments were separated using a 0.32 × 200 mm fused silica capillary column (Novotny, 1988) packed with Vydac C₁₈. For mass spectrometric measurements a Finnigan TSQ 7000 tandem mass spectrometer equipped with a microelectrospray ion source (Kele *et al.*, 1998) was used. The amino acid sequence of the most intense peptide fragments was determined by HPLC-MS/MS measurements. Molecular masses obtained in HPLC-MS experiments were compared with theoretical tryptic digestion profiles of candidate proteins generated by a computer program (Mann *et al.*, 1993).

Southern hybridization. Chromosomal DNA samples from 12 *Streptomyces* strains were digested to completion with *Sac*II. The DNA fragments were size-fractionated by gel electrophoresis on a 1% (w/v) TAE agarose gel and transferred to a Hybond-N (Amersham) membrane using the manufacturer's instructions. Hybridization was carried out at 42 °C in a mixture containing 50% (v/v) formamide, 2 × SSC, 5 × Denhardt's solution, 0.2 mg denatured salmon sperm DNA ml⁻¹ and 0.1% (w/v) SDS. The highest stringency washing was in 0.1 × SSC, 0.1% (w/v) SDS at 65 °C for 30 min. The 860 bp *Eco*RV-*Sal*I fragment (nucleotide positions 930-1790 in Fig. 1), was used as the probe. The DNA was radiolabelled by random priming with nonamers and the Klenow fragment of *E. coli* DNA polymerase I, using a Megaprime Kit from Amersham to incorporate [α -³²P]dCTP.

Construction of pSGF4, a low-copy-number vector containing *facC*. The 2.1 kb *Apal* fragment from pBZ3, harbouring the

entire factor C structural gene and approximately 560 bp of upstream region, was inserted into *Sma*I-digested pIJ2925 (Janssen & Bibb, 1993) lacking the *Bam*HI site, after filling in the 5' protruding ends of the insert with the Klenow fragment of DNA polymerase I and dNTPs. The resulting vector was designated pSGF2. To introduce *facC* into *S. griseus* 52-1 we used pHJL401, an *E. coli*-*Streptomyces* shuttle vector containing the *E. coli* pUC19 and *Streptomyces* SCP2* (Lydiat *et al.*, 1985) origins of replication (Larson & Herschberger, 1986) giving 100-200 copies per cell in *E. coli* and approximately 10 copies per chromosome in *Streptomyces*. For this purpose, the 2.1 kb *Eco*RI-*Hind*III insert of pSGF2 was cloned into pHJL401 digested with the same enzymes, resulting in pSGF4.

RESULTS AND DISCUSSION

N-terminal amino acid sequence of factor C and PCR amplification of the 5' end of the gene, *facC*

N-terminal amino acid sequencing of the purified, electrophoretically homogeneous factor C protein (Biró *et al.*, 1980) resulted in the following tentative order for the first 29 amino acids: Ala-Val-Pro-Ala-Thr-Lys-Arg-Phe-Ser-Leu-Thr-Glu-Pro-Ser-His-(Asp/Phe)-Leu-Phe-Arg-His-Ala-Lys-Leu-His-Asp-(Gly/Ala)-Arg-Val-Gln. The residues at positions 16 (Asp/Phe) and 26 Gly/Ala) were not unambiguously identified, and were not used in primer constructions. The underlined amino acids were chosen to design two PCR primers. Since all glycine and alanine codons begin with G, the first base of the ambiguous Gly/Ala was also included in the reverse

```

857 CCACCGCCGCGCCGCACTCCGGTGCCCGGTGGCGGACCCTGGACAAGCCCCGTCCGCACG
917 GCAACGATGCCGATATCGCACCACCTTGTGCTCCCACCATCGGAAGTGAGACAACGTGAGT
      RBS
977 GACCTCCGTGACGTACCGGAGCGGGACGGCGTGCCGGCGGCTACGACCTCGTCGCCGAGC
      ValProAlaAlaThrThrSerSerProSer
      secretion signal sequence
1037 CGGCGCGGGCTGCTCGGGCTCGGCGGGCGGGCTCGTCGCCGCTCCGCCTTCGGCCTGAGC
      ArgArgGlyLeuLeuArgLeuGlyGlyGlyLeuValAlaAlaSerAlaPheGlyLeuSer
1097 TTCGGCGCGGGCAGTGGCGGGCGGGCGGTACCGGCCACCAAGCGTTTCAGCCTGACCGAG
      PheGlyAlaGlySerAlaAlaAlaAlaValProAlaThrLysArgPheSerLeuThrGlu
1157 CCGTCCCACGACCTGTTCCGGCACGCAAGCTGCACGACGCGGGTCCAGCAGTCGTTTC
      ProSerHisAspLeuPheArgHisAlaLysLeuHisAspAlaArgValGlnGlnSerPhe
1217 ACCTTCGACATCGTCAACCGGGCGGCTGTTCTGTCGGCCAGTTGAAGTCGGGCTCGCCGAC
      ThrPheAspIleValAsnArgArgLeuPheValAlaGlnLeuLysSerGlySerProAsp
1277 GACAGCGGCGACCTGTGCATCACCCAGCTGGACTTCTCGGGCAACAAGCTGGGCCACATG
      AspSerGlyAspLeuCysIleThrGlnLeuAspPheSerGlyAsnLysLeuGlyHisMET
1337 TACCTGCTCGGCTTCGGCCACGGCGTCTCCATCGGCGCCAGCCGGTCGGCGCGGACACG
      TyrLeuLeuGlyPheGlyHisGlyValSerIleGlyAlaGlnProValGlyAlaAspThr
1397 TACCTGTGGACGGAGGTGGACGTCAACAGCAACGCGCGTGGCACCCGGCTGGCCCGGTTTC
      TyrLeuTrpThrGluValAspValAsnSerAsnAlaArgGlyThrArgLeuAlaArgPhe
1457 AAGTGAACAACGGCGCGACGCTCAGCCGACCTCCTCGGCACTCGCCAAGCACCAGCCG
      LysTrpAsnAsnGlyAlaThrLeuSerArgThrSerSerAlaLeuAlaLysHisGlnPro
1517 GTCCCCGGCGCCACCGAGATGACCTGCGCCATCGACCCGGTCAACAACCGCATGGCGATC
      ValProGlyAlaThrGluMETThrCysAlaIleAspProValAsnAsnArgMETAlaIle
1577 CGCTACCTCACCGCTCGGGCCGCGCTACGGCATCTACAACGTCGCCGACATCGCCGCC
      ArgTyrLeuThrAlaSerGlyArgArgTyrGlyIleTyrAsnValAlaAspIleAlaAla
1637 GCGTCTACGACAAGCCGCTCTCCGACGTCCCGCACCCACGGGCTGGGCACCTTCCAG
      GlyValTyrAspLysProLeuSerAspValProHisProThrGlyLeuGlyThrPheGln
1697 GGGTACGCGTGTACGGCAGCTACGTCTACAGCTCACCGGCAACCCGTACGGCCCGGAC
      GlyTyrAlaLeuTyrGlySerTyrValTyrGlnLeuThrGlyAsnProTyrGlyProAsp
1757 AACCCGAACCCGGGCAACTCCTACGTCTCCAGCGTCGACGTCAACACCGCGCCCTGGTC
      AsnProAsnProGlyAsnSerTyrValSerSerValAspValAsnThrGlyAlaLeuVal
1817 CAGCGGGCCTTCAACCGCGCCGGTTCCACGCTGACCTTCGCGAGCCCGAGGGCATGGGC
      GlnArgAlaPheThrArgAlaGlySerThrLeuThrPheArgGluProGluGlyMETGly
1877 ATCTACCGGACGGCCGCGGCGAGGTCCGCTCTCTCTGGGCTTCGCTCGGGTGTGGCA
      IleTyrArgThrAlaAlaGlyGluValArgLeuPheLeuGlyPheAlaSerGlyValAla
1937 GGGGACCGCCCGCAGCAATCTGTTCTACAAGAACGTCCTGATCTGA
      GlyAspArgArgSerAsnLeuPheTyrLysAsnValLeuIle---

```

Fig. 1. Nucleotide and amino acid sequence of the relevant portion (nt 857–1981) of the sequenced 2194 bp DNA fragment. The nucleotide sequence of the RBS and secretion signal are singly underlined. The amino acid sequences of peptide fragments generated by tryptic digestion of the protein and verified by HPLC/ESI-MS/MS analysis are underlined by single, double, dotted and boldface lines in order to distinguish the neighbouring and (due to partial digestion) overlapping peptides.

primer. The forward and reverse primers contained five and two degenerate positions, respectively (Table 1), designed according to the biased codon usage of *Streptomyces* strains (Bibb *et al.*, 1984). All but one triplet of the chromosomal gene, corresponding to the degenerate positions of the primers, as expected, turned out to contain G or C in the third position. The chromosomal triplet of the second amino acid (valine) of the forward primer is GTA but this mismatch did not affect the usefulness of the primer. In the PCR reaction *SacI*-digested *S. griseus* 45H chromosomal DNA was used as the template. Three short PCR products were amplified (approx. 38, 76 and 152 bp; data not shown). Amplification of a 76 bp fragment was expected, based on the N-terminal sequence. That fragment was extracted from the gel, subcloned into pUC19 and sequenced. Sequence analysis verified that the DNA fragment indeed corresponded to the known N-terminal amino acid sequence.

Cloning of *facC*

Instead of screening a complete chromosomal gene library, we used a 'mini-genebank', which was constructed in pBluescript II KS+ by cloning the gel-purified *SacII* fragments of about 2.8 kb in size, corresponding to the signal obtained by Southern hybridization. The 39-mer oligonucleotide flanked by the two primers in the 76 bp amplified fragment was used as probe. Colony hybridization identified one clone (pBZ3) with a 2.9 kb insert that repeatedly gave a positive signal in high stringency hybridizations. The DNA sequence of this clone was determined.

Sequence analysis of pBZ3

Sequencing was started from primers complementary to the T3 and T7 promoters flanking the cloned DNA fragment in the vector. The 39-mer DIG-labelled oligo-

Fig. 2. Coding region analysis by FramePlot version 2.2.1 (developed by Ishikawa Jun). Window size 40 nt, step 5 nt, minimum ORF 20 nt. The abscissa shows the analysed part (nt 857–2100) of the sequence. The ordinate shows the G + C content of the DNA at the third codon position. Only the three forward frames are shown. Above the figure the numbered lines represent the ORFs in the six possible (three forward and three backward) frames. > corresponds to start codons and | to stop codons. The ORF of factor C is in the first reading frame.

nucleotide served as an internal primer. Gaps were filled by directed sequencing from 'walking primers'. Both strands of the DNA were sequenced at least twice. FramePlot and coding region analysis by PC/GENE programs identified a 975 bp ORF, typical of *Streptomyces*, with 96.9% GC in the third letter position and an overall G + C content of 70.7 mol% (Fig. 2). The coding sequence starts with a GTG codon at nt 1007 and extends to the TGA stop codon at nt 1978 (Fig. 1). It encodes a protein of 324 amino acids with a predicted molecular mass of 34523 Da. Computer analysis of this protein located a sequence Ser-Ala-Ala-Ala/Ala/Val-Pro-Ala that contains two potential secretory signal cleavage sites at the positions indicated by slashes. Since the known N-terminal sequence began with the sequence Ala-Val-Pro it follows that the factor C propeptide contains a 38 amino acid secretion signal conforming to the consensus for prokaryotic signal sequences. It has the usual three domains: a positively charged N domain, a hydrophobic H domain required to initiate translocation across the cytoplasmic membrane, and a C domain preceding the cleavage site (Pugsley, 1993). The signal peptidase cleavage site also conforms to the $-3, -1$ rules (von Heijne, 1986). Since factor C was identified and isolated from the culture fluid of *S. griseus* 45H, and most or all of the factor-C-like antigen was found in the fermentation liquid, secretion of the protein was predictable. The mature protein contained 286 amino acids and its molecular mass was calculated to be 31038 Da, in good agreement with the 34500 Da estimated by SDS-PAGE (Biró *et al.*, 1980), especially if we consider that factor C is a strongly basic protein. Its calculated isoelectric point was 9.59, which is close to our previous estimate of 9.9 (Szeszák *et al.*, 1991).

When the N-terminal 29 amino acid sequence was compared to known sequences in the databases, signifi-

cant homology to several zinc-finger-type regulatory proteins was detected. It was also known that factor C binds to single- and double-stranded DNA, and that its cytomorphological effect can be increased by low concentrations of zinc (S. Biró and others, unpublished results). This line of evidence was further strengthened by the successful purification of factor C on a zinc affinity column (Szeszák *et al.*, 1997). From these results we believed the factor C to be a zinc-finger-type regulatory protein. However, the deduced complete amino acid sequence of the protein did not support our previous assumption. Interestingly, a hypothetical 1 nt insertion into the sequence at nt 1211 would result in an alternative similarly sized ORF whose product would contain about 20 His residues, most of them positioned regularly, 5–6 amino acids apart. FramePlot analysis of this alternative ORF did not show a typical *Streptomyces* coding sequence, but the difficulties in sequencing high-G + C-content *Streptomyces* DNA prompted us to determine the amino acid sequence of factor C by independent methods. Therefore, we determined the molecular mass of the protein and of tryptic digests by HPLC/ESI-MS/MS techniques.

Molecular mass determination of factor C and amino acid sequencing of its tryptic fragments

Only a small amount of factor C is produced by *S. griseus* 45H. Typically a few micrograms of purified protein can be obtained from 10 litres of culture broth. Therefore we used capillary chromatography/microelectrospray mass spectrometry, a method that required only minute amounts of protein for molecular mass determination of factor C and sequence verification of tryptic fragments. The molecular mass obtained for the mature factor C was 31045 ± 9 Da, very close to the value of 31038 Da calculated from the sequence.

After tryptic digestion the most intense peptide peaks were selected for sequence analysis in the HPLC-MS/MS measurements. The peptides identified by MS/MS are underlined in Fig. 1. Although the identified peptide fragments do not cover the whole sequence, the fragments are well distributed over the N-terminal, middle and C-terminal regions. Therefore, we concluded that the ORF identified by FramePlot analysis and the deduced amino acid sequence of the factor C protein are correct. One possible explanation for our previous, firmly established finding that factor C binds very specifically to zinc affinity columns, is that binding occurs in association with another protein. To study this possibility and detect any other protein that eluted with, or later than, factor C, we are examining proteins in *S. griseus* 45H cultures.

Factor C shows no significant homology to other proteins in the databases

Comparison of the amino acid sequence of the mature factor C protein with proteins in the databases using BLAST Search (Altschul *et al.*, 1997) showed a low level of

Fig. 3. Southern blot of various *Streptomyces* strains. (a) Ethidium-bromide-stained gel; (b) hybridization pattern. The probe was the ³²P-labelled 860 bp *EcoRV*–*SacI* fragment (nucleotide positions 930–1790 in Fig. 1), covering 80% of the coding region and a short (75 nt) upstream sequence. Lanes: λ, phage λ DNA digested with *HindIII*. All the other DNAs were digested with *SacII*. 1, pBZ3 plasmid DNA; 2, *S. clavuligerus*; 3, *S. parvulus*; 4, *S. griseoviridans*; 5, *S. coelicolor* M145; 6, *S. albus* 391; 7, *S. griseobrunneus*; 8, *S. flavofungini*; 9, *S. albus* R-55; 10, *S. griseus* 45H; 11, *S. griseus* 52-1; 12, *S. lividans* 1326; 13, *S. fradiae*.

similarity to teichoic acid biosynthesis protein C (*tagC*) of *B. subtilis* 168 (29% amino acid identity with 6 gaps). This seemed at first to be an interesting homology, because teichoic acid is a cell wall component of Gram-positive bacteria and expression of the *tag* genes is sporulation specific (Mauël *et al.*, 1991, 1994). However, recent reports have shown that *tagC* corresponds to *dinC* (Cheo *et al.*, 1991, 1993), and thus belongs to the SOS regulon. Therefore, the relevance of the low similarity between factor C and the TagC protein, if any, is unclear.

***facC* is present in several *Streptomyces* strains that sporulate in submerged culture**

To test for the presence of the factor C gene in other *Streptomyces* strains, Southern blots of chromosomal digests of 12 strains from 10 species were hybridized with a ³²P-labelled DNA fragment spanning nt positions 930–1790 (Fig. 1), covering approximately 75 bp of upstream region (Zs. Birkó and others, unpublished results) and about 80% of the coding region. As expected, a strong hybridization signal was observed in the lane with DNA from *S. griseus* 45H (Fig. 3). The lanes containing DNA from *S. albus* R-55, *S. flavofungini* and *S. albus* 391 each gave a strong signal, suggesting the presence of a gene with high homology to *facC*. No comparable band was observed with the DNA from any other strains tested, including our factor-C-sensitive test strain *S. griseus* 52-1. This contrasts with our previous results obtained using monoclonal antibody raised against factor C, which showed the presence of a factor-C-like antigen in all tested *Streptomyces* strains, including some of those that failed to hybridize in the present experiment (Szeszák *et al.*, 1990). This might be explained by the high stringency of washing. In other experiments when the DIG-labelled 39-mer oligo-

nucleotide (Table 1) was hybridized with chromosomal digests of the same strains, and a lower stringency washing was used, we could detect a hybridizing band or bands in all tested strains (*S. Biró*, unpublished results). The meaning and significance of these results await further clarification but our findings are reminiscent of those of Onaka *et al.* (1998), who found two low homologues of the *S. griseus* A-factor receptor protein, which has a regulatory role in secondary metabolism and morphogenesis, in *S. coelicolor*. The four *Streptomyces* species shown to harbour a high homologue of *facC* are known to sporulate in submerged culture (Kendrick & Ensign, 1983; Vitális *et al.*, 1981; Daza *et al.*, 1989; *S. Biró*, unpublished results). This points to a possible involvement of factor C in the onset of submerged sporulation.

Recently *ssgA* has been recognized as a gene involved in submerged sporulation of *S. griseus* B2682 (Kawamoto & Ensign, 1995; Kawamoto *et al.*, 1997). Like factor C, the *ssgA* gene product is a regulatory protein that shows no relevant similarity to any known protein. Factor C and the intracellular *ssgA* gene product might form parts of a regulatory mechanism unique to those *Streptomyces* species able to sporulate in submerged culture.

Expression of *facC* in *S. griseus* 52-1

To study the phenotypic effect of the cloned gene, an *E. coli*–*Streptomyces* SCP2*-based shuttle vector containing the entire structural gene for factor C and approximately 560 bp of upstream region was constructed; it was designated pSGF4. It is present at approximately 10 copies per chromosome in *Streptomyces*. Preliminary transcriptional analysis revealed that the 560 bp upstream region contains the *facC* promoter (Zs. Birkó and others, unpublished results), allowing expression of *facC* from its own promoter. In submerged culture,

Fig. 4. Anoptral (negative phase-contrast) micrographs of 72 h cultures of: (a) *S. griseus* 45H, (b) *S. griseus* 52-1, (c) *S. griseus* 52-1 grown with exogenously added factor C, (d) *S. griseus* 52-1/pSGF4. Magnification $\times 1700$.

under our standard test conditions, *S. griseus* 52-1 forms long, smooth, non-branching vegetative hyphae. It fails to sporulate and lacks the reproductive, frequently branching hyphae containing club-like thickening at their ends (Fig. 4b), characteristic of the factor C producer *S. griseus* 45H (Fig. 4a; Vitális *et al.*, 1963). Addition of purified factor C to submerged cultures of *S. griseus* 52-1 resulted in a morphology very similar to that of *S. griseus* 45H (Fig. 4c). Introducing pSGF4 into *S. griseus* 52-1 resulted in a phenotype (Fig. 4d) almost identical to that of the *facC*⁺ strain *S. griseus* 45H. The transformant showed this typical cytomorphology irrespective whether it was grown in the presence or absence of thiostrepton.

In summary, we believe that factor C plays an essential role in the development of *S. griseus* and is probably involved in the onset of sporulation in submerged culture.

ACKNOWLEDGEMENTS

We are grateful to Anikó Pay for the automated DNA sequencing, to Wolfgang Piepersberg for the N-terminal protein sequencing and to János Aradi for synthesis of the PCR primers. Andrea Kissné and Katalin Szabó are thanked

for technical assistance. This work was supported by research grants T 016796 and T 025553 by the National Scientific and Research Fund (OTKA) and by the Hungarian Academy of Sciences (to S.B.), and also by research grants F 025602 by OTKA (to P.T.Sz.) and T 026624 by OTKA (to T.J.). Zs.B. was the recipient of a short term TEMPUS PHARE JEP student mobility grant.

REFERENCES

- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402.
- Bibb, M. J., Findlay, P. R. & Johnson, M. W. (1984). The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein coding sequences. *Gene* **30**, 157–166.
- Biró, S., Békési, I., Vitális, S. & Szabó, G. (1980). A substance affecting differentiation in *Streptomyces griseus*. *Eur J Biochem* **103**, 359–363.
- Chater, K. F. (1998). Taking a genetic scalpel to the *Streptomyces* colony. *Microbiology* **144**, 1465–1478.
- Cheo, D. L., Bayles, K. W. & Yasbin, R. E. (1991). Cloning and characterization of DNA damage-inducible promoter regions from *Bacillus subtilis*. *J Bacteriol* **173**, 1696–1703.

- Cheo, D. L., Bayles, K. W. & Yasbin, R. E. (1993). Elucidation of regulatory elements that control damage induction and competence induction of the *Bacillus subtilis* SOS system. *J Bacteriol* **175**, 5907–5915.
- Daza, A., Martin, J. F., Dominguez, A. & Gil, J. A. (1989). Sporulation of several species of *Streptomyces* in submerged culture after a nutritional downshift. *J Gen Microbiol* **135**, 2483–2491.
- von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res* **14**, 4683–4690.
- Hopwood, D. A., Bibb, M. J., Chater, K. F., Kieser, T., Bruton, C. J., Lydiate, D. J., Smith, C. P., Ward, J. M. & Schrempf, H. (1985). *Genetic Manipulation of Streptomyces*. A Laboratory Manual. Norwich: John Innes Foundation.
- Horinouchi, S. (1996). *Streptomyces* genes involved in aerial mycelium formation. *FEMS Microbiol Lett* **141**, 1–9.
- Horinouchi, S. & Beppu, T. (1992). Autoregulatory factors and communication in Actinomycetes. *Annu Rev Microbiol* **46**, 377–398.
- Janssen, G. R. & Bibb, M. J. (1993). Derivatives of pUC18 that have *Bgl*II sites flanking a modified multiple cloning site and that retain the ability to identify recombinant clones by visual screening of *Escherichia coli* colonies. *Gene* **124**, 133–134.
- Kawamoto, S. & Ensign, J. C. (1995). Cloning and characterization of a gene involved in regulation of sporulation and cell division in *Streptomyces griseus*. *Actinomycetes* **9**, 136–151.
- Kawamoto, S., Watanabe, H., Hesketh, A., Ensign, J. C. & Ochi, K. (1997). Expression of the *ssgA* gene product, associated with sporulation and cell division in *Streptomyces griseus*. *Microbiology* **143**, 1077–1086.
- Kele, Z., Janáky, T., Mészáros, T., Fehér, A., Dudits, D. & Szabó, P. T. (1998). Capillary chromatography/microelectrospray mass spectrometry used for the identification of putative cyclin-dependent kinase inhibitory protein in *Medicago*. *Rapid Commun Mass Spectrom* **12**, 1564–1568.
- Kendrick, K. E. & Ensign, J. C. (1983). Sporulation of *Streptomyces griseus* in submerged culture. *J Bacteriol* **155**, 357–366.
- Khokhlov, A. S. (1991). *Microbial Autoregulators*. Chur: Harwood Academic Publishers.
- Larson, J. L. & Herschberger, C. L. (1986). The minimal replicon of a streptomycete plasmid produces an ultrahigh level of plasmid DNA. *Plasmid* **15**, 199–209.
- Lydiate, D. J., Malpartida, F. & Hopwood, D. A. (1985). The *Streptomyces* plasmid SCP2*: its functional analysis and development into useful cloning vectors. *Gene* **35**, 223–235.
- Mann, M., Hojrup, P. & Roepstorff, P. (1993). Use of mass spectrometric molecular weight information to identify proteins in sequence database. *Biol Mass Spectrom* **22**, 338–345.
- Mauël, C., Young, M. & Karamata, D. (1991). Genes concerned with synthesis of poly(glycerol phosphate), the essential teichoic acid in *Bacillus subtilis* strain 168, are organized in two divergent transcription units. *J Gen Microbiol* **137**, 929–941.
- Mauël, C., Young, M., Monsutti-Grecescu, A., Marriott, S. A. & Karamata, D. (1994). Analysis of *Bacillus subtilis* tag gene expression using transcriptional fusions. *Microbiology* **140**, 2279–2288.
- Messing, J., Crea, R. & Seeburg, P. H. (1981). A system for shotgun DNA sequencing. *Nucleic Acids Res* **9**, 309–321.
- Novotny, M. (1988). Recent advances in microcolumn liquid chromatography. *Anal Chem* **60**, 500A–510A.
- Onaka, H., Nakagawa, T. & Horinouchi, S. (1998). Involvement of two A-factor receptor homologues in *Streptomyces coelicolor* A3(2) in the regulation of secondary metabolism and morphogenesis. *Mol Microbiol* **28**, 743–753.
- Pugsley, A. P. (1993). The complete general secretory pathway in gram-negative bacteria. *Microbiol Rev* **57**, 50–108.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*. 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5467.
- Szabó, G., Barabás, G. & Vályi-Nagy, T. (1961). Comparison of *Streptomyces griseus* strains which produce streptomycin and those which do not. *Arch Mikrobiol* **40**, 261–274.
- Szabó, G., Vályi-Nagy, T. & Vitális, S. (1962). A new factor regulating life cycle of *Streptomyces griseus*. In *Genetics of Microorganisms, Proceedings of a Symposium on Heredity and Variability of Microorganisms*, pp. 282–292. Edited by V. D. Timakova. Moscow: State Publishing House of Medical Literature.
- Szabó, G., Biró, S., Trón, L., Valu, G. & Vitális, S. (1984). Mode of action of factor C upon the differentiation process of *Streptomyces griseus*. In *Biochemical and Biomedical Aspects of Actinomycetes*, pp. 197–214. Edited by L. Ortiz-Ortiz, L. F. Bojalil & V. Yakoleff. Orlando, FL: Academic Press.
- Szabó, G., Szeszák, F., Vitális, S. & Tóth, F. (1988). New data on the formation and mode of action of factor C. In *Biology of Actinomycetes '88*, pp. 324–329. Edited by Y. Okami, T. Beppu & H. Ogawara. Tokyo: Japan Scientific Societies Press.
- Szabó, G., Vitális, S., Szeszák, F. & Biró, S. (1995). Developmental heterogeneity of *Streptomyces griseus* conidia. In *Sekundärmetabolismus bei Mikroorganismen*, pp. 109–122. Edited by W. Kuhn & H.-P. Fiedler. Tübingen: Attempto Verlag.
- Szeszák, F. & Szabó, G. (1973). Alteration of RNA synthesis *in vitro* with an endogenous regulating factor of cytodifferentiation of *Streptomyces griseus*. *Acta Biol Acad Sci Hung* **24**, 11–17.
- Szeszák, F., Vitális, S. & Szabó, G. (1989). Factor C, a regulatory protein of *Streptomyces griseus*, induces release of potassium from the mycelium. *J Basic Microbiol* **29**, 233–240.
- Szeszák, F., Vitális, S., Tóth, F., Valu, G., Facht, J. & Szabó, G. (1990). Detection and determination of factor C – a regulatory protein – in *Streptomyces* strains by antiserum and monoclonal antibody. *Arch Microbiol* **154**, 82–84.
- Szeszák, F., Vitális, S. & Szabó, G. (1991). Presence of factor C in streptomycetes and other bacteria. In *Genetics and Product Formation in Streptomyces*, pp. 11–18. Edited by S. Baumberg, H. Krügel & D. Noack. New York: Plenum.
- Szeszák, F., Vitális, S., Biró, S. & Dalmi, L. (1997). Amino acid sequence homology of factor C produced by *Streptomyces griseus* with regulatory proteins of zinc finger type. *Acta Biol Acad Sci Hung* **48**, 265–273.
- Vitális, S. & Szabó, G. (1969). Cytomorphological effect of factor C in submerged cultures on the hyphae of *Streptomyces griseus* strain No. 52-1. *Acta Biol Acad Sci Hung* **20**, 85–92.
- Vitális, S., Szabó, G. & Vályi-Nagy, T. (1963). Comparison of the morphology of streptomycin producing and non-producing strains of *Streptomyces griseus*. *Acta Biol Acad Sci Hung* **14**, 1–15.
- Vitális, S., Biró, S., Vargha, G., Békési, I. & Szabó, G. (1981). Differentiation and its regulation in submerged culture of *Streptomyces griseus*. In *Actinomycetes (Zentbl Bakteriell Suppl*

11), pp. 153–156. Edited by K. P. Schaal & G. Pulverer. New York: Springer.

Vitális, S., Valu, G., Békési, I., Szeszák, F. & Szabó, G. (1988). Changes of protein pattern of *Streptomyces griseus* strains that characterize differentiation of mycelia in submerged culture. *J Basic Microbiol* **28**, 393–407.

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

.....
Received 5 March 1999; revised 13 May 1999; accepted 1 June 1999.