Ph.D. THESIS

Molecular genetic analysis of the gene of a pleiotropic autoregulatory protein in *Streptomyces griseus*

Zsuzsanna Birkó

Supervisor: Dr. Sándor Biró

University of Debrecen
Medical and Health Science Center
Department of Human Genetics

INTRODUCTION

Streptomycetes are Gram-positive, aerobic soil bacteria, which belong to the order of Actinomycetales. The study of Streptomycetes is particularly attractive, because of their mycelial life cycle that finally results in sporulation and the production of secondary metabolites. Streptomycetes produce approximately 70 % of the more than 12000 known antibiotics, including many with important applications in human medicine and agriculture.

The life cycle of *Streptomyces* strains on solid medium

Early in the life cycle of *Streptomyces* strains on solid medium germ tube emerges from the unigenomic spore and its cell wall grows mainly at the tip. Later new tips arise trough lateral branching while several DNA replications take place. Hyphal compartments contain many copies of the genome because the formation of the vegetative septa is infrequent. Vegetative hyphae grow into the agar medium therefore they are called substrate mycelia. After three or four days as older parts of the substrate mycelium produce aerial mycelium in response to nutrient limitation or other physiological stresses, most cells in the substrate mycelium die. After elongation septa are produced at regular intervals along the aerial hyphae to form unigenomic compartments, each are destined to become spores. Subsequently the walls thicken and round up to form the ovoid spores in chains.

The life cycle of *Streptomyces* strains in liquid medium

Some *Streptomyces* species like *S. griseus* have the ability to undergo sporulation when grown in liquid culture. During the vegetative phase hyphal branching allows quasiexponential growth kinetics. Subsequently in the reproductive phase at about 30 to 36 h secondary metabolite production begins and the hyphae start to differentiate in two different ways. As a result of it old vegetative and reproductive (sporogenic) hyphae are distinguishable. After growth stops the reproductive hyphae undergo multiple cell division to give unigenomic compartments each of which develops into spores.

Regulation of the life cycle in *S. griseus*

In the regulation of aerial mycelium and spore formation different regulatory cascades are involved. Regulatory cascades for differentiation and secondary metabolism in Actinomycetes include autoregulatory factors which are low-molecular-weight molecules (like A-factor in *S. griseus*) or proteins (like factor C in *S. griseus*) effective at very low concentrations.

The model for A-factor regulatory cascade leading to aerial mycelium formation

A-factor (2-isocapryloyl-3R-hydroxymethyl-γ-butyrolactone) is a signal molecule that acts at a concenration of 10^{-9} M as a switch for streptomycin production and aerial mycelium formation in *S. griseus*. Consistent with the idea that A-factor is a microbial hormone an A-factor-specific receptor protein (ArpA) is present in the cytoplasmic fraction. During the early stage of growth in the absence of A-factor, ArpA itself acts as a repressor-type regulator, binding in a dimeric form to a particular stretch of DNA. A-factor produced in a growth-dependent manner is accumulated gradually until the end of quasiexponential growth and at a certain intracellular concentration (32-160 nM) relieves the repression. Through a so far unknown mechanism this leads to induction of AmfR production. *amfR* encodes a response regulator of two-component regulatory system typical of prokaryotes and plays an important role in aerial mycelium formation. The phosphorylated form of AmfR conceivably acts as a transcriptional activator for some genes required for aerial mycelium formation.

The previous results of investigation of factor C

Factor C was detected and isolated as a protein from the culture fluid of *S. griseus* 45H, a strain that readily sporulates in liquid medium. Factor C is the first described and studied autoregulator protein that plays a key role in cellular communication and cytodifferentiation in *S. griseus*. Its molecular mass was estimated to be about 34 kDa. Factor C induced the formation of preconidia characteristic for the reproductive phase of the life cycle in submerged culture of the "test" strain *S. griseus* 52-1 (that is otherwise blocked in submerged sporulation) at concentration as low as 1 ng ml^{-1}. It was found that factor C very quickly disappeared from the liquid medium, and at about 48 to 60 h it reappeared again in higher concentration than it was added. It was also known that the cytomorphologycal effect of factor C can be increased by low concentrations of zinc and it can be purified on a zinc affinity column.

By using polyclonal and monoclonal antibodies raised against factor C, a factor C like antigen was detected in all investigated *Streptomyces* strain.

Regulation of streptomycin production in *S. griseus*

Early in growth, when the concentration of A-factor is still low, an A-factor receptor protein (ArpA) binds and represses *adpA* the key gene necessary for the onset of secondary metabolism. As the concentration of the A-factor reaches a critical level (32 nM) binds to ArpA and relieves repression of *adpA*. AdpA is a regulator of *strR* that encodes pathway-specific transcriptional activator for streptomycin-biosynthesis. StrR activates transcription of all the streptomycin-biosynthetic genes.

AIM OF THE STUDY

Extracellular regulatory molecules -called autoregulators- play a key role in controlling cellular differentiation and secondary metabolism in streptomycetes. One of them is factor C, an extracellular pleiotropic signal protein produced by *S. griseus* 45H. For further elucidation of the mechanism of regulation displayed by factor C we have planned:

1. Cloning, sequencing of the factor C gene (*facC*) and identification of its regulatory regions.
2. Test for the presence of *facC* in other *Streptomyces* strain with Southern hybridization and comparison of the amino acid sequence of the mature factor C protein with proteins in databases.
3. Establish the time of gene expression.
4. Expression of *facC* from low and high copy plasmids in wild type and mutant *S. griseus* strains on solid and liquid medium respectively and study of its effect on the morphological differentiation.
5. Search for protein(s) that interact(s) with factor C.
6. Study the effect of the expression of *facC* on streptomycin production in *S. griseus.*

METHODS

Growth conditions

Submerged cultures of *Streptomyces* strains were grown in soybean, YEME, standard minimal medium (MM) or tryptone-soya broth medium for 48 or 72 h on a rotary shaker (250 rpm) at 30 ^{o}C. R2YE was used as a solid medium and the plates were incubated at 30 ^{o}C for 5 or 6 days.

DNA manipulations

Plasmid DNA isolation from *E. coli*, cloning, gel electrophoresis, transformation of *E. coli*, and Southern hibridization was performed by standard procedures. Plasmid, chromosomal DNA and RNA isolation from *Streptomyces*, transformation of *Streptomyces*, nuclease S1 mapping was carried out according to Practical Streptomyces Genetics and Genetic Manipulation of Streptomycetes, a laboratory manual.

Construction and screening the mini gene-library of *S. griseus* 45H

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. The membrane was hybridized with the 5'-digoxigenin labelled 39-mer oligonucleotide, coding for the known N-terminal sequence of the mature factor C protein. Hybridizing bands were detected by immunostaining using a DIG nucleic acid detection KIT. On the basis of the hybridization signal, the *Sac*II fragments in the 2.8 kb region from a parallel gel were purified and cloned in *Sac*II-digested pBluescript II KS+. 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 oligonucleotide probe.

DNA sequencing and sequence analysis

Plasmids were sequenced either manually by the dideoxy chain-termination method using the T7 Sequencing Kit (Pharmacia Biotech) and [α- ^{35}S] dCTPαS 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). Sequence data were analysed with the PC/GENE, FramePlot, BLAST, DAS, Tmpred softwares.

Protein manipulation

SDS-PAGE gel electrophoresis, gel staining with Coomassie brillant blue, and Western blotting were performed by standard procedures.

Expression of His-tagged factor C in *E. coli*

In this experiment we expressed factor C without its secretional signal sequence. To generate the C-terminal His-tag we inserted six histidine codons in front of the translational stop codon of *facC*. This artifical factor C gen was ligated into the Qiagen expression plasmid pQE70, resulting pSB90. The recombinant plasmid was transformed into *E. coli* M15(pREP4). The purification of the His-tagged protein from IPTG-induced *E. coli* M15(pREP4)/pSB90 transformant cells was according to standard procedures of Qiagen.

Protein-protein interaction

For study of protein-protein interactions we used membrane fraction and supernatant of 72 h old TSB cultures of *S. griseus* 45H, *S. griseus* 52-1, *S. griseus* 52-1/pSB90. In the first two cases His-tag labelled factor C protein was added to the membrane fraction and to the supernatant separately. Incubation was carried out at 4 ^{o}C overnight. Subsequently we mixed the samples with Ni-NTA suspension and incubated them at 4 ^{o}C for 2 h. After this factor C together with its interacting protein(s) was eluted.

Measurement of antibiotic concentration

For the measurement of antibiotic production of *Streptomyces griseus* strains we used the agar diffusion method with *B. subtilis* ATCC 6633 as a test strain.

RESULTS AND DISCUSSION

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 *S. griseus* 45H

*Sac*II fragments of about 2.8 kb in size, corresponding to the signal obtained by Southern hybridization. A 5'-digoxigenin labelled 39-mer oligonucleotide, designed on the bases of the N-terminal amino acid sequence of the secreted factor C protein was used as a probe. Colony hibridization identified one clone (pBZ3) with a 2.9 kb insert that repeatedly gave a positive signal in high stringency hybridization. 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 and continued from walking primers. Both strands of DNA were sequenced at least twice. PC/GENE and Frame Plot analysis identified a 975 bp ORF, typical of *Streptomyces* with 96.9% G+C content in the third letter position and an overall G+C content of 71%. The factor C propeptide has a 38 amino acid secretion signal containing a twin-arginine motif, which directs the protein to the twin-arginine translocation (Tat) pathway. The Tat pathway translocates folded proteins across the cytoplasmic membrane which are more protected against proteases. This is very important when a regulatory protein like factor C is produced in small amount.

From the earlier results we believed factor C to be a zinc-finger-type protein. However the deduced complete amino acid sequence of the protein did not support our previous assumption. One possible explanation for our previous results is that factor C forms a complex with a zinc-finger type protein before secretion or binds zinc as a cofactor.

Analysis of the sequence of factor C by several computer programs including Tmpred and DAS suggested the presence of a transmembrane domain extending from the amino acid positions 69 to 90 with its N-terminus positioned inside the cell and its C-terminus facing outward. The added factor C rapidly disappears from the culture fluid probably taken up by cells and integrated it into the cytoplasmic membrane.

Detection of transcription start site by nuclease S1 mapping

To analyse transcription of *facC* and to localize its promoter(s) we performed nuclease S1 mapping experiment with RNA isolated from *S. griseus* 45H, using a 547 bp ^{32}P end-labelled *facC* as a probe. Cultures of *S. griseus* 45H were allowed to grow in TSBS until an OD_{550} of 0.7 and transferred to MM with mannitol as the carbon source. Cultures were incubated at 30 ^{o}C and RNA was isolated after 0, 30, 60, 120, 240 and

420 min. At the time just before the shift down no transcript could be identified. Shortly afterwards bands of 244 and 245 nt appeared, corresponding to transcripts initiated from a site 74-75 nt upstream of the start of *facC*. Transcript levels increased over time and were high more than 4 h after shift down so factor C is expressed in a life-cycle dependent manner. The promoter region was identified on the basis of transcriptional start site. The -35 sequence (TGGACA) is highly similar to the consensus -35 sequence for major promoters, recognized by the sigma factor σ^{hrdB} while the -10 promoter sequence (AACGAT) shows only low homology to the consensus -10 sequence. Deviation from the consensus sequences usually results in reduced promoter activity. It is true that factor C is expressed at low level and similarly to other autoregulatory proteins it is effective in low concentration. Addition of factor C at concentration as low as 0.5 ng ml^{-1} to the non-sporulating liquid cultures of *S. griseus* 52-1, induces the formation or spore-like bodies.

In vivo* promoter probing of *facC* in *S. coelicolor

To determine the presence and approximate location of possible promoters upstream of *facC* we used the promoter-probe vector pIJ2587, which contains the promoterless *redD* gene, encoding the transcriptional activator of the biosynthesis genes for the red-pigmented antibiotic (Red) undecylprodigiosin. In this way Red production becomes completely dependent on the promoter inserted in front of *redD*. A 547 bp DNA fragment of *facC* was amplified and cloned into pIJ2587 resulting in pIJ2587-*facC*p. Introduction of this construct into *S. coelicolor* M512 led to a low level of Red production in aerial mycelium 4-5 days after transformation while control transformants remained white. These data suggest that a promoter is present on the 547 bp fragment. Furthermore, regulation of this putative promoter appears differentiation-dependent, such that its activity correlates temporally to a time point just after the onset of aerial mycelium formation. The pink color of aerial mycelium refers to the low-level expression of *facC*.

Factor C shows no significant homology to other proteins in databases

Comparison of the amino acid sequence of the mature factor C with proteins in databases using Blast Search showed a low level of similarity to teichoic acid biosynthesis protein TagC of *B. subtilis* 168 (29 % amino acid identity with 6 gaps). Recent reports have shown that *tagC* corresponds to *dinC* and thus belongs to the SOS

regulon. The transcription of *dinC* is regulated by binding of DinR to DinR box presented in the promoter region of *dinC*. However the promoter region of *facC* lacks the consensus sequence of DinR box identified in Gram positive bacteria. Therefore the relevance of the low similarity between factor C and TagC is doubtful.

Detection of *facC* in different *Streptomyces* strains

To test the presence of *facC* in *Streptomyces* strains Southern blots of chromosomal digests of 12 strains were hybridized with a ^{32}P-labelled 860 bp DNA fragment covering about 80 % of the coding region. As expected a strong hybridization signal was observed in the lane with DNA from *S. griseus* 45H. 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 at least 87 % homology to *facC*. 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. The four *Streptomyces* strains shown to harbour a high homologue of *facC* are known to sporulate in submerged culture. This points to a possible involvement of factor C in the onset of submerged sporulation.

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

To study the phenotypic effect of factor C we expressed it in *S. griseus* 52-1 from a low and a high copy vector separately. In submerged culture the factor C non producing *S. griseus* 52-1 forms long, smooth, non-branching vegetative hyphae. Introducing the cloned *facC* gene on a low copy plasmid resulted in frequently branching hyphae containing club-like thickening at their ends, characteristic of the factor C producer *S. griseus* 45H. When *facC* gene was expressed from high copy plasmid, the transformant phenotype was identical to that of the *facC* non-producer *S. griseus* 52-1. After this we have tried to express *facC* from the very strong *ermE* promoter in *S. griseus* 52-1 but we could not get any transformants. The failure in transforming the test strain may indicate that high concentration of factor C cause complete inhibition of growth.

Expression of *facC* in *bald* mutants of *S. griseus* NRRL B-2682

We tested the effect of factor C in A factor non-producer *bald* mutant (*S. griseus* NRRL B-2682 BAFN), which strain does not form aerial mycelium on solid medium but

produces spores in the substrate mycelium. This *bald* mutant was transformed with low- and high-copy-number vectors, both harbouring the *facC* gene. Introduction of either construct restored aerial mycelium and mature spore formation. The parental strain *S. griseus* NRRL B-2682 sporulates well in submerged culture, however we failed to identify a homologue of *facC* neither in wild type nor in its A factor non producer *bald* mutant strain, analysing their chromosomal DNAs with Southern hybridization. We also tested an A factor producer *bald* mutant of *S. griseus* NRRL B-2682 (*S. griseus* NRRL B-2682 BAFP) which does not form any aerial mycelium or spore. In this case only the low copy *facC* restored the aerial mycelium and spore formation.

Restoration of aerial mycelium and spore formation to *S. griseus* NRRL B-2682 *bald* mutants might be explained by a mutation in a putative functional (but not sequence) homologue of factor C, which like factor C functions as an "upstream" regulator in the signal transduction pathway. Alternatively factor C may be a suppressor of the sporulation defect, for example by inducing silent sporulation genes. Both possibilities imply the presence of a complex regulatory network consisting of multiple signal transfer systems acting independently or having complex interaction.

Construction and expression of C-terminally His-tagged factor C in *Streptomyces* strains

We tried to overproduce the C-terminally His-tagged factor C in *S. lividans* TK24 and *S. griseus* 52-1. To our surprise the amount of protein was very low and could not be detected in the culture fluids. In the case of cell extracts just faint bands of the proper size were visible. The lack of high expression in the transformants indicates that the high expression cannot be tolerated by the cells. This is also supported by our observation that the producer *S. griseus* 45H strain could not be transformed with high copy number plasmid harbouring *facC* gene. Introducing of *facC* into the producer strain in low copy number, the resulted transformants were not viable. They died after a few days and never resulted in sporulating colonies. Therefore to achieve high expression of the His-tagged factor C we turned to the unrelated genus *E. coli*.

Construction and expression of C-terminally His-tagged factor C in *E. coli*

In the expression-system of *E. coli* we managed to overproduce C-terminally His-tagged factor C without its natural secretion signal sequence. By one-step Ni^{+}-affinity chromatography we could routinely isolate 5-10 mg of 90-95% pure His-tagged factor

C from 1 l of *E. coli* culture. We also examined the biological activity of the His-tag-labelled protein. This involved the exogenous addition of the purified protein to the liquid culture of the *S. griseus* 52-1 strain. The recombinant protein was biologically fully active in our *in vivo* test, judged by the formation of preconidia at concentrations as low as that observed with the unmanipulated native protein produced by *S. griseus* 45H (0.5 ng ml^{-1}).

Isolation of interacting protein with factor C

In isolation interacting protein(s) with factor C we used its biologically fully active C-terminally His-tagged form. The cell-free extract and the supernatant of 72 h old liquid cultures of three *Streptomyces griseus* strains were tested, respectively. However, significant difference from the protein pattern of the control (not incubated with His-tagged factor C) was found only in the sample from the cytosolic fraction of *S. griseus* 52-1, treated with His-tagged factor C. The molecular mass of the interacting proteins purified together with factor C on Ni-NTA column was about 55 kD and 110 kD respectively. One of the other two studied strains was the factor C producer *S. griseus* 45H. In this case the negative result could be explained by competition between the unmanipulated native and the recombinant factor C protein during the protein-protein interaction The third tested strain was *S. griseus* 52-1/pSB90. One possible reason for the negative result in this case is that the expression of the His-tagged protein is strictly controlled in the cell and produced at a very low amount.

Effect of *facC* on antibiotic production of *S. griseus* strains

The antibiotic production in connection with the morphological differentiation is also affected by factor C in the following way: When the studied strains do not produce any streptomycin (*S. griseus* NRRL B-2682 BAFN, *S. griseus* NRRL B-2682 BAFP), introduction of *facC* in low copies increased the antibiotic production in both strains, while high copy of *facC* increased only in *S. griseus* NRRL B-2682 BAFN. Introduction of *facC* into streptomycin producer *S. griseus* 52-1 had no effect on the antibiotic production. Consequently, factor C influences the level of the antibiotic synthesis, independently from the presence or the absence of the A factor.

SUMMARY

We cloned the gene coding for factor C and determined its nucleotide as well as amino acid sequences. The deduced complete amino acid sequence of the protein did not support our previous assumption that factor C was a zinc finger-type regulatory protein. Factor C is secreted from the cell through the Tat secretion pathway, which translocates folded proteins across the cytoplasmic membrane. Analysis of the sequence of factor C by several computer programs showed the presence of a transmembrane domain. This suggests that when factor C is taken up by other cells, it is actually integrated into the cytoplasmic membrane with its N-terminus positioned inside the cell and its C-terminus facing outward. Factor C showed no significant homology to other proteins in

databases. In contrast with previous results that showed the presence of a C factor-like antigen in all tested *Streptomyces* strains we could identify a gene with high homology to *facC* only in those *Streptomyces* strains which sporulate in submerged culture. Transcription of *facC* takes place from a single promoter located approximately 75 nt upstream of the gene in a differentiation-specific manner preceding sporulation.

Factor C plays a key role in morphological differentiation on solid-grown cultures of A factor positive and A factor negative *bald* mutants of factor C non-producing *S. griseus* NRRL B-2682. It restores their aerial mycelium formation and normal sporulation. Furthermore, factor C in low-copy number, or added to the culture fluid of the responsive strain *S. griseus* 52-1 (that does not produce factor C and is blocked in submerged sporulation) induces the formation of spore-like bodies similarly to that observed in the factor C producer *S. griseus* 45H that sporulates well in submerged culture. Restoration of aerial mycelium formation and the process of normal sporulation on solid medium to both *bald* mutants or the induction of spore-like body formation in liquid culture of *S. griseus* 52-1 transformant suggest that factor C acts as a suppressor of the aerial mycelium formation and sporulation defect. We constructed and had expressed C-terminally His-tagged factor C in *E. coli* at a relatively high level that was not attainable in *Streptomyces* strains. The recombinant protein was biologically fully active so we could use it in isolation of interacting proteins. Two proteins of about 55 kD and 110 kD were purified together with factor C from the cytosolic fraction of *S. griseus* 52-1.

The streptomycin synthesis in certain of the investigated *S. griseus* strains is affected by factor C, independently from the A factor regulatory pathway but the mode of its action is still not clear.

List of publications related to the thesis

Birkó Zs., Sümegi A., Vinnai A., van Wezel G., Szeszák F., Vitális S., Szabó P. T., Kele Z., Janáky T., Biró S. (1999) Characterization of the gene for factor C, an extracellular signal protein involved in morphological differentiation of *Streptomyces griseus*. *Microbiology* 145: 2245-2253 (IF: 2.7).

Szabó P. T., Kele Z., **Birkó Zs**., Szeszák F., Biró S., Janáky T. (1999) Identification of factor C protein from *Streptomyces griseus* by microelectrospray mass spectrometry. *Journal of Mass Spectrometry* 34: 1312-1316 (IF: 3.167).

Biró S., **Birkó ZS**., van Wezel G. P. (2000) Transcriptional and functional analysis of the gene for factor C, an extracellular signal protein involved in cytodifferentiation of *Streptomyces griseus*. *Antonie van Leeuwenhoek J. Gen. and Mol. Microbiol*. 78: 277-285 (IF: 1.423).

Birkó Zs., Schauwecker F., Pfennig F., Szeszák F., Vitális S., Keller U., Biró S. (2001) Expression and rapid one-step purification of biologically active His-tagged factor C by Ni^{+} affinity column chromatography. *FEMS Microbiology Letters* 196: 223-227 (IF: 1.673).

Other paper

Sümegi A., **Birkó Zs**., Szeszák F., Vitális S. and Biró S. (1997) A short GC-rich sequence involved in deletion formation of cloned DNA in *E. coli*. *Acta Biologica Hungarica* 48(3), 275-279 (IF: 0.219).

Posters and lectures

Birkó Zs., van Wezel G., Biró S. (1999) A C faktor gén transzkripciójának vizsgálata és expressziójának hatása a differenciálódásra. *A Magyar Biokémiai Társaság Molekuláris Biológia Szakosztályának 4. Munkaértekezlete*, Eger (poster).

Birkó Zs., van Wezel G., Biró S. (1999) Egy sejtdifferenciálódást szabályozó fehérje transzkripciójának vizsgálata és expressziójának a hatása a *Streptomyces griseus* differenciálódására. *IV. Magyar Genetikai Kongresszus*, Siófok (poster).

Birkó Zs., Szeszák F., Biró S. (2000) A C faktor génjének transzkripciós és funkcionális analízise *Streptomyces griseus*ban. *A Magyar Biokémiai Társaság Molekuláris Biológia Szakosztályának 5. Munkaértekezlete*, Sopron (lecture).

Birkó Zs. és Biró S. (2001) A differenciálódás szabályozásának vizsgálata *Streptomyces*ekben. *A Magyar Biokémiai Társaság Molekuláris Biológia Szakosztályának 6. Munkaértekezlete*, Sárospatak (poster).