Control of streptokinase gene expression in group A & C streptococci by two-component regulators

Horst Malke & Kerstin Steiner

Institute for Molecular Biology, Friedrich Schiller University Jena, D-07745 Jena, Germany

Received August 6, 2003

Background & objectives: Group A streptococci (GAS) and human isolates of group C streptococci (GCS) have the stable capacity to produce the plasminogen activator streptokinase, albeit with varying efficiency. This property is subject to control by two two-component regulatory systems, FasCAX and CovRS, which act as activator and repressor, respectively. The present work aims at balancing these opposing activities in GAS and GCS, and at clarifying the phylogenetic position of the FasA response regulator, the less understood regulator of the two systems.

Methods: The GCS strain H46A and GAS strain NZ131 were used. Escherichia coli JM 109 was used as host for plasmid construction. Streptokinase activity of various wild type and mutant strains was measured. Phylogenetic trees of streptococcal FasA homologues were established.

Results: The streptokinase activities of the GAS strain NZ131 and the GCS strain H46A were attributable to more efficient CovR repressor action in NZ131 than in H46A. The FasA activator, on the other hand, functioned about equally efficient in the two strains. Phylogenetically, FasA homologues clustered distinctly in the proposed FasA-BlpR-ComE family of streptococcal response regulators and used the LytTR domain for DNA binding.

Interpretation & conclusion: Assessing the apparent streptokinase activity of streptococcal strains require the dissection of the activities of the cov and fas systems. Although experimental evidence is still missing, FasA is closely related to a widely distributed family of streptococcal response regulators that is involved in behavioral processes, such as quorum sensing.

Key words CovR - FasA - phylogenetic tree of FasA - Streptococcus dysgalactiae subsp. equisimilis - Streptococcus pyogenes

The gene for the plasminogen activator streptokinase appears to be consistently present in all group A (GAS) and human isolates of group C streptococci (GCS). It is monocistronically expressed from a highly preserved chromosomal region in which it is interspersed among five unrelated genes transcribed in the opposite direction. Despite the omnipresence of the gene, its expression levels can vary considerably among strains. Thus, individual isolates may differ in their streptokinase-determined capacity to generate an optimized proteolytic habitat in the infected host and, consequently, may exhibit different degrees of invasiveness.

Recent investigations have begun to shed some light on the regulatory systems involved in the expression control of this important virulence factor. Regarding the characterization of cis-active sites, S1 nuclease experiments have identified the core promoter and the major transcription initiation site.

Circular permutation analysis combined with determination of the activity of nested deletions in the promoter-upstream region identified an intrinsic DNA bending locus which has a pivotal role in streptokinase (SK) gene expression. In addition, the use of reporter gene constructs in allele swap experiments between GAS and GCS strains revealed that the host genetic background dictates the SK gene expression levels. This suggested the existence of trans-acting factor(s) with
strain-specific activity that contact cis-active sites, thereby modulating streptokinase gene expression. Subsequent work carried out by a number of different laboratories identified such trans-acting factors as components of two independent two-component signal transduction systems, covRS regulating streptokinase gene expression negatively\(^5,6\) and fasCAX involved in positive regulation of this gene\(^7\). Whereas initially the action of these systems was studied independently of one another, the balance between their opposing actions in the GAS strain NZ131 to that in the GCS strain H46A was compared in the present study. The distribution and phylogenetic relationships of response regulator FasA homologues in the finished and unfinished genome sequences of various species of the genus Streptococcus were also analysed.

**Material & Methods**

*Bacterial strains and growth:* The GCS strain H46A and the GAS strains NZ131 and SF370 used in these experiments were grown in ambient air at 37 °C without agitation in brain heart infusion (BHI) broth (Difco, USA). *E. coli* JM109 was used as host for plasmid constructions and was grown at 37 °C in rotary flasks in standard Luria broth (LB). If appropriate, antibiotics were used at the following concentrations: chloramphenicol, 3 µg/ml for streptococci and 100 µg/ml for *E. coli*; erythromycin, 2.5 µg/ml for streptococci and 200 µg/ml for *E. coli*; kanamycin, 100 µg/ml for streptococci and 50 µg/ml for *E. coli*; spectinomycin, 100 µg/ml for both streptococci and *E. coli*.

*Construction of cov and fas plasmids:* General nucleic acids techniques have been described previously\(^1\). Oligonucleotide primers were used in polymerase chain reactions (PCR) for the construction of recombinant plasmids (Table I). Plasmids were electrotransformed into strains H46A or NZ131 to insertionally inactivate the specified cov and fas genes, or to complement the mutations. The construction and characteristics of the resultant strain H46A and strain NZ131 derivatives are described in Table II.

*Streptokinase activity assay:* The plasminogen activation assay on microtiter plates\(^14\) was used to measure streptokinase activity in BHI culture supernatant fluids of the various wild type and mutant strains. The release of para-nitroaniline from the chromogenic substrate H-D-valyl-leucyl-lysine p-nitroaniline (Sigma, USA) was measured at OD\(_{405}\) over time, and activity rates were calculated from the linear parts of absorbance versus time plots. Standard streptokinase was procured from Sigma, with 1 unit (U) being defined as the protein activity capable of liquifying a standard clot of fibrinogen, plasminogen and thrombin at pH 7.5 and at 37 °C in 10 min.

**Table I.** PCR primers used for the construction of cov and fas plasmids

<table>
<thead>
<tr>
<th>Primer pair</th>
<th>Sequence(^a)</th>
<th>Template DNA</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>covR</em>F1</td>
<td>ccggaattcCAAGGGTTGTTTGTAGAATA</td>
<td>SF370</td>
<td>pMcovR(_{SF}) construction</td>
</tr>
<tr>
<td><em>covR</em>R1</td>
<td>ccggaattcATGACTTATTTTCTCAC</td>
<td>SF370</td>
<td>pMcovR(_{SF}) construction</td>
</tr>
<tr>
<td><em>covR</em>R1</td>
<td>ccggaattcCAAGGGTTGTTTGTAGAATA</td>
<td>NZ131</td>
<td>pVAcovRS(_{NZ}) construction</td>
</tr>
<tr>
<td><em>covSR</em>F1</td>
<td>cgcctgcagCTTAAGCTACTCTAATCTTC</td>
<td>NZ131</td>
<td>pFC1 construction</td>
</tr>
<tr>
<td><em>fasA</em>F1</td>
<td>cgcggatatCCAGCAGACAGCATAGAATC</td>
<td>HF46A</td>
<td>pFW5fasA1 and pFW15fasA1 construction</td>
</tr>
<tr>
<td><em>fasA</em>R2</td>
<td>cgcggatatCGGTGCTACTGCTTTGAATCAG</td>
<td>HF46A</td>
<td>pFW14fasA2 construction</td>
</tr>
<tr>
<td><em>fasX</em>R3</td>
<td>cgcggatatCCAGCAGACAGCATAGAATC</td>
<td>SF370</td>
<td>pFasAX(_{SF}) construction</td>
</tr>
</tbody>
</table>

\(^a\) Lower case letters indicate sequences added to facilitate cloning
Sequence analysis of streptococcal FasA homologues: Similarity searches with varying parameters against publicly available databases (www.ncbi.nlm.nih.gov; www.sanger.ac.uk; www.tigr.org) containing finished and unfinished genomic sequences of Streptococcus species were performed using the TBLASTN program (www.ncbi.nlm.nih.gov) with Streptococcus dysgalactiae subsp. equisimilis FasA as a query. The multiple alignment of the retrieved sequences and their phylogenetic trees were generated by distance matrix analysis using the AllAll program (cbrg.inf.ethz.ch/Server/AllAll.html). The consensus sequence of the DNA-binding domain derived from a multiple alignment of 21 streptococcal FasA homologues was determined using the SeqLogo program in the implementation by SE Brenner. Secondary structure predictions were performed using the programs of the PredictProtein server (cubic.bioc.cam.ac.edu/predictprotein).

Results

Balancing the action of the cov and fas systems in GCS and GAS: Streptokinase activities of cell free...
culture fluids obtained from saturated BHI cultures of strains H46A and NZ131 were about 80 and 3 U/ml, respectively. Since wild type H46A possesses a naturally acquired amber mutation at codon position 102 of covR and so actually proved to be a derepressed mutant for streptokinase production, the question was whether or not the great difference in the streptokinase activities between H46A and NZ131 reflected solely the state of their covR alleles. Since the streptokinase alleles, skc and ska, respectively, of the two strains were also subject to positive control by the fas system, a comparative mutational approach was used to analyse the differential contributions of the two regulatory systems to streptokinase production. This approach involved the creation of all possible combinations of wild type and mutant covR and fasA alleles in the two strains, including complementation of mutant alleles to rule out possible effects of polar mutations (Table II).

The results of the streptokinase activity assays of the various strains in cultures with comparable cell density are given in Fig. 1. First, restoration of CovR repressor activity in H46A by introduction of the covR<sub>NZ131</sub> allele decreased its streptokinase activity to approximately 50 per cent. Compared to the low streptokinase activity of wild type NZ131, H46A (Cov<sup>+</sup> Fas<sup>+</sup>) released 13 times more streptokinase than NZ131. Inactivation of the CovR repressor of NZ131 resulted in a greater than 10-fold increase of its streptokinase activity which, however, was still approximately 50 per cent lower than that of H46A (Cov<sup>+</sup> Fas<sup>+</sup>). This difference might be attributable to a slightly lower stimulatory effect of the fas system in NZ131, as suggested by the streptokinase activity ratios of Fas<sup>+</sup> versus Fas<sup>-</sup> strains in a Cov<sup>-</sup> background (1.5 in NZ131 vs. 2.0 in H46A). Taken together, these results showed that the opposing activities of the cov and fas systems were about equal in H46A whereas in NZ131 the repressive CovR activity excelled the stimulatory FasA activity by a factor of about 9. In the absence of both repression and activation, i.e., in a Cov<sup>-</sup> Fas<sup>-</sup> background, the constitutive streptokinase activities did not differ substantially between the two strains, and, expectedly, both strains showed their lowest activities in a Cov<sup>-</sup> Fas<sup>-</sup> background (Fig. 1).

**Evolutionary relationships among streptococcal FasA response regulators:** The starting point of recent investigations that led to the identification of the fas regulatory system was the observation that there was one region in the S. pyogenes SF370 genome that exhibited similarity values >34 per cent to the Staphylococcus aureus accessory gene regulator AgrAC and the S. pneumoniae competence regulatory system ComDE, both of which were involved in quorum sensing. Recently, sequence analysis of bacterial genomes has placed the AgrA and ComE response regulators in a family of transcriptional regulators that bind DNA with a novel domain, designated LytTR, which is distinct from the classical DNA-binding helix-turn-helix or winged helix domain of the overwhelming majority of the response regulators (including CovR) of the bacterial two-component signal transduction systems. It was of considerable interest, therefore, to study the distribution and phylogenetic relationships of streptococcal FasA homologues.
TBLASTN searches of the databases with FasA<sub>Strep</sub> as a query retrieved, from 12 Streptococcus species, a set of 21 proteins (as of September 01, 2002) that showed >54 per cent sequence similarity (random expectation value, E<7x10<sup>-29</sup>) to the query sequence. The indicated threshold values were observed for sequence similarity between FasA<sub>Strep</sub> and AgrA. The unrooted phylogenetic tree (Fig. 2) based on the distance matrices between the sequences had a fitting index of 0.57, indicating that it approximated the original distances between each pair of nodes in a satisfactory manner. The tree also showed that the PAM (per cent accepted mutations) distance between AgrA and any other streptococcal sequence was greater than the divergence between the most distantly related streptococcal sequences. The most interesting topological feature of the tree was its branching pattern, which showed that the streptococcal sequences form 3 distinct clusters,
referred to as the BlpR, ComE, and FasA cluster, after the *S. pneumoniae* BlpR and ComE, and the *S. pyogenes* FasA response regulators, involved in bacteriocin biosynthesis and virulence factor regulation (BlpR), competence development (ComE), and the regulation of the production of streptokinase, streptolysin S and fibronecin/fibrinogen-binding proteins (FasA). The *fasX* gene, thought to encode a non-translated RNA functioning as the terminal effector of the *fasCAX* operon had no homologues in the operons of the ComE and BlpR clusters, underscoring the separate position of the FasA regulators.

The distribution of members of the ComE cluster appeared to be restricted to the transformable species of the mitis-anginosus group; members of the FasA cluster occurred preferentially in the pyogenic group, and the proteins of the BlpR cluster seemed to be most widely distributed, with occurrences in species of the pyogenic, mitis and mutans group. It was interesting to note, however, that as a member of the pyogenic group, *S. agalactiae* did not appear to contain FasA, as indicated by the complete genome sequences of two different serotypes, both of which contained members of the BlpR cluster. Similarly interesting was the observation, that all finished and unfinished genomic sequences (www.sanger.ac.uk/Projects/S_pyogenes) of different *S. pyogenes* strains contained a FasA protein but not necessarily a BlpR protein, as indicated by its absence from the genome of the M1 strain SF370. The BlpR-like protein designated Spy_2 in Fig. 2 is identical to SilA, a response regulator in the *sil* locus which has been found to be responsible for the invasive properties of JS95, an M-type 14 GAS strain isolated from a case of necrotizing fasciitis.

---

**Fig. 3.** Consensus sequence of the DNA-binding domain of the streptococcal FasA-BlpR-ComE family of response regulators, based on a multiple alignment of these domains from the proteins included in Fig. 2. Residue number 1 corresponds to Glu138 of the *S. dysgalactiae* subsp. *equisimilis* H46A FasA protein (AY075107), designated Sdy_1 in Fig. 2. Triangles point to the positions of highly conserved positively charged residues.
Concerning the domain composition of the FasA-BlpR-ComE response regulators, the N-terminal halves of all proteins included in the phylogenetic tree (Fig. 2) contained the response regulator receiver domain archetypically represented by the CheY domain of *E. coli* (NCBI protein database at www.ncbi.nlm.nih.gov/Structure/cdd/). As a peculiarity, the CheY-homologous domain of the RgfA protein (designated Sag_1 in Fig. 2) of *S. agalactiae* strain O90R[^21] was N-terminally truncated, lacking about 36 amino acid residues when compared to the other FasA-BlpR-ComE regulators. This deletion was not present in the full-length BlpR-homologues (Sag_2) of the two completely sequenced *S. agalactiae* strains, NEM316[^16] and 2603V/R[^17]. A more extensive deletion was seen in the BlpS protein from *S. pneumoniae* (acc. number, AAK04633; excluded from the tree in Fig. 2) which lacked the receiver domain completely.

The output domain of the bacterial response regulators was typically a DNA-binding domain contained in their C-terminal portions. Sequence analysis of the C-terminal halves of the streptococcal FasA-BlpR-ComE proteins generated the consensus sequence presented as a sequence logo in Fig. 3. It showed that among the amino acids with the highest degree of conservation (indicated by the height of the letters) were many of the positively charged residues, the position of which also tended to coincide with high columns, reflecting the statistical importance of the given position. The possibility existed that some of these positions were directly involved in DNA binding. Secondary structure prediction made it likely that the DNA-binding domain of the FasA-BlpR-ComE family contained 3 beta-strands, 2 alpha-helices, a beta strand, and an additional alpha-helix, in that order (Fig. 3). The order of these structures was similar to that predicted for the LytTR domain[^15], providing corroborative evidence for the novelty of this domain.

**Discussion**

Dissecting the two opposing regulatory systems involved in the modulation of streptokinase activity by mutation enabled us to balance their activity and provide possible explanations for different streptokinase activities observed in field strains. Several reasons may be advanced to explain the different strength of the CovR repressor in H46A and NZ131. First, there existed sequence differences between the two strains in their wider promoter regions, which might influence CovR binding. The proposed short consensus sequence (ATTARA) for CovR binding to the hasA promoter[^22] was seen only once relatively far upstream of the streptokinase core promoter in both H46A and NZ131. However, there were 4 more ATTA tetranucleotides, in which the thymin pair was found to be necessary for CovR binding[^22] in NZ131 than in H46A. Thus, the ska promoter region of NZ131 might present a better target for CovR repressor action than the corresponding H46A region. However, differential regulation could also be caused by different affinity of CovR to individual binding sites, or by different expression levels of CovR which autoregulates its own gene[^5].

It is noteworthy that H46A contains a naturally acquired nonsense mutation in *covR* and thus adds to the repertoire of strains that have been shown to carry spontaneous *covRS* mutations[^23]. At present, it is not clear whether this locus is hypermutable or mutations occur at random but there is strong *in vivo* selection for the loss of *covR* expression. Either possibility is intriguing, particularly in view of the fact that CovR is a global regulator that targets several virulence genes[^1,5,6]. In either case, the size of a more virulent subpopulation *in vivo* would thus depend on the point in time where mutation occurs in the course of infection and thus presumably contribute to the outcome of streptococcal infections. The potentiation of the mode of pathogenicity would appear to be most severe in strains that, as *covRS* wild types, carry efficiently repressed virulence genes. Given the high frequency of *covRS* mutations, vigilance should be used when clinical strains are established from single colony isolates and data generated with them are extrapolated to a clinical situation potentially determined by a diverse population of bacteria.

Compared to what is known about the *covRS* system, our knowledge of the *fasCAX* system is much less advanced. Its growth phase-dependent control activity over multiple streptococcal virulence factors, which appears to be influenced by the nutritional
environment, makes it an important factor involved in the pathogenesis of streptococcal disease. Although initial investigations failed to find evidence for its involvement in quorum sensing, encouraged by a way of reasoning known as guilt by association, one should keep an open mind regarding its participation in complex behavioral responses. It will be interesting to find out whether the distinct clustering of FasA within the well-defined FasA-BlpR-ComE family of streptococcal response regulators is reflected by a similarly distinct range of target genes.

Acknowledgment

Authors acknowledge the financial support from the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

References


*Reprint requests:* Dr Horst Malke, FSU Jena, Institute for Molecular Biology, Winzerlaer Strasse 10, D-07745 Jena, Germany

e-mail: hmalke@imb-jena.de