

# Structure of a group C streptococcal protein that binds to fibrinogen, albumin and immunoglobulin G via overlapping modules

Susanne R. TALAY,\* Melanie P. GRAMMEL and Gursharan S. CHHATWAL

Department of Microbiology, Technical University Braunschweig/GBF-National Research Centre for Biotechnology, D-38124 Braunschweig, Germany

Pathogenic streptococci express surface proteins that bind to host serum proteins. A novel multiple-ligand-binding protein has now been identified in a species belonging to serotype C streptococci. This protein binds to fibrinogen, albumin and IgG and was therefore designated FAI protein. The structure of the *fai* gene has been determined, and deletion analysis and expression of FAI fusion polypeptides revealed that the binding domain for fibrinogen and IgG is located within the non-repetitive N-terminal half of the protein. A 93-amino acid peptide retained the ability to bind both proteins, whereas a 56-amino acid subpeptide only bound fibrinogen. IgG-binding activity required the complete fibrinogen-binding domain and an additional 37 amino acids C-terminal to it, and albumin-binding

activity was only obtained with a polypeptide reflecting the complete surface-exposed region of FAI protein indicating that the binding sites for each ligand were located on overlapping modules. Signal sequence, C repeat region and C-terminus revealed high homology to group A streptococcal M proteins whereas the N-terminal region containing the fibrinogen/IgG-binding domains is completely different and exhibits no similarity to any other previously characterized protein. Thus FAI protein exhibits a framework structure that might have evolved in group C streptococci via fusion of unrelated sequences, thereby generating an albumin-binding domain in the functional context of multiple-ligand-binding activity.

## INTRODUCTION

Pathogenic streptococci have evolved surface proteins that enable them to bind host factors. Such proteins are considered as virulence factors since they can promote adherence, prevent phagocytosis and allow evasion of immune defence mechanisms by masking the bacteria with host proteins [1–3]. Human fibrinogen is bound by M protein, the antiphagocytic molecule on the surface of group A streptococci [4]. M or M-like genes of *Streptococcus pyogenes* also code for receptors for immunoglobulins and bind to IgG and IgA in a non-immune manner via the Fc region of the molecule [5–8]. Group G and C streptococci express Protein G, an IgG/Fc receptor which has evolved convergently with other IgG/Fc receptor proteins [9]. Protein L, an IgG-binding protein of *Peptostreptococcus magnus*, binds to the immunoglobulin in a unique manner by recognizing the  $\kappa$  light chains of the protein [10].

Some of the proteins described exhibit dual binding function in the way that they interact with more than one mammalian protein. Protein G, Protein L and Protein H are structurally distinct molecules that originate from different species but share the ability not only to bind IgG but also human serum albumin [11–13]. On the other hand, streptococci often express two or more distinct proteins that appear to fulfil similar functions to Fc receptors and fibronectin-binding proteins of *S. pyogenes* [14–16]. However, in group C streptococci Protein G and Protein G-related molecules are the only IgG/albumin receptors that have so far been characterized at the molecular level [17–20].

In this study, a novel group C streptococcal IgG- and albumin-binding protein with an additional ability to bind fibrinogen has

been identified and designated FAI (fibrinogen/albumin/IgG receptor). Sequence analysis revealed segments that are highly homologous to *S. pyogenes* M proteins, interrupted by a foreign module with fibrinogen- and IgG-binding activity. Moreover, the M-like region was necessary but not sufficient for albumin-binding activity, supporting the idea that recombination of unrelated sequences allows the generation of successful and multivalent ligand-binding proteins.

## EXPERIMENTAL

### Bacterial strains, bacteriophage and plasmids

The group C streptococcal isolate C20 was obtained from the Institut für Mikrobiologie und Tierseuchen der Tierärztlichen Hochschule, Hannover, Germany, and has been isolated from a horse. The strain was identified as belonging to Lancefield group C by using a Slidex Strepto-Kit (api bioMérieux, Nürtingen, Germany). The German Culture Collection, Germany, also identified the strain as a typical example of group C streptococci. *Escherichia coli* strains HB101 and JM101 [21] were used for plasmid propagation and expression of fusion proteins respectively. *E. coli* strains LE392 and P2392, the  $\lambda$  vector EMBL3 and packaging extracts were purchased from Stratagene (Heidelberg, Germany). *E. coli* plasmid pJLA601 has been used as a temperature-inducible expression system [22], and the pGEX2T glutathione S-transferase (GST) fusion and expression system was obtained from Pharmacia (Uppsala, Sweden). Streptococci were grown in Todd Hewitt broth (Oxoid, Basingstoke, Hants., U.K.) and *E. coli* in Luria–Bertoni broth. Phage propagation was carried out as described previously [21].

Abbreviation used: GST, glutathione S-transferase.

\* To whom correspondence should be addressed.

The nucleotide sequence for FAI protein will appear in EMBL/GenBank/DDBJ Nucleotide Sequences Databases under the accession number X84989.

### Construction of the streptococcal library and screening of recombinant plaques

Chromosomal DNA from the clinical group C streptococcal isolate C20 was purified, partially digested with *Sau3A* and size-fractionated on a NaCl gradient before ligation to *Bam*HI-digested  $\lambda$  EMBL3 vector arms. Ligation and packaging were performed according to the manufacturer's (Stratagene) protocol. The library was plated and screened by overlaying recombinant plaques with nitrocellulose membranes (Bio-Rad, Munich, Germany) to absorb proteins. Filters were blocked in 0.1 M PBS, pH 7.5, containing 10% skimmed milk followed by incubation in PBS containing 10  $\mu$ g/ml peroxidase-conjugated human fibrinogen. After being washed in PBS, filters were developed in PBS containing 1 mg/ml 4-chloro-1-naphthol and 1% H<sub>2</sub>O<sub>2</sub>.

### DNA preparations, cloning techniques and sequence determination

Recombinant  $\lambda$  DNA was purified from phage lysates as described by Sambrook et al. [21]. Isolation of plasmid DNA was achieved by the use of the QIAwell Plasmid Kit (Qiagen, Hilden, Germany) and PCR fragments were purified from the reaction mixture by use of Quick Spin Tubes (Qiagen) before subsequent cloning or sequencing. DNA restriction endonuclease digestion, ligation and chemical transformation were performed by standard techniques [21]. Nucleotide sequencing was carried out by the dideoxy chain-termination method [23] using fluorescent dye terminators (Applied Biosystems Inc., Weiterstadt, Germany) in cycle sequencing reactions. Samples were electrophoresed, detected and analysed on a DNA Sequencer (Applied Biosystems). Initial sequencing reactions were performed with appropriate vector primers, and subsequent sequencing was carried out with primers designed on the basis of the sequencing data already obtained.

### PCR and construction of fusion-protein-expressing plasmids

PCR was performed on a thermocycler (Hyaid, Teddington, Surrey, U.K.) with 100 ng of chromosomal DNA using standard protocols for amplification. Oligonucleotide primers were designed to generate DNA fragments with flanking *Bam*HI and *Eco*RI restriction sites that allow in-frame cloning of selected *fai* regions C-terminal to the GST coding sequence of the pGEX2T expression system. Primers for the respective clones were A/H (pMG3), F/H (pMG4), A/K (pMG6), D/K (pMG10), C/K (pMG12), C/L (pMG13), C/M (pMG14) and D/I (pMG15). The sequences of coding strand primers were (A) 5'-GACAGT-AGGATCCAGTGAATATTCGG-3', (C) 5'-CTAAACTTAA-AGGATCCGCTGAGGCGC-3', (D) 5'-CTTACAGGATCCA-ACGAGGCTTTATCAAC-3' and (F) 5'-GAAAGAAATAAA-GGATCCGAAGCAAGTCG-3'. The sequences of non-coding strand primers were (H) 5'-TCCAGGGAATTCTGGTTTT-CTGG-3', (I) 5'-AAGTTTAGCGAATTCTGCATCAACTGC-3', (K) 5'-GGTACGACTGAATTCAGAACTTTATTT-3', (L) 5'-TGATATGATCGAATTCCTTTAATTTGCC-3' and (M) 5'-TTATCTAGAATTCCTTGATAAAGCCTCG-3'.

### Overexpression and purification of recombinant FAI fusion proteins

Cells containing pMG constructs were grown to  $A_{600} \sim 1.0$  at 37 °C, induced by adding 1 vol. of fresh medium containing 2 mM isopropyl  $\beta$ -D-thiogalactopyranoside, and further incubated with vigorous shaking at 30 °C for 6 h. Cells were harvested by centrifugation, washed in PBS, pH 7.5, resuspended in 0.2 vol. of PBS, and lysed via sonication. The lysate was cleared by centrifugation and 10 ml of sonic lysate was applied to a 4 ml bed

vol. glutathione-Sepharose column (Pharmacia). The column was washed with 10 bed vol. of PBS before elution of bound fusion protein with 50 mM Tris/HCl, pH 8.0, containing 10 mM glutathione. Eluted fractions were checked for purity by SDS/PAGE and Coomassie Blue staining of the gel. The procedure yielded between 5 and 10 mg of purified fusion protein per litre of culture.

### Generation of antibodies and F(ab)<sub>2</sub> fragments

Antibodies against recombinant FAI fusion proteins were raised in rabbits. The immunization was carried out subcutaneously with 500  $\mu$ g of protein in 500  $\mu$ l of physiological saline and an equal volume of Freund's complete adjuvant. A booster was given after 14 days with the same amount of protein but incomplete adjuvant. The rabbits were bled 2 weeks after the booster. IgG fractions were isolated from serum by affinity chromatography on Protein A-agarose. The IgG fraction was depleted from anti-GST antibodies by affinity chromatography on GST-agarose. F(ab)<sub>2</sub> fragments were generated via pepsin digestion and separated from Fc fragments by standard procedures [24]. F(ab)<sub>2</sub> fragments were labelled with horseradish peroxidase (Sigma) according to a standard protocol [24].

### Proteins and labelling of proteins

All proteins were of human origin, unless otherwise stated. Fibrinogen was purchased from Kabi (Munich, Germany), and contaminating fibronectin was removed by preabsorption on gelatin-agarose. Albumin was from Serva (Heidelberg, Germany) and IgG was isolated from serum on Protein A-agarose. Peroxidase-labelled goat anti-human IgG antibodies and peroxidase-labelled rabbit IgG were from Nordic and Dako (Hamburg, Germany) respectively. Fibrinogen, albumin and immunoglobulins were labelled with horseradish peroxidase (Sigma) according to a standard protocol [24].

### Other methods

SDS/PAGE and Western blotting were performed as described previously [14], and dot-blot analysis was carried out using nitrocellulose membranes with 1–0.1  $\mu$ g of protein absorbed on the filter. Blocking, filter washing and development were as in the plaque lift method described above. Computer database searches and sequence homology determination were performed on a VAX system using our own programs.

## RESULTS

### Detection of FAI protein in a group C streptococcus

A protein with multiple-ligand-binding activity was detected in the supernatant of group C streptococcal cultures. First, secreted proteins were concentrated by (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> precipitation. Western-blot analysis of this material revealed fibrinogen-, albumin- and IgG-binding activity associated with a single protein band of apparent molecular mass 46 kDa. We assumed that one protein was responsible for this multiple-binding activity and thus the protein was termed FAI protein. To verify this and to determine its structural features and its functional epitopes we attempted to clone its gene and express it in *E. coli*.

### Cloning and sequencing of *fai* gene

The  $\lambda$  replacement vector EMBL3 was used to construct a genomic library of the group C streptococcal isolate C20. Recombinant phages were screened via a plaque lift method for

```

AGACAGCTCTGGAATGAGACCTCTCTAAAAAATCTAGGCAAAAAGAAAATAAGGAGAACAA 60
M A R K N K T K Q Y S I R K L S V G A A 20
ATGGCTAGAAGAATAAAACAAAACAGTACTCAATAGAAAATTGAGTGTAGGCGCGGCT 120
S V L V A T G L I G G A C T V S A S E Y S 40
TCAGTCTTGGTTGCGACAGTCTTATTGGTGGAGCGACAGTAAAGTGCAGTGAATATTCG 180
D T I T T T A1 G T S N G H I T S S G I G S 60
GACACCATTACTACTGGTGTACAAGTAATGGTCTATATTACATCTTCTGGTATTGGAAGT 240
S T I V G G L A K P A S N I I I A2 G T T 80
AGTACAATTGTCGGTGGATTAGCGAAACCTGCAAGCAATATTATTATTGGAGGAAGTACT 300
A G T I G S S G T I G G T I V S G Y S B1 100
CGGGGCACTATTGGTCTAGCGGTACTGAGGTGGCACTATTGTTAGTGGTTATTCTGGA 360
S N T I A S G T V A G Q A G T T H D B2 S 120
AGCAATACCATTGCAAGTGAAGTGTAGCTGGTCAAGCTGGAAGTACTCATGATGGTAGC 420
N T I A S G V E E N K D L L K E V A E L 140
AACACTATTGCTAGTGGTGTGTAAGAAAATAAAGTTTACTAAAAGAAGTTGCTGAACCT 480
IgG+
T E A N E A L S T I L D K T I E E K I K 160
ACAGAAGCGAAGCGGCTTATCAACAAATTTAGATAAAACAATCGAGGAAAAGATCAAG 540
S E Y K Y K E E I G K L K E L D H I N N 180
TCAGAGTACAAGTACAAGAAGAATTTGGCAAAATTAAGAAGTCTGATCATATCAACAA 600
N L L G N A K D Q L D K L T T E K E S L 200
AATCTTTTGGGAAATGCTAAAGATCAATTAGATAAACTTACTACAGAAAAGAAAGTTTG 680
S K D K E A L D E R N K C1 L E A S R T R 220
TCTAAGATAAAGAAGCATTAGATGAAAGAAATAAAGTTCTGGAAGCAAGTCCACCGCT 720
T N R D L E A A R D A K K A T E A E L A 240
ACAACCGTGATTAGAGGCGAGCTCGTGACGCTAAGAAAGCCACAGAAGCAGAATTAGCT 780
E T N A K V D K L E E E K Q C2 L E A S R 260
GAAAACAACCGTAAAGTTGACAAGCTTGAAGAAGAAAACAAGTCTTGAAGCCAGCGCT 840
K R T N R D L E A A R D A K K A T E A E 280
AAACGTACAACCGTGATTAGAGGCGAGCTCGTGACGCTAAGAAAGCCACAGAAGCAGAA 900
L A K A N E L N Q C3 L E A S R T R T N R 300
TTGGCTAAAGCTAATGAGCTAAACCAATCTTAGAAGCAAGCCGACACGACCCACCGCT 960
D L E A A R D A K K A V D A E L A K L K 320
GATTTAGAGGCGAGCTCGTGACGCTAAGAAAGCAGTGTGATGAGCAACTTGCTAAACTTAA 1020
A E A E G A L K E Q L L A K Q A Q E I E K L 340
GCTGAAGCTAAGGCGCTTAAAGAGCAATTTGGCTAAACAAAGCTCAAGAAATGAAAAACT 1080
K E S K E K A P E A T Q T P E K P E V P 360
AAAGAATCAAAGAAAAAGCACCAGAAGCAACTCAAACTCCAGAAAACCCAGAAGTACTCT 1140
G K P S M P W T G L T P A T P I A A1b K D R 380
GGTAAACCAAGCATGCTTGGACAGGATTAAACACCTGCTACACCAATCGCTAAAGACAGA 1200
K K T D V K P A A K A N M V P T D V K K 400
AAGAARACTGATGTTAAACCTCCGCTAAGCAACCAACATGGTGCCTACTGACGTTAAGAAA 1260
D E K K L P S T G E T V N P F F T A A G 420
GATGAGAAGAACTTCCATCAACTGGCGAAACTGTGAACCCATTCTTACAGCAGCAGGA 1320
M A G M A T A G V V A V G K R K E E N * 439
ATGGCAGGTATGGCAACAGCTGCTGTAGTGTGGAAAACGCAAGAAGAAAATCAA 1380

```

**Figure 1** Nucleotide sequence and derived amino acid sequence of the group C streptococcal *fai* gene

The vertical arrow indicates the signal sequence cleavage site. The first amino acid of each of the A, B and C repeats is boxed and the type of repeat is indicated above. The start and end of the fibrinogen- (fgn), albumin- (alb) and IgG-binding regions are noted below the sequence. A typical ribosomal-binding site and the membrane anchor consensus sequence are underlined.

expression of proteins that have the ability to bind biotinylated fibrinogen. All positive clones also reacted with IgG and albumin, indicating the presence of the *fai* gene in all constructs. One of these clones was chosen for further experiments. First, the complete 12.9 kb insert was excised from the  $\lambda$  DNA as a *SalI* fragment and cloned into the plasmid vector pJLA601. *E. coli* cells harbouring the resulting plasmid pMG1 expressed multiple-ligand-binding activity under non-induced conditions of the vector promoters, indicating insert-encoded promoter activity. A restriction map of pMG1 was created to allow the construction of subsequent deletion clones. Deletion of an 11.4 kb *SphI*-*SalI* fragment created pMG13, the smallest subclone in which expression of multiple plasma-binding activity was maintained under non-induced conditions of the vector promoters. First,

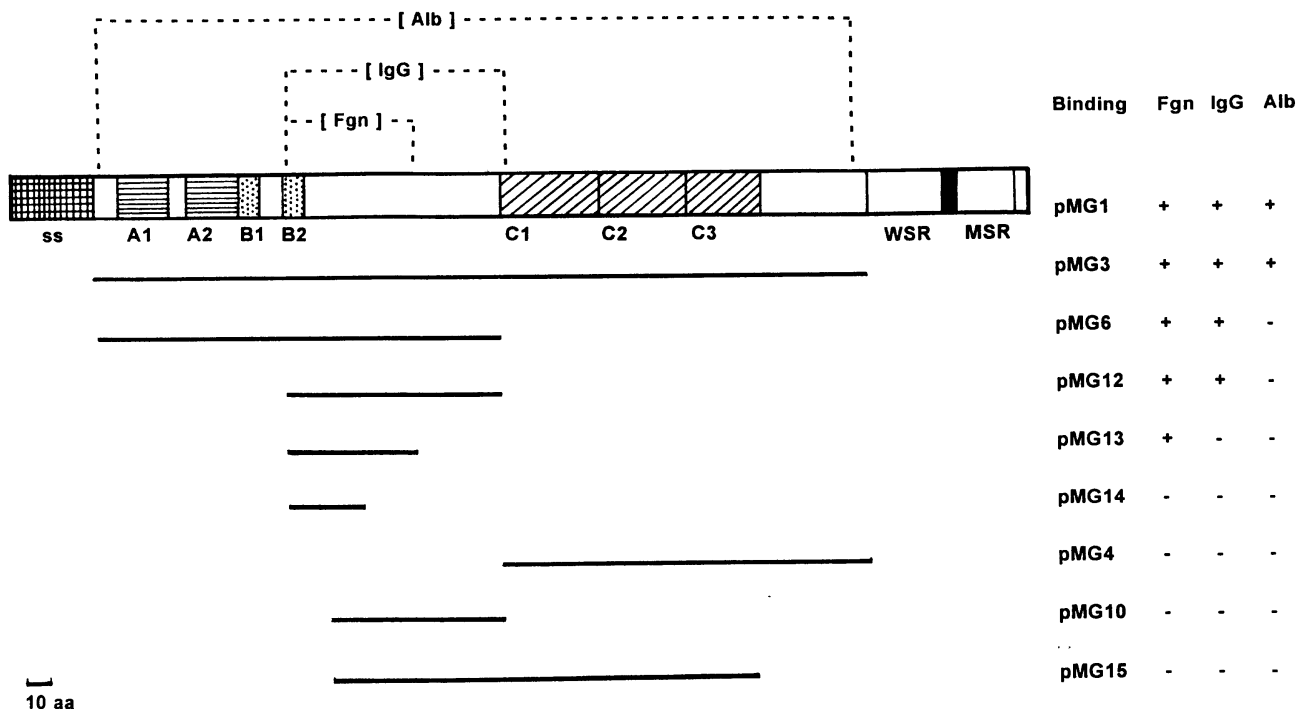
both strands of the 1.5 kb insert were completely sequenced. Since an internal region of 110 bp within a repetitive region appeared to be unstable because of high frequency of recombination in *E. coli*, the insert was amplified from streptococcal DNA and the resulting PCR product was directly sequenced. Analysis of the confirmed sequence revealed the presence of only one open reading frame of 1317 nucleotides which constitutes the *fai* gene (Figure 1).

### Structure of FAI protein

The open reading frame of 1317 nucleotides can be translated into a 439-amino acid residue comprising preprotein with a calculated molecular mass of 46.5 kDa. This deduced primary sequence of FAI protein exhibits typical features of streptococcal surface proteins: it starts with an N-terminal signal sequence of 36 amino acid residues and ends with an anchor cleavage signature sequence followed by a membrane-spanning region and a charged tail (Figure 1). Both signal sequence and membrane-spanning region are of hydrophobic character, whereas the central part of the FAI protein, which is expected to be surface exposed, is mainly hydrophilic. Inspection of the primary sequence of the processed form of the FAI protein revealed the presence of three distinct repeat regions consisting of two or three direct repeats. Two A and two B repeats are located within the N-terminal half of the molecule, whereas three C repeats are located within the C-terminal half of the protein. The two A repeats are 23 amino acids long, share 50% identical residues and are separated by a short spacer region. The A repeats are immediately followed by the B repeats comprising eight residues which are completely identical and are also separated by a short spacer region. A long non-repetitive spacer is followed by the C repeats which comprise 42, 35 and 39 amino acid residues respectively and share between 70 and 80% identical residues. Secondary-structure prediction revealed an extended structure interrupted by turns for the region containing A and B repeats, whereas the non-repetitive region and the C repeats are predicted to be  $\alpha$ -helical.

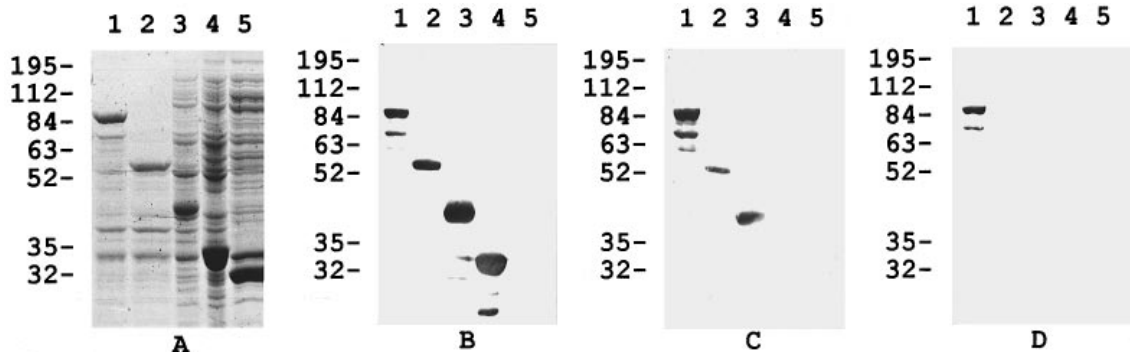
### Expression of fusion proteins and localization of ligand-binding domains

In order to determine the regions responsible for binding fibrinogen, IgG and albumin, distinct parts of the FAI molecule were expressed as in-frame fusions by the pGEX2T expression system. Fusion proteins were analysed in stained gels after SDS/PAGE, and their binding activities tested by Western blot. Figure 2 summarizes the range of fusion proteins, and their ligand-binding activity is expressed by the respective constructs. Similar results were obtained with dot-blot experiments (results not shown). As expected, deletion of the signal sequence as well as the wall- and membrane-associated parts of the FAI molecule did not influence the binding pattern of the protein expressed by pMG3 (Figure 3) compared with that of pMG1. Fusion proteins expressed by pMG6 and pMG4 indicate that the binding domains for fibrinogen and IgG were both located within the N-terminal half of the molecule, whereas neither the N- nor the C-terminal half alone of FAI protein were sufficient for albumin binding. C-Terminal-truncated forms of the fibrinogen/IgG-binding protein expressed by pMG6 (Figure 3) allowed a more precise discrimination between the fibrinogen- and the IgG-binding region. The protein expressed by pMG12 reflected 94 amino acids of FAI protein including one B repeat and the non-repetitive region up to the C repeats. This protein was the smallest molecule in which both binding activities were maintained (Figure 3), since further C-terminal (pMG13, pMG14) and N-terminal (pMG10)



**Figure 2** Schematic topology of FAI protein and its overlapping binding regions

Truncated derivatives were expressed as GST fusion proteins by the constructs given on the right and the individual binding activity of each protein is indicated. The abbreviations used in this Figure are: SS, signal sequence; A1 to C3, A, B and C repeats; WSR, wall-spanning region; MSR, membrane-spanning region; Fgn, fibrinogen; Alb, human serum albumin.



**Figure 3** Western-blot analysis of ligand-binding activities of selected FAI fusion proteins

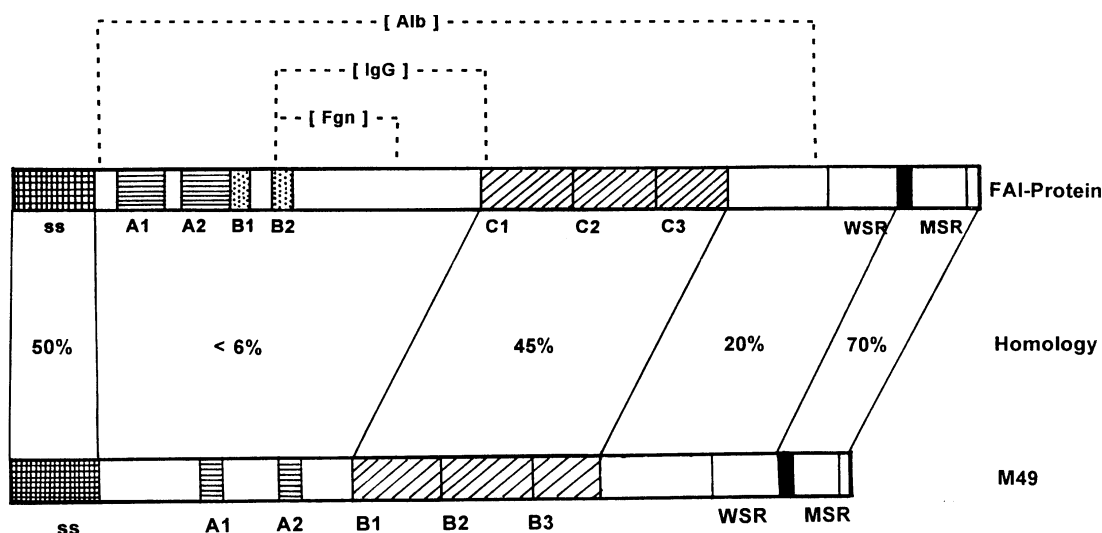
Crude cell extracts of *E. coli* cells expressing GST-FAI fusion proteins pMG3, pMG6, pMG12, pMG13 (lanes 1–4), as well as non-fusion GST (lane 5) as negative control were separated on SDS/polyacrylamide gels. The proteins were stained with Coomassie Blue (A) or blotted on to membranes that were probed with peroxidase-labelled fibrinogen (B), IgG (C) or albumin (D) for detection of ligand-binding activity. The molecular mass (kDa) of standard proteins (Sigma) is given on the left.

deletions abolished IgG-binding and the dual binding ability respectively. The pMG13 protein contained 56 amino acids of the FAI molecule and represented the only protein in which fibrinogen-binding was maintained without IgG binding (Figure 3). Furthermore the pMG10 construct revealed that the full-size fibrinogen-binding domain, or at least major parts of it, appeared to be necessary for the IgG-binding phenotype to be exhibited. An even larger region was needed for binding of albumin, since neither the N-terminal (pMG6) nor the C-terminal half (pMG4) or an overlapping central fragment (pMG15) of FAI protein had the ability to bind albumin. Thus restoration of the albumin-

binding activity required the complete surface-exposed part of the FAI protein including the fibrinogen/IgG-binding domain.

#### Homologies to M proteins of *S. pyogenes*

No homology to M proteins or any other protein could be found within the N-terminal half of the mature FAI protein which contains the fibrinogen/IgG-binding domain and is also essential for albumin binding. However, sequence homology searches of the databases revealed significant primary sequence homology of



**Figure 4** Modular structure of FAI protein and its homology to group A streptococcal M49 protein

Group C streptococcal FAI protein and *S. pyogenes* M49 protein are represented as bars, the positions of the signal sequence (ss), A, B and C repeats and wall- (WSR) and membrane-spanning (MSR) regions are indicated. Homology is expressed as percentage identical amino acid residues shared by the two proteins within the respective regions.

other regions of FAI protein to group A streptococcal M proteins. The regions that were found to be homologous were the N-terminal signal sequence and the complete C-terminal half of the FAI protein including the C repeats, the wall-associated region and the membrane-spanning domain. Homologous sequences were found in M49 [25], the M protein of an *S. pyogenes* serotype 49 strain, MRP4 [26], a fibrinogen- and IgG-binding protein of *S. pyogenes* serotype M4, ARP4 [27], an *S. pyogenes* IgA receptor protein, as well as the M5, M6 and M12 proteins originating from the respective group A streptococcal serotype strains [28–30]. The degree of M protein homology exhibited by the FAI modules is shown in Figure 4, in which an exemplary comparison of FAI protein with the M49 protein was performed. The highest primary sequence homology with 70 and 50% identical amino acid residues was found within the membrane-spanning region and the signal sequence respectively. Then an abrupt loss of similarity was found within the N-terminal region that spans the area between the signal sequence and C repeats of FAI protein. The C repeats display 45% sequence identity with the B repeats of M49, and between 40 and 50% identity with the analogue C repeats of M12 [30]. The C repeat region of M12 has recently been shown to constitute the binding domain of human serum albumin [31]. Thus the similar C repeats of the FAI protein might also be involved in albumin binding, although it was shown that they are, unlike the M12 repeats, not sufficient for binding albumin.

Sequence comparisons revealed no similarities of FAI protein to group C streptococcal surface proteins, such as the M protein of *Streptococcus equi* [32] or Protein G [33,34], except within the membrane anchor consensus sequence.

#### Specific antibodies recognize FAI protein in group C streptococci

Recombinant FAI fusion protein purified from *E. coli* cultures harbouring the expression plasmid pMG3 was used to generate specific anti-FAI antibodies. F(ab)<sub>2</sub> fragments of the antibodies were produced to avoid non-immunogenic Fc-receptor-mediated binding of the immunoglobulins. Peroxidase-labelled anti-FAI F(ab)<sub>2</sub> strongly reacted with whole cells of the group C strepto-

coccal isolate C20 spotted on nitrocellulose membranes. In Western blots anti-FAI F(ab)<sub>2</sub> reacted with concentrated C20 culture supernatant and FAI-expressing *E. coli* lysates, but not with *E. coli* control lysates or purified Protein A (results not shown). Since FAI protein possesses signal sequence and membrane-spanning region, the common feature for surface proteins of Gram-positive cocci, it is most likely that FAI protein is expressed on the surface of group C streptococcal strain C20.

#### DISCUSSION

A common feature of group A and C streptococci is their ability to express surface components that bind to host serum factors such as IgG [35], fibrinogen [36,37] and albumin [38]. In group A streptococci, members of the M protein family constitute the receptor molecules for those serum factors, whereas in group C streptococci, the structurally unrelated Protein G has been described as being responsible for albumin and/or IgG-binding activity of this species [9,11,39,40]. The molecular nature of the fibrinogen-binding component of group C streptococci is as yet unknown, but distinct receptor molecules have been suggested to be responsible for binding fibrinogen and albumin [41,42]. Hence, the primary objective of this study was to characterize the group C streptococcal fibrinogen receptor structurally, which led to the identification of a novel multiple-ligand-binding protein of group C streptococci which is distinct from the well-characterized Protein G.

FAI protein is the first group C streptococcal surface protein in which the binding regions for fibrinogen, albumin and IgG are united in one molecule. An interesting feature of FAI protein is the fact that the different binding functions are localized on overlapping modules. In M12 protein of *S. pyogenes* the C repeat region has been shown to be sufficient for binding albumin [31]. The C repeats of FAI protein, although similar to the C repeats of M12 protein, are necessary but not sufficient for albumin binding. Nearly all of the surface-exposed fraction of the molecule is required to exhibit the albumin-binding phenotype. This also includes the fibrinogen/IgG-binding region which can be localized to a 94-amino acid N-terminal fragment of FAI. Here again,

for IgG binding the complete fibrinogen-binding domain and an additional stretch of 38 amino acids is required. Only the fibrinogen-binding region could be separated from the other binding functions. It consists of 56 amino acids and is the first precisely mapped binding region for this ligand.

FAI protein appears to be related to group A streptococcal M proteins in terms of its structure. Localization of the fibrinogen/IgG-binding region of FAI, however, allowed a more precise comparison of the structure-function relationship to M protein which revealed surprising results. The fibrinogen-binding domain of FAI protein had no similarity to M protein or other proteins characterized so far. The more extended region necessary for binding IgG, which includes the fibrinogen-binding sequence, had no similarity to immunoglobulin-binding sites of M or M-like proteins or any type of Fc receptor. Since the major part of the FAI N-terminus is not homologous to M proteins but is flanked by regions abruptly exhibiting significant homology such as the signal sequence and C repeats, this region can be regarded as a foreign functional module in a structurally M-like environment. Similar features were exhibited by protein PAB, a recently identified and characterized mosaic protein of *Peptostreptococcus magnus* [43]. Protein PAB represents a protein L homologue in which a foreign albumin-binding (GA) module has been identified. There is strong evidence that the GA module originated from group G streptococcal Protein G and had been transferred via interspecific gene transfer. Recent results also showed that horizontal transfer of *emm* genes among group A streptococci led to the spread of *emm* genes and gene mosaics [44,45]. In contrast with protein PAB, the origin of the fibrinogen/IgG-binding module of FAI protein is still unknown and the mode of transfer can only be hypothesized. However, the presence of an M-like protein in a group C *Streptococcus* itself indicates horizontal transfer of genes across species borders, as has been postulated for the spread of *emm* genes from group A to group G streptococci [46].

Horizontal spread and intra- as well as inter-genic recombination are powerful mechanisms for the evolution of well-adapted surface proteins. Reorganization of functional units may not only lead to the generation of novel multifunctional molecules [47] but also to the formation of structural variants that enable the bacteria to evade the immune system of the host. Thus the FAI protein might have been evolved by such mechanisms, and the group C *Streptococcus* carrying its gene represents a potential donor within the streptococcal community.

We thank Dr. Peter Valentin-Weigand and Dr. K. N. Timmis for helpful discussions. G.S.C. thanks the Fonds der Chemischen Industrie for generous support.

## REFERENCES

- Goward, C. R., Scawen, M. D., Murphy, J. P. and Atkinson, T. (1993) Trends Biochem. Sci. **4**, 136–140
- Westerlund, B and Korhonen, T. K. (1993) Mol. Microbiol. **9**, 687–694
- Foster, T. J. and McDevitt, D. (1994) FEMS Microbiol. Lett. **118**, 199–206
- Fischetti, V. A. (1989) Clin. Microbiol. Rev. **2**, 285–314
- Colins, C. M., Kimura, A. and Bisno, A. L. (1992) Infect. Immun. **60**, 3689–3696
- Gomi, H., Hozumi, T., Hattori, S., Tagawa, C., Kishimoto, F. and Björck, L. (1990) J. Immunol. **144**, 4046–4052
- Stenberg, L., O'Toole, P. and Lindahl, G. (1992) Mol. Microbiol. **6**, 1185–1194
- Stenberg, L., O'Toole, P., Mestecky, J. and Lindahl, G. (1994) J. Biol. Chem. **269**, 13458–13464
- Frick, I. A., Akesson, P., Cooney, J., Sjöbring, U., Schmidt, K. H., Gomi, H., Hattori, S., Tagawa, C., Kishimoto, F. and Björck, L. (1994) Mol. Microbiol. **12**, 143–151
- Nilson, B. H. K., Solomon, A., Björck, L. and Akerström, B. (1992) J. Biol. Chem. **267**, 2234–2239
- Björck, L., Kastern, W., Lindahl, G. and Wiedebäck, K. (1987) Mol. Immunol. **24**, 1113–1122
- Murphy, J. P., Duggleby, C. J., Atkinson, M. A., Trowern, A. R., Atkinson, T. and Goward, C. R. (1994) Mol. Microbiol. **12**, 911–920
- Frick, I. M., Wikström, M., Forsen, S., Drakenberg, T., Gomi, H., Sjöbring, U. and Björck, L. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 8532–8536
- Talay, S. R., Ehrenfeld, E., Chhatwal, G. S. and Timmis, K. N. (1991) Mol. Microbiol. **5**, 1727–1734
- Talay, S. R., Valentin-Weigand, P., Timmis, K. N. and Chhatwal, G. S. (1994) Mol. Microbiol. **13**, 531–539
- Kreikemeyer, B., Talay, S. R. and Chhatwal, G. S. (1995) Mol. Microbiol. **17**, 137–145
- Sjöbring, U., Björck, L. and Kastern, W. (1991) J. Biol. Chem. **266**, 399–405
- Jonsson, H., Frykberg, L., Rantamäki, L. and Guss, B. (1994) Gene **143**, 85–89
- Jonsson, H. and Müller, H.-P. (1994) Eur. J. Biochem. **220**, 819–826
- Jonsson, H., Lindmark, H. and Guss, B. (1995) Infect. Immun. **63**, 2968–2975
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Schauder, B., Blöcker, H., Frank, R. and McCarthy, J. E. G. (1987) Gene **52**, 279–283
- Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 4767–4771
- Harlow, E. and Lane, D. (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Haanes, E. J. and Cleary, P. (1989) J. Bacteriol. **171**, 6397–6408
- O'Toole, P., Stenberg, L., Rissler, M. and Lindahl, G. (1992) Proc. Natl. Acad. Sci. U.S.A. **89**, 8661–8665
- Frithz, E., Heden, L.-O. and Lindahl, G. (1989) Mol. Microbiol. **3**, 1111–1119
- Miller, L., Gray, L., Beachey, E. and Kehoe, M. (1988) J. Biol. Chem. **263**, 5668–5673
- Hollingshead, S. K., Fischetti, V. A. and Scott, J. R. (1986) J. Biol. Chem. **261**, 1677–1686
- Robbins, J. C., Spanier, J. G., Jones, S. J., Simpson, W. J. and Cleary, P. P. (1987) J. Bacteriol. **169**, 5633–5640
- Retnoningrum, D. S. and Cleary, P. P. (1994) Infect. Immun. **62**, 2387–2394
- Timoney, J. F., Walker, J., Zhou, M. and Ding, J. (1995) Infect. Immun. **63**, 1440–1445
- Fahnestock, S. R., Alexander, P., Nagle, J. and Filpula, D. (1986) J. Bacteriol. **167**, 870–880
- Olsson, A., Eliasson, M., Guss, B., Nilsson, B., Hellman, U., Lindberg, M. and Uhlen, M. (1987) Eur. J. Biochem. **168**, 319–324
- Kronvall, G. (1973) J. Immunol. **111**, 1401–1406
- Chhatwal, G. S. and Blobel, H. (1986) in Fibrinogen-Fibrin Formation and Fibrinolysis (Lane, D. A., Henschen, A. and Jasani, M. K., eds.), vol. 4, pp. 239–249, Walter de Gruyter & Co., Berlin and New York
- Kronvall, G., Schonbeck, C. and Myhre, E. B. (1979) Acta Pathol. Microbiol. Scand. Sect. B **87**, 303–310
- Myhre, E. B. and Kronvall, G. (1980) Infect. Immun. **27**, 6–14
- Reis, K. J., Ayoub, E. M. and Boyle, M. D. P. (1984) J. Immunol. **132**, 3091–3097
- Otten, R. A. and Boyle, M. D. P. (1991) J. Microbiol. Methods **13**, 185–200
- Chhatwal, G. S., Dutra, I. S. and Blobel, H. (1985) Microbiol. Immunol. **29**, 973–980
- Widebäck, K. and Kronvall, G. (1982) Infect. Immun. **38**, 1154–1163
- de Chateau, M. and Björck, L. (1994) J. Biol. Chem. **269**, 12147–12151
- Bessen, D. E. and Hollingshead, S. K. (1994) Proc. Natl. Acad. Sci. U.S.A. **91**, 3280–3284
- Whatmore, A. M. and Kehoe, M. A. (1994) Mol. Microbiol. **11**, 363–374
- Simpson, W. J., Musser, J. M. and Cleary, P. P. (1992) Infect. Immun. **60**, 1890–1893
- Podbielski, A., Hawitzky, J., Pack, T. D., Flosdorff, A. and Boyle, M. D. P. (1994) Mol. Microbiol. **12**, 725–763