The EMBO Journal vol.15 no.5 pp.989-998, 1996

A two-domain mechanism for group A streptococcal adherence through protein F to the extracellular matrix

Vered Ozeri, Aviva Tovi, Israel Burstein, Shira Natanson-Yaron, Michael G.Caparon[1], Kenneth M.Yamada[2], Steven K.Akiyama[2], Israel Vlodavsky[3] and Emanuel Hanski[4]

Department of Clinical Microbiology, The Hebrew University-Hadassah Medical School, Jerusalem 91010, Israel, [1]Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110, [2]Laboratory of Developmental Biology, NIDR, National Institutes of Health, Bethesda, MD 20892, USA and [3]Department of Oncology and Medicine B, Hadassah University Hospital, Jerusalem 91120, Israel

[4]Corresponding author

Streptococcus pyogenes **binds to the extracellular matrix (ECM) and a variety of host cells and tissues, causing diverse human diseases. Protein F, a *S.pyogenes* adhesin that binds fibronectin (Fn), contains two binding domains. A repeated domain (RD2) and an additional domain (UR), located immediately N-terminal to RD2. Both domains are required for maximal Fn binding. In this study, we characterize RD2 and UR precisely and compare their functions and binding sites in Fn. The minimal functional unit of RD2 is of 44 amino acids, with contributions from two adjacent RD2 repeats flanked by a novel 'MGGQSES' motif. RD2 binds to the N-terminal fibrin binding domain of Fn. UR contains 49 amino acids, of which six are from the first repeat of RD2. It binds to Fn with higher affinity than RD2, and recognizes a larger fragment that contains fibrin and collagen binding domains. Expression of UR and RD2 independently on the surface-exposed region of unrelated streptococcal protein demonstrates that both mediate adherence of the bacteria to the ECM. We describe here a mechanism of adherence of a pathogen that involves two pairs of sites located on a single adhesin molecule and directed at the same host receptor.**

Keywords: bacterial adherence/extracellular matrix/fibronectin/group A streptococcus/protein F

Introduction

Fibronectin (Fn) is a multifunctional glycoprotein present in soluble form in plasma and other body fluids, and in insoluble form in extracellular matrices (ECMs) and basement membranes (Mosher, 1989; Hynes, 1990). Fn plays prominent roles in a number of cellular processes, including cell–cell and cell–substrate adhesion, cell spreading, differentiation, phagocytosis and wound healing. This is principally due to its ability to bind specifically to a variety of receptors and substrate molecules, including collagen, fibrin, heparin, DNA, actin, complement component, ganglioside and others (Mosher, 1989; Hynes, 1990). By developing the ability to bind Fn, many bacterial pathogens, including staphylococci, *Escherichia coli*, *Treponema pallidium*, mycobacteria and streptococci have exploited the binding properties of Fn in order to acquire a mechanism for adherence, colonization and subsequent invasion of host tissues (Ólsen *et al.*, 1993; Westerlund and Korhonen, 1993). Thus, an understanding of the molecular details of bacteria–Fn interactions will not only provide considerable insight into the pathogenesis of many bacterial diseases, but will also serve as a powerful model system for structural and functional analyses of Fn itself.

Streptococcus pyogenes, a group A streptococcus (GAS) is an example of a pathogen that has evolved the ability to bind Fn. GAS is the etiologic agent of a diverse collection of human diseases that range from self-limiting suppurative infections of the pharynx and the skin, to more serious invasive infections, a toxic shock-like syndrome, life-threatening sepsis and post-infection sequelae of acute glomerulonephritis or rheumatic fever (Stevens *et al.*, 1989; Bisno, 1991). A common feature of these diverse diseases is that each infection is initiated by the interaction of GAS with epithelial cells of the pharynx and skin (Wannamaker, 1970). Considerable evidence has accrued in recent years to suggest that the organism's ability to bind Fn plays an important role in the adherence process (Hasty *et al.*, 1992).

Several surface components of GAS have been implicated in Fn binding (Hasty *et al.*, 1992). Perhaps the best evidence for the role of a specific surface protein in promoting the ability of GAS to bind Fn has come from studies of a family of highly homologous proteins which include sfb (Talay *et al.*, 1991) and protein F (Hanski and Caparon, 1992). These proteins bind Fn when expressed in *Escherichia coli* (Hanski and Caparon, 1992; Talay *et al.*, 1992), can inhibit the binding of Fn to intact streptococcal cells (Hanski and Caparon, 1992; Talay *et al.*, 1992) and can inhibit the adherence of GAS to several epithelial cells (Talay *et al.*, 1992). Insertional inactivation of the gene encoding protein F (*prt*F), generated a mutant which lost Fn binding activity and the ability to adhere to respiratory epithelial cells (Hanski and Caparon, 1992). Moreover, expression of *prt*F in *Enterococcus faecalis* conferred upon these organisms the ability to bind Fn and to adhere to respiratory epithelial cells (Hanski *et al.*, 1992). These studies have provided strong evidence for the role of these proteins in promoting Fn binding and adherence.

DNA sequence analyses of sfb and protein F have revealed that they share a number of features in common with the Fn binding proteins that have been identified in *Staphylococcus aureus* (Signäs *et al.*, 1989; Jönsson *et al.*, 1991) and *Streptococcus dysgalactiae* (McGavin *et al.*, 1993). A prominent shared feature is a tandem repeat

element that is found adjacent to the conserved C-terminal cell attachment domain. It consists of up to six repeats of 32–44 amino acids, shows a high degree of homology between these different proteins and is considered to mediate the binding to Fn (McGavin *et al.*, 1993; Talay *et al.*, 1994). However, in the case of protein F, construction, purification and analysis of a series of mutant molecules revealed that an additional domain, located immediately N-terminal to the repeat domain, is essential for maximal Fn binding activity (Sela *et al.*, 1993).

In this study, we have defined precisely the additional Fn binding domain and have examined the relationship between these two Fn binding domains in detail. By expression of the repeat domain (RD2) and the additional binding domain (UR) on the surface of the streptococcal cell using the secretion and anchoring domains of unrelated streptococcal protein we showed that: (i) UR is composed of a 49 amino acid binding domain and mediates high affinity binding of protein F to Fn; (ii) the minimal functional unit of RD2 does not correspond to a single repeat as previously described but rather consists of 44 residues of two contiguous repeats; and (iii) UR and RD2 bind to two different fragments of Fn, and each domain is capable of mediating the adherence of GAS to the ECM.

This study provides the first detailed mapping of two independent functional sites on a single bacterial adhesin that can bind independently to two different domains on the same host receptor, i.e. Fn. This strategy provides the bacterium with an efficient, high affinity mechanism of adherence, and may account for the successful interaction of GAS with such a variety of human cells and tissues.

Results

Identification of UR, the high affinity Fn binding domain

We previously reported the existence of a second Fn binding domain of 43 amino acids located immediately N-terminal to RD2, which we termed UFBD (upstream Fn binding domain). UFBD and RD2 are both required for maximal Fn binding to the bacteria (Sela *et al.*, 1993). To characterize further the properties of these two domains, we constructed a series of proteins using the expression system described in Materials and methods that contained either UFBD or RD2 on separate molecules, or a single molecule that contained both domains. It was discovered that a single protein composed of UFBD + RD2 (five repeats) was much more potent in inhibiting Fn binding to the bacteria than either domain alone, or when the two domains on separate molecules were admixed (not shown). Furthermore, a protein which was composed of UFBD and a single repeat of RD2 was as efficient at inhibiting binding as a protein that consisted of UFBD and all five RD2 repeats (pPTF54, Figure 1C; Sela *et al.*, 1993). Since a single RD2 repeat by itself has no Fn binding activity (see below), these results suggest that a high affinity Fn binding site resides on this stretch of amino acids that contains UFBD and a single RD2 repeat (Figure 1A). To localize further this Fn binding domain, which was termed UR, a series of proteins were constructed that were truncated at the C-terminal end of the RD2 repeat (Figure 1A and B), and tested for their abilities to block Fn binding to the bacterium. The smallest protein which retained full activity (pUR-4; Figure 1A and C) was composed of 49 amino acids of which 43 are located towards the N-terminus of the first RD2 repeat (previously termed UFBD) and the rest are the first six amino acids of this repeat. The high affinity binding was lost upon truncation of these six amino acids (pUR-1, Figure 1A and C), indicating that some or all of these amino acids are essential for the high affinity binding of UR. It is of interest to note that these amino acids contain the conserved 'EDT' motif present in the repeat domains of Fn binding proteins from *S.aureus* and *S.dysgalactiae* (McGavin *et al.*, 1993), and the RD2 domain of *S.pyogenes* (see below).

Identification of the functional unit of the RD2-type repeats

The observation that UR contains a region derived from RD2 suggested that a functional binding unit of RD2 itself might be more complex than originally believed. A related question concerns the importance of the number of repeats contained by RD2, since a survey of *prt*F-related genes in 109 different strains revealed that the number of RD2-type repeats varies from one to six (Natanson *et al.*, 1995). One strain, (SS914 of M-type 59) bound high levels of Fn but contained only two repeats. Subsequent sequencing of the RD2 domain derived from this strain revealed the presence of a complete and a partial repeat (Figure 2A), which have an amino acid sequence almost identical to the first two repeats of protein F (Sela *et al.*, 1993). Expression of these two repeats as described in Materials and methods produced a polypeptide that blocked Fn binding to JRS145 (expressing protein F) with similar efficiency to the protein expressing five repeats (pPTF51, Figure 2C; Sela *et al.*, 1993). However, a protein expressing the first 37 amino acids, comprising the complete repeat unit (pPRD2-2), failed to bind Fn (Figure 2A–C). Since the segment consisting of two contiguous repeats produced an active protein (pRD2-1; Figure 2A–C), these results suggested that a functional Fn binding unit of RD2 includes elements of both repeats. To identify the precise composition of a functional RD2 unit, we constructed and purified a series of mutant proteins containing various N- or C-terminal truncations (Figure 2A and B), and tested their ability to block Fn binding to the intact bacterium. As shown in Figure 2C, the smallest active protein (pRD2-8; Figure 2A and C) was found to be composed of 44 amino acids, and overlapped the junction between the two contiguous repeat units. The presence of the motif 'MGGQSES' at both ends of the functional binding unit was essential for Fn binding activity.

Although the relative sizes of the purified proteins expressing UR or RD2 fit their composition (Figures 1A and B and 2A and B), their apparent molecular weights (M_r), as determined by SDS–PAGE, are higher than the calculated M_r. A discrepancy between calculated and determined M_r was also observed for the entire molecule of protein F (Sela *et al.*, 1993) and the Fn binding protein of *S.aureus* (Signäs *et al.*, 1989).

Expression of UR and RD2 within the surface-exposed region of M protein

To test the structure–function relationship of the Fn binding domains directly and independently of the other portions

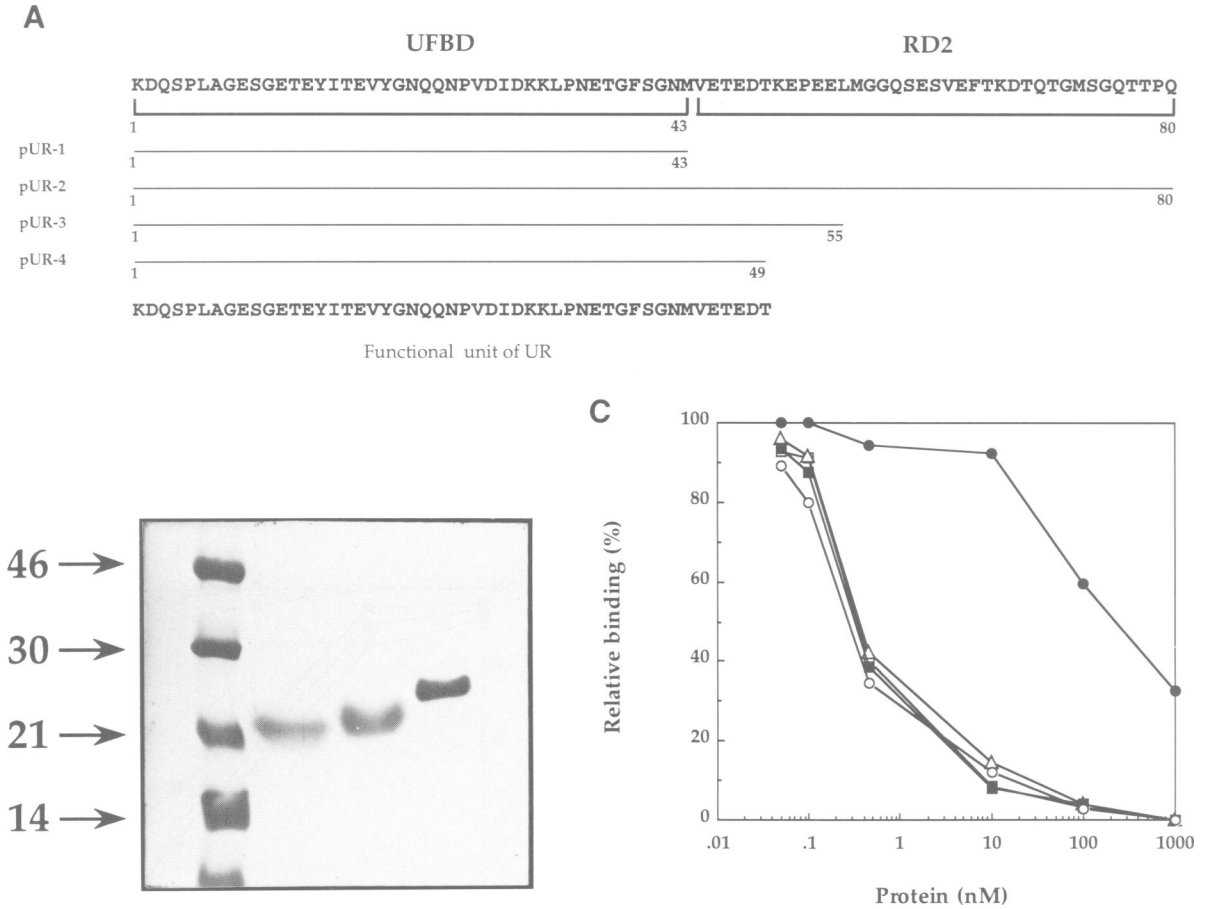


Fig. 1. Precise characterization of a functional UR. (**A**) The amino acid sequence of the UR domain composed of UFBD and a single RD2 repeat is shown. For convenience, the first amino acid of UR is denoted as number 1. The plasmids pUR-1–pUR-4 harboring the indicated DNA segments were derived as described in Materials and methods. (**B**) Coomassie brilliant blue stain of 12.5% SDS–polyacrylamide gel on which the following proteins were loaded: molecular weight markers (in kDa) and proteins expressed by the plasmids (from left to right): pUR-4; pUR-3; and pUR-2. The extent of purity of proteins expressed by the other plasmids described in (A) was similar. (**C**) Inhibition of Fn binding by purified proteins. The proteins expressed by the following plasmids were compared for their ability to inhibit Fn binding to JRS145: pUR-1, ●; pUR-2, □; pUR-3, △; pUR-4, ■; pPTF54 (Sela *et al.*, 1993, see text) ○. Total binding of Fn by JRS145 was 115 000 c.p.m. and the non-specific binding was 2800 c.p.m.; specific binding is presented as 100%.

of protein F, and to rule out the possibility that an individual repeat may be too small to fold independently and thus to acquire an active conformation, we have expressed these domains on the streptococcal cell surface using fusion proteins with the secretion and anchoring domains of M protein. The chimeric proteins were expressed in an *S.pyogenes* strain (SAM2) in which the genes encoding both protein F and M6 protein have been inactivated (Norgren *et al.*, 1989; Hanski and Caparon, 1992). The procedure for the construction and expression of the chimeric proteins is summarized in Materials and methods and is detailed elsewhere (Hanski *et al.*, 1995).

The composition of the regions of protein F that were expressed within the surface-exposed region of M protein is described in Figure 3A. Since M protein is a highly expressed surface protein of GAS, it was not surprising that a 50-fold lower concentration of bacterial cells of a strain which expresses a M/F chimera, SAM19 (containing the entire binding region of protein F), was required to produce a dose–response curve approaching saturation, similar to the dose–response curve of the strain which expresses native protein F (JRS145). Under these conditions, the number of Fn binding sites, as determined for SAM19, was 65-fold higher than that of JRS145 (Figure 3B), suggesting that the M protein is 65-fold more abundant on the streptococcal surface than is protein F. However, the affinity for Fn of JRS145 was similar to that of SAM19, both having apparent K_d values of ~1 nM.

In contrast to the results obtained for SAM19, SAM17 (which expresses a protein composed of only the five RD2 repeats, Figure 3A) bound Fn at much lower affinity and its binding did not approach saturation even at the highest concentration tested (Figure 3B). From the inhibition experiment shown in Figure 4A, it appears that the apparent affinity of RD2 is in the 50 nM range. No binding to SAM13, which expresses a chimeric protein with only a single RD2 repeat, was detected under these conditions (Figure 3B), or even when a 50-fold higher concentration of the bacterium was used (not shown). This indicates that a single repeat is inactive even when expressed as a large fusion protein. Significantly, the dose–response curve of SAM25, expressing UFBD plus the first entire RD2 repeat (Figure 3A), was almost identical to that of SAM19 (not shown). Since the only active determinant in SAM25 is the functional UR unit, it

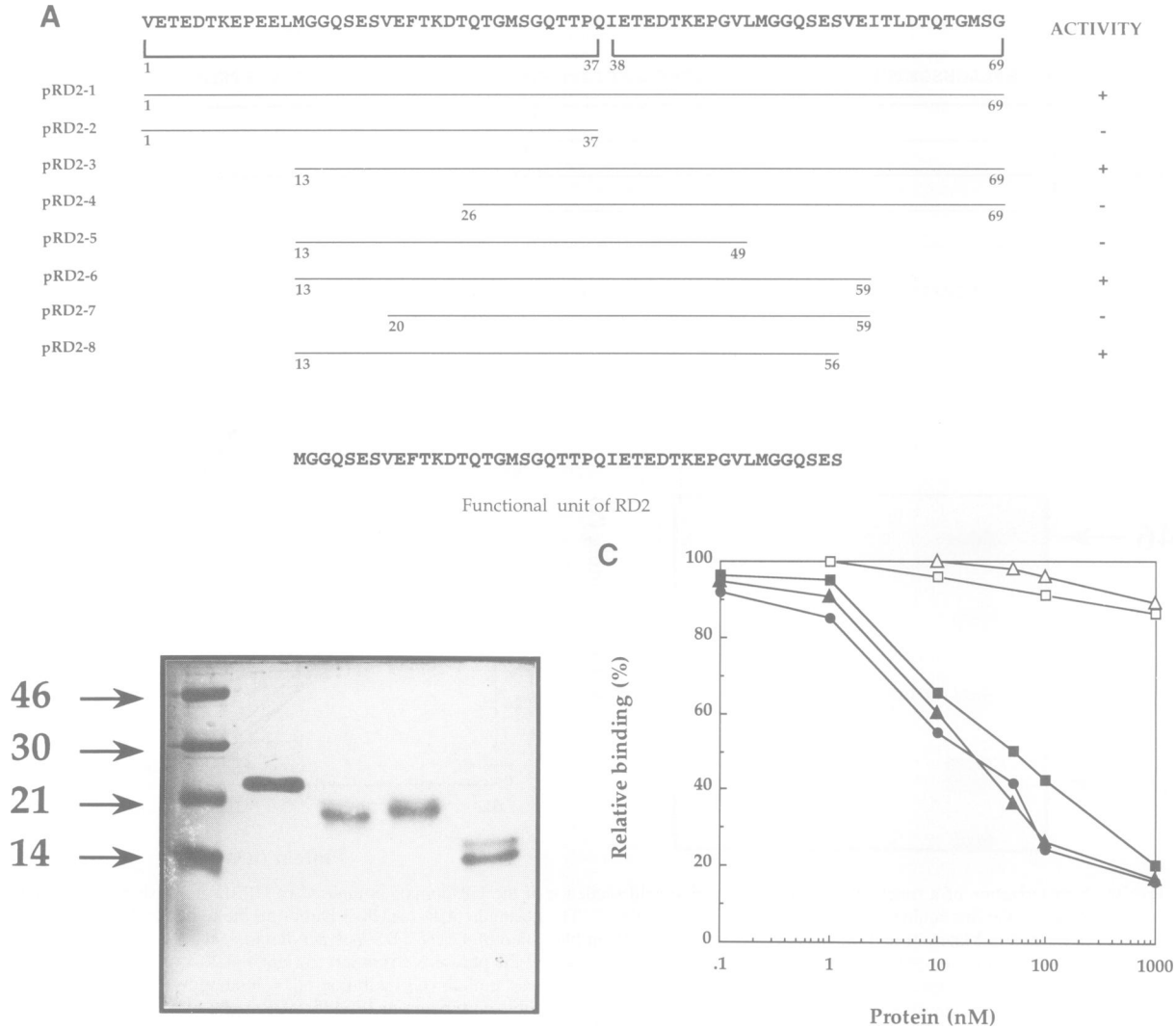


Fig. 2. Precise characterization of a functional RD2 repeat. (**A**) The RD2 domain in strain SS914 of M-type 59 is composed of two repeats: one of 37 and the other of 32 amino acids. For convenience, the initial amino acid of the first repeat is denoted here as number 1. The plasmids pRD2-1–pRD2-8, harboring the indicated DNA segments, were derived from the RD2 domain of SS914 as described in Materials and methods. The symbols + and – represent the presence or absence of Fn binding by the proteins expressed from the DNA segments. (**B**) Coomassie brilliant blue stain of 15% SDS–polyacrylamide gel on which the following proteins were loaded: molecular weight markers (in kDa) and proteins expressed by the following plasmids (from left to right): pRD2-3; pRD2-6; pRD2-8; and pRD2-2. The purity of proteins expressed by the other plasmids described in (A) was similar. (**C**) Inhibition of Fn binding by purified proteins. The proteins expressed by the following plasmids were compared for their ability to inhibit the binding of iodinated Fn to JRS145: pRD2-2, □; pRD2-7, △; pRD2-8,▲; pRD2-1, ■; pPTF51 (Sela *et al.*, 1993, see text) ●. For the sake of clarity, we have omitted curves for the proteins expressed by the remainder of the plasmids presented in (A). Proteins possessing Fn binding activity displayed curves similar to that of pRD2-8, whereas proteins lacking activity displayed curves similar to that of pRD2-2. Total binding of Fn by JRS145 was 125 000 c.p.m., and the non-specific binding was 2500 c.p.m.; specific binding is presented as 100%.

indicates that UR is the major determinant mediating binding of Fn to protein F.

Localization of RD2 and UR binding sites on Fn

We demonstrated previously that RD2 effectively blocked the binding of the radiolabeled N-terminal 29 kDa fibrin binding domain of Fn (Fib. 1) to GAS, whereas no inhibition of binding was observed for UFBD. These results led us to propose that RD2 binds to Fib. 1 and UFBD binds to a different region of Fn (Sela *et al.*, 1993). This hypothesis was validated by the experiment depicted in Figure 4A, showing that even though Fib. 1 effectively blocked the binding of Fn to SAM17 (expressing RD2 only), it had no effect on the binding of Fn to SAM25 (expressing a functional UR), and reduced by only 20% or less the binding of Fn to JRS145 (containing both domains). The latter observation underscores the conclusion that the binding of soluble Fn to protein F is mediated primarily by the UR domain.

Our initial attempts to block Fn binding to SAM25 by admixing six distinct fragments of Fn produced by the procedure of Borsi *et al.* (1986) failed, even though the mixture included Fib. 1, collagen binding (Col.), cell binding, heparin binding (two fragments of 38 and 29 kDa) and the C-terminal fibrin binding domains. However, a mixture of purified Fib. 1 and the 43 kDa Col. fragments completely blocked the binding of Fn to JRS145, although the latter had no effect by itself (Figure 4B). In contrast, as has already been mentioned, Fib. 1 + Col. did not affect the binding of Fn to SAM25

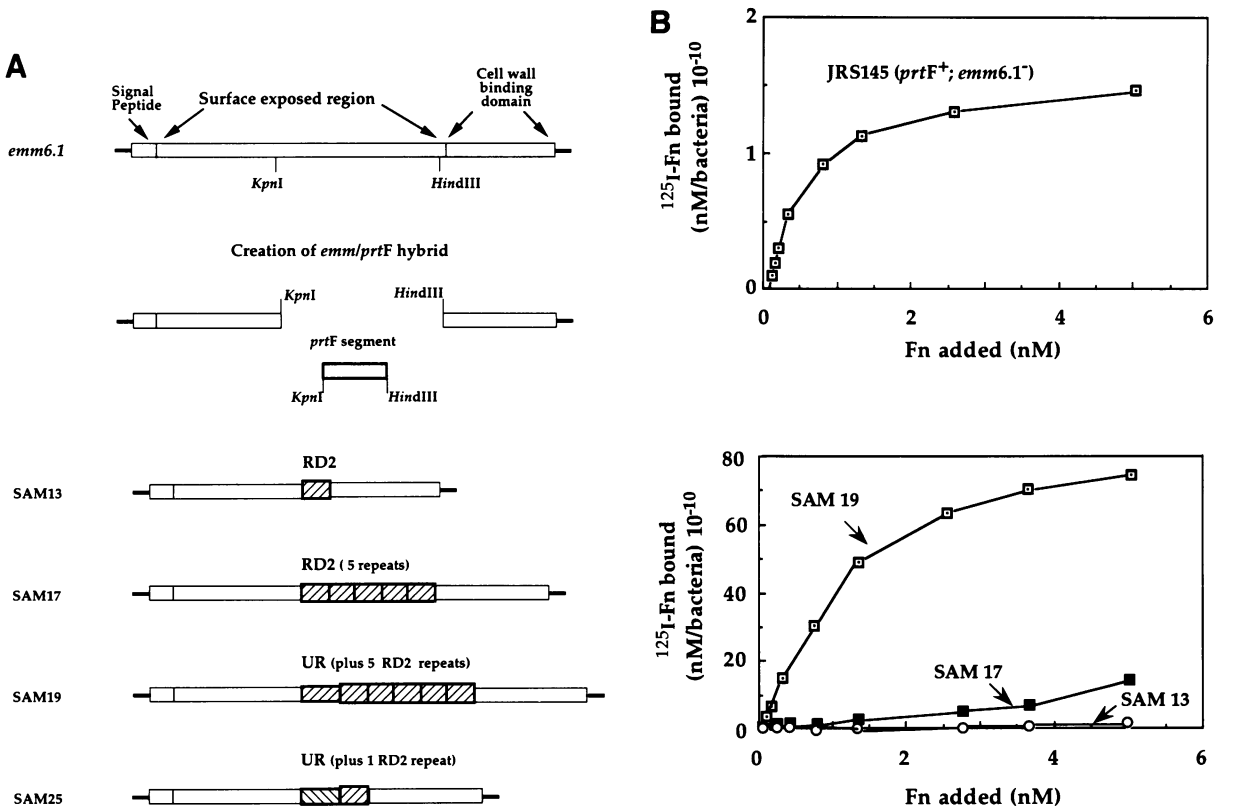


Fig. 3. Expression of protein F domains in hybrid proteins exposed on the streptococcal cell surface and their response to Fn. (**A**) A schematic representation that illustrates the construction of the *emm/prtF* hybrids. The procedure for the creation of the hybrids SAM13–25 is detailed elsewhere (Hanski *et al.*, 1995). (**B**) Dose–response curves for binding of iodinated Fn to JRS145, SAM19, SAM17 and SAM13 were determined as detailed in Materials and methods.

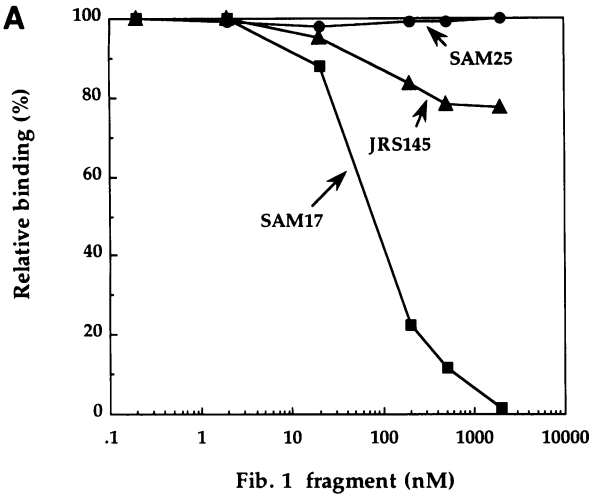
(Figure 4C). These results suggested the hypothesis that the 70 kDa N-terminal fragment of Fn, composed of Fib. 1 and the collagen domains, might be responsible for binding of Fn to UR. The binding of Fib. 1 to RD2, which is adjacent to UR on protein F, might facilitate the binding of the Fib. 1 and Col. domains of Fn to UR, even when added separately. Indeed, a mixture of the two fragments fully displaced Fn binding to JRS145, but with a potency that was ~10-fold lower than that of Fn (Figure 4B and C). This result also argues against the possibility that UR recognizes a small stretch of amino acids located between the Fib. 1 and the Col. domains in Fn and is cleaved or lost during their preparations by restricted proteolysis.

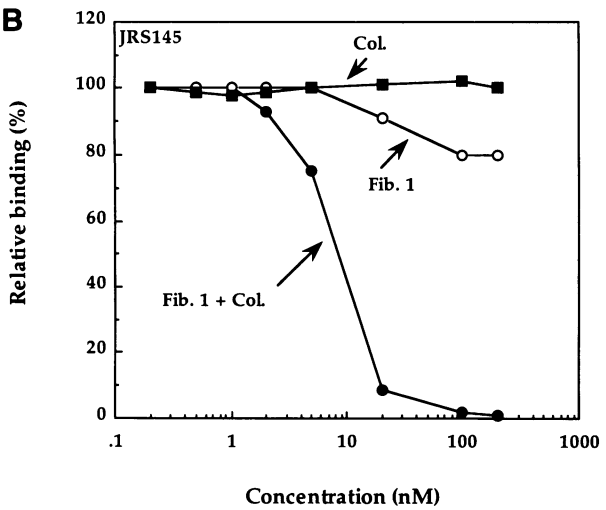
To test this hypothesis further, we produced the 70 kDa N-terminal fragment of Fn (McKeown-Longo and Mosher, 1985) and examined its ability to block Fn binding to SAM25. As shown in Figure 4C, the 70 kDa fragment blocked the binding of iodinated Fn with a potency similar to that of a non-iodinated Fn. These results establish that the fragment of Fn that binds most effectively to UR is composed of Fib. 1 and Col. domains linked together in a single fragment. A schematic representation of the interactions of UR and RD2 with the various fragments of Fn is depicted in Figure 5.

GAS adhere to the ECM through RD2 and UR

In order to study the role of UR and RD2 in adherence of GAS to the host, we examined the abilities of the strains expressing either of the two domains to attach to an ECM. We first compared mutants which express either protein F or M protein for their adherence to ECM produced by bovine corneal endothelial cells (Ishai-Michaeli *et al.*, 1992). It was found that the protein F-producing but M protein-deficient strain, JRS145, adhered well to cover slips coated with ECM (Figure 6A). In contrast, a mutant which expresses M protein but does not express protein F (SAM1), and a mutant which does not express either protein (SAM2), attached only very poorly (Figure 6A). These results demonstrate that the adherence of GAS to this ECM is mediated by protein F.

Since M protein did not contribute to the adherence in this assay, we could compare the abilities of various strains which expressed M/F chimeric proteins to adhere to this ECM, since all the observed adherence would be conferred by the protein F domain. Indeed, SAM19 expressing the chimeric protein containing the entire Fn binding domain of protein F adhered to the ECM as did JRS145 (Figure 6A). To compare the adherence of the various strains on a more quantitative basis, bacteria were radiolabeled with tritiated adenine, and their adherence to the ECM was determined. Strains which express native protein F (JRS145), the five repeats of RD2 alone (SAM17), the UR and five RD2 repeats (SAM19) or UR and a single RD2 repeat (SAM25), all attached very well (Figure 6B). However, strains which lacked protein F (SAM1, SAM2), or expressed only a single RD2 repeat (SAM13), adhered very poorly (Figure 6B). Thus, UR and RD2 are both responsible for the ability of protein F to promote the adherence of GAS to the host.





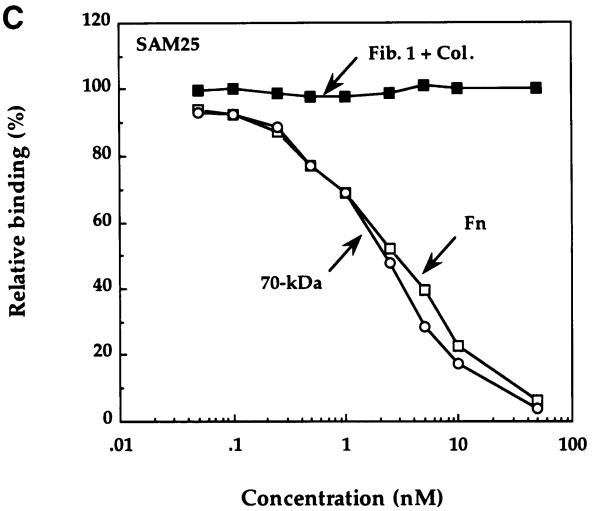


Fig. 4. UR and RD2 bind to the 70 and the 29 kDa N-terminal fragments of Fn, respectively. (**A**) Displacement of Fn by the 29 kDa N-terminal fragment (Fib. 1) from the indicated bacteria. The specific binding values of SAM17, SAM25 and JRS145 in the absence of Fib. 1 were: 10 000, 134 000 and 110 000 c.p.m., respectively. These values were denoted as 100% binding. (**B**) Displacement of Fn from JRS145 by Fib. 1, 40 kDa collagen binding fragment (Col.) and a mixture of the two fragments. The specific binding of JRS145 that was denoted as 100% was 115 000 c.p.m. (**C**) Displacement of Fn from SAM25 by a mixture of Fib. 1 + Col., Fn and the 70 kDa N-terminal fragments of Fn containing the Fib. 1 and the Col. fragments. The specific binding of SAM25 which was denoted as 100% binding was 138 000 c.p.m.

Discussion

In this study, we further characterized the structure–function relationship between the two Fn binding domains of protein F. The major new findings in this study are: (i) protein F contains a 49 amino acid binding domain termed UR which mediates high affinity binding to Fn and which accounts for most of the binding of protein F to soluble Fn; (ii) UR binds to the 70 kDa N-terminal region of Fn, which must contain both the fibrin and collagen binding domains linked together; (iii) a lower affinity binding site on protein F that binds to the fibrin binding domain of Fn was also mapped and found to be a 44 amino acid sequence that unexpectedly consists not of a single RD2 motif but, instead, of two adjacent segments of two contiguous repeats flanked by a novel 'MGGQSES' motif; and (iv) UR and RD2 can each function independently to mediate the adherence of the bacteria to the ECM. This novel mechanism involving dual binding domains in both protein F and Fn would be of obvious advantage to the bacterial pathogen by permitting efficient attachment to tissues. These mechanisms may be common to other bacterial systems, and the chimeric surface protein test system that we describe here should be useful for characterization of these and other cell adhesive interactions.

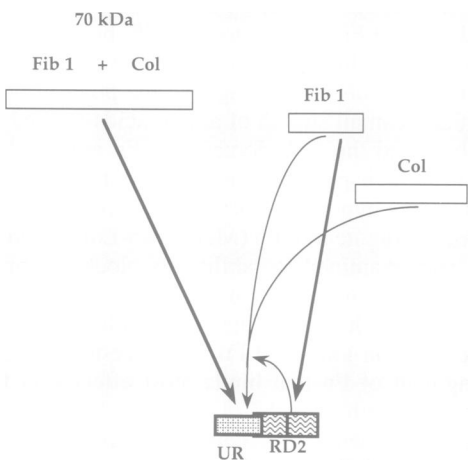


Fig. 5. Schematic representation of the various interactions that occur between UR, RD2 and the fibrin and collagen binding domains of Fn. The Fn binding domain of protein F is composed of UR and a functional RD2 unit. UR binds to the 70 kDa N-terminal fragment of Fn, containing the fibrin (Fib.1) and the collagen (Col.) binding domains, only when these domains are linked together. RD2 binds to Fib.1, but does not interact with Col. The model hypothesizes that upon occupation of RD2 by Fib.1, binding of Fib. 1 and Col. to UR is possible, even when these fragments of Fn are added separately.

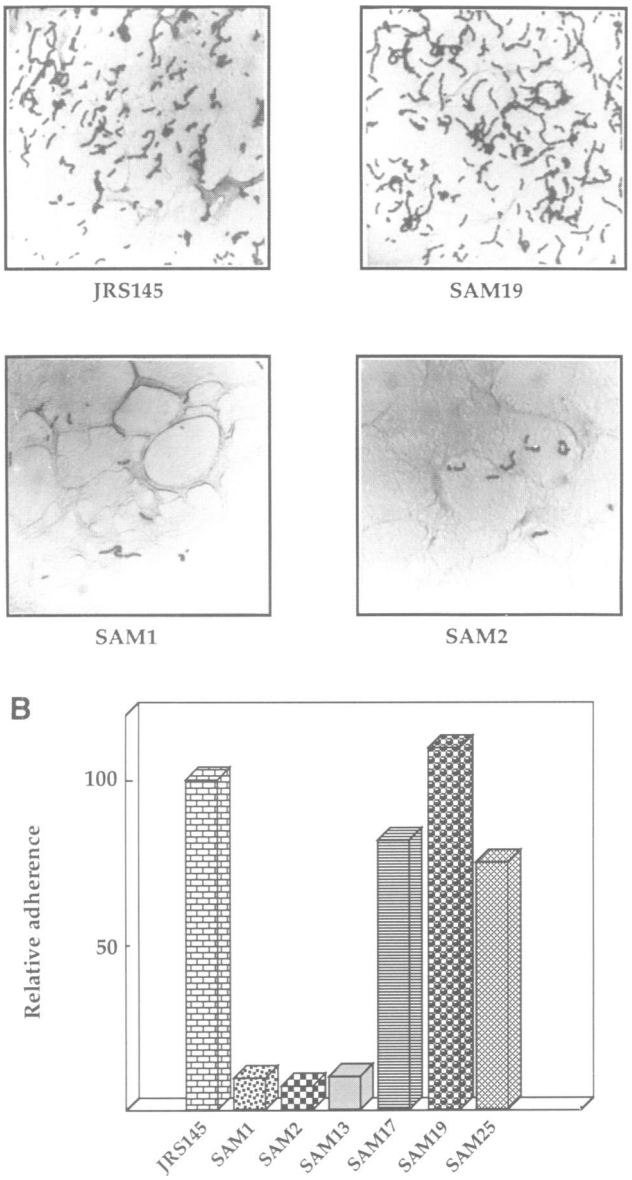


Fig. 6. UR and RD2 confer adherence of bacteria to ECM when expressed in a surface-exposed region of M protein. (**A**) The adherence of various bacteria to ECM was investigated as described in Materials and methods. The streptococci are the small, darkly stained objects on the surface of the ECM. Magnification, 250×. (**B**) Adhesion of radiolabeled bacteria. The adherence assay using radiolabeled streptococci was conducted as described in Materials and methods. 100% adherence represents the adherence of JRS145.

Indeed, the staphylococcal Fn binding protein seems also to recognize two domains in Fn (Bozzini *et al.*, 1992); however, this suggestion was not fully corroborated.

The Fn binding protein of *S.aureus* has been characterized extensively and numerous studies have established that three repeats of a 38 amino acid motif bind to the N-terminal fibrin binding domain of Fn (Mosher and Proctor, 1980; Flock *et al.*, 1987). A 32–37 amino acid long repetitive element found in two different surface proteins of *S.dysgalactiae* have also been implicated to bind to the same domain of the Fn molecule (McGavin *et al.*, 1993). These repetitive elements share conserved features with each other and with the RD2 repeat domain of protein F (McGavin *et al.*, 1993). However, as shown here, the binding of Fn to protein F is evidently more complex than it is in these other proteins, since in protein F the repetitive element is not the dominant binding domain of the molecule. This contention was supported initially by the observation that it was not possible to block the binding of Fn to protein F-expressing bacteria using the N-terminal fibrin binding domain of Fn (Fib. 1). An inhibition of only 20–30% was achieved at Fib. 1 concentrations >100 nM (Figure 4A and B), while the apparent K_d of protein F to Fn is in the low nM range. This observation suggested the existence of a second high affinity Fn binding domain on protein F.

We demonstrate here that this high affinity domain, termed UR, mediates most of the binding of Fn to protein F. UR binds to the 70 kDa N-terminal fragment of Fn with an affinity similar to that of intact protein F for Fn. Polypeptides expressing UR were several orders of magnitude more effective in inhibiting Fn binding to protein F than were polypeptides expressing the RD2 domain. Recently, F-like proteins that vary in the number of RD2 repeats have been identified (Natanson *et al.*, 1995). We have begun the analysis of an Fn binding strain (SS636) expressing UR with a single RD2 repeat (V.Ozeri and E.Hanski, unpublished observation). Since a single RD2 repeat has no Fn binding activity, UR is solely responsible for Fn binding activity in this strain. Taken together with those findings described above, these findings demonstrate that UR accounts for most of the binding of Fn to GAS strains expressing F-like proteins.

Screening of other Fn binding proteins for homology to UR revealed that complete identity was present only in GAS strains expressing F-like proteins. However, it was particularly interesting to discover that two Fn binding proteins of *S.dysgalactia* and the Fn binding protein of *Streptococcus equisimilis* possess the motif 'SGLSGETGQS' located N-terminal to the repeat domains of their respective proteins and displaying a high degree of structural homology with the sequence 'SPLAGESGET' found at the N-terminus of the UR. This finding may suggest that the Fn binding proteins of *S.dysgalactia* and *S.equisimilis* also contain a second Fn binding domain.

A number of observations suggesting that binding of Fn to RD2 itself was more complex than originally believed led us to investigate and define the composition of the minimum functional unit of the RD2 repeat. These include the observations that: (i) UR includes amino acids from the first repeat of RD2; (ii) the three repeats of *S.aureus*, when expressed in tandem, have an affinity for Fn that is several orders of magnitude higher than the affinity of the isolated individual repeats (Flock *et al.*, 1987); and (iii) synthetic peptides corresponding to specific repeats from both *S.aureus* and *S.dysgalactiae* proteins differ in their effectiveness in blocking Fn binding to their respective bacteria (Signäs *et al.*, 1989; McGavin *et al.*, 1993). Our analysis revealed that a single RD2 repeat unit of 37 amino acids was incapable of binding Fn, even when expressed within the surface-exposed region of an unrelated streptococcal protein. This is in contrast to another report that a single RD2 repeat could block the binding of Fn to GAS and adherence of the bacteria to cultured epithelial cells (Talay *et al.*, 1992). The reason for this discrepancy is not clear, but it may result from

the use of a different technique for the construction of the fusion protein with protein F. Our results demonstrate instead that an active unit of the RD2 repeat consists of 44 amino acids located on a fragment composed of two adjacent repeats. The sequence 'MGGQSES' must be present at both the N- and C-termini of this domain in order to preserve its Fn binding activity.

It was found that both Fn binding domains of protein F were independently capable of conferring adherence to the ECM when expressed within an unrelated surface protein. The multisite interaction of protein F with both Fn and ECM offers an obvious advantage to the bacterium for increasing its chances for successful adherence to the ECM or the host cells. This advantage may be particularly important, since Fn, both as a soluble molecule and as a component of the ECM, interacts with many additional host molecules. Thus, there may be a competition between the infecting microorganisms and these host molecules for a limited number of binding sites. The presence of multiple binding sites for Fn on protein F assures productive interaction of the pathogen with a wide variety of host tissues.

To determine the relative contributions to adherence of UR and RD2, additional experiments are required. For example, it would be necessary to compare the abilities of soluble UR and RD2 proteins to block adherence of GAS strains to the ECM and epithelial cells. This would indicate whether the observations described here, such as the relative affinities of UR and RD2 for Fn and their interactions with different fragments of Fn, can be extrapolated to the interaction of GAS with the ECM and cells.

Materials and methods

Bacterial strains

Streptococcus pyogenes JRS145 is derived from JRS4 (M^+F^+) by insertional inactivation of the gene encoding the M-type 6 protein, *emm6.1* (Norgren *et al.*, 1989). SAM1 and SAM2 are derived from JRS4 and JRS145, respectively, by insertion of an ΩKm-2 element into *prt*F, the gene encoding protein F. *S.pyogenes* SS636-M13 and SS914-M59 were obtained from Dr A.Beck (Streptococcal Reference Laboratory, Ministry of Health, Jerusalem, Israel). *E.coli* DH5a served as the host in molecular cloning experiments, HB101 was used in Fn binding assays (Hanski and Caparon, 1992) and SG13009 (pREP4) served for expression of cloned proteins. *S.pyogenes* was grown in Todd–Hewitt medium (Difco, Detroit, MI) supplemented with 0.2% yeast extract as previously described (Hanski and Caparon, 1992) and *E.coli* was cultured in Luria–Bertani broth. Antibiotics were used at the following concentrations: kanamycin, 500 μg/ml for *S.pyogenes* and 25 μg/ml for *E.coli*; streptomycin, 1 mg/ml for *S.pyogenes*; erythromycin, 1 μg/ml for *S.pyogenes* and 750 μg/ml for *E.coli*; and ampicillin, 100 μg/ml for *E.coli*.

DNA procedures

PCRs were performed on a thermocycler (M.J.Research) as detailed elsewhere (Natanson *et al.*, 1995). The primers were designed to preserve the reading frame of the different amplified protein F segments, utilizing the *Kpn*I restriction site of the pQE30 vector (Qiagen). The primers were also designed so that the insertion of the amplified fragments between the *Kpn*I and *Hin*dIII sites of *emm6.1* will preserve the *emm6.1* reading frame across both the *Kpn*I and *Hin*dIII junctions (see below). The list of primers used in this study and the combinations of primers and templates used to express the various Fn binding domains are described in Tables I and II, respectively. PCR products were purified using commercial reagents (Wizard, Promega; QIAEX, Qiagen), cloned into this vector and sequenced using a modified T7 polymerase (Sequenase), according to the directions of the manufacturer (US Biochemicals). The QIAexpress sequencing primers PQ2 and PQ4 (Qiagen), were used for sequencing of all the plasmids indicated in Table II, and also for cloning two of these plasmids.

Expression of Fn binding domains

Expression of Fn binding domains was conducted as previously described (Sela *et al.*, 1993; Hanski *et al.*, 1995), using the expression vector pQE30 (Qiagen). This vector encodes a short affinity tag of six histidine residues at the N-terminus of the expressed protein, which enables its purification in a single step by affinity chromatography on an immobilized

Table II. Templates and primers used for construction of plasmids

Plasmid	Primers used for amplification	Template
pRD2-1	P5-3:P3-7	pUR-5[a]
pRD2-2	P5-2:P3-3	pUR-5
pRD2-3	P5-4:P3-7	pUR-5
pRD2-4	P5-6:P3-7	pUR-5
pRD2-5	P5-4:P3-4	pUR-5
pRD2-6	P5-4:P3-6	pUR-5
pRD2-7	P5-5:PQ4	pRD2-6
pRD2-8	PQ2:P3-5	pRD2-6
pUR-1	P5-1:P3-1	pUR-5
pUR-2	P5-1:P3-3	pUR-5
pUR-3	P5-1:P3-4	pUR-2
pUR-4	P5-1:P3-2	pUR-2
pUR-5	P5-1:P3-7	SS914[b]

The plasmids designated pRD2-1–pRD2-8 and pUR-1–pUR-4 are described in Figures 2A and 1A, respectively.
[a]pUR-5, which is not described in the figures, harbors a DNA segment containing UFBD and two RD2 repeats, of which the second is partial.
[b]Chromosomal DNA of the strain SS914.

Table I. Oligonucleotide sequence and location in *prt*F of the primers used in this study

Primer	Sequence	Position (bp)
P5-1	5′-GGGGGGGTACCAAAGACCAATCACCTCTAGCGGGT-3′	1502–1525
P5-2	5′-CAGGTTTTTCAGGTACCATGGTTGAGACAG-3′	1611–1640
P5-3	5′-GGGGGGTACCTTTTCAGGAAATATGGTTGAGACA-3′	1616–1639
P5-4	5′-GGGGGGGTACCATGGGAGGTCAAAGTGAGTCTGTT-3′	1667–1690
P5-5	5′-GGGGGGGTACCGTTGAATTTACTAAAGACACTCAA-3′	1688–1711
P5-6	5′-GGGGGGTACCCAAACAGGCATGAGTGGTCAAACA-3′	1709–1732
P3-1	5′-GGGGGGAAGCTTTCATATTTCCTGAAAAACCTGTTTC-3′	1630–1607
P3-2	5′-GGGGGGAAGCTTTCGTATCTTCTGTCTCAACCATATT-3′	1648–1625
P3-3	5′-GGGGGAAGCTTCCTGAGGAGTTGTTTGACCACT-3′	1741–1721
P3-4	5′-GGGGGGAAGCTTTCAACACTTCTGGCTCTTTCGT-3′	1777–1757
P3-5	5′-GGGGGGAAGCTTTAGACTCACTTTGACCTCCCAT-3′	1798–1778
P3-6	5′-GGGGGGAAGCTTTAAATTCAACAGACTCACTTTG-3′	1807–1787
P3-7	5′-GGGGGGAAGCTTTCGCCGTTTCACTGAAACCACTCAT-3′	2185–2162

nickel column. The expressed proteins were purified, under denaturing conditions, according to the directions of the manufacturer (Qiagen).

Expression of Fn binding domains within the surface-exposed region of M protein

Methods for the expression of protein F domains have been described in detail elsewhere (Hanski *et al.*, 1995). Briefly, the desired domain of protein F is amplified by PCR so that it can produce an in-frame fusion when used to replace a *Kpn*I–*Hin*dIII internal region of *emm6.1* (see Figure 3A), which is present on the streptococcal–*E.coli* shuttle vector pJRS233 (Perez-Casal *et al.*, 1993). Since the pJRS233 vector is temperature-sensitive for replication in *S.pyogenes*, transformation of SAM2, followed by growth at the non-permissive temperature (37°C), results in the integration of the plasmid into the *emm* locus of the streptococcal chromosome by homologous recombination (Perez-Casal *et al.*, 1993), and places the chimera under the control of the *emm6.1* regulatory apparatus (Okada *et al.*, 1993). This was shown to be the most efficient method for generating high level expression of mutant alleles of M protein (Perez-Casal *et al.*, 1993). Recombination into the correct locus for each strain was confirmed by PCR analyses using multiple primers, by Southern blot analyses and by Western blotting analyses using polyclonal antibodies directed against the corresponding antigens of protein F (not shown).

Analysis of Fn binding

Fn binding activity was determined by: (i) measurement of the amounts of iodinated Fn specifically bound to GAS strains expressing protein F or M/F hybrids, or (ii) a competition assay in which the ability of purified proteins expressing Fn binding domains of protein F to compete with various strains for iodinated Fn was determined. The assays were performed as described previously (Hanski and Caparon, 1992), except that in the competition assay the proteins expressing the Fn binding domains were admixed with the bacteria and iodinated Fn at the same time, and the incubation was performed for 1.5 h. Iodinated Fn binding to JRS145 was conducted at a bacterial suspension of 0.1–0.2 OD at 600 nm, whereas the binding to strains expressing M/F hybrid proteins was conducted at a bacterial suspension of 0.005 OD. Background values, obtained from tubes that contained no streptococci, were determined for each concentration of iodinated Fn, and were subtracted from the values obtained with streptococci. Results of a typical binding curve ranged from 8000 to 150 000 c.p.m. while background binding ranged between 300 and 4000 c.p.m. All results are means of duplicate or triplicate determinations that differed by <5%.

Adherence of GAS to ECM

Preparation of glass cover slips and plates coated with ECM was performed as described previously (Ishai-Michaeli *et al.*, 1992). Briefly, bovine corneal endothelial cells were seeded onto glass cover slips or 24 well plates at an initial density of 50 000 cells/ml and maintained in DMEM supplemented with 10% bovine calf serum, 5% fetal calf serum, 5% dextran T-40 and antibacterial agents. Six to eight days after reaching confluency, the cells were dissolved by Triton X-100 and the remaining ECM was washed extensively and kept at 4°C until further use. The main constituents of the corneal endothelial ECM were: Fn, laminin, collagen types I, III and IV, elastin and sulfated proteoglycans (Ishai-Michaeli *et al.*, 1992). Bacterial cells were grown overnight in THY, then washed three times in DMEM containing 0.2% Tween-20 and 10 mM HEPES at pH 7.4 and adjusted to OD = 0.2 at 600 nM. Bacterial suspensions of 0.5 ml were incubated at 4°C for 4 h with ECM-coated cover slips or plates, then washed six times with 1 ml of PBS. Washed cover slips were Gram-stained, and the degree of bacterial adherence to ECM was analyzed by light microscopy. For determination of adherence to ECM-coated plates, bacteria were grown in the presence of 3 μCi/ml [2-^{3}H]adenine (Amersham Life Science, Amersham Place, UK). Specific radioactivity was calculated for each experiment by adjusting the bacterial suspension to an OD of 0.2 at 600 nM and determining bacterial radioactivity by liquid scintillation spectrometry. Typically, 0.5 ml of bacterial suspension of OD = 0.2 contained 3×10^5 d.p.m., of which ~30% specifically adhered to ECM-coated plates. All results are means of triplicate determinations that differed by <5%.

Fn and Fn fragments

Fn was purified from human plasma according to the procedure of Vuento and Vaheri (1979). The 29 kDa Fib. 1, the 40 kDa Col., the 110 kDa cell binding and the 38 and 29 kDa heparin binding fragments were prepared from Fn by digestion of purified Fn with thermolysin according to the procedure of Borsi *et al.* (1986). The 70 kDa N-terminal fragment of Fn containing the Fib. 1 and Col. binding domain was prepared by digestion of purified Fn with cathepsin D followed by its isolation on gelatin–Sepharose as described by McKeown-Longo and Mosher (1985). The 20 kDa C-terminal fibrin binding domain was kindly provided by Dr K.C.Ingham from the Holland Laboratory, American Red Cross, Rockville, MD. All other reagents used in this study were of the highest available quality.

Acknowledgements

We thank Dr K.C.Ingham from the Holland Laboratory, American Red Cross, Rockville, MD for providing the C-terminal fibrin binding domain of fibronectin and Dr A.Beck from the Streptococcal Reference Laboratory, Ministry of Health, Jerusalem, Israel for providing the streptococcal strains SS636 of M-type 13 and SS914 of M-type 59. This work was supported in part by United States–Israel Binational Science Foundation, Ministry of Science and Arts and Ministry of Health Foundations (to E.H.).

References

Bisno,A.L. (1991) Group A streptococcal infections and acute rheumatic fever. *N. Engl. J. Med.*, **325**, 783–793.

Borsi,L., Castellani,P., Balza,E., Siri,A., Pellecchia,C., De Scalzi,F. and Zardi,L. (1986) Large-scale procedure for the purification of fibronectin domains. *Anal. Biochem.*, **155**, 335–345.

Bozzini,S., Visai,L., Pignatti,P., Petersen,T.E. and Speziale,P. (1992) Multiple binding sites in fibronectin and staphylococcal fibronectin receptor. *Eur. J. Biochem.*, **207**, 327–333.

Flock,J.-I. *et al.* (1987) Cloning and expression of the gene for a fibronectin-binding protein from *Staphylococcus aureus*. *EMBO J.*, **6**, 2351–2357.

Hanski,E. and Caparon,M.G. (1992) Protein F, a fibronectin-binding protein, is an adhesin of the group A streptococcus—*Streptococcus pyogenes*. *Proc. Natl Acad. Sci. USA*, **89**, 6172–6176.

Hanski,E., Horwitz,P.A. and Caparon,M.G. (1992) Expression of protein F, the fibronectin-binding protein of *Streptococcus pyogenes* JRS4, in heterologous streptococcal and enterococcal strains promotes their adherence to respiratory epithelial cells. *Infect. Immun.*, **60**, 5119–5125.

Hanski,E., Fogg,G., Tovi,A., Okada,N., Burstein,I. and Caparon,M.G. (1995) Molecular analysis of *Streptococcus pyogenes* adhesion. *Methods Enzymol.*, **253**, 269–310.

Hasty,D.L., Ofek,I., Courtney,H.S. and Doyle,R.J. (1992) Multiple adhesins of streptococci. *Infect. Immun.*, **60**, 2147–2152.

Hynes,R.O. (1990) *Fibronectins*. Springer-Verlag, Berlin.

Ishai-Michaeli,R., Svahn,C.M., Weber,M., Chajek-Shaul,T., Korner,G., Ekre,H.-P. and Vlodavsky,I. (1992) Importance of size and sulfation of heparin in release of basic fibroblast growth factor from the vascular endothelium and extracellular matrix. *Biochemistry*, **31**, 2080–2088.

Jönsson,K., Signäs,C., Müller,H.-P. and Lindberg,M. (1991) Two different genes encode fibronectin binding proteins in *Staphylococcus aureus*. *Eur. J. Biochem.*, **202**, 1041–1048.

McGavin,M.J., Gurusiddappa,S., Lingren,P.-E., Lindberg,M., Raucci,G. and Höök,M. (1993) Fibronectin receptors from *Streptococcus dysgalactiae* and *Staphylococcus aureus*. *J. Biol. Chem.*, **268**, 23946–23953.

McKeown-Longo,P.J. and Mosher,D.F. (1985) Interaction of the 70,000-mol-wt amino-terminal fragment of fibronectin with the matrix-assembly receptor of fibroblasts. *J. Cell Biol.*, **100**, 364–374.

Mosher,D.F. (1989) *Fibronectin*. Academic Press. New York.

Mosher,D.F. and Proctor,R.A. (1980) Binding and factor XIIIa-mediated cross-linking of a 27-kilodalton fragment of fibronectin to *Staphylococcus aureus*. *Science*, **209**, 927–929.

Natanson,S., Sela,S., Moses,A.E., Musser,J.M., Caparon,M.G. and Hanski,E. (1995) Distribution of fibronectin-binding proteins among Group A streptococci of different M-types. *J. Infect. Dis.*, **171**, 871–878.

Norgren,M., Caparon,M.G. and Scott,J.R. (1989) A method for allelic replacement that uses the conjugative transposon Tn*916*: deletion of the *emm6.1* allele in *Streptococcus pyogenes* JRS4. *Infect. Immun.*, **57**, 3846–3850.

Okada,N., Geist,R.T. and Caparon,M.G. (1993) Positive transcriptional control of *mry* regulates virulence in group A streptococcus. *Mol. Microbiol.*, **7**, 893–903.

Ólsen,A., Hanski,E., Normark,S. and Caparon,M.G. (1993) Molecular characterization of fibronectin binding proteins in bacteria. *J. Microbiol. Methods*, **18**, 213–226.

Perez-Casal,J., Price,J.A., Maguin,E. and Scott,J.R. (1993) An M protein with a single C repeat prevents phagocytosis of *Streptococcus pyogenes*: use of a temperature sensitive shuttle vector to deliver homologous sequences to the chromosome of *S.pyogenes*. *Mol. Microbiol.*, **8**, 809–819.

Sela,S., Aviv,A., Tovi,A., Burstein,I., Caparon,M.G. and Hanski,E. (1993) Protein F: an adhesin of *Streptococcus pyogenes* binds fibronectin via two distinct domains. *Mol. Microbiol.*, **10**, 1049–1055.

Signäs,C., Raucci,G., Jönsson,K., Lingren,P.-E., Anantharamaiah,G.M., Höök,M. and Lindberg,M. (1989) Nucleotide sequence of the gene for fibronectin-binding protein from *Staphylococcus aureus*: use of this peptide sequence in the synthesis of biologically active peptides. *Proc. Natl Acad. Sci. USA*, **86**, 699–703.

Stevens,D.L, Tanner,M.H, Winship,J., Swarts,R., Ries,K.M., Schlievert, P.M. and Kaplan,E.L. (1989) Severe group A streptococcal infections associated with toxic shock-like syndrome and scarlet fever toxin A. *N. Engl. J. Med.*, **321**, 1–7.

Talay,S.R., Ehrenfeld,E., Chhatwal,G.S. and Timmis,K.N. (1991) Expression of the fibronectin-binding components of *Streptococcus pyogenes* in *Escherichia coli* demonstrates that they are proteins. *Mol. Microbiol.*, **5**, 1727–1734.

Talay,S.R., Valentin-Weigand,P., Jerlström,P.G., Timmis,K.N. and Chhatwal,G.S. (1992) Fibronectin-binding protein of *Streptococcus pyogenes*: sequence of the binding domain involved in adherence of streptococci to epithelial cells. *Infect. Immun.*, **60**, 3837–3844.

Talay,S.R., Valentin-Weigand,P., Timmis,K.N. and Chhatwal,G.S. (1994) Domain structure and conserved epitopes of Sfb protein, the fibronectin-binding adhesin of *Streptococcus pyogenes*. *Mol. Microbiol.*, **13**, 531–539.

Vuento,M. and Vaheri,A. (1979) Purification of fibronectin from human plasma by affinity chromatography under non-denaturing conditions. *Biochem. J.*, **183**, 331–337.

Wannamaker,L.W. (1970) Differences between streptococcal infections of the throat and skin. *N. Engl. J. Med.*, **282**, 23–30.

Westerlund,B. and Korhonen,T.K. (1993) Bacterial proteins binding to mammalian extracellular matrix. *Mol. Microbiol.*, **9**, 687–694.

Received on September 5, 1995; revised on November 2, 1995